The Value of Information and Dispersion*

Chang-Koo (CK) Chi†
University of Wisconsin-Madison

March 22, 2014

Abstract

This paper studies the value of information in theory of decision under uncertainty. I introduce a novel way to rank signals based on *dispersion* and integrate it with the previous three important signal orderings: (1) Lehmann (1988) precision, (2) effectiveness in statistical decision theory, and (3) informativeness in Bayesian decision theory. By incorporating this new ordering into the model, I establish the *equivalence* of these four different orderings within *each* of three classes of payoff functions: supermodular, single-crossing, and interval dominance order. As the first consequence of this equivalence theorem, I show that the Lehmann precision is both necessary and sufficient for one signal to be more valuable than another to both statisticians and Bayesian decision makers. Second, I exactly characterize the relationship between more precise signals and higher dispersion: a more precise signal generates more dispersed predictions about the true state of the world. This justifies another signal orderings used in the previous literature. Third, I illustrate how this result can be applied to strategic settings, by analyzing the effects of more precise information in three standard economic environments: auctions, bilateral contracts, and delegation.

JEL Classification: C44, C60, D81, D44, D86.
Keywords: Decision Theory, Value of Information, Lehmann precision, Effectiveness, Informativeness, Dispersion, Signal Orderings, Monotone Comparative Statics, Auctions, Contracts, Delegation.

*I am grateful to Daniel Quint for his advice and encouragement. I also thank Bill Sandholm and Lones Smith. The paper benefited from discussion with Joseph Halpern, Sang Yoon (Tim) Lee, Antonio Penta, Marzena Rostek, Ricardo Serrano-Padial, Bruno Strulovici, Marek Weretka, Jiabin Wu, and especially Jianrong Tian. All remaining errors are mine.
†Mail: William H. Sewell Social Science Building 1180 Observatory Dr. Madison, WI 53706-1393, E-mail: cchi2@wisc.edu. Phone: +1-608-263-3866.

1. Introduction

Many decisions get made based on imperfect information about relevant variables. An investor chooses investment portfolios with partial information about the future return of each financial asset. An employer hires workers with partial information about their innate skills. People get dressed in the morning with partial information about the weather. In all of these cases, the decision maker may have several possible sources of information, and before making a decision, can choose one of them to gather information relevant to her choices. In this paper I characterize how to value information in economic decision problems under uncertainty, establishing necessary and sufficient conditions for one source of information—called a *signal* hereafter—to be more valuable than another.[1]

In general, there are two distinctive approaches to deriving a binary relation over a set of possible signals. The first, the *preference-based ordering*, ranks signals based on the values they generate to a decision maker. Given a category of preference, it is natural to think of one signal being more valuable than another if it leads to higher expected payoffs for all decision makers whose preferences fall into the category. The second approach, the *statistical ordering*, focuses directly on the statistical characteristics of signals and rank s signals according to an apt statistical notion. The question of when one signal can be judged to be more valuable than another goes back to the pioneering works of Blackwell (1951, 1953), who linked this question to the statistical notion of sufficiency and showed the equivalence between the statistical ordering based on sufficiency and the preference-based ordering without any structures on preferences. Lehmann (1988) developed a more complete statistical ordering, compared to sufficiency, based on the notion of precision and established the equivalence between the two approaches within a certain category of preferences, large enough to include most decision problems of interest in statistics. However, their applications to economics have been limited in that Blackwell sufficiency is an extremely partial order and Lehmann precision is tailored to statistical contexts.

This paper attempts to fill this gap for the case of economic decision problems. I focus on three rich classes of preferences: supermodular, single-crossing, and interval dominance order. All of these three classes exhibits a complementarity between the decision maker's choice variable and the payoff-relevant state of the world.[2] More importantly, the classes I consider are large enough to encompass a variety of the most important economic applications including a firm's production planning under demand or cost uncertainty, an investor's portfolio decision under uncertainty about the return of a risky asset, and the matching problem in a model of marriage. Given each of these three preferences, I look for a statistical condition under which one signal is more valuable than another for all decision makers within the class.

[1]Whereas the source of information is called a statistical "experiment" in statistics following Blackwell (1951, 1953), it is typically referred to as a "signal" or "information structure" in economics.

[2]The class of single-crossing payoff functions, posited in Milgrom and Shannon (1994), exhibits an ordinal complementarity which is independent of order-preserving transformations. The class of supermodular payoff functions, first introduced by Topkis (1978), exhibits a cardinal complementarity which is not preserved under all such transformations. The class of interval dominance order payoff functions, recently developed by Quah and Strulovici (2009), reflects a even weaker ordinal complementarity than the one of single-crossing functions. Therefore, the class of interval dominance order functions includes single-crossing functions, which in turn themselves include supermodular functions.

For this purpose, this paper develops a novel statistical ordering based on the notion of *dispersion*. More precisely, let α and β be the two signals we want to compare. This new signal ordering concerns the decision maker's predictions about the state:

(D) α generates more dispersed predictions about the state of the world than β ($\alpha \succ_D \beta$) if for every nondecreasing function defined on the state, its expected value conditional on α is more variable than the expected value on β.

The main contributions of this paper are threefold. First, I show that this statistical ordering provides the necessary and sufficient condition of α being more valuable than β. The value of a signal, however, hinges upon the decision maker's decision principle as well as her primitive payoff function. For an illustration, let $\mathcal{U}^\star$ be a class of payoff functions satisfying some property $\star$ and $u \in \mathcal{U}^\star$ be a decision maker's payoff function. Suppose that the decision maker adopts the minimax principle, a major method of making a decision in statistical decision theory, so she acts to maximize the expected payoff evaluate on the worst-case basis. In order for α to be more valuable than β in this context, the two signals should meet the following criterion:

(E) α is more effective than β with respect to a class of payoff functions $\mathcal{U}^\star$ ($\alpha \succ_E^\star \beta$) if any expected payoffs attainable with β is also attainable with α for every decision maker within the class.

In economics, however, it is customary to think of a decision maker as Bayesian who acts to maximize expected payoffs given her beliefs. Since this different decision principle alludes to the Bayesian (or economic) value of a signal being different from its statistical value, we need a different criterion than the effectiveness above:

(I) α is more informative than β with respect to a class of payoff functions $\mathcal{U}^\star$ ($\alpha \succ_I^\star \beta$) if α leads to higher expected payoffs than β for every decision maker within the class.

I show that the dispersion-based ordering provides the sufficient and necessary condition for α to be more effective *and* informative than β within *each* of the three classes: supermodular $\mathcal{U}^{spm}$, single-crossing $\mathcal{U}^{sc}$, and interval dominance order $\mathcal{U}^{ido}$. Consequently, in common with Blackwell (1951, 1953) and Lehmann (1988), I establish the equivalence in both Bayesian and statistical decision theory between ordering signals based on a statistical notion and ordering signals based on the value to the decision maker.

One striking implication of this equivalence theorem is that we can rank two signals based on the same statistical notion, despite the fact that $\mathcal{U}^{sc}$ is a strict superset of $\mathcal{U}^{spm}$. Athey and Levin (2001) demonstrated that when we expand the family of decision problems, a more restrictive statistical condition is required for the preference-based ordering.[3] In contrast to their finding, the results of this paper assert that a more restrictive condition is unnecessary: The class $\mathcal{U}^{spm}$ is rich enough that if α is preferred to β for all decision makers within $\mathcal{U}^{spm}$, β is never preferred to α for the decision maker within $\mathcal{U}^{sc}$. Hence it is immaterial to comparing

[3]The idea can be traced back to Lehmann (1988). Unlike Blackwell, who does not place any limitations on the decision maker's preferences, Lehmann restricts the scope of payoff functions to the set of Karlin-Rubin Monotone payoff functions. As a result, his statistical notion of precision is less restrictive than Blackwell's sufficiency. See Section 2 for its details.

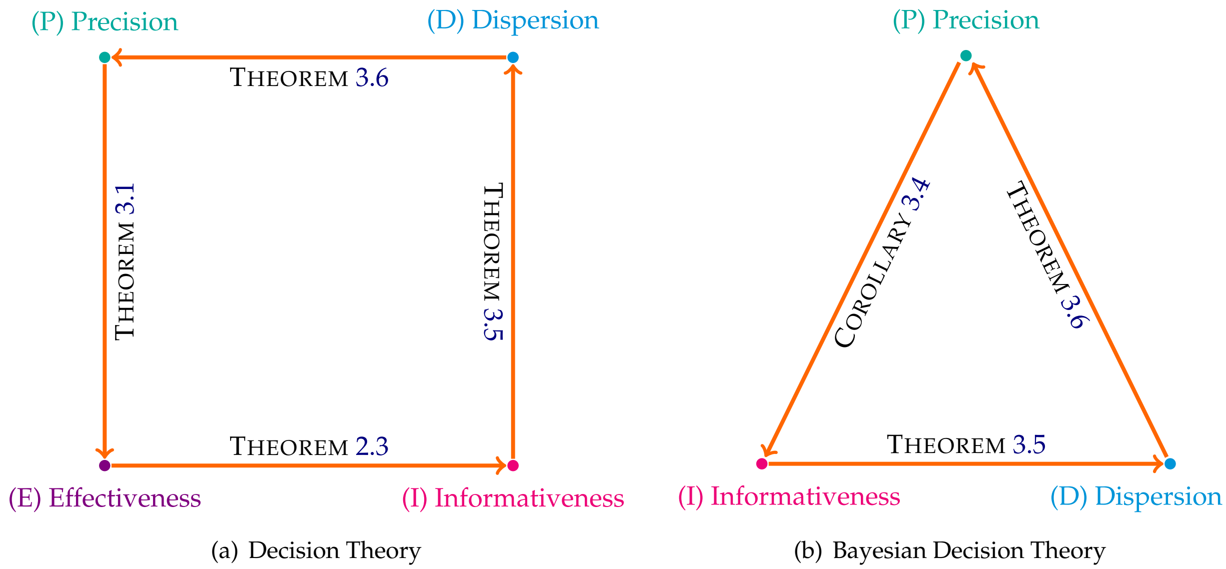


Figure 1: The Value of Information and Dispersion - The equivalence of four different signal orderings in theory of decision under uncertainty within each class of payoff functions: supermodular, single-crossing, and interval dominance order.

the values of signals whether the complementarity between the decision maker's choice and the state is a cardinal or an ordinal concept.

The second contribution of this paper is the characterization of this new signal ordering in terms of the statistical notion developed by Lehmann (1988):

(P) α is more Lehmann precise than β ($\alpha \succ_L \beta$) if α is more likely to generate high outcomes than β when the state is high.

The Dispersion Theorem presented in Section 4 demonstrates that this statistical ordering, called ”Lehmann precision” hereafter, is equivalent to the dispersion-based signal ordering. When a decision maker receives information from a more precise signal, she will put more weight on the signal's outcome, since the signal conveys more information about the state of the world. Hence her prediction about the state will change more depending on the outcome. The theorem states that the other implication is also true: if one signal generates more dispersed predictions about the states than another, it has to be more precise in Lehmann's notion. In fact, there are a variety of signal orderings developed in the previous literature based on this intuition, and they measure the signal's precision level by the variability of predictions. The theorem exactly captures this idea by equating $\alpha \succ_L \beta$ to $\alpha \succ_D \beta$, and thus helps to justify the previous signal orderings in a unified way.

To my knowledge, this theorem is the most general formulation of the link between Lehmann precision and the dispersion. Theorem 5.2 in Lehmann (1988) provides a similar characterization result when the associated density functions with signals are logconcave. However, the theorem of this paper does not impose any conditions on the signal's structures we want to compare, so it can be regarded as an extension of his result.

Putting the first two results together, this paper demonstrates that the above signal orderings based on four different concepts are mutually equivalent within each of the three classes of payoff functions as illustrated in Figure 1-(a). Under the Bayesian frameworks, this result reduces to the equivalence between the three orderings—Lehmann precision, informativeness, and dispersion—as depicted in Figure 1-(b). This abridged version is reminiscent of the classic work of Rothschild and Stiglitz (1970). They argued that two lotteries, X and Y, over monetary payoffs can be compared based on relative riskiness in three different ways. (i) A statistical ordering: $Y = X + \epsilon$ where $\mathbb{E}[\epsilon|X] = 0$, (ii) a preference-based ordering: $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for every concave Bernoulli payoff function u, and (iii) a dispersion-based ordering: the probability distribution of Y is the mean preserving spread of that of X. Hence the result in Figure 1-(b) extends their fundamental idea of comparing two lotteries in state-independent payoff functions to the idea of comparing two signals in state-dependent payoff functions.

The third contribution is to analyze the effect of more precise information in several strategic environments by utilizing the Dispersion Theorem. The first application is to auctions. I analyze the effects of more precise information on efficiency, revenue, and bidders' information rents. The second application is to the principal-agent model with hidden action. I show that a more precise contractible variable is necessary and sufficient for the principal to be able to control the agent's hidden action with less cost. The third application is to a delegation problem without bilateral money transfers. I provide a sufficient condition under which more precise private information for the agent is more valuable to the principal.

The remainder of the paper is organized as follows: Section 2 presents the mathematical preliminaries. It begins with the basic elements of decision problems under uncertainty and documents the four different signal orderings, along with some well-known results about the value of information in decision theory. Section 3 presents the main result of this paper. I shall demonstrate that the four signal orderings above are mutually equivalent in order illustrated in Figure 1-(a). First, I show in Theorem 3.1 that Lehmann precision is sufficient for ordering signals based on their statistical values. For this purpose, I exploit one theorem of Quah and Strulovici (2009) but give a more succinct and constructive proof based on the idea that when payoff functions exhibit complementarities, there exists some action which weakly dominates every nonincreasing strategy. This idea provides a unified way to construct a payoff-improving strategy with a more precise signal for the three classes of payoff functions I consider.[4] After, it is shown in Theorem 3.5 and Theorem 3.6 that the criterion of informativeness induces the statistical signal ordering based on the dispersion which, in turn, induces the signal ordering based on Lehmann precision.

In Section 4 we discuss the three by-products of the equivalence theorem, with emphasis on the connection between Lehmann precision and another signal orderings developed by previous literature and the dispersion theorem. In section 5, I illustrate how the dispersion theorem is applied to auctions, bilateral contracts, and delegation. Section 6 concludes.

[4]Moreover, I show in Appendix A that the improvement principle, the way to find such a strategy, is essentially equivalent to the principle developed by Lehmann (1988).

2. Preliminaries

2.1. Basic Elements

1. *State Space and Information Structures*: Let $\langle \Theta, \mathcal{F} \rangle$ denote a measurable space that represents the unknown state of the world.[5] A *signal* α (or a statistical experiment in statistics) is a random quantity X and a set $\{G^{\alpha}(\cdot|\theta)\}_{\theta\in\Theta}$ of possible distributions of X conditioned on the true state of the world θ. By purchasing the signal α, a decision maker can observe a sample $x \in \mathcal{X}$ of the random quantity X, where $\mathcal{X}$—the set of all conceivable outcomes—is called the sample space. Let $\beta = \langle Y, \{G^{\beta}(\cdot|\theta)\}_{\theta\in\Theta} \rangle$ be another signal concerning the same state of the world θ. Similarly, $y \in \mathcal{Y}$ denotes the observed sample of the random quantity Y. We assume that both sample spaces $\mathcal{X}$ and $\mathcal{Y}$ are subsets of real numbers. Additionally, we shall assume *without any loss of generality* that the state space Θ is an interval of real numbers $[\underline{\theta}, \overline{\theta}]$.[6] It is supposed that the cumulative distribution of X for each state θ is absolutely continuous and thus it takes a form of

$$G^{\alpha}(x|\theta) = \Pr(X \leq x|\theta) = \int_{\mathcal{X}} \mathbb{1}_{\{s\leq x\}}(s) g^{\alpha}(s|\theta) ds \ \text{ for each } x \in \mathcal{X},$$

where g^{α} is the density function with respect to Lebesgue measure.[7]

Throughout the remainder of this paper, we shall be concerned about the signal α with $g^{\alpha}(x|\theta)$ being atomless and possessing the *monotone likelihood ratio property* (MLRP). The MLRP requires that $g^{\alpha}(x|\theta)$ obey the following inequality:

$$\text{(MLRP)} \quad g^{\alpha}(x_1|\theta_1)g^{\alpha}(x_2|\theta_2) - g^{\alpha}(x_1|\theta_2)g^{\alpha}(x_2|\theta_1) \geq 0, \quad \forall\, x_1 > x_2 \quad \text{and} \quad \theta_1 > \theta_2.$$

In a nutshell, the property states that the signal α gives a higher outcome more likely in state θ_1 rather than in state θ_2.[8] As is well-known by Karlin and Rubin (1956), the set of probability distributions with the MLRP involves the class of probability distributions of the exponential type, to which most of the distributions studied in statistics and economics

[5]Each element θ in Θ is a complete description of exogenous variables of the model considered. The symbol Θ will be used to denote the set of all possible states of the world.

[6]It is notable that this assumption does not cause any loss of generality. Lemma B in Karlin and Rubin (1956) proves that we can restrict our attention to a *statistically connected* state space under the environment with the monotone likelihood ratio property.

[7]In case X is discrete, G^{α} need not be continuous. However, as Lehmann (1988) argues, one can construct a new variable X^* with a continuous distribution, which is statistically equivalent to X. For instance, suppose that X takes only a finite number of distinct values $\{x_1, \cdots, x_n\}$ with probability of $\Pr(X = x_k|\theta, \alpha) = \pi_k$ for each $k = 1, \cdots, n$. One can define X^* by the continuous distribution of

$$G^{\alpha}(x|\theta) = \begin{cases} \sum_{i=1}^{k-1} \pi_i & \text{if } x \in (x_{k-1}, x_k) \\ \sum_{i=1}^{k-1} \pi_i + \pi_k U_k & \text{if } x = x_k, \end{cases}$$

where $\{U_k\}$ is the i.i.d sequence of random variables uniformly distributed on $(0,1)$. Then the new variable X^* is continuous and equivalent to X.

[8]This interpretation coincides with the interpretation of the first-order stochastic dominance. The difference is, however, while the former imposes a statistical restriction on the density function, the latter on the distribution function. Milgrom (1981) shows their equivalence when all possible prior beliefs on Θ are taken into account.

belong.[9] We label by E the set of signals of our interest. Therefore, we shall be concerned with $E \triangleq \{\alpha = \langle X, \{G^\alpha(x|\theta)\}_{\theta\in\Theta}\rangle \,|\, g^\alpha(x|\theta)$ obeys the MLRP$\}$.

2. *Action Space and Decision Rules*: Let A indicate a set of allowed and undominated actions that the decision maker can choose in the face of the uncertainty θ.[10] For the sake of simplicity, we shall assume that A is a compact subset of $\Re$.[11] In a decision problem with a single signal α, a decision rule or a strategy d^α is a measurable mapping from the sample space $\mathcal{X}$ into the action space A. We label by $\mathcal{D}^\alpha$ the set of all permissible decision rules based on α. Similarly, $\mathcal{D}^\beta = \{d : \mathcal{Y} \to A|\text{measurable}\}$ is the set of all permissible decision rules based on another signal β.

3. *Payoff Functions*: Following Wald (1950), a decision problem can be modeled by a payoff function. A payoff function $u(a, \theta)$, specifying the gain to the decision maker when θ is the true state and a is the action taken, is a real-valued function defined on $A \times \Theta$. We assume that u is measurable with respect to θ for each $a \in A$ and the family of payoff functions $\{u(\cdot, \theta)\}_{\theta\in\Theta}$ for all possible states is equicontinuous in a.[12] We denote by $\mathcal{U}$ the *class* of all payoff functions satisfying these two properties.

Most of decision problems of interest in economics exhibit *complementarity* between the decision maker's action choice and the state of the world, in the sense that if a higher action a' is desirable at some state θ rather than a, the action a' would remain desirable at every state $\theta' \geq \theta$. For this reason, we shall be concerned with the following three subclasses of payoff functions of $\mathcal{U}$, which are classified by the strength of complementarity between a and θ.

DEFINITION 2.1 (SPM Family). *A payoff function $u(a, \theta)$ is supermodular (SPM, and we write $u \in \mathcal{U}^{spm}$) in $(a; \theta)$ provided the incremental return from taking higher actions is nondecreasing in θ. Namely, for every $a' > a$ in A, ${}^u\Delta(\theta) \triangleq u(a', \theta) - u(a, \theta)$ is nondecreasing in θ.*[13]

[9] A probability distribution is said to be of the exponential type if the density function is represented by

$$g(x|\theta) = \beta(\theta)e^{x\theta}r(x) = \frac{e^{x\theta}r(x)}{\int_X e^{s\theta}r(s)ds},$$

where $r : X \to \Re_+$ is measurable and nonnegative. Examples of this type are the binomial and Poisson for the discrete case, and the normal with known variance and chi square for the continuous case. Refer to Milgrom (1981) for the application of the MLRP to several economic environments.

[10] A dominated action a is the action for which there exists $a' \in A$ satisfying $u(a', \theta) \geq u(a, \theta)$ for all $\theta \in \Theta$. We do not consider the action space with such an element.

[11] In decision theory or monotone comparative statics literature, the action space A is frequently given by a partially ordered set (poset) with a binary relation $\geq$ on A. In this case, the sufficient and necessary condition for A to be compact (with respect to the interval topology) is that A is a *complete lattice*, that is, every nonempty subset of A has a supremum and an infimum in A (See Topkis (1998)). For simple analysis, however, we assume in this paper that A is a subset of $\Re$. All the results below can be extended into the case that A is a complete lattice.

[12] A family $\{u(\cdot, \theta)\}_{\theta\in\Theta}$ of real-valued functions $u(\cdot, \theta) : A \to \Re$ is said to be equicontinuous at $a_0 \in A$ if for every $\epsilon > 0$, there exists a $\delta(a_0) > 0$ such that $d(a, a_0) < \delta(x_0)$ implies $|u(a_0, \theta) - u(a, \theta)| < \epsilon$ for all $\theta \in \Theta$. Here d is a metric endowed with the action space A in case A is a metric space. The family of functions is said to be equicontinuous if it is equicontinuous at every $a_0 \in A$. The equicontinuity of the family guarantees the existence of the optimal decision rule in Bayesian decision theory.

[13] Since $u(a, \theta)$ is defined on the product of the two ordered sets, the concepts of supermodularity and increasing differences coincide (Topkis (1998)).

DEFINITION 2.2 (SC Family). *A payoff function $u(a,\theta)$ obeys the single-crossing property (SC, $u \in \mathcal{U}^{sc}$) in $(a;\theta)$ provided the incremental return from taking higher actions satisfies the single crossing property in θ. That is, for every $a' > a$ in A, ${}^u\Delta(\theta) \triangleq u(a',\theta) - u(a,\theta) \geq 0$ implies ${}^u\Delta(\theta') \geq 0$ for all $\theta' \geq \theta$.*

The SCP payoff function expresses a weaker complementarity than the SPM one. While the SCP is an ordinal property that is preserved by any order-preserving transformations, the SPM is a cardinal property that is not preserved by all such transformations. Note that $\mathcal{U}^{spm} \subset \mathcal{U}^{sc}$ as every nondecreasing function ${}^u\Delta(\theta)$ clearly crosses the horizontal axis from below at most once. The next class of payoff functions exhibits an even weaker complementarity than the SCP, and thus it is the largest among the three classes.

DEFINITION 2.3 (IDO Family, Quah and Strulovici (2009)). *A payoff function $u(a,\theta)$ obeys the interval dominance order (IDO, $u \in \mathcal{U}^{ido}$) property provided for every $a'' > a'$ in A,*

$$u(a'',\theta) - u(a,\theta) \geq 0, \;\; \forall\, a \in [a',a''] \;\; \textit{implies} \;\; u(a'',\theta') - u(a',\theta') \geq 0, \;\; \forall\, \theta' \geq \theta.$$

The IDO property requires the higher action a'' to be desirable for every $\theta' \geq \theta$ only if it is more desirable than any other actions belonging to the interval $[a',a''] = \{a \in A | a' \leq a \leq a''\}$ at θ. Whereas both the IDO and SCP properties have the same consequent, the former has a weaker antecedent than the latter. Hence $\mathcal{U}^{ido}$ includes $\mathcal{U}^{sc}$. Indeed, $\mathcal{U}^{ido}$ is large enough to encompass the most decision problems of interest, not only in economics, but even in statistics such as the problems of testing a hypothesis or doing point and interval estimation of the unknown parameter θ.[14]

4. *Decision Problem and its Process*: A decision problem concerning θ with a single signal is specified by the collection of the above elements—the state space, the signal structure, the action space, and the payoff function—$(\Theta, \langle X, \{G^\alpha(\cdot|\theta)\}_{\theta\in\Theta}\rangle, A, u(a,\theta))$. Since we are concerned about comparison of signals with holding Θ and A fixed, we label the decision problem by (α, u) for the simple exposition.

The decision problem (α, u) proceeds as follows: A decision maker observes an outcome x after purchasing an informative signal α. Then she chooses an action a from a given action space A, on the basis of the observed outcome, before the state of the world is realized. At the end of the day, the decision maker's payoff is determined by $u(a,\theta)$.

5. *The Value of Information in Decision Theory*: Theory of decision under uncertainty has two large branches depending on the assumption of prior information and the decision principle used. In *Bayesian decision theory*, the decision maker is endowed with prior beliefs on Θ and she acts to maximize her Bayesian expected payoff (the conditional Bayes principle). *Statistical decision theory*, however, does not call for such a prior belief. Also, the decision maker acts to

[14]Most of statistical decision problems belong to the Karlin-Rubin Monotone (KRM) payoff functions, extensively studied by Karlin and Rubin (1956) and Lehmann (1988), but the IDO family involves the KRM family. In Appendix A and B, we will look more in detail at the inclusive relationship among the several classes of payoff functions studied in past literature.

protect her payoff against the worst possible state of the world (the minimax or minimax regret principle).

Comparison of signals (or statistical experiments), pioneered by the celebrated works by Blackwell (1951, 1953), characterizes a binary preference relation on the signal space E for a certain class of payoff functions $\mathcal{U}^\star \subset \mathcal{U}$ according to the *value* generated by each signal. Formally, the field of inquiry attempts to derive an exact statistical condition under which the decision problem (α, u) yields more expected values than (β, u) for every payoff function $u \in \mathcal{U}^\star$. As is highlighted above, however, the differences in the assumption of priors and decision principles between statistical and Bayesian decision theory give rises to different values of a signal. Hence we need two preference-based orderings tailored to each branch. Below, we will go into the details under the statistical and Bayesian frameworks.

2.2. Effectiveness in Statistical Decision Theory

Analysis of statistical decision problems using the minimax or minimax regret principle calls for one concept of an expected payoff function. For an illustration, consider a given statistical decision maker (or a statistician) facing the decision problem (α, u). The *expected payoff function of decision d in state θ* is the function ρ^α defined on $\mathcal{D}^\alpha \times \Theta$ such that [15]

$$\rho^\alpha(d, \theta) \triangleq \int_{\mathcal{X}} u(d(x), \theta) dG^\alpha(x|\theta).$$

Then $\inf_{\theta\in\Theta} \rho^\alpha(d, \theta)$ represents the least expected payoff that can happen if the decision rule d is used. The minimax principle—the major principle of choice in statistical decision theory—requires the statistician to choose $d \in \mathcal{D}^\alpha$ so as to maximize $\inf_{\theta\in\Theta} \rho^\alpha(d, \theta)$.[16] We define the value function

$$V(\alpha, u) = \max_{d\in\mathcal{D}^\alpha} \inf_{\theta\in\Theta} \rho^\alpha(d, \theta)$$

as the *statistical value* generated by the signal α. For the decision problem (β, u) with another signal β, we define the value function $V(\beta, u)$ in an analogous manner, together with the set of decision rules $\mathcal{D}^\beta$ based on β.

Therefore, given a certain class of payoff functions $\mathcal{U}^\star$, it is natural to say that the signal α is more valuable than β to every statistician within $\mathcal{U}^\star$ if $V(\alpha, u) \geq V(\beta, u)$ for every $u \in \mathcal{U}^\star$. The next condition provides a signal ordering in this context:

DEFINITION 2.4 (Effectiveness). *The signal α is more effective than another signal β with respect to a class of payoff functions $\mathcal{U}^\star$ (in symbols, $\alpha \succ_E^\star \beta$) provided, for every decision rule $d \in \mathcal{D}^\beta$, there exists a decision rule $d^\alpha \in \mathcal{D}^\alpha$ for which $\rho^\alpha(d^\alpha, \theta) \geq \rho^\beta(d, \theta)$ for every $\theta \in \Theta$.*

This criterion—the signal ordering based on the statistical values—states that any expected

[15] This formulation is to treat a decision problem as a (statistical) game against nature, where $d \in \mathcal{D}^\alpha$ and θ are regarded as the action chosen by the statistician and by nature, respectively.

[16] An alternative decision principle—the minimax regret procedure—requires the statistician to choose d so as to maximize $\inf_\theta [\rho^\alpha(d, \theta) - \max_d \rho^\alpha(d, \theta)]$.

payoff function attainable with β is also attainable with the more effective signal α regardless of the state θ. It is immediate that a more effective signal generates more statistical values for every payoff function concerning θ.

The other way is also true, when the class of payoff functions is invariant to the addition of every function in the set $\mathcal{H} = \{h : \Theta \to \Re|$ measurable$\}$, i.e., $u \in \mathcal{U}^\star$ implies $u + h \in \mathcal{U}^\star$ for all $h \in \mathcal{H}$. To see this, let $d^\alpha \in \mathcal{D}^\alpha$ represent a minimax strategy for some payoff function $u \in \mathcal{U}^\star$. Then $V(\alpha, u) \geq V(\beta, u)$ $\forall u \in \mathcal{U}^\star$ implies

$$\begin{aligned}\inf_{\theta} \{\rho^\alpha(d^\alpha, \theta) + h(\theta)\} &\geq \max_{d \in \mathcal{D}^\beta} \inf_{\theta} \{\rho^\beta(d, \theta) + h(\theta)\}, \quad \forall\, h \in \mathcal{H} \\ &\geq \inf_{\theta} \{\rho^\beta(d, \theta) + h(\theta)\}, \quad \forall\, d \in \mathcal{D}^\beta \text{ and } h \in \mathcal{H},\end{aligned}$$

which is equivalent to $\rho^\alpha(d^\alpha, \theta) \geq \rho^\beta(d, \theta)$ $\forall$ θ and $\forall$ d. Therefore, for every class $\mathcal{U}^\star$ invariant to the addition of $\mathcal{H}$, α is more effective than β with respect to $\mathcal{U}^\star$ if and only if α presents more statistical values than β for every statistician with the payoff function $u \in \mathcal{U}^\star$.

Observe that the property of complementarity between a and θ is independent of the addition of $h \in \mathcal{H}$. Therefore, for every class of payoff functions introduced in the previous subsection, the effectiveness becomes the necessary and sufficient condition for α to generate more statistical values than β.

LEMMA 2.1. *For each property $\star \in \{SPM, SC, IDO\}$ and for two signals α and β, $\alpha \succ_E^\star \beta$ if and only if $V(\alpha, u) \geq V(\beta, u)$ for all u within the given class.*

Another salient property that the three classes of payoff functions possess is, discovered by Quah and Strulovici (2009), that one can restrict attention to a reduced set of decision rules for identifying the minimax decision rule. Formally, given the signal α and a class of payoff functions $\mathcal{U}^\star$, a subset $\mathcal{D}'$ of the set of all permissible decision procedures $\mathcal{D}^\alpha$ is called an *essentially complete class* provided for every decision rule $d \in \mathcal{D}^\alpha \sim \mathcal{D}'$, there exists a $d' \in \mathcal{D}'$ such that $\rho^\alpha(d', \theta) \geq \rho^\alpha(d, \theta)$ for every θ. By definition, the minimax strategy d^α must belong to the essentially complete class $\mathcal{D}'$.

THEOREM 2.2 (Essentially Complete Class). *For the every decision problem (α, u) with $u \in \mathcal{U}^{ido}$, the set of monotone decision rules $\mathcal{D}^{\alpha,M} = \{d : \mathcal{X} \to A|$ measurable and nondecreasing$\}$ constitutes an essentially complete class.*

In statistical decision theory, the essentially complete class theorem is of utmost importance in ranking two signals based on the criterion of effectiveness. By virtue of this theorem, the problem of proving the effectiveness of α greatly simplifies into the one of finding a monotone decision rule $d^\alpha \in \mathcal{D}^{\alpha,M}$ such that $\rho^\alpha(d^\alpha, \theta) \geq \rho^\beta(d, \theta)$ for every monotone $d \in \mathcal{D}^{\beta,M}$. That is, we need only look at the set of monotone decision rules, without any loss of generality, rather than the set of all permissible decision rules, so the problem becomes much easier.

2.3. Informativeness in Bayesian Decision Theory

Bayesian decision theory stems from the assumption that a given decision maker is *Bayesian rational*: (i) The decision maker forms a specific prior concerning θ, captured by a probability measure π on all possible relevant events $\mathcal{F}$. We denote by $P(\Theta)$ the set of all possible prior beliefs on Θ. Note that, given a signal α, each prior belief π yields the marginal distribution of X by Bayes' rule, which will be labeled by $M^\alpha(x) = \Pr(X \leq x)$, if necessary.[17] (ii) After observing an outcome x from the signal α, she updates her prior beliefs using Bayes' rule whenever possible. The updated beliefs—posterior beliefs—will be described by $F^\alpha(t|x) = \Pr(\theta \leq t | X = x)$, i.e., the conditional probability of the event $\{\theta \leq t\} \in \mathcal{F}$ on the outcome $X = x$. To be consistent with the prior, the following condition must hold:

$$\text{(1)} \qquad \mathbb{E}_X\left[F^\alpha(t|X)\right] = \pi(\{\theta \leq t\}) \;\; \forall\, t \in \Theta,$$

where $\mathbb{E}_X$ is the mathematical expectation over the random quantity X. (iii) Lastly, the decision maker chooses $a \in A$ so as to maximize her expected payoffs given her posterior beliefs:

$$\delta^\alpha(x) \in \operatorname*{argmax}_{a \in A} \int_\Theta u(a,\theta) dF^\alpha(\theta|x).$$

To distinguish it from the minimax decision rule, we label the (optimal) Bayesian decision rule by $\delta^\alpha : \mathcal{X} \to A$. Define the corresponding value function as

$$\mathcal{V}^\pi(\alpha, u) \triangleq \mathbb{E}_\theta\left[\int_X u(\delta^\alpha(x),\theta) dG^\alpha(x|\theta)\right],$$

where $\mathbb{E}_\theta$ represents the mathematical expectation over Θ with respect to the prior beliefs π. We refer to $\mathcal{V}^\pi(\alpha, u)$ as the *Bayesian value* of the signal α. Notice that, unlike the statistical value $V(\alpha, u)$, the Bayesian value of α hinges upon the decision maker's prior beliefs $\pi \in P(\Theta)$. Similarly, define by $\mathcal{V}^\pi(\beta, u)$ the Bayesian value corresponding to another signal β.

In Bayesian decision theory, therefore, it is natural to say that α is more valuable than β if $\mathcal{V}^\pi(\alpha, u) \geq \mathcal{V}^\pi(\beta, u)$. The next criterion extends this concept to all Bayesian decision makers within a certain class.

DEFINITION 2.5 (Informativeness). *The signal α is more informative than another signal β with respect to a class of payoff functions $\mathcal{U}^\star$ (in symbols, $\alpha \succ_I^\star \beta$) provided $\mathcal{V}^\pi(\alpha, u) \geq \mathcal{V}^\pi(\beta, u)$ for all $\pi \in P(\Theta)$ and $u \in \mathcal{U}^\star$.*

Note that free of the decision maker's prior information is the criterion of informativeness adopted for ranking signals. In order to meet this criterion, α should yield more ex ante payoffs to Bayesian decision makers independent of their priors. This prior-free facet of informativeness is predicated upon the following two reasons. First, we compare two signals based on their values not for a single decision maker but for every every decision maker falling within

[17] Although the marginal distribution of a signal is determined by the prior beliefs as well, I drop the notation π for the sake of simplicity.

a certain category of preferences. In view of the fact that prior information is the *subjective* assessment of a decision maker, therefore, it is hard to support that all of them have the identical prior beliefs.[18]

In addition, prior information is generally formed by past experience about similar decision problems. Hence such information is not available in situations where the decision maker encounters the source of uncertainty θ for the first time and thus statistical investigation—purchasing a signal or conducting an experiment—is probably the unique source of gathering information about θ. The informativeness is designed to address these two issues.[19]

So far, we presented the way of ordering signals based on the values. In light of Definition 2.4 and 2.5, however, it is immediate that informativeness is implied by effectiveness. Moreover, it is worth noting that this proposition holds no matter what class of payoff functions we have in mind.

THEOREM 2.3. *For every class of payoff functions $\mathcal{U}^\star$, $\alpha \succ_E^\star \beta$ implies $\alpha \succ_I^\star \beta$.*

PROOF OF THEOREM 2.3 : Let $\delta^\beta \in \mathcal{D}^\beta$ represent a Bayesian decision rule that maximizes her expected payoffs. By using Fubuni's Theorem, we can write the Bayesian values of β in terms of ρ^β,

$$\mathcal{V}^\pi(\beta, u) = \mathbb{E}_Y\left[\int_\Theta u(\delta^\beta(Y), \theta) dF^\beta(\theta|Y)\right] = \int_\Theta \rho^\beta(\delta^\beta, \theta) d\pi(\theta)$$

Since the effectiveness guarantees the existence of a better decision rule $d^\alpha \in \mathcal{D}^\alpha$ for which $\rho^\alpha(d^\alpha, \delta) \geq \rho^\beta(\delta^\beta, \theta)$ regardless of θ, the informativeness follows by integrating the inequality with respect to the measure π. $\square$

Before turning to next signal ordering, we remark one more crucial property that all of the three classes of payoff functions possess. In order to state this property, we define the set of optimal actions for each sample $x \in \mathcal{X}$ as $A^\alpha(x) = \{a^* \in A|\ a^* \in \text{argmax}_{a\in A}\ \mathbb{E}_\theta\left[u(a,\theta)|X = x\right]\}$.

THEOREM 2.4 (Monotone Comparative Statics). *For every decision problem (α, u) with $u \in \mathcal{U}^{ido}$ and for every prior $\pi \in P(\Theta)$, the set of optimal actions is nondecreasing in the strong set order. That is, for $x' > x$, the set $A^\alpha(x')$ is larger than $A^\alpha(x)$ in the strong set order.*[20]

Like the essentially complete class theorem, the monotone comparative statics theorem is useful in ranking two signals based on the criterion of informativeness. By the aid of this result, we can select a monotone strategy $\delta^\alpha : \mathcal{X} \to A$ from the set $\mathcal{D}^{\alpha,M}$ to write the Bayesian value

[18]In some contexts, however, it is reasonable to assume that they have the same prior information. For example, experienced employers may very well possess an objective prior about employee's unknown skills in practice; investment consultants with expertise may have a definite prior information about the returns of a stock, in particular, issued by S&P 500 companies. Athey and Levin (2001) analyzed the value of information in Bayesian settings in this regard. We discuss their results in Section 4.

[19]It should not be construed from the above statements that a Bayesian decision maker does not have an explicit prior information. Each decision maker does possess an explicit prior, but we leave open possibilities that her prior beliefs might be different from others although they have a common source of uncertainty.

[20]Given two subsets D_1 and D_2 of $A \subset \Re$, we say that D_1 is larger than D_2 in the strong set order (or in the induced set order) and it is written $D_1 \trianglerighteq D_2$ provided for every $d_1 \in D_1$ and $d_2 \in D_2$, $\max\{d_1, d_2\} \in D_1$ and $\min\{d_1, d_2\} \in D_2$. In case D_1 and D_2 are subsets of a lattice A with an order topology $\geq$, $D_1 \trianglerighteq D_2$ provided $d_1 \vee d_2 \in D_1$ and $d_1 \wedge d_2 \in D_2$.

of α as

$$\mathcal{V}^{\pi}(\alpha, u) = \mathbb{E}_X\left[\mathbb{E}_\theta\left[u(\delta^\alpha(X), \theta)|X\right]\right] = \int_\Theta \rho^\alpha(\delta^\alpha, \theta) d\pi(\theta)$$

In order to prove that α is more informative than β within $\mathcal{U}^{ido}$, therefore, it is enough to construct a monotone strategy $d^\alpha \in \mathcal{D}^{\alpha,M}$ such that $\rho^\alpha(\delta^\alpha, \theta) \geq \rho^\beta(d, \theta)$ for every $d \in \mathcal{D}^{\beta,M}$.

2.4. Lehmann Precision

The above discussion suggests that, for $\alpha = \langle X, \{G^\alpha(\cdot|\theta)\}_{\theta\in\Theta}\}\rangle$ to be more effective or informative than $\beta = \langle Y, \{G^\beta(\cdot|\theta)\}_{\theta\in\Theta}\}\rangle$ with respect to a class of decision problems, the random quantity X has to be more statistically correlated with the unknown state θ than Y. However, the degree of correlation necessary for α to be more valuable than β is also dependent upon the class of payoff functions $\mathcal{U}^\star$ we consider. Intuitively, as we expand the scope of payoff functions $\mathcal{U}^\star$, we need stronger statistical correlation between X and θ for α to be more effective or informative to every decision problem within $\mathcal{U}^\star$.

Blackwell sufficiency—the most standard statistical signal ordering—provides a way to rank two signals based on the statistical concept of sufficiency. In the influential papers, Blackwell (1951, 1953) showed, without any structures on the class of payoff functions, that α is Blackwell-sufficient for β if and only if α is more effective than β to every statistician within $\mathcal{U}$. Consequently, although powerful, the signal ordering based on sufficiency is too restrictive to provide a reasonable order in the signal space as pointed out by Lehmann (1988).[21] This leaves open possibility that a more complete and intuitive ordering can be found by narrowing down the scope of payoff functions.[22]

DEFINITION 2.6 (Lehmann (1988)). *The signal α is more Lehmann-precise than β (and we write $\alpha \succ_L \beta$) provided for each outcome $y \in \mathcal{Y}$, there exist an increasing function $T_y : \Theta \to \mathcal{X}$ such that*

(P) $\quad G^\alpha(T_y(\theta)|\theta) = G^\beta(y|\theta).$

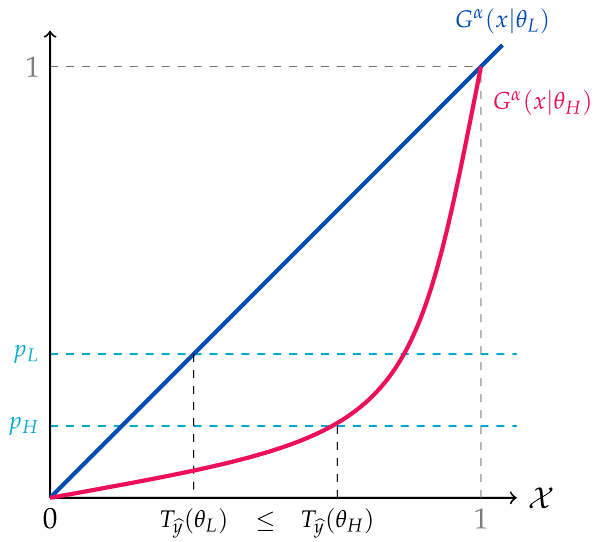
To better understand why the monotone T-transformation leads to a more statistically precise signal, consider a simple case of dichotomy : $\Theta = \{\theta_L, \theta_H\}$ with $\theta_L < \theta_H$. Furthermore, we suppose that both sample spaces $\mathcal{X}$ and $\mathcal{Y}$ are a unit interval and $G^\alpha(c|\theta_L) = G^\beta(c|\theta_L)$ for each $c \in [0,1]$. Refer to Figure 2 where we display the distributions of X and Y. Consider a distribution $G^\beta(\cdot|\theta_H)$ of Y when the state is θ_H. By the MLRP, it must dominate $G^\beta(\cdot|\theta_L)$ in the

[21] Theorem 3.1 in his paper tells us that when we compare the following two location distributions with

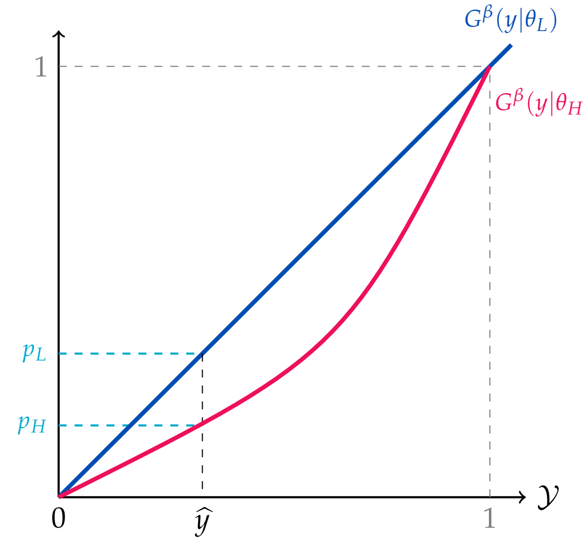
$$G^\alpha(\cdot|\theta) \overset{d}{=} \mathrm{U}\left[\theta - \frac{\alpha}{2},\ \theta + \frac{\alpha}{2}\right] \quad \text{and} \quad G^\beta(\cdot|\theta) \overset{d}{=} \mathrm{U}\left[\theta - \frac{1}{2},\ \theta + \frac{1}{2}\right],$$

the sufficient and necessary condition for α to be Blackwell-sufficient for β is $1/\alpha \in \{1, 2, 3, \cdots\}$, although α conveys more information on θ for every $0 < \alpha < 1$. However, for every $0 < \alpha < 1$, it can be shown that α is more Lehmann-precise than β. See Example 2.2.

[22] Restricting attention to the class of *quasiconcave with increasing peaks* payoff functions (See Appendix B for its definition.), Lehmann (1988) proves that α is more effective than β within the reduced class if and only if α is more Lehmann-precise than β.



(a) Signal α



(b) Signal β

Figure 2: Lehmann (1988) Precision and T-transformation

first-order stochastic dominance (FOSD). Then for each sample $\hat{y} \in \mathcal{Y}$, the signal β assigns two quantiles $p_L \equiv G^{\beta}(\hat{y}|\theta_L)$ and $p_H \equiv G^{\beta}(\hat{y}|z_H)$ to the event $\{Y \leq \hat{y}\}$ when the state is θ_L and θ_H, respectively.

Given the quantile p_L, there is a sample $x \in \mathcal{X}$ at which $G^{\alpha}(x|\theta_L) = p_L$, which we label by $T_{\hat{y}}(\theta_L)$.[23] Since we assume the identical distributions for θ_L, we have $T_{\hat{y}}(\theta_L) = \hat{y}$. When the state is θ_H, however, the monotone T-transformation calls for $T_{\hat{y}}(\theta_H) \geq \hat{y}$. Hence the graph of $G^{\alpha}(\cdot|\theta_H)$ will be uniformly below the graph of $G^{\beta}(\cdot|\theta_H)$ like the FOSD.[24] As a consequence, the signal α assigns relatively more densities to higher outcomes than β when the state is θ_H, so α is more statistically precise than β.

Another simple example of the T-transformation can be found when each signal is generated by adding a noise.

Example 2.1 (Gaussian Learning[25]). *Suppose that a decision maker observes $x = \theta + \epsilon_{\alpha}$ or $y = \theta + \epsilon_{\beta}$ from the signal α or from β where ϵ_{α} and ϵ_{β} are normally distributed, independent of θ, as $N(0, \sigma_{\alpha}^2)$ and $N(0, \sigma_{\beta}^2)$, respectively. Then converting each conditional distribution into the standard normal distribution yields*

$$G^{\alpha}(x|\theta) = \Phi\left(\frac{x-\theta}{\sigma_{\alpha}}\right) \quad \textit{and} \quad G^{\beta}(y|\theta) = \Phi\left(\frac{y-\theta}{\sigma_{\beta}}\right),$$

[23] In general, Lehmann precision imposes no restrictions on $G^{\alpha}(\cdot|\theta_L)$ in case of dichotomy. In light of this fact, Jewitt (2007) showed the equivalence between Blackwell sufficiency and Lehmann precision in the two states of the world.

[24] This statement is indeed true. It can be shown that $G^{\alpha}(\cdot|\theta_H) \leq G^{\beta}(\cdot|\theta_H)$ is equivalent to the Lehmann-precision, assuming the identical distributions for θ_L and the binary states.

[25] This signal structure is frequently used, in Morris and Shin (2002) and Angeletos and Pavan (2007), for analyzing the effect of the precise signal on social welfare in macroeconomic models.

where Φ, the cumulative distribution function of $N(0,1)$, is bijective. Hence the associated T-transformation with the two signals is $T_y(\theta) = \sigma_\alpha/\sigma_\beta(y-\theta)+\theta$. Therefore, $\alpha \succ_L \beta$ if and only if $\sigma_\alpha < \sigma_\beta$.

The next lemma furnishes us a simple but useful characterization of Lehmann precision.[26]

LEMMA 2.5. *$\alpha \succ_L \beta$ if and only if $G^\beta(y|\theta) - G^\alpha(x|\theta)$ satisfies the single crossing property in θ for every pair $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.*

PROOF OF LEMMA 2.5 : Suppose $\alpha \succ_L \beta$. By definition, there exists an increasing transformation $T_y : \Theta \to \mathcal{X}$ for which $G^\beta(y|\theta) = G^\alpha(T_y(\theta)|\theta)$ for each y. Assume $G^\beta(y|\theta_0) - G^\alpha(x|\theta_0) \geq 0$ for some θ_0. Since the distribution function $G^\alpha(\cdot|\theta)$ is nondecreasing, it must be the case that $x \leq T_y(\theta_0)$. Thus, for $\theta \geq \theta_0$

$$G^\beta(y|\theta) - G^\alpha(x|\theta) \geq G^\beta(y|\theta) - G^\alpha(T_y(\theta_0)|\theta) \geq G^\beta(y|\theta) - G^\alpha(T_y(\theta)|\theta) = 0,$$

where the last inequality is due to $T_y(\theta) \geq T_y(\theta_0)$.

To prove the converse, define the transformation $T_y : \Theta \to \mathcal{X}$ for each outcome y such that $G^\beta(y|\theta) = G^\alpha(T_y(\theta)|\theta)$ holds. Continuity of G^α and G^β guarantees the existence of such T_y. We need to show that T_y is increasing with θ. Given a pair of outcomes x and y, suppose that $G^\beta(y|\theta_0) - G^\alpha(x|\theta_0) = 0$ for some $\theta_0 \in \Theta$. Then $T_y(\theta_0) = x$ by construction. Due to the single crossing property, we have for every $\theta \geq \theta_0$

$$G^\beta(y|\theta) - G^\alpha(x|\theta) = G^\beta(y|\theta) - G^\alpha(T_y(\theta_0)|\theta) \geq G^\beta(y|\theta) - G^\alpha(T_y(\theta)|\theta) = 0,$$

and thus we obtain $T_y(\theta) \geq T_y(\theta_0)$. $\square$

This lemma provides a simple sufficient condition for $\alpha \succ_L \beta$. Recall that if the density function $g(x|\theta)$ satisfies the MLRP, then the corresponding cumulative distribution $G(x|\theta)$ is decreasing in θ, i.e., for $\theta \geq \theta_0$, $G(x|\theta)$ dominates $G(x|\theta_0)$ in the FOSD. Therefore, if for every x and y, $G^\alpha(x|\theta)$ is declining in θ with a larger rate than $G^\beta(y|\theta)$, then $G^\beta(y|\theta) - G^\alpha(x|\theta)$ clearly satisfies the single crossing property in θ, so α becomes more Lehmann-precise than β.

EXAMPLE 2.2 (Uniform Distributions). *Consider the two signals with uniform distributions*

$$X \sim U\left[\theta - \frac{\alpha}{2},\, \theta + \frac{\alpha}{2}\right] \quad \text{and} \quad Y \sim U\left[\theta - \frac{\beta}{2},\, \theta + \frac{\beta}{2}\right].$$

Note that for every pair $x \in \mathcal{X}$ and $y \in \mathcal{Y}$

$$\frac{\partial}{\partial\theta} G^\alpha(x|\theta) = -\frac{1}{\alpha} < -\frac{1}{\beta} = \frac{\partial}{\partial\theta} G^\beta(y|\theta) \ \text{ if and only if } \ \alpha < \beta.$$

Therefore, $\alpha \succ_L \beta$ if and only if $\alpha < \beta$.

[26] In some statistics textbook, the Lehmann precision is defined as the statement in Lemma 2.5.

EXAMPLE 2.3 (Truth-or-Noise). *Suppose that the signal α perfectly reveals the true state of the world with probability of α, but is independently drawn from a distribution F with the support Θ otherwise. Similarly, $y = \theta$ with probability of β but $y \sim F$ otherwise. That is, both parameters α and β measure the precision of each signal. Note that the cumulative distribution of each signal is written $G^{\alpha}(x|\theta) = \alpha\mathbb{1}_{\{x\geq\theta\}}(x) + (1-\alpha)F(x)$ and $G^{\beta}(y|\theta) = \beta\mathbb{1}_{\{y\geq\theta\}}(y) + (1-\beta)F(y)$, respectively. We use Lemma 2.5 to show that $\alpha \succ_L \beta$ if and only if $\alpha > \beta$.*

Suppose that $\alpha > \beta$. We want to show that $\Delta^{x,y}(\theta) \equiv G^{\beta}(y|\theta) - G^{\alpha}(x|\theta)$ satisfies the SCP in θ for every pair $x, y \in \Theta$. Case 1: $x < y$. Note that $\Delta^{x,y}(\theta)$ is strictly positive for $\theta \geq y$ and that is nondecreasing for $\theta < y$; it jumps by α at $\theta = x$. Hence the SCP is satisfied. Case 2: $x > y$. When $\Delta^{x,y}(\theta)$ is negative for $\theta \geq x$, the function becomes negative everywhere. When $\Delta^{x,y}(\theta)$ is positive for $\theta \geq x$, it is enough for the SCP to show that $\Delta^{x,y}(\theta)$ is negative for $\theta < y$. In case $\theta < y < x$, the function assumes

$$\Delta^{x,y}(\theta) = \beta + (1-\beta)F(y) - \alpha - (1-\alpha)F(x) = (\beta-\alpha)(1-F(x)) + (1-\beta)(F(y)-F(x)).$$

Therefore, $\alpha > \beta$ and $y < x$ implies $\Delta^{x,y}(\theta) < 0$. Case 3: When $x = y$, $\Delta^{x,x}(\theta)$ crosses the horizontal axis from below at $\theta = x$.

To prove the converse, suppose that $\Delta^{x,y}$ satisfies the SCP for all x, y. When $x = y$, the function takes on the value $(\beta-\alpha)(1-F(x))$ for $\theta \leq x$ and $(\alpha-\beta)F(x)$ for $\theta > x$. Hence $\alpha > \beta$ is necessary for the SCP of $\Delta^{x,x}(\theta)$. $\square$

Compared to Blackwell sufficiency, Lehmann precision is easier to check as seen in the examples above. More importantly, this signal ordering is applicable to more signal structures than sufficiency, so that it appeared in many economic settings for analyzing the value of information.

2.5. Dispersion

The main contribution of this paper is to integrate the above three different concepts of ranking the signals—effectiveness, informativeness, and Lehmann precision—with another concept based on *dispersion*. The main idea behind the link to the dispersion can be best understood by the relationship between the precision of a signal and the dispersion of posterior beliefs: As a signal gets more statistically precise, its outcome becomes more trustworthy, and thus a decision maker's posterior beliefs about θ become more dispersed depending on the outcome. In order to incorporate this simple idea into our setup, we need introduce a few stochastic orderings to which the dispersion gives rise.

Let X and Y be two random variables such that $\mathbb{E}[\sigma(X)] \leq \mathbb{E}[\sigma(Y)]$ for all increasing convex functions σ, provided that the both expectations exist. Then we say that X is smaller than Y *in the increasing convex order*, and we write $X \leq_{\text{icx}} Y$. Roughly speaking, if $X \leq_{\text{icx}} Y$, then X is "less variable" than Y in the stochastic sense. The increasing concave order $\leq_{\text{icv}}$ is defined in a similar manner, and $X \leq_{\text{icv}} Y$ implies that X is more variable than Y.[27]

[27] Note that a function $f(x)$ is increasing and convex if and only if $-f(-x)$ is increasing and concave. Hence the two variability orders are closely related to each other in the following sense: $X \leq_{\text{icx}} Y$ if and only if $-X \geq_{\text{icv}} -Y$. For a

The following lemma is simply the reflection of the two facts that $(x-c)\vee 0 \triangleq \max\{x-c,0\}$ is an increasing and convex function for all constants c and every increasing convex function belongs to the closed convex cone generated by the family of functions $\{(x-c)\vee 0 | c\in\Re\}$ up to constants.[28]

LEMMA 2.6. *Let X and Y be two random variables with probability distributions F_X and F_Y, respectively. Then $X \leq_{icx} Y$ if and only if*

$$\int_{\kappa}^{\infty}[1-F_X(x)]dx \leq \int_{\kappa}^{\infty}[1-F_Y(y)]dy \quad \textit{for all } \kappa\in\Re. \tag{2}$$

The lemma associates with the familiar notion in economics; from the inequality (2) and Rothschild and Stiglitz (1970), it follows that $X \leq_{\mathrm{icv}} Y$ is equivalent to domination of X by Y in the second-order stochastic dominance. See footnote 28.

In case the two random variables have the same mean, the two variability orders $\geq_{\mathrm{icx}}$ and $\geq_{\mathrm{icv}}$ reduce into $\geq_{\mathrm{cx}}$ and $\geq_{\mathrm{cv}}$, respectively, where $X \geq_{\mathrm{cx}} Y$ implies $\mathbb{E}[\sigma(X)] \geq \mathbb{E}[\sigma(Y)]$ for every convex function σ. It is worth noting that whereas $\geq_{\mathrm{cx}}$ and $\geq_{\mathrm{cv}}$ are dependent on the locations of random variables in the sense that the orders can be used only for comparing the variability of random variables with the same mean, both $\geq_{\mathrm{icx}}$ and $\geq_{\mathrm{icv}}$ are independent of the locations of random variables.[29] Therefore, with $\geq_{\mathrm{icx}}$ and $\geq_{\mathrm{icv}}$, we are able to compare much more distributions in terms of variability.

Let Θ^* be the set of bounded, nondecreasing and measurable functions from Θ to $\Re$. Given a signal α and prior beliefs $\pi \in P(\Theta)$, we define the linear operator $J^\alpha : \Theta^* \times \mathcal{X} \to \Re$ as follows:

$$J^\alpha[\psi](x) \triangleq \int_\Theta \psi(\theta)dF^\alpha(\theta|x). \tag{3}$$

That is, the operator J^α represents the conditional expectation of ψ on the basis of the signal's outcome $x\in\mathcal{X}$. Note that J^α is the integral of ψ with respect to the posterior beliefs. Hence the value of J^α relies on the prior beliefs π as well, although we suppress π for simple exposition. In addition, note that the operator $J^\alpha[\psi](X)$, when we replace the sample x with the random quantity X, is a random variable for each ψ. With a different signal β but the same prior beliefs π, we define $J^\beta[\psi](Y)$ in the same way.

Now we are ready to state the main concept of this paper:

DEFINITION 2.7 (Dispersion). *The signal α generates a more dispersed prediction about θ than another*

comprehensive discussion of many notions of variability (dispersion) order, refer to Shaked and Shanthikumar (2007).

[28]Likewise, every increasing concave function can be approximated by $\{(x-c)\wedge 0 | c\in\Re\}$ where $(x-c)\wedge 0 \triangleq \min\{x-c,0\}$. Hence one could get the counterpart to Lemma 2.6 for the increasing concave order: $X \leq_{\mathrm{icv}} Y$ if and only if

$$\int_{-\infty}^{\kappa} F_X(x)dx \geq \int_{-\infty}^{\kappa} F_Y(y)dy \quad \text{for all } \kappa\in\Re.$$

[29]To see why the convex order $\geq_{\mathrm{cx}}$ requires the same mean between two random variables, note that both x and $-x$ are convex. Hence $X \leq_{\mathrm{cx}} Y$ implies $\mathbb{E}[X] \leq \mathbb{E}[Y]$ and $\mathbb{E}[-X] \leq \mathbb{E}[-Y]$ so $\mathbb{E}[X] = \mathbb{E}[Y]$. Also note that $X \leq_{\mathrm{cx}} Y$ if and only if $X \geq_{\mathrm{cv}} Y$.

signal β (and we write $\alpha \succ_D \beta$) provided $J^\alpha[\psi](X) \geq_{icx} J^\beta[\psi](Y)$ for every $\psi \in \Theta^*$ and $\pi \in P(\Theta)$.

The dispersion furnishes us another way to rank the two signals in terms of the variability of the operator J. $\alpha \succ_D \beta$ requires the estimated value of every nondecreasing function ψ based on α to be "more variable" than the estimated value of ψ based on β, for a given prior information on θ.

I wish to highlight that the set of nondecreasing functions defined on Θ is large enough that it subsumes some important predictions about θ as special cases. As an illustration of what can be gained by taking every nondecreasing function on Θ, note that when $\psi(\theta) = \theta$ the operator J^α reduces to $\mathbb{E}[\theta|X]$. Hence $\alpha \succ_D \beta$ leads to more variable conditional expectations of θ. For another example, note that $\psi(\theta) = 1 - \mathbb{1}_{\{\theta \leq t\}}(\theta)$ is nondecreasing with θ for every constant $t \in \Theta$. In this case, J^α assumes

$$J^\alpha[\psi](X) = 1 - \int_\Theta \mathbb{1}_{\{\theta \leq t\}}(\theta) dF^\alpha(\theta|X) = 1 - F^\alpha(t|X)$$

Hence $J^\alpha[\psi](X) \geq_{\text{icx}} J^\beta[\psi](Y)$ implies $1 - F^\alpha(t|X) \geq_{\text{icx}} 1 - F^\beta(t|Y)$, which in turn implies $F^\alpha(t|X) \geq_{\text{cx}} F^\beta(t|Y)$. Namely, the signal α larger than β in this order generates more dispersed posterior beliefs.

The above examples are predicated on another signal orderings employed in past literature. We will revisit them and discuss the connection with Lehmann precision in the next section.

3. The Main Result

We now present the main result of this paper, which states the equivalence of the four different signal orderings we introduced in the previous section.

THEOREM (Equivalence). *For each class of payoff functions—$\mathcal{U}^{spm}$, $\mathcal{U}^{sc}$, or $\mathcal{U}^{ido}$—concerning θ and for two signals α and β, the following conditions are mutually equivalent:*

(P) α is more Lehmann-precise than β.

(E) α is more effective than β with respect to the given class of payoff functions.

(I) α is more informative than β with respect to the given class of payoff functions.

(D) α generates more dispersed predictions about θ than β.

The dispersion theorem, associating (P) a statistically precise signal with (D) high dispersion of predictions concerning θ, plays a key role in two aspects. First, sufficiency of the Lehmann precision provides the relationship between Lehmann precision and other signal orderings developed by previous literature. Second, necessity helps to derive the exact condition under which a signal α is more valuable than another signal β in both Bayesian and statistical decision theory.

In this section, we prove the equivalence of these four properties for $\mathcal{U}^{spm}$ and $\mathcal{U}^{sc}$ only. The class of $\mathcal{U}^{ido}$, the largest of the three classes, is relegated to Appendix B. We organize the proof in the following order: (P) $\to$ (E) $\to$ (I) $\to$ (D) $\to$ (P), where we have already verified in Theorem 2.3 that (E) $\to$ (I) holds irrespective of the class of payoff functions.

3.1. Precision implies Effectiveness

We begin with the following theorem:[30]

THEOREM 3.1. *If* $\alpha \succ_L \beta$, *then* $\alpha \succ_E^{sc} \beta$.

For the proof of the theorem, we will utilize the next two lemmas. The first lemma demonstrates that given a decreasing decision procedure d, there exists a *single action* which is weakly dominant over d for every payoff function $u \in \mathcal{U}^{sc}$. In Appendix B, we show that the IDO family also possesses the same property.

LEMMA 3.2. *Let* $u \in \mathcal{U}^{sc}$. *Then for every decreasing decision rule* $d : \Theta \to A$, *there exists an action* $a^\star \in A$ *such that* $u(a^\star, \theta) \geq u(d(\theta), \theta)$ *for all* $\theta \in \Theta$.

PROOF OF LEMMA 3.2 : Let $A_d = \{d(\theta) \in A | \theta \in \Theta\}$ denote the image of the given decision d. Since $u(a, \theta)$ satisfies the SCP in $(a; \theta)$, we infer from the Monotonicity Theorem in Milgrom and Shannon (1994) that there exists an increasing decision rule $d^\star$ such that

$$d^\star(\theta) \in \underset{a \in A_d}{\operatorname{argmax}} \ u(a, \theta).$$

Notice that the two strategies d and $d^\star$ cross only once.[31] Let $\theta^\star \in \Theta$ indicate the state at which they intersect. Then $d^\star(\theta) \geq d(\theta)$ for $\theta \geq \theta^\star$ and $d^\star(\theta) \leq d(\theta)$ otherwise. Set $a^\star = d^\star(\theta^\star)$, i.e., the optimal action at the intersection chosen by $d^\star$. Then $u(a^\star, \theta^\star) \geq u(d(\theta), \theta^\star)$ for all θ. For all $\theta > \theta^\star$, $a^\star \geq d(\theta)$ since the decision d is decreasing. Hence it follows from the SCP that

$$u(a^\star, \theta^\star) \geq u(d(\theta), \theta^\star) \ \text{ implies } \ u(a^\star, \theta) \geq u(d(\theta), \theta) \quad \forall\, \theta > \theta^\star.$$

The case $\theta < \theta^\star$ can be shown in an analogous way. □

The next lemma—dubbed the *improvement principle*—plays a central role in the proof that the signal ordering based on the Lehmann precision guarantees the two signal orderings based on both effectiveness and informativeness. The principle elaborates the way to construct a monotone decision rule based on a more precise signal, given a monotone decision rule based on another signal, which benefits the decision maker's payoffs regardless of θ.[32]

LEMMA 3.3 (Improvement Principle). *Suppose that* $u \in \mathcal{U}^{sc}$ *and* $\alpha \succ_L \beta$. *Then for every monotone decision rule* $d^\beta \in \mathcal{D}^{\beta,M}$, *there exists a monotone decision rule* $d^\alpha \in \mathcal{D}^{\alpha,M}$ *for which*

$$u(d^\alpha \circ T_y(\theta), \theta) \geq u(d^\beta(y), \theta) \quad \forall\, y \in \mathcal{Y},\ \theta \in \Theta, \tag{4}$$

[30] Quah and Strulovici (2009) prove the same statement for the IDO payoff functions, but I present a more succinct and intuitive proof. It helps us compare the improvement principle for $\mathcal{U}^{sc}$ with the principle for $\mathcal{U}^{qc}$ in Lehmann (1988).

[31] The Monotonicity Theorem in Milgrom and Shannon (1994) states that if the payoff function obeys the single crossing property in $(a; \theta)$, the set of maximizers $A^\star(\theta) \triangleq \operatorname{argmax}_{a \in A} u(a, \theta)$ is nondecreasing in the induced set order. Hence the theorem enables us to select a nondecreasing strategy from the maximizer. In case the maximizer is a singleton, there is nothing to prove.

[32] Refer to Appendix A for the key idea of this principle.

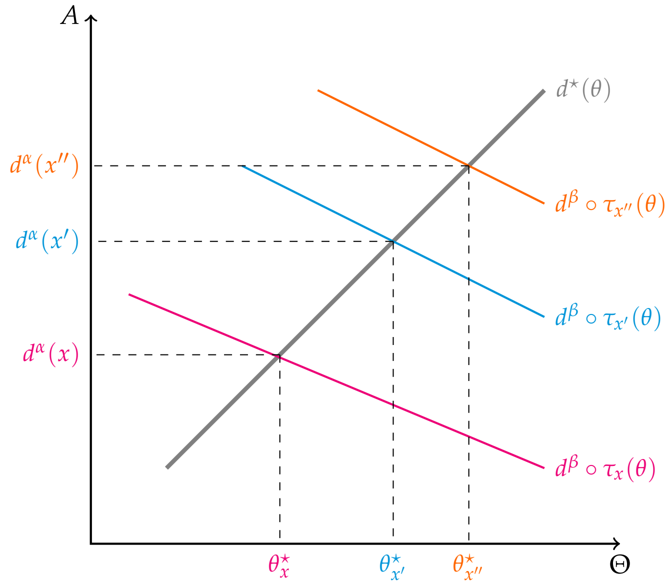


Figure 3: Improvement Principle - While the decision rule $d^\beta \circ \tau_x(\theta)$ is decreasing in θ for each x, it is increasing in x for each θ. This property leads to the increasing sequence of intersections $\{d^\alpha(x)\}_{x\in\mathcal{X}}$ with $d^\star$.

where the transformation T_y is from (P) in Definition 2.6.

PROOF OF LEMMA 3.3 : Recall that $\alpha \succ_L \beta$ requires T_y to be increasing in θ for each outcome $y \in \mathcal{Y}$. Also, for each state θ, there is a unique $x \in \mathcal{X}$ for which $x = T_y(\theta)$, since the distribution function $G^\alpha(\cdot|\theta)$ is strictly increasing due to the assumption that its density function is atomless. Since $T_y : \Theta \to \mathcal{X}$ is bijective, we can write $y = \tau_x(\theta)$ where τ_x is decreasing in θ. Hence, the inequality (4) is equivalent to

$$(5) \qquad u(d^\alpha(x), \theta) \geq u(d^\beta \circ \tau_x(\theta), \theta) \quad x \in \mathcal{X},\ \theta \in \Theta.$$

Note that since d^β is assumed to be increasing, the composition $d^\beta \circ \tau_x(\theta)$ is decreasing. Hence it follows from the preceding lemma that there exists an action $a^\star$ such that $u(a^\star, \theta) \geq u(d^\beta \circ \tau_x(\theta), \theta)$ for all states θ. For each x, $d^\alpha(x) = a^\star$ will lead to the desired inequality (5).

It remains to show that the decision rule d^α we just constructed is increasing. Note that $d^\beta \circ \tau_{x'}(\theta) \geq d^\beta \circ \tau_x(\theta)$ for all θ and for $x' > x$, since both functions d^β and τ of the composition is increasing with x. Hence the intersection with the optimal decision rule $d^\star(\theta)$ in Lemma 3.2 is also increasing (See Figure 3). It proves the lemma. □

PROOF OF THEOREM 3.1 : Note that for every monotone decision rule d^β permissible under β,

$$\begin{aligned} \rho^\beta(d^\beta, \theta) = \int u(d^\beta(y), \theta) dG^\beta(y|\theta) &\leq \int u(d^\alpha \circ T_y(\theta), \theta) dG^\beta(y|\theta) \\ &= \int u(d^\alpha(x), \theta) dG^\alpha(x|z) = \rho^\alpha(d^\alpha, \theta), \end{aligned}$$

where the inequality follows from the preceding improvement principle and the next equality follows from the change of variable $x = T_y(\theta)$, whose distribution is $G^\alpha(\cdot|\theta)$ for each θ. Theorem 2.2 allows us restrict attention to the set of monotone strategies for effectiveness. Therefore, $\alpha \succ_L \beta$ is sufficient for α to be more effective than β with respect to the class $\mathcal{U}^{sc}$. $\square$

Note that the essentially complete class theorem for $\mathcal{U}^{sc}$, due to Quah and Strulovici (2009), plays a central role in the proof of Theorem 3.1. The theorem allows us to focus on the set of monotone decision procedures, in which the task of finding a better decision rule is much easier than in the set of all permissible decision procedures. As a result, we derived a simple improvement principle in Lemma 3.3.

Along with Theorem 2.3, one immediate consequence of the preceding theorem is that the Lehmann precision induces the signal ordering based on the Bayesian values as well.

COROLLARY 3.4. *If $\alpha \succ_L \beta$, then $\alpha \succ_I^{sc} \beta$.*

Without the aid of the essentially complete class theorem, however, one can use Theorem 2.4 and Lemma 3.3 to prove the same result in a direct way, not through the effectiveness. To this end, we first infer from Theorem 2.4 that there exists a monotone decision rule d^β for which $\mathcal{V}^\pi(\beta, u) = \mathbb{E}_Y\left[\mathbb{E}_\theta[u(d^\beta(Y), \theta)|Y]\right]$. By virtue of Lemma 3.3, we can assure that there exists a decision rule d^α for which

$$\mathbb{E}_\theta[\mathbb{E}_Y[u(d^\beta(Y), \theta)|\theta]] \leq \mathbb{E}_\theta\left[\mathbb{E}_X[u(d^\alpha(X), \theta)|\theta]\right] = \mathbb{E}_X\left[\mathbb{E}_\theta[u(d^\alpha(X), \theta)|X]\right].$$

But the last expression is smaller than $\mathbb{E}_X\left[\mathbb{E}_\theta[u(\delta^\alpha(X), \theta)|X]\right] = \mathcal{V}^\pi(\alpha, u)$, since the Bayesian decision rule δ^α maximizes the expected payoff at each outcome $x \in \mathcal{X}$. Here, the key observation is that δ^α is a *pointwise* optimal action for each x.

The discussion presented so far in this subsection tells us that, for every permissible decision rule based on β, the decision maker with $u \in \mathcal{U}^{sc}$ is able to find a decision rule based on $\alpha \succ_L \beta$ that makes her better off regardless regardless of the decision principle she has in mind. Since $\mathcal{U}^{spm}$ is included in $\mathcal{U}^{sc}$, the above results also hold for the class $\mathcal{U}^{spm}$.

3.2. Informativeness implies Dispersion

Now we link the informativeness to the main concept of this paper, demonstrating that the informativeness within $\mathcal{U}^{spm}$ leads to the signal ordering based on the dispersion. Since $\mathcal{U}^{spm} \subset \mathcal{U}^{sc}$, the informativeness within $\mathcal{U}^{sc}$ induces the same signal ordering.

THEOREM 3.5. *If $\alpha \succ_I^{spm} \beta$, then $\alpha \succ_D \beta$.*

PROOF OF THEOREM 3.5 : Consider an action space $A = \{0, 1\}$. For a constant $\kappa \in \Re$ and a nondecreasing function $\psi \in \Theta^*$ defined on Θ, we define a payoff function $u_{\psi,\kappa} : A \times \Theta \to \Re$ as

$$u_{\psi,\kappa}(a, \theta) = a \cdot (\psi(\theta) - \kappa).$$

Note that the payoff function $u_{\psi,\kappa}$ is supermodular in $(a;\theta)$ for every constant κ and every $\psi \in \Theta^*$ so the family of such payoff functions $\{u_{\psi,\kappa}(\cdot,\theta)\}_{\kappa\in\Re,\ \psi\in\Theta^*,\ \theta\in\Theta}$ is a subset of the class $\mathcal{U}^{spm}$. It is a routine task to derive the (unique) Bayesian decision rule for the decision problem $(\alpha, u_{\psi,\kappa})$:

$$\delta^{\alpha}(x) \;=\; \begin{cases} 0 & \text{if } J^{\alpha}[\psi](x) \leq \kappa \\ 1 & \text{otherwise.} \end{cases}$$

Then we can write the Bayesian value of the signal α as

$$\begin{aligned} \mathcal{V}^{\pi}(\alpha, u_{\psi,\kappa}) &= \mathbb{E}_{\theta}\left[\int_{\mathcal{X}} u_{\psi,\kappa}(\delta^{\alpha}(x),\theta) dG^{\alpha}(x|\theta)\right] \\ &= \int_{\mathcal{X}} \mathbb{1}_{\{J^{\alpha}[\psi](x)>\kappa\}}(x) \cdot \left(\int_{\Theta}[\psi(\theta)-\kappa]\, dF^{\alpha}(\theta|x)\right) dM^{\alpha}(x) \\ &= \int_{\mathcal{X}} \mathbb{1}_{\{J^{\alpha}[\psi](x)>\kappa\}}(x) \cdot (J^{\alpha}[\psi](x)-\kappa)\, dM^{\alpha}(x). \end{aligned}$$

Then we use the *layer cake representation* to rewrite the bottom line as[33]

$$\text{(6)} \qquad \mathcal{V}^{\pi}(\alpha, u_{\psi,\kappa}) = \int_{\kappa}^{\infty}[1-H^{\alpha}(\xi)]d\xi,$$

where $H^{\alpha}(\xi) \triangleq \Pr\left(J^{\alpha}[\psi](X) \leq \xi\right)$ stands for the probability distribution generated by the operator $J^{\alpha}[\psi]$. Since α is more informative than β with respect to $\mathcal{U}^{spm}$, it must be the case that $\mathcal{V}^{\pi}(\alpha, u_{\psi,\kappa}) \geq \mathcal{V}^{\pi}(\beta, u_{\psi,\kappa})$ for all κ, ψ, and all prior beliefs $\pi \in P(\Theta)$. By using the representation result (6), therefore, the informativeness leads us to

$$\int_{\kappa}^{\infty}[1-H^{\alpha}(\xi)]d\xi \geq \int_{\kappa}^{\infty}[1-H^{\beta}(\xi)]d\xi, \quad \forall\, \kappa,\ \psi,\ \pi.$$

Lastly, it is immediate from Lemma 2.6 that $J^{\eta}[\psi] \geq_{\text{icx}} J^{\theta}[\psi]$ for all ψ and π. Therefore, the informativeness induces the new signal ordering $\alpha \succ_D \beta$ based on the dispersion. $\square$

3.3. Dispersion implies Precision

The next result completes the proof of the main theorem.

THEOREM 3.6. *If $\alpha \succ_D \beta$, then $\alpha \succ_L \beta$.*

PROOF OF THEOREM 3.6 : For a pair of outcomes $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, suppose that $\Delta^{x,y}(\theta_0) = G^{\beta}(y|\theta_0) - G^{\alpha}(x|\theta_0) \geq 0$ for some θ_0. We want to show that $\Delta^{x,y}(\theta) \geq 0$ for every $\theta \geq \theta_0$, which is equivalent to $\alpha \succ_L \beta$ by Lemma 2.5. To verify this claim, we suppose that there exists a state $\theta_1 > \theta_0$ at which $\Delta^{x,y}(\theta_1) < 0$, and derive a contradiction.

[33] The representation theorem (Folland (1999)) states that for a σ-finite Borel measure ν on $\Re_+ = [0,\infty)$, $\phi(t) \triangleq \nu[0,t)$, and a nonnegative measurable function f defined on Θ, we have

$$\int_{\Theta} \phi(f(\theta))\, d\pi(\theta) = \int_0^{\infty} \pi\{\theta \in \Theta \mid f(\theta) > t\} d\nu(t).$$

Let $\pi \in P(\Theta)$ be a prior belief that assigns probability π_0 and π_1 to two possible states θ_0 and θ_1, respectively. Consider a function ψ defined on Θ taking on the value 0 for $\theta \leq \widehat{\theta}$ and the value $c \in \Re$ otherwise, where $c > 0$ is an arbitrary constant and $\widehat{\theta} \in (\theta_0, \theta_1)$ is the jump point of ψ. Note that ψ is bounded, measurable, and nondecreasing for every $c > 0$, and thus it belongs to Θ^*. With the given outcome x and ψ, set $\kappa = J^\alpha[\psi](x)$, i.e., the value of the operator $J^\alpha[\psi]$ at x. Then $\kappa \leq c$ is immediate.

Recall that $\alpha \succ_D \beta$ implies $J^\alpha[\psi] \geq_{\text{icx}} J^\beta[\psi]$. Hence we have

$$\mathbb{E}_X\left[(J^\alpha[\psi](X) - \kappa) \vee 0\right] \geq \mathbb{E}_Y[(J^\beta[\psi](Y) - \kappa) \vee 0], \tag{7}$$

as $(x - \kappa) \vee 0$ is increasing and convex. With the prior belief π, however, we can rewrite the left-hand side of (7) as

$$\mathbb{E}_X\left[(J^\alpha[\psi](X) - \kappa) \vee 0\right] = \int_\Theta \int_{\mathcal{X}} \mathbb{1}_{\{\xi \geq x\}}(\xi)\left[\psi(\theta) - \kappa\right] dG^\alpha(\xi|\theta) d\pi(\theta)$$

Note that the event of the indicator function above $\{\xi \geq x\}$ is the same as the event $\{J^\alpha[\psi](\xi) \geq \kappa\}$. Thus $\mathbb{1}_{\{\xi \geq x\}}(\xi)$ has the same functional structure as the Bayesian decision rule δ^α described in the proof of the preceding theorem. Consequently, $\mathbb{E}_X\left[(J^\alpha[\psi](X) - \kappa) \vee 0\right] = \mathcal{V}^\pi(\alpha, u_{\psi,\kappa})$.

Applying the same argument to the signal β gives

$$\mathbb{E}_X\left[(J^\beta[\psi](Y) - \kappa) \vee 0\right] \geq \int_\Theta \int_{\mathcal{Y}} \mathbb{1}_{\{\xi \geq y\}}(\xi)\left[\psi(\theta) - \kappa\right] dG^\beta(\xi|\theta) d\pi(\theta),$$

since the operator $J^\beta[\psi]$ may not assume κ at the given outcome y. Simplifying the two iterated integrals above and substituting them into (7) leads to

$$\begin{aligned}
&-\pi_0\kappa(1 - G^\alpha(x|\theta_0)) + \pi_1(c - \kappa)(1 - G^\alpha(x|\theta_1)) \\
&\qquad\qquad \geq -\pi_0\kappa(1 - G^\beta(y|\theta_0)) + \pi_1(c - \kappa)(1 - G^\beta(y|\theta_1)).
\end{aligned}$$

Since $\Delta^{x,y}(\theta_0) \geq 0$ and $\Delta^{x,y}(\theta_1) < 0$ reverses the inequality above, we arrive at a contradiction to $J^\alpha[\psi](X) \geq_{\text{icx}} J^\beta[\psi](Y)$. This completes the proof. □

4. Discussion

This section will be devoted to the discussion of the main result we established in the previous section. The equivalence theorem gives rise to numerous important results both in statistical and Bayesian decision theory. Moreover, the theorem helps to consolidate several signal orderings in previous literature.

The first implication of the main theorem is that for each class of payoff functions, Lehmann precision is both necessary and sufficient for every statistician within the class to prefer one signal to another.

COROLLARY 4.1. *For each property $\star \in \{SPM, SC\}$, $\alpha \succ_L \beta$ if and only if $\alpha \succ_E^\star \beta$.*

This result is an extension of Lehmann (1988) to the two classes of payoff functions. Focusing on the class of Karlin-Rubin monotone payoff functions $\mathcal{U}^{KRm}$ (See Appendix A), he develops a statistical signal ordering—Lehmann precision "$\succ_L$"—such that α is larger than β in that order if and only if every statistician with $u \in \mathcal{U}^{KRm}$ prefers α to β. Corollary 4.1 extends his result to $\mathcal{U}^{spm}$ and $\mathcal{U}^{sc}$, respectively.

EXAMPLE 4.1 (Lehmann (1988)). *Consider a quintessential one-tail hypothesis testing problem—$H_0 : \theta \leq \theta^*$ and $H_1 : \theta > \theta^*$—for a state $\theta^* \in \Theta$. The associated primitive payoff function with this test can be written*

$$u(a_0, \theta) = \begin{cases} u_0 & \text{if } \theta \leq \theta^* \\ e_{II} & \text{otherwise} \end{cases} \quad \text{and} \quad u(a_1, \theta) = \begin{cases} u_1 & \text{if } \theta > \theta^* \\ e_I & \text{otherwise,} \end{cases}$$

where e_I and e_{II} ($e_I, e_{II} < 0$) represent disutility from committing type I and II error, respectively, and the action a_0 and a_1 represent admitting and rejecting the null hypothesis, respectively. In the action space $A = \{a_0, a_1\}$ endowed with an order structure $a_1 > a_0$, it can be easily shown that this payoff function belongs to $\mathcal{U}^{spm}$, the smallest class of payoff functions we consider.

Due to Theorem 2.2, we can restrict attentions to the set of monotone decision rules, and each $d \in \mathcal{D}^{\alpha,M}$ can be described by a single point x in $\mathcal{X}$ at which the assigned action changes from a_0 to a_1. Hence, given a decision rule $d = \{x\}$ and signal α, we can write the expected payoffs of d at θ as

$$\rho^\alpha(d, \theta) = \begin{cases} u_0 G^\alpha(x|\theta) + e_I[1 - G^\alpha(x|\theta)] & \text{if } \theta \leq \theta^*, \\ e_{II} G^\alpha(x|\theta) + u_1[1 - G^\alpha(x|\theta)] & \text{if } \theta > \theta^*. \end{cases}$$

In order α to be more effective than β in this decision problem, namely, in order for $\rho^\alpha(d^\alpha, \theta) \geq \rho^\beta(d, \theta)$ for all $d \in \mathcal{D}^{\beta,M}$, there should exist a $x^* \in \mathcal{X}$, for each $y \in \mathcal{Y}$, such that

$$G^\alpha(x^*|\theta) \geq G^\beta(y|\theta) \;\; \forall\, \theta \leq \theta^* \;\; \text{and} \;\; G^\alpha(x^*|\theta) \leq G^\beta(y|\theta) \;\; \forall\, \theta > \theta^*, \tag{8}$$

which is equivalent to $\alpha \succ_L \beta$ by Lemma 2.5. Therefore, Lehmann precision is necessary for satisfying the criterion of effectiveness with respect to $\mathcal{U}^{spm}$. $\square$

It follows from Lemma 2.1 that α generates more statistical values $V(\alpha, u)$ than β if and only if α is more effective than β. Therefore, Corollary 4.1 tells us that $V(\alpha, u) \geq V(\beta, u)$ for all $u \in \mathcal{U}^{spm}$ if and only if the same inequality holds for all $u \in \mathcal{U}^{sc}$. Recall that $\mathcal{U}^{sc}$ is a strict superset of $\mathcal{U}^{spm}$, which makes this result somewhat contrary to the intuition that when we narrow down the scope of payoff functions, we can rank signals with a less restrictive condition.

To understand the insight behind this result, consider a simple decision problem with the payoff function $u(a_0, \theta) = 0$ and $u(a_1, \theta) = {}^u\Delta(\theta)$ defined on two possible actions. For α to be more effective than β for this payoff function with ${}^u\Delta(\theta)$ nondecreasing, it must hold that

$$\rho^\alpha(d^\alpha, \theta) = {}^u\Delta(\theta)[1 - G^\alpha(x|\theta)] \geq {}^u\Delta(\theta)[1 - G^\beta(y|\theta)] = \rho^\beta(d^\beta, \theta) \quad \text{for every } \theta,$$

equivalent to the condition (8). Hence, even if ${}^{u}\Delta(\theta)$ satisfies the single-crossing property, we arrive at the same condition since the single-crossing property of ${}^{u}\Delta(\theta)$ solely cannot reverse the above inequality.

Relating (i) Lehmann precision to (iii) the criterion of informativeness from the Bayesian perspective, we can obtain the second implication of the main theorem:

COROLLARY 4.2. *For each property* $\star \in \{SPM, SC\}$, $\alpha \succ_L \beta$ *if and only if* $\alpha \succ_I^\star \beta$.

Although the statistical value of a signal is generically different from its Bayesian value as discussed in Section 2, Lehmann precision gives a uniform signal ordering necessary and sufficient for representing a Bayesian decision maker's preference on the signal space.

The equivalence between $\alpha \succ_L \beta$ and informativeness within $\mathcal{U}^{spm}$ is worthy of remark. Athey and Levin (2001) explore the value of information in a Bayesian decision problem with a fixed prior belief $\pi \in P(\Theta)$ and $\mathcal{U}^{spm}$. They develop another signal ordering based on posterior beliefs, named "MIO-ND" in their paper, sufficient and necessary for α to generate more Bayesian values than β within $\mathcal{U}^{spm}$.[34] Formally, given a prior belief π, $\mathcal{V}^\pi(\alpha, u) \geq \mathcal{V}^\pi(\beta, u)$ for every $u \in \mathcal{U}^{spm}$ if and only if $\alpha \succ^\pi_{\text{MIO-ND}} \beta$, defined by

$$\text{(MIO-ND)} \quad \alpha \succ^\pi_{\text{MIO-ND}} \beta \ \text{ if } \ F^\alpha(t|M^\alpha(x) > c) \leq F^\beta(t|M^\beta(y) > c), \quad \forall\, t \in \Theta \text{ and } \forall\, c \in [0,1].$$

In words, the posterior distribution of the event $\{\theta \leq t\}$ conditional on the event $M^\alpha(x) > c$ dominates the posterior distribution of the same event conditional on $M^\beta(y) > c$ in the FOSD.[35] As a result, given that a decision maker receives a large signal outcome, the posterior distribution generated by α will assign more probabilities to high states compared to β.

To comprehend the relationship between (MIO-ND) condition and Lehmann precision, observe that $\alpha \succ_L \beta$ guarantees $\mathcal{V}^\pi(\alpha, u) \geq \mathcal{V}^\pi(\beta, u)$ for every $u \in \mathcal{U}^{spm}$ *and* for every π. Hence we infer from Corollary 4.2 that (MIO-ND) is implied by Lehmann precision.

PROPOSITION 4.3. *For every* $\pi \in P(\Theta)$, $\alpha \succ_L \beta$ *implies* $\alpha \succ^\pi_{MIO-ND} \beta$.

PROOF OF PROPOSITION 4.3 : Notice that the event $\{x \in \mathcal{X} | M^\alpha(x) > c\}$ has the probability of $1-c$ for each $c \in [0,1]$. Let $x^\alpha_c = \sup\{x \in \mathcal{X} | M^\alpha(x) \leq c\}$. Then we use Fubini's Theorem and Bayes' rule to rewrite the posterior distribution in (MIO-ND) as

$$F^\alpha(t|M^\alpha(x) > c) = \frac{1}{1-c} \int_\Theta \mathbb{1}_{\{\theta \leq t\}}(\theta) \cdot [1 - G^\alpha(x^\alpha_c|\theta)]\, d\pi(\theta).$$

Hence the difference of the two posteriors $F^\alpha - F^\beta$ reduces into

$$\frac{1}{1-c} \int_\Theta \mathbb{1}_{\{\theta \leq t\}}(\theta) \cdot \left[G^\beta(y^\beta_c|\theta) - G^\alpha(x^\alpha_c|\theta)\right] d\pi(\theta).$$

[34] (MIO-ND) is the abbreviation of "Monotone Information Order for payoff functions with NonDecreasing incremental returns". Note that we need attach the superscript π to the order, unlike Lehmann precision, since their statistical condition is dependent upon the decision maker's prior beliefs. Thus, applications of their result call for exact knowledge of π.

[35] Jewitt (2007) shows that the signal ordering (MIO-ND) is equivalent to another signal ordering based on a statistical notion—"*concordance*"—developed by Tchen (1980).

We infer from Lemma 2.5 that the expression in the bracket satisfies the single-crossing property in θ and from the definition of x_c^α and x_c^β that

$$\int_\Theta \left[G^\beta(y_c^\beta|\theta) - G^\alpha(x_c^\alpha|\theta)\right] d\pi(\theta) = M^\beta(y_c^\beta) - M^\alpha(x_c^\alpha) = 0.$$

Additionally, the indicator function $\mathbb{1}_{\{\theta \le t\}}(\theta)$ is nonincreasing in θ. Hence it follows from the *folk single-crossing lemma* that $F^\alpha - F^\beta$ is nonpositive.[36] $\square$

Conversely, if (MIO-ND) holds for every prior beliefs π in $P(\Theta)$, the criterion of informativeness within $\mathcal{U}^{spm}$ is satisfied. Hence Corollary 4.2 leads us to the result that if $\alpha \succ^\pi_{\text{MIO-ND}} \beta$ for every π, α must be more Lehmann precise than β. In light of this observation and the preceding proposition, we obtain the equivalence between Lehmann precision and (MIO-ND).

COROLLARY 4.4. *$\alpha \succ^\pi_{\text{MIO-ND}} \beta$ for every $\pi \in P(\Theta)$ if and only if $\alpha \succ_L \beta$.*

Just as we discussed above, Corollary 4.2 tells us that α is more valuable than β for every Bayesian decision maker with $u \in \mathcal{U}^{spm}$ if and only if the same relation holds for everyone with $u \in \mathcal{U}^{sc}$. While the implication $\alpha \succ^{sc}_I \beta \Rightarrow \alpha \succ^{spm}_I \beta$ is trivial, its converse is somewhat surprising.

From Bayesian perspectives, their equivalence is predicated on the idea that the enriched set of prior beliefs *nullifies* the difference between $\mathcal{U}^{spm}$ and $\mathcal{U}^{sc}$ when comparing two signals based on their Bayesian values. To be more concrete, go back to the simple decision problem with $u(a_0, \theta) = 0$ and $u(a_1, \theta) = {}^u\Delta(\theta)$. Suppose that α is more informative than β for every nondecreasing function ${}^u\Delta(\theta)$. By virtue of Theorem 2.4, It can be stated that for every $y \in \mathcal{Y}$, there exists a $x \in \mathcal{X}$ such that

$$\int_\Theta {}^u\Delta(\theta)[1 - G^\alpha(x|\theta)]d\pi(\theta) \ge \int_\Theta {}^u\Delta(\theta)[1 - G^\beta(y|\theta)]d\pi(\theta) \quad \forall\, \pi \in P(\Theta),$$

so we arrive at the same condition (8). Therefore, if there is a ${}^u\widehat{\Delta}(\theta)$ crossing the axis at some point $\widehat{\theta}$ but violating the inequality above, there exists a ${}^u\Delta(\theta)$ nondecreasing, crossing the axis at the same point $\widehat{\theta}$ and violating the inequality.

EXAMPLE 4.2 (Portfolio Decision Problem). *Suppose that there are two assets, a safe asset with a return of s per dollar invested and a set of risky assets $\mathcal{R} = \{r : \Theta \to \Re | \text{continuously differentiable}\}$ where the unknown return of each risky asset is denoted $r(\theta) \in \mathcal{R}$ per dollar invested. Consider an investor having initial wealth w to invest and an increasing Bernoulli utility function v. Let $a \in [0, w]$ denote the amounts of money invested in the risky asset with $r \in \mathcal{R}$. Then his utility function is written $u_r(a, \theta) = v(a[r(\theta) - s] + ws)$. Within this family of decision problems, $\mathcal{U}^{sc}$ is characterized by the set of risky assets $\mathcal{R}^{sc} = \{r(\theta) | r(\theta) - s \text{ obeys the single-crossing property in } \theta\}$ and $\mathcal{U}^{spm}$ by $\mathcal{R}^{spm} = \{r(\theta) | \partial^2 u_r / \partial a \partial \theta \ge 0\}$ due to Topkis (1978).*

[36] The folk single-crossing lemma states that if a function $h : \Theta \to \Re$ satisfies the single-crossing property in θ and

$$\int_\Theta h\, d\pi = 0 \text{ for a measure } \pi \text{ on } (\Theta, \mathcal{F}), \text{ then } \int_\Theta \phi_1 h\, d\pi \ge 0 \text{ and } \int_\Theta \phi_2 h\, d\pi \le 0$$

for every $\phi_1 : \Theta \to \Re$ nondecreasing and $\phi_2 : \Theta \to \Re$ nonincreasing.

Suppose that the investor receives a signal α or β by hiring an investment consultant with expertise. By Corollary 4.2, if the investor faces a risky asset with the return $\widehat{r}(\theta) \in \mathcal{R}^{sc}$ but prefers β to α, there exists another investor facing a risky asset with $r(\theta) \in \mathcal{R}^{spm}$ who also prefers β to α, and vice versa. $\square$

The next result we can derive from the main theorem is the heart of a characterization of Lehmann precision, which we call the dispersion theorem hereafter.

COROLLARY 4.5 (Dispersion Theorem). *$\alpha \succ_L \beta$ if and only if $\alpha \succ_D \beta$, i.e., for every increasing and convex function σ,*

(ICX) $\mathbb{E}_X \left[\sigma \circ J^{\alpha}[\psi](X)\right] \geq \mathbb{E}_Y[\sigma \circ J^{\beta}[\psi](Y)]$

Corollary 4.5 provides a characterization of Lehmann-precision in terms of the dispersion: The signal α is more statistically precise than β if and only if the former generates a more dispersed prediction about θ than the latter. Intuitively, the decision maker with α puts more weights on the upper or lower tails of the distribution of θ than the one with β, so the distributions of the linear operator $J^{\alpha}[\psi]$, generated by every $\psi \in \Theta^*$, will be more spread out than those of $J^{\beta}[\psi]$, since the monotone property of the function $\psi \in \Theta^*$ will preserve this dissemination motion of θ to each tail.

Theorem 5.2 in Lehmann (1988) provides another characterization result of $\alpha \succ_L \beta$ based on the dispersion, but the result is specialized to the location problem with log-concave densities where the distributions of signals take a form of $G^{\alpha}(x|\theta) = G^{\alpha}(x - \theta)$ and $G^{\beta}(y|\theta) = G^{\beta}(y - \theta)$ and their densities are log-concave in x for each θ. Since the above dispersion theorem does not utilize any additional structures on the signal's primitive distributions, however, it is applicable to more distributions than the Lehmann's result.

The dispersion theorem furnishes us with the link between Lehmann precision and several signal orderings, based on the dispersion of a certain variable, in previous literature. This link can be found by making a clever choice of the nondecreasing function $\psi : \Theta \to \Re$. For instance, by setting $\psi(\theta) = \theta$, the theorem along with the law of iterated expectation—$\mathbb{E}_X[J^{\alpha}[\theta](X)] = \mathbb{E}_X[\mathbb{E}_{\theta}[\theta|X]] = \mathbb{E}_Y[\mathbb{E}_{\theta}[\theta|Y]]$—gives the second-order stochastic dominance of conditional expectations:[37]

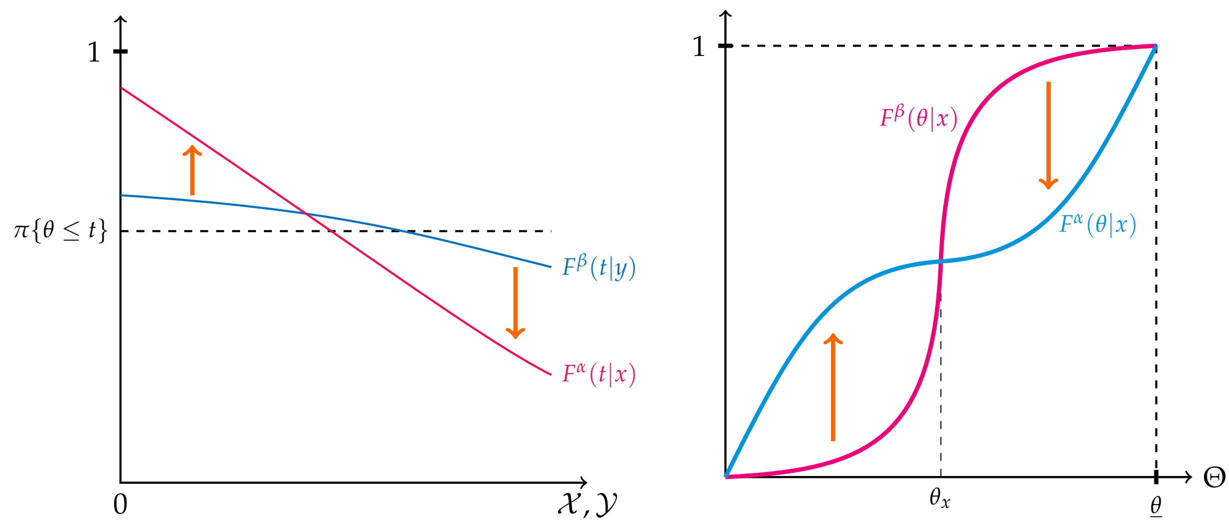
COROLLARY 4.6 (Conditional Expectations). *If $\alpha \succ_L \beta$, $\mathbb{E}[\theta|X]$ is dominated by $\mathbb{E}[\theta|Y]$ in the second-order stochastic dominance. That is, $\mathbb{E}[\theta|X] \geq_{\text{cx}} \mathbb{E}[\theta|Y]$.*

Thus, the signal ordering based on the dispersion of conditional expectations is induced by Lehmann precision. Using this ordering, Ganuza and Penalva (2010) study the effect of releasing more precise information by the seller on efficient allocation and the seller's expected revenue in a symmetric second-price auction.[38] We use the dispersion theorem to revisit their questions in Section 5.1 and generalize some results to an asymmetric second-price auction.

The theorem can be utilized for the connection between Lehmann precision and the dispersion of posterior beliefs. As a signal becomes more precise in Lehmann's sense, it generates

[37]Some alternative proofs can be found in Mizuno (2006) and Hernando-Veciana (2009).

[38]The signal ordering is named "*integral precision*" in their paper. That is, α is more integral precise than β if $\mathbb{E}[\theta|X] \geq_{\text{cx}} \mathbb{E}[\theta|Y]$.



(a) MPS of Posterior Beliefs

(b) Rotation Orders

Figure 4: Mean Preserving Spread of Posterior Beliefs and Rotation Orders of Johnson and Myatt (2006) - In subfigure (a), $\pi\{\theta \leq t\}$ indicates the prior beliefs of the event $\{\theta \leq t\}$. The two curves display the posterior beliefs of the same event as a function of outcomes x and y. In subfigure (b), however, the two curves display the posterior beliefs as a function of θ conditional on the same outcome x.

more variable posterior beliefs on Θ as depicted in Figure 4-(a). Put differently, when the decision maker observes an outcome from a less precise signal, her posterior beliefs of the event $\{\theta \leq t\}$ would not make a big difference with her prior beliefs $\pi\{\theta \leq t\}$.

COROLLARY 4.7 (Posterior Beliefs). *If $\alpha \succ_L \beta$, $F^{\alpha}(\theta|\cdot)$ is the mean-preserving spread of $F^{\beta}(\theta|\cdot)$ for each $\theta \in \Theta$.*

PROOF OF COROLLARY 4.7 : Define $\psi(\theta) = \mathbb{1}_{\{\theta \geq t\}}(\theta)$ for each $t \in \Theta$. Note that the indicator function is nondecreasing in θ and the linear operator $J^{\alpha}[\psi](x) = 1 - F^{\alpha}(t|x)$ gives rise to the posterior beliefs of $\{\theta \leq t\}$. In this case, the dispersion theorem presents

$$1 - F^{\alpha}(t|X) \geq_{\text{icx}} 1 - F^{\beta}(t|Y), \text{ equivalently, } F^{\alpha}(t|X) \leq_{\text{icv}} F^{\beta}(t|Y).$$

See footnote 27. Since $\mathbb{E}_X[F^{\alpha}(t|X)] = \mathbb{E}_Y[F^{\beta}(t|Y)] = \pi\{\theta \leq t\}$ by Bayes' rule, the above increasing concave order reduces to $F^{\alpha}(t|X) \leq_{\text{cv}} F^{\beta}(t|Y)$. □

Ranking signals based on the dispersion of posterior beliefs, Lewis and Sappington (1994) study the incentive of a monopolist to provide more accurate private information with potential buyers. The preceding result verifies that Lehmann precision is sufficient for inducing their signal ordering.

This result also gives a link to the rotation order which has been developed by Johnson and Myatt (2006) and applied by Shi (2012) for characterizing an optimal auction with endogenous information. Formally, $\alpha \succ_{RO} \beta$ requires that the posterior beliefs rotate clockwise on some point θ_x as we change the signal from β to α, that is, for each $x \in \mathcal{X} \cap \mathcal{Y}$, there exists a point

$\theta_x \in \Theta$ such that

(RO) ${}^F\Delta_x(\theta) \triangleq F^\beta(\theta|x) - F^\alpha(\theta|x)$ crosses the horizontal axis once from below at $\theta = \theta_x$.

Refer to Figure 4-(b). Since $\alpha \succ_L \beta$ implies $F^\beta(\theta|\cdot) \leq_{\text{cx}} F^\alpha(\theta|\cdot)$ only, the rotation order is stronger than Lehmann precision.

In addition, The third corollary of the dispersion theorem concerns the distribution of likelihood ratio. To ascertain the impact on the likelihood ratio, we note that, given a signal α, if the density function $g^\alpha(x|\theta)$ satisfies the MLRP, then the posterior density function $f^\alpha(\theta|x)$ also exhibits the MLRP. To see this, for every $x' > x$, Bayes' rule gives an alternative expression of the likelihood ratio function

$$\frac{f^\alpha(\theta|x')}{f^\alpha(\theta|x)} = \frac{m^\alpha(x)}{m^\alpha(x')} \cdot \frac{g^\alpha(x'|\theta)}{g^\alpha(x|\theta)}$$

Hence both of the likelihood ratios move in the same direction as θ changes. For this result only, we assume that f^α is differentiable with respect to x and let f^α_x denote the derivative. Then substituting nondecreasing function $\psi(\theta) = \mathbb{1}_{\{f^\alpha_x/f^\alpha(\theta|x) \geq \kappa\}}$ into (ICX) presents the next result :

COROLLARY 4.8 (Likelihood Ratio). *If $\alpha \succ_L \beta$, the distribution of the likelihood ratio with α*

$$L^\alpha(\kappa) \triangleq \text{PR}\left(\frac{f^\alpha_x}{f^\alpha}(\theta|x) \leq \kappa\right)$$

is the mean-preserving spread of $L^\beta(\kappa)$.

In a principal-agent model with transferable utilities, Kim (1995) asserts that, in order for the principal to control the agent's hidden action more efficiently, she has to offer an incentive scheme based on the contractible variable X with better prediction about the likelihood ratio, and provides an appropriate order on the experiment tailored to the bilateral contract problem. Although we will examine the exact connection between Lehmann preciseness and his criterion in Section 5.2, the corollary captures a simple idea that, under more precise experiment, the principal is able to predict the likelihood ratio (hence, the agent's hidden action) more accurately.

Lastly, the dispersion theorem provides an alternative way to prove the equivalence between Lehmann precision and informativeness for $\mathcal{U}^{spm}$. The theorem enables us to present a simple proof without the aid of the improvement principle even, although the result is immediate from Corollary 3.4 as $\mathcal{U}^{spm} \subset \mathcal{U}^{sc}$. Furthermore, it plays a key role in proving the necessity of $\alpha \succ_L \beta$ and thus in establishing the equivalence of ordering experiments.

COROLLARY 4.9. *The signal α is more informative than β with respect to the class of decision problems $\mathcal{U}^{spm}$ if and only if $\alpha \succ_L \beta$.*

PROOF OF COROLLARY 4.9 : For sufficiency, see Appendix C. Here we establish the necessity of $\alpha \succ_L \beta$ for being more informative within $\mathcal{U}^{spm}$. Note that the payoff function $u_{\psi,\kappa}$ presented in the proof of Theorem 3.5 belongs to $\mathcal{U}^{spm}$ for every $\psi \in \Theta^*$ and $\kappa \in \Re$. The definition of α

being more informative than β for the decision maker with $u_{\psi,\kappa}$ implies $J^{\eta}[\psi](X) \geq_{\text{icx}} J^{\theta}[\psi](Y)$ for every $\psi \in \Theta^*$, which is equivalent to $\alpha \succ_L \beta$ by Corollary 4.5. $\square$

5. Applications

In this section we use the results established above to study the value of precise information in several important economic environments with strategic interactions; auctions, bilateral contracts, and delegation.

5.1. Auctions

The aim of this subsection is to analyze the value of precise information to an auctioneer when she is able to supply more precise private information on the object for sale with bidders and such an information disclosure is observable to every participant.[39] The question is closely related to Ganuza and Penalva (2010). We use the dispersion theorem to extend their results to the asymmetric auction in this subsection. Before proceeding, one should note that, unlike the *covert* information acquisition, gathering precise information is not necessarily valuable to each participant in case such activity is observable, because the change in distribution can alter the bidding strategies of others; this second effect may hurt himself in the end.[40]

Consider an allocation problem where a single indivisible object is auctioned off to N risk-neutral bidders, indexed by i, through the second-price auction.[41] The true *private value* of the object to bidder i is denoted by $v_i(s, \theta_i)$, where $\theta_i \in \Theta_i$ represents his idiosyncratic component reflecting his preference and $s \in S$ [to be added]. We assume that $v_i(s, \theta_i)$ is nondecreasing in θ_i and unknown till he wins the auction. The prior distribution $\pi(\theta_1, \cdots, \theta_N)$ over the space $\Theta_1 \times \cdots \times \Theta_N$ is common knowledge among players, and it satisfies *independence* across bidders. That is,

$$\pi(\theta_1, \cdots, \theta_N) = \prod_{i \in N} \pi_i(\theta_i).$$

Prior to bidding, each bidder observes an outcome $x_i \in \mathcal{X}_i$ affiliated with θ_i. The signal's cumulative distribution is $G_i^{\alpha}(x_i|\theta_i)$, where η_i measures precision level of the signal x_i for bidder i, and the posterior beliefs about θ_i conditional on $X_i = x_i$ is $F_i^{\alpha_i}(\theta_i|x_i)$. The winning bidder's

[39] Literature on auctions with endogenous information is voluminous. Refer to Bergemann and Välimäki (2006) and its reference for survey. Following Shi (2012), the related literature segments broadly into four groups according to whether the information acquisition activity is centralized or decentralized, and according to whether the activity is covert or overt. Our setup falls into the category where the information acquisition is centralized—the information structure is controlled by the seller—and the acquisition is overt. For the recent related papers, see Bergemann and Pesendorfer (2007) and Ganuza and Penalva (2010).

[40] In Appendix D, the paper examines the decentralized and covert information acquisition case in a general mechanism design setting and proves that the precise information is always valuable in every incentive compatible auction.

[41] When the information acquisition activity is overt some results depend on the number of bidders. See Proposition 5.3. When the activity is covert, however, the results are independent of the number of bidders so one can focus on $N = 2$ case—like Persico (2000)—without loss of generality.

primitive payoff function is quasilinear: $u_i(v_i, t_i) = v_i - t_i$, where t_i is a money transfer to the auctioneer.

We assume that the auctioneer is able to select a N-tuple vector $\alpha = (\alpha_1, \cdots, \alpha_N)$ by releasing some information about the object. We call this activity the auctioneer's information policy in this section.[42] It is common knowledge that such information release is observed by every bidder, and thus the information structure $\{G^\alpha(x_i|\theta_i), \mu_i\}_{i \in N}$ of the environment is common knowledge among the bidders and the auctioneer. Examples of such information disclosures we have in mind include advertising, self-inspection, and marketing. Their primary function is to allow the bidders privately to learn of their personal match with the product, i.e., to receive more accurate information about θ_i.[43,44]

Our first result addresses the effect of the information policy on efficiency. To state the result, we define by $v_i^\alpha(\cdot)$ the bidder i's value estimate under the information policy α. Since, in the second price auction, the unique dominant symmetric strategy for each bidder is to bid his own value estimate, we can write the ex-ante surplus generated by α as

$$\mathcal{E}(\alpha) \triangleq \mathbb{E}\left[\max\{v_1^\alpha, \cdots, v_N^\alpha\}\right] \tag{9}$$

I wish to remark that the following result also holds in every symmetric auction with independent private value : [45]

PROPOSITION 5.1 (Efficiency). *More Lehmann-precise information generates more ex ante surplus.*

The proof is immediate from the next lemma:

LEMMA 5.2 (Shaked and Shanthikumar (2007)). *Let $\{X_1, \cdots, X_N\}$ be a collection of independent random variables. Let $\{Y_1, \cdots, Y_N\}$ be another collection of independent random variables. Suppose $X_i \geq_{icx} Y_i$ for each i. Then we have*

$$\sigma(X_1, \cdots, X_N) \geq_{icx} \sigma(Y_1, \cdots, Y_N)$$

for every increasing and componentwise convex function $\sigma : \Re^N \to \Re$.

PROOF OF LEMMA 5.2 : We proceed by induction on N. For $N = 1$, the result is straightforward. In order to advance the induction step, we prove that if the result holds for $N = k - 1$ it

[42] In this case the information policy η is more-Lehmann precise than θ (denoted $\eta \succ_L \theta$) if $\alpha_i \succ_L \beta_i$ for all i.

[43] The assumption on the independent prior can be justified in the situation where the "objective quality" of the product is unambiguously well-known to every bidder but its "subjective quality" is uncertain.

[44] As pointed out by Johnson and Myatt (2006), most advertising inherits the two characteristics: (1) *promotional hype*; it highlights the product's existence, price, availability, and any objective quality, and increases the bidder's value estimate in the sense of First-Order Stochastic Dominance; (2) *provision of real information*; it helps the bidders to evaluate their subjective preferences. Thus, we concentrate on its second role.

[45] The symmetric auction implies that every bidder is symmetric in the sense that the common prior and information structures are identical across bidders, i.e., $\pi_i = \pi$ and $G_i^\alpha(x_i|\theta_i) = G^\alpha(x_i|\theta_i)$ for all $i \in N$ and it is common knowledge. In such a symmetric environment, it is natural to focus on the equilibrium where each bidder adopts a symmetric bidding strategy $b_i(\cdot) = b(\cdot)$, that is increasing. Hence the bidder placing the highest bid will be awarded the object, and therefore the ex ante social surplus takes the same form as (9).

also holds for $N = k$. To this end, let $\psi : \Re \to \Re$ be increasing and convex. Note that

$$\begin{aligned}\mathbb{E}[\psi(\sigma(X_1, X_2, \cdots, X_k))|X_1 = x] &= \mathbb{E}[\psi(\sigma(x, X_2, \cdots, X_k))] \\ &\geq \mathbb{E}[\psi(\sigma(x, Y_2, \cdots, Y_k))] \\ &= \mathbb{E}[\psi(\sigma(X_1, Y_2, \cdots, Y_k))|X_1 = x].\end{aligned}$$

The first equality follows from the independence assumption and the inequality from the induction assumption. Integrating over X_1 gives $\sigma(X_1, X_2, \cdots, X_k) \geq_{\text{icx}} \sigma(X_1, Y_2, \cdots, Y_k)$. To complete the proof, we take the same steps but now condition on $Y_2, \cdots, Y_k$. Since we have already seen that the result holds for $N = 1$, we have

$$\sigma(X_1, X_2, \cdots, X_k) \geq_{\text{icx}} \sigma(X_1, Y_2, \cdots, Y_k) \geq_{\text{icx}} \sigma(Y_1, Y_2, \cdots, Y_k).$$

The proof is complete. □

PROOF OF PROPOSITION 5.1 : Suppose $\alpha \succ_L \beta$. For each bidder i, since $v_i(s, \theta_i)$ is assumed to be increasing in θ_i, $v_i^\alpha \geq_{\text{icx}} v_i^\beta$ follows from Theorem 4.5. Note that the function $\max_{i \in N}\{v_1^\alpha, \cdots, v_N^\alpha\}$ is increasing and componentwise convex. Since the ICX order is preserved by such a function, we obtain

$$\max_{i \in N} \{v_1^\alpha, \cdots, v_N^\alpha\} \geq_{\text{icx}} \max_{i \in N} \{v_1^\beta, \cdots, v_N^\beta\}.$$

Taking expectation of both sides gives $\mathcal{E}(\alpha) \geq \mathcal{E}(\beta)$. □

The intuition behind the result is quite straightforward: More precise information improves the efficient allocation by increasing the probability that the object is awarded to the bidder with the highest true value, so the total surplus increases.

Unlike the efficiency, on the other hand, the effect on the expected revenue is ambiguous. To go into the details, note that the revenue is determined by the second-highest value estimate in the second-price auction. However, we cannot arrange it in a variability order unlike the highest value estimate; the variability order of the distribution of the second-order statistic varies depending on the number of bidders and the information structures. Put it in another way, the dissemination of precise information increases the total surplus as we discussed above, but it also typically increases each bidder's information rents since his value estimate is more dispersed.[46] Since the revenue is determined by their difference but they increase with different ratios, we need figure out which one grows faster to study the effect on the revenue.

We should make a remark regarding the linkage principle of Milgrom and Weber (1982). The principle demonstrates that when bidders' true values are affiliated with additional information obtained by the seller, public disclosures of such information would increase the expected revenue in every standard auction. This striking result (*"Honesty is the best policy."*) is predicated upon the idea that, by releasing the additional information publicly, the auction-

[46]The idea of relating high dispersion to more information rents can be traced back to Lewis and Sappington (1994).

eer can have each bidder's value estimates get closer to others'. As a result, the auction will generate more intense Bertrand price competition between the bidders, which results in higher expected revenue.[47] However, our model assumes the independent information structures across bidders' true values. So more information causes the value estimates to move apart and thus enhances privacy of each bidder's information. Therefore, it raises the bidder's information rent so the expected revenue can go either way in the independent private value model.

The discussion above provides an incentive for the auctioneer to promote more participation, since a large number of bidders helps to reduce the winning bidder's information rent. In light of this fact, Ganuza and Penalva (2010) found that the auctioneer can benefit from releasing more precise information in the second-price auction when the number of bidders is sufficiently large. In contrast, the revenue (denoted by $R(\alpha)$ hereafter) is decreasing in α when there are only two bidders as the next result shows:

PROPOSITION 5.3 (Revenue I). *In case $N = 2$, more Lehmann-precise information is detrimental to expected revenue.*

PROOF OF PROPOSITION 5.3 : Since $v_i^\alpha \geq_{\text{icx}} v_i^\beta$ and their ex ante means are the same by the law of iterated expectation, we have $v_i^\alpha \geq_{\text{cx}} v_i^\beta$ for each $i = 1, 2$, equivalently, $v_i^\alpha \leq_{\text{cv}} v_i^\beta$. It implies $v_i^\alpha \leq_{\text{icv}} v_i^\beta$. Therefore, by the counterpart of Lemma 5.2 to the ICV order, we have

$$\min\{v_1^\alpha, v_2^\alpha\} \leq_{\text{icv}} \min\{v_1^\beta, v_2^\beta\}.$$

Taking expectation of both sides gives the result. □

The next example shows that for any number of bidders, the auctioneer's expected profit can fall as she releases more precise information.[48]

EXAMPLE 5.1. *Let $\Theta_i = \{\theta_1, \theta_2, \theta_3\}$ for all i, and $v(\theta_1) = 0$, $v(\theta_2) = 0.9$, and $v(\theta_3) = 1.1$. The bidders' common prior is $\pi\{\theta_1\} = 1 - \epsilon$ and $\pi\{\theta_2\} = \pi\{\theta_3\} = \epsilon/2$ for $\epsilon > 0$. With the auctioneer's information policy β, each bidder has a partition on Θ $\{\{\theta_1\}, \{\theta_2, \theta_3\}\}$. The partition corresponding to information policy α is $\{\{\theta_1\}, \{\theta_2\}, \{\theta_3\}\}$. That is,*

$$g^\alpha(x_k|\theta_k) = 1 \text{ for each } k = 1,2,3, \quad \text{and} \quad g^\beta(y_1|\theta_1) = 1,\ g^\beta(y_k|\theta_l) = \frac{1}{2} \text{ for } k,l = 2,3.$$

Thus, while each bidder receives a perfectly informative signal under α, he cannot distinguish θ_2 and θ_3 when he receives y_2 or y_3. Note that we can write the expected profit generated by β as

$$R(\beta) = \sum_{n=2}^{N} p(n) \cdot \mathbb{E}[R|n, \beta] = \sum_{n=2}^{N} \binom{N}{n} \epsilon^n (1-\epsilon)^{N-n},$$

where $p(n)$ is the probability that n of N bidders receive an imperfect signal—either y_2 or y_3—and $\mathbb{E}[R|n, \beta]$ is the conditional expected revenue given the event n out of N bidders receive the imperfect

[47]In the words of McLean and Postlewaite (2002), the public disclosure of additional information can reduce the *informational size* of the private signal for each bidder, so it alleviates the incentive compatibility for truthful revelation.

[48]I am deeply indebted to Daniel Quint for this example.

signals. Since the bidders cannot distinguish θ_2 with θ_3, they will place a bid as much as the expected value estimate 1. So $\mathbb{E}[R|n] = 1$ for all $n \geq 2$. Similarly, the expected revenue under η can be written

$$R(\alpha) = \sum_{n=2}^{N} p(n) \cdot \mathbb{E}[R|n, \alpha] = \sum_{n=2}^{N} \begin{pmatrix} N \\ n \end{pmatrix} \epsilon^n (1-\epsilon)^{N-n} \cdot \mathbb{E}[R|n, \alpha].$$

Note that $\mathbb{E}[R|2, \alpha] = 0.95 < \mathbb{E}[R|2, \beta] = 1$, $\mathbb{E}[R|3, \alpha] = \mathbb{E}[R|3, \beta] = 1$, and $\mathbb{E}[R|n, \alpha] > \mathbb{E}[R|n, \beta]$ for $n \geq 4$. Therefore, for arbitrary number of bidders N, there exists a sufficiently small $\bar{\epsilon}$ such that $\epsilon < \bar{\epsilon}$ leads to $R(\beta) > R(\alpha)$. [49] $\square$

Lastly, we explore the effect on the bidders' information rents in case $N \geq 3$.[50] While it seems quite clear that more precise information increases the rent, Lehmann precision is not enough for monotone information rents in case of overt information acquisition. To see this, we first dissect the bidder's information rents. Given the information policy α, let $Y_i^\alpha = \max_{j \neq i}\{v_j^\alpha\}$ denote the highest value estimates among bidder i's opponents and $H_i^\alpha(y) = \Pr(Y_i^\alpha \leq y)$ its distribution function. With slight abuse of notations, let $F_i^\alpha(v) = \Pr(v_i^\alpha \leq v)$ denote the distribution of bidder i's value estimate. We can decompose the bidder i's expected payoff generated by α, denoted $\mathcal{U}_i(\alpha)$, into three components as follows:

$$\begin{aligned} U_i(\alpha) \triangleq \mathbb{E}\left[(v_i^\alpha - Y_i^\alpha) \vee 0\right] &= \int \left\{ \int_{\{v \geq y\}} (1 - F_i^\alpha(v)) dv \right\} dH_i^\alpha(y) \\ &= \int \left\{ \mathbb{E}[V_i^\alpha] - y + \int_{\{v \leq y\}} F_i^\alpha(v) dv \right\} dH_i^\alpha(y) \\ &= \mathbb{E}[v_i^\alpha] - \mathbb{E}[Y_i^\alpha] + \mathbb{E}[\sigma_i^\alpha(Y_i^\alpha)], \end{aligned}$$

where the increasing and convex function σ_i^α is

$$\sigma_i^\alpha(y) \triangleq \int_{\{v \leq y\}} F_i^\alpha(v) dv.$$

By the dispersion theorem, $\alpha \succ_L \beta$ along with $\mathbb{E}[v_i^\alpha] = \mathbb{E}[v_i^\beta]$ implies $v_i^\alpha \geq_{\text{cx}} v_i^\beta$, which in turn implies $\sigma_i^\alpha(y) \geq \sigma_i^\beta(y)$ for every y in the support of H_i^α. Additionally, note that $Y_i^\alpha \geq_{\text{icx}} Y_i^\beta$ follows from Lemma 5.2 and that the function σ_i^β is increasing and convex. Putting them together, we have

$$\mathbb{E}[\sigma_i^\alpha(Y_i^\alpha)] \geq \mathbb{E}[\sigma_i^\beta(Y_i^\alpha)] \geq \mathbb{E}[\sigma_i^\beta(Y_i^\beta)], \text{ i.e., } \mathbb{E}[\sigma_i^\alpha(Y_i^\alpha)] \geq \mathbb{E}[\sigma_i^\beta(Y_i^\beta)].$$

The last inequality above can be interpreted as the bidder i's gain from more precise information; since more Lehmann-precise information causes his value estimate to be more dispersed,

[49] We should make a note that this example does not contradict with the result in Ganuza and Penalva (2010). Theorem 5 in their paper states that given two information structures, i.e., given $\epsilon > 0$, there exists $\overline{N}$ such that if the number of bidders exceed $\overline{N}$ the revenue is increasing.

[50] As we already have shown in Proposition 5.3, the revenue falls in case $N = 2$. As its immediate consequence, the information rent will increase in this case.

it is advantageous to the information rent. However, this also applies to his opponents. Hence $\mathbb{E}[Y_i^\alpha] \geq \mathbb{E}[Y_i^\beta]$, i.e., precise information would cause his opponents to bid more aggressively, which becomes disadvantageous to the bidder i's payoffs.

In addition to Lehmann precision, therefore, the next inequality is essential for the monotone information rents:

$$\mathbb{E}[\sigma_i^\alpha(Y_i^\alpha)] - \mathbb{E}[\sigma_i^\beta(Y_i^\beta)] \geq \mathbb{E}[Y_i^\alpha] - \mathbb{E}[Y_i^\beta]. \tag{10}$$

It reduces into a simple inequality in the symmetric second-price auction where $\pi_i = \pi$ and $G_i^\alpha = G^\alpha$ for all i, as the next proposition shows.

PROPOSITION 5.4 (Bidder's Information Rent). *In the symmetric second-price auction more Lehmann-precise information generates more each bidder's expected payoff, provided the following inequality holds:*

$$\int_{\{v \geq \alpha(p)\}} (1 - F^\alpha(v))dv \geq \int_{\{v \geq \beta(p)\}} (1 - F^\beta(v))dv, \quad \forall\, p \in [0,1], \tag{11}$$

where $\alpha(p)$ and $\beta(p)$ are the p-quantiles of the cumulative distribution functions F^α and of F^β, respectively. That is, $F^\alpha(\alpha(p)) = F^\beta(\beta(p)) = p$.

PROOF OF PROPOSITION 5.4 : Note that, in the symmetric auction, the difference between the benefits and costs to the information rents can be written by

$$\begin{aligned}
\mathbb{E}[\sigma_i^\alpha(Y_i^\alpha)] - \mathbb{E}[Y_i^\alpha] &= (N-1)\int \left\{ \int_{\{v \leq y\}} F^\alpha(v)dv - y \right\} [F^\alpha(y)]^{N-2} f^\alpha(y)dy \\
&= (N-1)\int_0^1 \left\{ \int_{\{v \leq \alpha(p)\}} F^\alpha(v)dv - \alpha(p) \right\} p^{N-2} dp,
\end{aligned}$$

where the bottom line is due to the change of variable $p = F^\alpha(y)$. Now we show that the given condition (11) is equivalent to

$$\int_{\{v \leq \alpha(p)\}} F^\alpha(v)dv - \alpha(p) \geq \int_{\{v \leq \beta(p)\}} F^\beta(v)dv - \beta(p) \quad \text{for each } p \in [0,1], \tag{12}$$

which is clearly sufficient for (10) to hold. To see this, we integrate the left-hand side of the inequality (12) by parts to obtain

$$\int_{\{v \leq \alpha(p)\}} F^\alpha(v)dv - \alpha(p) = \int_{\{v \geq \alpha(p)\}} (1 - F^\alpha(v))dv - \mathbb{E}[v_i^\alpha].$$

Since $\mathbb{E}[v_i^\alpha] = \mathbb{E}[v_i^\beta]$ by the iterated law of expectation, the desired result follows. □

It is worth comparing the condition (11) with Lehmann precision. To interpret the condition, note that a simple integration by parts gives

$$\int_{\{v \geq \alpha(p)\}} (1 - F^\alpha(v))dv = \mathbb{E}[(v_i^\alpha - \alpha(p)) \vee 0].$$

Hence the condition is equivalent to $\mathbb{E}[(v_i^\alpha - \alpha(p)) \vee 0] \geq \mathbb{E}[(v_i^\beta - \beta(p)) \vee 0]$ for every $p \in [0,1]$.[51] But it is in accord with high dispersion at least, since the inequality requires that the random variable v_i^α have more weight in the upper tail than v_i^β. The next result shows that the condition is connected with one variability order:

COROLLARY 5.5 (Ganuza and Penalva (2010)). *Suppose that the index set E is the dispersive-ordered, i.e., $\alpha > \beta$ implies $v_i^\alpha \geq_{disp} v_i^\beta$. Then the bidder's information rent is nondecreasing in α.*

PROOF OF COROLLARY 5.5 : Recall that

$$v_i^\alpha \geq_{\text{disp}} v_i^\beta \text{ provided } F^\alpha(q)^{-1} - F^\alpha(p)^{-1} \geq F^\beta(q)^{-1} - F^\beta(p)^{-1}$$

for all $p, q \in [0,1]$ and $p \leq q$.[52] Given p, let $\alpha(p) = F^\alpha(p)^{-1}$ and $\beta(p) = F^\beta(p)^{-1}$. Then we integrate the inequality above with respect to $q \geq p = F^\alpha(\alpha(p)) = F^\beta(\beta(p))$ to get

$$\int_{F^\alpha(\kappa_1)}^1 (F^\alpha(q)^{-1} - \kappa_1)dq \geq \int_{F^\beta(\kappa_2)}^1 (F^\beta(q)^{-1} - \kappa_2)dq$$

Changing the variable $F^e(q)^{-1} = x$ for each $e = \alpha, \beta$ and integrating by parts yields the condition (11). Since $v_i^\alpha \geq_{\text{disp}} v_i^\beta$ together with their same mean implies $v_i^\alpha \geq_{\text{icx}} v_i^\beta$, it follows from the dispersion theorem that α is more Lehmann-precise than β. Therefore, the monotone information rents follow from the preceding proposition. $\square$

It is a routine task to check that the condition (11) is not necessary for the dispersive order of the value estimates. Hence Proposition 5.4 is a generalization of Ganuza and Penalva (2010). We conclude this subsection with the following remarks:

REMARK 1 : When the value of estimates are arranged in the dispersive order, Ganuza and Penalva (2010) show that the auctioneer's revenue function $R(\alpha; N)$ obeys the supermodularity in $(\alpha; N)$, i.e., there is strategic complementarity between α and N from the auctioneer's point of view. Consequently, unless the cost $C(\alpha)$ the auctioneer has to incur for releasing information α relies on N, the auctioneer will provide more precise information as more bidders participate.

REMARK 2 : The second remark concerns the auctioneer's incentive to supply more information. If the condition (11) holds, the auctioneer would release less information compared to the efficient precision level since the condition guarantees that $\mathcal{E}(\alpha) - R(\alpha) = \sum_i U_i(\alpha)$ is nondecreasing in α; that is, the marginal revenue from more information would be smaller than

[51]Recall from the dispersion theorem that $\alpha \succ_L \beta$ implies $v_i^\alpha \geq_{\text{icx}} v_i^\beta$, but it is equivalent to $\mathbb{E}[(v_i^\alpha - c) \vee 0] \geq \mathbb{E}[(v_i^\beta - c) \vee 0]$ for every common constant c, while the given condition is based on every p-th quantile. Jewitt (1989) extended the mean-preserving spread of Rothschild and Stiglitz (1970) to present a location-independent variability order (also known as *dilation order* in statistics literature), but it is also based on the quantiles of each distribution.

[52]In Ganuza and Penalva (2010), the dispersive order is called "*supermodular precision*".

the corresponding marginal efficiency. Consequently, for any arbitrary cost function $C(\eta)$, the following monotone comparative static result holds.

$$\underset{\alpha\in E}{\operatorname{argmax}}\, \mathcal{E}(\alpha) - C(\alpha) \trianglerighteq \underset{\eta\in E}{\operatorname{argmax}}\, R(\alpha) - C(\alpha).$$

This result is readily extended to every symmetric standard auction due to the revenue equivalence theorem.

REMARK 3 : Bergemann and Välimäki (2002) showed that when bidders gather information covertly on their own before participating in an IPV auction, the second-price auction provides the (socially) efficient incentives for information acquisition with the bidders, provided the total cost incurred by them is additively separable, i.e., $C(\alpha_1, \cdots, \alpha_N) = \sum_i c_i(\alpha_i)$. One important implication from the preceding remark is, therefore, if a mechanism designer could decide on the information policy in the IPV auction, the decentralized covert information acquisition is more beneficial to the efficient information structure than the centralized overt information acquisition.

5.2. The Principal-Agent Problem

In this subsection we revisit the efficiency of information systems in a principal-agent model. The main inquiry is to rank two contractible variables for a principal who wants to control an agent's unobservable action, and it has been studied by Kim (1995) in the context of the first-order approach. The aim of this subsection is to examine the relationship between Lehmann-precision and the MPS criterion developed in his paper. It is shown that these two statistical notions are equivalent, in case the two contractible variables have a sufficiently smooth distribution function.

Kim (1995) considers a bilateral contract problem between a principal (she) and an agent (he). The scheme of the contract game is a typical moral hazard problem; the principal designs a contract to make him an offer, and the agent decides whether to accept, and if he accepts the contract he chooses a productive but hidden action; as a result, both the participants' payoffs are realized. Before she decides what contract to offer, however, she must ponder which variable the contract has to be contingent on.

To go into the details, suppose that there are two random variables X and Y, both of which are contractible and verifiable. Let $F(x|a)$ and $G(y|a)$ denote the X's and Y's distribution, respectively, given the agent's effort $a \in A$. In this subsection, we assume that the supports of X and Y—denoted $\mathcal{X}$ and $\mathcal{Y}$, respectively—are a unit interval $[0, 1]$. We assume that both distributions are absolutely continuous and thus they are endowed with the density functions $f(x|a)$ and $g(y|a)$ with respect to Lebesgue measure, respectively. In addition, we shall assume that the density functions obey the MLRP and they are differentiable with respect to a.

Let $u(w, a) = v(w) - c(a)$ denote the agent's payoff function, where the utility $v(w)$ from monetary transfer $w \in \Re_+$ is increasing and concave whereas the disutility $c(a)$ the agent incurs from exerting effort $a \in A$ is increasing and convex. For each distribution $F(x|a)$, in order to implement an action a at minimum cost, the principal must solve the following optimization problem:[53]

$$\min_{s \in \Psi} \rho^F(s,a) \triangleq \int_0^1 s(x) dF(x|a)$$

subject to

$$\text{(IR)} \quad \int_0^1 u(s(x)) dF(x|a) - c(a) \geq \overline{U}$$

$$\text{(IC)} \quad a \in \operatorname*{argmax}_{a' \in A} \int_0^1 u(s(x)) dF(x|a') - c(a').$$

Let $s_f^a \in \Psi$ denote the efficient incentive scheme implementing a under information system F, i.e., the solution to the optimization problem above. Similarly, let s_g^a be the solution under G. In this context, we say that the information system F is more efficient than G provided $\rho^F(s_f^a, a) \leq \rho^G(s_g^a, a)$ for all $a \in A$. That is, the more efficient information system alleviates the agent's incentive problem associated with his hidden effort.

In the differentiable environment where the first-order approach is valid, Holmström (1979) showed that the efficient incentive scheme s_f^a under F must satisfy the following first-order condition:[54]

$$\text{(13)} \quad \frac{1}{u'(s_f^a(x))} = \lambda_f + \mu_f \cdot l_f^a(x),$$

where λ_f and μ_f is the Lagrange multiplier for the (IR) and (IC) condition, respectively, and l_f^a is the likelihood ratio function:[55]

$$l_f^a(x) = \frac{\partial}{\partial a} \log f(x|a) = \frac{f_a}{f}(x|a).$$

In this circumstance, the MPS criterion developed by Kim (1995) states that if l_f^a is the mean-preserving spread of l_g^a for all $a \in A$, then F is more efficient than G.[56] In terms of a variability order, we can simply rewrite the criterion as $l_f^a(X) \geq_{\text{cx}} l_g^a(Y)$ for all a since those two functions have the identical mean 0; $\mathbb{E}[l_f^a(X)] = \mathbb{E}[l_g^a(Y)] = 0$. To understand the intuition behind this comparative static result, observe from (13) that the likelihood ratio function, rather than the density function itself, directly conveys information on the agent's action to the principal.[57] Hence, for F to be more efficient in a moral hazard problem, the corresponding contractible

[53] By Grossman and Hart (1983), the principal-agent problem with hidden action should be split into two parts: First, the principal computes the least expected cost necessary for the agent to take an action a^*, and then she considers which $a^* \in A$ should be implemented based on her payoff function.

[54] See Jewitt (1988) for the sufficient and necessary conditions on $F(x|a)$ under which the agent's expected payoff in (IC) is strictly concave in a', and thus the stationary point becomes the solution to (IC).

[55] When the density function $f(x|a)$ is differentiable with respect to a, Milgrom (1981) shows that $f(x|a)$ has the MLRP if and only if the likelihood ratio $l_f^a(x)$ is increasing in x $\forall$ a.

[56] Jewitt (2007) shows that the MPS criterion is also necessary for efficiency.

[57] Grossman and Hart (1983) showed that F is more efficient than G if F is Blackwell-sufficient for G. However, the sufficiency puts a condition on the density function.

variable X to F allows the principal to estimate the likelihood ratio l_f^a more accurately.

The following proposition shows that Lehmann-preciseness is equivalent to the MPS criterion. As a consequence, more Lehmann-precise contractible variable is sufficient and necessary for reducing the cost of the contract.

PROPOSITION 5.6. *F is more Lehmann-precise than G if and only if $l_f^a \geq_{cx} l_g^a$ for all $a \in A$.*

PROOF OF PROPOSITION 5.6 : (**Sufficiency**) Recall from Definition 2.6 that $F \succ_L G$ implies that for each $y \in [0,1]$ there exists an increasing function $T_y : A \to [0,1]$ for which $F(T_y(a)|a) = G(y|a)\ \forall\ a \in A$. Using Leibniz integral rule, we take the derivative of it with respect to a to obtain

$$\int_0^y g_a(t|a)dt = f(T_y(a)|a) \cdot \frac{\partial}{\partial a} T_y(a) + \int_0^{T_y(a)} f_a(t|a)dt.$$

Since $T_y(a)$ is increasing with a, the equation above leads us to:

$$\int_0^y l_g^a(t)g(t|a)dt \geq \int_0^{T_y(a)} l_f^a(t)f(t|a)dt.$$

Let $c_y = l_g^a(y)$. Adding $c_y[1 - G(y|a)] = c_y[1 - F(T_y(a)|a)]$ to both sides of the last inequality and integrating the left-hand side by parts yields

$$\int_0^1 \left(l_g^a(t) \wedge c_y\right) g(t|a)dt \geq c_y[1 - F(T_y(a)|a)] + \int_0^{T_y(a)} l_f^a(t)f(t|a)dt,$$

Since the likelihood ratio l_f^a is increasing, for all $\tau \in [0,1]$, we have[58]

$$c_y[1 - F(\tau|a)] + \int_0^\tau l_f^a(t)f(t|a)dt \geq \int_0^1 \left(l_f^a(t) \wedge c_y\right) f(t|a)dt.$$

Putting the two inequalities together, we obtain $\mathbb{E}\left[l_g^a(Y) \wedge c_y\right] \geq \mathbb{E}\left[l_f^a(X) \wedge c_y\right]$ for each $c_y \in \Re$. Since every concave function lies in the closed convex hull of the set $\{x \wedge c | c \in \Re\}$ up to constants, we have $l_g^a \geq_{\text{cv}} l_f^a$. Since both the likelihood ratio functions have the same mean, $l_f^a \geq_{\text{cx}} l_g^a$ follows.

(**Necessity**) To show the converse, suppose on the contrary that $l_f^a \leq_{\text{cv}} l_g^a$ but F is not more Lehmann-precise than G. Then it follows from Lemma 2.5 that for some $x, y \in [0,1]$, there exists an $\epsilon > 0$ such that for all $a^* \in (a, a+\epsilon)$, $G(y|a) = F(x|a)$ but $G(y|a^*) < F(x|a^*)$. Note

[58] To see this, define the function $\psi : [0,1] \to \Re$ as

$$\psi(\tau) = c_y\left[1 - F(\tau|a)\right] + \int_0^\tau l_f^a(t)f(t|a)dt.$$

Then its derivative $\psi'(\tau)$ is $f(\tau|a)(l_f^a(\tau) - c_y)$, and thus in case $l_f^a(\cdot)$ is increasing, it satisfies the single-crossing property in τ and crosses the horizontal axis when $l_f^a(\tau) = c_y$, where the function ψ assumes the minimum value, $\mathbb{E}[l_f^a(X) \wedge c_y]$.

that

$$G(y|a^*) - G(y|a) = \int_0^y \left(\frac{g(t|a^*)}{g(t|a)} - 1 \right) g(t|a)dt. \tag{14}$$

Let $c_x = f(x|a^*)/f(x|a) - 1$. By the assumptions we made above, we have

$$c_x\left[1 - G(y|a)\right] + G(y|a^*) - G(y|a) < c_x\left[1 - F(x|a)\right] + F(x|a^*) - F(x|a).$$

Using (14), we can rewrite the inequality as

$$\begin{aligned} &c_x\left[1 - G(y|a)\right] + \int_0^y \left(\frac{g(t|a^*)}{g(t|a)} - 1 \right) g(t|a)dt \\ &\qquad < c_x\left[1 - F(x|a)\right] + \int_0^x \left(\frac{f(t|a^*)}{f(t|a)} - 1 \right) f(t|a)dt. \end{aligned} \tag{15}$$

Note that, by the Lebesgue Dominated Convergence Theorem, the right-hand side of the inequality converges to $\mathbb{E}[l_f^a(X) \wedge c_x]$ as ϵ goes to 0. Moreover, the left-hand side is greater than $\mathbb{E}[l_g^a(Y) \wedge c_x]$ as we have pointed out in footnote 58. Hence, (15) gives $\mathbb{E}[l_g^a(Y) \wedge c_x] < \mathbb{E}[l_f^a(X) \wedge c_x]$, a contradiction. The proof is complete. □

We conclude the value of information in an agency model with the following remarks:

REMARK 1 : Proposition 5.6 generalizes the characterization result (Proposition 5) in Kim (1995), which demonstrates that Blackwell sufficiency implies the MPS criterion. Additionally, Kim showed with a counterexample that Blackwell sufficiency is not necessary for the MPS to hold. However, one can easily check that the counterexample satisfies Lehmann precision.

REMARK 2 : Kim (1995) illustrates via the counterexample that a Bayesian decision problem is generically different from an agency model in the sense that while the unknown state in a decision problem is to be *estimated* by the decision maker, the unobservable agent's action in the agency model is to be *controlled* by the principal. As Lehmann (1988) highlighted, however, Blackwell sufficiency is too restrictive to compare signals in many situations. Proposition 5.6 is, therefore, in substantial accord with the idea of Grossman and Hart (1983) that more precise information helps the principal to alleviate the incentive compatibility issue.[59]

REMARK 3 : When the first-order approach is not valid, Lehmann precision is not a suitable ordering in an agency model. To see this, recall from the example in Grossman and Hart (1983) that the MLRP *per se* does not guarantee monotonicity of the efficient incentive scheme for every $a \in A$. Intuitively, for a finite action space $A = \{a_1, a_2, a_3\}$ with $a_1 < a_2 < a_3$, if the principal wishes to implement an intermediate level of effort (a_2), but if the associated IC conditions with a_1 and a_3 are simultaneously binding, she is willing to pay less for higher

[59] Grossman and Hart (1983) show that the Blackwell sufficient signal guarantees a more efficient information system in a general environment where the first-order approach is not valid.

outcomes.[60] Consequently, $s_f^a(x)$ is not monotone in x, where Lehmann-precise information is not necessarily more valuable to the principal. Refer to Jewitt (2007) for a concrete example.

5.3. Delegation

The final application concerns the value of information in an organization where internal contingent transfer is infeasible.[61] Suppose that a principal and an agent have to make a collective decision under uncertainty. The principal has the *formal authority* (a legal right) to make the decision, but she lacks some of the relevant information required to make the decision. In this setting, she often delegates the decision process to the agent and allows him to take an action on her behalf, so that she can utilize the agent's information.[62] The problem is, however, that their preferences are not necessarily aligned. The delegation problem, initiated by Holmström (1977, 1984), deals with such a tradeoff between the agent's superior information and his biased preference.

To state the problem of our interest formally, consider a project selection problem. Let $A = \{a_1, \cdots, a_n\}$ denote the set of n possible projects and $\Theta = [\underline{\theta}, \overline{\theta}]$ represent the space of unknown states. The principal's and the agent's payoff functions are $v(a, \theta)$ and $u(a, \theta)$, respectively. For each selected project a_k, both $v(a_k, \theta)$ and $u(a_k, \theta)$ are assumed to be measurable mappings from Θ to $\Re$. Furthermore, we shall assume that $v(a, \theta)$ is supermodular in $(a; \theta)$ and $u(a, z)$ satisfies the single-crossing property in $(a; \theta)$.

Let $\pi \in P(\Theta)$ be the common prior on Θ. Without any loss of generality, we assume that for each $a_k \in A$,

$$\pi \left\{ \theta \in \Theta |\, a_k \in \underset{a \in A}{\operatorname{argmax}}\ u(a, \theta) \right\} > 0 \ \text{ and } \ \pi \left\{ \theta \in \Theta | a_k \in \underset{a \in A}{\operatorname{argmax}}\ v(a, \theta) \right\} > 0;$$

that is, every project is *rationalizable* for each party in a subset of Θ with a strictly positive measure.[63]

Suppose that the agent purchasing the signal α privately receives an outcome $x \in [0, 1]$ affiliated with θ, which is drawn from a distribution $G^\alpha(x|\theta)$, before selecting a project. Suppose, in addition, that the marginal distribution of x is uniform on the interval $[0, 1]$ for each of the two signals we compare.[64] Then the agent uses Bayes' rule to update his belief about

[60]When the first-order approach holds, however, only IC for the next lower action (a_1) is binding. Furthermore, the monotone property of the optimal scheme immediately follows from the first-order condition (13).

[61]It might be either because the two parties do not have any contractible variables or because such a transfer rule is illegal itself. For more specific examples, see Alonso and Matouschek (2008).

[62]In this case, the principal delegates the agent the *real authority* (effective control over decision) with commitment. Refer to Aghion and Tirole (1997) for detailed discussion on the formal and real authorities. If the principal does not have such a commitment power, the game belongs to a *cheap-talk game*, developed into a theory by Crawford and Sobel (1982).

[63]It is worth noting that this assumption does not cause any loss of generality: let $A_P \subset A$ denote the set of projects rationalizable for the principal, and $A_S \subset A$ for the agent. In a standard delegation problem, it is the principal who determines the set of possible projects but the agent with real authority will choose one in A_S only. Therefore, any supersets of $A_P \cap A_S$ will provide the same expected payoffs with the principal.

[64]The assumption that the marginal distribution is independent of the signals is crucial in this subsection. Under this assumption, we regard the cdf $M^\eta(x)$ as the outcome rather than x itself. If the density function of M^η is atomless, then

θ which is denoted $F^{\alpha}(\cdot|x)$, given the signal realization x. Then his optimal project-selection rule $\delta^{\alpha} : [0,1] \to A$ can be represented by a $n-1$-tuple vector $\{x_1, \cdots, x_{n-1}\}$ with $x_0 = 0$ and $x_N = 1$; for $x_{k-1} \le x \le x_k$, the project a_k maximizes the interim expected payoffs:

$$a_k \in \operatorname*{argmax}_{a \in A} \int_{\Theta} u(a,\theta) dF^{\alpha}(\theta|x).$$

In this context, we can write the principal's ex ante payoffs from the delegation as

$$R(\alpha) \triangleq \mathbb{E}_{\theta} \left[\int_{\mathcal{X}} v(\delta^{\alpha}(x), \theta) dG^{\alpha}(x|\theta) \right].$$

Similarly, the agent with the signal β privately observes an outcome $y \in [0,1]$ drawn from a distribution $G^{\beta}(y|\theta)$, and his optimal selection rule δ^{β} can be described by another $n-1$-tuple vector $\{y_1, \cdots, y_{n-1}\}$. Let $R(\beta)$ denote the principal's ex ante payoffs from delegating authority to the agent with β.

The central question in this subsection is when the principal benefits from a more informed agent in the delegation? In other words, we are interested in the condition for $R(\alpha) \ge R(\beta)$ where the signal α is more Lehmann-precise than β. To obtain this comparative static result, it seems natural that both the parties' payoff functions are to be sufficiently aligned.

For simple exposition, we define, for each $k = 1, \cdots, N-1$, the expected incremental returns from selecting the next project (a_{k+1}) conditional on x as

$$^{u}\Delta_k^{\alpha}(x) = \int_{\Theta} \left[u(a_{k+1},\theta) - u(a_k,\theta) \right] dF^{\alpha}(\theta|x) \;\text{ and }\; ^{v}\Delta_k^{\alpha}(x) = \int_{\Theta} \left[v(a_{k+1},\theta) - v(a_k,\theta) \right] dF^{\alpha}(\theta|x)$$

to the agent and to the principal, respectively.

DEFINITION 5.1 (Aligned Preferences). *Given the two signals α, β and the agent's projection selection rules $\delta^{\alpha} = \{x_1, \cdots, x_{n-1}\}$ and $\delta^{\beta} = \{y_1, \cdots, y_{n-1}\}$, we say that the principal's preference is aligned with the agent's preference provided for each $k = 1, \cdots, n-1$,*

$$\textit{for } x_k < y_k,\;\; ^{u}\Delta_k^{\alpha}(x) \ge 0 \;\Rightarrow {}^{v}\Delta_k^{\alpha}(x) \ge 0 \;\; \textit{but for } x_k > y_k,\;\; ^{v}\Delta_k^{\alpha}(x) \ge 0 \;\Rightarrow {}^{u}\Delta_k^{\alpha}(x) \ge 0.$$

Recall that the function $^{u}\Delta_k^{\alpha} : [0,1] \to \Re$ changes its sign from negative to positive at the cutoff point x_k, and that both $^{u}\Delta_k^{\alpha}$ and $^{v}\Delta_k^{\alpha}$ satisfy the SCP in x. For $x_k < y_k$, the aligned preferences, therefore, require that the function $^{v}\Delta_k^{\alpha}$ crosses the horizontal axis before $^{u}\Delta_k^{\alpha}$ does as is illustrated in Figure 5. In contrast, for $x_k > y_k$, the function $^{u}\Delta_k^{\alpha}$ has to cross the axis first for the aligned preferences.

The next proposition shows that the condition is sufficient for the principal to gain from the better informed agent in the delegation.

PROPOSITION 5.7 (Value of Information in Delegation). *Suppose that the two parties' payoff functions are aligned. Then the principal benefits from the agent with more Lehmann-precise information.*

both $M^{\eta}(x)$ and x contain the same information. For this reason, many literature (for instance, Szalay (2009)) works with $M^{\eta}(x)$ instead of x itself. Nevertheless, the assumption restricts either the set of possible experiments or the set of priors.

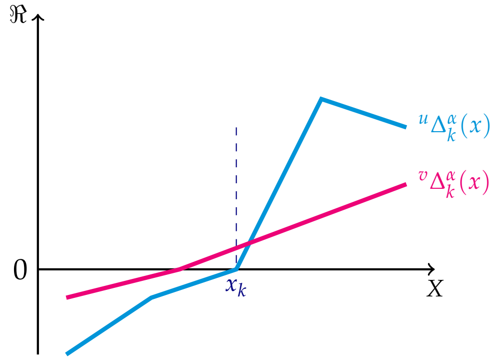


Figure 5: Aligned Preferences for $x_k < y_k$

Proof of Proposition 5.7 : See Appendix E.

Proposition 5.7 grasps the basic intuition that when the principal's and the agent's interest are sufficiently congruent, the principal can gain from the better informed agent. The proof exploits an immediate result of the dispersion theorem that if the principal faces a more informed agent, more dispersed is the expected incremental returns from taking the next higher project; ${}^v\Delta_k^\alpha(X) \geq_{\text{icx}} {}^v\Delta_k^\beta(Y)$. In other words, the choice between the two adjacent projects becomes riskier. Since a more informed agent will put more weights on the signal realization, his optimal selection rule becomes more sensitive to change in x. Consequently, for such a strategy to be more beneficial to the principal, their preferences must be aligned.

However, the result is deficient to the extent that it requires a substantial amount of structure on the primitive payoff functions, since the condition is based on the change in the agent's optimal rule responding to change in the information structure. In fact, unless the preferences are perfectly aligned, a general discussion of the value of information is very problematic without formulating how the agent's optimal rule changes. The example below illustrates this point:

Example 5.2. *Let θ be uniformly distributed on the unit interval $(0,1)$. From the signal β, the agent has a partition on Θ as follows:*

$$\left\{\left(0,\frac{1}{2}\right),\left(\frac{1}{2},\frac{3}{4}\right),\left(\frac{3}{4},1\right)\right\}$$

The information partition corresponding to the signal α is

$$\left\{\left(0,\frac{1}{4}\right),\left(\frac{1}{4},\frac{1}{2}\right),\left(\frac{1}{2},\frac{3}{4}\right),\left(\frac{3}{4},1\right)\right\}$$

On the other hand, the principal's information partition is $\{(0,1)\}$. Note that $\alpha \succ_L \beta$.[65]

[65] For $\theta_1 \in (0,1/4)$ and $\widehat{y} \in (0,1/4)$, the T-transformation is$T_{\widehat{y}}(\theta_1) = \frac{1}{2}\widehat{y}$. For $\theta_2 \in (1/4,1/2)$, however, $T_{\widehat{y}}(\theta_2) = \frac{2\widehat{y}+1}{4}$. Accordingly, $T_{\widehat{y}}$ is increasing in θ and thus α is more precise in Lehmann's notion. In fact, α is Blackwell sufficient

Let $A = \{a_1, a_2, a_3\}$ with $a_1 < a_2 < a_3$. The payoffs are given in the table below, where $\theta_i \in \left(\frac{i-1}{4}, \frac{i}{4}\right)$ $i \in \{1,2,3,4\}$. In each entry, the first number indicates the principal's payoffs and the second number indicates the agent's payoffs.

	States			
Projects	θ_1	θ_2	θ_3	θ_4
a_1	(10,10)	(10,4)	(0,2)	(0,0)
a_2	(5,3)	(5,7)	(5,5)	(7,7)
a_3	(0,0)	(0,5)	(1,4)	(10,10)

Table 1: A better information may hurt the principal.

It is easy to see that each payoff function belongs to $\mathcal{U}^{spm}$, and that the agent's optimal project selection rule is $\delta^\beta = \{1/2, 3/4, 1\}$ under β but $\delta^\alpha = \{1/4, 3/4, 1\}$ under α. Note that $1/4 = x_1 < y_1 = 1/2$. The aligned preference requires, in this circumstance, that ${}^u\Delta_1^\alpha(x) \geq 0$ imply ${}^v\Delta_1^\alpha(x) \geq 0$. However, for $x_2 \in (1/4, 1/2)$,

$$ {}^u\Delta_1^\alpha(x_2) = u(a_2, \theta_2) - u(a_1, \theta_2) \geq 0 \;\; \textit{but} \;\; {}^v\Delta_1^\alpha(x_2) = v(a_2, \theta_2) - v(a_1, \theta_2) < 0. $$

Now we show that the principal gets worse off as the agent gets more informed, even though she is able to optimally choose a delegation set D^*, an element of the power set of A.

Under the signal β, the agent's optimal selection rule is perfectly aligned with the principal's. Thus, full delegation $D^* = A$ is optimal. The expected payoff from the full delegation is $R(\beta) = 8.75$. Under the more precise signal α, it can be shown that full delegation is still optimal, although the agent's optimal project is a_2 for $x \in (1/4, 1/2)$ which is misaligned with the principal's. The signal α, however, yields $R(\alpha) = 7.5$ to the principal. Therefore, the marginal value of precise information for the principal $R(\alpha) - R(\beta)$ is negative. $\square$

6. Conclusion

The main purpose of this paper has been to develop a new approach to ordering signals in a rich class of economic decision problems and integrate it with the three well-known signal orderings. As a result, I established the equivalence of the four different ways to rank signals between Lehmann precision, the effectiveness, the informativeness, and the dispersion. I used this equivalence theorem to find the sufficient and necessary condition for one signal to be more valuable than another in the two branches of decision theory within several classes of preferences, frequently used in nearly every subfield of economics and statistics.

Also, I derived from the theorem the exact characterization of precise signals in terms of dispersion, associating the precision of signals with high dispersion of predictions about uncertainty. I showed that this characterization result justifies, in a unified way, other signal orderings developed by previous literature. Furthermore, I showed how the result can be ap-

for β.

plied to several strategic settings, in each of which I ascertained the impact of precise signals on players' payoffs.

This paper provides a natural guide to future research. Built upon the binary preference relation over the signal space, one natural question ripe for study is the demand for information. To this purpose, an essential prerequisite is the systematic analysis of the *marginal value* of information. Initial efforts have been done by Persico (2000), but practical applications of his result are rare at best since it calls for a great deal of structures on payoff functions and even the decision maker's optimal decision rules.

While I talked about the value of information from Bayesian perspectives, I in fact adopted the strong criterion such that for one signal to be more valuable than another, the signal should yield more ex ante values to every decision maker irrespective of her prior information. However, in many economic applications, decision makers possess prior beliefs with a common property such as unimodal densities, and thus it is reasonable to utilize the additional structure on the set of priors. It will be interesting to investigate how such a structure of prior beliefs affects the decision maker's preference over the signal.

Another subject is the value of information in the presence of strategic interaction between decision makers. Although the dispersion theorem provides a method for analysis in some environments, a new approach for ordering signals should be developed to accommodate the strategic interaction. A recent study by Bergemann and Morris (2013) addresses this issue in a general game environment, in line with Blackwell (1951, 1953), by introducing a new signal ordering (named the *individual sufficiency*). A natural question that will arise in mind is what if we restrict our attention to a subclass of games in the same spirit with Lehmann (1988). The most interesting class would be the games with strategic complementarities (the supermodular games) developed by Milgrom and Roberts (1990).

Appendices

A. The Karlin-Rubin Monotone Property

In this appendix, we introduce another class of payoff functions, to which most of primitive functions in statistics—such as hypothesis testing and estimation problems—belong. We also revisit Lehmann's Theorem to compare his improvement principle with the principle developed in Lemma 3.3. Also, we establish the equivalence of the four signal orderings within this class.

We begin with a formal definition of the class of payoff functions to be studied.

DEFINITION A.1 (KRM Family, Karlin and Rubin (1956)). *A family $\{u(\cdot,\theta)\}_{\theta\in\Theta}$ of payoff functions obeys the Karlin-Rubin Monotone property (*KRM*, and we write $u \in \mathcal{U}^{KRm}$) provided (*MCS*) the set of maximizers $A^*(\theta) = \text{argmax}_{a\in A}\, u(a,\theta)$ is nondecreasing with θ in the strong set order and (*SP*) $u(a,\theta)$ is nonincreasing with a away from the $A^*(\theta)$ for each θ.*

The KRM family stems from the two properties: (MCS) for monotone comparative statics and (SP) for the single peak. Here, the condition (SP) reduces to a quasiconcave payoff function in a for each θ when the action space is given by a convex set.[66]

Before proceeding, it is worth noting that the KRM family is distinct with the SCP family as is pointed by Jewitt (2007). To see that the KRM does not necessarily satisfy the SCP, refer to the subfigure (a) in Figure 6, where $u(a,\theta)$ is described as a function of a for each state $\theta \in \{\theta_L, \theta_H\}$ with $\theta_L < \theta_H$. For each state θ, $u(a,\theta)$ has a single peak and the peak at θ_H—labeled by $a^*(\theta_H)$—is larger than $a^*(\theta_L)$. Hence it satisfies the two given conditions so $u \in \mathcal{U}^{KRm}$. However, note that the incremental returns from taking a'' instead a' is nonnegative in θ_L but is negative in θ_H, so the payoff function violates the SCP.

Likewise, the SCP does not imply the KRM. The subfigure (b) displays an example, where for each of three possible actions, the payoff function is sketched as a function of θ.[67] The example provides two possibilities depending on whether the payoff when a_2 is taken is either $u(a_2,\theta)$ or $\widehat{u}(a_2,\theta)$, holding $u(a_1,\theta)$ and $u(a_3,\theta)$ fixed. In both cases, one can easily verify that the SCP is satisfied.[68] However, in case $u(a_2,\theta)$ is below the intersection between $u(a_1,\theta)$ and $u(a_3,\theta)$, the payoff function is not a KRM family since it violates the condition (SP) for the state $\theta \in (\theta_2, \theta_1)$.

We first establish the next two features of the KRM property in a finite action space.

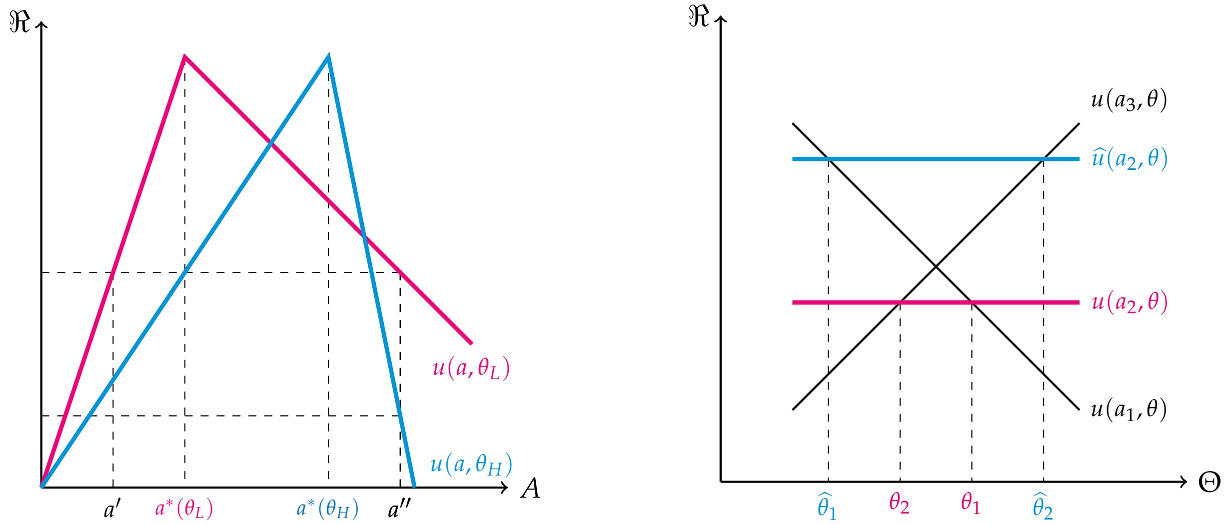
LEMMA A.1. *Suppose that $A = \{a_1, \cdots, a_n\}$ is the set of n possible actions. For every $u \in \mathcal{U}^{KRm}$, the incremental returns from the next higher action rather than a_k, denoted ${}^u\Delta_k(\theta) \triangleq u(a_{k+1},\theta) - u(a_k,\theta)$, satisfy the single crossing property in θ.*[69]

[66]In light of this fact together with (MCS), this family is dubbed "*quasiconcave payoff functions with increasing peaks*" in Quah and Strulovici (2009). The term "Karlin-Rubin Monotone" is taken from Jewitt (2007).

[67]Quah and Strulovici (2009) provides another interesting example of production planning with a state-dependent cost where the producer's profit function satisfies the KRM but violates the SCP.

[68]Note that both of them are supermodular in $(a;\theta)$. Therefore, the examples in Figure 6 also tell us that there is no relationship between $\mathcal{U}^{spm}$ and $\mathcal{U}^{KRm}$, either.

[69]It should be clear that this property does not imply $\mathcal{U}^{KRm} \subset \mathcal{U}^{sc}$, since the single-crossing property does not



(a) A convex action space

(b) A finite action space $A = \{a_1, a_2, a_3\}$

Figure 6: The KRM and SCP Family - There is no relationship between the two classes. The both classes are included in the IDO class, however.

PROOF OF LEMMA A.1 : Suppose that ${}^u\Delta_k(\theta_k) \geq 0$ for some θ_k. By (SP) in Definition A.1, it must be the case that $\min A^*(\theta_k) \geq a_{k+1} > a_k$, where $\min A^*(\theta_k)$ indicates the smallest element in $A^*(\theta_k)$. For every $\theta \geq \theta_k$, (MCS) implies $\min A^*(\theta) \geq \min A^*(\theta_k) \geq a_{k+1}$. Hence ${}^u\Delta_k(\theta) \geq 0$ is immediate from (SP). $\square$

An intrinsic feature of $\mathcal{U}^{KRm}$ that $\mathcal{U}^{sc}$ does not possess is the following:

LEMMA A.2. *Let $u \in \mathcal{U}^{KRm}$ and $A = \{a_1, \cdots, a_n\}$. For the three adjacent actions $a_{k+2} > a_{k+1} > a_k$ in A, suppose that ${}^u\Delta_k$ and ${}^u\Delta_{k+1}$ cross the horizontal axis at θ_k and θ_{k+1}, respectively. Then $\theta_{k+1} \geq \theta_k$.*

PROOF OF LEMMA A.2 : We argue by contradiction. Suppose to the contrary that $\theta_{k+1} < \theta_k$. Then there exists a state $\theta^* \in (\theta_{k+1}, \theta_k)$ at which ${}^u\Delta_{k+1}(\theta^*) \geq 0$ but ${}^u\Delta_k(\theta^*) < 0$. ${}^u\Delta_{k+1}(\theta^*) \geq 0$ implies $a_k < a_{k+1} < a_{k+2} \leq \min A^*(\theta^*)$, which in turn implies ${}^u\Delta_k(\theta^*) \geq 0$, a contradiction. $\square$

The preceding lemma clarifies why the payoff function depicted in Figure 6-(b) is not an element of $\mathcal{U}^{KRm}$; when $u(a_2,\theta)$ is below the intersection of $u(a_1,\theta)$ and $u(a_3,\theta)$, we have $\theta_2 < \theta_1$ where θ_1 and θ_2 denote the points at which ${}^u\Delta_1$ and ${}^u\Delta_2$ change the sign from negative to positive, respectively.[70]

Focusing on this class of payoff functions, Lehmann (1988) proves that the signal α is more effective than another signal β with respect to the class $\mathcal{U}^{KRm}$ if and only if $\alpha \succ_L \beta$. Put differ-

necessarily hold for any other higher actions than a_{k+1}. In a binary action space, however, the inclusion $\mathcal{U}^{KRm} \subset \mathcal{U}^{sc}$ holds. Furthermore, since (MCS) is implied by the SCP and (SP) is always true in a binary action space, the two classes are equivalent. Therefore, this result provides an alternative proof of necessity of the SCP, in a binary action space, for the monotone comparative statics in Milgrom and Shannon (1994).

[70]However, by reassigning higher values than the intersection to $u(a_2,\theta)$ or by simply ruling it out, the KRM property can be obtained.

ently, he develops a statistical signal ordering which is equivalent to the signal ordering based on the statistical values within $\mathcal{U}^{KRm}$.[71] Like the proofs of Theorem 3.1 for $\mathcal{U}^{sc}$, he suggested one way to find a monotone decision rule d^{α} that makes every statistician within the class better off for every monotone decision rule $d \in \mathcal{D}^{\beta,M}$, and then exploited the essentially complete class theorem to complete the proof.

Before turning to the next result, it is instructive to compare his improvement principle with the principle developed in Lemma 3.3. To go into details, we work with a finite action space $A = \{a_1, \cdots, a_n\}$. Given a payoff function $u \in \mathcal{U}^{KRm}$, let $\widehat{\theta}_k \in \Theta$ denote the state at which the function ${}^{u}\Delta_k(\theta)$ changes its sign for each $k = 1, \cdots, n-1$. By Lemma A.2, it follows $\widehat{\theta}_1 \leq \cdots \leq \widehat{\theta}_{n-1}$.

Now consider a monotone decision rule $d^{\beta} : \mathcal{Y} \to A$ based on β. In every finite action space, such a monotone decision rule can be represented by a nondecreasing sequence of $n-1$ cutoff points in $\mathcal{Y}$, $\{y_1, \cdots, y_{n-1}\}$, where d^{β} assigns the action a_k to $y \in (y_{k-1}, y_k)$ for each i. The improvement principle of Lehmann (1988) argues that the monotone decision rule $d^{\alpha} = \{x_1, \cdots, x_{n-1}\}$ based on α, dominant over d^{β}, can be found by the T-transformation in Definition 2.6 and the set of the change points $\{\widehat{\theta}_k\}_{k=1}^{n-1}$ as follows:

$$x_k = T_{y_k}(\widehat{\theta}_k) \quad \text{for each } k = 1, 2, \cdots, n-1.$$

The constructed decision rule is indeed monotone, since

$$x_k = T_{y_k}(\widehat{\theta}_k) \leq T_{y_k}(\widehat{\theta}_{k+1}) \leq T_{y_{k+1}}(\widehat{\theta}_{k+1}) = x_{k+1},$$

where the inequalities are due to the monotone property of the T-transformation.

The above principle, however, does not apply to $u \in \mathcal{U}^{sc}$ since the SCP imposes no ordering structures on the change points $\{\widehat{\theta}_k\}$.[72] Instead, we can take advantage of the property that there exists an increasing best response to the primitive payoff function, $d^{\star}(\theta) \in \operatorname{argmax}_{a \in A} u(a, \theta)$, which comes up with a partition of $\Theta = [\underline{\theta}, \overline{\theta}]$—$\underline{\theta} = \theta_0^* \leq \theta_1^* \leq \cdots \leq \theta_{n-1}^* \leq \theta_n^* = \overline{\theta}$—such that $d^{\star}(\theta) = a_k$ for $\theta \in (\theta_{k-1}^*, \theta_k^*)$. Furthermore, it should be clear from the proof of Lemma 3.3 that $d^{\alpha}(x) = a_k$ if and only if $d^{\star}(\theta) = d^{\beta} \circ \tau_x(\theta) = a_k$. From $d^{\star}(\theta) = a_k$ and $d^{\beta} \circ \tau_x(\theta) = a_k$, we infer that θ belongs to $(\theta_{k-1}^*, \theta_k^*)$ and $\tau_x(\theta)$ to (y_{k-1}, y_k), equivalently, $x \in (T_{y_{k-1}}(\theta), T_{y_k}(\theta))$. Putting them together, we have

$$d^{\alpha}(x) = a_k \ \text{ if and only if } \ x \in (T_{y_{k-1}}(\theta_{k-1}^*), T_{y_k}(\theta_k^*)).$$

Therefore, d^{α} is monotone and, in addition, the way to find a such strategy for both $\mathcal{U}^{sc}$ and $\mathcal{U}^{KRm}$ is virtually identical, except one difference that the partition of Θ is determined by the inherent nature of the primitive function for $\mathcal{U}^{KRm}$ on the basis of Lemma A.2, while it is determined by the monotone property of $d^{\star}$ for $\mathcal{U}^{sc}$.

[71] Like the three classes of payoff functions, even $\mathcal{U}^{KRm}$ is invariant to the addition of $h \in \mathcal{H}$ in Lemma 2.1. Therefore, the criterion of effectiveness validates the signal ordering based on the statistical values within $\mathcal{U}^{KRm}$ as well.

[72] Recall the example in the subfigure (b) above. Despite $u(a_2, \theta)$ being lower than the intersection, the SCP is preserved.

We conclude our discussion on the class $\mathcal{U}^{KRm}$ with the next result:

COROLLARY A.3. *Within the class of payoff functions $\mathcal{U}^{KRm}$, the four signal orderings stated in the main theorem are equivalent.*

PROOF OF COROLLARY A.3 : Since (i) Precision $\Rightarrow$ (ii) Effectiveness is done by Lehmann (1988) and both implications—(ii) $\Rightarrow$ (iii) Informativeness and (iv) Dispersion $\Rightarrow$ (i)—hold universally, it suffices to establish (iii) $\Rightarrow$ (iv). To this end, we need show that the primitive $u_{\psi,\kappa} \in \mathcal{U}^{KRm}$ defined in the proof of Theorem 3.5 for every $\kappa \in \Re$ and $\psi \in \Theta^*$. But it is straightforward from Lemma A.1. $\square$

Corollary A.3 is a generalization of the Lehmann's Theorem with another two equivalent signal orderings, one based on the Bayesian values and another based on the dispersion.

B. The Interval Dominance Order Property

In this section, we extend the equivalence theorem to the interval dominance order (IDO) family, the largest of all classes considered in this paper.

To illustrate the IDO family in an arbitrary compact action space, we must introduce the concept of an *interval*. To state formally, consider a compact action space A and let $a', a'' \in A$ with $a' < a''$. The interval $[a', a'']$ between the two actions indicates $[a', a''] = \{a \in A |\ a' \leq a \leq a''\}$. Definition 2.3 requires, therefore, that if a'' is dominant over every other action within the interval $[a', a'']$ in state θ, a'' remains dominant over a' in every $\theta' > \theta$.

By comparing their definitions, $\mathcal{U}^{sc} \subset \mathcal{U}^{ido}$ is straightforward, since the two classes have the same consequent but the latter places stronger restrictions on the antecedent. Note that $\mathcal{U}^{sc}$ is a proper subset of $\mathcal{U}^{ido}$. The example displayed in the Figure 6-(a) clarifies this fact.

The IDO class also contains the KRM class as well. To see $\mathcal{U}^{KRm} \subset \mathcal{U}^{ido}$ formally, consider a payoff function $u \in \mathcal{U}^{KRm}$. When there is no action within the interval $(a', a'') = \{a \in A | a' < a < a''\}$, the IDO property is immediate from Lemma A.1. Thus we suppose that (a', a'') is nonempty and for some θ

$$(\bigstar) \quad u(a'', \theta) \geq u(a, \theta) \quad \forall\, a \in [a', a''].$$

Let $a^*(\theta) = \min A^*(\theta)$ at each state θ. When $a'' \notin A^*(\theta)$, the inequality ($\bigstar$) leads us to $a' < a'' < a^*(\theta) \leq a^*(\theta')\ \forall\ \theta' > \theta$ due to the (MCS) and (SP) conditions of $\mathcal{U}^{KRm}$. Hence $u(a'', \theta') \geq u(a', \theta')$ follows by the condition (SP). When $a'' \in A^*(\theta)$, suppose to the contrary that $u(a'', \theta') < u(a', \theta')$ for some $\theta' > \theta$. Then $a'' \notin A^*(\theta')$. Since $A^*(\theta') \trianglerighteq A^*(\theta)$ and $a'' \in A^*(\theta)$, however, it must be the case that $a'' < a^*(\theta')$. Thus the inequality $u(a'', \theta') < u(a', \theta')$ contradicts with $u \in \mathcal{U}^{KRm}$ so we prove that $\mathcal{U}^{ido}$ is the largest class among the ones taken into account.

The first result of this section shows that $\mathcal{U}^{ido}$ is equivalent to $\mathcal{U}^{KRm}$ under the additional assumption of the condition (SP). For this purpose only, we denote by $\mathcal{U}^{SP}$ the set of payoff functions satisfying (SP).

Action Space	Payoff Function Space	Inclusion
$A = \{a_1, a_2\}$	$\mathcal{U}$	$\mathcal{U}^{spm} \subset \mathcal{U}^{sc} = \mathcal{U}^{KRm} = \mathcal{U}^{ido} = \mathcal{U}^{MCS}$
Compact A	$\mathcal{U}^{SP}$	$\mathcal{U}^{spm} \subset \mathcal{U}^{sc} \subset \mathcal{U}^{KRm} = \mathcal{U}^{ido} = \mathcal{U}^{MCS}$
	$\mathcal{U}$	$\mathcal{U}^{spm} \subset \mathcal{U}^{sc} \subset \mathcal{U}^{ido}$ and $\mathcal{U}^{KRm} \subset \mathcal{U}^{ido}$

Table 2: Payoff Functions Relationship between SPM, SCP, KRM, and IDO

PROPOSITION B.1. *Within the class of payoff functions $\mathcal{U}^{SP}$, the condition* (MCS) *holds if and only if u obeys the IDO property.*

PROOF OF PROPOSITION B.1 : We provide the proof under the additional assumption that A is a convex set. In case A is a finite action space, the result can be shown in an analogous way. Recall, in this circumstance, that $u \in \mathcal{U}^{SP}$ is equivalent to saying that u is quasiconcave in a for each θ.

By Theorem 1 in Quah and Strulovici (2009), the IDO property is sufficient for (MCS) without the aid of the quasiconcavity. To prove the converse, given a state θ, we suppose that $u(a',\theta) \geq u(\hat{a},\theta)$ for all $\hat{a} \in [a,a']$. Then it follows from the quasiconcavity that $u(\lambda a + (1-\lambda)a',\theta) \geq u(a,\theta) \wedge u(a',\theta) = u(a,\theta)$ for all $\lambda \in [0,1]$. Thus, $u(a,\theta)$ is nondecreasing with a on the interval $[a,a']$.

Now we claim that $u(a,\theta')$ is also nondecreasing on the same interval $[a_1,a_2]$ for every $\theta' > \theta$, which is a sufficient condition for $u(a,\theta)$ to satisfy the IDO property. We prove this claim by contradiction. Suppose to the contrary that there exist a pair of actions $a_1 < a_2$ in $[a,a']$ for which

$$u(a_2,\theta') < u(a_1,\theta') \leq u(a^*,\theta'), \quad \text{where } a^* \in A^*(\theta') \text{ and } a^* \geq a'.$$

Note that $A^*(\theta') \trianglerighteq A^*(\theta)$ guarantees the existence of a^*. Hence there exists some $\lambda^* \in [0,1]$ for which $a_2 = \lambda^* a^* + (1-\lambda^*)a_1$ since we assume that the set A is convex. Then, it follows from the quasiconcavity of u that $u(a_2,\theta') \geq u(a^*,\theta') \wedge u(a_1,\theta') = u(a_1,\theta')$, which is a contradiction. Accordingly, $u(\cdot,\theta)$ is nondecreasing over the interval $[a,a']$ and therefore $u(a',\theta') \geq u(a,\theta')$ for all $\theta' > \theta$. □

Table 2 summarizes the discussion about the classes of payoff functions concerning θ. In a binary action space, Lemma A.1 tells us that the three classes—SCP, KRM, and IDO—are mutually equivalent to the class of payoff functions—labeled by $\mathcal{U}^{MCS}$—satisfying the condition (MCS). When we focus on the class of payoff functions with (SP) defined on a compact action space, the preceding proposition demonstrates that the three classes with KRM, IDO, and MCS are mutually equivalent. The example in Figure 6-(a) also shows that the class of SCP is a proper subset of them. Without any structures on A and $\mathcal{U}$, however, there is no relationship between the two classes KRM and SCP as discussed above.

Now we turn to the equivalence of the signal orderings within $\mathcal{U}^{ido}$. The next result is an extension of Lemma 3.2 to $\mathcal{U}^{ido}$.

LEMMA B.2. *Let $u \in \mathcal{U}^{ido}$. Then for every decreasing function $d : \Theta \to A$, there is an action $a^* \in A$ for which $u(a^*, \theta) \geq u(d(\theta), \theta)$ for every state $\theta \in \Theta$.*

PROOF OF LEMMA B.2 We infer from Theorem 1 in Quah and Strulovici (2009) that for every $u \in \mathcal{U}^{ido}$, there is a nondecreasing strategy $d^\star(\theta) \in A^*(\theta) \equiv \text{argmax}_{a \in A} u(a, \theta)$. Suppose that d and $d^\star$ intersect at a point $(\theta^\star, a^\star)$. Then $d^\star(\theta^\star) = a^\star$ implies

$$u(a^\star, \theta^\star) \geq u(a, \theta^\star) \quad \forall\, a \in A.$$

In particular, we have $u(a^\star, \theta^\star) \geq u(a, \theta^\star)$ on each interval $[a', a^\star]$ for every $a' \leq a^\star$. Hence, it follows from the IDO property that $u(a^\star, \theta) \geq u(a, \theta)$ for all $a \in [\underline{a}, a^*]$ and $\theta > \theta^\star$, equivalently, $u(a^\star, \theta) \geq u(d(\theta), \theta)$ for all $\theta > \theta^\star$ since d is decreasing. Similarly, since $u(a^\star, \theta^\star) \geq u(a, \theta^\star)$ for all $a \in [a^\star, \overline{a}]$, it follows from the IDO property that $u(a^\star, \theta) \geq u(a, \theta)$ for all $a \in [a^\star, \overline{a}]$ and $\theta < \theta^\star$, namely, $u(a^\star, \theta) \geq u(d(\theta), \theta)$ for all $\theta < \theta^\star$. □

Therefore, the improvement principle (Lemma 3.3) holds even for $\mathcal{U}^{ido}$, and therefore, we extend the equivalence theorem to the IDO class. This result completes the proof of the main theorem in Section 3.

COROLLARY B.3. *Within the class of payoff functions $\mathcal{U}^{ido}$, the four signal orderings stated in the main theorem are equivalent.*

PROOF OF COROLLARY B.3 : It suffices to show that (i) Precision ⇒ (ii) Effectiveness and (iii) Informativeness ⇒ (iv) Dispersion. (i) ⇒ (ii) follows from Theorem 2.2 and the preceding lemma. (iii) ⇒ (iv) also holds since the set of payoff functions $\{u_{\psi,\kappa} | \psi \in \Theta^*,\ \kappa \in \Re\}$ is a subset of $\mathcal{U}^{ido}$. □

C. Informativeness for Supermodular Payoff Functions

In this section, we use the dispersion theorem (Corollary 4.5) to provide a more concise proof of the informativeness for a decision maker with supermodular payoff functions. To this end, we shall work with a finite action space, $A = \{a_1, a_2, \cdots, a_n\}$. The extension to a general infinite action space follows via a limiting argument.

Suppose that a decision maker has a payoff function $u(a, \theta)$ supermodular in $(a; \theta)$. Recall that supermodularity is preserved under stochastic integrations with the MLRP. Hence, given a signal α, there is an optimal Bayesian decision rule $\delta^\alpha : \mathcal{X} \to A$ nondecreasing which maximizes the expected payoffs $\mathbb{E}_\theta [u(a, \theta) | X = x]$ for each $x \in \mathcal{X}$. In a finite action space, therefore, δ^α can be represented by a $(n-1)$-tuple vector $\{x_1, \cdots, x_{n-1}\}$; to each outcome $x \in (x_{k-1}, x_k)$ δ^α assigns a_k. It enables us to write the Bayesian value of α as

$$\mathcal{V}^\pi(\alpha, u) = \mathbb{E}_\theta \left[\int_{\mathcal{X}} u(\delta^\alpha(x), \theta) dG^\alpha(x|\theta) \right]$$

$$= \mathbb{E}_\theta \left[\sum_{k=1}^{N} \int_{x_{k-1}}^{x_k} u(a_k, \theta) dG^\alpha(x|\theta) \right]$$

$$= \mathbb{E}_\theta \left[u(a_1, \theta) + \sum_{k=1}^{N-1} \int_{\mathcal{X}} \mathbb{1}_{\{x \geq x_k\}}(x) \left[u(a_{k+1}, \theta) - u(a_k, \theta) \right] dG^\alpha(x|\theta) \right]$$

$$= E_\theta \left[u(a_1, \theta) \right] + \sum_{k=1}^{N-1} \int_{\mathcal{X}} \mathbb{1}_{\{x \geq x_k\}}(x) \left\{ \int_\Theta \left[u(a_{k+1}, \theta) - u(a_k, \theta) \right] dF^\alpha(\theta|x) \right\} dM^\alpha(x).$$

For simple exposition, we define the inner integral of the iterated integral above as

$$^u\Delta_k^\alpha(x) \triangleq \int_\Theta \left[u(a_{k+1}, \theta) - u(a_k, \theta) \right] dF^\alpha(\theta|x).$$

For every $u \in \mathcal{U}^{spm}$, the incremental return $u(a_{k+1}, \theta) - u(a_k, \theta)$ is nondecreasing in θ. Thus, it is immediate from the dispersion theorem that $\alpha \succ_L \beta$ implies $^u\Delta_k^\alpha(X) \geq_{\text{icx}} {}^u\Delta_k^\beta(Y)$. Observe that the function $^u\Delta_k^\alpha : \mathcal{X} \to \Re$ crosses the horizontal-axis at the cutoff point x_k.[73] Hence we can simplify the iterated integral into

$$\int_{\mathcal{X}} \mathbb{1}_{\{x \geq x_k\}}(x) {}^u\Delta_k^\alpha(x)\, dM^\alpha(x) = \mathbb{E}\left[{}^u\Delta_k^\alpha(X) \vee 0 \right]. \tag{16}$$

Since the function $\sigma(x) = x \vee 0$ is increasing and convex, $^u\Delta_k^\alpha(X) \geq_{\text{icx}} {}^u\Delta_k^\beta(Y)$ gives us

$$\mathbb{E}\left[{}^u\Delta_k^\alpha(X) \vee 0 \right] \geq \mathbb{E}\left[{}^u\Delta_k^\beta(Y) \vee 0 \right].$$

Consequently,

$$\mathcal{V}^\pi(\alpha, u) - \mathcal{V}^\pi(\beta, u) = \sum_{k=1}^{N-1} \mathbb{E}\left[{}^u\Delta_k^\alpha(X) \vee 0 \right] - \mathbb{E}\left[{}^u\Delta_k^\beta(Y) \vee 0 \right] \geq 0 \quad \forall\, \pi \in P(\Theta),\ u \in \mathcal{U}^{spm},$$

and therefore, α is more informative than β with respect to the class $\mathcal{U}^{spm}$. □

D. Covert Information Acquisition in Auctions

This section addresses the value of information to bidders in an auction, where each bidder *covertly* obtains more precise information on the object for sale prior to participating in the auction. We show that more Lehmann-precise information is valuable in every incentive-compatible auction. Although this result is well-known in some standard auction mechanisms (such as the first- and second-price auction), the previous proof heavily depends on the single-crossing condition (SCC) of the bidder's payoff function in each auction setting.[74] Rather than the SCC condition, we use the dispersion theorem to extend this result to all incentive-

[73] Even in the case it does not cross the axis at a single point, the function $^u\Delta_k^\alpha$ will assume 0 at the cutoff point and so the equation (16) must hold.

[74] The single-crossing condition requires in game environments that each player i's payoff function satisfies the single-crossing property in $(a_i; \theta_i)$ if his opponents adopt a nondecreasing strategy. Athey (2001) develops this condition to prove the existence of pure-strategy Nash equilibria in games with incomplete information.

compatible auctions.

For an illustration, suppose that there are two risk-neutral bidders who compete for a single indivisible object.[75] The actual value of the object to bidder $i = 1, 2$ is given by $v_i(\theta_i, \theta_j)$, where $(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2 = \Theta$ denotes the *informational variable* pertaining to the value of the object. We assume that the function v_i is increasing in each argument.

Prior to the bidding stage, each bidder i is able to acquire information on θ_i by purchasing a signal α. As a result, he privately observes an outcome $x_i \in \mathcal{X}_i = [\underline{x}_i, \overline{x}_i]$, which is drawn from the distribution $G_i^\alpha(x_i|\theta_i)$. Let π denote the common prior on the space Θ, $\pi_i(\cdot|\theta_i)$ the conditional belief on Θ_j on θ_i, and $F_i^\alpha(\cdot|x_i)$ the posterior belief on Θ_i given $X_i = x_i$. Given the information structure of α, we write the expected value of the object to bidder i conditional on x_i as

$$v_i^\alpha(x_i) = \mathbb{E}\left[v_i(\theta_1, \theta_2)|X_i = x_i\right] = \iint\limits_{\Theta} v_i(\theta_1, \theta_2)\, d\pi_i(\theta_j|\theta_i)\, dF_i^\alpha(\theta_i|x_i).$$

An *auction mechanism* is characterized by a pair $\langle \mathbf{p}, \mathbf{t} \rangle$ where $\mathbf{p} = (p_1, p_2)$ is a *assignment scheme* with $0 \leq p_i \leq 1$ and $\sum_i p_i \leq 1$ indicating the probability that the bidder i is awarded the object, and $\mathbf{t} = (t_1, t_2)$ is a *payment scheme* with $t_i \in \Re$ indicating the transfer from the bidder i to the auctioneer.

Due to the celebrated revelation principle, we can focus on the direct mechanism in which both the assignment and the transfer schemes are mappings defined a vector of announced types. That is, $p_i : \mathcal{X}_1 \times \mathcal{X}_2 \to [0, 1]$ and $t_i : \mathcal{X}_1 \times \mathcal{X}_2 \to \Re$ for each i. Given the direct auction mechanism $\langle \mathbf{p}, \mathbf{t} \rangle$, we define for each bidder i the conditional expected assignment and payment schemes on the signal's outcome $x_i \in \mathcal{X}_i$ as follows:

$$\widehat{p}_i(x_i) \triangleq \mathbb{E}\left[p_i(x_i, X_j)|X_i = x_i\right] = \iint\limits_{\Theta} \mathbb{E}_{X_j}\left[p_i(x_i, X_j)|\theta_j\right] d\pi_i(\theta_j|\theta_i) dF_i^\alpha(\theta_i|x_i),$$

$$\widehat{t}_i(x_i) \triangleq \mathbb{E}\left[t_i(x_i, X_j)|X_i = x_i\right] = \iint\limits_{\Theta} \mathbb{E}_{X_j}\left[t_i(x_i, X_j)|\theta_j\right] d\pi_i(\theta_j|\theta_i) dF_i^\alpha(\theta_i|x_i).$$

These two functions help us simplify the bidder i's expected payoffs from $\langle \mathbf{p}, \mathbf{t} \rangle$, conditioned on x_i, into

$$u_i^\alpha(\tau_i, x_i) \triangleq \mathbb{E}\left[\, p_i(\tau_i, X_j) \cdot v_i^\alpha(x_i) - t_i(\tau_i, X_j) \middle|\, X_i = x_i\right] = \widehat{p}_i(\tau_i) \cdot v_i^\alpha(x_i) - \widehat{t}_i(\tau_i),$$

when he reports another outcome $\tau_i \in \mathcal{X}_i$.

We say that the direct auction mechanism $\langle \mathbf{p}, \mathbf{t} \rangle$ is *individual rational* if (IR) $u_i^\alpha(x_i) \triangleq u_i^\alpha(x_i, x_i) \geq 0$ for all $x_i \in \mathcal{X}_i$ and $i = 1, 2$, and *incentive compatible* if (IC) $x_i \in \operatorname{argmax}_{\tau_i \in \mathcal{X}_i} u_i^\alpha(\tau_i, x_i)$.

By the *constraint simplification theorem*,[76] the mechanism satisfies the (IC) condition if and

[75] Like Persico (2000), we consider the two-bidder case only for clean exposition. In case of overt information acquisition, however, the result depends on the number of bidders as illustrated in Section 5.1.

[76] Refer to p105 in Milgrom (2004).

only if (M) $\widehat{p}_i(\cdot)$ is nondecreasing for all i and the next envelope formula holds:

$$\text{(ET)}\quad u_i^\alpha(x_i) = u_i^\alpha(\underline{x}_i) + \int_{\underline{x}_i}^{x_i} \frac{d}{d\tau} v_i^\alpha(\tau)\ \widehat{p}_i(\tau)\ d\tau.$$

Then we apply the change of variable, $\overline{p}_i(v_i^\alpha(\cdot)) = \widehat{p}_i(\cdot)$, to rewrite (ET) in terms of the estimated values v_i^α. For this purpose, we note that $v_i^\alpha(\cdot)$ is increasing, since logsupermodularity of the posterior beliefs implies the First Order Stochastic Dominance. It implies in turn that $v_i^\alpha(x_i)$ is a sufficient statistic for x_i. Hence the direct auction mechanism $\langle \widehat{\mathbf{p}}, \widehat{\mathbf{x}} \rangle$ is equivalent to $\langle \overline{\mathbf{p}}, \overline{\mathbf{x}} \rangle$.

Hence we can characterize the (IC) condition as follows: the auction mechanism is incentive compatible if and only if (M') $\overline{p}_i(\cdot)$ is increasing for all i and the next envelope theorem holds:

$$\text{(ET')}\quad U_i(v_i^\alpha) = U_i(\underline{v}_i^\alpha) + \int_{\underline{v}_i^\alpha}^{v_i^\alpha} \overline{p}_i(\tau)\ d\tau,$$

where $\underline{v}_i^\alpha = v_i^\alpha(\underline{x}_i)$ is the lower bound of value estimates to bidder i and $U_i(v_i^\alpha(x_i)) = u_i^\alpha(x_i)$ for each x_i. We write the *ex ante* payoffs to the bidder i, generated by the signal α, as

$$\mathcal{U}_i(\alpha) \triangleq \mathbb{E}[U_i(v_i^\alpha)] = \int U_i(v)\ dH_i^\alpha(v),$$

where $H_i^\alpha(v) \triangleq \Pr(v_i^\alpha(X) \leq v) = M^\eta\{x_i \in \mathcal{X}_i | v_i^\alpha(x_i) \leq v\}$ is the distribution function generated by v_i^α. We are now ready to state the following result:

PROPOSITION D.1 (Value of Information to Bidders). *If $\alpha \succ_L \beta$, then $\mathcal{U}_i(\alpha) \geq \mathcal{U}_i(\beta)$.*

PROOF OF PROPOSITION D.1 : Notice that, by (ET'), in the incentive compatible auction mechanism $\langle \mathbf{p}, \mathbf{x} \rangle$ the interim expected payoff $U_i(\cdot)$ is the define integral of the monotone function $\overline{p}_i$. Hence it is twice differentiable almost everywhere and is increasing convex. By Corollary 4.5, $\alpha \succ_L \beta$ implies $v_i^\alpha \geq_{\text{icx}} v_i^\beta$, and thus $\mathcal{U}_i(\alpha) \geq \mathcal{U}_i(\beta)$. □

Intuitively, more precise information about each bidder's true value results in the higher dispersion of his value estimates and thus it will increase his *information rents*. The envelope theorem plays a central role in deriving this result along with the dispersion theorem. The theorem provides an alternative expression to the bidder's information rents as an increasing convex function of the value estimates in every incentive compatible auction mechanism.

E. Proof of Proposition 5.7

The proof begins with a simple property of the $\geq_{\text{icx}}$ order:

LEMMA E.1. *Let X be a random variable with the distribution F and $\sigma_1, \sigma_2 : \Re \to \Re$ be nondecreasing. Suppose that $\sigma_1(X) \geq_{icx} \sigma_2(X)$. Then for every c in the support of F, we have*

$$\int \mathbb{1}_{\{x \geq c\}}(x)\sigma_1(x) dF(x) \geq \int \mathbb{1}_{\{x \geq c\}}(x)\sigma_2(x) dF(x).$$

PROOF OF LEMMA E.1 : Given a constant c in the support of F, define $c_1 = \sigma_1(c)$. Since $\sigma_1(X) \geq_{\text{icx}} \sigma_2(X)$, $\mathbb{E}\left[\sigma_1(X) \vee c_1\right] \geq \mathbb{E}\left[\sigma_2(X) \vee c_1\right]$. Using integration by parts, we can rewrite the inequality as

$$c_1 F(c) + \int \mathbb{1}_{\{x \geq c\}}(x)\sigma_1(x)dF(x) \geq c_1 F(\tau) + \int \mathbb{1}_{\{x \geq \tau\}}(x)\sigma_2(x)dF(x),$$

where the constant τ is determined by $\sigma_2(\tau) = c_1$. Moreover, since σ_2 is nondecreasing, the integral on the right-hand side is larger than

$$c_1 F(c) + \int \mathbb{1}_{\{x \geq c\}}(x)\sigma_2(x)dF(x).$$

Therefore, the desired result follows. $\square$

PROOF OF PROPOSITION 5.7 : Using the same methods in Appendix C, we can write the expected payoff of the principal from α as

$$R(\alpha) = E_\theta\left[v(a_1, \theta)\right] + \sum_{k=1}^{N-1} \int_{x_k}^{1} {}^{v}\Delta_k^\alpha(x)dx,$$

where the sequence of the cutoff points $\{x_1, \cdots, x_{N-1}\}$ represents the optimal strategy of the agent with the signal α. Similarly, the expected payoff of the principal from another signal β can be written

$$R(\beta) = E_\theta\left[v(a_1, \theta)\right] + \sum_{k=1}^{N-1} \int_{y_k}^{1} {}^{v}\Delta_k^\beta(y)dy.$$

Note that for each $k = 1, \cdots, N-1$, the above integral in $R(\alpha)$ can be decomposed into

$$\int_{x_k}^{1} {}^{v}\Delta_k^\alpha(x)dx = \int_{y_k}^{1} {}^{v}\Delta_k^\alpha(x)dx + \int_{x_k}^{y_k} {}^{v}\Delta_k^\alpha(x)dx.$$

Observe that the aligned preference is sufficient for the second integral of the right-hand side to be nonnegative. Furthermore, since $v \in \mathcal{U}^{spm}$, it follows from Corollary 4.5 that $\alpha \succ_L \beta$ implies ${}^{v}\Delta_k^\alpha(X) \geq_{\text{icx}} {}^{v}\Delta_k^\beta(Y)$. By Lemma E.1, therefore, we have

$$\int_{x_k}^{1} {}^{v}\Delta_k^\alpha(x)dx \geq \int_{y_k}^{1} {}^{v}\Delta_k^\alpha(x)dx \geq \int_{y_k}^{1} {}^{v}\Delta_k^\beta(y)dy.$$

As a result, the marginal value of precise information for the principal

$$R(\alpha) - R(\beta) = \sum_{k=1}^{N-1} \left\{ \int_{x_k}^{1} {}^{v}\Delta_k^\alpha(x)dx - \int_{y_k}^{1} {}^{v}\Delta_k^\beta(y)dy \right\}$$

is nonnegative. The proof is complete. $\square$

References

AGHION, P., AND J. TIROLE (1997): "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105(1), 1–29.

ALONSO, R., AND N. MATOUSCHEK (2008): "Optimal Delegation," *Review of Economic Studies*, 75(1), 259–293.

ANGELETOS, G.-M., AND A. PAVAN (2007): "Efficient Use of Information and Social Value of Information," *Econometrica*, 75(4), 1103–1142.

ATHEY, S. (2001): "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information," *Econometrica*, 69(4), 861–89.

ATHEY, S., AND J. LEVIN (2001): "The Value of Information in Monotone Decision Problems," Working Papers 01003, Stanford University, Department of Economics.

BERGEMANN, D., AND S. MORRIS (2013): "The Comparison of Information Structures in Games: Bayes Correlated Equilibrium and Individual Sufficiency," Cowles Foundation Discussion Papers 1909, Cowles Foundation for Research in Economics, Yale University.

BERGEMANN, D., AND M. PESENDORFER (2007): "Information structures in optimal auctions," *Journal of Economic Theory*, 137(1), 580–609.

BERGEMANN, D., AND J. VÄLIMÄKI (2002): "Information Acquisition and Efficient Mechanism Design," *Econometrica*, 70(3), 1007–1033.

——— (2006): "Information in Mechanism Design," CEPR Discussion Papers 5494, C.E.P.R. Discussion Papers.

BLACKWELL, D. (1951): *Comparison of Experiments*. Proceedings of the Second Berkeley Symposium on Mathematics, Statistics and Probability.

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

CRAWFORD, V. P., AND J. SOBEL (1982): "Strategic Information Transmission," *Econometrica*, 50(6), 1431–51.

FOLLAND, G. B. (1999): *Real Analysis : Modern Techniques and their Applications*. Wiley Interscience.

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

GROSSMAN, S. J., AND O. D. HART (1983): "An Analysis of the Principal-Agent Problem," *Econometrica*, 51(1), 7–45.

HERNANDO-VECIANA, . (2009): "Information acquisition in auctions: Sealed bids vs. open bids," *Games and Economic Behavior*, 65(2), 372–405.

HOLMSTRÖM, B. (1977): "On Incentives and Control in Organizations," Ph.D. thesis, Standford University.

——— (1979): "Groves' Scheme on Restricted Domains," *Econometrica*, 47(5), 1137–44.

——— (1984): "On The Theory of Delegation," Discussion Papers 438, Northwestern University, Center for Mathematical Studies in Economics and Management Science.

JEWITT, I. (1988): "Justifying the First-Order Approach to Principal-Agent Problems," *Econometrica*, 56(5), 1177–90.

——— (1989): "Choosing Between Risky Prospects: The Characterization of Comparative Statics Results, and Location Independent Risk," *Management Science*, 35(1), 60–70.

——— (2007): "Information Order in Decision and Agency Problems," Discussion paper, Working Paper.

JOHNSON, J. P., AND D. P. MYATT (2006): "On the Simple Economics of Advertising, Marketing, and Product Design," *American Economic Review*, 96(3), 756–784.

KARLIN, S., AND H. RUBIN (1956): "The Theory of Decision Procedures for Distributions with Monotone Likelihood Ratio," *The Annals of Mathematical Statistics*, 27(2), 272–299.

KIM, S. K. (1995): "Efficiency of an Information System in an Agency Model," *Econometrica*, 63(1), 89–102.

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

LEWIS, T. R., AND D. E. M. SAPPINGTON (1994): "Supplying Information to Facilitate Price Discrimination," *International Economic Review*, 35(2), 309–27.

MCLEAN, R., AND A. POSTLEWAITE (2002): "Informational Size and Incentive Compatibility," *Econometrica*, 70(6), 2421–2453.

MILGROM, P. (2004): *Putting Auction Theory to Work*. Cambridge University Press: Cambridge.

MILGROM, P., AND J. ROBERTS (1990): "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58(6), 1255–77.

MILGROM, P., AND C. SHANNON (1994): "Monotone Comparative Statics," *Econometrica*, 62(1), 157–80.

MILGROM, P. R. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12(2), 380–391.

MILGROM, P. R., AND R. J. WEBER (1982): "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50(5), 1089–1122.

MIZUNO, T. (2006): "A Relation between Positive Dependence of Signal and the Variability of Conditional Expectation Given Signal," *Journal of Applied Probability*, 43(4), 1181–1185.

MORRIS, S., AND H. S. SHIN (2002): "Social Value of Public Information," *American Economic Review*, 92(5), 1521–1534.

PERSICO, N. (2000): "Information Acquisition in Auctions," *Econometrica*, 68(1), 135–148.

QUAH, J. K.-H., AND B. STRULOVICI (2009): "Comparative Statics, Informativeness, and the Interval Dominance Order," *Econometrica*, 77(6), 1949–1992.

ROTHSCHILD, M., AND J. E. STIGLITZ (1970): "Increasing risk: I. A definition," *Journal of Economic Theory*, 2(3), 225–243.

SHAKED, M., AND J. G. SHANTHIKUMAR (2007): *Stochastic Orders and Their Applications*. Academic Press, Inc.

SHI, X. (2012): "Optimal Auctions with Information Acquisition," *Games and Economic Behavior*, 74(2), 666–686.

SZALAY, D. (2009): "Contracts with endogenous information," *Games and Economic Behavior*, 65(2), 586–625.

TCHEN, A. H. (1980): "Inequalities for Distributions with Given Marginals," *The Annals of Probability*, 8(4), 814–827.

TOPKIS, D. M. (1978): "Minimizing a Submodular Function on a Lattice," *Operations Research*, 26, 305–321.

——— (1998): *Supermodularity and Complementarity*. Princeton University Press.

WALD, A. (1950): *Statistical Decision Functions*. John Wiley & Sons, New York.