THE VALUE OF INFORMATION AND DISPERSION

by

Chang Koo Chi

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Economics)

at the

UNIVERSITY OF WISCONSIN–MADISON

2014

Date of final oral examination: June 24, 2014

The dissertation is approved by the following members of the Final Oral Committee:
Daniel Quint, Associate Professor of Economics, Advisor
Lones Smith, Professor of Economics
William H. Sandholm, Professor of Economics
Marzena Rostek, Associate Professor of Economics
Briana Chang, Assistant Professor of School of Business, Finance

to my parents

ACKNOWLEDGMENTS

Many people have contributed to the completion of this dissertation. Even at this moment, a lot of people coursed through my mind. It is a pleasure to acknowledge my indebtedness to them, although I've never been the type of person to express how I feel verbally.

First and foremost I would like to thank my advisor Daniel Quint for his intellectual guidance. Daniel managed to put up with my struggles for almost 7 years, remained confident in my ability to write a good paper when I got bogged down, tried to make time for the meeting even when he was engaged, and taught me modern economic theory in ECON 805 class. I deeply appreciate his guidance, encouragement, supervision, and tolerance. This dissertation would never come out without his efforts he has done for me so far.

I would also like to show my appreciation to Lones Smith and Bill Sandholm as main committee members who helped me complete this dissertation. Lones not only taught me a lot of concepts and mathematical tools, directly related to this dissertation, but also inspired me with a love of economic theory and showed me how to be a successful theorist. Bill taught me game theory in the first two years and was pivotal in my decision to study economic theory. Their guidance has been in valuable over the years.

I have benefited from conversations with many professors in Madison. Discussion with Ricardo Serrano-Padial and Antonio Penta right after my job practice talk helped me polish up the main result of this dissertation. Marzena Rostek and Marek Weretka encouraged me to prepare for the job market with confidence. I am also grateful to Steven Durlauf, Chao Fu, Amit Ghandi, Bruce Hansen, Kenneth Hendricks, John Kennan, and Jack Porter. Especially, when I was about to give up this research, Jack's emails and words of encouragement gave me strength to carry on and helped me along the way.

During my graduate school life, I was fortunate to befriend many amazing people. Ignacio Monzon and Michael Rapp were my TAs for microeconomics and kindly listened to my previous researches over the years. Carl Sanders, Peter

Newberry, Jeff Traczynski, Camila Casas, Mai Seki, Wallice Ao, and Xiaodong Fan were my valued classmates and officemates who shared and rode out many hardships in Madison, and all of them are now pursuing academic career in a variety of top research schools and institutions. Beyond my classmates, I was lucky to meet Jiabin Wu, Francisco Franchetti, Fumihiko Suga, Jianrong Tian, Yu Zhu, Lihan Liu, Serafin Grundl, Hector Sandoval, and Suphanit Piyapromdee. I am especially grateful to Jianrong. Many contents of this dissertation were built upon numerous discussions with him.

I also thank my valuable Korean friends; Hojun Lee, Sang Yoon Lee, Kyuchol Jung, Kyungjin Choi, Heewoo Kang, and Seojeong Lee. Without their friendship and help I could not have completed my degree. Hojun was a good role model as well as a great mentor. Sang Yoon spared no effort to help me throughout my graduate school life. Heewoo made a lot of time to listen to my research while preparing for the job market together and allowed me to share his hotel room for two days in the ASSA meeting.

Finally, I wish to dedicate this dissertation to my parents. This work was possible only thanks to their constant love and support. There are no words that can express my gratitude towards them. I will never forget their selfless love and great sacrifices.

Last and most to So Jung for everything.

Chi, Chang Koo
June, 2014

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In a standard decision problem under uncertainty, the decision maker gathers information about a payoff-relevant state of the world from a set of possible signals. The dissertation aims to shed light on the nature of better information in a decision problem, establishing sufficient and necessary conditions for one signal to be more valuable than another. This analysis—called "*comparison of signals (or experiments)*"—is carried out for a general class of decision problems with complementarity between the decision maker's choice and the state of uncertainty, which arise in many economic decision models.

The first chapter is preliminary. It deals with definitions in which one signal can be judged to be better than another and proposes two distinct approaches to formulating a partial ordering on the set of signals. One, the preference-based approach, treats a signal as a good that generates value to the decision maker, quantifies the value for each signal, and arranges signals by their value. Two signal orderings—based on the criterion of effectiveness and informativeness—are considered within this approach. The other, the statistical approach, treats a signal as a statistical device that imparts relevant information to the decision maker, focuses on the statistical characteristics of signals, and rank them by a statistical notion. Taking this approach, another two signal orderings—based on Lehmann precision and dispersion—are studied with an emphasis on their relationship.

The second chapter presents the main result of the dissertation. It is shown that the four different signal orderings presented in the previous chapter are mutually equivalent within each class of the three payoff functions: supermodular, single-crossing, and interval dominance order. This equivalence theorem provides several important implications. First, Lehmann precision is both sufficient and necessary for one signal to be more effective and more informative than another within each of the three classes. This clarifies how to value information in both statistical and Bayesian decision theory by establishing that Lehmann precision constitutes a statistical ordering representing a statistical and Bayesian decision maker's preference-based orderings. Second, the theorem exactly characterizes the relationship between more

precise signals and higher dispersion by coupling the precision-based ordering with the dispersion-based ordering. This result—called the dispersion theorem—justifies another signal orderings used in the previous literature.

The third chapter examines the impact of precise information in equilibrium models by utilizing the dispersion theorem. Although the first implication above fails in the presence of strategic interactions, the second implication—the equivalence between the two statistical orderings—holds and can be applied to several settings. The first application is to auctions. It deals with how the seller's activity of supplying precise information to participants impinges on efficiency, the expected revenue, and the bidder's information rent. The second is to the principal-agent model with hidden action. It is shown 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 is to delegation. A sufficient condition is provided for the value of delegating the principal's right to make a decision to be monotonic in the level of agent's expertise.

1 INTRODUCTION

1.1 Overview of the Results

Many economic decisions get made based on imperfect information about payoff-relevant variables. An investor chooses investment portfolios with partial information about the future return of financial assets. A firm hires workers with partial information about their innate talents. A buyer places a bid with imperfect information about the true value of an object for sale. 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 garner information relevant to her choices. This paper studies 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]

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). Given a category of preferences, 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. This signal ordering, as induced by the decision maker's primitive preference, is referred to as the *preference-based ordering*. Without any structures on preferences, Blackwell linked the question of ordering signals to the statistical notion of sufficiency and showed the equivalence between the preference-based ordering and the statistical ordering based on sufficiency. Subsequently, Lehmann (1988) developed a more complete statistical ordering based on the notion of precision than sufficiency and established the equivalence between the two orderings within a certain category of preferences in statistical contexts. 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 general economic environments. To model an economic decision problem, I consider three large classes of preferences: supermodular, single-crossing, and interval dominance order payoff functions. All

[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.

of these three classes exhibit a complementarity between the decision maker's choice variable and the payoff-relevant state of the world in the sense that if an increase in the choice variable is desirable at some state, it remains desirable at every higher state.[2] The preference classes I consider are large enough to encompass a variety of the 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 derive a sufficient and necessary condition under which one signal is more valuable than another for all decision makers within the class.

For this purpose, I develop a novel statistical ordering based on the notion of *dispersion*. To state it more precisely, let α and β be two available signals we want to compare and let θ represent the state variable. The dispersion-based order concerns the decision maker's predictions about θ:

(D) The signal α generates more dispersed predictions about θ than β ($\alpha \succ_D \beta$ in symbols) if for every nondecreasing function ψ defined on θ, the expected value of ψ conditional on α is more variable than the expected value of ψ on β in the second-order stochastic dominance.

To see how the dispersion-based ordering leads to higher expected payoffs, consider a simple example in which a firm decides whether to invest K in R&D to improve its production technology. The returns to R&D are uncertain and given by a function ψ nondecreasing in θ, which reflects complementarity between the firm's investment decision and the state. Normalizing the payoffs from action "Not Invest" to 0, we can write the firm's ex ante expected payoffs generated by the signal α as

[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.

$\mathbb{E}[\max\{\mathbb{E}[\psi(\theta)|\alpha] - K, 0\}]$, where $\mathbb{E}[\psi(\theta)|\alpha]$ represents the expected returns given α. As $\mathbb{E}[\psi(\theta)|\alpha]$ becomes more diffuse or dispersed, therefore, the firm's expected payoffs would increase since the operator max is convex.

The main contributions of this paper are threefold. First, I show that the dispersion-based 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 the 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 that she acts to maximize the expected payoff for a worst case scenario. 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 $\mathcal{U}^\star$.

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 the 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 $\mathcal{U}^\star$.

I show that the dispersion-based order 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 $\mathcal{U}^{ido}$ payoff functions. 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 with an arbitrary but fixed prior belief, a more restrictive statistical condition is required for the preference-based ordering.[3] In contrast to their finding, insofar as we enrich the set of prior beliefs by allowing the decision maker to have any prior, 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 the values of signals whether complementarity between the decision maker's choice and the state is of a cardinal or an ordinal property.

The second contribution of this paper is the characterization of $\alpha \succ_D \beta$ 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 3.5 demonstrates that this statistical ordering, referred to as "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

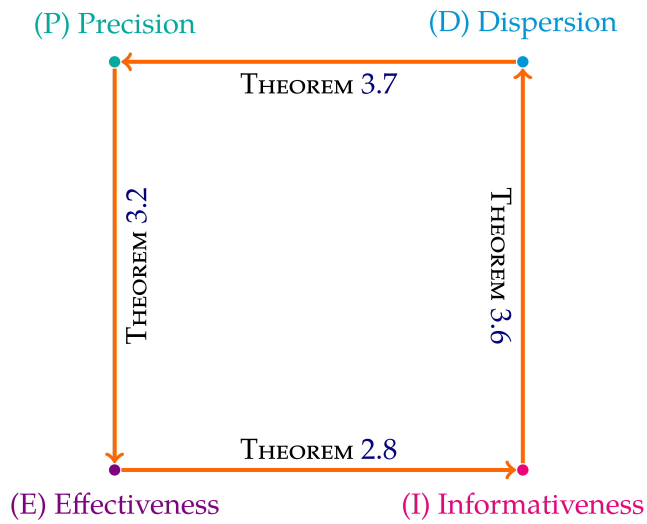
[3]The idea can be traced back to Lehmann (1988). Unlike Blackwell, who does not place any limitations on the decision maker's preferences, Lehmann restricts the scope of payoff functions to the set of Karlin-Rubin Monotone payoff functions. As a result, his statistical notion of precision is less restrictive than Blackwell's sufficiency. See Section 2.5 for its details.

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.

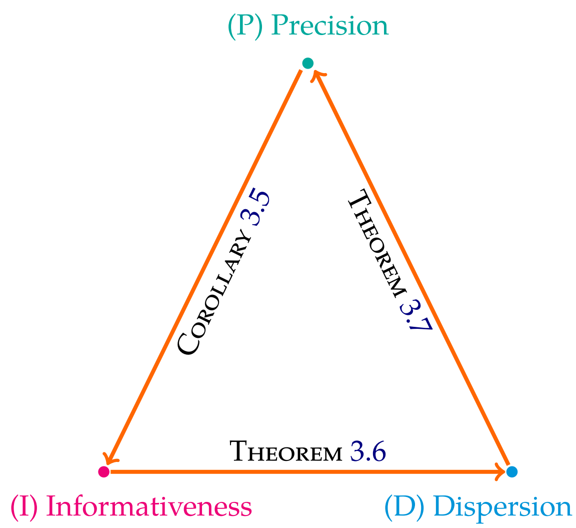
To my knowledge, this characterization theorem is the most general formulation of the link between Lehmann precision and the dispersion. Theorem 5.2 in Lehmann (1988) provides a similar characterization result when the associated density functions with signals are logconcave. However, the theorem of 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 displayed in Figure 1.1-(a). Under the Bayesian framework, this result reduces to the equivalence between the three orderings—Lehmann precision, informativeness, and dispersion—as depicted in Figure 1.1-(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: $\mathbb{E}[u(X)] \geqslant \mathbb{E}[u(Y)]$ 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 1.1-(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



(a) Decision Theory



(b) Bayesian Decision Theory

Figure 1.1: 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.

Preference	Statistical Order		Preference-based Order
SPM	$\alpha \succ_{MIO-ND} \beta$	$\xleftrightarrow{\text{Athey and Levin (2001)}}$	$\alpha \succ_{I}^{spm} \beta$
SC	$\alpha \succ_{L} \beta$	$\xrightarrow{\text{Quah and Strulovici (2009)}}$	$\alpha \succ_{E}^{sc} \beta$, $\alpha \succ_{I}^{sc} \beta$
IDO	$\alpha \succ_{L} \beta$	$\xrightarrow{\text{Quah and Strulovici (2009)}}$	$\alpha \succ_{E}^{ido} \beta$, $\alpha \succ_{I}^{ido} \beta$

Table 1.1: Literature Review

without bilateral money transfers. I provide a sufficient condition under which more precise private information for the agent is more valuable to the principal.

1.2 Review of Precedent Literature

This paper is related to large literature on the value of information in two strands of decision theory. The classical theory of characterizing a binary relation over a set of signals has been initiated by Blackwell (1951, 1953) in statistics. In the general decision-making framework with a discrete state space, he developed the statistical ordering based on sufficiency and showed that α is Blackwell-sufficient for β if and only if α is more effective and informative than β regardless of the decision maker's preference. As a natural consequence, Blackwell sufficiency is a double-edged statistical ordering: although being powerful, it is too restrictive to hold in practice. In light of this weakness Lehmann (1988) restricted attention to the class of Karlin-Rubin monotone (KRM) payoff functions and developed the statistical notion of precision, establishing the equivalence between the statistical ordering based on precision and the preference-based ordering based on effectiveness within the class. This paper extends Lehmann Theorem to several classes of payoff functions used in economics.

The equivalence theorem presented in Section 3 complements several work by Persico (2000), Athey and Levin (2001), Jewitt (2007), and Quah and Strulovici

(2009).[4] In the Bayesian framework, Athey and Levin (2001) suggested a novel statistical ordering (called "MIO-ND") based on the posterior beliefs and proved that α is larger than β in this order if and only if α is more informative than β within the class of supermodular payoff functions. My result extends their equivalence theorem to the effectiveness within the same scope, and relates their statistical ordering to Lehmann precision and the dispersion-based ordering.

The main contribution of this paper can be best understood by comparison with Quah and Strulovici (2009). Within a new class of IDO payoff functions including other classes considered in this paper, they showed that Lehmann precision is sufficient for ordering signals based on the informativeness and effectiveness. The dispersion-based ordering plays two important roles of extending their result. First, it enables us to show necessity of Lehmann precision and thereby to generalize the Lehmann's equivalence theorem to their IDO class. Second, the ordering is used to show that this equivalence theorem still holds true even in the supermodular class. Put differently, Lehmann precision is essential and irreplaceable for one signal being more valuable than another as long as complementarity between the decision maker's choice and the state is incorporated into the payoff function.

1.3 Outline of the Dissertation

The dissertation is organized as follows: Chapter 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. Chapter 3 presents the main result: I prove that the four signal orderings above are mutually equivalent in the order illustrated in Figure 1.1-(a). First, I show in Theorem 3.2 that Lehmann precision is sufficient for ordering signals based on their statistical

[4] Applications of Lehmann precision to economic environments has been initiated by Persico (2000) within the class of single-crossing payoff functions. Jewitt (2007) first pointed out the difference between the single-crossing and the KRM payoff functions used in Lehmann (1988). Also he employed the envelop theorem to extend Lehmann Theorem to the single-crossing payoff functions in the Bayesian framework.

Preference	Statistical Order		Preference-based Order
SPM	$\alpha \succ_D \beta \Leftrightarrow \alpha \succ_L \beta$	$\longleftrightarrow$	$\alpha \succ_E^{spm} \beta, \alpha \succ_I^{spm} \beta$
SC	$\alpha \succ_D \beta \Leftrightarrow \alpha \succ_L \beta$	$\longleftrightarrow$	$\alpha \succ_E^{sc} \beta, \alpha \succ_I^{sc} \beta$
IDO	$\alpha \succ_D \beta \Leftrightarrow \alpha \succ_L \beta$	$\longleftrightarrow$	$\alpha \succ_E^{ido} \beta, \alpha \succ_I^{ido} \beta$

Table 1.2: For each class of preferences, the dispersion-based ordering $\alpha \succ_D \beta$ the signals in accordance with the decision maker's preference. The Dispersion Theorem associates (D) $\alpha \succ_D \beta$ with (P) $\alpha \succ_L \beta$, establishing their equivalence.

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.[5] Subsequently, I show in Theorem 3.6 and Theorem 3.7 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. I discuss the three by-products of the main result, with emphasis on the connection between Lehmann precision and another signal orders developed by previous literature. Chapter 4 illustrates how the dispersion theorem is applied to auctions, bilateral contracts, and delegation. Chapter 5 contains concluding remarks with some suggestions for future research.

[5]Moreover, I show in Appendix A.1 that the improvement principle, the way to find the payoff-improving strategy, is essentially equivalent to the principle developed by Lehmann (1988).

2 DECISION THEORY AND SIGNAL ORDERINGS

2.1 Introduction

Consider a standard decision problem under uncertainty. A given decision maker is uncertain about a payoff-relevant state of the world θ, and before making a decision, can observe a realization of one random variable which conveys partial information about θ. A signal α is defined as the random variable and its possible distributions according to θ. In this chapter, we begin our study of comparing two signals α and β by considering which one is more valuable in the decision-making process. In Section 2.2, we describe fundamental elements formulating a decision problem under uncertainty and introduce a few mathematical notions and notations which will be employed in this and subsequent chapters. In the remaining sections, we present the formal definitions of the four different partial orderings on a set of signals set out in Chapter 1.[1]

There are two distinct approaches to formalizing the intuitive concept of ordering signals. The first approach, which we introduce in Section 2.3 and 2.4, treats a signal as a good that generates value for the decision maker, quantifies the value for each signal, and specifies a binary relation over the set of signals by their values. We refer to this approach as the *preference-based* signal ordering, since it is induced by the decision maker's primitive payoff function. Having completed the description of a decision problem, the theory is developed by first specifying a principle of decision-making and providing a rational framework for dealing with the problem. Under the two major—statistical and Bayesian—frameworks, we define the criterion of effectiveness and informativeness and illustrate why each criterion furnishes a reasonable signal ordering.

The second approach, which we introduce in Section 2.5 and 2.6, treats a signal as a statistical device that provides a sample correlated with θ to the decision maker and specifies a binary relation by the primitive characteristics of signals, which we refer to as the *statistical* signal ordering. The theory proceeds with a statistical notion. In Section 2.5 we explore the statistical ordering on the basis of Lehmann Precision and provide a simple characterization lemma that will be used in the

[1] A partial ordering $\succ$ on a set S is a reflexive, transitive, and antisymmetric binary relation.

subsequent chapters. Section 2.6 contains a formal statement of our main concept, *dispersion*, and gives an intuitive motivation for what it means for one signal to generate more dispersed predictions about θ than another.

Understanding the relationship between these two approaches is an important problem in theory of decision under uncertainty. As a matter of fact, the problem of comparison of two signals—how to value information—is to identify a statistical ordering that represents the preference-based ordering induced by a class of decision problems. More technically, for a class of payoff functions concerning θ, we look for a statistical ordering $\alpha \succ_{\diamond} \beta$ based on a statistical notion $\diamond$ that is equivalent to rendering α to be more effective or informative than β for every payoff function within the class. This is a main issue in Chapter 3

2.2 Basic Elements of Decision Theory

Information Structures

Let $\langle \Theta, \mathcal{F} \rangle$ denote a measurable space that represents the unknown state of the world. Each element θ in Θ is a complete description of exogenous variables of the model considered, where the state space Θ will be used to denote the set of all possible states. 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 of signal α—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 an observed sample of the random quantity Y.

In this paper we shall assume that both sample spaces $\mathcal{X}$ and $\mathcal{Y}$ are subsets of real numbers. Additionally, we assume *without any loss of generality* that the state space Θ is an interval on the real line $[\underline{\theta}, \overline{\theta}]$.[2] It is supposed that the cumulative

[2]It is worthy of note that the interval state space will not cause any loss of generality. Lemma B in Karlin and Rubin (1956) proves that one can restrict attention to a *statistically connected* state space

distribution of X for each state θ is absolutely continuous and thus it takes a form of

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

where $g^{\alpha}(\cdot|\theta)$ is the density function with respect to Lebesgue measure.[3]

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 property requires that $g^{\alpha}(x|\theta)$ obey the following inequality:

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

In a nutshell, the above inequality means that X assigns more densities to large outcomes in state θ_1 rather than in state θ_2.[4] As is well-known by Karlin and Rubin (1956), the set of probability distributions with the MLRP is large enough that it encompasses the class of probability distributions of the exponential type, to which most of the distributions studied in statistics and economics belong.[5] We label by $\mathcal{S}$ the set of signals of our interest. Therefore, we shall be concerned with

under the environment with the monotone likelihood ratio property.

[3] In case X is discrete, the distribution G^{α} is not necessarily 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 $\Pr(X = x_k|\theta) = g_k(\theta)$ for each $k = 1, \cdots, n$. One can define X^* by the continuous distribution of

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

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

[4] 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) has shown their equivalence when all possible prior beliefs on Θ are taken into account.

[5] A probability distribution is said to be of the exponential type if the density function takes a form of

$$g(x|\theta) = h(\theta)e^{x\theta}r(x) = \frac{e^{x\theta}r(x)}{\int_{\mathcal{X}} e^{s\theta}r(s)ds},$$

$S \triangleq \{\alpha = \langle X, \{G^{\alpha}(x|\theta)\}_{\theta\in\Theta}\rangle |\, g^{\alpha}(x|\theta) \text{ obeys the MLRP}\}$.

Action Space and Decision Rules

Let A indicate a set of undominated actions that the decision maker is allowed to choose in the decision problem concerning θ.[6] For the sake of simplicity, we assume that A is a compact subset of $\mathfrak{R}$.[7] In a decision problem endowed with a signal α, a decision rule or a strategy d^{α} is defined as a measurable mapping from the sample space $\mathcal{X}$ into the action space A. We denote by $\mathcal{D}^{\alpha} = \{d : \mathcal{X} \to A|\text{measurable}\}$ the set of all permissible decision rules based on α. Similarly, $\mathcal{D}^{\beta}$ is the set of all permissible decision rules based on another signal β.

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.[8] We denote by $\mathcal{U}$ the *class* of all payoff functions satisfying these two properties.

---

where $r : \mathcal{X} \to \mathfrak{R}_+$ 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.

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

[7] In decision theory or monotone comparative statics, the action space A is frequently given by a partially ordered set (poset) with a binary relation $\geqslant$ 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 $\mathfrak{R}$. All the results below can be extended into the case that A is a complete lattice.

[8] A family $\{u(\cdot, \theta)\}_{\theta\in\Theta}$ of real-valued functions $u(\cdot, \theta) : A \to \mathfrak{R}$ 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.

Many 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 one action $a' > a$ is desirable at some state θ rather than a, it remains desirable at every state $\theta' \geqslant \theta$. For example, a standard matching function between firms and employees in matching theory assumes complementarities between the attributes of the firm (say, the firm size) and those of the employee (abilities). In this respect we shall focus on 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*,* $u(a', \theta) - u(a, \theta)$ *is nondecreasing in* θ*.*[9]

Definition 2.2 (SC Family, Milgrom and Shannon (1994))**.** *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*,* $u(a', \theta) - u(a, \theta) \geqslant 0$ *implies* $u(a', \theta') - u(a, \theta') \geqslant 0$ *for all* $\theta' \geqslant \theta$*.*

The SC payoff function expresses a weaker complementarity than the SPM one. Since every nondecreasing function clearly crosses the horizontal axis from below at most once, we have $\mathcal{U}^{spm} \subset \mathcal{U}^{sc}$. Observe that the SC property is ordinal so that the property is preserved by any order-preserving transformations while the supermodularity is cardinal. In addition, the SPM family satisfies the closed convex cone property—if $u_1, u_2 \in \mathcal{U}^{spm}$ then $\lambda_1 u_1 + \lambda_2 u_2 \in \mathcal{U}^{spm}$ for positive λ_1 and λ_2, while the SC family does not have possess this property.[10]

The next class of payoff functions exhibits an even weaker complementarity than $\mathcal{U}^{sc}$ and thus is the largest among the three classes.

---

[9]Since $u(a, \theta)$ is defined on the product of the two ordered sets, the concepts of supermodularity and increasing differences coincide (Topkis (1998)). In case $u(a, \theta)$ is differentiable, $u \in \mathcal{U}^{spm}$ if and only if $\partial^2 u(a, \theta)/\partial a \partial \theta \geqslant 0$. Refer to Milgrom and Roberts (1990a) for applications of supermodularity to economic environments.

[10]Due to this fact, Quah and Strulovici (2012) refined the SC family and introduced a new class of payoff functions, referred to as the *single-ratio monotonicity* functions, which is preserved under aggregation.

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*,*

$$u(a'',\theta) - u(a,\theta) \geqslant 0 \ \textit{for all } a \in [a', a''] \ \textit{implies} \\ u(a'',\theta') - u(a',\theta') \geqslant 0 \ \textit{for all } \theta' \geqslant \theta.$$

The IDO property requires the higher action a'' to be desirable for every $\theta' \geqslant \theta$ only if it is more desirable than any other actions belonging to the interval $[a', a''] = \{a \in A | a' \leqslant a \leqslant a''\}$ at θ. Whereas both the IDO and SC 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 null hypothesis concerning θ or doing point and interval estimation of the unknown parameter θ.[11]

Decision Problem and its Process

A decision problem concerning θ with a single signal $\alpha \in S$ is specified by the collection of the above elements; the state space, an information structure, an action space, and a payoff function;

$$(\Theta, \ \langle X, \{G^{\alpha}(\cdot|\theta)\}_{\theta\in\Theta}\rangle, \ A, \ u(a,\theta)) \, .$$

Since we are mainly concerned about comparison of signals within the above three classes of payoff functions, we denote the decision problem by (α, u) for the simple exposition.

[11] 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 contains the KRM family. In Appendix A.1 and A.2, we will look more in detail at the inclusive relationship among the several classes of payoff functions studied in past literature.

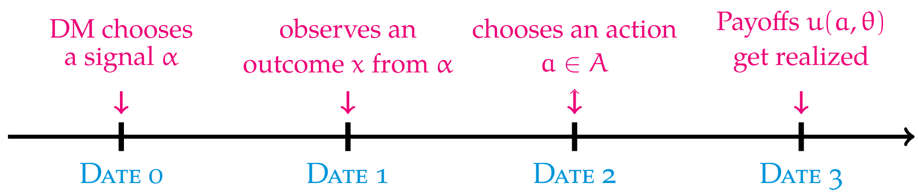


Figure 2.1: Decision-Making Process under Uncertainty

The decision problem (α, u) proceeds as depicted in Figure 2.1: A decision maker observes a realization x of X from the signal α. Then she chooses an action a from the given action space A, on the basis of the observed x, before the state of the world is realized. At the end of the day, the decision maker's payoff is determined by $u(a, \theta)$.

Value of Information

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 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 S 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 higher expected value than (β, u) for every payoff function $u \in \mathcal{U}^\star$. As is highlighted in Section 2.1, 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.

2.3 Effectiveness

Analysis of statistical decision problems using the minimax or minimax regret decision 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

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

This formulation is to treat a decision problem as a two-person statistical game against "nature", where $d \in \mathcal{D}^\alpha$ and $\theta \in \Theta$ are regarded as the action chosen by the statistician and by nature, respectively. 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 $\inf_{\theta \in \Theta} \rho^\alpha(d, \theta)$.[12] 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 $\mathcal{D}^\beta = \{d : \mathcal{Y} \to A | \text{measurable}\}$ based on β.

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 $V(\alpha, u) \geqslant V(\beta, u)$ for

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

every $u \in \mathcal{U}^\star$. The next condition provides a signal ordering in this context:

Definition 2.4 (Effectiveness)**.** *A signal α is more effective than another signal β with respect to a class of payoff functions $\mathcal{U}^\star$ (in symbols, $\alpha \succ^\star_E \beta$) provided for every decision rule $d \in \mathcal{D}^\beta$, there exists a decision rule $d^\alpha \in \mathcal{D}^\alpha$ such that*

$$\rho^\alpha(d^\alpha, \theta) \geqslant \rho^\beta(d, \theta) \quad \textit{for all } \theta \in \Theta. \tag{E}$$

The criterion of effectiveness—the signal ordering based on the statistical value—states that any expected payoff function attainable with β is also attainable with the more effective signal α regardless of the state θ. It is an immediate consequence of the definition that a more effective signal generates higher statistical value 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 $\mathcal{H} = \{h : \Theta \to \mathfrak{R} | \text{ measurable}\}$, i.e., $u \in \mathcal{U}^\star$ implies $u + h \in \mathcal{U}^\star$ for all $h \in \mathcal{H}$. To see this, let $d^\alpha \in \mathcal{D}^\alpha$ indicate a minimax strategy for some payoff function $u \in \mathcal{U}^\star$. Then $V(\alpha, u) \geqslant V(\beta, u)\ \forall\ u \in \mathcal{U}^\star$ implies

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

which is equivalent to $\rho^\alpha(d^\alpha, \theta) \geqslant \rho^\beta(d, \theta)\ \forall\ \theta \in \Theta$ and $\forall\ d \in \mathcal{D}^\beta$. 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 generates higher statistical value 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 section, the criterion of effectiveness becomes the necessary and sufficient condition for α to generate higher statistical value than β.

Lemma 2.5. *For each property $\star \in \{SPM, SC, IDO\}$ and for two signals α and β, $\alpha \succ^\star_E \beta$ if and only if $V(\alpha, u) \geqslant V(\beta, u)$ 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 $d \in \mathcal{D}^\alpha - \mathcal{D}'$, there exists a $d' \in \mathcal{D}'$ such that $\rho^\alpha(d', \theta) \geqslant \rho^\alpha(d, \theta)$ for every θ.

Theorem 2.6 (Essentially Complete Class, Quah and Strulovici (2009))**.** *For every decision problem* (α, u) *with* $u \in \mathcal{U}^{ido}$*, the set of monotone decision rules*

$$\mathcal{D}^{\alpha,M} = \{d : \mathcal{X} \to A | \textit{ measurable and nondecreasing}\}$$

constitutes an essentially complete class.

Proof of Theorem 2.6 : See Appendix A.2. □

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, it is sufficient for the proof of $\alpha \succ_E^\star \beta$ for each class $\star \in \{SPM, SC, IDO\}$ to find a monotone decision rule $d^\alpha \in \mathcal{D}^{\alpha,M}$ such that $\rho^\alpha(d^\alpha, \theta) \geqslant \rho^\beta(d, \theta)$ for every monotone $d \in \mathcal{D}^{\beta,M}$. That is, rather than the set of all permissible decision rules, we need only look at the set of monotone decision rules so the problem becomes much more tractable.

2.4 Informativeness

Bayesian decision theory stems from the assumption that a given decision maker is *Bayesian rational*:

(i) The decision maker has a specific prior belief 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 for each signal α, the marginal distribution of X is derived from the prior belief π: $\Pr(X \leqslant x) = \int_\Theta G^\alpha(x|\theta)d\pi(\theta)$. The marginal distribution will be labeled by $M^\alpha(x)$, if necessary.[13]

(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

$$F^\alpha(t|x) = \Pr(\theta \leqslant t|X = x) = \frac{\int_\Theta \mathbb{1}_{\{\theta \leqslant t\}}(\theta) g^\alpha(x|\theta)d\pi(\theta)}{\int_\Theta g^\alpha(x|\theta)d\pi(\theta)},$$

i.e., the conditional probability of the event $\{\theta \leqslant t\} \in \mathcal{F}$ on the outcome $X = x$. To be consistent with the prior, $\mathbb{E}_X[F^\alpha(t|X)] = \pi(\{\theta \leqslant t\})$ must hold for all $t \in \Theta$, 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 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

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

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 $\pi \in P(\Theta)$. With u and π being fixed, define by $\mathcal{V}^\pi(\beta, u)$ the Bayesian value corresponding to another signal β in a similar manner.

[13] 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.

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

Definition 2.7 (Informativeness)**.** *A 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*

$$\mathcal{V}^{\pi}(\alpha, u) \geqslant \mathcal{V}^{\pi}(\beta, u) \quad \textit{for all } \pi \in P(\Theta) \textit{ and } u \in \mathcal{U}^{\star}. \tag{I}$$

Note that the criterion of informativeness does not depend on the decision maker's prior. In order to meet this criterion, α should yield more ex ante payoffs to Bayesian decision makers for every prior. 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 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.[14]

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.[15]

[14]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 3.5.

[15]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.

So far, we presented the way of ordering signals based on the values. In light of Definition 2.4 and 2.7, 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 2.8. *For every class of payoff functions $\mathcal{U}^\star$, $\alpha \succ_E^\star \beta$ implies $\alpha \succ_I^\star \beta$.*

Proof of Theorem 2.8 : 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 $\rho^\alpha(d^\alpha, \delta) \geqslant \rho^\beta(\delta^\beta, \theta)$ 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 $A^\alpha(x) = \left\{a^* \in A |\ a^* \in \operatorname{argmax}_{a \in A} \mathbb{E}_\theta\left[u(a, \theta)|X = x\right]\right\}$.

Theorem 2.9 (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.*[16]

Like the essentially complete class theorem, the monotone comparative statics theorem is useful in ranking two signals based on the criterion of informativeness.

[16]Given two subsets D_1 and D_2 of $A \subset \mathfrak{R}$, 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 $\geqslant$, $D_1 \trianglerighteq D_2$ provided $d_1 \vee d_2 \in D_1$ and $d_1 \wedge d_2 \in D_2$.

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 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 $\delta^{\alpha} \in \mathcal{D}^{\alpha,M}$ such that $\rho^{\alpha}(\delta^{\alpha}, \theta) \geqslant \rho^{\beta}(d, \theta)$ for every $d \in \mathcal{D}^{\beta,M}$.

2.5 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 and informative than β to every decision maker. 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).[17] This

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

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

the sufficient and necessary condition for α to be Blackwell-sufficient for β is $1/\alpha \in \{1, 2, 3, \cdots\}$, leaves open possibility that a more complete and intuitive ordering can be found by narrowing down the scope of payoff functions.[18]

Definition 2.10 (Lehmann (1988))**.** *A 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*

$$G^{\alpha}(T_y(\theta)|\theta) = G^{\beta}(y|\theta). \tag{P}$$

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

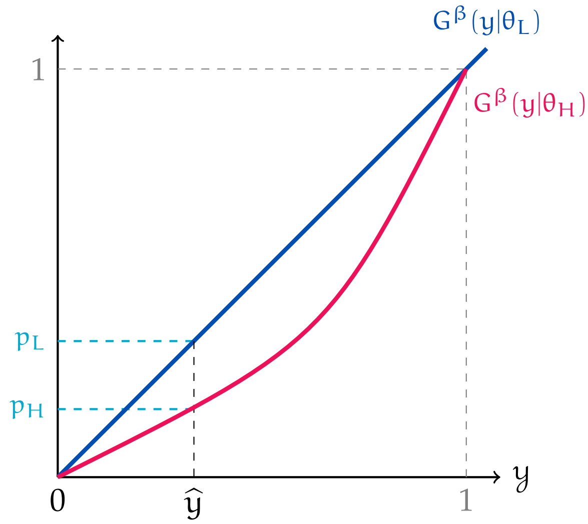
Given the quantile p_L, there is a sample $x \in \mathcal{X}$ at which $G^{\alpha}(x|\theta_L) = p_L$, which we label by $T_{\widehat{y}}(\theta_L)$.[19] Since we assume the identical distributions for θ_L, we have $T_{\widehat{y}}(\theta_L) = \widehat{y}$. When the state is θ_H, however, the monotone T-transformation calls for $T_{\widehat{y}}(\theta_H) \geqslant \widehat{y}$. Hence the graph of $G^{\alpha}(\cdot|\theta_H)$ will be uniformly below the graph of $G^{\beta}(\cdot|\theta_H)$ like the FOSD.[20] As a consequence, the signal α assigns relatively more densities to higher outcomes than β when the state is θ_H, and hence α is more statistically precise than β.

although α conveys more information on θ for every $0 < \alpha < 1$. However, it will be shown in 2.13 that every $\alpha \in (0, 1)$ is ranked higher than β in the notion of Lehmann (defined shortly).

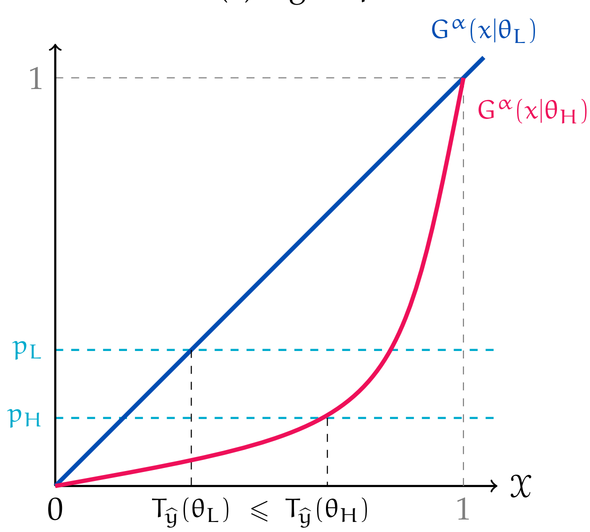
[18] Restricting attention to the class of Karlin-Rubin monotone payoff functions (See Appendix A.1 for its definition.), Lehmann (1988) proved that α is more effective than β within this reduced class if and only if α is more Lehmann-precise than β.

[19] 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.

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



(a) Signal β



(b) Signal α

Figure 2.2: Lehmann (1988) Precision and T-transformation

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

Example 2.11 (Gaussian Learning[21]). *Assume that θ is unknown and the prior of θ is $N(\theta_0, \sigma_0^2)$, where θ_0 and σ_θ are prior mean and standard deviation, respectively. 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*

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

where Φ, the cumulative distribution function of $N(0,1)$, is bijective. Hence the associated T-transformation with these two signal structures is

$$T_y(\theta) = \frac{\sigma_\alpha}{\sigma_\beta}(y-\theta) + \theta,$$

which is nondecreasing in θ if and only if $\sigma_\alpha < \sigma_\beta$. Therefore, it is immediate from Definition 2.10 that $\alpha \succ_L \beta$ if and only if $\sigma_\alpha < \sigma_\beta$.

The next lemma furnishes us a simple but useful characterization of Lehmann precision.[22]

Lemma 2.12. *$\alpha \succ_L \beta$ 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 2.12 : 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 $G^\beta(y|\theta_0) - G^\alpha(x|\theta_0) \geqslant 0$ for some θ_0. Since the distribution function $G^\alpha(\cdot|\theta)$ is nondecreasing, it must be the case that $x \leqslant T_y(\theta_0)$. Thus, for $\theta \geqslant \theta_0$

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

[21] 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.

[22] In some statistics textbook, the Lehmann precision is defined as the statement in Lemma 2.12.

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

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 $\theta \geqslant \theta_0$

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

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

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 $\theta \geqslant \theta_0$, $G(x|\theta)$ 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 2.13 (Uniform Distributions)**.** *Consider the two signals with uniform distributions*

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

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

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

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

Example 2.14 (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.*

$$\text{Signal } \alpha : x \begin{cases} \xrightarrow{1-\alpha} x \sim F(\cdot) \\ \xrightarrow{\alpha} x = \theta \end{cases} \qquad \text{Signal } \beta : y \begin{cases} \xrightarrow{1-\beta} y \sim F(\cdot) \\ \xrightarrow{\beta} y = \theta \end{cases}$$

Note that the cumulative distribution of each signal is written $G^{\alpha}(x|\theta) = \alpha \mathbb{1}_{\{x \geqslant \theta\}}(x) + (1-\alpha)F(x)$ and $G^{\beta}(y|\theta) = \beta \mathbb{1}_{\{y \geqslant \theta\}}(y) + (1-\beta)F(y)$, respectively. We use Lemma 2.12 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 $\theta \geqslant y$ 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 $\theta \geqslant x$, the function becomes negative everywhere. When $\Delta^{x,y}(\theta)$ is positive for $\theta \geqslant x$, 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

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

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 $\theta \leqslant x$ 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 we have seen in the examples above. More importantly, this statistical signal ordering is applicable to more distributions than sufficiency so that it has appeared in recent literature on endogenous information acquisition since Persico (2000) for analyzing the value of information in a variety of economic settings.

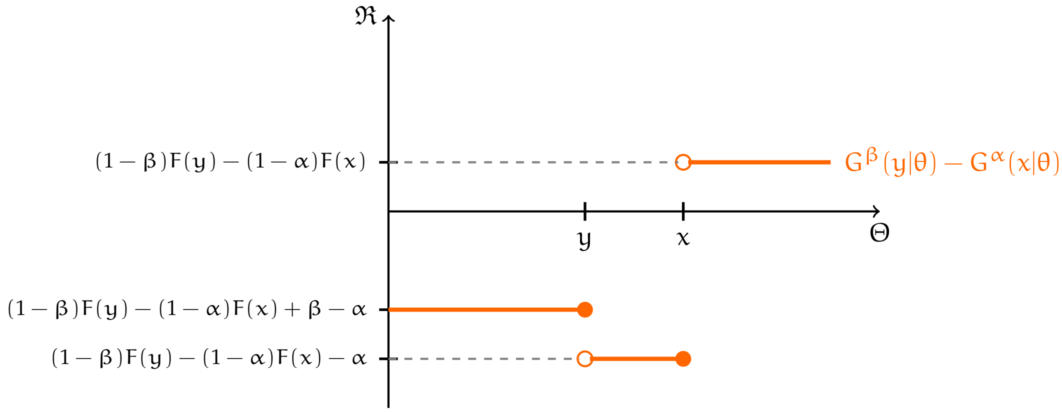


Figure 2.3: When $\alpha > \beta$, the function $G^{\beta}(y|\theta) - G^{\alpha}(x|\theta)$ satisfies the single-crossing property in θ so $\alpha \succ_L \beta$.

2.6 Dispersion

The main contribution of this article is to integrate the above three different ways of ranking signals—effectiveness, informativeness, and Lehmann precision—with another signal ordering based on *dispersion*. The central idea behind this link to the dispersion-based ordering can be best understood by the relationship between the precision of a signal and the dispersion of posterior beliefs. The next motivating example illustrates this relationship.

Example 2.15 (Precision and Dispersion)**.** *Suppose that the state of the world is dichotomy* $\Theta = \{\theta_L, \theta_H\}$ *with a prior belief* $q = Pr(\theta = \theta_L)$*. At the outset, a decision maker observes an outcome* $x \in \{x_L, x_H\}$ *from signal* α*. Its probability distribution is given by the following two-by-two symmetric matrix:*

$$g^{\alpha}(x|\theta) \quad \begin{array}{c|c|c|} & \theta_L & \theta_H \\ \hline x_L & \alpha & 1-\alpha \\ \hline x_H & 1-\alpha & \alpha \\ \hline \end{array}$$

Here $\alpha \in (1/2, 1)$ *measures the precision level of signal* α*. Let* $\pi_1 = Pr(\theta = \theta_L | x_L)$ *and* $\pi_2 = Pr(\theta = \theta_H | x_H)$ *denote the posterior beliefs of* $\{\theta = \theta_L\}$ *and* $\{\theta = \theta_H\}$ *conditioned on*

$x = x_L$ *and* $x = x_H$*, respectively. It is a routine task to check that*

$$\pi_1 = Pr(\theta = \theta_L | x_L) = \frac{\alpha q}{\alpha q + (1-\alpha)(1-q)} \text{ is increasing in } \alpha \text{ but}$$
$$1 - \pi_2 = Pr(\theta = \theta_L | x_H) = \frac{(1-\alpha) q}{(1-\alpha) q + \alpha(1-q)} \text{ is decreasing in } \alpha.$$

As the precision level α *increases, therefore, the posterior beliefs of the event* $\{\theta = \theta_L\}$ *will be spread out around the prior belief* q.

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 formalize this simple idea, we need introduce a few stochastic orderings to which the dispersion gives rise.

Let X and Y be two random variables such that $\mathbb{E}[\sigma(X)] \leqslant \mathbb{E}[\sigma(Y)]$ 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 write $X \leqslant_{icx} Y$. Roughly speaking, if $X \leqslant_{icx} Y$, then X is both lower and "less variable" than Y in some stochastic sense. The increasing concave order $\leqslant_{icv}$ is defined in a similar manner, and $X \leqslant_{icv} Y$ implies that X is both lower and more variable than Y.[23]

The following lemma follows from the two facts that (i) $(x - c) \vee 0 \triangleq \max\{x - c, 0\}$ is an increasing and convex function for all constants c and that (ii) every increasing convex function belongs to the closed convex cone generated by the family of functions $\{(x - c) \vee 0 | c \in \mathfrak{R}\}$ up to constants.[24]

---

[23] Note that a function $\sigma(x)$ is increasing and convex if and only if $-\sigma(-x)$ is increasing and concave. Hence $X \leqslant_{icx} Y$ if and only if $-X \geqslant_{icv} -Y$. For a comprehensive discussion of several stochastic orders based on variability, refer to Shaked and Shanthikumar (2007).

[24] Likewise, every increasing concave function can be uniformly approximated by $\{(x - c) \wedge 0 | c \in \mathfrak{R}\}$, where $(x - c) \wedge 0 \triangleq \min\{x - c, 0\}$. Hence we obtain the counterpart to Lemma 2.16 for the increasing concave order: $X \leqslant_{icv} Y$ if and only if

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

Lemma 2.16. *Let* X *and* Y *be two random variables with probability distributions* F_X *and* F_Y*, respectively. Then* $X \leqslant_{icx} Y$ *if and only if*

$$\int_{\kappa}^{\infty} [1 - F_X(x)]dx \leqslant \int_{\kappa}^{\infty} [1 - F_Y(y)]dy \quad \textit{for all } \kappa \in \mathfrak{R}. \tag{2.1}$$

The lemma associates with the familiar notion in economics; from the inequality (2.1) and Rothschild and Stiglitz (1970), it follows that $X \leqslant_{\text{icv}} Y$ is equivalent to domination of X by Y in the second-order stochastic dominance. See footnote 24.

In case the two random variables have the same mean, the two variability orders $\geqslant_{\text{icx}}$ and $\geqslant_{\text{icv}}$ reduce into $\geqslant_{\text{cx}}$ and $\geqslant_{\text{cv}}$, respectively, where $X \geqslant_{\text{cx}} Y$ implies $\mathbb{E}[\sigma(X)] \geqslant \mathbb{E}[\sigma(Y)]$ for every convex function σ. It is worth noting that whereas $\geqslant_{\text{cx}}$ and $\geqslant_{\text{cv}}$ are used to rank random variables according to their variability only, $\geqslant_{\text{icx}}$ and $\geqslant_{\text{icv}}$ are used to rank random variables according to both their mean and their variability.[25] Consequently, the latter are slightly more complete orderings than the former in that $X \geqslant_{\text{cx}} Y$ implies $X \geqslant_{\text{icx}} Y$ but not vice versa.

A description of the dispersion-based ordering requires one operator corresponding to the concept of "predictions about θ". Let Θ^* be the set of bounded, nondecreasing and measurable functions from Θ to $\mathfrak{R}$. Given a signal α and prior beliefs $\pi \in P(\Theta)$, we define the linear operator $J^{\alpha} : \Theta^* \times \mathcal{X} \to \mathfrak{R}$ as follows:

$$J^{\alpha}[\psi](x) \triangleq \int_{\Theta} \psi(\theta) dF^{\alpha}(\theta|x). \tag{2.2}$$

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

[25] To see why the convex order $\geqslant_{\text{cx}}$ requires the same mean between two random variables, note that both x and $-x$ are convex. Hence $X \leqslant_{\text{cx}} Y$ implies both $\mathbb{E}[X] \leqslant \mathbb{E}[Y]$ and $\mathbb{E}[-X] \leqslant \mathbb{E}[-Y]$, and thus $\mathbb{E}[X] = \mathbb{E}[Y]$. Also note that $X \leqslant_{\text{cx}} Y$ if and only if $X \geqslant_{\text{cv}} Y$.

the same way.

Now we are ready to state the new signal ordering:

Definition 2.17 (Dispersion)**.** *A signal α generates a more dispersed prediction about θ than another signal β (and we write $\alpha \succ_D \beta$) provided*

$$J^{\alpha}[\psi](X) \geqslant_{icx} J^{\beta}[\psi](Y) \quad \text{for every } \psi \in \Theta^* \text{ and } \pi \in P(\Theta). \tag{D}$$

The dispersion-based order furnishes us another way to rank 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 every 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 $\psi(\theta) = 1 - \mathbb{1}_{\{\theta \leqslant t\}}(\theta)$ is nondecreasing with θ for every constant $t \in \Theta$. In this case, J^{α} assumes

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

Hence $J^{\alpha}[\psi](X) \geqslant_{icx} J^{\beta}[\psi](Y)$ implies $1 - F^{\alpha}(t|X) \geqslant_{icx} 1 - F^{\beta}(t|Y)$, which in turn implies $F^{\alpha}(t|X) \geqslant_{cx} F^{\beta}(t|Y)$. 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 chapter.

3 THE EQUIVALENCE THEOREM

3.1 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 chapter.

Theorem 3.1 (Equivalence)**.** *For each class of payoff functions—$\mathcal{U}^{\mathrm{spm}}$, $\mathcal{U}^{\mathrm{sc}}$, or $\mathcal{U}^{\mathrm{ido}}$—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.*

In Sections 3.2 to 3.4, we prove the equivalence theorem for the two classes $\mathcal{U}^{\mathrm{spm}}$ and $\mathcal{U}^{\mathrm{sc}}$ only. The largest class $\mathcal{U}^{\mathrm{ido}}$ is relegated to Appendix A.2. We organize the proof in order (P) $\to$ (E) $\to$ (I) $\to$ (D) $\to$ (P), where we have already verified in Theorem 2.8 that (E) $\to$ (I) holds true irrespective of the class of payoff functions.

In Section 3.2, we first show that Lehmann precision $\alpha \succ_L \beta$ induces the preference-based ordering based on effectiveness. As we have discussed its analytical power in Section 2.3, Theorem 2.6 allows us to only look at the set of monotone strategies rather than the set of all permissible strategies so greatly simplifies our analysis. More precisely, to prove $\alpha \succ_E \beta$, it is sufficient to construct a payoff-improving strategy based on α for a given monotone strategy based on β. Lemma 3.4 provides the heart of this analysis, illustrating how the better strategy is to be constructed.

Section 3.3 contains the proof that if α is more informative than β for the SPM family, the former generates more dispersed predictions about θ. Since $\mathcal{U}^{\mathrm{spm}} \subset \mathcal{U}^{\mathrm{sc}}$, this statement is also true for the SC family. Section 3.4 completes the proof of the equivalence theorem by linking (D) to (P). Section 3.5 concludes this chapter with several important results of the theorem.

The statistical signal ordering based on dispersion plays a central role in three aspects. First, it acts as a medium that patches the two preference-based ordering through to the well-known statistical ordering based on Lehmann Precision to obtain the complete chain of implications. Second, the ordering is used to show that the equivalence theorem still holds even in $\mathcal{U}^{spm}$, the smallest family of payoff functions. In fact, Theorem 3.6 demonstrates that the dispersion-based ordering is induced by $\succ_I^{spm}$ which is the weakest partial preference-based ordering. Therefore, the chain of implications applies to $\mathcal{U}^{spm}$:

$$\alpha \succ_L \beta \rightarrow \alpha \succ_E^{spm} \beta \rightarrow \alpha \succ_I^{spm} \beta \rightarrow \alpha \succ_D \beta \rightarrow \alpha \succ_L \beta.$$

Finally, the dispersion-based ordering is used to establish the Dispersion Theorem, one corollary of the equivalence theorem, which associates the two statistical orderings and exactly characterizes more precise signals and higher dispersion. The theorem justifies another signal orderings used in the previous literature and will be exploited to analyze the effect of precise information in several game-theoretic models in the next chapter.

3.2 Precision leads to Effectiveness

We begin with the following theorem:[1]

Theorem 3.2. *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 A.2, we show that the IDO family also possesses the same property.

[1] 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 $\mathcal{U}^{KRm}$ in Lehmann (1988).

Lemma 3.3. *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* $u(a^\star, \theta) \geqslant u(d(\theta), \theta)$ *for all* $\theta \in \Theta$*.*

Proof of Lemma 3.3 : Let $A_d = \{d(\theta) \in A | \theta \in \Theta\}$ denote the image of the given decision d. Since $u(a, \theta)$ satisfies the SC property 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.[2] Let $\theta^\star \in \Theta$ indicate the state at which they intersect. Then $d^\star(\theta) \geqslant d(\theta)$ for $\theta \geqslant \theta^\star$ and $d^\star(\theta) \leqslant d(\theta)$ otherwise. Set $a^\star = d^\star(\theta^\star)$, i.e., the optimal action at the intersection chosen by $d^\star$. Then $u(a^\star, \theta^\star) \geqslant u(d(\theta), \theta^\star)$ for all θ. For all $\theta > \theta^\star$, $a^\star \geqslant d(\theta)$ since the decision d is decreasing. Hence it follows from the SC property that

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

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 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.[3]

Lemma 3.4 (Improvement Principle)**.** *Suppose that* $u \in \mathcal{U}^{sc}$ *and* $\alpha \succ_L \beta$*. Then for*

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

[3]Refer to Appendix A.1 for the key idea of this principle.

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*

$$u(d^\alpha \circ T_y(\theta), \theta) \geqslant u(d^\beta(y), \theta) \quad \forall\, y \in \mathcal{Y},\ \theta \in \Theta, \tag{3.1}$$

where the transformation T_y *is from (P) in Definition 2.10.*

Proof of Lemma 3.4 : 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 (3.1) is equivalent to

$$u(d^\alpha(x), \theta) \geqslant u(d^\beta \circ \tau_x(\theta), \theta) \qquad x \in \mathcal{X},\ \theta \in \Theta. \tag{3.2}$$

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 $u(a^\star, \theta) \geqslant u(d^\beta \circ \tau_x(\theta), \theta)$ for all states θ. For each x, $d^\alpha(x) = a^\star$ will lead to the desired inequality (3.2).

It remains to show that the decision rule d^α we just constructed is increasing. Note that $d^\beta \circ \tau_{x'}(\theta) \geqslant d^\beta \circ \tau_x(\theta)$ 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 3.3 is also increasing (See Figure 3.1). It proves the lemma. □

Proof of Theorem 3.2 : Note that for every monotone decision rule d^β permissible

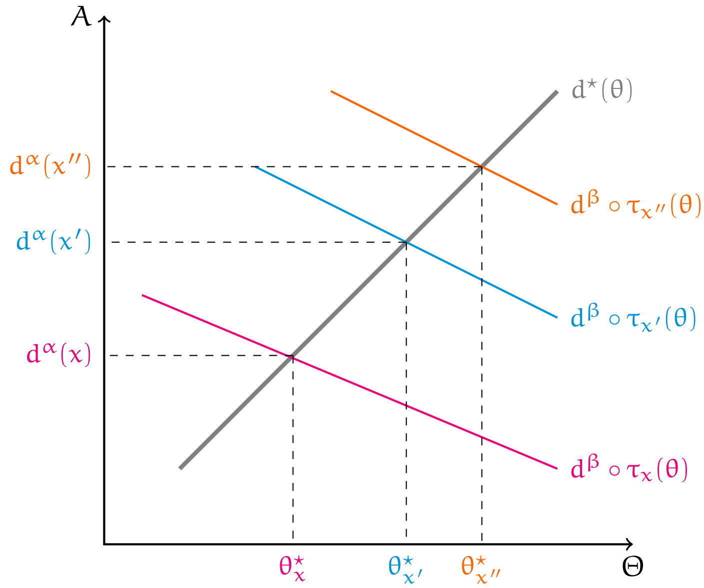


Figure 3.1: Improvement Principle - While the decision rule $d^\beta \circ \tau_x(\theta)$ is decreasing in θ for each x, it is increasing in x for each θ. This property leads to the increasing sequence of intersections $\{d^\alpha(x)\}_{x\in\mathcal{X}}$ with $d^\star$.

under β,

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

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 2.6 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 3.2. 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 3.4.

Along with Theorem 2.8, one immediate consequence of the preceding theorem is that the Lehmann precision induces the signal ordering based on the Bayesian values as well.

Corollary 3.5. *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 2.9 and Lemma 3.4 to prove the same result in a direct way, not through the effectiveness. To this end, we first infer from Theorem 2.9 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 3.4, we can assure that there exists a decision rule d^α for which

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

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 under β, the decision maker with $u \in \mathcal{U}^{sc}$ is able to find a decision rule based on a more Lehmann-precise signal α that makes her better off regardless of whether she is a statistician or Bayesian. Since $\mathcal{U}^{spm}$ is included in $\mathcal{U}^{sc}$, the above results also hold for the class $\mathcal{U}^{spm}$.

3.3 Informativeness leads to Dispersion

Now we link the criterion of informativeness to the dispersion-based ordering, demonstrating that if α is more informative than β with respect to $\mathcal{U}^{spm}$, α generates more dispersed predictions about θ than β. Since $\mathcal{U}^{spm} \subset \mathcal{U}^{sc}$ and thus $\alpha \succ_I^{sc} \beta$ implies $\alpha \succ_I^{spm} \beta$, $\alpha \succ_D \beta$ is induced by $\alpha \succ_I^{sc} \beta$ as well.

Theorem 3.6. *If* $\alpha \succ_I^{spm} \beta$, *then* $\alpha \succ_D \beta$.

Proof of Theorem 3.6 : Consider an action space $A = \{0, 1\}$. For a constant $\kappa \in \mathfrak{R}$ and a nondecreasing function $\psi \in \Theta^*$ defined on Θ, we define a payoff function $u_{\psi,\kappa} : A \times \Theta \to \mathfrak{R}$ 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 $\{u_{\psi,\kappa}(\cdot, \theta)\}_{\kappa\in\mathfrak{R},\, \psi\in\Theta^*,\, \theta\in\Theta}$ 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 $(\alpha, u_{\psi,\kappa})$:

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

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[4]

$$\mathcal{V}^{\pi}(\alpha, u_{\psi,\kappa}) = \int_{\kappa}^{\infty} [1 - H^{\alpha}(\xi)]\, d\xi, \tag{3.3}$$

where $H^{\alpha}(\xi) \triangleq \Pr\left(J^{\alpha}[\psi](X) \leqslant \xi\right)$ 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 $\mathcal{V}^{\pi}(\alpha, u_{\psi,\kappa}) \geqslant \mathcal{V}^{\pi}(\beta, u_{\psi,\kappa})$ for all κ, ψ, and all prior beliefs $\pi \in P(\Theta)$. By using the representation result (3.3), therefore, the informativeness leads us to

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

Lastly, it is immediate from Lemma 2.16 that $J^{\eta}[\psi] \geqslant_{icx} J^{\theta}[\psi]$ for all ψ and π. Therefore, the informativeness induces the new signal ordering $\alpha \succ_D \beta$ based on the dispersion. $\square$

3.4 Dispersion leads to Precision

The next result completes the proof of the main theorem.

Theorem 3.7. *If $\alpha \succ_D \beta$, then $\alpha \succ_L \beta$.*

Proof of Theorem 3.7 : For a pair of outcomes $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, suppose that $\Delta^{x,y}(\theta_0) = G^{\beta}(y|\theta_0) - G^{\alpha}(x|\theta_0) \geqslant 0$ for some θ_0. We want to show that $\Delta^{x,y}(\theta) \geqslant 0$ for every $\theta \geqslant \theta_0$, which is equivalent to $\alpha \succ_L \beta$ by Lemma 2.12. 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.

[4] The representation theorem (Folland (1999)) states that for a σ-finite Borel measure ν on $\mathfrak{R}_+ = [0, \infty)$, $\phi(t) \triangleq \nu[0, t)$, 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 $\pi \in \mathsf{P}(\Theta)$ 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 $\theta \leqslant \widehat{\theta}$ and the value $c \in \mathfrak{R}$ 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 $\kappa \leqslant c$ is immediate.

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

$$\mathbb{E}_X\left[(J^\alpha[\psi](X) - \kappa) \vee 0\right] \geqslant \mathbb{E}_Y[(J^\beta[\psi](Y) - \kappa) \vee 0], \tag{3.4}$$

as $(x - \kappa) \vee 0$ is increasing and convex. With the prior belief π, however, we can rewrite the left-hand side of (3.4) as

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

Note that the event of the indicator function above $\{\xi \geqslant x\}$ is the same as the event $\{J^\alpha[\psi](\xi) \geqslant \kappa\}$. Thus $\mathbb{1}_{\{\xi \geqslant x\}}(\xi)$ 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

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

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 (3.4) leads to

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

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

3.5 Discussion

This section will be devoted to the discussion of the main result we established in the previous sections. 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 3.8. *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 frequently used to model economic problems. Focusing on the class of Karlin-Rubin monotone payoff functions $\mathcal{U}^{KRm}$ (See Appendix A.1), he developed a statistical signal ordering based on Precision to establish 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 3.8 extends his result to $\mathcal{U}^{spm}$ and $\mathcal{U}^{sc}$, respectively.

Example 3.9 (Lehmann (1988))**.** *Consider a quintessential one-tail hypothesis testing problem—*$H_0 : \theta \leqslant \theta^*$ *and* $H_1 : \theta > \theta^*$*—for a state* $\theta^* \in \Theta$*. The associated primitive payoff function with this test can be written*

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

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* $\mathcal{U}^{spm}$*.*

Due to Theorem 2.6, 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*

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

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

$$G^{\alpha}(x^*|\theta) \geqslant G^{\beta}(y|\theta) \;\; \forall\, \theta \leqslant \theta^* \;\; \textit{and} \;\; G^{\alpha}(x^*|\theta) \leqslant G^{\beta}(y|\theta) \;\; \forall\, \theta > \theta^*, \tag{3.5}$$

which is equivalent to $\alpha \succ_L \beta$ *by Lemma 2.12. Therefore, Lehmann precision is necessary for satisfying the criterion of effectiveness with respect to* $\mathcal{U}^{spm}$*.* □

It follows from Lemma 2.5 that α generates more statistical values $V(\alpha, u)$ than β if and only if α is more effective than β. Therefore, Corollary 3.8 tells us that $V(\alpha, u) \geqslant V(\beta, u)$ 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 by narrowing down the scope of payoff functions we are able to rank signals with a less restrictive condition.

The example above imparts the main insight underlying this result. To formalize it, 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 $\{a_0, a_1\}$. For α to be more effective than β for this payoff function with ${}^{u}\Delta(\theta)$ nondecreasing, it must hold that

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

where $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ indicate the cutoff points at which the minimax strategies d^{α} and d^{β} jump from a_0 to a_1, respectively.[5] When ${}^{u}\Delta(\theta) = u(a_1, \theta) - u(a_0, \theta)$ is

[5] Recall that there exists a monotone minimax strategy due to Theorem 2.6.

nondecreasing, therefore, the inequality above boils down to the condition (3.5), namely, $\alpha \succ_L \beta$.

Relaxing the assumption of monotonicity on ${}^u\Delta(\theta)$ and imposing a weaker condition that it obeys the single-crossing property in θ make no difference. Since the SC property of ${}^u\Delta(\theta)$ solely cannot reverse the above inequality, $\alpha \succ_E^{sc} \beta$ leads to the same condition as (3.5). Hence Lehmann precision is necessary for effectiveness in the SC family.

Corollary 3.10. $\alpha \succ_E^{sc} \beta$ *if and only if* $\alpha \succ_E^{spm} \beta$.

Relating (P) Lehmann precision to (I) the criterion of informativeness from the Bayesian perspective, we can obtain the second implication of the main theorem. It answers the question of when one signal is more valuable than another in Bayesian decision theory.

Corollary 3.11. *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 the Bayesian value as discussed in Chapter 2, Lehmann precision gives a uniform signal ordering necessary and sufficient for representing a Bayesian decision maker's preference over the signal space.

The equivalence between $\alpha \succ_L \beta$ and informativeness within $\mathcal{U}^{spm}$ is worthy of remark. Athey and Levin (2001) explored the value of information in a Bayesian decision problem with a fixed prior belief $\pi \in P(\Theta)$ and $\mathcal{U}^{spm}$. They developed a statistical 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 $\mathcal{U}^{spm}$.[6] Formally, given a prior belief π, $\mathcal{V}^\pi(\alpha, u) \geqslant \mathcal{V}^\pi(\beta, u)$ for every $u \in \mathcal{U}^{spm}$ if and only if $\alpha \succ^\pi_{\text{MIO-ND}} \beta$, defined by

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

[6](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.

(MIO-ND)

In words, the posterior distribution of the event $\{\theta \leqslant t\}$ 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.[7] 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 $\mathcal{V}^{\pi}(\alpha, u) \geqslant \mathcal{V}^{\pi}(\beta, u)$ for every $u \in \mathcal{U}^{spm}$ *and* for every π. Hence we infer from Corollary 3.11 that (MIO-ND) is implied by Lehmann precision.

Proposition 3.1. *For every* $\pi \in P(\Theta)$*,* $\alpha \succ_L \beta$ *implies* $\alpha \succ^{\pi}_{MIO-ND} \beta$.

Proof of Proposition 3.1 : 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 $x^{\alpha}_c = \sup\{x \in \mathcal{X} | M^{\alpha}(x) \leqslant c\}$. Then we use Fubini's Theorem and Bayes' rule to rewrite the posterior distribution in (MIO-ND) as

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

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

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

We infer from Lemma 2.12 that the expression in the bracket satisfies the single-crossing property in θ and from the definition of x^{α}_c and x^{β}_c that

$$\int_{\Theta} \left[G^{\beta}(y^{\beta}_c | \theta) - G^{\alpha}(x^{\alpha}_c | \theta) \right] d\pi(\theta) = M^{\beta}(y^{\beta}_c) - M^{\alpha}(x^{\alpha}_c) = 0.$$

[7] 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).

Additionally, the indicator function $\mathbb{1}_{\{\theta \leqslant t\}}(\theta)$ is nonincreasing in θ. Hence it follows from the *folk single-crossing lemma* that $F^{\alpha} - F^{\beta}$ is nonpositive.[8] $\square$

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 3.11 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 3.12. $\alpha \succ^{\pi}_{\text{MIO-ND}} \beta$ *for every* $\pi \in P(\Theta)$ *if and only if* $\alpha \succ_L \beta$.

Just as we discussed above, Corollary 3.11 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 2.9, $\alpha \succ^{spm}_{I} \beta$ can be stated that for every $y \in \mathcal{Y}$ there exists a $x \in \mathcal{X}$ such that

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

so we arrive at the same condition (3.5), which in turn implies $\alpha \succ^{sc}_{I} \beta$. Put differently, if the above inequality does not hold for some ${}^{u}\widehat{\Delta}(\theta)$ satisfying the SC

[8] The folk single-crossing lemma states that if a function $h : \Theta \to \mathfrak{R}$ satisfies the single-crossing property in θ and

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

for every $\phi_1 : \Theta \to \mathfrak{R}$ nondecreasing and $\phi_2 : \Theta \to \mathfrak{R}$ nonincreasing.

property and for some prior $\widehat{\pi}$, there exists a ${}^{u}\Delta(\theta)$ nondecreasing and another prior π for which the inequality would be reversed.

Example 3.13 (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* $\mathcal{R} = \{r : \Theta \to \mathfrak{R} |$*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* $\mathcal{R}^{sc} = \{r(\theta) | r(\theta) - s$ *obeys the single-crossing property in* $\theta\}$ *and* $\mathcal{U}^{spm}$ *by* $\mathcal{R}^{spm} = \{r(\theta) | \partial^2 u_r / \partial a \partial \theta \geqslant 0\}$ *due to Topkis (1978).*

Suppose that the investor receives a signal α *or* β *by hiring an investment consultant with expertise. By Corollary 3.11, 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 derived from the main theorem is the heart of a characterization of Lehmann precision, which we call "Dispersion Theorem" hereafter.

Corollary 3.14 (Dispersion Theorem)**.** $\alpha \succ_L \beta$ *if and only if* $\alpha \succ_D \beta$, *i.e., for every increasing and convex function* σ,

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

Corollary 3.14 provides a characterization of the statistical ordering $\succ_L$ in terms of another statistical ordering $\succ_D$: A signal α is more statistically precise than another β if and only if the former generates a more dispersed prediction about θ than the latter. As we have discussed in Section 2.6, the decision maker with the more precise signal α tends to put more weights on the upper or lower tails of the distribution of θ according to outcomes than the one with β, so her posterior beliefs about θ become more dispersed. By the same token, the posterior 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 monotonicity of ψ will preserve the 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 his 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 the outcome for each θ. In contrast, our characterization does not utilize any additional structures on the signal's primitive distributions so is applicable to more distributions than the Lehmann's result.

Furthermore, Dispersion Theorem furnishes us with a link between Lehmann precision and several signal orderings, based on the dispersion of a certain variable, used in previous literature. This link can be found by making a clever choice of the nondecreasing function $\psi : \Theta \to \mathfrak{R}$. For instance, by setting $\psi(\theta) = \theta$, the theorem along with the law of iterated expectations—$\mathbb{E}_X[J^\alpha[\psi](X)] = \mathbb{E}_X[\mathbb{E}_\theta[\theta|X]] = \mathbb{E}_Y[\mathbb{E}_\theta[\theta|Y]]$—gives the second-order stochastic dominance of conditional expectations:[9]

Corollary 3.15 (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,* $\mathbb{E}[\theta|X] \geqslant_{cx} \mathbb{E}[\theta|Y]$.

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.[10] We use the dispersion theorem to revisit their questions in Section 4.2 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

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

[10]The signal ordering is named "*integral precision*" in their paper. That is, α is more integral precise than β if $\mathbb{E}[\theta|X] \geqslant_{cx} \mathbb{E}[\theta|Y]$. Another application of this ordering to the optimal contract problem can be found in Szalay (2009).

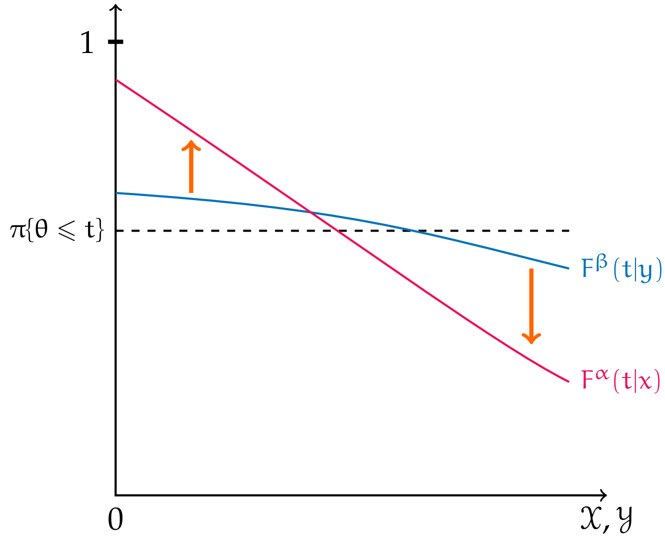


Figure 3.2: Mean-Preserving Spread of Posterior Beliefs - For two signals α and β with $\alpha \succ_L \beta$, the two curves depict the posterior distribution of the event $\{\theta \leqslant t\}$ as a function of outcomes. The posterior distributions of β will be more clustered toward the mean $\mathbb{E}_X[F^{\beta}(t|X)] = \pi\{\theta \leqslant t\}$ than those of α, hence $F^{\beta}(t|X)$ does not make a big difference from the prior.

sense, it generates more variable posterior beliefs on Θ as depicted in Figure 3.2-(a). Put differently, when the decision maker observes an outcome from a less precise signal, her posterior beliefs of the event $\{\theta \leqslant t\}$ would not make a big difference with her prior beliefs $\pi\{\theta \leqslant t\}$.

Corollary 3.16 (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 3.16 : Define $\psi(\theta) = \mathbb{1}_{\{\theta \geqslant t\}}(\theta)$ 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 $\{\theta \leqslant t\}$. In this case, the dispersion theorem presents

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

See footnote 23. Since $\mathbb{E}_X[F^\alpha(t|X)] = \mathbb{E}_Y[F^\beta(t|Y)] = \pi\{\theta \leqslant t\}$ by Bayes' rule, the above increasing concave order reduces to $F^\alpha(t|X) \leqslant_{cv} F^\beta(t|Y)$. □

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 their signal ordering is induced by Lehmann Precision but not vice versa.

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

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

Refer to Figure 3.2-(b). Since $\alpha \succ_L \beta$ implies $F^\beta(\theta|\cdot) \leqslant_{cx} F^\alpha(\theta|\cdot)$ only, the rotation

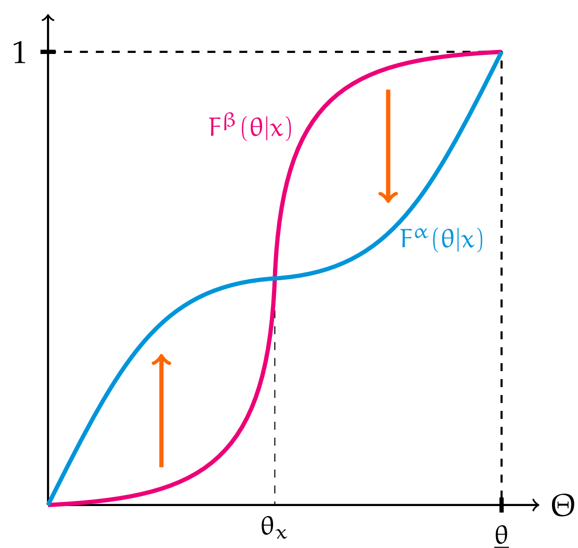


Figure 3.3: Rotation Orders - The two curves display the posterior distributions of α and β as a function of θ, with the outcome x held fixed.

order is stronger than Lehmann precision.

The third application of the dispersion theorem concerns the distribution of likelihood ratio. To ascertain the impact on the likelihood ratio, we first note that if the density function $g^\alpha(x|\theta)$ satisfies the MLRP, so does the posterior density function $f^\alpha(\theta|x)$. To see this, for every $x' > x$, Bayes' rule gives an alternative expression of the likelihood ratio function

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

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

Corollary 3.17 (Likelihood Ratio)**.** *If $\alpha \succ_L \beta$, the distribution of the likelihood ratio with α*

$$L^\alpha(\kappa) \triangleq \Pr\left(\frac{f^\alpha_x}{f^\alpha}(\theta|x) \leqslant \kappa\right)$$

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

In the principal-agent model with hidden actions, Kim (1995) has shown that, in order for the principal to control the agent's hidden action more efficiently, she has to design an incentive scheme based on the contractible variable generating more dispersed distributions of the likelihood ratios, and he provided a partial ordering on the set of contractible variables. We will examine the exact connection between Lehmann precision and this ordering in Section 4.3.

Lastly, the dispersion theorem provides an alternative way to prove that $\alpha \succ_I^{spm} \beta$ is induced by $\alpha \succ_L \beta$. Although this implication has been shown in Corollary 3.5, the theorem enables us to present a simple and direct proof without relying on effectiveness and the Improvement Principle. Furthermore, we will make use of some arguments in this proof in Section 4.4 to analyze the effect of precise information in delegation.

Corollary 3.18. *If* $\alpha \succ_L \beta$*, then* $\alpha \succ_I^{spm} \beta$.

Proof of Corollary 3.18 : See Appendix A.3. □

4 APPLICATIONS TO GAME-THEORETIC MODELS

4.1 Introduction

For decision making under uncertainty, precise information concerning the underlying state is valuable in the sense that it makes better decisions possible as we have discussed in the previous chapter. The Improvement Principle (Lemma 3.4) illustrated how such a decision rule is constructed on the basis of more precise information, and the method generally applies to a fairly large class of economic problems. From this starting point, we have established the equivalence between $\alpha \succ_L \beta$ and $\alpha \succ_I \beta$ in a Bayesian decision-making model.

When strategic interaction between players is embedded into the model, however, this equivalence theorem is invalidated as it has been first pointed out by Hirshleifer (1971). A decision maker is still able to make better decision from precise information, but the change in her strategy may alter her opponent's decision problem in games. This strategic effect can either help or harm the decision maker herself.

In this chapter we explore how precise information impinges on the players' decision problems in several economic environments. The main analytic tool for this exploration is the dispersion theorem, coupling the precision-based ordering with the dispersion-based ordering. Although other equivalence results established in the previous chapter break down, the equivalence between these two statistical orderings is valid even in the presence of strategic interactions.

The first application is to auctions. In Section 4.2, we study an allocation problem where the auctioneer is able to control the information structure from which potential buyers learn the true value of product. We investigate how the auctioneer's information policy influences efficiency, expected revenues, and buyers' payoffs.

The second application is to bilateral contracts. When the principal cannot perfectly monitor the agent's action, moral hazard problems arise so the agent would choose a different action from what the principal wishes to implement. As is well-documented, an mutually-observed signal serves as a proxy when writing an incentive contract. In Section 4.3, we study a signal ordering in this standard contract problem and attempt to answer which signal is more cost-effective in implementing the desired action.

The third is to delegation. We consider a project-choice problem where two parties—a principal and a subordinate—in an organization choose a project under a state of uncertainty and while the principal has a formal authority to make a final decision, the subordinate has the expertise to help the principal choose a better project. In Section 4.4, we provide a condition for the value of delegating the decision right to the subordinate to be monotone with the level of expertise.

4.2 Auctions

The aim of this section is to examine the effect of more precise information in auctions. In many cases, the auctioneer has superior information about the product for sale and is able to supply more information on the product for bidders through marketing or advertising activities. Such information provision by the auctioneer enables each bidder to learn better his true preference for the product and to enhance taste perception. We explore how the equilibrium is affected when the auctioneer publicly provides more precise information.[1] The question is closely related to Ganuza and Penalva (2010). We use the dispersion theorem to extend their results to the asymmetric auction. Before proceeding, it is worthy of note that, unlike the *covert* information acquisition, precise information is not necessarily valuable to each participant in case the information gathering process is observable, because the change in distributions can alter the bidding strategies of others. This second strategic effect may hurt himself in the end.[2]

[1]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 within the same category, see Bergemann and Pesendorfer (2007), Esö and Szentes (2007), and Ganuza and Penalva (2010).

[2]In Appendix A.4, 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.

Consider an allocation problem in which a single indivisible object is auctioned off to N risk-neutral bidders through the second-price auction.[3] Each bidder is indexed by $i \in \{1, 2, \cdots, N\}$. The *true value* of the object to bidder i is given by $v_i(s, \theta_i)$. Here $\theta_i \in \Theta_i$ represents the informational variable reflecting his idiosyncratic preferences for the object, and $s \in S$ represents the additional common-value component influencing the value of the object to all bidders. For example, in a takeover bid for a target company (Bulow et al. (1999)), the variable s may represent common-value information on the target—such as its accounting and earnings estimates—determining the target's financial value to all bidders, but the variable θ_i may correspond to private-value information about its other attributes affecting the post-takeover value of the target to bidder i only.

We assume that all participants are perfectly (or identically) informed about s while each bidder i is imperfectly informed about θ_i and thus the true value $v_i(s, \theta_i)$ is unknown till he wins the auction. $v_i(s, \theta_i)$ is nondecreasing in θ_i. The prior distribution $\pi(\theta_1, \cdots, \theta_N)$ over $\Theta_1 \times \cdots \times \Theta_N$ is common knowledge among players, and it satisfies *independence* across bidders.[4] That is,

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

Prior to the bidding stage, each bidder observes a realization $x_i \in [0, 1]$ of random variable X_i which is affiliated with θ_i. The signal's cumulative distribution is $G_i^{\alpha}(\cdot|\theta_i)$ with support $[0, 1]$ for all $\theta_i \in \Theta_i$, where α measures the precision level of X_i.[5] The posterior distribution about θ_i conditional on $X_i = x_i$ is $F_i^{\alpha}(\cdot|x_i)$. The winning bidder's primitive payoff function is quasilinear: $u_i(v_i, t_i) = v_i(s, \theta_i) - t_i$,

[3]When the information acquisition activity is overt, some results depend on the number of bidders. See Proposition 4.2. When the activity is covert, on the other hand, the results are independent of the number of bidders so one can focus on $N = 2$ case—like Persico (2000)—without loss of generality.

[4]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.

[5]The assumption that the support is fixed by $[0, 1]$ regardless of θ_i for all signals is essential for Proposition 4.2 and 4.3, but is not needed for Proposition 4.1.

where t_i is a money transfer to the auctioneer.

We assume that the auctioneer is able to choose the precision level $\alpha \in E$ of signals $(X_1, \cdots, X_N)$ by releasing relevant information about the object and the choice of α is observed by every participant. Here E, the set of possible α, is Lehmann-ordered so that $\alpha > \beta$ for $\alpha, \beta \in E$ implies $\alpha \succ_L \beta$. We call this activity the auctioneer's information policy in this section.[6] Consequently, the information structure $\{G_i^\alpha(x_i|\theta_i), \pi_i, s\}_{i \in N}$ of the environment is common knowledge. The 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 object for sale and thus to receive more accurate information about θ_i.[7]

Our first result addresses the effect of the information policy on efficiency. To state the result, we define the bidder i's value estimate under the information policy α by

$$v_i^\alpha(x_i) = \int_{\Theta_i} v_i(s, \theta_i) dF_i^\alpha(\theta_i|x_i).$$

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

$$\mathcal{E}(\alpha) \triangleq \mathbb{E}\left[\max\{v_1^\alpha, \cdots, v_N^\alpha\}\right] \tag{4.1}$$

I wish to remark that the following result also holds in every symmetric auction with independent private value.[8]

[6]Our formulation also covers the case in which the auctioneer reveals information relevant with θ_i only. In this case the auctioneer's information policy is given by N-tuple vector $\alpha = (\alpha_1, \cdots, \alpha_N)$ and $\alpha \succ_L \beta$ if $\alpha_i \succ_L \beta_i$ for some i.

[7]As is pointed out by Johnson and Myatt (2006), most advertising activities inherit the two characteristics: (i) *promotional hype*, which advertises the product's existence, price, availability, and any other objective qualities, and thus increases the bidder's value estimate in the First-Order Stochastic Dominance; and (ii) *provision of real information*, which helps the bidders to learn their subjective preferences. In this section we concentrate on the second role of information disclosures.

[8]The symmetric auction implies that every bidder is symmetric in the sense that the common

Proposition 4.1 (Efficiency)**.** *More Lehmann-precise information generates more ex ante surplus.*

The proof is immediate from the next lemma:

Lemma 4.1 (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* $X_i \geqslant_{icx} Y_i$ *for each* i*. Then we have*

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

for every increasing and componentwise convex function $\sigma : \mathfrak{R}^N \to \mathfrak{R}$*.*

Proof of Lemma 4.1 : 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 also holds for $N = k$. To this end, let $\psi : \mathfrak{R} \to \mathfrak{R}$ be increasing and convex. Note that

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

The first equality follows from the independence assumption and the inequality from the induction assumption. Integrating over X_1 gives $\sigma(X_1, X_2, \cdots, X_k) \geqslant_{icx} \sigma(X_1, Y_2, \cdots, Y_k)$. 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

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

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 this 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 (4.1).

The proof is complete. □

Proof of Proposition 4.1 : Suppose $\alpha \succ_L \beta$. For each bidder i, since $v_i(s, \theta_i)$ is assumed to be increasing in θ_i, $v_i^\alpha \geqslant_{icx} v_i^\beta$ follows from Corollarly 3.14. 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 by the preceding lemma, we obtain

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

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

Proposition 4.1 states that the total surplus increases with precision of the information released by the auctioneer. The basic intuition behind this result, captured by the dispersion theorem, is that greater precision of the information disclosure intensifies variability of the value estimates v_i^α for each bidder and thus the variability of the largest value estimates. As its consequence, the precise information policy increases the probability of assigning the object to the highest true value so improves efficiency.

Unlike the total surplus, on the other hand, the expected revenue may decrease or increase—the relationship between the precision level of information disclosures and the expected revenue is not monotonic. To go into the details, observe that the expected revenue accruing from the second-price auction is determined by the second-highest expected value. In contrast with the highest expected value, the second-order statistic of $\{v_1^\alpha, \cdots, v_N^\alpha\}$ is not arranged in a variability order since the second max operator is neither increasing nor convex.

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, with other strategic factors held fixed, since his value estimate v_i^α is more dispersed as α increases.[9] Since the expected revenue is determined by

[9] The idea of relating high dispersion to more information rents can be traced back to Lewis and Sappington (1994). In Appendix A.4 it is formally shown that greater precision of the signal

the difference between efficiency and the aggregate of information rents but they would increase with different rates, it is necessary for a comparative statics result to ascertain which one increases with the larger rate.

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 seller is able to render the value estimates of each bidder to be closer to those of his opponents. As a result, the auction implemented with more precise information will generate more intense Bertrand price competition between the bidders, which results in higher expected revenue.[10] However, our model centers on the independent information structures across bidders' true values. In this case, more information causes the value estimates to be spread out and thus enhances privacy of each bidder's information. Therefore, it would exacerbate the incentive compatibility issue so the expected revenue can go either way.

The following result asserts that the information rent would increase with the larger rate than the total surplus so the expected revenue, denoted $R(\alpha)$, decreases with α in case of two bidders.

Proposition 4.2 (Revenue I)**.** *In case* $N = 2$*, more Lehmann-precise information is detrimental to the expected revenue.*

Proof of Proposition 4.2 : Since $v_i^\alpha \geqslant_{icx} v_i^\beta$ and their ex ante means are the same by the law of iterated expectation, we have $v_i^\alpha \geqslant_{cx} v_i^\beta$ for each $i = 1, 2$, equivalently, $v_i^\alpha \leqslant_{cv} v_i^\beta$. It implies $v_i^\alpha \leqslant_{icv} v_i^\beta$. Therefore, by the counterpart of Lemma 4.1 to the

increases the bidder's expected payoff when the information acquisition is unobserved.

[10] 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 issue for truthful revelation.

ICV order, we have

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

Taking expectation of both sides yields $R(\alpha) \leqslant R(\beta)$. □

The preceding proof is based on a simple observation that the second-order statistic is determined by $\min\{v_1^\alpha, v_2^\alpha\}$ under which the increasing concave order is preserved. Thus a similar argument with Proposition 4.1 can be used to establish that $R(\alpha)$ decreases with α.

The discussion above tells us that there is an incentive for the auctioneer to solicit bids from more participants, since a large number of bidders helps to reduce the bidder's information rents by intensifying competition. In light of this fact, Ganuza and Penalva (2010) found that $R(\alpha, N)$ increases with α when the number of bidders N is sufficiently large.[11] In contrast with this result, the next example shows that for any fixed number of bidders, the expected revenue can fall as the auctioneer provides more precise information.[12]

Example 4.2. *Let* $\Theta_i = \{\theta_1, \theta_2, \theta_3\}$ *for every bidder* i *and let* $v(\theta_1) = 0$, $v(\theta_2) = 0.9$, *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* $\epsilon > 0$. *Under the auctioneer's information policy* β, *each bidder has an information partition* $\{\{\theta_1\}, \{\theta_2, \theta_3\}\}$. *The partition corresponding to the information policy* α *is* $\{\{\theta_1\}, \{\theta_2\}, \{\theta_3\}\}$. *Note that each information partition leads to the following probability distribution:*

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

[11] In the same spirit, Rezende (2013) has shown that the second max operator $T^{2:N}(v_1^\alpha, \cdots, v_N^\alpha) =$ second $\max\{v_1^\alpha, \cdots, v_N^\alpha\}$ is *asymptotically componentwise convex* in the sense that the sequence of measurable sets

$$\left\{(v, w) \in \mathfrak{R}_+^2 \,|\, \exists\, \lambda \in (0,1) \text{ such that } T^{2:N}(\lambda v + (1-\lambda)w, v_{-i}^\alpha) > \lambda T^{2:N}(v, v_{-i}^\alpha) + (1-\lambda)T^{2:N}(w, v_{-i}^\alpha)\right\}$$

converges to 0 in measure.

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

In a nutshell, 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 *under* β*. Thus* $\alpha \succ_L \beta$*.*

Note that we can write the seller's expected profit generated by β *as*

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

where $p(n)$ *indicates 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 that* n *out of* N *bidders receive the imperfect signals. Since the bidders cannot distinguish* θ_2 *with* θ_3 *under* β*, they will place a bid as much as* $\mathbb{E}[v(\theta)|y_2] = \mathbb{E}[v(\theta)|y_3] = 1$ *in the second-price auction. Hence* $\mathbb{E}[R|n, \beta] = 0$ *for* $n = 1$ *but* $\mathbb{E}[R|n, \beta] = 1$ *for all* $n \geqslant 2$*.*

Similarly, the expected revenue accruing from the information policy α *can be written*

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

where $p(n)$ *represents the probability of the event that* n *of* N *bidders receive* x_2 *or* x_3 *and* $\mathbb{E}[R|n, \alpha]$ *is the conditional expected revenue given the event and* α*.*

Comparing the conditional expected revenues from α *and* β *for each* $n = 2, 3, \cdots$*, we have* $\mathbb{E}[R|2, \alpha] = 0.95 < \mathbb{E}[R|2, \beta] = 1$*,* $\mathbb{E}[R|3, \alpha] = \mathbb{E}[R|3, \beta] = 1$*, and* $\mathbb{E}[R|n, \alpha] > \mathbb{E}[R|n, \beta]$ *for* $n \geqslant 4$*. Therefore, for arbitrary number of bidders* N*, there exists a sufficiently small* $\bar{\epsilon}$ *such that* $\epsilon < \bar{\epsilon}$ *leads to* $R(\beta) > R(\alpha)$*.*[13] $\square$

Now, we explore the effect on the bidder's expected payoff, denoted $U_i(\alpha)$, in case $N \geqslant 3$.[14] While it is tempting to say that $U_i(\alpha)$ is monotone in α, Lehmann precision itself is not enough due to the strategic (or equilibrium) effect: Once α

[13]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.

[14]As an immediate consequence of Proposition 4.2, the information rents would increase when $N = 2$.

changes, bidder i adjusts his bids, which changes his opponents' problems, and all effects need to be traced through to equilibrium.

To account for this effect more precisely, 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 $H_i^\alpha(y) = \Pr(Y_i^\alpha \leqslant y)$ its distribution function. With slight abuse of notations, let $F_i^\alpha(v) = \Pr(v_i^\alpha \leqslant v)$ denote the distribution of bidder i's value estimate. Armed with these variables, we can decompose $U_i(\alpha)$ into three components as follows:

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

where the increasing and convex function σ_i^α is

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

By the dispersion theorem, $\alpha \succ_L \beta$ along with the identical means leads to $v_i^\alpha \geqslant_{cx} v_i^\beta$, which in turn leads to $\sigma_i^\alpha(y) \geqslant \sigma_i^\beta(y)$ for every y in the support of H_i^α by Lemma 2.16. Additionally, note that $Y_i^\alpha \geqslant_{icx} Y_i^\beta$ follows from Lemma 4.1 and that the function σ_i^β is increasing and convex. Accordingly, we have

$$\left.\begin{array}{lll} \sigma_i^\alpha(y) \geqslant \sigma_i^\alpha(y) \ \ \forall\, y & \Rightarrow & \mathbb{E}[\sigma_i^\alpha(Y_i^\alpha)] \geqslant \mathbb{E}[\sigma_i^\beta(Y_i^\alpha)] \\ Y_i^\alpha \geqslant_{icx} Y_i^\beta & \Rightarrow & \mathbb{E}[\sigma_i^\beta(Y_i^\alpha)] \geqslant \mathbb{E}[\sigma_i^\beta(Y_i^\beta)] \end{array}\right\} \Rightarrow \ \mathbb{E}[\sigma_i^\alpha(Y_i^\alpha)] \geqslant \mathbb{E}[\sigma_i^\beta(Y_i^\beta)]$$

The last inequality above can be interpreted as the bidder i's gain from greater precision of the signal, ignoring the strategic effect. Since more Lehmann-precise information causes his value estimate to be more dispersed, it is advantageous to the information rent. However, a similar argument applies to his opponents: $Y_i^\alpha \geqslant_{icx} Y_i^\beta$ implies $\mathbb{E}[Y_i^\alpha] \geqslant \mathbb{E}[Y_i^\beta]$, that is, 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 $U_i(\alpha)$ to be monotone:

$$\mathbb{E}[\sigma_i^\alpha(Y_i^\alpha)] - \mathbb{E}[\sigma_i^\beta(Y_i^\beta)] \geqslant \mathbb{E}[Y_i^\alpha] - \mathbb{E}[Y_i^\beta]. \tag{4.2}$$

This inequality 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 4.3 (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:*

$$\int_{\{v \geqslant \alpha(p)\}} (1 - F^\alpha(v))dv \geqslant \int_{\{v \geqslant \beta(p)\}} (1 - F^\beta(v))dv, \quad \forall\, p \in [0, 1], \tag{4.3}$$

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

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

where the bottom line is due to the change of variable $p = F^\alpha(y)$. Now we show that the given condition (4.3) is equivalent to

$$\int_{\{v\leqslant\alpha(p)\}} F^{\alpha}(v)\,dv - \alpha(p) \geqslant \int_{\{v\leqslant\beta(p)\}} F^{\beta}(v)\,dv - \beta(p) \quad \text{for each } p \in [0,1], \tag{4.4}$$

which is clearly sufficient for (4.2) to hold. To see this, we integrate the left-hand side of the inequality (4.4) by parts to obtain

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

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 (4.3) with Lehmann precision. To interpret the condition, note that a simple integration by parts gives

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

Hence the condition is equivalent to $\mathbb{E}[(v_i^{\alpha} - \alpha(p)) \vee 0] \geqslant \mathbb{E}[(v_i^{\beta} - \beta(p)) \vee 0]$ for every $p \in [0,1]$. To compare it with Lehmann precision, recall that $\alpha \succ_L \beta$ implies $v_i^{\alpha} \geqslant_{\text{icx}} v_i^{\beta}$, equivalently, $\mathbb{E}[(v_i^{\alpha} - c) \vee 0] \geqslant \mathbb{E}[(v_i^{\beta} - c) \vee 0]$ for every constant c, while the condition (4.3) sets out a requirement on every p-th quantile.[15] But (4.3) 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 in statistics:

Corollary 4.3 (Ganuza and Penalva (2010))**.** *Suppose that the index set* E *is the dispersive-ordered, i.e.,* $\alpha > \beta$ *implies* $v_i^{\alpha} \geqslant_{\text{disp}} v_i^{\beta}$*. Then the bidder's information rent is nondecreasing in* α*.*

[15] Jewitt (1989) has extended the mean-preserving spread of Rothschild and Stiglitz (1970) to develop a location-independent variability order (also known as *dilation order* in statistics literature), but it also imposes a condition on the quantiles of each distribution.

PROOF OF COROLLARY 4.3 : Recall that

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

for all $p, q \in [0, 1]$ and $p \leqslant q$.[16] 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 $q \geqslant p = F^{\alpha}(\alpha(p)) = F^{\beta}(\beta(p))$ to get

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

Changing the variable $F^{e}(q)^{-1} = x$ for each $e = \alpha, \beta$ and integrating by parts yields the condition (4.3). Since $v_i^{\alpha} \geqslant_{disp} v_i^{\beta}$ together with their same mean implies $v_i^{\alpha} \geqslant_{icx} v_i^{\beta}$, it follows from the dispersion theorem that α is more Lehmann-precise than β. Therefore, the monotone information rents follow from the preceding proposition. □

It is easy to check that the condition (4.3) is not necessary for $v_i^{\alpha} \geqslant_{disp} v_i^{\beta}$. Hence Proposition 4.3 is a generalization of Ganuza and Penalva (2010). We conclude this section with the following remarks:

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

REMARK 2 : If the condition (4.3) holds, the auctioneer would release information with smaller precision compared to the efficient precision level. To see this, recall that (4.3) guarantees $\mathcal{E}(\alpha) - R(\alpha) = \sum_i U_i(\alpha)$ to be nondecreasing in α. Namely,

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

the marginal revenue from greater precision would be smaller than the corresponding marginal efficiency. Consequently, for any arbitrary cost function $C(\alpha)$, the following monotone comparative statics hold.

$$\operatorname*{argmax}_{\alpha \in E} \mathcal{E}(\alpha) - C(\alpha) \trianglerighteq \operatorname*{argmax}_{\alpha \in E} R(\alpha) - C(\alpha).$$

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 independent-private-value (IPV) auction, the second-price auction provides the bidders with the (socially) efficient incentives for information acquisition, provided the total cost incurred for gathering information 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 that if a mechanism designer could decide on the information policy in an IPV environment, the decentralized covert information acquisition is more beneficial to the efficient information structure than the centralized overt information acquisition.

4.3 The Principal-Agent Problem

We turn to the value of information in a principal-agent model. In a standard agency model where the principal cannot observe the agent's productive action, it is assumed that there are mutually observed signals (e.g., outputs) conveying partial information about the hidden action and one of the signals can be used as a proxy in the contract to induce the agent to take the desired action. A natural question that arises is which signal the principal should make use of for provision of incentives.[17]

[17] Several papers examined the value of information in a different environment. In the agency model where the first-order approach is valid, Holmström (1979) has shown that if additional signal Y is uninformative about the agent's action in the sense that the signal X is Blackwell-sufficient for

The fundamental principle from the traditional principal-agent model is that the main source of problems of moral hazard stems from the inability of perfectly monitoring the agent's action. Hence it is conceivable that the more accurate information about the agent's action one signal imparts, the less intense the incentive problem becomes. In this aspect an agency model is closely related to a statistical decision problem. One subtle but important difference to realize, however, is that a signal is used to *estimate* the unknown parameter in the latter but is used to *assess* whether the agent indeed takes the desired action in the former. In light of this difference, Kim (1995) introduced a new signal ordering appropriate to the agency model. The aim of this section is to examine the relationship between Lehmann-precision and the MPS criterion developed in his paper and to shed light on the value of information in the agency framework.

Consider a static bilateral contract problem between a principal and an agent. The scheme of the contract game is a typical moral hazard problem. The principal designs a contract to make him an offer, and then the agent decides whether to accept the offer. If the agent accepts it, he chooses a productive but hidden action $a \in A$. Both contracting parties observe a realization of the signal and the contract is enforced. Before the principal decides what contract to offer, however, she decides which signal the payouts are contingent on.

To go into the details, suppose that there are two signals X and Y, both of which provide imperfect information about the agent's action and are observable to both parties and verifiable in the court. Let $F(x|a)$ and $G(y|a)$ denote the conditional distributions of X and Y on the agent's action, respectively. In the remainder of this subsection we will refer to $F(x|a)$ as an information system. We assume that the supports of X and Y are a unit interval $[0, 1]$, and that both $F(\cdot|a)$ and $G(\cdot|a)$ are absolutely continuous for each $a \in A$ and thereby they are endowed with the density functions $f(x|a)$ and $g(y|a)$ with respect to Lebesgue measure, respectively.

(X, Y), Y needs not be used at all in writing the contract. Within the same framework Kim (1995) and Jewitt (2007) established the MPS criterion (See below), a sufficient and necessary condition for X to be more informative than Y. In case the first-order approach is invalid, Grossman and Hart (1983) showed that if X is Blackwell-sufficient for Y regarding the agent's action X is more informative than Y, but not vice versa.

In addition, we shall assume that both $f(x|a)$ and $g(y|a)$ satisfy the MLRP and 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 $w \in \mathfrak{R}_+$ is increasing and concave while the disutility $c(a)$ the agent incurs from exerting effort a is increasing and convex. The additively separable payoff function implies that the agent's preferences over monetary transfer are independent of his action.

Given the information system $F(x|a)$, in order to implement an action a at the least cost, the principal must solve the following optimization problem:[18]

$$\min_{s \in \Psi} \ \sigma^F(s, a) \triangleq \int_0^1 s(x) f(x|a) dx$$

subject to

$$\int_0^1 u(s(x)) f(x|a) dx - c(a) \geqslant \overline{U} \tag{IR}$$

$$a \in \operatorname*{argmax}_{a' \in A} \int_0^1 v(s(x)) f(x|a') dx - c(a'). \tag{IC}$$

Let $s_f^a \in \Psi \equiv \{s : [0, 1] \to \mathfrak{R}_+ | \text{measurable}\}$ be the optimal sharing rule under the information system F and $\rho^F(s_f^a, a)$ be the corresponding value function, indicating the least expected cost of getting the agent to choose the recommended action. Under the information system G define s_g^a and $\rho^G(s_g^a, a)$ in the same manner. With these notions, we say that F is more efficient than G provided $\rho^F(s_f^a, a) \leqslant \rho^G(s_g^a, a)$ for all $a \in A$. In other words, the more efficient information system alleviates the problem of moral hazard in the sense that the principal can induce the agent to pick every action at a less expected cost.

In the differentiable environment where the first-order approach is valid, Holm-

[18] Grossman and Hart (1983) has shown that when the agent's payoff function is additively separable, a principal-agent problem with hidden action can be split into two parts: For each action $a \in A$, the principal first computes the minimum expected cost necessary for the agent to take a, and then she decides from the cost-benefit analysis which action to implement based on her payoff function.

ström (1979) has shown that the optimal sharing rule s_f^a under F must satisfy the following first-order condition:[19]

$$\frac{1}{\nu'(s_f^a(x))} = \lambda_f + \mu_f \cdot l_f^a(x), \tag{4.5}$$

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:[20]

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

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. In terms of a variability order, we can simply rewrite the criterion as $l_f^a(X) \geqslant_{cx} l_g^a(Y)$ for all a since both likelihood ration functions have the identical mean 0; $\mathbb{E}[l_f^a(X)] = \mathbb{E}[l_g^a(Y)] = 0$.

To understand the idea behind this comparative static result, observe that $s_f^a(x)$ hinges on the likelihood ratio function rather than on the density function itself: The first-order condition (4.5) tells us that the incentive scheme is based on the likelihood that the recommended action is indeed taken. Put differently, the more dispersed likelihood ratio function conveys more accurate information about the agent's hidden action, and thereby the principal can implement the desired action more easily.

Proposition 4.4 (MPS Criterion, Kim (1995))**.** *If* $l_f^a(X) \geqslant_{cx} l_g^a(Y)$, *then* F *is more efficient than* G *in the agency model.*[21]

Proof of Proposition 4.4 : Our proof is based on the fact that the principal's optimization problem is convex in the likelihood ratio function.

[19]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).

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

[21]Kim (1995) proved sufficiency of the MPS criterion and Jewitt (2007) proved necessity by taking the dual approach of the optimization problem.

Dropping some superfluous scripts, let $q_g = \lambda_g + \mu_g \cdot l_g$ and let $s(q_g)$ be the optimal sharing rule satisfying the first order condition (4.5) under the information system G. Define

$$L(q_g) \triangleq \nu(s(q_g))q_g - s(q_g).$$

Note that the second derivative of this function is $L''(q_g) = \nu'(s(q_g))s'(q_g) \geqslant 0$ and q_g is linear with respect to l_g, meaning that L is convex in l_g.[22] Hence it follows from the MPS criterion that $\mathbb{E}[L(\lambda_g + \mu_g \cdot l_f)] \geqslant \mathbb{E}[L(q_g)]$. Also note that the expectation on the right-side of the inequality reduces to

$$\begin{aligned}
\mathbb{E}[L(q_g)] &= \int_0^1 [\nu(s(q_g))q_g - s(q_g)]\, g(y|a)dy \\
&= -\int_0^1 s(q_g)g(y|a)dy + \lambda_g \int_0^1 \nu(s(q_g))g(y|a)dy \\
&\qquad\qquad + \mu_g \int_0^1 \nu(s(q_g))\frac{g_a}{g}(y|a)g(y|a)dy \\
&= \mathcal{L}(s_g^a) + \lambda_g c(a) + \mu_g c'(a),
\end{aligned}$$

where $\mathcal{L}(s_g^a)$ stands for the Lagrangian of the optimization problem under G evaluated at the optimum. We infer from the Kuhn-Tucker theorem that $\mathcal{L}(s_g^a) = -\rho^G(s_g^a, a)$. Substituting it back into the previous inequality gives

$$\mathbb{E}[L(\lambda_g + \mu_g \cdot l_f)] - \lambda_g c(a) - \mu_g c'(a) \geqslant -\rho^G(s_g^a, a).$$

The Lagrange multiplier theorem tells us that $\mathcal{L}(s_f^a) = -\rho^F(s_f^a, a)$ should be larger than the expression on the left-side of the last inequality. Therefore, the MPS criterion is sufficient for one information system to be more efficient than another.[23] □

[22]Note that the first derivative of L is $\nu'(s(q_g))s'(q_g)q_g + \nu(s(q_g)) - s'(q_g)$ by the product rule. Since $\nu'(s(q_g))q_g = 1$ by the first-order condition (4.5), however, the derivative simplifies to $\nu(s(q_g))$. The monotone property of $s(q_g)$ follows from (4.5) and the MLRP.

[23]It appears that necessity of the MPS criterion can be established by the same argument. Since the proof is more complicated, however, I will not pursue the necessity here.

The following result relates the MPS criterion to Lehmann precision, showing that the two statistical orderings are essentially equivalent when $f(x|a)$ and $g(x|a)$ are differentiable with respect to a.

Proposition 4.5. F *is more Lehmann-precise than* G *if and only if* $l_f^a \geqslant_{cx} l_g^a$ *for all* $a \in A$.

Proof of Proposition 4.5 : (**Sufficiency**) Recall from Definition 2.10 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:

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

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

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

Since the likelihood ratio l_f^a is increasing, for all $\tau \in [0,1]$, we have[24]

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

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

(**Necessity**) To show the converse, suppose on the contrary that $l_f^a \leqslant_{cv} l_g^a$ but F is not more Lehmann-precise than G. Then it follows from Lemma 2.12 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 that

$$G(y|a^*) - G(y|a) = \int_0^y \left(\frac{g(t|a^*)}{g(t|a)} - 1\right) g(t|a) dt. \tag{4.6}$$

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

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

Using (4.6), we can rewrite the inequality as

$$\begin{aligned} c_x\,[1 - G(y|a)] + \int_0^y \left(\frac{g(t|a^*)}{g(t|a)} - 1\right) g(t|a) dt \\ < c_x\,[1 - F(x|a)] + \int_0^x \left(\frac{f(t|a^*)}{f(t|a)} - 1\right) f(t|a) dt. \end{aligned} \tag{4.7}$$

[24] To see this, define the function $\psi : [0,1] \to \mathfrak{R}$ as

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

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, $\mathbb{E}[l_f^a(X) \wedge c_y]$.

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 24. Hence, (4.7) 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. □

Proposition 4.5 generalizes the characterization result of the MPS criterion in Kim (1995). Proposition 4 in his paper related the MPS criterion to Blackwell sufficiency, showing that if F is Blackwell sufficient for G the likelihood ratio function of F is the mean preserving spread of that of G. Subsequently, Kim demonstrated with one counterexample that the converse is not true. It can be easily checked that the information systems given in the example indeed satisfies Lehmann precision, though.

The two results above in this subsection state that Lehmann precision is also applicable to ordering information systems in favor of the principal in a principal-agent model. Proposition 4.5 reflects the fundamental insight of moral hazard in Holmström (1979) and Grossman and Hart (1983) that more precise information helps the principal to alleviate the incentive problem, and Proposition 4.4 discovers that this insight holds true to the extent that (1) the first-order approach is valid and (2) the agent's preferences over monetary income are independent of his actions.

When the first-order approach is not valid, however, Lehmann precision is not a suitable way of ordering information systems. To see this, recall from the example in Grossman and Hart (1983) that the MLRP *per se* does not guarantee monotonicity of the optimal sharing rule for every $a \in A$. Intuitively, in 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), she is willing to pay less for higher outcomes to render a_3 less attractive and thereby to render both associated incentive compatibility conditions simultaneously binding.[25] Hence the payouts would not be monotone in the signal's outcomes, in which case Lehmann precision is not a sufficient condition for the

[25]When the first-order approach holds, however, only the incentive compatibility condition for the lower action (a_1) is binding. Therefore, monotonicity of the optimal scheme immediately follows from the first-order condition (4.5).

signal to be informative.[26]

4.4 Delegation

The final application concerns the value of information in an organization where internal contingent transfer is infeasible.[27] 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, the principal often delegates the decision process to the agent and allows the agent to take an action on her behalf, so that she can utilize the agent's information.[28] 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 $\mathfrak{R}$. Furthermore, endowing the set of projects with an order structure $a_{k+1} \geqslant a_k$, we shall assume that $v(a, \theta)$ is supermodular in $(a; \theta)$ and $u(a, \theta)$ satisfies the single-crossing property in $(a; \theta)$. This preference structure captures minimally aligned preference between both parties in the sense that if the project a_{k+1} is more desirable to both parties at some state θ than a_k, it remains more desirable at every state $\theta' \geqslant \theta$. With such

[26] Jewitt (2007) imparted a concrete example where more Lehmann precise information system hurts the principal when the first-order approach does not hold.

[27] 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).

[28] 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).

complementary between a and θ, if a signal α is more valuable than β to the agent, so is it to the principal.

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.[29]

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 $[0,1]$.[30] Then the agent uses Bayes' rule to update his belief about θ given the signal realization x. The posterior belief is denoted $F^\alpha(\cdot|x)$. The agent's 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 $x_{k-1} \leqslant x \leqslant x_k$, the project a_k maximizes the interim expected payoffs:

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

[29]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.

[30]The assumption that the marginal distribution $M^\alpha(x)$ of x is uniform, regardless of the signals, is crucial in this section. This assumption can be justified the value of $M^\alpha(\cdot)$ as the outcome rather than x itself. If the density function of M^α is atomless, then both $M^\alpha(x)$ and x contain the same information. For this reason, many literature (for instance, Szalay (2009)) works with $M^\alpha(x)$ instead of x itself. Nevertheless, the assumption restricts either the set of possible experiments or the set of priors.

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

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

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.

This section attempts to answer the question: When the principal benefits from a more informed agent in delegation? In other words, we derive a sufficient condition for $R(\alpha) \geqslant R(\beta)$ 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.

To formalize this idea, 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

$$\begin{aligned} {}^u\Delta_k^\alpha(x) &= \int_\Theta [u(a_{k+1}, \theta) - u(a_k, \theta)]\, dF^\alpha(\theta|x) \text{ and} \\ {}^\nu\Delta_k^\alpha(x) &= \int_\Theta [\nu(a_{k+1}, \theta) - \nu(a_k, \theta)]\, dF^\alpha(\theta|x) \end{aligned}$$

to the agent and to the principal, respectively.

Definition 4.4 (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 $k = 1, \cdots, N-1$,*

$$\begin{aligned} &\textit{when } x_k < y_k \;\; {}^u\Delta_k^\alpha(x) \geqslant 0 \;\; \textit{implies} \;\; {}^\nu\Delta_k^\alpha(x) \geqslant 0 \;\; \textit{and} \\ &\textit{when } x_k > y_k \;\; {}^\nu\Delta_k^\alpha(x) \geqslant 0 \;\; \textit{implies} \;\; {}^u\Delta_k^\alpha(x) \geqslant 0. \end{aligned}$$

Recall that the function ${}^u\Delta_k^\alpha : [0,1] \to \mathfrak{R}$ changes its sign from negative to positive at the cutoff point x_k, and that both ${}^u\Delta_k^\alpha(x)$ and ${}^\nu\Delta_k^\alpha(x)$ satisfy the SC

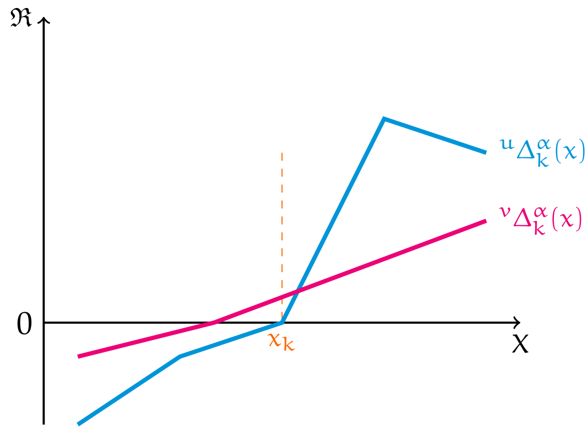


Figure 4.1: Aligned Preferences for $x_k < y_k$

property in x. When $x_k < y_k$, 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 4.1. When $x_k > y_k$, on the other hand, the function ${}^{u}\Delta_k^{\alpha}$ has to cross the axis first rather than ${}^{v}\Delta_k^{\alpha}$.

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

Proposition 4.6 (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.*

The proof begins with a simple property of the increasing convex order.

Lemma 4.5. *Let X be a random variable with the distribution F and $\sigma_1, \sigma_2 : \mathfrak{R} \to \mathfrak{R}$ be nondecreasing. Suppose that $\sigma_1(X) \geqslant_{icx} \sigma_2(X)$. Then for every c in the support of F, we have*

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

Proof of Lemma 4.5 : Given a constant c in the support of F, define $c_1 = \sigma_1(c)$. Since $\sigma_1(X) \geqslant_{\text{icx}} \sigma_2(X)$, $\mathbb{E}[\sigma_1(X) \vee c_1] \geqslant \mathbb{E}[\sigma_2(X) \vee c_1]$. Using integration by parts, we can rewrite the inequality as

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

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

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

Therefore, the desired result follows. □

Proof of Proposition 4.6 : Using the same methods in Appendix A.3, 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

3.14 that $\alpha \succ_L \beta$ implies ${}^{\nu}\Delta_k^{\alpha}(X) \geqslant_{icx} {}^{\nu}\Delta_k^{\beta}(Y)$. By Lemma 4.5, therefore, we have

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

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

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

is nonnegative. The proof is complete. □

Proposition 4.6 grasps a 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; ${}^{\nu}\Delta_k^{\alpha}(X) \geqslant_{icx} {}^{\nu}\Delta_k^{\beta}(Y)$. In other words, the principal's expected payoff function becomes more sensitive to the agent's private information. However, since the more informed agent will put more weight on the signal's outcome, his optimal selection rule becomes more sensitive to his information. Consequently, for such a strategy to be more beneficial to the principal, their preferences must be sufficiently aligned.

However, our result is limited to the extent that the aligned preference requires a substantial amount of structure on the primitive payoff functions, since the condition depends on the change in the agent's optimal rule from y_k to x_k responding to the change in the agent's information structure. In fact, a general discussion about the value of information in delegation is very difficult without formulating how the agent's optimal rule varies with the change in distributions. Although the agent's bias is very small, a better-informed agent may hurt the principal's payoff in delegation. The following example illustrates this point.[31]

[31] This is a slight variation of one example presented in Holmström (1977).

Example 4.6. *Let θ be uniformly distributed on the unit interval $(0,1)$. From the signal β, the agent has an informational 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 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 $\alpha \succ_L \beta$.[32]

Let $A = \{a_1, a_2, a_3\}$ with $a_1 < a_2 < a_3$. The payoffs of both parties from each project in θ_i are displayed 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 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 4.1: A better information may hurt the principal.

It is easy to see that the payoff function of each party belongs to $\mathcal{U}^{spm}$ and 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 ${}^{u}\Delta_1^{\alpha}(x) \geqslant 0$ imply ${}^{v}\Delta_1^{\alpha}(x) \geqslant 0$. However, for $x_2 \in (1/4, 1/2)$,

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

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

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 value of precise information* $R(\alpha) - R(\beta)$ *is negative.* $\square$

5 EPILOGUE

The main purpose of this dissertation has been to attempt to answer the question: "What is better information in an economic decision problem under a state of uncertainty?" The past literature on "comparison of signals" in statistics and economics has formulated several statistical signal orderings that represent a given decision maker's preference-based ordering. To achieve this within several classes of payoff functions used in economics, we developed a new statistical ordering based on the notion of dispersion and proved that this partial ordering provides a sufficient and necessary condition for one signal to be more valuable than another under both statistical and Bayesian frameworks. Furthermore, we characterized it in terms of another statistical ordering based on Lehmann Precision. As a consequence, we have established the equivalence of the four different ways to rank signals—Lehmann Precision, Effectiveness, Informativeness, and Dispersion.

Although the analysis we have pursued in this dissertation has been abstract and theoretical, some results of the equivalence theorem can be directly applied to practice. Above all things, the Dispersion Theorem is very useful. We have seen in Section 3.5 that the theorem can be used to justify in a unified way and consolidate several signal orderings developed by previous literature. Also we have illustrated in Chapter 4 how it can be applied to several strategic settings by carrying out the equilibrium analysis about the impact of precise signals.

This dissertation provides a natural guide to future research. Built upon the binary preference relation over a 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 in that it calls for a great deal of structures on the decision maker's primitive payoff functions and even the optimal decision rules.

When we dealt with the value of information from the Bayesian perspective, we in fact adopted a 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, there are many economic settings in which decision makers possess prior beliefs with a common property such as unimodal densities and thus some additional structure on the set of priors can be utilized. It will be interesting to investigate how such a structure of prior beliefs affects the decision maker's preference-based ordering on signals.

Another subject is the value of information in equilibrium models. Although the dispersion theorem provides a method for equilibrium analysis of the effect of precise information 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 (1990b).

A APPENDICES

A.1 The Karlin-Rubin Monotone Property

In this appendix we study another family of payoff functions, to which most of primitive functions in statistics—such as hypothesis testing and estimation problems—belong. We derive a couple of properties of the payoff functions that help to clarify the difference from the SC family. We also revisit Lehmann's Theorem to compare his improvement principle with the principle developed in Lemma 3.4. In addition, we establish the equivalence of the four signal orderings within this class.

We start with the formal definition 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*) for each* $\theta \in \Theta$*,* $u(a, \theta)$ *as a function of* a *is nonincreasing away from the* $A^*(\theta)$*.*

The KRM family stems from the two properties: (MCS) for monotone comparative statics and (SP) for the single peak. Refer to Figure A.1. When the action space is given by a convex set, the single peak condition is equivalent to quasiconcavity of the payoff function in a for each θ. In light of this fact together with (MCS), the family is dubbed "*quasiconcave payoff functions with increasing peaks*" in Quah and Strulovici (2009).[1]

Before proceeding, it is worth noting that the KRM family is distinct with the SC family as is pointed by Jewitt (2007). To see that the KRM payoff function does not necessarily satisfy the single-crossing property, observe the example depicted in Figure A.1. For a binary state space $\{\theta_L, \theta_H\}$ with $\theta_L < \theta_H$, the payoff function $u(a, \theta)$ is described as a function of a. 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, the incremental returns from taking

[1] The term "Karlin-Rubin Monotone" is taken from Jewitt (2007).

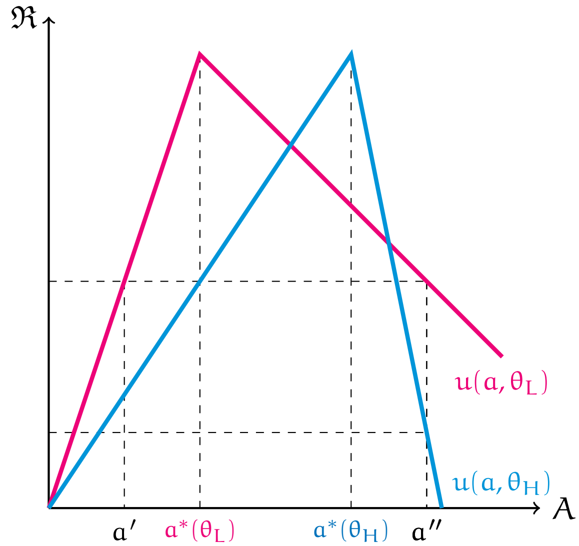


Figure A.1: The displayed function $u(a, \theta)$ satisfies the KRM property but does not satisfy the SCP.

a'' rather than a' is nonnegative in θ_L but negative in θ_H, so the displayed payoff function violates the SC property.

Likewise, the SC property does not imply the KRM property. Figure A.2 displays an example, where for each of three possible actions the payoff function is sketched as a function of θ.[2] 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.[3] 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 $\theta \in (\theta_2, \theta_1)$.

The payoff function displayed in Figure A.2 provides a clue to the difference between $\mathcal{U}^{KRm}$ and $\mathcal{U}^{sc}$. In order to formalize it, we first establish the following

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

[3] Note that both of them are supermodular in $(a; \theta)$. Therefore, the examples in Figure A.2 also tell us that there is no relationship between $\mathcal{U}^{spm}$ and $\mathcal{U}^{KRm}$, either.

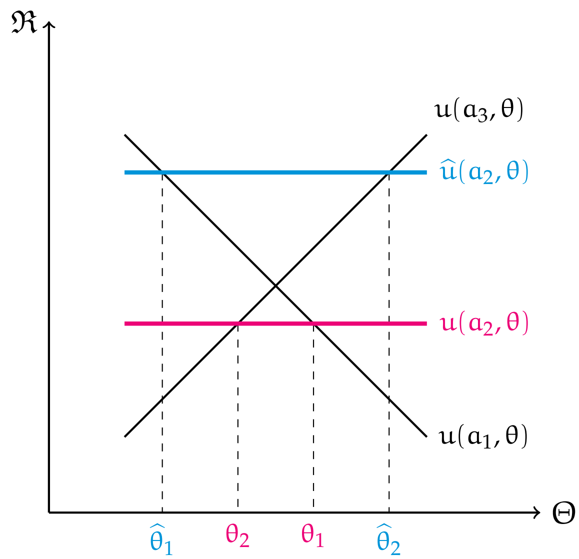


Figure A.2: For both $\hat{u}(a_2, \theta)$ and $u(a_2, \theta)$ the payoff function satisfies the SC property. For $u(a_2, \theta)$, however, it violates the KRM property.

lemma.

Lemma A.2. *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*.*

Proof of Lemma A.2 : Suppose that ${}^u\Delta_k(\theta_k) \geqslant 0$ for some θ_k. By the single peak condition of $\mathcal{U}^{KRm}$, it must be the case that $\min A^*(\theta_k) \geqslant a_{k+1} > a_k$, where $\min A^*(\theta_k)$ indicates the smallest element in $A^*(\theta_k)$. For every $\theta \geqslant \theta_k$, the condition (MCS) implies $\min A^*(\theta) \geqslant \min A^*(\theta_k) \geqslant a_{k+1}$. Hence ${}^u\Delta_k(\theta) \geqslant 0$ is immediate from (SP). □

It should be clear that the preceding lemma does not imply $\mathcal{U}^{KRm} \subset \mathcal{U}^{sc}$, since the single-crossing property does not necessarily hold for any other higher actions than a_{k+1} unlike $\mathcal{U}^{sc}$. In a binary action space, however, the lemma tells us $\mathcal{U}^{KRm} \subset \mathcal{U}^{sc}$. Furthermore, since (MCS) is implied by the SC property and (SP) is always true in a binary action space, we have $\mathcal{U}^{sc} \subset \mathcal{U}^{KRm}$, that is, the two classes are equivalent.

Corollary A.3. *In a binary action space* $\mathcal{U}^{KRm} = \mathcal{U}^{sc}$.

With a binary action space, Milgrom and Shannon (1994) showed that the SC property of $u(a, \theta)$ is sufficient and necessary for the set of maximizers to be nondecreasing in the strong set order. The equivalence between $\mathcal{U}^{KRm}$ and $\mathcal{U}^{sc}$ provides an alternative proof of necessity of the SC property for (MCS) to hold.

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

Lemma A.4. *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* $\theta_{k+1} \geqslant \theta_k$.

Proof of Lemma A.4 : 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 ${}^u\Delta_{k+1}(\theta^*) \geqslant 0$ but ${}^u\Delta_k(\theta^*) < 0$. ${}^u\Delta_{k+1}(\theta^*) \geqslant 0$ implies $a_k < a_{k+1} < a_{k+2} \leqslant \min A^*(\theta^*)$, which in turn implies ${}^u\Delta_k(\theta^*) \geqslant 0$, a contradiction. □

The preceding lemma clarifies why the payoff function depicted in Figure A.2 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.[4]

Focusing on this class of payoff functions, Lehmann (1988) proved that a 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 differently, he develops a statistical signal ordering which is equivalent to the signal ordering based on the statistical values within $\mathcal{U}^{KRm}$.[5] Like the proofs of Theorem 3.2 for $\mathcal{U}^{sc}$, Lehmann's proof is based on the improvement principle—the way to find a monotone decision rule admissible under α that makes every statistician within $\mathcal{U}^{KRm}$ better off for every monotone decision rule admissible

[4]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.

[5]Like the three classes of payoff functions, even $\mathcal{U}^{KRm}$ is invariant to the addition of $h \in \mathcal{H}$ in Lemma 2.5. Therefore, the criterion of effectiveness validates the signal ordering based on the statistical values within $\mathcal{U}^{KRm}$ as well.

under β—and on the essentially complete class theorem established by Karlin and Rubin (1956).

Before turning to the next result, it is instructive to compare his improvement principle with the principle developed in Lemma 3.4. To go into details, we shall 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 from negative to positive for each $k = 1, \cdots, n-1$. By Lemma A.4, it must be the case that $\widehat{\theta}_1 \leqslant \cdots \leqslant \widehat{\theta}_{n-1}$.

Now consider a monotone decision rule $d^\beta : \mathcal{Y} \to A$ permissible under β. In every finite action space, a monotone strategy can be represented by a nondecreasing sequence of $n-1$ cutoff points in the sample space, $\{y_1, \cdots, y_{n-1}\}$, where the strategy d^β assigns the action a_k to $y \in (y_{k-1}, y_k)$ for each k. The improvement principle in Lehmann (1988) argues that if $\alpha \succ_L \beta$, there exists a monotone strategy $d^\alpha = \{x_1, \cdots, x_{n-1}\}$ permissible under α which yields more expected payoffs than those generated by the monotone strategy d^β regardless of θ. The payoff-improving strategy d^α can be constructed by the T-transformation in Definition 2.10. Given the monotone strategy $d^\beta = \{y_1, \cdots, y_{n-1}\}$, d^α is given by the set of change points $\{x_1, \cdots, x_n\}$, where we define

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

The constructed strategy is indeed monotone, since

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

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

To see that d^α leads to more expected payoffs than d^β, note that the difference between $\rho^\alpha(d^\alpha, \theta)$ and $\rho^\beta(d^\beta, \theta)$ is simplified into

$$\begin{aligned}
\rho^\alpha(d^\alpha, \theta) - \rho^\beta(d^\beta, \theta) &= \sum_{k=1}^{n-1} {}^u\Delta_k(\theta)\,[1 - G^\alpha(x_k|\theta)] - \sum_{k=1}^{n-1} {}^u\Delta_k(\theta)\left[1 - G^\beta(y_k|\theta)\right] \\
&= \sum_{k=1}^{n-1} {}^u\Delta_k(\theta)\left[G^\beta(y_k|\theta) - G^\alpha(x_k|\theta)\right] \\
&= \sum_{k=1}^{n-1} {}^u\Delta_k(\theta)\left[G^\beta(y_k|\theta) - G^\alpha(T_{y_k}(\widehat{\theta}_k)|\theta)\right]
\end{aligned}$$

Consequently, every term in the bottom line above is nonnegative for every θ and thus $\rho^\alpha(d^\alpha, \theta) \geqslant \rho^\beta(d^\beta, \theta)$.

The above principle, however, does not apply to $u \in \mathcal{U}^{sc}$ since the SC property imposes no order structures on the set of cutoff points $\{\widehat{\theta}_k\}$.[6] 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 $\Theta = [\underline{\theta}, \overline{\theta}]$—$\underline{\theta} = \theta_0^* \leqslant \theta_1^* \leqslant \cdots \leqslant \theta_{n-1}^* \leqslant \theta_n^* = \overline{\theta}$—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 3.4 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.4, while it is determined by the monotone property of $d^\star$ for $\mathcal{U}^{sc}$.

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

[6] Recall the example in Figure A.2. Despite $u(a_2, \theta)$ being lower than the intersection, the payoff function satisfies the SC property.

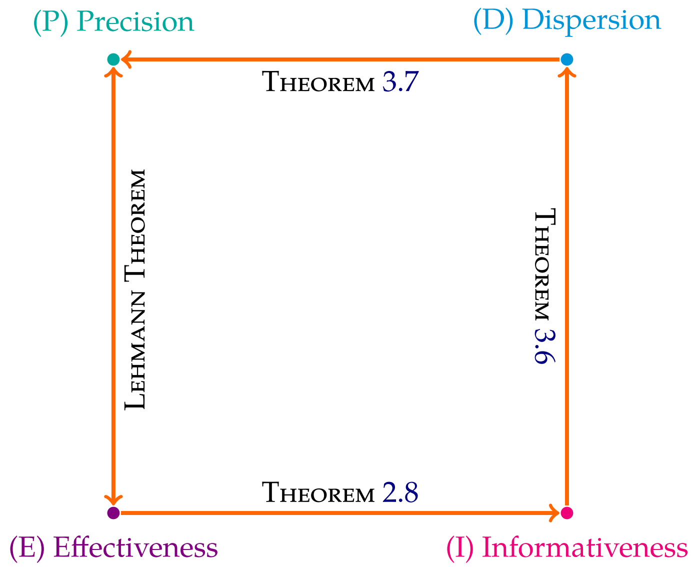


Figure A.3: Lehmann (1988) has shown the equivalence between (P) and (E) within $\mathcal{U}^{\mathsf{KRm}}$. Corollary A.5 extends his equivalence theorem into Bayesian decision theory with the dispersion-based signal ordering.

Corollary A.5. *Within the class of payoff functions $\mathcal{U}^{\mathsf{KRm}}$, the four signal orderings stated in the main theorem are equivalent.*

Proof of Corollary A.5 : Since (P) Precision $\Rightarrow$ (E) Effectiveness is done by Lehmann (1988) and both implications—(E) $\Rightarrow$ (I) Informativeness and (D) Dispersion $\Rightarrow$ (P)—hold universally, it suffices to establish (I) $\Rightarrow$ (D). To this end, we need show that the primitive $u_{\psi,\kappa} \in \mathcal{U}^{\mathsf{KRm}}$ defined in the proof of Theorem 3.6 for every $\kappa \in \mathfrak{R}$ and $\psi \in \Theta^*$. But it is straightforward from Lemma A.2. □

Corollary A.5 is a generalization of the Lehmann's Theorem into another two equivalent signal orderings, one based on the Bayesian values and another based on the dispersion, within the KRM family. Refer to Figure A.3.

A.2 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 this concept 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 $[a', a''] = \{a \in A |\ a' \leqslant a \leqslant a''\}$. 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 Figure A.1 clarifies this fact.

The IDO class also contains the KRM class as well. To see this, 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.2. Thus we suppose that (a', a'') is nonempty and for some θ

$$u(a'', \theta) \geqslant u(a, \theta) \quad \forall\ a \in [a', a'']. \tag{$\bigstar$}$$

Let $a^*(\theta) = \min A^*(\theta)$ at each state θ. When $a'' \notin A^*(\theta)$, the inequality ($\bigstar$) leads us to $a' < a'' < a^*(\theta) \leqslant a^*(\theta')\ \forall\ \theta' > \theta$ due to the (MCS) and (SP) conditions of $\mathcal{U}^{KRm}$. Hence $u(a'', \theta') \geqslant u(a', \theta')$ 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 of payoff functions considered in this paper.

Lemma A.6. *$\mathcal{U}^{sc} \cup \mathcal{U}^{KRm} \subset \mathcal{U}^{ido}$.*

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 result only, we denote by $\mathcal{U}^{SP}$

the set of payoff functions satisfying (SP).

Proposition A.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 A.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 $u(a', \theta) \geqslant u(\widehat{a}, \theta)$ for all $\widehat{a} \in [a, a']$. Then it follows from the quasiconcavity that $u(\lambda a + (1-\lambda)a', \theta) \geqslant u(a, \theta) \wedge u(a', \theta) = u(a, \theta)$ 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

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

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 $u(a_2, \theta') \geqslant u(a^*, \theta') \wedge u(a_1, \theta') = u(a_1, \theta')$, which is a contradiction. Accordingly, $u(\cdot, \theta)$ is nondecreasing over the interval $[a, a']$ and therefore $u(a', \theta') \geqslant u(a, \theta')$ for all $\theta' > \theta$. $\square$

Table A.1 summarizes our discussion about the classes of payoff functions concerning θ. In a binary action space, Lemma A.2 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

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 A.1: Payoff Functions Relationship between SPM, SCP, KRM, and IDO

proposition demonstrates that the three classes with KRM, IDO, and MCS are mutually equivalent. The example in Figure A.2-(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 3.3 to $\mathcal{U}^{ido}$.

Lemma A.7. *Let $u \in \mathcal{U}^{ido}$. Then for every decreasing function $d : \Theta \to A$, there is an action $a^* \in A$ for which $u(a^*, \theta) \geqslant u(d(\theta), \theta)$ for every state $\theta \in \Theta$.*

Proof of Lemma A.7 : We infer from Theorem 1 in Quah and Strulovici (2009) that for every $u \in \mathcal{U}^{ido}$, there is a nondecreasing strategy $d^\star(\theta) \in A^*(\theta) \equiv \operatorname{argmax}_{a \in A} u(a, \theta)$. Suppose that d and $d^\star$ intersect at a point $(\theta^\star, a^\star)$. Then $d^\star(\theta^\star) = a^\star$ implies

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

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

One immediate consequence of the preceding lemma is that the improvement principle, Lemma 3.4, applies to $\mathcal{U}^{\text{ido}}$. This in turn leads to the following result:

Lemma A.8. *Suppose* $\alpha \succ_L \beta$. *Then for every strategy* $d \in \mathcal{D}^{\beta,M}$, *there exists* $d^\alpha \in \mathcal{D}^{\alpha,M}$ *such that* $\rho^\alpha(d^\alpha,\theta) \geqslant \rho^\beta(d,\theta)$ *for all* $\theta \in \Theta$.

Proof of Lemma A.8 : For each $y \in \mathcal{Y}$, let $T_y : \Theta \to \mathcal{X}$ represent the T-transformation associated with α and β. The improvement principle guarantees the existence of d^α for which

$$u(d^\alpha \circ T_y(\theta),\theta) \geqslant u(d(y),\theta), \quad \text{for all } y \in \mathcal{Y} \text{ and } \theta \in \Theta.$$

Taking the integral of both sides of the above inequality with respect to the probability distribution $G^\beta(\cdot|\theta)$ gives

$$\begin{aligned}\int_{\mathcal{Y}} u(d^\alpha \circ T_y(\theta),\theta)dG^\beta(y|\theta) &= \int_{\mathcal{X}} u(d^\alpha(x),\theta)dG^\alpha(x|\theta) = \rho^\alpha(d^\alpha,\theta)\\ &\geqslant \int_{\mathcal{Y}} u(d(y),\theta)dG^\beta(y|\theta) = \rho^\beta(d,\theta). \quad \square\end{aligned}$$

Lemma A.8 plays a central role in establishing the following important result.

Theorem A.9 (Essentially Complete Class)**.** *For each decision problem* (α, u) *with* $\alpha \in S$ *and* $u \in \mathcal{U}^{\text{ido}}$, *the collection of monotone strategies* $\mathcal{D}^{\alpha,M}$ *constitutes an essentially complete class.*

Proof of Theorem A.9 : Dropping the superscript α for a simple exposition, let $G(\cdot|\theta)$ denote the probability distribution of X. Furthermore, we assume that the support of X is $\mathcal{X} = [0,1]$ and the action space is finite $A = \{a_1, \cdots, a_n\}$. The extension to a general compact space follows via a limiting argument.

The proof will proceed in three steps. Given an arbitrary strategy $d \in \mathcal{D}$ and a set of distributions $\{G(\cdot|\theta)\}_{\theta\in\Theta}$, we first construct a monotone strategy $d^\star : [0,1] \to A$ and another set of distributions $\{\overline{G}(\cdot|\theta)\}_{\theta\in\Theta}$ *payoff-equivalent* to the given pair of

d and $\{G(\cdot|\theta)\}_{\theta\in\Theta}$. We then show that the original set of distributions G is more Lehmann-precise than $\overline{G}$. Finally, we use the preceding lemma to complete the proof.

For each $k = 1, 2, \cdots, n$ and $\theta \in \Theta$, let $\overline{G}(k/n|\theta) = \Pr(d(X) \leqslant a_k|\theta) = \int_0^1 \mathbb{1}_{\{d(x)\leqslant a_k\}}(x)dG(x|\theta)$ and define $x_k(\theta) \in [0,1]$ to be a point at which $G(x_k(\theta)|\theta) = \overline{G}(k/n|\theta)$. For $y \in ((k-1)/n, k/n)$ such that $y = \lambda \cdot (k-1)/n + (1-\lambda)\cdot k/n$ for some $\lambda \in (0,1)$, let $\overline{G}(y|\theta) = G(\lambda x_{k-1}(\theta) + (1-\lambda)x_k(\theta)|\theta)$. Regarding the new strategy $d^\star$, let $d^\star(y) = a_k$ for $y \in ((k-1)/n, k/n)$. This completely characterizes both $d^\star$ and $\overline{G}$.

To see that the pair $d^\star$ and $\overline{G}$ generate the same expected payoffs as d and G, i.e., $\rho^G(d,\theta) = \rho^{\overline{G}}(d^\star,\theta)$ for all θ, note that

$$\begin{aligned}
\int_0^1 u(d(x),\theta)dG(x|\theta) &= \sum_{k=1}^{n} u(a_k,\theta)\cdot \Pr(d(X) = a_k|\theta)\\
&= \sum_{k=1}^{n} u(a_k,\theta)\cdot [G(x_k(\theta)|\theta) - G(x_{k-1}(\theta)|\theta)]\\
&= u(a_1,\theta) + \sum_{k=1}^{n-1} \Delta_k(\theta)\cdot[1 - G(x_k(\theta)|\theta)]\\
&= u(a_1,\theta) + \sum_{k=1}^{n-1} \Delta_k(\theta)\cdot\left[1 - \overline{G}(k/n|\theta)\right]\\
&= \int_0^1 u(d^\star(y),\theta)d\overline{G}(y|\theta),
\end{aligned}$$

where $\Delta_k(\theta) = u(a_{k+1},\theta) - u(a_k,\theta)$.[7]

Now we show that G is more Lehmann-precise than $\overline{G}$. Recall that $\overline{G}(k/n|\theta) = \int_0^1 \mathbb{1}_{\{d(x)\leqslant a_k\}}(x)dG(x|\theta) = G(x_k(\theta)|\theta)$. That is, for all θ, we have

$$\overline{G}(k/n|\theta) - G(x_k(\theta)|\theta) = \int_0^1 \left[\mathbb{1}_{\{d(x)\leqslant a_k\}}(x) - \mathbb{1}_{\{x\leqslant x_k(\theta)\}}(x)\right] g(x|\theta)dx = 0.$$

[7] The third equality follows from the fact that $\sum_{k=1}^{n} a_k(b_k - b_{k-1}) = a_1 + \sum_{k=1}^{n-1}(a_{k+1} - a_k)(b_n - b_k)$ for two finite sequences $\{a_k\}_{k=1}^n$ and $\{b_k\}_{k=1}^n$ with $b_k \leqslant b_{k+1}$.

Observe that the above stochastic integral obeys

$$\int_c^1 \left[\mathbb{1}_{\{d(x)\leqslant a_k\}}(x) - \mathbb{1}_{\{x\leqslant x_k(\theta)\}}(x)\right] g(x|\theta)dx \geqslant 0 \quad \forall\, c \in [0,1].$$

Hence it follows from the Banks' inequality that[8]

$$\int_0^1 \left[\mathbb{1}_{\{d(x)\leqslant a_k\}}(x) - \mathbb{1}_{\{x\leqslant x_k(\theta)\}}(x)\right] g(x|\theta')dx \geqslant 0 \quad \forall\, \theta' \geqslant \theta. \tag{A.1}$$

Since $\overline{G}(k/n|\theta') - G(x_k(\theta')|\theta') = 0$ and $G(\cdot|\theta)$ is nondecreasing, it follows from (A.1) that $x_k(\theta)$ is nondecreasing in θ for each k. Note that for each y there exist $\lambda \in [0,1]$ and $k = 1,2,\cdots,n$ for which $\overline{G}(y|\theta) = G(\lambda x_{k-1}(\theta) + (1-\lambda)x_k(\theta)|\theta)$. Since $\lambda x_{k-1}(\theta) + (1-\lambda)x_k(\theta)$ is nondecreasing in θ, it follows from Definition 2.10 that G is more Lehmann-precise than $\overline{G}$.

So far we showed that for an arbitrary strategy $d \in \mathcal{D}$ and a signal structure G, there exist a monotone strategy $d^\star \in \mathcal{D}^{\overline{G},M}$ and another signal structure $\overline{G}$ dominated by G in the Lehmann-precision order such that $\rho^G(d,\theta) = \rho^{\overline{G}}(d^\star,\theta)$ for all θ. Therefore, by Lemma A.8, there exists a monotone strategy d^M such that $\rho^G(d,\theta) \leqslant \rho^G(d^M,\theta)$. The proof is complete. □

Now we are ready to extend the equivalence theorem to $\mathcal{U}^{ido}$, the largest class of payoff functions. This result completes the proof of the main theorem presented in Section 3.

Corollary A.10. *Within the class of payoff functions $\mathcal{U}^{ido}$, the four signal orderings stated in the main theorem are equivalent.*

Proof of Corollary A.10 : It suffices to show that (P) Precision ⇒ (E) Effectiveness and (I) Informativeness ⇒ (D) Dispersion. (P) ⇒ (E) follows from Lemma A.7 and

[8]The inequality in Banks (1963) states that if a function $h : X \to \mathfrak{R}$ satisfies $\int_X \mathbb{1}_{\{x\leqslant t\}\cap X}(x)h(x)d\pi(x) \geqslant 0$ for a positive measure π, then $\int_X \phi(x)h(x)d\pi(x) \geqslant 0$ for all non-negative and nondecreasing functions $\phi : X \to \mathfrak{R}$.

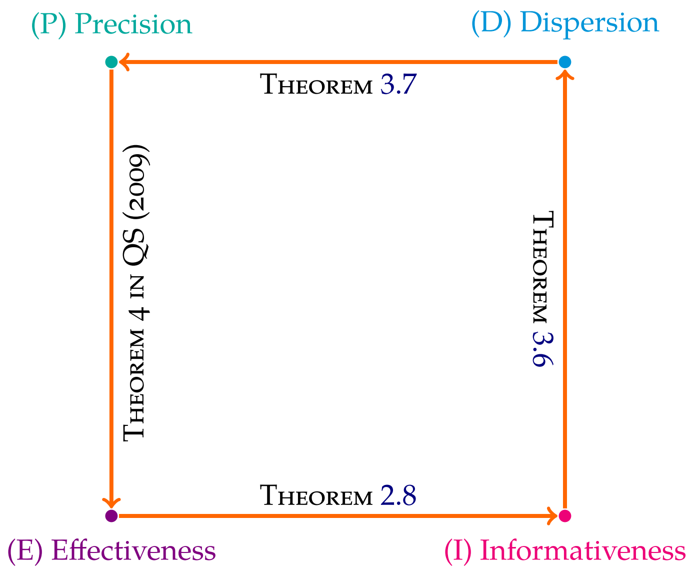


Figure A.4: Quah and Strulovici (2009) has shown (P) $\Rightarrow$ (E) $\Rightarrow$ (I) within $\mathcal{U}^{ido}$. The dispersion-based signal order extends their result to complete the value of information in decision theory.

Theorem 3.2. (I) $\Rightarrow$ (D) holds since the set of payoff functions $\{u_{\psi,\kappa} | \psi \in \Theta^*, \ \kappa \in \mathfrak{R}\}$ in Theorem 3.6 is included in $\mathcal{U}^{ido}$. $\square$

A.3 Informativeness for $\mathcal{U}^{spm}$

In this section, we use the dispersion theorem (Corollary 3.14) to provide a more concise proof of relating (P) Lehmann Precision to (I) Informativeness within the class of 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's preference is represented by a payoff function $u(a, \theta) \in \mathcal{U}^{spm}$. 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 pointwise maximizes the expected payoffs $\mathbb{E}_\theta[u(a,\theta)|X=x]$. In a finite action space, therefore, δ^α can be represented by a $(n-1)$-tuple vector $\{x_1,\cdots,x_{n-1}\}$ such that to each outcome $x\in(x_{k-1},x_k)$ δ^α assigns a_k for each $k=1,2,\cdots,n-1$. This formulation enables us to write the Bayesian value of α as

$$\begin{aligned}
\mathcal{V}^\pi(\alpha,u) &= \mathbb{E}_\theta\left[\int_{\mathcal{X}} u(\delta^\alpha(x),\theta)dG^\alpha(x|\theta)\right]\\
&= \mathbb{E}_\theta\left[\sum_{k=1}^{N}\int_{x_{k-1}}^{x_k} u(a_k,\theta)dG^\alpha(x|\theta)\right]\\
&= \mathbb{E}_\theta\left[u(a_1,\theta)+\sum_{k=1}^{N-1}\int_{\mathcal{X}}\mathbb{1}_{\{x\geqslant x_k\}}(x)\left[u(a_{k+1},\theta)-u(a_k,\theta)\right]dG^\alpha(x|\theta)\right]\\
&= E_\theta\left[u(a_1,\theta)\right]\\
&\quad+\sum_{k=1}^{N-1}\int_{\mathcal{X}}\mathbb{1}_{\{x\geqslant x_k\}}(x)\left\{\int_\Theta\left[u(a_{k+1},\theta)-u(a_k,\theta)\right]dF^\alpha(\theta|x)\right\}dM^\alpha(x),
\end{aligned}$$

where $M^\alpha(\cdot)$ is the marginal distribution of x under the signal α.

For simple exposition, we define the integral in the above curly bracket 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 $u\in\mathcal{U}^{spm}$, recall that the incremental return $u(a_{k+1},\theta)-u(a_k,\theta)$ is nondecreasing in θ. Thus, it follows from the dispersion theorem that $\alpha\succ_L\beta$ implies $^u\Delta_k^\alpha(X)\geqslant_{icx}{}^u\Delta_k^\beta(Y)$ where

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

Observe that the function $^u\Delta_k^\alpha:\mathcal{X}\to\mathfrak{R}$ crosses the horizontal-axis at the cutoff point x_k.[9] Hence we can simplify the iterated integral into

$$\int_{\mathcal{X}} \mathbb{1}_{\{x \geqslant x_k\}}(x)\, {}^{u}\Delta_k^{\alpha}(x)\ dM^{\alpha}(x) = \mathbb{E}\left[{}^{u}\Delta_k^{\alpha}(X) \vee 0\right]. \tag{A.2}$$

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

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

Consequently,

$$\mathcal{V}^{\pi}(\alpha, u) - \mathcal{V}^{\pi}(\beta, u) = \sum_{k=1}^{N-1} \left\{ \mathbb{E}\left[{}^{u}\Delta_k^{\alpha}(X) \vee 0\right] - \mathbb{E}\left[{}^{u}\Delta_k^{\beta}(Y) \vee 0\right] \right\} \geqslant 0$$

for every prior $\pi \in P(\Theta)$ and $u \in \mathcal{U}^{spm}$ and therefore, α is more informative than β with respect to the class $\mathcal{U}^{spm}$. $\square$

A.4 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.[10] Rather than the SCC condition,

[9]Even in the case it does not cross the axis at a single point, the function ${}^{u}\Delta_k^{\alpha}$ will assume 0 at x_k and so the equation (A.2) must hold.

[10]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.

we use the dispersion theorem to extend this result to all incentive-compatible auctions.

For an illustration, suppose that there are two risk-neutral bidders who compete for a single indivisible object.[11] 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 $0 \leqslant p_i \leqslant 1$ and $\sum_i p_i \leqslant 1$ indicating the probability that the bidder i is awarded the object, and $\mathbf{t} = (t_1, t_2)$ is a *payment scheme* with $t_i \in \mathfrak{R}$ 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 $t_i : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathfrak{R}$ 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

[11] 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 4.2.

$x_i \in \mathcal{X}_i$ 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) $u_i^{\alpha}(x_i) \triangleq u_i^{\alpha}(x_i, x_i) \geqslant 0$ 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*,[12] the mechanism satisfies the (IC) condition if and only if (M) $\widehat{p}_i(\cdot)$ is nondecreasing for all i and the next envelope formula holds:

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

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

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

envelope theorem holds:

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

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

Proposition A.2 (Value of Information to Bidders)**.** *If* $\alpha \succ_L \beta$*, then* $\mathcal{U}_i(\alpha) \geqslant \mathcal{U}_i(\beta)$*.*

Proof of Proposition A.2 : 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 3.14, $\alpha \succ_L \beta$ implies $v_i^\alpha \geqslant_{icx} v_i^\beta$, and thus $\mathcal{U}_i(\alpha) \geqslant \mathcal{U}_i(\beta)$. □

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.

REFERENCES

Aghion, Philippe, and Jean Tirole. 1997. Formal and real authority in organizations. *Journal of Political Economy* 105(1):1–29.

Alonso, Ricardo, and Niko Matouschek. 2008. Optimal delegation. *Review of Economic Studies* 75(1):259–293.

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

Athey, Susan. 2001. Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. *Econometrica* 69(4):861–89.

Athey, Susan, and Jonathan Levin. 2001. The value of information in monotone decision problems. Working Papers 01003, Stanford University, Department of Economics.

Banks, Dallas. 1963. An integral inequality. *Proceedings of the American Mathematical Society* 14(5):823–828.

Bergemann, Dirk, and Stephen 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, Dirk, and Martin Pesendorfer. 2007. Information structures in optimal auctions. *Journal of Economic Theory* 137(1):580–609.

Bergemann, Dirk, and Juuso 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, David. 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.

Bulow, Jeremy, Ming Huang, and Paul Klemperer. 1999. Toeholds and takeovers. *Journal of Political Economy* 107(3):427–454.

Crawford, Vincent P, and Joel Sobel. 1982. Strategic information transmission. *Econometrica* 50(6):1431–51.

Esö, Peter, and Balazs Szentes. 2007. Optimal information disclosure in auctions and the handicap auction. *Review of Economic Studies* 74(3):705–731.

Folland, Gerald B. 1999. *Real analysis : Modern techniques and their applications*. Wiley Interscience.

Ganuza, Juan-JosÈ, and JosÈ S. Penalva. 2010. Signal orderings based on dispersion and the supply of private information in auctions. *Econometrica* 78(3):1007–1030.

Grossman, Sanford J, and Oliver D Hart. 1983. An analysis of the principal-agent problem. *Econometrica* 51(1):7–45.

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

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

Holmström, Bengt. 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, Ian. 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. Tech. Rep., Working Paper.

Johnson, Justin P., and David P. Myatt. 2006. On the simple economics of advertising, marketing, and product design. *American Economic Review* 96(3):756–784.

Karlin, Samuel, and Herman Rubin. 1956. The theory of decision procedures for distributions with monotone likelihood ratio. *The Annals of Mathematical Statistics* 27(2):272–299.

Kim, Son Ku. 1995. Efficiency of an information system in an agency model. *Econometrica* 63(1):89–102.

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

Lewis, Tracy R, and David E M Sappington. 1994. Supplying information to facilitate price discrimination. *International Economic Review* 35(2):309–27.

McLean, Richard, and Andrew Postlewaite. 2002. Informational size and incentive compatibility. *Econometrica* 70(6):2421–2453.

Milgrom, Paul. 2004. *Putting auction theory to work*. Cambridge University Press: Cambridge.

Milgrom, Paul, and John Roberts. 1990a. The economics of modern manufacturing: Technology, strategy, and organization. *American Economic Review* 80(3):511–28.

———. 1990b. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica* 58(6):1255–77.

Milgrom, Paul, and Chris Shannon. 1994. Monotone comparative statics. *Econometrica* 62(1):157–80.

Milgrom, Paul R. 1981. Good news and bad news: Representation theorems and applications. *Bell Journal of Economics* 12(2):380–391.

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

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

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

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

———. 2012. Aggregating the single crossing property. *Econometrica* 80(5):2333–2348.

Rezende, Leonardo. 2013. Mid-auction information acquisition. Tech. Rep., PUC-Rio, Department of Economics.

Rothschild, Michael, and Joseph E. Stiglitz. 1970. Increasing risk: I. a definition. *Journal of Economic Theory* 2(3):225–243.

Shaked, Moshe, and J. George Shanthikumar. 2007. *Stochastic orders and their applications*. Academic Press, Inc.

Shi, Xianwen. 2012. Optimal auctions with information acquisition. *Games and Economic Behavior* 74(2):666–686.

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

Tchen, Andre H. 1980. Inequalities for distributions with given marginals. *The Annals of Probability* 8(4):814–827.

Topkis, Donald M. 1978. Minimizing a submodular function on a lattice. *Operations Research* 26:305–321.

———. 1998. *Supermodularity and complementarity*. Princeton University Press.

Wald, Abraham. 1950. *Statistical decision functions*. John Wiley & Sons, New York.