The Value of Information in Monotone Decision Problems*

Susan Athey
Stanford and NBER

Jonathan Levin
Stanford University

First Draft: September 1997
This Draft: December 2001

Abstract

This paper studies decision problems under uncertainty where a decision-maker observes an imperfect signal about the true state of the world. We analyze the information preferences and information demand of such decision-makers, based on properties of their payoff functions. We restrict attention to "monotone decision problems," whereby the posterior beliefs induced by the signal can be ordered so that higher actions are chosen in response to higher signal realizations. Monotone decision problems are frequently encountered in economic modeling. We provide necessary and sufficient conditions for all decision makers with different classes of payoff functions to prefer one information structure to another. We also provide conditions under which two decision-makers in a given class can be ranked in terms of their marginal value for information and hence information demand. Applications and examples are given.

JEL Classification: C44, C60, D81.

Keywords: Bayesian decision problems, value of information, stochastic dominance, stochastic orderings, decision-making under uncertainty.

*We would like to thank Scott Ashworth, Kyle Bagwell, Dirk Bergemann, Glenn Ellison, John Geanakoplos, Bengt Holmström, Eric Lehmann, Richard Levin, Meg Meyer, Paul Milgrom, Stavros Panageas, Ben Polak, and Lones Smith, as well as seminar participants at Harvard, Michigan, MIT, Northwestern, UCLA, Stanford, Yale, the 1998 Summer Meetings of the Econometric Society, and the Pacific Institute for Mathematical Sciences for valuable discussions. We have both benefited from the hospitality of the Cowles Foundation at Yale University. Athey acknowledges the support of NSF Grant SBR-9631760.

1 Introduction

In many economic models, a decision-maker faces uncertainty about her marginal returns to some action and obtains information about these returns. It is frequently appropriate and useful in such models to assume that the decision problem has an order structure, in the sense that the potential beliefs the agent could arrive at can be ranked from "less" to "more" optimistic, where more optimistic beliefs are beliefs that induce a higher action. Examples of such *monotone decision problems* arise in many contexts, including problems of production under uncertainty about marginal costs or about demand elasticity, financial and capital investment, auctions, contracting, adverse selection, coordination under uncertainty, and search.

In this paper, we provide definitions of "more information" that are tailored to different classes of monotone decision problems. Every agent faced with such a problem will prefer one information structure to another if and only if they can be ranked according to our conditions. We also provide conditions under which the incentives of two agents to acquire better information in such an environment can be ranked. As an illustration, we show that a monopolist has a lower demand for marginal cost information than a social planner.

The stochastic environment we consider is composed of two real-valued random variables: an unknown state of the world W and a signal X. The decision-maker has a prior belief about the distribution of W. She observes the realization of X, say x, and chooses her action to maximize her expected payoff. Her payoff $u(\omega, a)$ depends on her action a and ω, the realization of W. This decision problem is *monotone* if observing a higher signal realization induces a higher action.

A *family* of decision-makers is defined by *(i)* a set of possible priors, and *(ii)* a set of possible payoff functions. We consider sets of payoff functions that are alike in how the incremental returns to higher actions change with ω. An example is the set of *supermodular* payoff functions, for which the returns to increasing the action are nondecreasing in ω. A signal X leads to a monotone decision problem for each agent in the family if the posterior beliefs about W induced by observing X can be ranked in an appropriate stochastic order. The stochastic order is chosen so that higher posterior beliefs (in the given order) induce all agents in the family to choose higher actions. For the family of supermodular payoff functions with a given prior, the requires that the posterior beliefs induced by X can be ordered by first-order stochastic dominance (FOSD).

Now, consider two signals, X and X', that lead to monotone decision problems for each agent in a given family of decision-makers. We look for conditions under which, for all decision-makers in the family, observing X' is more valuable than observing X. Our answer, stated roughly, is that X' is more informative than X for these decision-makers if "on average" the posteriors induced by high realizations of X' are "higher" than the posteriors induced by high realizations of X, and conversely for low realizations. Here, posteriors are higher or lower in the appropriate stochastic order, e.g. FOSD for the family with supermodular payoff functions. Intuitively, better information allows for a more accurate match between beliefs and actions.

Our second result concerns relative demands for information. Extending the techniques from our first result, we provide conditions under which two payoff functions u and v can be ranked in terms of their marginal value for better information, and show how this result can be used to obtain comparative statics results concerning the demand for information.

Our results extend a line of inquiry begun by Lehmann (1988) and continued by Persico (2000). Lehmann and Persico study decision problems in which the payoff function has a single-crossing property and posterior beliefs have the monotone likelihood ratio property. Problems of this form were first analyzed by Karlin and Rubin (1956) and are a special case of what we refer to as monotone decision problems. Lehmann (1988) obtains an elegant informativeness ordering for this particular subset of monotone problems. His analysis exploits the sign-preservation properties of distributions with a monotone likelihood ratio. To characterize informativeness for more general classes of monotone problems, we rely on a very different set of convex cone arguments. Nevertheless, our results are related. First, a by-product of our main theorem is an alternative derivation of Lehmann's order. Our derivation is less direct, but shows that his order is necessary for a higher payoff in the class of Bayesian decision problems with the single crossing property (rather than for the class of statistical hypothesis tests considered in his paper). Moreover, in Section 5, we show a close connection between his ordering for single crossing decision problems and our informativeness condition for supermodular problems. These two classes of problems play a central role in information-theoretic modeling.

Our results in Section 4 build directly on Persico's (2000) work. For the class of problems studied by Lehmann, Persico identified when the marginal value of information is higher for one payoff function rather than another. We show that his approach generalizes and provide analogous conditions to compare the marginal value of information in other types of monotone problems.

More broadly, our analysis relates to the classic work of Blackwell (1951, 1953). Blackwell showed that a signal X' is more valuable than a signal X for *every* decision problem if and only if X' is statistically sufficient for X. Our problem is different in that we seek notions of valuable information that are tailored to specific types of economic contexts, rather than a general statistical property that holds across all environments. In addition to showing how a problem's structure can be incorporated to order more pairs of signals, our results also identify the relevant consequences of statistical sufficiency for specific classes of problems, consequences that may be useful for further analysis (e.g. comparative statics on information acquisition or market equilibria).

2 Monotone Decision Problems

2.1 The Set-Up

The stochastic environment is composed two random variables: an unknown state of the world W, with typical realization $\omega \in \Omega \subset \mathbb{R}$, and a signal X with typical realization $x \in \mathcal{X} \subset \mathbb{R}$. Given a prior, $H \in \Delta(\Omega)$, the distribution of the signal induces a joint distribution over signals and states, $F : \Omega \times \mathcal{X} \to [0, 1]$. We refer to the joint distribution F as an *information structure*. We allow the random variables W, X to be either discrete or continuous.

Given the joint distribution F, it is useful to define the following marginal and conditional distributions. Let $F_X(\cdot|\omega)$ be the signal distribution conditional on $W = \omega$. Let $F_X(\cdot)$ be the marginal distribution of the signal, where $F_X(x) = \mathbb{E}_W [F_X(x|W)]$ is computed using the prior distribution on W. Let $F_W(\cdot|x)$ denote the conditional distribution of W given $X = x$. If an agent observes a realization $X = x$, her poster belief will be $F_W(\cdot|x)$. Of course, posterior beliefs must be consistent with the prior. For all $\omega \in \Omega$, $\mathbb{E}_X [F_W(\omega|X)] = H(\omega)$. We use f, f_W, and f_X to denote the corresponding probability mass functions or densities.

The decision-maker first observes a signal realization, then chooses an action $a \in A$. We take the set of actions A to be a compact subset of $\mathbb{R}$. A payoff function is a mapping $u : \Omega \times A \to \mathbb{R}$. Throughout, we assume payoff functions are continuous in a. The decision-maker acts to maximize her expected payoff. Thus, we can write the ex ante *value* of the decision problem $\langle F, u \rangle$ as:

$$V(F, u) = \mathbb{E}_X \left[\max_{a \in A} \int_\Omega u(\omega, a) dF_W(\omega|X) \right]. \tag{1}$$

We use $\alpha^*(x)$ to denote an optimal decision rule that achieves this ex ante value.

A *family of decision-makers* is defined by a pair of sets (Λ, U), where $\Lambda \subseteq \Delta(\Omega)$, and U is a set of payoff functions $u : \Omega \times A \to \mathbb{R}$. Each member of the family has a prior belief $H \in \Lambda$

and a payoff function $u \in U$. We ask when a signal X' is preferred to another signal X for all decision-makers in a given family (Λ, U).

Observe that expanding either the set of allowable priors or the set of allowable payoff functions will lead to a larger family of decision-makers and hence a more restrictive ordering on signals. For this reason, we initially take the set of prior beliefs be an arbitrary singleton, H, and analyze orderings over signals taking the prior as given. This allows us to focus on the effects of considering different sets of payoff functions. It is also natural in economic contexts where agents have objective information about the prior distribution (i.e., the distribution of worker abilities in a population is known, but not the ability of the worker being screened for a job). Indeed, economic theory typically models agents as Bayesian decision-makers with knowledge of the environment (or a common prior among players in a Bayesian game). For such applications, standard practice is to begin by fixing the prior, and then specify general properties of the payoff function, as motivated by the economic problem. Our approach is tailored to this type of modeling strategy.

An alternative approach is to let Λ be the set of all priors. This "classical" perspective is useful in separating the properties of the signal from the context in which it will be used; however, because any pair of states of the world could potentially be the only relevant ones, this approach is sensitive to small perturbations. If the states affected by a perturbation would be in the tails of all relevant prior distributions, this approach would be "too sensitive."[1] In Section 5.2, we analyze conditions under which enlarging the set of priors leads to more restrictive orderings.

2.2 Monotone Decision Problems

A decision problem is monotone if it has an optimal decision rule $\alpha^*(x)$ that is monotone. We characterize monotonicity in terms of a relationship between the payoff function's incremental return to higher action and posterior beliefs. Section 2.4 provides examples.

Let $\mathcal{R} = \{g \mid g : \Omega \to \mathbb{R},\ g \text{ bounded, measurable}\}$.

Definition 1 *Given $R \subset \mathcal{R}$, a payoff function u has R-incremental returns if for any $a' > a$, the incremental return function $r(\omega) = u(\omega, a') - u(\omega, a) \in R$.*

Using this definition, we can associate to any set $R \subset \mathcal{R}$, a set U^R of payoff functions with R-incrememtal returns. For example, let R be the set of nondecreasing functions. Then

[1] Indeed, this concern has motivated work in statistics on "approximate sufficiency"; see Le Cam (1964).

U^R is the set of supermodular payoff functions $u(\omega, a)$ (subject to the additional restriction that for all $a' > a$, $u(\cdot, a') - u(\cdot, a)$ be a bounded and measurable function of ω).

In addition to defining a set of payoff functions, a set $R \subset \mathcal{R}$ induces a stochastic order on $\Delta(\Omega)$ (e.g. on posterior beliefs) as follows. For $P, Q \in \Delta(\Omega)$, write $Q \succ_R P$ if

$$\forall r \in R: \quad \int_\Omega r(\omega) dP(\omega) \geq 0 \Rightarrow \int_\Omega r(\omega) dQ(\omega) \geq 0.$$

The stochastic order $\succ_R$ represents a notion of *single crossing.* A function $g : \mathbb{R} \to \mathbb{R}$ satisfies single crossing if $g(x) \geq 0$ implies $g(x') \geq 0$ for all $x' > x$. Single crossing is an important property, since it is necessary and sufficient for comparative statics predictions in many contexts.[2] The order $\succ_R$ is weaker than standard stochastic dominance with respect to the set R. Stochastic dominance means that for all $r \in R$, $\int r dQ \geq \int r dP$. In some interesting examples, however, $\succ_R$ is in fact equivalent to stochastic dominance with respect to R (see Athey (1998) or Lemma 2, below). For instance, if R is the set of nondecreasing functions, $Q \succ_R P$ means that Q dominates P in the sense of first order stochastic dominance.

We use the $\succ_R$ order to develop a concept of monotonicity for information structures.

Definition 2 *Given $R \subset \mathcal{R}$, an information structure F is R-ordered if $\succ_R$ is a complete order on $\{F_W(\cdot|x)\}_{x \in \mathcal{X}}$, that is, for any $x' > x$, $F_W(\cdot|x') \succ_R F_W(\cdot|x)$.*

Our first result show that if a decision-maker with payoff function $u \in U^R$ is faced with an R-ordered information structure, her decision problem is monotone. That is, she has an optimal decision rule that is monotone.

Lemma 1 *Given $R \subset \mathcal{R}$, if u has R-incremental returns, and F is R-ordered, then there exists a function $\alpha^*(x) \in \arg\max \int_\Omega u(\omega, a) dF_W(\omega|x)$ that is nondecreasing in x.*

Proof. Define $U(x, a) = \int_\Omega u(\omega, a) dF_W(\omega|x)$. Then if $a' > a$, by the R-order, $U(x, a') - U(x, a) \geq 0$ implies $U(x', a') - U(x', a) \geq 0$ for any $x' > x$. This implies Shannon's (1995) weak single crossing property, from which the result follows. *Q.E.D.*

Now consider a family of decision-makers (Λ, U^R), for some set of payoff functions U^R. We can associate with this family a *class of monotone decision problems* by considering all signals X with the property that for any decision-maker $(H, u) \in (\Lambda, U^R)$, their decision problem given the signal X is monotone. Until Section 5.2 we fix $\Lambda = \{H\}$, and refer to the

[2] There are variety of alternative (i.e., weak and strong) notions of single crossing; see Milgrom-Shannon (1994) or Shannon (1995).

family of decision-makers by the set of payoff functions U^R. The induced class of monotone decision problems then admits all signals X where the corresponding information structure F is R-ordered.

2.3 A Characterization Lemma

In what follows, we will make use of an alternative characterization of the stochastic order introduced above. To obtain this, we place some structure on the sets $R \subset \mathcal{R}$ used to define sets of payoff functions. We say a set of functions $R \subset \mathcal{R}$ and a probability distribution $P \in \Delta(\Omega)$ satisfy Condition C if the following holds.

Condition C R is a closed convex cone, and contains some function $\hat{r} : \Omega \to \mathbb{R}$ and its negative $-\hat{r}$, where $\int \hat{r} dP \neq 0$.

This condition is relatively mild. If R is the set of nondecreasing functions (or the set of concave functions), Condition C is satisfied with $\hat{r}, -\hat{r}$ as the constant functions: $\hat{r} \equiv 1$, $-\hat{r} \equiv -1$; other examples will be given below. For our purposes, an important implication of Condition C is that if $u \in U^R$, then adding a benefit $K_a \hat{r}(\omega)$ $(K_a \in \mathbb{R})$ to each action $a \in A$ results in a new payoff function that is also in U^R. We will use this below to establish that our order on information structures is not just sufficient, but necessary, for all decision-makers in a given class to prefer one signal to another.

When Condition C is satisfied, the stochastic order $\succ_R$ admits an alternative representation.[3]

Lemma 2 *Suppose $R \subset \mathcal{R}$ and $P, Q \in \Delta(\Omega)$, and (R, P) satisfy Condition C. Then $Q \succ_R P$ if and only if*

$$\forall r \in R, \quad \int_\Omega r dQ \geq \lambda \int_\Omega r dP \tag{2}$$

for some $\lambda \geq 0$, where $\lambda = \left(\int \hat{r} dQ\right) / \left(\int \hat{r} dP\right)$.

Proof. Suppose $Q \succ_R P$, and consider the problem of choosing $r \in R$ to minimize $\int r dQ$ subject to the constraint that $\int r dP \geq 0$. The minimized value must be nonnegative. Moreover, there is some $\lambda \geq 0$ such that this linear program is equivalent to choosing $r \in R$

[3] This result is proved in Athey (1998); the approach builds on Jewitt (1986) and Gollier and Kimball (1995). Here we provide an alternative proof for completeness, since we will rely on the Lemma throughout the paper.

to minimize $\int rdQ - \lambda \int rdP$. Thus, $Q \succ_R P$ implies $\int rdQ \geq \lambda \int rdP$ for all $r \in R$. But if $\hat{r}, -\hat{r} \in R$ (the existence of such an $\hat{r}$ is ensured by Condition C), then we must have $\lambda = \left(\int \hat{r} dQ\right) / \left(\int \hat{r} dP\right)$. The other direction is immediate. *Q.E.D.*

Lemma 2 can be used to relate the $\succ_R$ order to standard stochastic dominance. If R contains the constant functions (i.e. we can choose $\hat{r} \equiv 1$, and $-\hat{r} \equiv -1$), then $\lambda = 1$. Consequently, $\succ_R$ coincides exactly with standard stochastic dominance.

2.4 Examples

We now describe two sets of payoff functions, and the induced classes of monotone decision problems, captured by our framework. More examples will be introduced in Section 3.3.

1. *Payoff functions with nondecreasing incremental returns.* A payoff function $u(\omega, a)$ with nondecreasing incremental returns is *supermodular* in (ω, a). The property that the marginal returns to action are nondecreasing in some unknown variable is a common feature of economic problems (see Milgrom and Roberts (1990), Topkis (1998) and Cooper (1999) for many examples). A classic example is a firm choosing an output plan subject to uncertainty about its marginal cost. If ω parametrizes marginal costs, so that $C_q(q, \omega)$ is decreasing in ω, the firm's profits $P(q)q - C(q, \omega)$ will be supermodular in the action q and the unknown parameter ω. A related example is a competitive firm that must choose its output under uncertainty about the eventual market clearing price p. Its profits $pq - C(q)$ are also supermodular in the action q and the unknown parameter p.

The relevant stochastic order for such problems is first order stochastic dominance. To see this, observe that the relevant set of functions R_{ND} is the set of nondecreasing functions, which contain $\hat{r} = 1$. In applying Lemma 2 to characterize $\succ_{R_{ND}}$, we have $\lambda = 1$, so $\succ_{R_{ND}}$ is exactly FOSD. If posterior beliefs are ordered by FOSD, higher signal realizations will lead to higher actions when u is supermodular.

2. *Payoff functions with concave incremental returns.* Consider a risk-averse agent who can choose between two projects, one with a safe return (normalized to 0) and one with return ω. The agent's payoff is $u(\omega, a) = v(a\omega)$, where v is concave, and the incremental returns to investment, $v(\omega) - v(0)$, are concave in ω. For problems with concave incremental returns, the relevant stochastic order is second order stochastic dominance (Rothschild and Stiglitz, 1970). The agent will be more likely to invest if she believes the return is less risky. In our

formal notation, if we let R_{CV} denote the set of concave functions, then Lemma 2 (with $\hat{r} \equiv 1$ so $\lambda = 1$) shows that $\succ_{R_{CV}}$ is SOSD.

We introduce several more classes of problems below. Of these, one set of payoff functions captures a fairly general formulation of the single-dimensional portfolio problem. Another class is the single-crossing payoff functions studied by Lehmann (1988) and Persico (2000). A third class we consider arise in first-price auctions and lie between supermodular and single-crossing. The theory can also be applied to sets of payoff functions whose incremental returns satisfy tighter properties. For example, if R is the set of affine functions, $\succ_R$ compares distribution means, while if R contains quadratic functions, $\succ_R$ is a mean-variance order (and could apply to investment problems with mean-variance preferences over gambles).

3 Monotone Information Orders

We now derive an exact condition under which all decision-makers in a family $\left(\{H\}, U^R\right)$, will prefer a signal X' to another signal X given that the corresponding information structures F' and F are R-ordered. That is, we identify the statistical condition that is equivalent to having $V(F', u) \geq V(F, u)$ for every payoff function u with R-incremental returns.

3.1 The Main Result

The key step that allows us to compare information structures is to identify each signal realization with a real number between zero and one. Given a prior H and a class of functions $R \subset \mathcal{R}$ satisfying Condition C, together with a signal X, define:[4]

$$T(x) = \frac{\int_\Omega \hat{r}(\omega) \Pr\left[X \leq x \mid \omega\right] dH(\omega)}{\int_\Omega \hat{r}(\omega) dH(\omega)}. \tag{3}$$

If the information structure F is R-ordered, then T is a nondecreasing function mapping $\mathcal{X} \to [0, 1]$. An indexing function T' can be similarly defined for an alternative information structure F'.

In general, T (and T') need not be continuous or strictly increasing. However, Lehmann (1988) shows that for any signal X, it is possible to define information-equivalent random variable X^* with associated information structure F^* such that the corresponding T^* *is*

[4] If there is more than one $\hat{r}$ satisfying Condition C, simply choose one. The choice will not affect the results that follow.

continuous and strictly increasing.[5] Thus we will treat T and T' as continuous and strictly increasing in what follows. Note that so long as T, T' are strictly increasing, one can think of $T(X)$ and $T'(X')$ as random variables with identical information content to X and X' — i.e. as re-scaled versions of the original signals.

To illustrate the indexing, observe that if R contains the constant functions, $T(x) \equiv F_X(x)$ — signal realizations are identified with their ex ante percentile. More generally, if R contains a function $\hat{r}(\omega) = \mathbf{1}_{\{\omega\in\Phi\}}$ and its negative for some measurable set Φ, then

$$T(x) = \Pr\left(X \leq x \mid W \in \Phi\right).$$

We require one technical condition to establish the *necessity* part of our main result. Condition Z is always satisfied when we can choose $\hat{r} \equiv 1$, and appears to mild in other cases as well.

Condition Z Given $R \subset \mathcal{R}$ and a prior H that satisfy Condition C, the information structure F satisfies Condition Z if for all $x \in \mathcal{X}$, $\mathbb{E}[\hat{r}(W)|X = x] \neq 0$.

Theorem 1 *Consider $R \subset \mathcal{R}$ and a prior H such that (R, H) satisfy Condition C. Let X and X' be signals with corresponding information structures F and F' that are R-ordered, and suppose that F' satisfies Condition Z. Then $V(F', u) \geq V(F, u)$ for all $u \in U^R$ if and only if*

$$F'_W(\cdot|T'(X') > z) \succ_R F_W(\cdot|T(X) > z) \qquad \textit{for all } z \in [0,1]. \tag{MIO}$$

If the conditions of the Theorem are met, we say that X' is more informative than X for all decision-makers in the family $\left(\{H\}, U^R\right)$, or alternatively that $F' \succ_{MIO-R} F$.

The (MIO) condition compares averages of posterior beliefs, where the averages are computed according to our indexing function. It says that high posteriors under F' are, on average, higher than high posteriors under F — where "high" refers to the stochastic order $\succ_R$ induced by R. Because the average of all posteriors is the prior distribution H,

$$\Pr(T(X) \leq z)\, F_W(\cdot|T(X) \leq z) + \Pr(T(X) > z)\, F_W(\cdot|T(X) > z) \;=\; H(\cdot). \tag{4}$$

Using this, it is straightforward to show that (MIO) is equivalent to:[6]

$$F'_W(\cdot|T'(X') \leq z) \prec_R F_W(\cdot|T(X) \leq z) \qquad \text{for all } z \in [0,1], \tag{5}$$

[5] To see an example, suppose that X takes on only two realizations, 0 and 1. One can define an information-equivalent X^* that is uniformly distributed on $[0, 1/2]$ when $X = 0$, and on $[1/2, 1]$ when $X = 1$.

[6] When $\mathbb{E}[\hat{r}(W)|T(X) > z] \neq 0$, this can be shown using the characterization from Lemma 2 for (MIO), and applying (4). Otherwise, we use the fact that Lemma 2 holds when $\int \hat{r}dP = 0$ with the modification that there is no simple closed form for λ.

in words, that low posteriors are on average lower under F' than under F. Put simply, *informativeness corresponds to posterior beliefs being more spread out in a given stochastic order.*

Proof of Theorem 1. We proceed in three steps. First, we use Lemma 2 to obtain an alternative characterization of (MIO). Second, we show that if (MIO) holds, and $u \in U^R$, then for any monotone policy based on observing X, there is an alternative policy based on observing X' that gives an expected payoff at least as high. This establishes that $V(F', u) \geq V(F, u)$ for all $u \in U^R$. Finally, we show that if $F' \not\succ_{MIO-R} F$, there will be some $u \in U^R$ such that $V(F', u) < V(F, u)$.

STEP 1: Fix $z \in [0, 1]$. Suppose for the moment that $\int \hat{r}(\omega) dF_W(\omega|T(X) > z) \neq 0$. By Lemma 2, (MIO) is equivalent to

$$\begin{aligned} \forall r \quad \in \quad R : \int_\Omega r(\omega) dF_W(\omega|T(X) > z) &\geq \lambda_z \int_\Omega r(\omega) dF_W(\omega|T'(X') > z), \\ \text{where } \lambda_z = \frac{\int \hat{r}(\omega) dF'_W(\omega|T'(X') > z)}{\int \hat{r}(\omega) dF_W(\omega|T(X) > z)} &= \frac{\Pr(T(X) > z)}{\Pr(T'(X') > z)}. \end{aligned}$$

The second equality goes beyond Lemma 2 and requires two steps. First, observe that the transformations T and T' are defined so that:

$$\int \hat{r}(\omega) dF_W(\omega|T(X) \leq z) \cdot \Pr(T(X) \leq z) = \int \hat{r}(\omega) dF_W(\omega|T'(X') \leq z) \cdot \Pr(T'(X') \leq z). \tag{6}$$

Second, the identity (4) implies that (6) also holds when the inequalities in the conditioning statements are reversed. Combining these two observations gives the second characterization of λ_z. Note that in the event that $\hat{r} \equiv 1$, we have $\lambda_z = 1$.

In the case where $\int \hat{r}(\omega) dF_W(\omega|T(X) \leq z) = 0$, Condition C implies that $\int \hat{r}(\omega) dF_W(\omega|T(X) > z) \neq 0$. We can use essentially the same argument so long as we replace (MIO) with the alternative representation (5). We conclude that (MIO) holds if and only if for all $z \in [0, 1]$ and $r \in R$,

$$\mathbb{E}_W \left[r(W) \mid T'(X') > z \right] \Pr \left(T'(X') > z \right) \geq \mathbb{E}_W \left[r(W) \mid T(X) > z \right] \Pr \left(T(X) > z \right). \tag{7}$$

STEP 2: This step uses the characterization of Step 1 to establish sufficiency. Suppose F, F' are R-ordered and that (MIO) holds. And suppose $A = \{a_1, ..., a_n\}$ is finite. Consider an arbitrary payoff function u with R-incremental returns. Note that we can write

$$u(\omega, a_k) = u(\omega, a_1) + \sum_{i=2}^{k} r_i(\omega),$$

where $r_i(\omega) = u(\omega, a_i) - u(\omega, a_{i-1}) \in R$.

Consider an arbitrary monotone policy $\alpha : \mathcal{X} \to A$ for use with F. It is defined by a set of "cut points": $x_1 \leq x_2 \leq ... \leq x_{n+1}$, with $\alpha(x) = a_i$ when $x_i < x < x_{i+1}$, and $\alpha(x) = \max_{j:x_j=x} a_j$ when $x = x_i$ for some i. The cut points also can be represented by their indices $\{z_i\}_{i=1}^{n+1}$ where $z_i \equiv T(x_i)$. The ex-ante payoff using this policy with F is:

$$\begin{aligned} V(F, u, \alpha) &= \int_{\mathcal{X}} \int_{\Omega} u(\omega, \alpha(x)) dF_W(\omega, x) \\ &= \mathbb{E}_W [u(W, a_1)] + \sum_{i=2}^{n} \mathbb{E}_W [r_i(W) \mid T(X) > z_i] \Pr(T(X) > z_i). \end{aligned} \tag{8}$$

We now construct an alternative policy α' for use with F' that does better than the policy α used in conjunction with F. This new policy is given by "cut points" $\{x'_i\}$, where we define $T'(x'_i) = T(x_i) = z_i$ for all $i = 1, ..., n+1$. Then, e.g., $\alpha'(x) = a_i$ when $x'_i < x < x'_{i+1}$. The payoff to this new policy under F' is:

$$V(F', u, \alpha') = \mathbb{E}_W [u(W, a_1)] + \sum_{i=2}^{n} \mathbb{E}_W \left[r_i(W) \mid T'(X') > z_i\right] \Pr(T'(X') > z_i). \tag{9}$$

It follows immediately from (7) that if (MIO) holds, then $V(F', u, \alpha') \geq V(F, u, \alpha)$. Since there is a monotone policy under F that is optimal, we are done.

The case where A is compact follows from a limiting argument. Any monotone policy $\alpha(x)$ used under F can be approximated by a sequence of step functions $\alpha^1(x), \alpha^2(x), ...$ converging to $\alpha(x)$, and for each step function, we can construct a policy $\alpha^{k\prime}(x)$ for use with F' such that $V(F', u, \alpha^{k\prime}) \geq V(F, u, \alpha^k)$. Moreover, $\alpha^{k\prime}(x) \to \alpha'(x)$, for some monotone policy $\alpha'(x)$. Since u is continuous in a, it follows that $V(F', u, \alpha') \geq V(F, u, \alpha)$.

STEP 3: This step establishes necessity. Suppose F, F' are R-ordered, and (MIO) fails. Then, using the characterization from Step 1, there exists some $r \in R$ and $z \in [0, 1]$ such that:

$$\mathbb{E}_W \left[r(W) \mid T'(X') > z\right] \Pr\left(T'(X') > z\right) \; < \; \mathbb{E}_W \left[r(W) \mid T(X) > z\right] \Pr\left(T(X) > z\right).$$

Now define a payoff function $u_{rz} \in U^R$ as follows:

$$u_{rz}(\omega, a) = \begin{cases} 0 & \text{if } a < \overline{a} \\ r(\omega) - K_{rz}\hat{r}(\omega) & \text{if } a \geq \overline{a} \end{cases}$$

where $\overline{a}$ is some arbitrary element of A, and $K_{rz} = \mathbb{E}\left[r(\omega) \mid T'(X') = z\right] / \mathbb{E}\left[\hat{r}(\omega) \mid T'(X') = z\right]$ is a fixed constant. This is the (only) place where we need Condition Z: by Condition Z, the denominator is nonzero.

Consider the decision problem $\langle u_{rz}, F' \rangle$. The incremental returns to choosing $a \geq \overline{a}$ rather than $a < \overline{a}$ given a signal realization x are equal to:

$$\mathbb{E}\left[r(\omega) - K_{rz}\hat{r}(\omega) \mid T'(X') = T'(x)\right].$$

As a function of x, these incremental returns start negative and cross zero once (because F' is R-ordered) at the signal realization x such that $T'(x) = z$. So the optimal policy is to choose $a \geq \overline{a}$ if and only if $T'(X') > z$. Denote this policy by $\alpha' : \mathcal{X} \to A$. Thus,

$$V(F', u_{rz}) = \mathbb{E}_W[r(W) - K_{rz}\hat{r}(W) | T'(X') > z] \Pr(T'(X') > z).$$

From above, we know that:

$$\mathbb{E}_W[\hat{r}(W) \mid T(X) > z] \Pr(T(X) > z) = \mathbb{E}_W[\hat{r}(W) \mid T'(X') > z] \Pr(T'(X') > z)$$

Combining these two facts with the failure of (MIO) at r, z, we have

$$\begin{aligned} V(F, u_{rz}, \alpha') &= \mathbb{E}_W[r(W) - K_{rz}\hat{r}(W) | T(X) > z] \Pr(T(X) > z). \\ &> \mathbb{E}_W[r(W) - K_{rz}\hat{r}(W) | T'(X') > z] \Pr(T'(X') > z) = V(F', u_{rz}). \end{aligned}$$

Since α' may not be an optimal policy under information structure F, $V(F, u_{rz}) \geq V(F, u_{rz}; \alpha)$. We conclude that F' cannot be more valuable than F for all decision-makers with R-incremental returns. *Q.E.D.*

3.2 Discussion

There are three steps to establishing the result. The first step is to link our *one-dimensional* stochastic order over posterior beliefs to a *two-dimensional* stochastic order over the joint distributions of the state W and the indexed signals $T(X)$, and $T(X')$. Our proof shows that $F' \succ_{MIO-R} F$, then for all $r \in R$, and $z \in [0, 1]$:

$$\int_{\mathcal{X}} \int_{\Omega} \mathbf{1}_{\{T'(x) > z\}} r(\omega) dF'(\omega, x) \geq \int_{\mathcal{X}} \int_{\Omega} \mathbf{1}_{\{T(x) > z\}} r(\omega) dF(\omega, x). \tag{10}$$

That is, integrating any $r \in R$ with respect to the joint distribution F' over the upper rectangle defined by $T'(x) > z$, gives a higher total than integrating r with respect to F over the rectangle defined by $T(x) > z$. This step is fairly subtle when $\hat{r}$ is not a constant function.

The second step is to show that if F' and F are ranked in this two-dimensional stochastic order, then all decision-makers with R-incremental returns will prefer F' to F. To understand this, consider a binary choice problem where $a \in \{0,1\}$, and $u(\omega, a) = ar(\omega)$. An optimal policy under F will simply be a cut-point x_0. The value function for the problem will be the integral of r with respect to F over the rectangle defined by $x > x_0$ (equivalently $T(x) > T(x_0)$). If (10) holds, this same decision-maker can do *at least as well* with information F' by choosing $a = 1$ whenever $X' > x'_0$, where $T(x_0) = T'(x'_0)$. (Here, we understand T and T' as the indices that make policies comparable across signals.) Equation (10) implies that the expected payoff from this is greater than the problem's value under F. The only remaining issue is to show that even with richer action spaces and utility functions, monotonicity allows the problem to be reduced to a series of simple binary comparisons.

The final step is to show that if F', F cannot be compared under the two-dimensional order (10), there will be at least one decision-maker with R-incremental returns who prefers F. The insight here is that the space spanned by value functions of the form $v(\omega, x) = u(\omega, \alpha(x))$, where $\alpha(\cdot)$ is an optimal policy under F' is large enough relative to the space of functions spanned by functions of the form $\mathbf{1}_{\{T'(x)>z\}} r(\omega)$. What we show is that that the space spanned by value functions $v(\omega, x)$ contains the space spanned by functions of the form $\mathbf{1}_{\{T'(x)>z\}} \{r(\omega) + K_{rz}\hat{r}(\omega)\}$ for particular constants K_{rz}. This allows us to show that if (10) fails, there will some decision-maker who can do at least as well using F as she does using F'.

It is important to observe the crucial role played by the indexing transformations T,T' at each stage. As we observed above, there is no gain or loss in information content to re-scaling, or indexing signal realizations. The origin of our particular indices lies in the first step of the proof. Since $\hat{r}$ *and* $-\hat{r}$ are in R, clearly (10) must hold with equality for $r = \hat{r}$. We define the indexes T, T' to make this so. Only then can we reduce the two-dimensional stochastic order for comparing ex ante values of decision problems to a one-dimensional stochastic order over posterior beliefs.

It is worth commenting briefly on the relationship of our result to Blackwell's comparison of information structures. Blackwell's approach also starts by deriving a stochastic ordering over posterior beliefs. In particular, he observes that for any payoff function u, the function $\bar{u}(P) = \max_{a \in A} \int_\Omega u(\omega, a) dP(\omega)$ mapping $\Delta(\Omega) \to \mathbb{R}$ will be convex in the decision-maker's belief P.[7] Defining the value of a decision problem $V(F, u)$ as above, the problem of comparing X and X' can be framed as a problem of comparing the distributions over posteriors (denoted

[7] Convexity requires $\max_{a \in A} \int_\Omega u(\omega, a) d(\gamma P + (1-\gamma)Q) \leq \gamma \max_{a \in A} \int_\Omega u(\omega, a) dP + (1-\gamma) \max_{a \in A} \int_\Omega u(\omega, a) dQ$, which follows by linearity of the integral and because the agent can do better by optimizing for each posterior rather than choosing the optimal action for the convex combination of posteriors.

$\mu(\cdot)$ and $\mu'(\cdot)$) generated by the two signals. Specifically, if $\int \bar{u}(P)d\mu'(P) \geq \int \bar{u}(P)d\mu(P)$ for all convex functions $\bar{u} : \Delta(\Omega) \to \mathbb{R}$, then any agent will prefer X' to X. The former condition is that μ' dominates μ according to the "convex stochastic dominance order." The last step of Blackwell's analysis is to characterize this stochastic dominance order (which is typically quite complicated) in terms of the signals X' and X, showing that it is equivalent to requiring that X be a "garbling" of X'.

Our approach is related in that we use the consequences of optimality (here, monotone policies) together with properties of the payoff function to derive a (somewhat complicated) ordering over the joint distribution of signals and states. We are then able to reduce this to a one-dimensional ordering over posterior beliefs that it is far less complex than the multi- (or infinite-) dimensional convex stochastic dominance order. We return to further characterizations of the order over posteriors (analogous to Blackwell's final step) in Section 5.

3.3 Examples

1. *Payoff functions with nondecreasing incremental returns.* If a decision-maker's incremental returns are nondecreasing in the state variable, she wants to match high actions with high beliefs about W, and low actions with low beliefs. All such decision-makers with prior H will prefer a signal X' to another X, given that the induced posterior beliefs can be ordered by FOSD, if for all $z \in [0,1]$, $F'_W(\cdot|F'_X(X') > z)$ dominates $F_W(\cdot|F_X(X) > z)$ by FOSD—that is, if high realizations of X' induce, on average, higher beliefs (in the sense of FOSD) than high realizations of X.[8] Intuitively, this allows for a better match between actions and beliefs.

The following example illustrates this idea. Fix a prior $H \in \Delta(\Omega)$ with density h and define, for $\theta \in [0, \overline{\theta}]$, a signal X^θ with support $[0,1]$ such that (W, X^θ) have joint density:

$$f^\theta(\omega, x) = h(\omega) + \theta k(\omega) l(x),$$

where $k : \Omega \to \mathbb{R}$ and $l : [0,1] \to \mathbb{R}$ are bounded increasing functions with $\int_\Omega k(\omega)d\omega = \int_0^1 l(x)dx = 0$, and $f^\theta \geq 0$. Note that this example is constructed so that $T^\theta(x) = F_X^\theta(x) = x$. Further, for any θ, the posteriors of $W|X^\theta = x$ are ordered by FOSD (typically, they will not have the MLRP). Finally, an increase in θ makes X^θ more informative for supermodular

[8] It can further be shown that in this case, (MIO) is equivalent to requiring that for all $(\omega, x) \in (\Omega, \mathcal{X})$, $F'(\omega, F'_X(x)) \geq F(\omega, F_X(x))$. This, in turn, is equivalent to the supermodular stochastic order (see Meyer (1990) and Shaked and Shantikumar (1997)) for the comparison between $(W, F'_X(X'))$ and $(W, F_X(X))$.

decision-makers. However, it will generally not be the case that $X^{\theta'}$ is sufficient for X^θ when $\theta' > \theta$, or even that $X^{\theta'}$ dominates X^θ in Lehmann's order (defined below in (13)). To better understand this example, observe that for high (low) values of x, $l(x) > (<)0$, and an increase in θ puts more positive (negative) weight on the nondecreasing function k, thus shifting probability weight towards (away from) high values of ω.

Below, in Section 5, we provide further characterizations of this information order in terms of "marginal-preserving spreads."

2. *Payoff functions with concave incremental returns.* A decision-maker with incremental returns that are concave as a function of the state variable wants to match high actions with less risky beliefs. So X' is more informative than X under *(MIO)* if, for every $z \in [0, 1]$, $F'_W(\cdot|F'_X(X') > z)$ dominates $F_W(\cdot|F_X(X) > z)$ by SOSD. That is, high realizations of the signal lead, on average, to less risky posteriors under F' than F, and low realizations of the signal lead to more risky posteriors. Again, (MIO) ensures a better match between actions and belief. In Section 5, we characterize this information order in terms of "mean and marginal-preserving spreads."

3. *Payoff functions with incremental returns that cross zero at $\omega = \omega_0$.* Consider the standard portfolio problem where an agent with increasing utility over wealth $v(\cdot)$ and endowment e must choose between a risky asset with return ω and a safe asset with return ω_0. Letting a denote the invesment level, the decision-maker's payoff is $u(\omega, a) = v(a\omega + (e - a)\omega_0)$. Her incremental returns to investing are negative for $\omega < \omega_0$, and positive for $\omega > \omega_0$. Payoff functions with this property also arise in more general investment problems (Athey, 2002).

To study this class of problems, we consider a series of sets $R^n_{SC(\omega_0)} = \{r : \Omega \to \mathbb{R} : r(\omega) \leq 0 \text{ if } \omega < \omega_0 - \frac{1}{n} \text{ and } r(\omega) \geq 0 \text{ if } \omega > \omega_0 + \frac{1}{n}\}$. In the limiting case, we identify the set $R_{SC(\omega_0)} = \{r : \Omega \to \mathbb{R} : r(\omega) \leq 0 \text{ if } \omega < \omega_0 \text{ and } r(\omega) \geq 0 \text{ if } \omega > \omega_0\}$.[9] The stochastic order $\succ_{R_{SC(\omega_0)}}$ for this set is a weakening of the monotone likelihood ratio property (MLRP). Recall that a conditional density f_W has the MLRP if for $x' > x$, $f_W(\cdot|x')/f_W(\cdot|x)$ is a nondecreasing function. Specifically, if $f_W(\omega_0|x) > 0$ for all x, then $F_W(\cdot|x') \succ_R F_W(\cdot|x)$ if $f_W(\cdot|x') - [f_W(\omega_0|x')/f_W(\omega_0|x)]\ f_W(\cdot|x)$ is single crossing.[10] A more familiar statement, if

[9] The reason for considering $R^n_{SC(\omega_0)}$ rather than $R_{SC(\omega_0)}$ directly is that $\hat{r}^n = \mathbf{1}_{\{\omega \in [\omega_0 - 1/n, \omega_0 + 1/n]\}}$ will be positive on a set of positive H-measure so long as $h(\omega_0) > 0$ and h is continuous at ω_0.

[10] This follows from applying Lemma 2 to each $R^n_{SC(\omega_0)}$. It may seem restrictive to assume $f(\omega_0|x) > 0$. However, to satisfy Condition C requires $h(\omega_0) > 0$. If $f_W(\cdot|x) = 0$ for some x, and if $F_W(\cdot|x') \succ_R F_W(\cdot|x)$ for all $x' > x$, then it would have to be the case that $f_W(\cdot|x') = 0$ for all $x' > x$, and indeed for all $x' > x$, the support of $W|x'$ would have to lie entirely above ω_0. This would imply that for every $x' > x$, the incremental

$f_W(\cdot|x) > 0$ on Ω, is that $f_W(\cdot|x')/\ f_W(\cdot|x)\ -\ f_W(\omega_0|x')/f_W(\omega_0|x)$ must be single crossing as a function of ω. We consider informational comparisons just below.

4. *Payoff functions with single-crossing incremental returns.* A payoff function u has single crossing incremental returns if the incremental returns to higher action are single crossing as a function of ω. This is the class of problems studied by Lehmann (1988) and Persico (2000). For example, if a firm chooses price to maximize its expectation of $u(p,\omega) = (p-c)D(p,\omega)$, then the payoff function $u(p,\omega)$ will have single crossing incremental returns if an increase in ω reduces demand elasticity. More generally, a sufficient condition is that a strictly increasing transformation of u, such as $\ln(u)$, be supermodular; indeed, the set of log-supermodular functions induces the same order $\succ_R$ (Athey, 2002). Persico considers this class because single crossing incremental returns arise in auction models; Athey (2002) provides a number of other economic examples. Formally, we define R_{SC} by considering the union of $R_{SC(\omega_0)}$ sets. That is, $r \in R_{SC}$ is there is some ω_0 such that $r \in R_{SC(\omega_0)}$.[11] Assuming posterior densities or mass functions are positive throughout Ω, and requiring $\succ_{R_{SC(\omega_0)}}$ to hold for all ω_0, we find that the order $\succ_R$ induced by R_{SC} is the MLRP.

To make informational comparisons, consider two information structures F and F' whose posteriors are ordered by the MLRP. For the class of payoff functions whose incremental returns cross zero at a particular point ω_0, we can express (MIO) as requiring that for all $z \in [0,1]$,

$$\omega \geq \omega_0 \quad \Leftrightarrow \quad \Pr\left(F'_X(X'|\omega_0) > z \mid \omega\right) \geq \Pr\left(F_X(X|\omega_0) > z \mid \omega\right). \tag{11}$$

This condition has a nice interpretation in terms of hypothesis testing. Consider a statistical test of the hypothesis $H_0 : \omega \geq \omega_0$ against the alternative $\omega < \omega_0$. An optimal size z test will imply rejection if and only if the realization of X is greater than x, where $F_X(x|\omega_0) = z$. Thus (MIO) states that X' provides a *uniformly more powerful test* than X — for any size, it is more likely to reject when the true state is below ω_0. Lehmann shows that (MIO) is also a *necessary* condition for X' to yield a uniformly more powerful test than X for testing problems of this form. It is not obvious how this relates to Bayesian decision problems.[12] Our

returns to higher action are necessarily positive, and distinguishing between these signals is unimportant from a decision-making standpoint.

[11] Note that R_{SC} will not satisfy Condition C, just as (depending on the prior H) $R_{SC(\omega_0)}$ may not. However, it is relatively easy to show that all of our results still apply.

[12] The subtlety is that one must construct decision problems that yield as optimal policies the most powerful test for every given size.

result shows that (MIO) is actually *necessary* for X' to be preferred to X for all Bayesian decision makers with single crossing payoffs.

Re-writing (11), (MIO) requires that for all $z \in [0,1]$,

$$\omega \geq \omega_0 \quad \Leftrightarrow \quad F_X\left(F_X^{-1}(z|\omega_0) \mid \omega\right) \geq F'_X\left(F'^{-1}_X(z|\omega_0) \mid \omega\right). \tag{12}$$

For F' to be more informative than F for all payoff functions with single crossing incremental returns, this condition must hold for *all* ω_0. Applying the monotone transformation $F'^{-1}_X(\cdot|\omega)$ to each side, and substituting $x = F_X^{-1}(z|\omega_0)$, we obtain the equivalent condition that:

$$\forall x \in \mathcal{X},\ F'^{-1}_X(F_X(x|\omega)|\omega) \text{ is nondecreasing in } \omega. \tag{13}$$

This is the condition derived by Lehmann (1988).

This condition also has a convenient representation as a likelihood ratio ordering on posterior beliefs.

Corollary 1 *Suppose that F, F' are information structures with differentiable densities that satisfy the MLRP. Then $F' \succ_{MIO-SC} F$ if and only if for all $\omega \in \Omega$ and all $z \in [0,1]$,*

$$\frac{\frac{\partial}{\partial\omega} f'_W\left(\omega \mid F'_W(X') > z\right)}{f'_W\left(\omega \mid F'_W(X') > z\right)} \geq \frac{\frac{\partial}{\partial\omega} f_W\left(\omega \mid F_W(X) > z\right)}{f_W\left(\omega \mid F_W(X) > z\right)}. \tag{14}$$

This basically says that the high posteriors under F' dominate the high posteriors under F in the sense of having higher local likelihood ratios. The proof of this result is in the Appendix.

5. *Payoff functions with that are nondecreasing on $\omega \leq \omega_0$, and zero thereafter.* Payoff functions that take this form arise in the study of first-price auctions with affiliated values, where W represents the opponent's signal.[13] Formally, define $R_{NDZ(\omega_0)}$ as follows: $r \in R_{NDZ(\omega_0)}$ if r is nondecreasing on $\omega \leq \omega_0$ and $r(\omega) = 0$ for all $\omega > \omega_0$. This class satisfies Condition C, taking $\hat{r}(\omega) \equiv \mathbf{1}_{\{\omega \leq \omega_0\}}$. The stochastic order $\succ_R$ induced by $R_{NDZ(\omega_0)}$ requires that $F_W(\cdot|x') \succ_R F_W(\cdot|x)$ if $F_W(\cdot|x')/\ F_W(\cdot|x) \leq F_W(\omega_0|x')/F_W(\omega_0|x)$. This is FOSD conditional on the event $\omega \leq \omega_0$. Taking the set of payoff functions $R_{NDZ} = \cup_{\omega_0 \in \Omega} R_{NDZ(\omega_0)}$

[13] In a two-bidder first-price auction, if bidder 2 uses a strictly monotone bidding strategy $\beta(\cdot)$, the payoffs to bidder 1 for a given bid b and realization of the opponent's signal are given by $u(b,\omega) = (E[V|S=s, W=\omega] - b)\mathbf{1}_{\{b>\beta(\omega)\}}$. As a function of ω, the returns to choosing b_H rather than $b_L < b_H$ are first negative (and constant) at $b_L - b_H$, then (in the region where higher bids cause bidder 1 to go from losing to winning by increasing the bid) nondecreasing in ω, and finally the returns are equal to 0 (when both b_L and b_H are losing bids). See Athey (forthcoming) for generalizations and further discussion.

induces as $\succ_R$ the "monotone probability ratio order," which requires that $F_W(\cdot|x')/\ F_W(\cdot|x)$ is nondecreasing for $x' > x$. This order is weaker than the MLRP but stronger than FOSD.[14]

The analysis of MIO for this set is analogous to that for single crossing functions. By essentially the same argument as above (first considering $R = R_{NDZ(\omega_0)}$, and then concatenating requirements), we find that (MIO) for the set R_{NDZ} requires that:

$$\forall x \in \mathcal{X},\ F_X'^{-1}(F_X(x|W \leq \omega)|W \leq \omega) \text{ is nondecreasing in } \omega. \tag{15}$$

One can also derive a condition on posterior beliefs as in Corollary 1 above.

Athey (2000) applies this order to investment problems. She shows that under a few additional regularity conditions, this informational ranking is a sufficient condition for all risk-averse investors to prefer one information structure to another.

4 Demand for Information

This section considers the *relative* demand for information of two different decision-makers. This problem has been studied by Persico (2000) for the case of decision-makers with single-crossing incremental returns. The following Theorem allows comparisons to be made for many other classes of problems.

Following Persico, we consider small changes in information content. To this end, say that for a given prior H, a family of signals $\{X^\theta\}_{\theta\in\Theta}$ is *smoothly parametrized* by θ if the mapping from $\Theta \subset \mathbb{R}$ into information structures F^θ is continuously differentiable in θ.

Theorem 2 *Consider $R \subset \mathcal{R}$ and a prior H satisfying Condition C. Let $\{X^\theta\}_{\theta\in\Theta}$ be smoothly parametrized family of signals such that for each θ, the information structure F^θ is R-ordered, and $\{F^\theta\}_{\theta\in\Theta}$ is ordered by (MIO). Define*

$$w(\omega, x) = u(\omega, \alpha^{\theta,u}(x)) - v(\omega, \alpha^{\theta,v}(x)),$$

where $\alpha^{\theta,u}$ denotes an optimal decision rule for the decision problem $\langle F^\theta, u\rangle$. If for all $x' > x$, $w(\omega, x') - w(\omega, x) \in R$, then $\frac{\partial}{\partial\theta}V(\theta, u) \geq \frac{\partial}{\partial\theta}V(\theta, v)$.

Proof. To proceed, it is convenient to work with a re-scaling of the signals. For each θ, we have a signal X^θ and corresponding information structure F^θ. Define the informationally

[14] Athey (2000) shows that R_{NDZ} induces the same order $\succ_R$ as the one induced by considering the incremental returns to investment for the class of risk-averse investors making an investment, so that $u(a,\omega) = v(\pi(a,\omega))$, where v is concave and π is supermodular. The order $\succ_{R_{NDZ}}$ is also the same as the order induced by the set of payoffs r that are single crossing and quasi-concave (Athey (forthcoming)).

equivalent signal $Z^\theta = T^\theta(X^\theta)$. Then Z^θ corresponds to an information structure G^θ, where $G^\theta(\omega, z) = F^\theta(\omega, T^{\theta^{-1}}(z))$. We will think of the decision-makers observing Z^θ rather than X^θ. Define $w(\omega, z) = u(\omega, a^{\theta,u}(T^{\theta^{-1}}(z)) - v(\omega, a^{\theta,v}(T^{\theta^{-1}}(z))$. By the Envelope Theorem,

$$\frac{\partial}{\partial\theta}V(\theta, u) - \frac{\partial}{\partial\theta}V(\theta, v) = \int_\Omega \int_{[0,1]} w(\omega, z) d\frac{\partial}{\partial\theta}\left\{G^\theta(\omega, z)\right\}.$$

Note that the support of each Z^θ is $[0, 1]$. Assuming A is finite, there will be some series of cut-points $\{z_i\}_{i=1}^n$, with $z_1 \leq z_2 \leq ... \leq z_{N+1}$, such that $w(\omega, z) = w(\omega, z_i)$ on $[z_i, z_{i+1})$. Let $r_i(\omega) = w(\omega, z_i) - w(\omega, z_{i-1})$, and note that $r_i(\omega) \in R$ for all $i = 2, ..., N$. Then, $\frac{\partial}{\partial\theta}V(\theta, u) - \frac{\partial}{\partial\theta}V(\theta, v) \geq 0$ if

$$\sum_{i=2}^n \mathbb{E}_W\left[r_i(W) \mid Z^{\theta+d\theta} > z_i\right] \Pr\left(Z^{\theta+d\theta} > z_i\right) - \sum_{i=2}^n \mathbb{E}_W\left[r_i(W) \mid Z^\theta > z_i\right] \Pr\left(Z^\theta > z_i\right) \geq 0,$$

which holds by (MIO). The case where A is continuous follows via a limiting argument. *Q.E.D.*

An immediate consequence is that under the conditions of Theorem 2, we can make a comparison between levels of costly information acquisition.

Theorem 3 *Suppose the conditions of Theorem 2 hold. For $C : \Theta \to \mathbb{R}$, let*

$$\theta^*(u) = \arg\max_{\theta\in\Theta} V(\theta, u) - C(\theta).$$

Then $\theta^*(u) \geq \theta^*(v)$ (in the strong set order).

Proof. By Theorem 2, $[V(\theta, u) - C(\theta)] - [V(\theta, v) - C(\theta)]$ is nondecreasing in θ. So by Topkis' Monotonicity Theorem (e.g. Topkis, 1998), $\theta^*(u) \geq \theta^*(u)$ in the strong set order. *Q.E.D.*

Theorem 2 provides a starting point for obtaining comparative statics of information acquisition. Applying the result requires a fair amount of structure, however, since the critical condition depends on the properties of the two objective functions evaluated at their respective optima. To some extent, we see this as unavoidable in comparing information preferences. Typically, any change in preferences has at least two effects in a decision problem under uncertainty. First, it changes the optimal behavior and hence the responsiveness to different realizations of information. Second, it changes the preferences over the residual risk faced after a decision is made. Thus, to study how the marginal value of information will adjust, one generally must deal with subtle comparisons of the curvature of the payoff functions. The details vary depending on the environment. Below we illustrate by considering the standard monopoly model.

4.1 Small Changes in Preferences

If we consider small changes in the payoff function, the requirements of Theorem 2 can be characterized in terms of primitive properties of payoff function. We parameterize $u : \Omega \times \mathcal{A} \times \Gamma \to \mathbb{R}$ by some γ, and define $\bar{u} : \mathcal{A} \times \mathcal{X} \times \Gamma \to \mathbb{R}$ by $\bar{u}(a, x; \gamma) = \int u(\omega, a; \gamma) dF_W(\omega|x)$. We let subscripts denote partial derivatives.

Corollary 2 *Suppose the conditions of Theorem 2 hold, and that (a)* $\{1, -1\} \in R$*; (b)* $\Gamma, \mathcal{A}$ *are compact, convex subsets of* $\mathbb{R}$*; (c) for each* ω*,* $u(\omega, \cdot; \cdot)$ *is* C^3*; and (d)* $\bar{u}$ *is quasi-concave in* a *and* C^2*. Then* $\frac{\partial^2}{\partial\theta\partial\gamma} V(\theta, u(\cdot; \gamma)) \geq 0$ *if, for each* (a, γ)*,* $u(\cdot, a; \gamma)$ *satisfies: (i)* $u(\cdot, \cdot; \gamma) \in U^R$*; (ii)* $u_\gamma(\cdot, a; \gamma) \in U^R$*, (iii)* $u_{aa\gamma}(\cdot, a; \gamma) \geq 0$*, and (iv) either* $u_{aa}(\cdot, a; \gamma)$ *is a constant function of* ω*, or else* $u_a(\cdot, \cdot; \gamma) \in U^R$ *and* $u_{a\gamma}(\cdot, a; \gamma) \geq 0$.

Proof. By Theorem 2, the result follows if $\frac{\partial^2}{\partial\gamma\partial x} u(\cdot, \alpha^{\theta,\gamma}(x); \gamma) \in R$. Differentiating yields:

$$\frac{\partial^2}{\partial\gamma\partial x} u(\cdot, \alpha^{\theta,\gamma}(x); \gamma) = u_{a\gamma}(\cdot, a; \gamma)\alpha_x^{\theta,\gamma}(x) + u_a(\cdot, a; \gamma)\alpha_{x\gamma}^{\theta,\gamma}(x) + u_{aa}(\cdot, a; \gamma)\alpha_x^{\theta,\gamma}(x)\alpha_\gamma^{\theta,\gamma}(x),$$

evaluated at $a = \alpha^{\theta,\gamma}(x)$. Since F^θ is R-ordered, $\alpha_x^{\theta,\gamma} \geq 0$ by (i). Thus, the first term is in R by (ii). If u_{aa} is constant in ω, the last term is constant in ω (and thus in R); otherwise, it is in R if $u_{aa} \in R$ and $\alpha_\gamma^{\theta,\gamma} \geq 0$, which follows if $u_{a\gamma} \geq 0$ (as in (iv)). The second term is in R by (i), so long as $\alpha_{x\gamma}^{\theta,\gamma} \geq 0$. To evaluate this, we apply the implicit function theorem, yielding:

$$\frac{\partial^2}{\partial x\partial\gamma} \alpha^{\theta,\gamma}(x) = \left[-\bar{u}_{aa}(\cdot)\bar{u}_{ax\gamma}(\cdot) + \bar{u}_{ax}(\cdot)\bar{u}_{aa\gamma}(\cdot)\right] \Big/ (\bar{u}_{aa}(\cdot))^2 \Big|_{a=\alpha^{\theta,\gamma}(x)}$$

At the optimum, $\bar{u}_{aa} < 0$, and $\bar{u}_{ax} \geq 0$ since F^θ is R-ordered; $\bar{u}_{aa\gamma} \geq 0$ by (iii). Since $\{1, -1\} \in R$, $F_W^\theta(\cdot|x)$ is ordered by stochastic dominance for R; then $u_{a\gamma} \in R$ implies $\bar{u}_{ax\gamma} \geq 0$. *Q.E.D.*

Each of conditions *(ii)-(iv)* has a natural interpretation and ensures that a particular effect works in the right direction. Consider an example as follows. A firm must choose an investment a, with increasing and convex cost $c(a)$, that has returns $v(\omega, a)$ parametrized by ω, where an increase in ω means higher returns to investment. Then $u(\omega, a) = v(\omega, a) - c(a)$ is supermodular in (ω, a).

- For condition *(ii),* suppose γ affects the scale of returns. Let $u(\omega, a; \gamma) = \gamma a r(\omega) - c(a)$. Conditions (iii) and (iv) are trivial. An increase in γ increases information demand because it makes the marginal returns to investment more sensitive to changes in ω (i.e. "more nondecreasing" in ω).

- For condition *(iii)*, suppose that γ multiplies the costs of investment. Let $u(\omega, a; \gamma) = ar(\omega) - (\overline{\gamma} - \gamma)c(a)$. Conditions (ii) and (iv) are trivial. Now, an increase in γ increases information demand by making the investment problem less concave in a, and hence the policy function more responsive to information about returns (i.e. $\frac{\partial}{\partial x}\alpha^{\theta,\gamma}$ is increasing in γ).

- Condition *(iv)* becomes relevant when we generalize the functional form for investment. Let $u(\omega, a; \gamma) = v(\omega, a) - (\overline{\gamma} - \gamma)c(a)$. Since c is nondecreasing, an increase in γ increases the optimal policy. In turn, this increases the demand for information if it makes marginal returns more sensitive to ω (i.e. condition (iv) requires that $v_{a\omega}$ be increasing in a). To see an application, suppose that a is an input (e.g. labor). If $u(\omega, a; \gamma) = v(\omega, a) + \gamma a - c(a)$, if $v_{a\omega}$ is increasing in a, the firm invests more (less) in information gathering in response to a wage subsidy (tax).

4.2 Application: Production under Uncertainty

A growing literature considers the value of information to firms under imperfect competition (see, e.g., Mirman, Samuelson, and Schlee (1994) and references therein). Here, we prove a simple but new result using Theorem 2: under standard conditions, a monopolist will not only produce less but will gather less information about production costs than is socially efficient.

To model this, let $P(q)$ be the inverse demand curve. Suppose the cost of producing q units is $C(q, \omega)$, where (letting subscripts denote partial derivatives) C_q is nonincreasing in ω. The monopolist's payoff is $u^M(\omega, q) = qP(q) - C(q, \omega)$, while the social planner's payoff is $u^S(\omega, q) = \int_0^q P(t)dt - C(q, \omega)$. Both payoff functions are supermodular, so consider $\{F^\theta\}_{\theta\in\Theta}$ satisfying the hypotheses of Theorem 2 for R_{ND}. The cost of information is $c(\theta)$.

By Theorem 1, both monopolist and social planner prefer more information to less according to our definition of information. To show that $\frac{\partial}{\partial\theta}V(\theta, u^S) \geq \frac{\partial}{\partial\theta}V(\theta, u^M)$, we ask when $\frac{\partial}{\partial x}u^S(\omega, q^S(x)) - \frac{\partial}{\partial x}u^M(\omega, q^M(x))$ is nondecreasing in ω. Eliminating terms that do not depend on ω, we express this difference as

$$-C_q(q^S(x), \omega)q_x^S(x) + C_q(q^M(x), \omega)q_x^M(x). \tag{16}$$

Proposition 1 *In the production problem with cost uncertainty, a social planner has a higher demand for information than a monopolist, if for any θ, (i) marginal costs are increasing in q and submodular in (q, ω), (i.e. $C_{qq} \geq 0$, $C_{qq\omega} \leq 0$), (ii) the marginal revenue curve is* *downward sloping ($qP''(q)+P'(q) \leq 0$ for all q), and (iii) the marginal social surplus function ($P(q) - \mathbb{E}[C_q(q^S, W)|X = x]$) is convex in q.*

The basic intuition for the result is easily grasped. Different realizations of X shift the perceived marginal cost curve. Under relatively standard conditions, the demand curve faced by a social planner will be flatter than the marginal revenue curve faced by a monopolist, so the social planner will be more responsive to shifts in marginal cost. Consequently, the social planner benefits more from improving the match between perceived and actual marginal costs. A formal proof obtains by using the implicit function theorem to express q_x^M and q_x^S, and checking that (16) is indeed nondecreasing in ω.

5 Further Characterizations

This section provides further characterizations of our informativeness conditions. We first characterize informativeness in terms of "marginal-preserving spreads." We then discuss the restrictions implied by our criteria on the conditional signal distributions when we allow for sets of prior beliefs.

5.1 Marginal-Preserving Spreads

Above, we stated (MIO) in terms of average posterior beliefs. We now provide an alternative characterization in terms of the joint distributions on the signal and the state. For concreteness, we consider two sets of payoff functions: those with nondecreasing incremental returns ($U^{R_{ND}}$), and those with concave incremental returns ($U^{R_{CV}}$).

Let H be the prior, and X, X' two signals. With essentially no loss of generality, we assume that X, X' both have uniformly distributed marginal distributions $F_X(X), F'_X(X')$.[15] To minimize notation, we take Ω and $\mathcal{X}$ to be finite, and work with discrete random variables. The continuous case is analogous.

Let f, f' denote the probability mass functions associated with F, F', and define $\gamma(\omega, x) = f'(\omega, x) - f(\omega, x)$. We say that F' can be obtained from F by a single "marginal-preserving

[15] The advantage to this is that for nondecreasing and concave incremental returns, this will ensure that the T transformation is just the identity, $T(x) = x$. The reason there is no loss of generality is that given a signal X, with associated information structure F, we can always define an informationally equivalent signal $Z = F_X(X)$. If X is continuous, then Z has a uniform marginal. Even if X is discrete, then using Lehmann's (1988) technique, we can first construct a continuous random variable X^* that is informationally equivalent ot X, and then derive Z from X^*.

spread" (MgPS) if for some $\varepsilon > 0$, $\omega_1 < \omega_2$, and $x_1 < x_2$,

$$\gamma(\omega_1, x_1) = \gamma(\omega_2, x_2) = -\gamma(\omega_2, x_1) = -\gamma(\omega_1, x_2) = \varepsilon,$$

and otherwise $\gamma = 0$.[16] Similarly, F' can be obtained from F by a single "mean and marginal-preserving spread" (MMgPS) if for some $\delta, \eta > 0$, $\omega_1 < \omega_2 < \omega_3 < \omega_4$, and $x_1 < x_2$,

$$\begin{aligned} \gamma(\omega_1, x_1) &= \gamma(\omega_2, x_2) = -\gamma(\omega_2, x_1) = -\gamma(\omega_1, x_2) = \delta, \\ -\gamma(\omega_3, x_1) &= -\gamma(\omega_4, x_2) = \gamma(\omega_4, x_1) = \gamma(\omega_3, x_2) = \eta. \end{aligned}$$

where $\sum_i \omega_i \gamma(\omega_i, x_j) = 0$ for $j = 1, 2$, and otherwise $\gamma = 0$.

As Figure 1 illustrates, a marginal-preserving spread captures the idea of making two random variables more positively dependent. A mean and marginal-preserving spread adds positive dependence below $\mathbb{E}[W]$ and negative dependence above $\mathbb{E}[W]$.

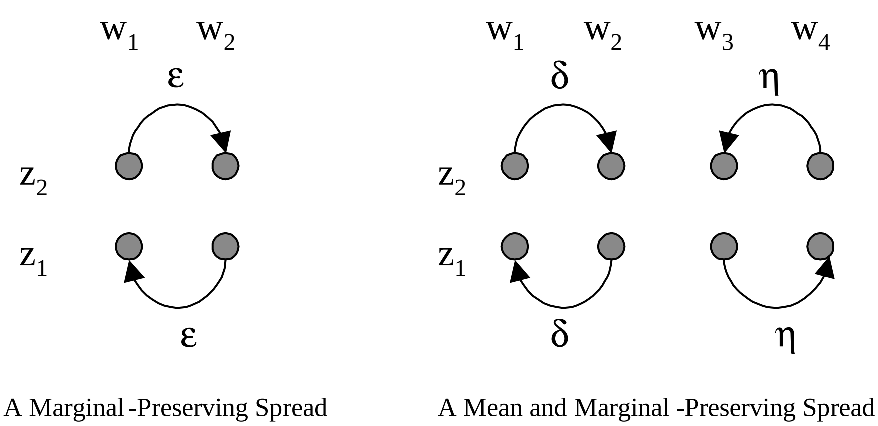

Figure 1: Illustration of Spreads

Proposition 2 *Consider a prior H, and two signals X, X' with uniform marginals, whose information structures F, F' are R-ordered (for the appropriate R). (i) $F' \succ_{MIO-ND} F$ if and only if F' can be obtained from F via a series of MgPSs; (ii) $F' \succ_{MIO-CV} F$ if and only if F' can be obtained from F via a series of MMgPSs.*

Proof: Part *(i)* follows a result in Meyer (1990). The proof of *(ii)* uses Rothschild and Stiglitz's (1970) characterization of SOSD in terms of mean-preserving spreads. Suppose F'

[16] Hamada (1974), Tchen (1976), Epstein and Tanny (1980) and Meyer (1990) have all discussed this concept in studying positive dependence of random variables.

is obtained by F via a single MMgPS, where the MMgPS is carried out on $z_1 < z_2$. Then $W|F > x \stackrel{d}{=} W|X' > x$ for any $x > x_2$ and any $x \leq x_1$, and the distribution of $W|X > x$ is a mean-preserving spread of $W|X' > x$ for any $x_1 < x \leq x_2$. So $W|X' > x$ dominates $W|X > x$ by SOSD for any x, and $F' \succ_{MIO-CV} F$. With multiple MMgPSs, the argument is repeated. For the other direction, suppose X, X' take values on $\{x_1, ..., x_n\}$ and have uniform marginals, and that $F' \succ_{MIO-CV} F$. Then the distribution of $W|X' = x_1$ is SOSD dominated by the distribution of $W|X = x_1$ and hence must be obtained by a finite series of mean-preserving spreads. Starting with F, make a series of MMgPSs using x_1, x_2 that result in a distribution $F^{(1)}$ with $W|X^{(1)} = x_1 \stackrel{d}{=} W|X' = x_1$, and $W|X^{(1)} \leq x_2 \stackrel{d}{=} W|X \leq x_2$. These last two are SOSD dominated by $W|X' \leq x_2$, so in particular, $W|X^{(1)} = x_2$ is stochastically dominated by $W|X' = x_2$. Now, starting with $F^{(1)}$ make a series of MMgPSs using x_2, x_3 that result in $G^{(2)}$ with $W|X^{(2)} = x_2 \stackrel{d}{=} W|X' = x_2$. Continuing in this fashion, we obtain from G via a series of MMgPSs, a distribution $G^{(n-1)}$ such that for all $x < x_n$, $W|X^{(n-1)} = x \stackrel{d}{=} W|X' = x$. Since $G^{(n-1)}$ and G' have equivalent marginals, it must therefore be that $G^{(n-1)} = G'. Q.E.D.$

This result fully characterizes (MIO) when R is the set of nondecreasing or concave functions.[17]

Example: A MMgPS that Violates Sufficiency. The following example illustrates a series of MMgPSs; it also shows how two signals X' and X might be information-ranked for payoff functions with concave incremental returns without one being statistically sufficient for the other. Suppose $\Omega = \{-2, -1, 0, 1, 2\}$ and the prior on Ω is uniform. There are two signals, which take realizations on $\mathcal{X} = \{x_1, x_2\}$, $x_2 > x_1$. The joint signal-state distributions are:

$\Pr(W = \omega, X = x)$	-2	-1	0	1	2
x_1	$\frac{11}{80}$	$\frac{5}{80}$	$\frac{8}{80}$	$\frac{5}{80}$	$\frac{11}{80}$
x_2	$\frac{5}{80}$	$\frac{11}{80}$	$\frac{8}{80}$	$\frac{11}{80}$	$\frac{5}{80}$

$\Pr(W = \omega, X' = x)$	-2	-1	0	1	2
x_1	$\frac{14}{80}$	$\frac{4}{80}$	$\frac{4}{80}$	$\frac{4}{80}$	$\frac{14}{80}$
x_2	$\frac{2}{80}$	$\frac{12}{80}$	$\frac{12}{80}$	$\frac{12}{80}$	$\frac{2}{80}$

Simple calculation shows that the posteriors induced by X do not lie in the convex hull of the posteriors induced by X', so X' cannot be sufficient for X. Thus for some decision-makers, X is preferred to X'. However, the posteriors induced by both X and X' are ordered by SOSD, and the distribution of (W, X') is obtained from the distribution of (W, X) by a sequence of two MMgPSs (one on $(\{x_1, x_2\}; \{-1, 0, 0, 1\})$, the other on $(\{x_1, x_2\}; \{-2, -1, 1, 2\})$). So

[17] When we further restrict attention to two states and two signals (where posterior beliefs are ordered by FOSD), it can be shown that a MgPS is equivalent to two elementary linear bifurcations on the distribution over posteriors, as in Grant, Kajii and Polak (1998); these authors show that this implies statistical sufficiency, which in turn implies that (MIO) and sufficiency are equivalent for this case.

X' is preferred to X by all decision-makers whose payoff functions have concave marginal returns.

5.2 Information Orders and Prior Beliefs

While (MIO) and the marginal preserving spread conditions are intuitive and easy to state, they have the disadvantage of intermingling the prior belief and the conditional signal distributions. In many economic contexts, it is reasonable to assume that the prior is known to the analyst; for example, the distribution over previous asset returns, input costs, or employee test scores may be objectively measured. But in other cases, such information may be unavailable, and we would prefer a characterization of information that relies less on knowledge of the prior beliefs.

Put differently, we have characterized information preferences for families of decision-makers of the form $(\{H\}, U^R)$. To obtain an information ranking for a family of decision-makers (Λ, U^R) where Λ is some set of possible prior beliefs, we must check (MIO) for the entire set Λ. Typically, this implies a further restriction on the informativeness order. To see why, write T_H in place of T to highlight dependence on the prior, and rewrite (MIO) (using (7)) as:

$$\forall r \in R,\ z \in [0,1]: \quad \int_\Omega r(\omega) \Pr\left(T'_H(X') > z \,|\, \omega\right) dH(\omega) \geq \int_\Omega r(\omega) \Pr\left(T_H(X) > z \,|\, \omega\right) dH(\omega). \tag{17}$$

This version of (MIO) compares weighted averages of the conditional signal distributions, where the states are weighted according to the prior H and the return functions $r \in R$.

Clearly, the prior H plays a role. Consequently, X' might be R-information preferred to X for one prior H_1, but not another prior H_2. Intuitively, relative to X, X' might not distinguish well between two states, ω and ω'; but if both are unlikely, X' might still be R-information preferred to X. Clearly, this feature is desirable when the prior is known, but undesirable when we consider a set of priors where other members place large weight on ω and ω'. An additional complication is that even for a given signal X, it might be the case that the induced posteriors are R-ordered for one prior H_1 but not for another prior H_2.

Consider first the set from Example 3, $R_{SC(\omega_o)}$. Here, $T_H(X) = F_X(X|\omega_o)$, which does not depend on the prior; further, $r \cdot h \in R_{SC(\omega_o)}$ for all prior densities h, if and only if $r \in R_{SC(\omega_o)}$. Thus, the prior does not affect (MIO) in this case. In contrast, for nondecreasing functions R_{ND}, $T_H(X) = F_X(X)$, which depends on the prior; and further, if we consider an alternate prior density $\tilde{h}$, r nondecreasing implies only that $r \cdot \tilde{h}$ is single crossing. Thus, this analysis indicates that much of the structure imposed by commonly encountered economic restrictions (i.e. monotonicity, concavity) can be "undone" by allowing for a rich enough set of prior beliefs Λ. We formalize this discussion as follows (recalling R_{SC} contains single crossing functions, and $R_{ND} \subset R_{SC}$):

Proposition 3 *(i) The information structure F corresponding to X is R_{ND}-ordered (i.e., FOSD-ordered) for all prior beliefs $H \in \Delta(\Omega)$ if and only if it is R_{SC}-ordered (i.e., MLRP-ordered). (ii) Given two signals X' and X that are MLRP-ordered, X' is more informative that X for all decision-makers in $(\Delta(\Omega), U^{R_{ND}})$ if and only if X' is more informative that X for decision-makers in $(\Delta(\Omega), U^{R_{SC}})$.*

Proof: Part *(i)* is due to Milgrom (1981). For part *(ii)*, observe that for a fixed prior H, $X' \succ_{MIO-ND} X$ states that $F'_W(\omega|F'_X(X') > z) \leq F_W(\omega|F_X(X) > z)$ $\forall \omega \in \Omega, z \in [0,1]$. Applying Bayes' Rule and canceling terms yields $F'_X(F'^{-1}_X(z)|W \leq \omega) \leq F_X(F^{-1}_X(z)|W \leq \omega)$ $\forall \omega \in \Omega, z \in [0,1]$; substituting $z = F_X(x)$ and re-arranging:

$$F'^{-1}_X\left(F_X(x)\right) \leq F'^{-1}_X\left(F_X(x|W \leq \omega)|W \leq \omega\right) \qquad \forall \omega \in \Omega, x \in \mathcal{X}.$$

Clearly, $F'^{-1}_X(F_X(x|\omega)|\omega)$ nondecreasing in ω for all $x \in \mathcal{X}$ implies this last inequality regardless of the prior H. Moreover, for the last inequality to hold for all $H \in \Delta(\Omega)$, it must hold for all priors with two point support, from which it follows that $F'^{-1}_X\left(F_X(x|\omega)|\omega\right)$ must be nondecreasing in ω for all x. Finally, as discussed in Example 3, $F'^{-1}_X(F_X(x|\omega)|\omega)$ is nondecreasing in ω for all x if and only if $X' \succ_{MIO-SC} X$. *Q.E.D.*

Thus, we see that allowing for a rich enough set of prior beliefs renders useless the knowledge that r is nondecreasing rather than just single crossing. In contrast, the orderings for single crossing functions (as well as statistical sufficiency) are independent of the prior, and knowing the prior does not allow these conditions to be tightened. Intuitively, the sets of payoff functions being considered are so large that if X' does not distinguish well between ω and ω', then even if these states are unlikely, there is still some payoff function that cares *only* about the comparison between them. Thus, we have shown that if the set of payoff functions being studied is large, knowledge of the prior does not help in characterizing information preferences, while such information is essential to fully exploit the structure of smaller sets of payoff functions.[18]

[18] Jewitt (1997) shows that for two-point priors, where posteriors are ordered by the monotone likelihood ratio order, Lehmann's order and statistical sufficiency coincide.

Example With a given prior, X' might be preferred to X for all decision-makers with nondecreasing incremental returns (but not for all decision-makers with single crossing incremental returns); yet, another prior may disturb the comparison. Suppose $\Omega = \{\omega_1, \omega_2, \omega_3\}$ with $\omega_1 < \omega_2 < \omega_3$ and $\mathcal{X} = \{x_1, x_2\}$ with $x_1 < x_2$. Consider the following two joint distributions, where F' is obtained from F via a single MgPS:

$\Pr(W = \omega, X = x)$	ω_1	ω_2	ω_3
x_1	$\frac{6}{24}$	$\frac{3}{24}$	$\frac{3}{24}$
x_2	$\frac{2}{24}$	$\frac{5}{24}$	$\frac{5}{24}$

$\Pr(W = \omega, X' = x)$	ω_1	ω_2	ω_3
x_1	$\frac{6}{24}$	$\frac{4}{24}$	$\frac{2}{24}$
x_2	$\frac{2}{24}$	$\frac{4}{24}$	$\frac{6}{24}$

Both $F_W(\omega|x)$ and $F'_W(\omega|x)$ have the MLRP, but X' and X do not satisfy (13). Consider the following payoff function, where $A = \{0, 1\}$. $u(\omega, a) = ar(\omega)$, where $r(\omega_1) = -2$, $r(\omega_2) \in (1, 2)$, and $r(\omega_3) = 1$. It is easily checked that $r(\omega)$ satisfies single crossing but is not nondecreasing, and that the ex ante payoff for u is higher under F than under F'. The idea is that X' distinguishes well between ω_1 and $\{\omega_2, \omega_3\}$, between $\{\omega_1, \omega_2\}$ and $\{\omega_3\}$, and between ω_2 and ω_3, but not well between ω_1 and ω_2. By allowing for a larger set of prior beliefs or payoff functions, one can choose the prior or the payoff function to focus attention on this last fact.

6 Conclusion

In this paper, we established that the additional structure of monotone decision problems can be exploited to derive necessary and sufficient conditions for all agents in a family to prefer one signal to another. These conditions are typically less restrictive than statistical sufficiency (the ordering for all payoff functions) or Lehmann's order (for single crossing functions). Alternatively, our results can be interpreted as deriving additional consequences of statistical sufficiency and Lehmann's order in different monotone decision problems (e.g. in problems where the payoff functions is supermodular, and the prior distribution is fixed). These consequences lead directly to comparative statics results in decision problems and equilibrium models.

Finally, both our definitions of more information and our results on information demand may be useful for building models in which agents acquire their information at some cost prior to playing a game of incomplete information. The examples given here, and Persico's (2000) work on auctions, indicate that this may be a fruitful area for further inquiry.

7 Appendix

Proof of Corollary 1: By definition, $F_X^{-1}(F_X(x|\omega_0) \mid \omega_0) = x$. Thus,

$$\frac{\partial}{\partial \omega}\ F_X^{-1}(F_X(x|\omega_0) \mid \omega)\big|_{\omega=\omega_0} = -\frac{\frac{\partial}{\partial \omega} F_X(x|\omega_0)}{f_X(x \mid \omega_0)},$$

or, letting $z = F_X(x|\ \omega_0)$,

$$\frac{\partial}{\partial \omega_0} F_X^{-1}(z|\omega_0) = -\frac{\frac{\partial}{\partial \omega} F_X(F_X^{-1}(z|\omega_0) \mid \omega)\big|_{\omega=\omega_0}}{f_X(F_X^{-1}(z|\omega_0) \mid \omega_0)}.$$

Then, letting $z = F_X(x|\omega_0)$, we differentiate (13), yielding $\frac{\partial}{\partial \omega_0} F_{X'}^{-1}(F_X(x|\omega_0)|\ \omega_0) =$

$$-\frac{\frac{\partial}{\partial \omega} F_{X'}(F_{X'}^{-1}(z|\omega_0) \mid \omega)\big|_{\omega=\omega_0}}{f_{X'}(F_{X'}^{-1}(z|\omega_0) \mid \omega_0)} + \frac{\frac{\partial}{\partial \omega} F_X(F_X^{-1}(z|\omega_0) \mid \omega)\big|_{\omega=\omega_0}}{f_{X'}(F_{X'}^{-1}(z|\omega_0) \mid \omega_0)}.$$

Multiplying both sides by $f_{X'}(F_{X'}^{-1}(z|\omega_0) \mid \omega_0)$, this is nonnegative if

$$\begin{aligned} &\frac{\partial}{\partial \omega}\left(\frac{f_W(\omega|X \leq F_X^{-1}(z|\omega_0))\ F_X(F_X^{-1}(z|\omega_0)}{h(\omega)}\right)\Bigg|_{\omega=\omega_0} \\ \geq\ &\frac{\partial}{\partial \omega}\left(\frac{f_W(\omega|X' \leq F_{X'}^{-1}(z|\omega_0))\ F_{X'}(F_{X'}^{-1}(z|\omega_0)}{h(\omega)}\right)\Bigg|_{\omega=\omega_0}. \end{aligned}$$

Simplifying, we have

$$\begin{aligned} &\frac{F_X(F_X^{-1}(z|\omega_0))}{h(\omega_0)}\left[f_W'(\omega_0|X \leq F_X^{-1}(z|\omega_0)) - f_W(\omega_0|X \leq F_X^{-1}(z|\omega_0))\ \frac{h'(\omega_0)}{h(\omega_0)}\right] \\ \geq\ &\frac{F_{X'}(F_{X'}^{-1}(z|\omega_0))}{h(\omega_0)}\left[f_W'(\omega_0|X' \leq F_{X'}^{-1}(z|\omega_0)) - f_W(\omega_0|X' \leq F_{X'}^{-1}(z|\omega_0))\frac{h'(\omega_0)}{h(\omega_0)}\right] \end{aligned}$$

or, multiplying through by $h(\omega_0)$ and factoring out a term, this is:

$$\begin{aligned} &f_W(\omega_0|X \leq F_X^{-1}(z|\omega_0))\ F_X(F_X^{-1}(z|\omega_0))\left[\frac{f_W'(\omega_0|X \leq F_X^{-1}(z|\omega_0))}{f_W(\omega_0|X \leq F_X^{-1}(z|\omega_0))} - \frac{h'(\omega_0)}{h(\omega_0)}\right] \qquad (18) \\ \geq\ &f_W(\omega_0|X' \leq F_{X'}^{-1}(z|\omega_0))\ F_{X'}(F_{X'}^{-1}(z|\omega_0))\left[\frac{f_W'(\omega_0|X' \leq F_{X'}^{-1}(z|\omega_0))}{f_W(\omega_0|X' \leq F_{X'}^{-1}(z|\omega_0))} - \frac{h'(\omega_0)}{h(\omega_0)}\right]. \end{aligned}$$

But, using Bayes' rule, $f_W(\omega_0|X \leq F_X^{-1}(z|\omega_0))\ F_X(F_X^{-1}(z|\omega_0)) = zh(\omega_0)$, so (18) becomes

$$\frac{f_W'(\omega_0 \mid X \leq F_X^{-1}(z|\omega_0))}{f_W(\omega_0 \mid X \leq F_X^{-1}(z|\omega_0))} \geq \frac{f_W'(\omega_0 \mid X' \leq F_{X'}^{-1}(z|\omega_0))}{f_W(\omega_0 \mid X' \leq F_{X'}^{-1}(z|\omega_0))}.$$

Finally, using Bayes' rule and the fact that the expectation of the posteriors must equal the prior, the latter inequality is equivalent to the desired condition. ■

8 References

Athey, Susan (1998): "Characterizing Properties of Stochastic Objective Functions," MIT Working Paper 96-1R.

Athey, Susan (2002): "Comparative Statics under Uncertainty: Single Crossing Properties and Log-Supermodularity," forthcoming, *Quarterly Journal of Economics*.

Blackwell, David (1951): "Comparisons of Experiments," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics*, 93–102.

Blackwell, David (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

Cooper, Russell (1999): *Coordination Games: Complementarities and Macroeconomics*, Cambridge University Press.

Epstein, L. and S. Tanny (1980): "Increasing Generalized Correlation: A Definition and Some Economic Consequences," *Canadian Journal of Economics*, 13, 16-34.

Grant, Simon, Ben Polak and Atsushi Kajii (1998): "Intrinsic Preference for Information," *Journal of Economic Theory*, 83, 233–259.

Gollier, C. and M. Kimball (1995): "Toward a Systematic Approach to the Economic Effects of Uncertainty I: Comparing Risks," Mimeo, University of Toulouse.

Hamada, K. (1974): "Comment on 'Stochastic Dominance in Choice under Uncertainty"', in M. Balch, D. McFadden, and S. Wu, eds., *Essays on Economic Behavior under Uncertainty*. Amsterdam: North Holland.

Jewitt, Ian (1986): "A Note on Comparative Statics and Stochastic Dominance," *Journal of Mathematical Economics*, 15, 249–254.

Karlin, Samuel and Howard Rubin (1956): The Theory of Decision Procedures for Distributions with Monotone Likelihood Ratio," *Annals of Mathematical Statistics*, 27, 272–299.

Le Cam, Luc (1964): "Sufficiency and Approximate Sufficiency," *Annals of Mathematical Statistics,* 35,

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

Meyer, Margaret (1990): "Interdependence of Multivariate Distributions: Stochastic Dominance Theorems and an Application to the Measurement of Ex Post Inequality under Uncertainty," Nuffield College Discussion Paper.

Milgrom, Paul (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 380–391.

Milgrom, Paul and John Roberts (1990): "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58, 1255–1277.

Milgrom, Paul and Christina Shannon (1994): "Monotone Comparative Statics," *Econometrica*, 62, 157–180.

Mirman, L., L. Samuelson, and E. Schlee (1994): "Strategic Manipulation of Information in Duopolies." *Journal of Economic Theory,* 36, 364-384.

Persico, Nicola (2000): "Information Acquisition in Auctions," *Econometrica*.

Rothschild, Michael and Joseph Stiglitz (1970): "Increasing Risk I: A Definition," *Journal of Economic Theory,* 2, 223–243.

Shaked, Moshe and George Shantikumar (1997): "Supermodular Stochastic Order and Positive Dependence of Random Variables," *Journal of Multivariate Analysis*, 61, 86–101.

Shannon, Chris (1995), "Weak and Strong Monotone Comparative Statics," *Economic Theory* 5 (2), March, pp. 209-27.

Tchen, A. (1976): *Inequalities for Distributions with Given Marginals*, Ph.D. Thesis, Dept. of Statistics, Stanford, June 1976

Topkis, Donald (1998): *Supermodularity and Complementarity.* Princeton University Press.