COMPARATIVE STATICS, INFORMATIVENESS, AND

THE INTERVAL DOMINANCE ORDER

By John K.-H. Quah and Bruno Strulovici

Abstract: We identify a natural way of ordering functions, which we call the *interval dominance order* and develop a theory of monotone comparative statics based on this order. This way of ordering functions is weaker than the standard one based on the single crossing property (Milgrom and Shannon, 1994) and so our results apply in some settings where the single crossing property does not hold. For example, they are useful when examining the comparative statics of optimal stopping time problems. We also show that certain basic results in statistical decision theory which are important in economics - specifically, the complete class theorem of Karlin and Rubin (1956) and the results connected with Lehmann's (1988) concept of informativeness - generalize to payoff functions obeying the interval dominance order.

Keywords: single crossing property, interval dominance order, supermodularity, comparative statics, optimal stopping time, complete class theorem, statistical decision theory, informativeness.

JEL Classification Numbers: C61, D11, D21, F11, G11.

Authors' Emails: john.quah@economics.ox.ac.uk   bruno.strulovici@economics.ox.ac.uk

1. INTRODUCTION

A central concept in the theory of monotone comparative statics is the single crossing property.[1] Let X be a subset of the real line R, and $\{f(\cdot, s)\}_{s \in S}$ a family of functions mapping X to R and parameterized by s in $S \subseteq R$. We say that this family is ordered by the *single crossing property* if for all $s'' > s'$ and $x'' > x'$, the following holds:

$$f(x'', s') - f(x', s') \geq (>) \, 0 \implies f(x'', s'') - f(x', s'') \geq (>) \, 0. \tag{1}$$

To keep our discussion simple, assume that $f(\cdot, s)$ has a unique maximum in X for all s. The fundamental importance of the single crossing property (SCP) arises from this result (Milgrom and Shannon (1994)): *if the family $\{f(\cdot, s)\}_{s \in S}$ is ordered by the single crossing property, then $\operatorname{argmax}_{x \in X} f(x, s)$ is increasing in s.*

It is trivial to see that SCP is also necessary for comparative statics in the following sense: if we require $\operatorname{argmax}_{x \in Y} f(x, s'') \geq \operatorname{argmax}_{x \in Y} f(x, s')$ for all $Y \subset X$ whenever $s'' > s'$, then (1) must hold. (This is true since we could let Y be any set consisting of two points in X.) In other words, SCP is necessary if we require the monotone comparative statics conclusion to be robust to certain changes in the domain of the objective functions.

However, SCP is *not* necessary for monotone comparative statics if we only require $\operatorname{argmax}_{x \in Y} f(x, s'') \geq \operatorname{argmax}_{x \in Y} f(x, s')$ for $Y = X$ or for Y belonging to a particular subcollection of the subsets of X. To see this, consider Figures 1 and 2. In both cases, we have $\operatorname{argmax}_{x \in X} f(x, s'') \geq \operatorname{argmax}_{x \in X} f(x, s')$; furthermore, $\operatorname{argmax}_{x \in Y} f(x, s'') \geq \operatorname{argmax}_{x \in Y} f(x, s')$ where Y is any closed interval contained in X. In Figure 1, SCP is satisfied (specifically, (1) is satisfied) but this is not true in Figure 2. Consider the points x'' and x' as depicted in Figure 2. We see that $f(x', s') = f(x'', s')$ but $f(x'', s'') < f(x', s'')$, violating SCP.

---

[1]Early contributions to the literature on monotone comparative statics include Topkis (1978), Milgrom and Roberts (1990), Vives (1990), and Milgrom and Shannon (1994). For a textbook treatment see Topkis (1998).

The first objective of this paper is to develop a new way of ordering functions that guarantees monotone comparative statics, whether the situation is like the one depicted in Figure 1 or the one depicted in Figure 2. We call this new order the *interval dominance order*. This order is more general than the one based on the single crossing property. We show that it holds in some significant situations where SCP does not hold; at the same time it retains many of the nice comparative statics properties associated with SCP.

The second objective of this paper is to bridge the gap between the literature on monotone comparative statics and the closely related literature in statistical decision theory on informativeness. We refer in particular to Lehmann's concept of informativeness (1988) which in turn builds on the complete class theorem of Karlin and Rubin (1956).[2] In that setting, the state of the world is unknown and the agent has to take an action based on an observed signal, which conveys information on the true state. Karlin and Rubin identify conditions under which optimal decision rules (in some well-defined sense) must be monotone, i.e., the agent's action is higher when he receives a higher signal.[3] Lehmann shows how we may compare the informativeness of two families of signals (or *experiments*) when the optimal decision rules are monotone.

A crucial assumption in the Karlin-Rubin and Lehmann theorems is that the agent's payoff in state s when she takes action x, which we write as $f(x, s)$, has the following property: $f(\cdot, s)$ is a quasiconcave function of x, achieving a maximum at $\bar{x}(s)$, with $\bar{x}$ increasing in s (like the situation depicted in Figure 2). However, their results do not cover the case where $\{f(\cdot, s)\}_{s \in S}$ is ordered by the single crossing property. Amongst other things, SCP allows for non-quasiconcave payoff functions (see Figure 2); indeed this feature is crucial to many of its economics applications.

---

[2]Economic applications of Lehmann's concept of informativeness can be found in Persico (2000), Athey and Levin (2001), Levin (2001), Bergemann and Valimaki (2002), and Jewitt (2006).

[3]In other words, monotone decision rules form a *complete class*.

In short, the standard results on comparative informativeness accommodates the situation depicted in Figure 2 but not that in Figure 1, whereas the standard results on comparative statics accommodate the situation in Figure 1 but not that in Figure 2.

We generalize the Karlin-Rubin and Lehmann results by showing that their conclusions hold even when the payoff functions $\{f(\cdot, s)\}_{s \in S}$ satisfy the interval dominance order. In this way, we obtain a single condition on the payoff functions that is useful for *both* comparative statics and comparative informativeness, so results in one category extend seamlessly into results in the other.

The rest of this paper is organized as follows. In Section 2, we define the interval dominance order, explore its properties, and develop a comparative statics theorem. Section 3 is devoted to applications. In Section 4 we show that the concept of IDO can be easily extended to settings with uncertainty and that it is useful in that context. Lehmann's concept of informativeness is introduced in Section 5 and we demonstrate its relevance for payoff functions obeying the interval dominance order. Finally, we generalize Karlin and Rubin's complete class theorem in Section 6.

2. The Interval Dominance Order

We begin by showing how a situation like that depicted in Figure 2, involving a violation of SCP, can arise naturally in an economic setting.

Example 1. Consider a firm producing some good whose price we assume is fixed at 1 (either because of market conditions or for some regulatory reason). It has to decide on the production capacity (x) of its plant. Assume that a plant with production capacity x costs Dx, where D is a positive scalar. Let s be the state of the world, which we also identify with the demand for the good. The marginal cost of producing the good in state s is $c(s)$. We assume that, for all s, $D + c(s) < 1$. The firm makes its capacity decision *before* the state of the world is realized and its

4

production decision *after* the state is revealed.

Suppose it chooses capacity x and the realized state of the world (and thus realized demand) is $s \geq x$. In this case, the firm should produce up to its capacity, so that its profit $\Pi(x, s) = x - c(s)x - Dx$. On the other hand, if $s < x$, the firm will produce (and sell) s units of the good, giving it a profit of $\Pi(x, s) = s - c(s)s - Dx$. It is easy to see that $\Pi(\cdot, s)$ is increasing linearly for $x \leq s$ and thereafter declines, linearly with a slope of $-D$. Its maximum is achieved at $x = s$, with $\Pi(s, s) = (1 - c(s) - D)s$. Suppose $s'' > s'$ and $c(s'') > c(s')$. In other words, the state with higher demand also has higher marginal cost. Then it is clear that the situation depicted in Figure 2 can arise, with $f(\cdot, s') = \Pi(\cdot, s')$ and $f(\cdot, s'') = \Pi(\cdot, s'')$.

Let X be a subset of R and f and g two real-valued functions defined on X. We say that g dominates f by the *single crossing property* (which we denote by $g \succeq_{sc} f$) if for all x'' and x' such that $x'' > x'$, the following holds:

$$f(x'') - f(x') \geq (>) \, 0 \implies g(x'') - g(x') \geq (>) \, 0. \tag{2}$$

A family of real-valued functions $\{f(\cdot, s)\}_{s \in S}$, defined on X and parameterized by s in $S \subset R$ is referred to as an *SCP family* if the functions are ordered by SCP, i.e., whenever $s'' > s'$, we have $f(\cdot, s'') \succeq_{sc} f(\cdot, s')$.

In Figure 2, we have $\operatorname{argmax}_{x \in R_+} f(x, s'') > \operatorname{argmax}_{x \in R_+} f(x, s')$ even though $f(\cdot, s'')$ does not dominate $f(\cdot, s')$ by SCP. Notice, however, that violations of (2) can only occur if we compare points x' and x'' on opposite sides of the maximum point of $f(\cdot, s')$. This suggests that a possible way of weakening SCP, while retaining comparative statics, is to require (2) to hold only for a certain collection of pairs $\{x', x''\}$, rather than all possible pairs.

The set J is an *interval of X* if, whenever x' and x'' are in J, any element x in X such that $x' \leq x \leq x''$ is also in J.[4] Let f and g be two real-valued functions defined

---

[4]Note that X need not be an interval in the conventional sense, i.e., X need not be, using our

on X. We say that g dominates f by the *interval dominance order* (or, for short, g I-dominates f, with the notation $g \succeq_I f$) if (2) holds for x'' and x' such that $x'' > x'$ and $f(x'') \geq f(x)$ for all x in the interval $[x', x''] = \{x \in X : x' \leq x \leq x''\}$.

Clearly, the interval dominance order (IDO) is weaker than ordering by SCP. For example, in Figure 2, $f(\cdot, s'')$ I-dominates $f(\cdot, s')$ but $f(\cdot, s'')$ does not dominate $f(\cdot, s')$ by SCP.

For many results in the paper, we shall impose a mild regularity condition on the objective function. A function $f : X \to R$ is said to be *regular* if $\operatorname{argmax}_{x \in [x', x'']} f(x)$ is nonempty for any points x' and x'' with $x'' > x'$. Suppose the set X is such that $X \cup [x', x'']$ is always closed, and thus compact, in R (with the respect to the Euclidean topology). This is true, for example, if X is finite, if it is closed, or if it is a (not necessarily closed) interval. Then f is regular if it is upper semi-continuous with respect to the relative topology on X.

We are now ready to examine the relationship between the interval dominance order and monotone comparative statics. Theorem 1 gives the precise sense in which IDO is both sufficient and necessary for monotone comparative statics. To deal with the possibility of multiple maxima we need a way of ordering sets. The standard way of ordering sets in this context is the *strong set order* (see Topkis (1998)). Let S' and S'' be two subsets of R. We say that S'' dominates S' in the strong set order, and write $S'' \geq S'$ if for any for x'' in S'' and x' in S', we have $\max\{x'', x'\}$ in S'' and $\min\{x'', x'\}$ in S'. Suppose that S'' and S' both contain their largest and smallest elements. Then it is clear that if $S'' \geq S'$, the largest (smallest) element in S'' is greater than the largest (smallest) element in S'.[5]

<hr />

terminology, an *interval of* R. Furthermore, the fact that J is an interval of X does not imply that it is an interval of R. For example, if $X = \{1, 2, 3, \}$, then $J = \{1, 2\}$ is an interval of X, but of course neither X nor J are intervals of R.

[5]Throughout this paper, when we say that something is 'greater' or 'increasing', we mean to say that it is greater or increasing in the *weak* sense. Most of the comparisons in this paper are weak, so this convention makes sense. When we are making a strict comparison, we shall say so explicitly,

THEOREM 1: *Suppose that f and g are real-valued functions defined on $X \subset R$ and $g \succeq_I f$. Then the following property holds:*

($\star$) $\qquad \mathrm{argmax}_{x \in J} g(x) \geq \mathrm{argmax}_{x \in J} f(x)$ *for any interval J of X.*

Furthermore, if property ($\star$) holds and g is regular, then $g \succeq_I f$.

Proof: Assume that g I-dominates f and that x'' is in $\mathrm{argmax}_{x \in J} f(x)$ and x' is in $\mathrm{argmax}_{x \in J} g(x)$. We need only consider the case where $x'' > x'$. Since x'' is in $\mathrm{argmax}_{x \in J} f(x)$, we have $f(x'') \geq f(x)$ for all x in $[x', x''] \subseteq J$. Since $g \succeq_I f$, we also have $g(x'') \geq g(x')$; thus x'' is in $\mathrm{argmax}_{x \in J} g(x)$. Furthermore, $f(x'') = f(x')$ so that x' is in $\mathrm{argmax}_{x \in J} f(x)$. If not, $f(x'') > f(x')$ which implies (by the fact that $g \succeq_I f$) that $g(x'') > g(x')$, contradicting the assumption that g is maximized at x'.

To prove the other direction, we assume that there is an interval $[x', x'']$ such that $f(x'') \geq f(x)$ for all x in $[x', x'']$. This means that x'' is in $\mathrm{argmax}_{x \in [x', x'']} f(x)$. There are two possible violations of IDO. One possibility is that $g(x') > g(x'')$; in this case, by the regularity of g, the set $\mathrm{argmax}_{x \in [x', x'']} g(x)$ is nonempty but does not contain x'', which violates ($\star$). Another possible violation of IDO occurs if $g(x'') = g(x')$ but $f(x'') > f(x')$. In this case, the set $\mathrm{argmax}_{x \in [x', x'']} g(x)$ either contains x', which violates ($\star$) since $\mathrm{argmax}_{x \in [x', x'']} f(x)$ does not contain x', or it contains some element distinct from x'', which also violates ($\star$). $\hfill$ QED

For the interval dominance order to be useful in applications, we need some simple ways of checking that the property holds. For this purpose, the next result is crucial.

PROPOSITION 1: *Suppose that X is an interval of R and the functions f, $g : X \to R$ are absolutely continuous on compact intervals in X (and thus f and g are differentiable a.e.). If there is an increasing and positive function $\alpha : X \to R$ such that $g'(x) \geq \alpha(x) f'(x)$ a.e., then $g \succeq_I f$.*

If the function α in Proposition 1 is a constant $\bar{\alpha}$, then we obtain $g(x'') - g(x') \geq \bar{\alpha}(f(x'') - f(x'))$, which implies $g \succeq_{SC} f$. When α is not a constant, the functions f and g in Proposition 2 need not be related by SCP, as the following example shows.

---

as in 'strictly higher', 'strictly increasing', etc.

Let $f : [0, M] \to R$ be a differentiable and quasiconcave function, with $f(0) = 0$ and a unique maximum at x^* in $(0, M)$. Let $\alpha : [0, M] \to R$ be given by $\alpha(x) = 1$ for $x \leq x^*$ and $\alpha(x) = 1 + (x - x^*)$ for $x > x^*$. Consider $g : [0, M]$ satisfying $g(0) = f(0) = 0$ with $g'(x) = \alpha(x)f'(x)$ (as in Proposition 1). Then it is clear that $g(x) = f(x)$ for $x \leq x^*$ and $g(x) < f(x)$ for $x > x^*$. The function g is also quasiconcave with a unique maximum at x^* and g I-dominates f, but g does not dominate f by SCP.

Proposition 1 is a consequence of the following lemma.

LEMMA 1: *Suppose $[x', x'']$ is a compact interval of R and α and h are real-valued functions defined on $[x', x'']$, with h integrable and α increasing (and thus integrable as well). If $\int_x^{x''} h(t)dt \geq 0$ for all x in $[x', x'']$, then*

$$\int_{x'}^{x''} \alpha(t)h(t)dt \geq \alpha(x') \int_{x'}^{x''} h(t)dt. \tag{3}$$

Proof: We confine ourselves to the case where α is an increasing and differentiable function. If we can establish (3) for such functions, then we can extend it to all increasing functions α since any such function can be approximated by an increasing and differentiable function.

The function $H(t) = \alpha(t) \int_t^{x''} h(z)dz$ is absolutely continuous and thus differentiable a.e.; by the fundamental theorem of calculus, we have $H(x'') - H(x') = \int_{x''}^{x'} H'(t)dt$. Note that $H(x'') = 0$ and that by the product rule,

$$H'(t) = \alpha'(x) \int_t^{x''} h(z)dz - \alpha(t)h(t).$$

So

$$-H(x') = -\alpha(x') \int_{x'}^{x''} h(t)dt = \int_{x'}^{x''} \alpha'(x) \left(\int_t^{x''} h(z)dz \right) dt - \int_{x'}^{x''} \alpha(t)h(t)dt.$$

Note that the first term on the right of this equation is always nonnegative by assumption and so we obtain (3). QED

8

Proof of Proposition 1: Consider x'' and x' in X such that $x'' > x'$ and assume that $f(x) \leq f(x'')$ for all x in $[x', x'']$. Since f is absolutely continuous on $[x', x'']$, $f(x'') - f(x) = \int_x^{x^*} f'(t)dt$ (with an analogous expression for g). We then have

$$\int_x^{x''} g'(t)dt \geq \int_x^{x''} \alpha(t)f'(t)dt \geq \alpha(x') \int_{x'}^{x''} f'(t)dt,$$

where the second inequality follows from Lemma 1. So

$$g(x'') - g(x') = \alpha(x')(f(x'') - f(x')) \tag{4}$$

and $g(x'') \geq (>)g(x')$ if $f(x'') \geq (>)f(x')$. \hfill QED

Another familiar concept in the theory of monotone comparative statics, and one which is stronger than SCP, is the concept of *increasing differences* (see Milgrom and Shannon (1994) and Topkis (1998)). The function g dominates f by *increasing differences* (see Milgrom and Shannon (1994) and Topkis (1998)) if for any x'' and x' in X, with $x'' > x'$, we have

$$g(x'') - g(x') \geq f(x'') - f(x'). \tag{5}$$

We say that a function g dominates f by *conditional increasing differences* (and denote it by $g \succeq_{IN} f$) if (5) holds for all pairs x' and x'' in X such that $f(x'') \geq f(x)$ for x in $[x', x'']$. Clearly, if g dominates f by conditional increasing differences then g I-dominates f. The next result gives sufficient conditions under which g and f can be related in this manner.

PROPOSITION 2: *Suppose X is an interval of R and the functions f, $g : X \to R$ are absolutely continuous on compact intervals in X (and thus f and g are differentiable a.e.). If there is an increasing function $\alpha : X \to R$ with $\alpha(x) \geq 1$ for all x in X such that $g'(x) \geq \alpha(x)f'(x)$ a.e., then $g \succeq_{\text{IN}} f$.*

Proof: Retrace the proof of Proposition 1; the result follows from (4). \hfill QED

9

Many problems in economic theory involves the maximization of the difference between benefits and costs arising from some activity. The next result shows the relevance of conditional increasing differences to problems of this sort.

PROPOSITION 3: *Let $X \subset R$ and $b, \tilde{b},$ and c be real-valued functions defined on X, with c an increasing function. If $\tilde{b}$ dominates b by conditional increasing differences, then the function $\tilde{\Pi}$ dominates the function Π by conditional increasing differences, where $\tilde{\Pi}(x) = \tilde{b}(x) - c(x)$ and $\Pi(x) = b(x) - c(x)$. Consequently, $\mathrm{argmax}_{x \in X} \tilde{\Pi}(x) \geq \mathrm{argmax}_{x \in X} \Pi(x)$.*

Proof: Suppose $\Pi(x'') \geq \Pi(x)$ for x in $[x', x'']$. Since c is increasing, we also have $b(x'') \geq b(x)$ for x in $[x', x'']$. By conditional increasing differences $\tilde{b}(x'') - \tilde{b}(x') \geq b(x'') - b(x')$. Adding $-c(x'') + c(x')$ to both sides of this inequality, we obtain

$$\tilde{\Pi}(x'') - \tilde{\Pi}(x') \geq \Pi(x'') - \Pi(x')$$

as required. The final comparative statics statement follows from Theorem 1. QED

There are several issues relating to Proposition 3 that are worth highlighting. Firstly, the proposition does not assume that the benefit functions b and $\tilde{b}$ are increasing in x. It does, of course, require that they be related by conditional increasing differences; furthermore, this cannot be weakened to to SCP or I-dominance (see Milgrom and Shannon (1994)). The assumption that c is increasing is crucial to its proof. Indeed, without this assumption, the conclusion that $\mathrm{argmax}_{x \in X} \tilde{\Pi}(x) \geq \mathrm{argmax}_{x \in X} \Pi(x)$ is *only* possible by assuming that $\tilde{b}$ dominates b by increasing - rather than conditional increasing - differences (see Athey et al. (1998)).

In this section, we introduced the interval dominance order and gave a quick survey of its basic properties. This theory admits further development. In particular, we shall examine in Section 4 how the property arises in decision problems under uncertainty. There is also natural way in which the order may be defined in higher dimensions, with useful implications for comparative statics in that context. This issue, and others, are studied in detail in a companion paper (see Quah and Strulovici

(2007)). The next section is devoted to applying the theoretical results obtained so far; readers interested in a quick overview of the theory can skip the next section.

3. Applications of the IDO property

Example 2. A very natural application of Propositions 1 and 2 is to the comparative statics of optimal stopping time problems. We consider a simple deterministic problem here; there is a natural extension to the stochastic optimal stopping time problem which we examine in Quah and Strulovici (2007).

Suppose we are interested in maximizing $V_\delta(x) = \int_0^x e^{-\delta t} u(t) dt$ for $x \geq 0$, where $\delta > 0$ and the function $u : R_+ \to R$ is bounded on compact intervals and measurable. So x may be interpreted as the stopping time, δ is the discount rate, $u(t)$ the cash flow or utility of cash flow at time t (which may be positive or negative), and $V_\delta(x)$ is the discounted sum of the cash flow (or its utility) when x is the stopping time.

We are interested in how the optimal stopping time changes with the discount rate. It seems natural that the optimal stopping time will rise as the discount rate δ falls. This intuition is correct but it cannot be proved by the methods of concave optimization since V_δ need not be a quasiconcave function. Indeed, it will have a turning point every time u changes sign and its local maxima occur when u changes sign from positive to negative. Changing the discount rate does not change the times at which local maxima are achieved, but it potentially changes the time at which the *global* maximum is achieved, i.e., it changes the optimal stopping time. The next result gives the solution to this problem.

PROPOSITION 4: *Suppose that $\delta > \bar\delta > 0$. Then the following holds:*

(i) $V_{\bar\delta} \succeq_{IN} V_\delta$;

(ii) $\operatorname{argmax}_{x \geq 0} V_{\bar\delta}(x) \geq \operatorname{argmax}_{x \geq 0} V_\delta(x)$; and

(iii) $\max_{x \geq 0} V_{\bar\delta}(x) \geq \max_{x \geq 0} V_\delta(x)$.

Proof: The functions $V_{\bar\delta}$ and V_δ are absolutely continuous and thus differentiable

11

a.e.; moreover,

$$V_{\bar{\delta}}'(x) = e^{-\bar{\delta}x}u(x) = e^{(\delta-\bar{\delta})x}V_{\delta}'(x).$$

Note that the function $\alpha(x) = e^{(\delta-\bar{\delta})x}$ is increasing and greater than 1. So part (i) follows from Proposition 2 and part (ii) from Theorem 1. For (iii), let us suppose that $V_{\delta}(x)$ is maximized at $x = x^*$. Then for all x in $[0, x^*]$, $V_{\delta}(x) \leq V_{\delta}(x^*)$. Since $V_{\delta}(0) = V_{\bar{\delta}}(0) = 0$, the fact that $V_{\bar{\delta}} \succeq_{IN} V_{\delta}$ now guarantees that $V_{\bar{\delta}}(x^*) \geq V_{\delta}(x^*)$. Finally, note that $\max_{x \geq 0} V_{\bar{\delta}}(x) \geq V_{\bar{\delta}}(x*)$. $\hspace{2cm}$ QED

Example 3. Consider a firm that chooses output x to maximize profit, given by $\Pi(x) = xP(x) - C(x)$, where P is the inverse demand function and C is the cost function. Imagine that there is a change in market conditions, so that the both P and C are changed to $\tilde{P}$ and $\tilde{C}$ respectively. When can we say that $\text{argmax}_{x \geq 0}\tilde{\Pi}(x) \geq \text{argmax}_{x \geq 0}\Pi(x)$? By Theorem 1, this holds if $\tilde{\Pi}$ I-dominates Π. Intuitively, we will expect this to hold if the increase in the inverse demand is greater than any increase in costs. This idea can be formalized in the following manner.

Assume that all the functions are differentiable, that $\tilde{P}$ and P take strictly positive values, and that the cost functions are strictly increasing. Define $a(x) = \tilde{\Pi}(x)/\Pi(x)$. Then

$$
\begin{aligned}
\tilde{\Pi}'(x) &= a'(x)xP(x) + a(x)(xP(x))' - \tilde{C}'(x) \\
&\geq a(x)(xP(x))' - \left[\frac{\tilde{C}'(x)}{C'(x)}\right]C'(x),
\end{aligned}
$$

where the inequality follows since $xP(x) > 0$. Now suppose we make we assume that

$$a(x) = \frac{\tilde{P}(x)}{P(x)} \geq \frac{\tilde{C}'(x)}{C'(x)}, \hspace{3cm} (6)$$

then we obtain $\tilde{\Pi}'(x) \geq a(x)\Pi'(x)$. By Proposition 1, $\tilde{\Pi}$ I-dominates Π if a is increasing and (6) holds; in other words, the ratio of the inverse demand functions is increasing in x and greater than the ratio of the marginal costs.[6]

---

[6]Note that our argument does not require that P be decreasing in x.

Example 4. The perceptive reader will notice that Proposition 1 and Lemma 1 when presented in a somewhat different context are in fact very familiar since they are closely related to standard results on stochastic dominance. For example, recall that a density function γ defined on some interval $[a, b]$ in R is said to first order stochastically dominate another density function β defined on the same interval if $\int_a^x \beta(t)dt \geq \int_a^x \gamma(t)dt$ for all x in $[a, b]$. It is well known that this implies that for any increasing and bounded utility function u defined on $[a, b]$, we have

$$\int_a^b u(t)\gamma(t)dt \geq \int_a^b u(t)\beta(t)dt \; ; \tag{7}$$

in other words, the expected utility of γ is greater than that of β.

One could think of this basic result as an application of the theory developed here. If γ first order stochastically dominates β then $f : [a, b] \to R$ defined by $f(x) = \int_a^x \left(\gamma(t) - \beta(t)\right) dt$ has the following properties: $f(a) = f(b) = 0$ and $f(x) \leq 0 = f(b)$ for all x in $[a, b]$. By Proposition 1, the function g defined by

$$g(x) = \int_a^x u(t) \left(\gamma(t) - \beta(t)\right) dt$$

I-dominates f since $g'(x) = u(x)f'(x)$. So $g(b) \geq g(a) = 0$, which gives us (7).

4. The interval dominance order when the state is uncertain

Consider the following problem. Let $\{f(\cdot, s)\}_{s \in S}$ be a family of functions parameterized by s in S, an interval of R, with each function $f(\cdot, s)$ mapping Y, an interval of R, to R. Assume that all the functions are quasiconcave, with their peaks increasing in s; by this we mean that $\operatorname{argmax}_{x \in Y} f(x, s'') \geq \operatorname{argmax}_{x \in Y} f(x, s')$ whenever $s'' > s'$. (Note that since each function $f(\cdot, s)$ is quasiconcave, it either has a unique maximizer or they must form an interval.) We shall refer to such a family of functions as a QCIP family, where QCIP stands for *quasiconcave with increasing peaks*. Interpreting s to be the state of the world, an agent has to choose x under uncertainty, i.e., before

s is realized. We assume the agent maximizes the expected value of his objective; formally, he maximizes

$$F(x, \lambda) = \int_{s \in S} f(x, s)\lambda(s)ds,$$

where $\lambda : S \to R$ is the density function defined over the states of the world. It is natural to think that if the agent considers the higher states to be more likely, then his optimal value of x will increase. Is this true? More generally, we can ask the same question if the functions $\{f(\cdot, s)\}_{s \in S}$ form an IDO family, i.e., a family of regular functions $f(\cdot, s) : X \to R$, with $X \subseteq R$, such that $f(\cdot, s'')$ I-dominates $f(\cdot, s')$ whenever $s'' > s'$.

One way of formalizing the notion that higher states are more likely is via the monotone likelihood ratio (MLR) property. Let λ and γ be two density functions defined on the interval S of R and assume that $\lambda(s) > 0$ for s in S. We call γ an *MLR shift* of λ if $\gamma(s)/\lambda(s)$ is increasing in s. For density changes of this kind, there are two results that come close, though not quite, to addressing the problem we posed.

Ormiston and Schlee (1993) identify some conditions under which an upward MLR shift in the density function will raise the agent's optimal choice. Amongst other conditions, they assume that $F(\cdot; \lambda)$ is quasiconcave. This will hold if all the functions in the family $\{f(\cdot, s)\}_{s \in S}$ are concave but will not generally hold if the functions are just quasiconcave. Athey (2002) has a related result which says that an upward MLR shift will lead to a higher optimal choice of x provided $\{f(\cdot, s)\}_{s \in S}$ is an SCP family. As we had already pointed out in Example 1, a QCIP family need not be an SCP family.

The next result gives the solution to the problem we posed.

THEOREM 2: *Let S be an interval of R and $\{f(\cdot, s)\}_{s \in S}$ be an IDO family. Then $F(\cdot, \gamma) \succeq_I F(\cdot, \lambda)$ if γ is an MLR shift of λ. Consequently, $\operatorname{argmax}_{x \in X} F(x, \gamma) \geq \operatorname{argmax}_{x \in X} F(x, \lambda)$.*

Notice that since $\{f(\cdot, s)\}_{s \in S}$ in Theorem 2 is assumed to be an IDO family, we know (from Theorem 1) that $\operatorname{argmax}_{x \in X} f(x, s'') \geq \operatorname{argmax}_{x \in X} f(x, s')$. Thus Theorem

14

2 guarantees that the comparative statics which holds when s is known also holds when s is unknown but experiences an MLR shift.[7]

The proof of Theorem 2 requires a lemma (stated below). Its motivation arises from the observation that if $g \succeq_{SC} f$, then for any $x'' > x'$ such that $g(x') - g(x'') \geq (>) 0$, we must also have $f(x') - f(x'') \geq (>) 0$. Lemma 2 is the (less trivial) analog of this observation in the case when $g \succeq_I f$.

LEMMA 2: *Let X be a subset of R and f and g two regular functions defined on X. Then $g \succeq_I f$ if and only if the following property holds:*

(M)   if $g(x') \geq g(x)$ for x in $[x', x'']$ then

$$g(x') - g(x'') \geq (>) 0 \implies f(x') - f(x'') \geq (>) 0.$$

Proof: Suppose $x' < x''$ and $g(x') \geq g(x)$ for x in $[x', x'']$. There are two possible ways for property (M) to be violated. One possibility is that $f(x'') > f(x')$. By regularity, we know that $\operatorname{argmax}_{x \in [x', x'']} f(x)$ is nonempty; choosing x^* in this set, we have $f(x^*) \geq f(x)$ for all x in $[x', x^*]$, with $f(x^*) \geq f(x'') > f(x')$. Since $g \succeq_I f$, we must have $g(x^*) > g(x')$, which is a contradiction.

The other possible violation of (M) occurs if $g(x') > g(x'')$ but $f(x') = f(x'')$. By regularity, we know that $\operatorname{argmax}_{x \in [x', x'']} f(x)$ is nonempty, and if f is maximized at x^* with $f(x^*) > f(x')$, then we are back to the case considered above. So assume that x' and x'' are both in $\operatorname{argmax}_{x \in [x', x'']} f(x)$. Since $f \succeq_I g$, we must have $g(x'') \geq g(x')$, contradicting our initial assumption.

So we have shown that (M) holds if $g \succeq_I f$. The proof that (M) implies $g \succeq_I f$ is similar.   QED

Proof of Theorem 2: This consists of two parts. Firstly, we prove that if $F(x'', \lambda) \geq F(x, \lambda)$ for all x in $[x', x'']$, then, for any $\tilde{s}$ in S,

$$\int_{\tilde{s}}^{s^*} \left(f(x'', s) - f(x', s) \right) \lambda(s) ds \geq 0 \tag{8}$$

(where s^* denotes the supremum of S). Assume instead that there is $\bar{s}$ such that

$$\int_{\bar{s}}^{s^*} \left(f(x'', s) - f(x', s) \right) \lambda(s) ds < 0. \tag{9}$$

By the regularity of $f(\cdot, \bar{s})$, there is $\bar{x}$ that maximizes f in $[x', x'']$. In particular, $f(\bar{x}, \bar{s}) \geq f(x, \bar{s})$ for all x in $[\bar{x}, x'']$. Since $\{f(\cdot, s)\}_{s \in S}$ is an IDO family of regular functions, we also have $f(\bar{x}, s) \geq f(x'', s)$ for all $s \leq \bar{s}$ (using Lemma 2). Thus

$$\int_{s_*}^{\bar{s}} \left(f(\bar{x}, s) - f(x'', s) \right) \lambda(s) ds \geq 0, \tag{10}$$

where s_* is the infimum of S. Notice also that $f(\bar{x}, \bar{s}) \geq f(x, \bar{s})$ for all x in $[x', \bar{x}]$, which implies that $f(\bar{x}, s) \geq f(x', s)$ for all $s \geq \bar{s}$. Aggregating across s we obtain

$$\int_{\bar{s}}^{s^*} \left(f(\bar{x}, s) - f(x', s) \right) \lambda(s) ds \geq 0. \tag{11}$$

It follows from (9) and (11) that

$$
\begin{aligned}
\int_{\bar{s}}^{s^*} \left(f(\bar{x}, s) - f(x'', s) \right) \lambda(s) ds &= \int_{\bar{s}}^{s^*} \left(f(\bar{x}, s) - f(x', s) \right) \lambda(s) ds \\
&\quad + \int_{\bar{s}}^{s^*} \left(f(x', s) - f(x'', s) \right) \lambda(s) ds \\
&> 0.
\end{aligned}
$$

Combining this with (10), we obtain

$$\int_{s_*}^{s^*} \left(f(\bar{x}, s) - f(x'', s) \right) \lambda(s) ds > 0\,;$$

in other words, $F(\bar{x}, \lambda) > F(x'', \lambda)$ which is a contradiction.

Given (8), the function $H(\cdot, \lambda) : [s_*, s^*] \to R$ defined by

$$H(\tilde{s}, \lambda) = \int_{s_*}^{\tilde{s}} \left(f(x'', s) - f(x', s) \right) \gamma(s) ds$$

satisfies $H(s^*, \lambda) \geq H(\tilde{s}, \lambda)$ for all $\tilde{s}$ in $[s_*, s^*]$. Defining $H(\cdot, \gamma)$ in an analogous fashion, we also have $H'(\tilde{s}, \gamma) = [\gamma(s)/\lambda(s)]H'(\tilde{s}, \lambda)$ for $\tilde{s}$ in S. Since γ is an upward MLR shift of λ, the ratio $\gamma(s)/\lambda(s)$ is increasing in s. By Proposition 1, $H(\cdot, \gamma) \succeq_I H(\cdot, \lambda)$.

16

In particular, we have $H(s^*, \gamma) \geq (>)H(s_*, \gamma) = 0$ if $H(s^*, \lambda) \geq (>)H(s_*, \lambda) = 0$. Re-writing this, we have $F(x'', \gamma) \geq (>)F(x', \gamma)$ if $F(x'', \lambda) \geq (>)F(x', \lambda)$.   QED

Note that Theorem 2 remains true if S is not an interval; in Appendix A, we prove Theorem 2 in the case where S is a finite set of states.

We turn now to two applications of Theorem 2.

Example 1 continued. Recall that in state s, the firm's profit is $\Pi(x, s)$. It achieves its maximum at $x^*(s) = s$, with $\Pi(s, s) = (1 - c(s) - D)s$, which is strictly positive by assumption. The firm has to choose its capacity before the state of the world is realized; we assume that s is drawn from S, an interval in R, and has a distribution given by the density function $\lambda : S \to R$. We can think of the firm as maximizing its expected profit, which is $\int_S \Pi(x, s)\lambda(s)ds$, or more generally, let us assume that it maximizes the expected utility from profit, i.e., it maximizes $F(x, \lambda) = \int_S u(\Pi(x, s), s)\lambda(s)ds$, where, for each s, the function $u(\cdot, s) : R \to R$ is strictly increasing. The family $\{u(\Pi(\cdot, s), s)\}_{s \in S}$ consists of quasiconcave functions, the peaks of which are increasing in s. By Theorem 2, we know that an upward MLR shift of the density function will lead the firm to choose a greater capacity.

Example 5. Consider a firm that has to decide on when to launch a new product. The more time the firm gives itself, the more it can improve the quality of the product and its manufacturing process, but it also knows that there is a rival about to launch a similar product. In formal terms, we assume that the firm's profit (if it is not anticipated by its rival) is an increasing function of time $\bar{\pi} : R_+ \to R_+$. If the rival launches its product at time s, then the firm's profit falls to $w(s)$ (in R). In other words, the firm's profit in state s is $\pi(t, s) = \bar{\pi}(t)$ for $t \leq s$ and $w(s)$ for $t > s$, where $w(s) < \bar{\pi}(s)$. Clearly, each $\pi(\cdot, s)$ is a quasiconcave function and $\{\pi(\cdot, s)\}_{s \in S}$ is an IDO family. The firm decides on the launch date t by maximizing $F(t, \lambda) = \int_{s \in S} \pi(t, s)\lambda(s)ds$, where $\lambda : R_+ \to R$ is the density function over s. By Theorem 2, if the firm thinks that it is less likely that the rival will launch early, in the sense that there is an upward MLR shift in the density function, then it will

decide on a later launch date.

Note that we impose no restrictions on the function $w : R_+ \to R$, which gives the firm's profit should it be anticipated by its rival at time s. If w is an increasing function of s, then one can check that $\{\pi(\cdot, s)\}_{s \in S}$ is an SCP-ordered family, but Theorem 2 gives us the desired conclusion without making this stronger assumption.

5. Comparing Information Structures[8]

Consider an agent who, as in the previous section, has to make a decision before the state of the world (s) is realized, where the set of possible states S is a subset of R. In this section we assume that the decision x is picked from a *finite* set X in R. Suppose that, before he makes his decision, the agent observes a signal z. This signal is potentially informative of the true state of the world; we refer to the collection $\{H(\cdot|s)\}_{s \in S}$, where $H(\cdot|s)$ is the distribution of the signal z conditional on s, as the *information structure* of the decision maker's problem. (Whenever convenient, we shall simply call this information structure H.) We assume that, for every s, $H(\cdot|s)$ admits a density function and has the compact interval Z as its support. We say that H is *MLR-ordered* if $H(\cdot|s'')$ is an MLR shift of $H(\cdot|s')$ whenever $s'' > s'$.

We assume that the agent has a prior distribution λ on S. We allow either of the following: (i) S is a compact interval and λ admits a density function with S as its support or (ii) S is finite and λ has S as its support.

The agent's *decision rule* (under H) is a map from Z to X. We denote his posterior distribution (on S) upon observing z by λ_H^z; so the agent with a decision rule $\phi : Z \to X$ will have an ex ante utility given by

$$\mathcal{U}(\phi, H, \lambda) = \int_{z \in Z} \left[\int_{s \in S} u(\phi(z), s) d\lambda_H^z \right] dM_{H,\lambda} = \int_{Z \times S} u(\phi(z), s) \, dJ_{H,\lambda}$$

where $M_{H,\lambda}$ is the marginal distribution of z and $J_{H,\lambda}$ the joint distribution of (z, s) given H and λ. A decision rule $\hat{\phi} : Z \to X$ that maximizes the agent's (posterior)

---

[8]We are very grateful to Ian Jewitt for introducing us to the literature in this section and the next and for extensive discussions.

expected utility at each realized signal is called an *H-optimal* decision rule. We denote the agent's ex ante utility using such a rule by $\mathcal{V}(H, \lambda, u)$.

Consider now an alternative information structure given by the collection $\{G(\cdot|s)\}_{s \in S}$; we assume that $G(\cdot|s)$ admits a density function and has the compact interval Z as its support. What conditions will guarantee that the information structure H is more favorable than G in the sense of offering the agent a higher ex ante utility; in other words, how can we guarantee that $\mathcal{V}(H, \lambda, u) \geq \mathcal{V}(G, \lambda, u)$?

It is well known that this holds if H is *more informative* than G according to the criterion developed by Blackwell (1953); furthermore, this criterion is also necessary if one does not impose significant restrictions on u (see Blackwell (1953) or, for a recent textbook treatment, Gollier (2001)). We wish instead to consider the case where a significant restriction *is* imposed on u; specifically, we assume that $\{u(\cdot, s)\}_{s \in S}$ is an IDO family. We show that, in this context, a different notion of informativeness due to Lehmann (1988) is the appropriate concept.[9]

Our assumptions on H guarantee that, for any s, $H(\cdot|s)$ admits a density function with support Z; therefore, for any (z, s) in $Z \times S$, there exists a unique element in Z, which we denote by $T(z, s)$, such that $H(T(z, s)|s) = G(z|s)$. We say that H is *more accurate* than G if T is an increasing function of s.[10] Our goal in this section is to prove the following result.

THEOREM 3: *Suppose $\{u(\cdot, s)\}_{s \in S}$ is an IDO family, G is MLR-ordered, and λ is the agent's prior distribution on S. If H is more accurate than G, we obtain*

$$\mathcal{V}(H, \lambda, u) \geq \mathcal{V}(G, \lambda, u). \tag{12}$$

[9]For a discussion of the advantages of Lehmann's concept over Blackwell's, see Lehmann (1988) and Persico (2000). Some papers with economic applications of Lehmann's concept of informativeness are Persico (2000), Athey and Levin (2001), Levin (2001), Bergemann and Valimaki (2002), and Jewitt (2006). Athey and Levin's paper also explores other related concepts of informativeness and their relationship with the payoff functions.

[10]The concept is Lehmann's; the term *accuracy* follows Persico (2000).

This theorem generalizes a number of earlier results. Lehmann (1988) establishes a special case of Theorem 3 in which $\{u(\cdot, s)\}_{s \in S}$ is a QCIP family. Persico (1996) has a version of Theorem 3 in which $\{u(\cdot, s)\}_{s \in S}$ is an SCP family, but he requires the optimal decision rule to vary smoothly with the signal, a property that is not generally true without the *sufficiency* of the first order conditions for optimality. A proof of Theorem 3 for the general SCP case can be found in Jewitt's (2006) unpublished notes.[11]

To prove Theorem 3, we first note that if G is MLR-ordered, then the family of posterior distributions $\{\lambda_H^z\}_{z \in Z}$ is also MLR-ordered, i.e., if $z'' > z'$ then $\lambda_H^{z''}$ is an MLR shift of $\lambda_H^{z'}$.[12] Since $\{u(\cdot, s)\}_{s \in S}$ is an IDO family, Theorem 3 guarantees that the G-optimal decision rule can be chosen to be increasing with z. Therefore, Theorem 3 is valid if we can show that for *any* increasing decision rule $\psi : Z \to X$ under G there is a rule $\phi : Z \to X$ under H that gives a higher ex ante utility, i.e.,

$$\int_{Z \times S} u(\phi(z), s) dJ_{H,\lambda} \geq \int_{Z \times S} u(\psi(z), s) dJ_{G,\lambda}.$$

This inequality in turn follows from aggregating (across s) the inequality (13) below.

PROPOSITION 5: *Suppose $\{u(\cdot, s)\}_{s \in S}$ is an IDO family and H is more accurate than G. Then for any increasing decision rule $\psi : Z \to X$ under G, there is an increasing decision rule $\phi : Z \to X$ under H such that, at each state s, the distribution of utility induced by ψ and $H(\cdot|s)$ first order stochastically dominates the distribution of utility induced by ϕ and $G(\cdot|s)$. Consequently, at each state s,*

$$\int_{z \in Z} u(\phi(z), s) dH(z|s) \geq \int_{z \in Z} u(\psi(z), s) dG(z|s). \tag{13}$$

[11] However, there is at least one sense in which it is not correct to say that Theorem 3 generalizes Lehmann's result. The criterion employed by us here (and indeed by Persico (1996) and Jewitt (2006) as well) - comparing information structures with the ex ante utility - is different from the criterion Lehmann used. We return to this issue in the next section (specifically, Corollary 1), where we compare information structures using precisely the same criterion as Lehmann's.

[12] This is not hard to prove; indeed, the two properties are equivalent.

(At a given state s, a decision rule ρ and a distribution on z induces a distribution of utility in the following sense: for any measurable set U of R, the probability of $\{u \in U\}$ equals the probability of $\{z \in Z : u(\rho(z), s) \in U\}$. So it is meaningful to refer, as this proposition does, to the distribution of utility at each s.)

Our proof of Proposition 5 requires the following lemma.

LEMMA 3: *Suppose $\{u(\cdot, s)\}_{s \in S}$ is an IDO family and H is more accurate than G. Then for any increasing decision rule $\psi : Z \to X$ under G, there is an increasing decision rule $\phi : Z \to X$ under H such that, for all (z, s),*

$$u(\phi(T(z, s)), s) \geq u(\psi(z), s). \tag{14}$$

Proof: We shall only demonstrate here the way in which we construct ϕ from ψ; the fact that it is an *increasing* rule is proved in Appendix B. We shall confine ourselves here to the case where S is a compact interval; the proof can be modified in an obvious way to deal with the case where S is finite.

For every $\bar{t}$ in Z and s in S, there ia a unique $\bar{z}$ in Z such that $\bar{t} = T(\bar{z}, s)$. This follows from the fact that $G(\cdot|s)$ is a strictly increasing continuous function (since it admits a density function with support Z). We write $\bar{z} = \tau(\bar{t}, s)$; note that because T is increasing in both its arguments, the function $\tau : Z \times S \to Z$ is decreasing in s. Note also that (14) is equivalent to

$$u(\phi(t), s) \geq u(\psi(\tau(t, s)), s). \tag{15}$$

We will now show $\phi(t)$ may be chosen to satisfy (15). We may partition $S = [s^*, s^{**}]$ into the sets $S_1, S_2,..., S_M$, where M is odd, with the following properties: (i) if $m > n$, then any element in S_m is greater than any element in S_n; (ii) whenever m is odd, S_m is a singleton, with $S_1 = \{s^*\}$ and $S_M = \{s^{**}\}$; (iii) when m is even, S_m is an open interval; (iv) for any s' and s'' in S_m, we have $\psi(\tau(t, s')) = \psi(\tau(t, s''))$; and (v) for s'' in S_m and s' in S_n such that $m > n$, $\psi(\tau(t, s')) \geq \psi(\tau(t, s''))$.

21

In other words, we have partitioned S into finitely many sets, so that within each set, the same action is taken under ψ and the actions are decreasing. This is possible because X is finite, ψ is increasing in the signal and τ is decreasing in s. We denote the action taken in S_m by ψ_m, with $\psi_1 \geq \psi_2 \geq \psi_3 \geq ... \geq \psi_M$. Establishing (15) involves finding $\phi(t)$ such that

$$u(\phi(t), s_m) \geq u(\psi_m, s_m) \text{ for any } s_m \in S_m; \, m = 1, 2, ..., M. \qquad (16)$$

In the interval $[\psi_2, \psi_1]$, we pick the largest action $\hat{\phi}_2$ that maximizes $u(\cdot, s^*)$ in that interval. By the IDO property,

$$u(\hat{\phi}_2, s_m) \geq u(\psi_m, s_m) \text{ for any } s_m \in S_m; \, m = 1, 2. \qquad (17)$$

Recall that S_3 is a singleton; we call that element s_3. The action $\hat{\phi}_4$ is chosen to be the largest action in the interval $[\psi_4, \hat{\phi}_2]$ that maximizes $u(\cdot, s_3)$. Since ψ_3 is in that interval, we have $u(\hat{\phi}_4, s_3) \geq u(\psi_3, s_3)$. Since $u(\hat{\phi}_4, s_3) \geq u(\psi_4, s_3)$, the IDO property guarantees that $u(\hat{\phi}_4, s_4) \geq u(\psi_4, s_4)$ for any s_4 in S_4. Using the IDO property again (specifically, Lemma 2), we have $u(\hat{\phi}_4, s_m) \geq u(\hat{\phi}_2, s_m)$ for s_m in S_m ($m = 1, 2$) since $u(\hat{\phi}_4, s_3) \geq u(\hat{\phi}_2, s_3)$. Combining this with (17), we have found $\hat{\phi}_4$ in $[\psi_4, \hat{\phi}_2]$ such that

$$u(\hat{\phi}_4, s_m) \geq u(\psi_m, s_m) \text{ for any } s_m \in S_m; \, m = 1, 2, 3, 4. \qquad (18)$$

We can repeat the procedure finitely many times, at each stage choosing $\hat{\phi}_{m+1}$ (for m odd) as the largest element maximizing $u(\cdot, s_m)$ in the interval $[\psi_{m+1}, \hat{\phi}_{m-1}]$, and finally, choosing $\hat{\phi}_{M+1}$ as the largest element maximizing $u(\cdot, s^{**})$ in the interval $[\psi_M, \hat{\phi}_{M-1}]$. It is clear that $\phi(t) = \hat{\phi}_M$ will satisfy (16). QED

Proof of Proposition 5: Let $\tilde{z}$ denote the random signal received under information structure G and let $\tilde{u}^G$ denote the (random) utility achieved when the decision rule ψ is used. Correspondingly, we denote the random signal received under H by $\tilde{t}$, with $\tilde{u}^H$ denoting the utility achieved by the rule ϕ, as constructed in Lemma 3. Observe

that for any fixed utility level u' and at a given state s',

$$
\begin{aligned}
\Pr[\tilde{u}^H \le u'|s = s'] & = \Pr[u(\phi(\tilde{t}), s') \le u'|s = s'] \\
& = \Pr[u(\phi(T(\tilde{z}, s')), s') \le u'|s = s'] \\
& \le \Pr[u(\psi(\tilde{z}), s') \le u'|s = s'] \\
& = \Pr[\tilde{u}^G \le u'|s = s']
\end{aligned}
$$

where the second equality comes from the fact that, conditional on $s = s'$, the distribution of $\tilde{t}$ coincides with that of $T(\tilde{z}, s')$, and the inequality comes from the fact that $u(\phi(T(z, s')), s') \ge u(\psi(z), s')$ for all z (by Lemma 3).

Finally, the fact that, given the state, the conditional distribution of $\tilde{u}^H$ first order stochastically dominates $\tilde{u}^G$ means that the conditional mean of $\tilde{u}^H$ must also be higher than that of $\tilde{u}^G$.   QED

As a simple application of Theorem 3, consider again Example 1, where a firm has to decide on its production capacity before the state of the world is realized. Suppose that before it makes its decision, the firm receives a signal z from the information structure G. Provided G is MLR-ordered, we know that the posterior distributions (on S) will also be MLR-ordered (in z). It follows from Theorem 2 that a higher signal will cause the firm to decide on a higher capacity. Assuming the firm is risk neutral, its ex ante expected profit is $\mathcal{V}(\lambda, G, \Pi)$, where λ is the firm's prior on S. Theorem 3 tells us that a more accurate information structure H will lead to a higher ex ante expected profit; the difference $\mathcal{V}(\lambda, , H, \Pi) - \mathcal{V}(\lambda, , G, \Pi)$ represents what the firm is willing to spend for the more accurate information structure.

Our next example is a less straightforward application of Proposition 5.

Example 6. There are N investors, with investor i having wealth $w_i > 0$ and the strictly increasing Bernoulli utility function v_i. These investors place their wealth with a manager who has to decide on an investment policy; specifically, the manger must allocate the total pool of funds $W = \sum_{i=1}^{N} w_i$ between a risky asset, with return s in state s, and a safe asset with return $r > 0$. Denoting the fraction invested

in the risky asset by x, investor i's utility (as a function of x and s) is given by $u_i(x, s) = v_i((xs + (1 - x)r)w_i)$. It is easy to see that $\{u_i(\cdot, s)\}_{s \in S}$ is an IDO family. Indeed, we can say more: for $s > r$, $u_i(\cdot, s)$ is strictly increasing in x; for $s = r$, $u_i(\cdot, r)$ is the constant $v_i(rw_i)$; and for $s < r$, $u_i(\cdot.s)$ is strictly decreasing in x.

Before she makes her portfolio decision, the manager receives a signal z from some information structure G. She employs the decision rule ψ, where $\psi(z)$ (in $[0, 1]$) is the fraction of W invested in the risky asset. We assume that ψ is increasing in the signal. (We shall justify this assumption in the next section.) Suppose that the manager now has access to a superior information structure H. By Proposition 5, there is an increasing decision rule ϕ under H such that, at any state s, the distribution of investor k's utility under H and ϕ first order stochastically dominates the distribution of k's utility under G and ψ. In particular, (13) holds for $u = u_k$. Aggregating across states we obtain $\mathcal{U}_k(\phi, H, \lambda_k) \geq \mathcal{U}_k(\psi, G, \lambda_k)$, where λ_k is investor k's (subjective) prior; in other words, k's ex ante utility is higher with the new information structure and the new decision rule.

But even more can be said because, for any other investor i, $u_i(\cdot, s)$ is a strictly increasing transformation of $u_k(\cdot, s)$, i.e., there is a strictly increasing function f such that $u_i = f \circ u_k$. It follows from Proposition 5 that (13) is true, not just for $u = u_k$ but for $u = u_i$. Aggregating across states, we obtain $\mathcal{U}_i(\phi, H, \lambda_i) \geq \mathcal{U}_i(\psi, G, \lambda_i)$, where λ_i is investor i's prior.

To summarize, we have shown the following: though different investors may have different attitudes towards risk aversion and different priors, the greater accuracy of H compared to G allows the manager to implement a new decision rule that gives greater ex ante utility to *every* investor.

Finally, we turn to the following question: how important is the accuracy criterion to the results in this section? For example, we may wonder if the conclusion in Theorem 3 is, in a sense, too strong. Theorem 3 tells us that when H is more

accurate than G, it gives the agent a higher ex ante utility for any prior that he may have on S. This raises the possibility that the accuracy criterion may be weakened if we only wish H to give a higher ex ante utility than G for a *particular* prior. However, this is not the case, as the next result shows.

PROPOSITION 6: *Let S be finite, and H and G two information structures on S. If (12) holds at a given prior λ^* which has S as its support and for any SCP family $\{u(\cdot, s)\}_{s \in S}$, then (12) holds at any prior λ which has S as its support and for any SCP family.*

Proof: Given a prior λ with S as its support, and given the SCP family $\{u(\cdot, s)\}_{s \in S}$, we define the family $\{\tilde{u}(\cdot, s)\}_{s \in S}$ by $\tilde{u}(x, s) = [\lambda(s)/\lambda^*(s)]u(x, s)$. The ex ante utility of the decision rule ϕ under H, when the agent's utility is $\tilde{u}$, may be written as

$$\tilde{\mathcal{U}}(\phi, H, \lambda^*) = \sum_{s \in S} \lambda^*(s) \int \tilde{u}(\phi(z), s) dH(z|s).$$

Clearly,

$$\mathcal{U}(\phi, H, \lambda) \equiv \sum_{s \in S} \lambda(s) \int u(\phi(z), s) dH(z|s) = \tilde{\mathcal{U}}(\phi, H, \lambda^*).$$

From this, we conclude that

$$\mathcal{V}(H, \lambda, u) = \mathcal{V}(H, \lambda^*, \tilde{u}). \tag{19}$$

Crucially, the fact that $\{u(\cdot, s)\}_{s \in S}$ is an SCP family, guarantees that $\{\tilde{u}(\cdot, s)\}_{s \in S}$ is also an SCP family. By assumption, $\mathcal{V}(H, \lambda^*, \tilde{u}) \geq \mathcal{V}(G, \lambda^*, \tilde{u})$. Applying (19) to both sides of this inequality, we obtain $\mathcal{V}(H, \lambda, u) \geq \mathcal{V}(G, \lambda, u)$. QED

Loosely speaking, this result says that if we wish to have ex ante utility comparability for *any* SCP family (or, even more strongly, any IDO family), then fixing the prior does not lead to a weaker criterion of informativeness. A weaker criterion can only be obtained if we fix the prior *and* require ex ante utility comparability for a smaller class of utility families.[13]

---

[13]This possibility is explored in Athey and Levin (2001).

To construct a converse to Theorem 3, we assume that there are two states and two actions and that the actions are *non-ordered with respect to u* in the sense that x_1 is the better action in state s_1 and x_2 the better action in s_2, i.e., $u(x_1, s_1) > u(x_2, s_1)$ and $u(x_1, s_2) < u(x_2, s_2)$. This condition guarantees that information on the state is potentially useful; if it does not hold, the decision problem is clearly trivial since either x_1 or x_2 will be unambiguously superior to the other action. Note also that the family $\{u(\cdot, s_1), u(\cdot, s_2)\}$ is an IDO family. We have the following result.

PROPOSITION 7: *Suppose that $S = \{s_1, s_2\}$, $X = \{x_1, x_2\}$, and that the actions are non-ordered with respect to u. If H is MLR-ordered and* not *more accurate than G, then there is a prior $\bar{\lambda}$ on S such that $\mathcal{V}(H, \bar{\lambda}, u) < \mathcal{V}(G, \bar{\lambda}, u)$.*

Proof: Since H is not more accurate than G, there is $\bar{z}$ and $\bar{t}$ such that

$$G(\bar{z}|s_1) = H(\bar{t}|s_1) \ \text{ and } \ G(\bar{z}|s_2) < H(\bar{t}|s_2). \tag{20}$$

Given any prior λ, and with the information structure H, we may work out the posterior distribution and the posterior expected utility of any action after receipt of a signal.

We claim that there is a prior $\bar{\lambda}$ such that, action x_1 maximizes the agent's posterior expected utility after he receives the signal $z < \bar{t}$ (under H), and action x_2 maximizes the agent's posterior expected utility after he receives the signal $z \geq \bar{t}$. This result follows from the assumption that H is MLR-ordered and is proved in Appendix B.

Therefore, the decision rule ϕ such that $\phi(z) = x_1$ for $z < \bar{t}$ and $\phi(z) = x_2$ for $z \geq \bar{t}$ maximizes the agent's ex ante utility, i.e.

$$\begin{aligned}
\mathcal{V}(H, \bar{\lambda}, u) \ &= \ \mathcal{U}(\phi, H, \bar{\lambda}) \\
&= \ \bar{\lambda}(s_1) \left\{ u(x_1|s_1)H(\bar{t}|s_1) + u(x_2|s_1)[1 - H(\bar{t}|s_1)] \right\} \\
&\quad + \bar{\lambda}(s_2) \left\{ u(x_1|s_2)H(\bar{t}|s_2) + u(x_2|s_2)[1 - H(\bar{t}|s_2)] \right\}.
\end{aligned}$$

Now consider the decision rule ψ under G given by $\psi(z) = x_1$ for $z < \bar{z}$ and

$\psi(z) = x_2$ for $z \geq \bar{z}$. We have

$$
\begin{aligned}
\mathcal{U}(\psi, G, \bar{\lambda}) = {} & \bar{\lambda}(s_1) \left\{ u(x_1|s_1)G(\bar{z}|s_1) + u(x_2|s_1)[1 - G(\bar{z}|s_1)] \right\} \\
& + \bar{\lambda}(s_2) \left\{ u(x_1|s_2)G(\bar{z}|s_2) + u(x_2|s_2)[1 - G(\bar{z}|s_2)] \right\}.
\end{aligned}
$$

Comparing the expressions for $\mathcal{U}(\psi, G, \bar{\lambda})$ and $\mathcal{U}(\phi, H, \bar{\lambda})$, bearing in mind (20), and the fact that x_2 is the optimal action in state s_2, we see that

$$
\mathcal{U}(\psi, G, \bar{\lambda}) > \mathcal{U}(\phi, H, \bar{\lambda}) = \mathcal{V}(H, \bar{\lambda}, u).
$$

Therefore, $\mathcal{V}(G, \bar{\lambda}, u) > \mathcal{V}(H, \bar{\lambda}, u)$. QED

6. Statistical Decision Theory

Much of what is known and used in economics on comparative information have their origins in statistical decision theory, so it is appropriate for us to re-present and further develop the results of the last section in that context.[14] The main result in this section is a complete class theorem, which says that, in a certain sense and under certain conditions (which we will make precise), the statistician need only employ increasing decision rules.

We assume that the statistician conducts an experiment in which she observes the realization of a random variable (i.e., the outcome of the experiment) that takes values in a compact interval Z in R. The distribution of this random variable depends on the state s; we denote this distribution by $H(\cdot|s)$ and assume that it admits a density function and has Z as its support. The state s is chosen by Nature from a set S, also contained in R. As in the previous section, S may either be a compact interval or a finite set of points. Unlike the last section, we now adopt the perspective of a classical rather than Bayesian statistician, so we do not assume at the outset the existence of a (prior) probability distribution on S.

---

[14]For an introduction to statistical decision theory, see Blackwell and Girshik (1954) or Berger (1985).

Upon observing the experiment's outcome, the statistician takes a decision from a finite set X contained in R; formally, a decision function (or rule) is a measurable map ϕ from Z to X. Associated to each action and state is a loss; the *loss function* L maps $X \times S$ to R. Let D be the set of all decision functions. This experiment, which we shall call H, has the *risk function* $\mathcal{R}_H : D \times S \rightarrow R$, given by

$$\mathcal{R}_H(\phi, s) = \int_{z \in Z} L(\phi(z), s) dH(z|s).$$

For a classical statistician, the risk function $\mathcal{R}_H$ is the central concept with which decisions and experiments are compared. Note that our assumptions guarantee that this function is well defined.

Comparison of Experiments

Suppose now that the statistician may conduct another experiment G that also has outcomes in Z. At each state s, the distribution $G(\cdot|s)$ admits a density function and has Z as its support. We denote G's risk function by $\mathcal{R}_G$. Our first result of this section is a re-statement of Proposition 5 in the context of statistical decisions.

PROPOSITION 8: *Suppose $\{-L(\cdot, s)\}_{s \in S}$ is an IDO family and H is more accurate than G. Then for any increasing decision rule $\psi : Z \rightarrow X$ under G, there is an increasing decision rule $\phi : Z \rightarrow X$ under H such that*

$$\mathcal{R}_H(\phi, s) \leq \mathcal{R}_G(\psi, s) \quad \text{at each } s \in S. \tag{21}$$

This proposition considers the case when $\{-L(\cdot, s)\}_{s \in S}$ is an IDO family. It says that for any increasing decision rule employed in experiment G, there is a decision rule under H which gives lower risk at every possible state. For this result to be interesting, we must explain why allowing the statistician to employ other decision rules under G (besides increasing rules) is of no use to her. If we can show this, then experiment G is indeed superior to H (as measured by risk).

We have already implicitly given (in the previous section) the justification for the focus on increasing decision rules in the case of a Bayesian statistician. The Bayesian will have a prior on S, given by the probability distribution λ, and is interested in decision rules ψ that minimize the expected or *Bayes risk*, which is

$$\int_{s \in S} \mathcal{R}_G(\psi, s) d\lambda = \int_{z \in Z} \left[\int_{s \in S} L(\psi(z), s) d\lambda_G^z \right] dM_{G,\lambda}$$

where λ_G^z is the posterior distribution of s given z and $M_{H,\lambda}$ is the marginal distribution of z. Since $\{G(\cdot|z)\}_{z \in Z}$ - and thus $\{\lambda_G^z\}_{z \in Z}$ - is MLR-ordered (in z) and $\{-L(\cdot, s)\}_{s \in S}$ is an IDO family, ψ can indeed be chosen to be increasing.[15] Thus, when presented with the alternative experiments H and G obeying the conditions of Proposition 6, the Bayesian statistician will certainly prefer H since it follows immediately from (21) that H gives a lower Bayes risk.

For the classical statistician who chooses not to use Bayes risk as her criterion, a different justification must be given for the focus on increasing decision rules.

The completeness of increasing decision rules

We confine our attention to experiment G. The decision rule ψ is said to be *at least as good as* another decision $\tilde{\psi}$, if $\mathcal{R}_G(\psi, s) \leq \mathcal{R}_G(\tilde{\psi}, s)$ for all s in S. A subset D' of decision rules forms an *essentially complete class* if for any decision rule $\tilde{\psi}$, there is a rule ψ in D' that is at least as good as $\tilde{\psi}$. The following is the main result of this section.

THEOREM 4: *Suppose $\{-L(\cdot, s)\}_{s \in S}$ is an IDO family and G is MLR-ordered. Then the increasing decision rules form an essentially complete class.*

Theorem 4 is an example of a *complete class theorem*; it generalizes the complete class theorem of Karlin and Rubin (1956), which in turn generalizes Blackwell and Girshik (1954, Theorem 7.4.3). Karlin and Rubin (1956) establishes the essential

---

[15]This claim follows from Theorem 2. Note that our argument here is completely analogous to the one we used to establish Theorem 3.

completeness of increasing decision rules under the assumption that $\{L(\cdot, s)\}_{s \in S}$ is a family of quasiconvex functions, with $\operatorname{argmin}_{x \in X} L(x, s)$ increasing with s. Clearly, this is a special case of our assumption that $\{-L(\cdot, s)\}_{s \in S}$ forms an IDO family.[16]

For the classical statistician, Theorem 4 provides the justification for the restriction to increasing decision rules in Proposition 8. Combining these two results tells us that when an experiment H is more accurate than G, then H is capable of achieving lower risks at all states. So we have discovered something about the classical statistician that we already know about her Bayesian counterpart: she too regards H as superior to G.

COROLLARY 1: *Suppose $\{-L(\cdot, s)\}_{s \in S}$ is an IDO family, H is more accurate than G, and G is MLR-ordered. Then for any decision rule $\psi : Z \to X$ under G, there is an increasing decision rule $\phi : Z \to X$ under H such that*

$$\mathcal{R}_H(\phi, s) \leq \mathcal{R}_G(\psi, s) \ \ at\ each\ s \in S. \tag{22}$$

Proof: If ψ is not increasing, then by Theorem 4, there is an increasing rule $\bar{\psi}$ that is at least as good as ψ. Proposition 8 in turn guarantees that there is a decision rule under H that is at least as good as $\bar{\psi}$. QED

Corollary 1 is a generalization of Lehmann (1988) which establishes a version of this result in the case where $\{-L(\cdot, s)\}_{s \in S}$ is a QCIP family. Note that Corollary 1 is the classical analog to the Bayesian Theorem 3.

Proof of Theorem 4: The idea of the proof is to show that the statistician who uses a strategy that is not increasing is in some sense debasing the information made available to her by G.

Having assumed that Z is a compact interval, we can assume, without further loss of generality, that $Z = [0, 1]$. Suppose that ψ is a decision rule (not necessarily

---

[16]It is worth pointing out that Theorem 4 is not known even for the case where $\{-L(\cdot, s)\}_{s \in S}$ forms an SCP family.

increasing) under G where the actions taken are exactly x_1, x_2,...x_n (arranged in increasing order). We construct a new experiment $\bar{G}$ along the following lines. For each s, $\bar{G}(0|s) = 0$ and $\bar{G}(k/n|s) = \Pr_G[\psi(z) \leq x_k|s]$ for $k = 1, 2, ..., n$, where the right side of the second equation refers to the probability of $\{z \in Z : \psi(z) \leq x_k\}$ under the distribution $G(\cdot|s)$. We define $t_k(s)$ as the unique element in Z that obeys $G(t_k(s)|s) = \bar{G}(k/n|s)$. (Note that $t_0(s) = 0$ for all s.) Any z in $((k-1)/n, k/n)$ may be written as $z = \theta\,[(k-1)/n] + (1-\theta)\,[k/n]$ for some θ in $(0,1)$; we define

$$\bar{G}(z|s) = G(\theta t_{k-1}(s) + (1-\theta)t_k(s)|s). \tag{23}$$

This completely specifies the experiment $\bar{G}$.

Define a new decision rule $\bar{\psi}$ by $\bar{\psi}(z) = x_1$ for z in $[0, 1/n]$; for $k \geq 2$, we have $\bar{\psi}(z) = x_k$ for z in $((k-1)/n, k/n]$. This is an increasing decision rule under $\bar{G}$. It is also clear from our construction of $\bar{G}$ and $\bar{\psi}$ that, at each state s, the distribution of utility induced by $\bar{G}$ and $\bar{\psi}$ equals the distribution of utility induced by G and ψ.

We claim that G is more accurate than $\bar{G}$. Provided this is true, Proposition 5 says that there is an increasing decision rule ϕ under G that is at least as good as $\bar{\psi}$ under $\bar{G}$, i.e., at each s, the distribution of utility induced by G and ϕ first order stochastically dominates that induced by $\bar{G}$ and $\bar{\psi}$. Since the latter coincides with the distribution of utility induced by G and ψ, the proof is complete.

That G is more accurate than $\bar{G}$ follows from the assumption that G is MLR-ordered. We prove this in Appendix B. 　　　　　　　　　　QED

It is clear from our proof of Theorem 4 that we can in fact give a sharper statement of that result; we do so below.

THEOREM 4*: *Suppose $\{-L(\cdot, s)\}_{s \in S}$ is an IDO family and G is MLR-ordered. Then for any decision rule $\psi : Z \rightarrow X$, there is increasing decision rule $\phi : Z \rightarrow X$ such that, at each s, the distribution of utility induced by G and ϕ first order stochastically dominates the distribution of utility induced by G and ψ.*

Example 6 continued. Recall that we assumed in this application that the man-

ager's decision rule under G, which is ψ, is increasing in the signal. Provided G is MLR-ordered, Theorem 4* provides a justification for this rule.

Let $\tilde{\psi}$ be a (not necessarily increasing) decision rule. Theorem 4* tells us that for some investor k, there is an increasing decision rule $\psi : Z \to X$ such that, at each s, the distribution of u_k induced by G and ψ first order stochastically dominates the distribution of u_k induced by G and $\tilde{\psi}$. This implies that

$$\int_{z \in Z} u_k(\phi(z), s) dG(z|s) \geq \int_{z \in Z} u_k(\tilde{\psi}(z), s) dG(z|s). \tag{24}$$

Aggregating this inequality across states, we obtain $\mathcal{U}_k(\psi, G, \lambda_k) \geq \mathcal{U}_k(\tilde{\psi}, G, \lambda_k)$, i.e., the increasing rule gives investor k a higher ex ante utility.

However, we can say more because, for any other investor i, $u_i(\cdot, s)$ is just an increasing transformation of $u_k(\cdot, s)$, i.e., there is a strictly increasing function f such that $u_i = f \circ u_k$. Appealing to Theorem 4* again, we see that (24) is true if u_k is replaced with u_i. Aggregating this inequality across states gives us $\mathcal{U}_i(\psi, G, \lambda_i) \geq \mathcal{U}_i(\tilde{\psi}, G, \lambda_i)$. In short, we have shown the following: any decision rule admits an increasing decision rule that (weakly) raises the ex ante utility of every investor. This justifies our assumption that the manager uses an increasing decision rule.

Appendix A

Our objective in this section is to prove Theorem 2 in the case where S is a finite set. Suppose that $S = \{s_1, s_2, ..., s_N\}$ and the agent's objective function is

$$F(x, \lambda) = \sum_{i=1}^{N} f(x, s_i) \lambda(s_i)$$

where $\lambda(s_i)$ is the probability of state s_i. We assume that $\lambda(s_i) > 0$ for all s_i. Let γ be another distribution with support S. We say that γ is an MLR-shift of λ if $\gamma(s_i)/\lambda(s_i)$ is increasing in i.

Proof of Theorem 2 for the case of finite S: Suppose $F(x'', \lambda) \geq F(x, \lambda)$ for x in $[x', x'']$. We denote $(f(x'', s_i) - f(x', s_i))\lambda(s_i)$ by a_i and define $A_k = \sum_{i=k}^{N} a_i$. By

32

assumption, $A_1 = F(x'', \lambda) - F(x', \lambda) \geq 0$; we claim that $A_k \geq 0$ for any k (this claim is analogous to (8) in the proof of Theorem 2).

Suppose instead that there is $M \geq 2$ such that

$$A_M = \sum_{i=M}^{N} \left(f(x'', s_i) - f(x', s_i) \right) \lambda(s_i) < 0.$$

As in the proof of Theorem 2 in the main part of the paper, we choose $\bar{x}$ that maximizes $f(\cdot, s_M)$ in $[x', x'']$. By the IDO property and Lemma 2, we have

$$f(\bar{x}, s_i) - f(x'', s_i) \geq 0 \text{ for } i \leq M \text{ and} \tag{25}$$

$$f(\bar{x}, s_i) - f(x', s_i) \geq 0 \text{ for } i \geq M. \tag{26}$$

(These inequalities (25) and (26) are analogous to (10) and (11) respectively.) Following the argument used in Theorem 2, these inequalities lead to

$$A_1 = \sum_{i=1}^{N} (f(x'', s_i) - f(x', s_i)) \lambda(s_i) < 0, \tag{27}$$

which is a contradiction. Therefore $A_M \geq 0$.

Denoting $\gamma(s_i)/\lambda(s_i)$ by b_i, we have

$$F(x'', \gamma) - F(x', \gamma) = \sum_{i=1}^{N} a_i b_i.$$

It is not hard to check that[17]

$$A_1 b_1 + \sum_{i=2}^{N} A_i (b_i - b_{i-1}) = \sum_{1=1}^{N} a_i b_i. \tag{28}$$

Since γ is an MLR shift of λ, $b_i - b_{i-1} \geq 0$ for all i and so $\sum_{i=2}^{N} A_i (b_i - b_{i-1}) \geq 0$. Thus (28) guarantees that $\sum_{1=1}^{N} a_i b_i \geq A_1 b_1$; in other words,

$$F(x'', \gamma) - F(x', \gamma) \geq \frac{\gamma(s_1)}{\lambda(s_1)} \left[F(x'', \lambda) - F(x', \lambda) \right]. \tag{29}$$

It follows that the left hand side is nonnegative (positive) if the right side is nonnegative (positive), as required by the IDO property.   QED

---

[17]This is just a discrete version of integration by parts.

Proof of Lemma 3 continued: It remains for us to show that ϕ is an increasing rule. We wish to compare $\phi(t')$ against $\phi(t)$ where $t' > t$. Note that the construction of $\phi(t)$ first involves partitioning S into subsets $S_1, S_2, ...S_M$ obeying properties (i) to (v). In particular, (v) says that the actions (under ψ) associated to any two points in a set S_m are the same; we denoted that action by ψ_m. To obtain $\phi(t')$, we first partition S into disjoints sets $S_1', S_2',..., S_L'$, where L is odd, with the partition satisfying properties (i) to (v). We denote the action associated to S_k' by ψ_k'. Any s in S belongs to some S_m and some S_k'. The important thing to note is that

$$\psi_k' \geq \psi_m. \tag{30}$$

This follows immediately from that fact that $t' > t$ and ψ is increasing.

The construction of $\phi(t')$ involves the construction of $\hat{\phi}_2'$, $\hat{\phi}_4'$, etc. The action $\hat{\phi}_2'$ is the largest action maximizing $u(\cdot, s^*)$ in the interval $[\psi_2', \psi_1']$. Comparing this with $\hat{\phi}_2$, which is the largest action maximizing $u(\cdot, s^*)$ in the interval $[\psi_2, \psi_1]$, we know that $\hat{\phi}_2' \geq \hat{\phi}_2$ since (following from (30)) $\psi_2' \geq \psi_2$ and $\psi_1' \geq \psi_1$.

By definition, $\hat{\phi}_4'$ is the largest action maximizing $u(\cdot, s_3')$ in $[\psi_4', \hat{\phi}_2']$, where s_3' refers to the unique element in S_3'. Let $\bar{m}$ be the largest odd number such that $s_{\bar{m}} \leq s_3'$. (Recall that $s_{\bar{m}}$ is the unique element in $S_{\bar{m}}$.) By definition, $\hat{\phi}_{\bar{m}+1}$ is the largest element maximizing $u(\cdot, s_{\bar{m}})$ in $[\psi_{\bar{m}+1}, \hat{\phi}_{\bar{m}-1}]$. We claim that $\hat{\phi}_4' \geq \hat{\phi}_{\bar{m}+1}$. This is an application of Theorem 1. If follows from the following: (i) $s_3' \geq s_{\bar{m}}$, so $u(\cdot, s_3') \succeq_I u(\cdot, s_{\bar{m}})$; (ii) the manner in which $\bar{m}$ is defined, along with (30), guarantees that $\psi_4' \geq \psi_{\bar{m}+1}$; and (iii) we know (from the previous paragraph) that $\hat{\phi}_2' \geq \hat{\phi}_2 \geq \hat{\phi}_{\bar{m}-1}$.

So we obtain $\hat{\phi}_4' \geq \hat{\phi}_{\bar{m}+1} \geq \phi(t)$. Repeating the argument finitely many times (on $\hat{\phi}_6'$ and so on), we obtain $\phi(t') \geq \phi(t)$. QED

Proof of Proposition 6 continued: It remains for us to show how $\bar{\lambda}$ is constructed. We denote the density function of $H(\cdot|s)$ by $h(\cdot|s)$. It is clear that since the actions

are non-ordered, we may choose $\bar{\lambda}(s_1)$ and $\bar{\lambda}(s_2)$ such that

$$\bar{\lambda}(s_1)h(\bar{t}|s_1)\left[u(x_1, s_1) - u(x_2, s_1)\right] = \bar{\lambda}(s_2)h(\bar{t}|s_2)\left[u(x_2|s_2) - u(x_1, s_2)\right]. \tag{31}$$

Re-arranging this equation, we obtain

$$\bar{\lambda}(s_1)h(\bar{t}|s_1)u(x_1|s_1) + \bar{\lambda}(s_2)h(\bar{t}|s_2)u(x_1|s_2) = \bar{\lambda}(s_1)h(\bar{t}|s_1)u(x_2|s_1) + \bar{\lambda}(s_2)h(\bar{t}|s_2)u(x_2|s_2).$$

Therefore, given the prior $\bar{\lambda}$, the posterior distribution after observing $\bar{t}$ is such that the agent is indifferent between actions x_1 and x_2.

Suppose the agent receives the signal $z < \bar{t}$. Since H is MLR-ordered, we have

$$\frac{h(z|s_1)}{h(\bar{t}|s_1)} \geq \frac{h(z|s_2)}{h(\bar{t}|s_2)}.$$

This fact, together with (31) guarantee that

$$\bar{\lambda}(s_1)h(z|s_1)\left[u(x_1, s_1) - u(x_2, s_1)\right] \geq \bar{\lambda}(s_2)h(z|s_2)\left[u(x_2|s_2) - u(x_1, s_2)\right].$$

Re-arranging this equation, we obtain

$$\bar{\lambda}(s_1)h(z|s_1)u(x_1|s_1) + \bar{\lambda}(s_2)h(z|s_2)u(x_1|s_2) \geq \bar{\lambda}(s_1)h(z|s_1)u(x_2|s_1) + \bar{\lambda}(s_2)h(z|s_2)u(x_2|s_2).$$

So, after observing $z < \bar{t}$, the (posterior) expected utility of action x_1 is greater than that of x_2. In a similar way, we can show that x_2 is the optimal action after observing a signal $z \geq \bar{t}$. QED

Proof of Theorem 4 continued: We denote the density function associated to the distribution $G(\cdot|s)$ by $g(\cdot|s)$. The probability of $Z_k = \{z \in Z : \psi(z) \leq x_k\}$ is given by $\int \mathbf{1}_{Z_k}(z)g(z|s)dz$, where $\mathbf{1}_{Z_k}$ is the indicator function of Z_k. By the definition of $\bar{G}$, we have

$$\bar{G}(k/n|s) = \int \mathbf{1}_{Z_k}(z)g(z|s)dz.$$

Recall that $t_k(s)$ is defined as the unique element that obeys $G(t_k(s)|s) = \bar{G}(k/n)$; equivalently,

$$\bar{G}(k/n|s) - G(t_k(s)|s) = \int \left[\mathbf{1}_{Z_k}(z) - \mathbf{1}_{[0,t_k(s)]}(z)\right] g(z|s)dz = 0. \tag{32}$$

The function W given by $W(z) = \mathbf{1}_{Z_k}(z) - \mathbf{1}_{[0,t_k(s)]}(z)$ has the following single-crossing type condition: $z > t_k(s)$, we have $W(z) \geq 0$ and for $z \leq t_k(s)$, we have $W(z) \leq 0$.[18] Let $s' > s$; since $G(\cdot|s)$ is MLR-ordered, $g(z|s')/g(z|s)$ is an increasing function of z. By a standard result (see, for example, Athey (2002, Lemma 5)), we have

$$\bar{G}(k/n|s') - G(t_k(s)|s') = \int \left[\mathbf{1}_{Z_k}(z) - \mathbf{1}_{[0,t_k(s)]}(z) \right] g(z|s')dz \geq 0. \qquad (33)$$

This implies that $t_k(s') \geq t_k(s)$.

To show that G is more accurate than $\bar{G}$, we require $T(z,s)$ to be increasing in s, where T is defined by $G(T(z,s)|s) = \bar{G}(z|s)$. For $z = k/n$, $T(z,s) = t_k(s)$, which we have shown is increasing in s. For z in the interval $((k-1)/n, k/n)$, recall (see (23)) that $\bar{G}(z)$ was defined such that $T(z,s) = \theta t_{k-1}(s) + (1-\theta)t_k(s)$. Since both t_{k-1} and t_k are increasing in s, $T(z,s)$ is also increasing in s. \hfill QED

REFERENCES

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

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

ATHEY, S., P. MILGROM, AND J. ROBERTS (1998): *Robust Comparative Statics.* (Draft Chapters.) http://www.stanford.edu/ athey/draftmonograph98.pdf

BERGER, J. O. (1985): *Statistical Decision Theory and Bayesian Analysis* New York: Springer-Verlag

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

---

[18]This property is related to but not the same as the single crossing property we have defined in this paper; Athey (2002) refers to this property as SC1 and the one we use as SC2.

BLACKWELL, D. AND M. A. GIRSHIK (1954): *Theory of Games and Statistical Decisions.* New York: Dover Publications

GOLLIER, C. (2001): *The Economics of Risk and Time.* Cambridge: MIT Press.

JEWITT, I. (2006): "Information Order in Decision and Agency Problems," Personal Manuscript.

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

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

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

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

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

ORMISTON, M. B. AND E. SCHLEE (1993): "Comparative Statics under uncertainty for a class of economic agents," *Journal of Economic Theory,* 61, 412-422.

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

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

QUAH, J. K.-H. AND B. STRULOVICI (2007): "Comparative Statics with the Interval Dominance Order II," Personal Manuscript.

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

TOPKIS, D. M. (1998): *Supermodularity and Complementarity.* Princeton: Princeton University Press.

VIVES, X. (1990): "Nash Equilibrium with Strategic Complementarities," *Journal of Mathematical Economics*, 19, 305-21.
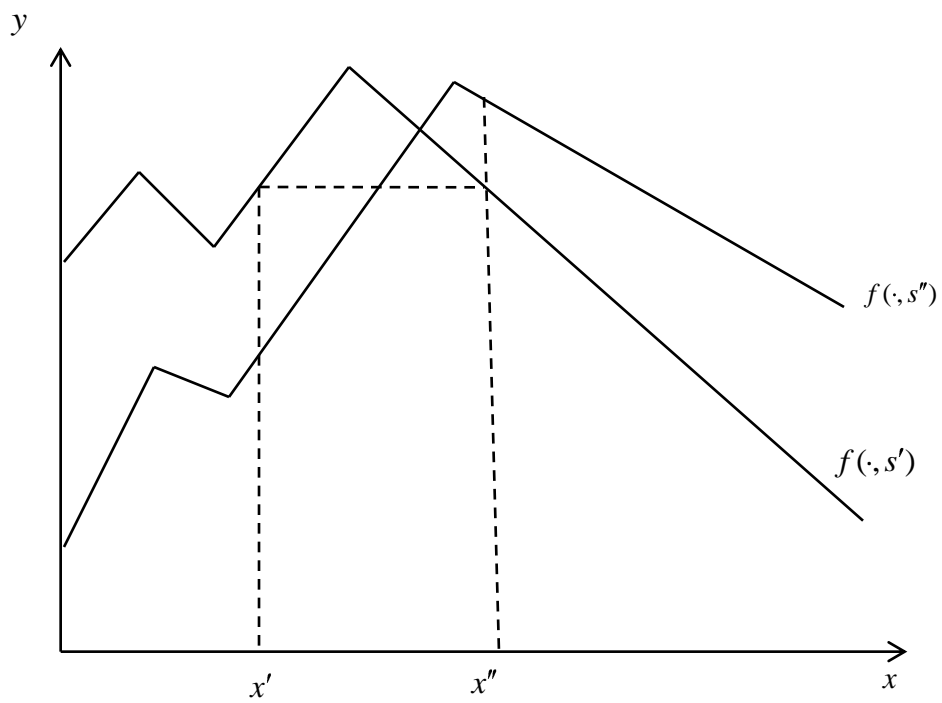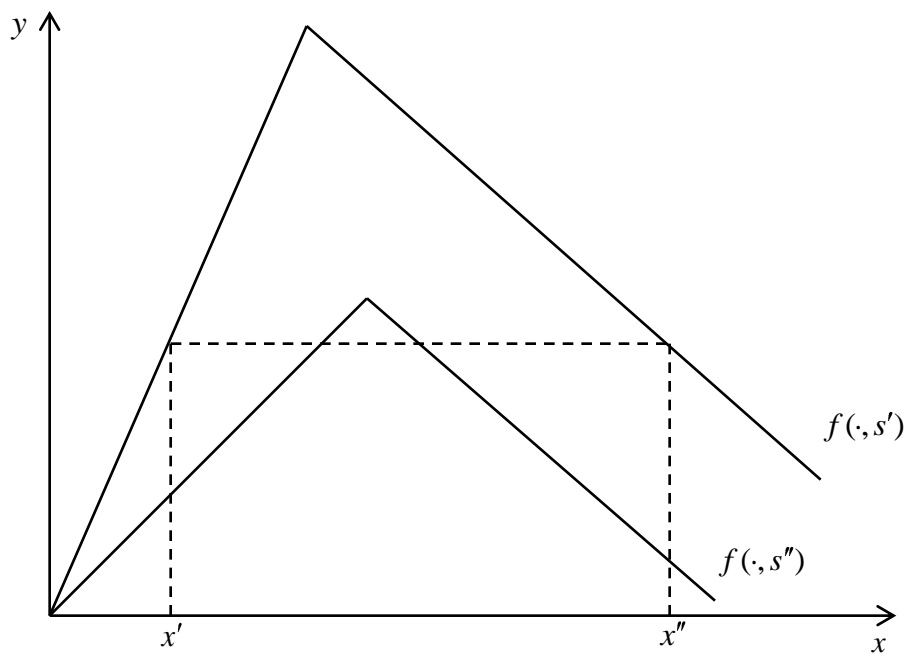
Figure 1

Figure 2