

Investment and Information Value for a Risk Averse Firm

Susan Athey, MIT

First Draft: July, 1998

This Draft: November, 2000

Abstract

This paper analyzes the problem faced by a risk-averse firm considering how much to invest in a risky project. The firm receives a signal about the value of the project. We derive necessary and sufficient conditions on the signal distribution such that (i) the agent's investment is nondecreasing in the realization of the signal, and (ii) different signals can be ranked according to their ex ante information value. Finally, we provide conditions under which it is possible to compare the incentive to acquire information across agents with different risk preferences, and we identify a class of utility functions for which agents who are less risk averse purchase more information.

JEL Classification Numbers: C44, C60, D81.

Keywords: Portfolio problem, value of information, stochastic orderings, comparative statics under uncertainty, monotone likelihood ratio order, monotone probability ratio order.

*Some of the comparative statics results in this paper originally appeared in an MIT working paper entitled "Comparative Statics Under Uncertainty: Single Crossing Properties and Log-Supermodularity," which has since been shortened to eliminate these results. I am grateful to Christian Gollier, Jonathan Levin, Paul Milgrom, John Roberts, and Ed Schlee for helpful discussions, and the National Science Foundation, graduate fellowship and Grants SBR-9631760 and SES-9983820, for financial support; I am also grateful to the Cowles Foundation at Yale University and the Hoover Institution for their hospitality and support. Corresponding address: MIT Department of Economics, E52-252C, 50 Memorial Drive, Cambridge, MA 02142. Email: athey@mit.edu, <http://web.mit.edu/athey/www/>.

1 Introduction

This paper analyzes the problem faced by a risk-averse firm who must choose an investment in a project with uncertain returns. The firm can potentially purchase a signal about the quality of the project. The paper consider three comparative statics questions. First, under what conditions on the information structure will higher realizations of the signal lead to higher investments on the part of the firm? Second, under what conditions can we ensure that one signal (i.e., one information source) will provide higher *ex ante* value than another for all such risk averse firms? Third, what conditions on preferences guarantee that one agent will be willing to pay more than another for better information?

It turns out that the answers to the three questions are closely related. Consider the first question. Formally, the firm faces uncertainty about the state of the world, W , with typical realization $\omega \in \Omega \subset \mathbb{R}$. The agent has a prior, $H(\cdot)$, and observes the realization of a partially informative signal X , with typical realization $x \in \mathcal{X} \subset \mathbb{R}$. The signal distribution, conditional on $W = \omega$, is given by $F_{X|W}(\cdot|\omega)$. The posterior $F_{W|X}(\cdot|x)$ is computed according to Bayes' rule. The firm's investment problem is given as follows:

$$\max_a \int u(\pi(a, \omega)) dF_{W|X}(\omega|x),$$

where u is a (typically concave) utility function, and π represents a return function which depends on the investment, $a \in \mathcal{A}$. The marginal returns to investment, $\frac{\partial}{\partial a}\pi(a, \omega)$, are nondecreasing in the state of the world (ω). The optimal policy is denoted $\alpha^*(x)$. In this context, the first comparative statics question concerns necessary and sufficient conditions on the signal distribution such that the agent's level of investment is nondecreasing in x .

For two extreme cases of assumptions on risk preferences, the answer to this question has been established in the existing literature. First, if the agent is risk neutral, the necessary and sufficient condition for investment to be nondecreasing is that x orders the posterior distribution $F_{W|X}(\cdot|x)$ according to First-Order Stochastic Dominance (FOSD).¹ Next suppose that risk preferences are unrestricted (not necessarily risk averse). In this case, a necessary and sufficient condition for the optimal choice of a to be nondecreasing in x is that x orders the posterior distribution by the Monotone Likelihood Ratio Order (MLR).² The MLR requires that the relative likelihood of high versus low realizations of W , $f_{W|X}(\omega_H|x)/f_{W|X}(\omega_L|x)$, is nondecreasing in x . Many authors have pointed out the disadvantages of the MLR order: it is very stringent, and it is equivalent to requiring

¹See Ormiston and Schlee (1992) or Athey (1999) for alternative proofs.

²See Ormiston and Schlee (1993) or Athey (forthcoming). More generally, this result holds imposing only the restriction that $\pi_a(a, \omega)$ crosses 0, at most once, from below, as a function of ω , rather than our stronger maintained assumption that π is supermodular (Athey, forthcoming).

that the distribution be ordered by FOSD conditional on any two-point set of states. Thus, a small change in probability mass affecting only two points in the distribution can upset the order.

Now consider incorporating risk aversion. Several important special cases of this problem have been considered in the literature. For example, the standard portfolio problem has $\pi(a, \omega) \equiv a\omega + (1 - a)\omega_0$, where ω_0 is the risk-free rate of return. Eeckhoudt and Gollier (1995) showed in the portfolio problem that a sufficient, *but not necessary*, condition on F for a risk-averse agent to increase her investment in response to a higher signal is that the posterior distribution, $F_{W|X}(\cdot|x)$, satisfies the “Monotone Probability Ratio Order” (MPR), which entails that the probability ratio for low versus high realizations of W , $F_{W|X}(\omega_H|x)/F_{W|X}(\omega_L|x)$, is nondecreasing in x .³ As expected, this ordering is more restrictive than FOSD but less restrictive than the MLR. Both the MPR and FOSD allow for some “averaging” over states and thus lead to more robust orderings.

Many other economic examples motivate a more general investment return function than that of the portfolio problem. In a generalized version of Sandmo’s (1971) model of a firm facing demand uncertainty, where $\pi(a, \omega) \equiv (P(a) + \omega)a - c(a)$, Milgrom (1994) shows that any comparative statics result that holds in the portfolio problem, also holds in the Sandmo model. This functional form is still fairly special. For example, we might like to generalize the model to allow for non-additive demand shocks, for example by allowing ω to make the demand curve less elastic.

Our first main result is that the finding from the portfolio problem can be generalized: if higher signals x increase the distribution in the MPR order, the optimal investment policy will be nondecreasing whenever the agent is risk averse. Furthermore, although the MPR is not a necessary condition for comparative statics in the portfolio problem, we show that it *is* a necessary condition when we allow for more general (supermodular) investment return functions π .

Now consider our second question, which concerns the agent’s preferences over different sources of information. For example, an investor might have access to “inside information” about asset returns, or she might purchase investment advice. A firm might engage in market research, or it might learn from experience or experimentation. Consumers might have access to different sources of information about product quality. Formally, in the ex ante information gathering problem, the firm chooses between information sources, indexed by a parameter θ . Each information source θ generates a signal X^θ , and after the agent observes the realization of this signal, she forms a posterior and chooses an action. Formally, the agent solves:

$$\max_{\theta \in \Theta} \int_{\mathcal{X}^\theta} \max_{a \in \mathcal{A}} \int_{\Omega} u(\pi(a, \omega)) dF_{W|X^\theta}(\omega|x) dF_{X^\theta}(x) - c(\theta).$$

³Gollier (1995) shows that a necessary condition for comparative statics in the portfolio problem is that the posterior distribution is ordered by “greater central riskiness.” We will not pursue that order here, because it does not apply directly to problems with a structure more general than the portfolio problem.

In this problem, we derive necessary and sufficient conditions on the joint distribution of signals and states, F_{W,X^θ} , such that higher values of θ increase the agent’s ex ante expected utility.

Our approach to analyzing the agent’s information preferences builds directly on Lehmann (1988) and Athey and Levin (2000) (henceforth AL). Lehmann (1988) and AL study the problem of deriving informativeness orders, when payoff functions are restricted to lie in some set U .⁴ These authors further restrict attention to “monotone decision problems.” These are decision problems where (for a given θ and U) all posteriors in the set $\{F_{W|X^\theta}(\cdot|x), x \in \mathcal{X}^\theta\}$ can be totally ordered such that, for *every* decision-maker in the class U under consideration, higher realizations of the signal lead to higher actions. Of course, it is just this comparative statics result that is the subject of the first question addressed in this paper.

Lehmann (1988) studies the particular case where U is the set of all payoffs that satisfy a single crossing property; Persico (1996) applies this order to several economic problems, including portfolio problems, but without fully exploiting the consequences of risk aversion. AL generalize Lehmann’s (1988) approach to allow for different classes of payoff functions. In particular, AL derive an information ordering for supermodular payoff functions, which would apply to the investment problem when the firm is risk neutral.

However, AL did not consider the special structure that arises in the case of a risk-averse firm. This paper provides a new information order for this case. Just as the order over posteriors for the risk averse firm, the MPR, is weaker than the MLR and stronger than FOSD, the information ranking for the risk-averse firm is weaker than Lehmann’s order and stronger than AL’s information ranking for supermodular payoff functions.

Our third question concerns conditions on agent preferences that imply that, if signals are ranked according to the information order identified above, one agent will be willing to pay more than another for a more informative signal. Notice that there is not a clear intuition for how risk aversion should affect willingness-to-pay for information. On the one hand, an agent who is more risk-averse might be willing to pay to avoid the uncertainty that comes with an uninformative signal. On the other hand, an agent who is more risk-averse may choose a policy that is less responsive to the realizations of signals, and thus may not find it as valuable to purchase better information.

Our approach to the comparative statics of information gathering builds on an approach taken

⁴The approach of Lehmann (1988) and AL contrasts with the approach taken by Blackwell (1951, 1953). Blackwell’s order is necessary and sufficient for *all* decision-makers to have a higher ex ante expected value from the decision problem. As many authors (see, for example, LeCam (1964), Lehmann (1988) and AL) have pointed out, Blackwell’s (1951) informativeness order is often too strong to be useful in applications. The approach of Lehmann (1988) and AL exploits additional structure that arises in economic problems, allowing for the ranking of many signals that are not comparable using Blackwell’s order.

by Persico (2000) for the case of single crossing functions, and generalized by AL to arbitrary sets of payoff functions U . This paper identifies (rather stringent) sufficient conditions on risk preferences so that one agent will purchase better information than another, and in particular, we identify a family of utility functions for which agents who are less risk averse are willing to pay more for information.

2 The Investment Problem

2.1 Stochastic Single Crossing and Comparative Statics

In this section, we introduce and characterize several stochastic orders that will be used throughout the paper. Let \mathcal{X} and Ω be compact subsets of \mathbb{R} , and let \mathcal{P} be the set of probability distributions on Ω . Let \mathcal{R} be the set of bounded, measurable functions mapping Ω to \mathbb{R} . For a given set $R \subset \mathcal{R}$, we define the set U^R as the set of functions with R -incremental returns. Formally,

$$U^R = \{u \mid u : \mathcal{X} \times \Omega \rightarrow \mathbb{R}, \text{ and } \forall x_H > x_L, u(x_H, \cdot) - u(x_L, \cdot) \in R\}.$$

For example, if R is the set of nondecreasing functions, U^R is the set of supermodular functions (with incremental returns to x that are bounded and measurable).

A function $f : \mathcal{X} \rightarrow \mathbb{R}$ is weak single crossing at x_0 if $f(x) \leq 0$ for all $x < x_0$ and $f(x) \geq 0$ for all $x > x_0$. The function is weak single crossing if there exists some x_0 such that $f(x)$ is single crossing at x_0 . Then, we have:⁵

Definition 1 *Let $P, Q \in \mathcal{P}$. Then Q dominates P in the stochastic single crossing order for R , if*

$$\int r(\omega)dP(\omega) \geq 0 \text{ implies } \int r(\omega)dQ(\omega) \geq 0 \text{ for all } r \in R.$$

Why are stochastic single crossing orderings useful? The following Lemma (which follows as an application of Shannon (1995)) shows that stochastic single crossing orders can be used to derive comparative statics results.

Lemma 1 *If, for all $x_H > x_L$, $F_{W|X}(\cdot|x_H)$ dominates $F_{W|X}(\cdot|x_L)$ in the stochastic single crossing order for R , then for all $g \in U^R$, there exists a nondecreasing selection $\alpha^*(x)$ from $\arg \max_a \int g(a, \omega)dF_{W|X}(\omega|x)$.*

⁵Given the assumptions we use in this paper, all of our results would also hold if we defined the stochastic single crossing order in terms of Milgrom and Shannon's (1994) single crossing property rather than the weak single crossing property. We use the weak single crossing property for simplicity.

Now let us define several important examples of stochastic single crossing orders, each corresponding to a different set of payoff functions R . We begin by defining the orders directly; subsequently, we will identify the sets R for which they are stochastic single crossing orders.

Suppose that P and Q are probability distributions, and suppose that $\text{supp}(P)=\text{supp}(Q)=\Omega$. First, consider FOSD. Formally, $Q \succ_{FOSD} P$ if $Q(\omega) \leq P(\omega)$ for all $\omega \in \Omega$.

Second, let μ be a measure on Ω such that P and Q are absolutely continuous with respect to μ with strictly positive densities p and q . P dominates Q according to single crossing of likelihood ratios at ω_0 , denoted $Q \succ_{LR-\omega_0} P$, if⁶

$$\frac{q(\omega)}{q(\omega_0)} - \frac{p(\omega)}{p(\omega_0)} \text{ is weak single crossing in } \omega \text{ at } \omega_0 \text{ a.e.-}\mu.$$

In other words, the likelihood of a state $\omega > (<)\omega_0$ relative to the likelihood of ω_0 is greater (smaller) with q than p . If $Q \succ_{LR-\omega_0} P$ for μ -almost all ω_0 , then $Q \succ_{LR} P$. In this language, requiring that x orders $F_{W|X}(\cdot|x)$ by MLR is equivalent to requiring that, for all $x_H > x_L$, $F_{W|X}(\cdot|x_H) \succ_{LR} F_{W|X}(\cdot|x_L)$. Equivalently, the density $f_{W|X}$ must be log-supermodular, so that $f_{W|X}(\omega|x^H)/f_{W|X}(\omega|x^L)$ is nondecreasing in ω .

Third, we can similarly define single crossing of probability ratios, replacing the density with the distribution: $Q \succ_{PR-\omega_0} P$ if

$$\frac{Q(\omega)}{Q(\omega_0)} - \frac{P(\omega)}{P(\omega_0)} \text{ is weak single crossing in } \omega \text{ at } \omega_0.$$

If $Q \succ_{PR-\omega_0} P$ for all ω_0 , then $Q \succ_{PR} P$. When $Q \succ_{PR} P$, then for every ω_0 , the likelihood of state ω_0 relative to the likelihood of all states $\omega < \omega_0$ is higher for Q than P . Requiring that x orders $F_{W|X}(\cdot|x)$ by MPR is equivalent to requiring that, for all $x_H > x_L$, $F_{W|X}(\cdot|x_H) \succ_{PR} F_{W|X}(\cdot|x_L)$. Equivalently, $F_{W|X}$ must be log-supermodular.

It can be verified directly that MPR is weaker than MLR (since log-supermodularity is preserved by integration) and stronger than FOSD (since the MPR requires that $F_{W|X}(\omega|x^H)/F_{W|X}(\omega|x^L)$ is nondecreasing in ω , and the upper bound of the ratio is 1). The following lemma provides alternative characterizations of the MLR, MPR, and FOSD, and the relationships between them.

Lemma 2 *Suppose that $\text{supp}(P)=\text{supp}(Q)=\Omega$. (i) $Q \succ_{FOSD} P$ if and only if $\int rdQ \geq \int rdP$ for all r nondecreasing. (ii) $Q(\cdot|W \in \{\omega_1, \omega_2\}) \succ_{FOSD} P(\cdot|W \in \{\omega_1, \omega_2\})$ for μ -almost all $\{\omega_1, \omega_2\} \in \Omega$, if and only if $Q \succ_{LR} P$. (iii) $Q(\cdot|W \leq \omega_0) \succ_{FOSD} P(\cdot|W \leq \omega_0)$ for all $\omega_0 \in \Omega$ if and only if $Q \succ_{PR} P$.*

⁶We say that a function $g : \Omega \rightarrow \mathbb{R}$ satisfies single crossing almost everywhere (a.e.- μ) if, for almost every (ω_L, ω_H) pair (using the measure induced by μ on Ω^2), g satisfies single crossing on $\{\omega_L, \omega_H\}$.

The idea of this Lemma can be stated in words: MLR is equivalent to FOSD conditional on any two-point set, while the MPR is equivalent to FOSD conditional on any lower interval of the form $\{\omega : \omega \leq \omega_0\}$. Parts (i) and (ii) of this Lemma are well-known, while part (iii) follows by checking the definitions. As a consequence of (i) and (iii), x orders $F_{W|X}(\cdot|x)$ by the MPR if and only if $\mathbb{E}_W[r(W)|x, W \leq \omega_0]$ is nondecreasing in x for all $\omega_0 \in \Omega$ and all r nondecreasing.

Note that because the MLR order must be checked for every pair of states, it is independent of the prior. More precisely, given a prior H with support equal to Ω and a signal distribution $F_{X|W}$, if x orders the posteriors $F_{W|X}(\cdot|x)$ by MLR, the same conclusion holds if H is replaced with any other prior distribution. In contrast, the MPR order might hold for the posteriors generated by a given prior H and a signal distribution $F_{X|W}$, while it might fail for the posteriors generated by another prior \tilde{H} and the *same* signal distribution $F_{X|W}$.⁷

Our next step is to define several classes of payoff functions of interest. R^{SC} is the set of all $r(\cdot)$ which are weak single crossing, while $R^{SC(\omega_0)}$ is the set of all $r(\cdot)$ which are weak single crossing at ω_0 ; R^{ND} is the set of all r nondecreasing; $R^{SCQ(\omega_0)}$ is the set of all $r(\cdot)$ which are single crossing and quasi-concave, where the function changes direction at ω_0 , while R^{SCQ} is the set of r which are in $R^{SCQ(\omega_0)}$ for some ω_0 . An important subset of $R^{SCQ(\omega_0)}$ is the set of all functions g such that, for some r nondecreasing, $g(\omega) \equiv \mathbf{1}_{\{\omega \leq \omega_0\}}(\omega) \cdot r(\omega)$. Clearly, $R^{ND} \subset R^{SCQ} \subset R^{SC}$.

The following Lemma characterizes the stochastic single crossing orders for these payoff functions.

Lemma 3 *Suppose that $\text{supp}(P) = \text{supp}(Q) = \Omega$.*

- (i) *Q dominates P in the stochastic single crossing order for R^{ND} , if and only if $Q \succ_{FOSD} P$.*
- (ii) *Suppose that either Ω is finite, or else P and Q have continuous, strictly positive densities with respect to a measure μ . (a) For $\omega_0 \in \Omega$, Q dominates P in the stochastic single crossing order for $R^{SC(\omega_0)}$, if and only if $Q \succ_{LR-\omega_0} P$. (b) Q dominates P in the stochastic single crossing order for R^{SC} , if and only if $Q \succ_{LR} P$.*
- (iii) (a) For $\omega_0 \geq \inf \Omega$, Q dominates P in the stochastic single crossing order for $R^{SCQ(\omega_0)}$, if and only if $Q \succ_{PR-\omega_0} P$. (b) Q dominates P in the stochastic single crossing order for R^{SCQ} , if and only if $Q \succ_{PR} P$.

The proof follows from Athey (forthcoming), Theorems 3 and 4, and Athey (1999).⁸

⁷However, under these conditions the MPR would hold for any prior distribution formed from H (using Bayes' rule) by conditioning on the event $W < a$.

⁸Athey (forthcoming) focuses on problems where R is the set of functions that satisfy Milgrom and Shannon's (1994) single crossing property, rather than the weak single crossing property used here (where r may become positive and then return to 0). In Athey (forthcoming), we may allow for $\text{supp}(Q)$ to be greater than $\text{supp}(P)$ (in the strong set

To connect part (iii) of Lemma 2 and part (iii) of Lemma 3, observe that if $r(\cdot)$ is nondecreasing, the function defined by $g(\omega) \equiv \mathbf{1}_{\{\omega \leq \omega_0\}}(\omega) \cdot r(\omega)$ is in $R^{SCQ(\omega_0)}$. Further, Lemma 2 implies that $Q \succ_{PR} P$ if and only if, for all r nondecreasing,

$$\frac{\int r(\omega) \cdot \mathbf{1}_{\{\omega \leq \omega_0\}}(\omega) dQ(\omega)}{Q(\omega_0)} \geq \frac{\int r(\omega) \cdot \mathbf{1}_{\{\omega \leq \omega_0\}}(\omega) dP(\omega)}{P(\omega_0)},$$

or, rearranging,

$$\int r(\omega) \cdot \mathbf{1}_{\{\omega \leq \omega_0\}}(\omega) dQ(\omega) \geq \frac{Q(\omega_0)}{P(\omega_0)} \cdot \int r(\omega) \cdot \mathbf{1}_{\{\omega \leq \omega_0\}}(\omega) dP(\omega).$$

This in turn implies

$$\int r(\omega) \cdot \mathbf{1}_{\{\omega \leq \omega_0\}}(\omega) dP(\omega) \geq 0 \implies \int r(\omega) \cdot \mathbf{1}_{\{\omega \leq \omega_0\}}(\omega) dQ(\omega) \geq 0.$$

In words, Q dominates P in the stochastic single crossing order for the set of functions of the form $\mathbf{1}_{\{\omega \leq \omega_0\}}(\omega) \cdot r(\omega)$, where r is nondecreasing, a subset of R^{SCQ} . Further, we can construct the counter-examples required for necessity in part (iii) of Lemma 3 using functions of the form $\mathbf{1}_{\{\omega \leq \omega_0\}}(\omega) \cdot r(\omega)$. The following characterization lemma, proved in the Appendix, builds on this logic.

Lemma 4 *Suppose that $\text{supp}(P) = \text{supp}(Q) = \Omega$. For $\omega_0 \geq \inf \Omega$, Q dominates P in the stochastic single crossing order for $R^{SCQ(\omega_0)}$ (equivalently, $Q \succ_{PR-\omega_0} P$) if and only if*

$$\int r(\omega) dQ(\omega) \geq \frac{Q(\omega_0)}{P(\omega_0)} \int r(\omega) dP(\omega) \text{ for all } r \in R^{SCQ(\omega_0)}. \quad (1)$$

2.2 Comparative Statics in the Investment Problem

Consider the investment problem of a risk-averse firm, as described in the Introduction, where the firm has access to the signal X and solves $\max_{a \in \mathcal{A}} \int u(\pi(a, \omega)) dF_{W|X}(\omega|x)$. In our analysis, we will refer to the following assumptions.

A1 For each $x \in \mathcal{X}$, $\text{supp}(F_{W|X}(\cdot|x)) = \Omega$.

A2 Either Ω is finite, or else, for each $x \in \mathcal{X}$, $F_{W|X}(\cdot|x)$ has a strictly positive and continuous density.

A3 u and π are C^2 , the derivatives are absolutely continuous, and \mathcal{A} , Ω , and \mathcal{X} are compact, convex subsets of \mathbb{R} .

order). When we consider the larger set R^{SC} , it can be shown that Q dominates P in the stochastic single crossing order for R^{SC} only if $\text{supp}(P) = \text{supp}(Q) = \Omega$. We maintain the assumption $\text{supp}(P) = \text{supp}(Q) = \Omega$ as a prerequisite rather than a conclusion in order to make it easier to state the definitions of the relevant ratio orders.

A4 For all θ , $\int u(\pi(a, \omega))dF_{W|X}(\omega|x)$ is quasi-concave in a and C^3 . Further, assume that for each x , the optimal choice of a is interior.

A5 $\pi(a, \omega)$ is nondecreasing in ω , and further, the expected value of the project, $\int \pi(a, \omega)dF_{W|X}(\omega|x)$, is nondecreasing in a .

Assumptions (A1) and (A2) simplify the analysis by allowing us to work with the MLR and MPR without treating separately the possibility of moving support. Further, (A2) allows us to assume that if Ω is not finite, the posterior distribution has a well-defined, continuous density, and that for each x , the likelihood ratio $f(\omega'|x)/f(\omega''|x)$ exists for almost every (ω', ω'') pair in Ω . The differentiability assumptions (A3) simplify the notation and analysis. Assumption (A4) allows us to ignore the possibility of multiple optima, and to assume that the first-order conditions characterize the optimal choice of investment (we clearly cannot impose the assumption for the portfolio problem when the firm is risk-neutral, however). (A3) and (A4) can be relaxed; see Athey (forthcoming) and AL for details. Assumption (A5) says that higher states are better, and that on average, the returns to investment are non-negative. These assumptions are satisfied in the portfolio problem, and they highlight the elements of the portfolio problem that are critical for comparative statics results. Finally, note that this paper focuses on the case where π is supermodular, incorporating the idea that ω represents the marginal returns to the investment.

Now consider the role of risk preferences in deriving orders over posteriors that lead to monotone decision problems. The following Proposition summarizes the comparative statics results for two extreme cases: firms whose risk preferences are unrestricted, and risk neutral firms.

Proposition 5 *Assume A1-A2, and suppose there is a unique optimal action for each x . (i) $\alpha^*(\cdot)$ is nondecreasing for all u nondecreasing and linear (that is, all risk neutral firms), and all π supermodular, if and only if $F_{W|X}(\cdot|x)$ is ordered by FOSD. (ii) $\alpha^*(\cdot)$ is nondecreasing for all u nondecreasing (that is, all firms with arbitrary risk preferences) and all π supermodular, if and only if $F_{W|X}(\cdot|x)$ is ordered by MLR.*

Proof. Under the assumptions of the proposition, weak single crossing of the incremental returns to a is necessary and sufficient for the comparative statics prediction. (i) Taking the class of functions $g(\omega, a) = u \circ \pi$, where u is nondecreasing and π is supermodular, generates the set U^{RSC} . Apply Lemma 3. (ii) Taking the class of linear, nondecreasing u returns the class of supermodular payoff functions. Apply Lemma 3. ■

Now consider the problem faced by a risk-averse firm. Notice that supermodularity is not preserved by nondecreasing, concave transformations. Thus, $g(\omega, a) = u \circ \pi$ is not necessarily supermodular when π is supermodular, unless u is linear, or else u is convex and π satisfies some

monotonicity restrictions. However, for any $g(\omega, a) = u \circ \pi$ where π is supermodular and u is nondecreasing and concave, $\frac{\partial}{\partial a}g(\omega, a)$ satisfies the single crossing property. But, if u is concave, not all functions with the single crossing property can be generated in this way.

Since the relevant set of payoff functions for risk averse firms is larger than the set of supermodular payoff functions, but smaller than the set of payoffs whose marginal returns satisfy the single crossing property, the comparative statics result for risk averse firms requires an intermediate condition on the distribution. It turns out this condition is equivalent to the stochastic single crossing order for the set of functions R^{SCQ} defined in the last subsection. This is somewhat surprising, given that $r(\omega) \equiv u'(\pi(a, \omega))\pi_a(a, \omega)$ is not necessarily in the set R^{SCQ} .

Proposition 6 *Assume A1-A5. Then $\alpha^*(\cdot)$ is nondecreasing in x for all u concave and nondecreasing and all π supermodular, if and only if x orders $F_{W|X}(\cdot|x)$ by MPR.*

The proof of this result is in the Appendix. As expected, the condition for comparative statics, MPR, is weaker than MLR and stronger than FOSD. The sufficiency part of this result generalizes Eeckhoudt and Gollier (1995), who establish the result for the portfolio problem. The proof of necessity builds on our discussion from the last subsection: the counter-examples we construct in Lemma 3 use payoff functions of the form $\mathbf{1}_{\{\omega \leq a\}}(\omega) \cdot r(\omega)$; but, for any such function we can find x , u concave, and π supermodular, so that the marginal returns to x , $u'(\pi(x, \omega))\frac{\partial}{\partial x}\pi(x, \omega)$, approximate this function. In other words, the set of functions of the form $g = u \circ \pi$, where u is concave and π is supermodular, contains (and is strictly larger than) $U^{R^{SCQ}}$.

3 The Value of Information

Now consider the problem of comparing the value of different signals. In particular, suppose the firm can choose from a family of signals $\{X^\theta\}_{\theta \in \Theta}$, where the quality of the signal is indexed by θ . Formally, the firm has a fixed prior, $H(\cdot)$, about the state of the world. Then, the firm chooses θ , which determines the set of possible posteriors $\{F_{W|X^\theta}(\cdot|x), x \in \mathcal{X}^\theta\}$ as well as the likelihood of each. After observing the realization of the signal $X^\theta = x$, the firm forms the posterior $F_{W|X^\theta}(\cdot|x)$ and solves $\max_{a \in \mathcal{A}} \int u(\pi(a, \omega))dF_{W|X^\theta}(\omega|x)$.

When evaluating the problem from an ex ante perspective, for each θ , the prior H and the signal distribution, $F_{X^\theta|W}$, determine a joint distribution over states and signal realizations, denoted $F_{W, X^\theta}(\cdot, \cdot)$. The joint distribution $F_{W, X^\theta}(\cdot, \cdot)$ is referred to as the *information structure*. The marginal distribution of the signal is denoted $F_{X^\theta}(\cdot)$. The notation $F_W(\omega|X^\theta \leq x)$ represents $\Pr(W \leq \omega|X^\theta \leq x)$, and $F_{X^\theta}(x|W \leq \omega)$ represents $\Pr(X^\theta \leq x|W \leq \omega)$. The corresponding probability mass functions or densities are denoted by f .

For a given payoff function $g : \mathcal{A} \times \Omega \rightarrow \mathbb{R}$ (such as $g = u \circ \pi$), the ex ante value of information is given by

$$V^*(\theta; g) \equiv \int_{\mathcal{X}^\theta} \left[\max_{a \in \mathcal{A}} \int_{\Omega} g(a, \omega) dF_{W|X^\theta}(\omega|x) \right] dF_{X^\theta}(x).$$

In the next subsection, we review and extend the results from AL that give necessary and sufficient conditions such that $\frac{\partial}{\partial \theta} V^*(\theta; u \circ \pi) \geq 0$ for all payoff functions in a given class.

3.1 Monotone Information Orders

AL develop a theory which can be used to order information structures for payoff functions of the form $g : \mathcal{A} \times \mathcal{X} \rightarrow \mathbb{R}$. A higher value of θ provides a “more informative” signal for a set of payoff functions U^R , if $\frac{\partial}{\partial \theta} V^*(\theta; g) \geq 0$ for all $g \in U^R$.

As in Lehmann (1988), AL’s approach considers only “Monotone Decision Problems.” That is, the set of admissible information structures must be “ R -ordered,” so that the stochastic single crossing order for R , denoted \succ_R , is a complete order on $\{F_W(\cdot|x)\}_{x \in \mathcal{X}}$. In other words, for all $x_H > x_L$, $F_W(\cdot|x_H) \succ_R F_W(\cdot|x_L)$. If F_{W, X^θ} is R -ordered, Lemma 1 implies that for all $g \in U^R$, there exists a selection $\alpha^*(x; \theta)$ from $\arg \max_{a \in \mathcal{A}} \int g(a, \omega) dF_{W|X^\theta}(\omega|x)$ such that $\alpha^*(x; \theta)$ is nondecreasing in x ; thus, throughout we assume that the agent chooses a nondecreasing policy. For each of the different sets of payoff functions R we have considered, Lemma 3 describes the restrictions implied for F_{W, X^θ} .

The restriction to monotone decision problems allows us to make use of our knowledge of the decision-maker’s preferences. In particular, when the decision policy is nondecreasing, $g(\alpha^*(x; \theta), \omega)$ inherits properties from $g(a, \omega)$. For example, if g is supermodular and $F_{W|X^\theta}(\cdot|x)$ is ordered by FOSD, then $\alpha^*(x; \theta)$ is nondecreasing in x and $g(\alpha^*(x; \theta), \omega)$ is supermodular in (x, ω) .

An important thing to note about the setup of the problem is that the value attached to a realization of X^θ has no intrinsic meaning to a decision-maker: all that matters after the signal is realized is the posterior distribution over the state of the world. For this reason, we are free to normalize the signal as we like, so long as the ordering is preserved. For example, we can consider a transformation $T : \mathcal{X}^\theta \times \Theta \rightarrow [0, 1]$ strictly increasing in x , and for each θ , consider the signal $Y^\theta \equiv T(X^\theta; \theta)$; having access to Y^θ must give the decision-maker the same payoffs as X^θ .

Why is such a normalization important? It plays a central role in our characterization of the information orders. Our approach to ranking information structures can be outlined as follows. Consider $\theta_H > \theta_L$, with corresponding signals X^H and X^L . We will provide conditions under which X^H is “more informative” than X^L . To do so, it is sufficient to show that for each $g \in U^R$, there exists a policy (not necessarily optimal) that a decision-maker can use with signal X^H , such

that expected payoffs are higher than when the decision-maker uses the optimal policy with signal X^L .

Consider how this might be accomplished. Consider $g \in U^R$, and take the optimal policy, $\alpha^*(\cdot; \theta_L) : \mathcal{X}^L \rightarrow \mathcal{A}$. This yields expected payoffs $V^*(\theta_L; g)$. Our goal is then to construct a policy, $\tilde{\alpha}(\cdot; \theta_H) : \mathcal{X}^H \rightarrow \mathcal{A}$, that (when the decision-maker has access to X^H rather than X^L) yields ex ante expected payoffs greater than $V^*(\theta_L; g)$. Notice that it is not immediately obvious how to construct this policy. For example, if $\mathcal{A} = [0, 1]$, $\mathcal{X}^L = [0, 1]$ and $\mathcal{X}^H = [-100, 100]$, if $\alpha^*(x; \theta_L) \equiv x$, the the function $\tilde{\alpha}(\cdot; \theta_H) = \alpha^*(\cdot; \theta_L)$ does not map \mathcal{X}^H to \mathcal{A} .

Instead, our approach is to choose a particular transformation, $T : \mathcal{X}^\theta \times \Theta \rightarrow [0, 1]$, and rescale each signal, defining two new signals $Y^L = T(X^L; \theta_L)$ and $Y^H = T(X^H; \theta_H)$. Next, we define $\beta : [0, 1] \rightarrow \mathcal{A}$ by $\beta(T(x; \theta_L)) \equiv \alpha^*(x; \theta_L)$, so that using $\beta(\cdot)$ with signal Y^L is equivalent to using $\alpha^*(\cdot; \theta_L)$ with signal X^L . Then, our approach to ranking information structures reduces to showing that an agent who has access to signal Y^H can do better than the agent with access to Y^L using the same policy, $\beta(\cdot)$. Since the realizations of both Y^L and Y^H lie in $[0, 1]$, this policy is well-defined for an agent who has access to Y^H .

Of course, we have not yet said anything about how this transformation, T , should be determined. It turns out that the answer depends on which class of payoff functions is under consideration. The following proposition includes the main results about information orderings; it is proved in the Appendix, and in the next subsection, we provide characterizations that restate these orders in more familiar terms.

Proposition 7 *Let H be a prior probability distribution over Ω , let $F_{X^\theta|W}$ be a smoothly parameterized signal distribution, and let F_{W, X^θ} be the corresponding family of information structures. Assume A1-A2.*

(i) *Suppose that for each θ , $F_{W|X^\theta}(\cdot|x)$ is ordered by FOSD in x . Then $\frac{\partial}{\partial \theta} V^*(\theta; g) \geq 0$ for all θ and all g supermodular, if and only if, letting $T(x; \theta) \equiv F_{X^\theta}(x)$,*

$$F_W(\cdot | X^\theta \geq T^{-1}(y; \theta)) \text{ is ordered by } \succ_{FOSD} \text{ in } \theta \text{ for all } y \in [0, 1]. \quad (\text{MIO-spm})$$

(ii) *Suppose that for each θ , $F_{W|X^\theta}(\cdot|x)$ is ordered by the MLR in x (that is, $F_{W|X^\theta}(\cdot|x)$ is ordered by $\succ_{LR-\omega_0}$ in x for each ω_0). Then $\frac{\partial}{\partial \theta} V^*(\theta; g) \geq 0$ for all θ and all $g \in U^{R^{SC}}$, if and only if, if we let $T(x; \omega, \theta) \equiv F_{X^\theta|W}(x|\omega)$,*

$$\text{For all } \omega_0 \in \Omega \text{ and } y \in [0, 1], F_W(\cdot | X^\theta \geq T^{-1}(y; \omega_0, \theta)) \text{ is ordered by } \succ_{LR-\omega_0} \text{ in } \theta. \quad (\text{MIO-sc})$$

(iii) *Suppose that for each θ , $F_{W|X^\theta}(\cdot|x)$ is ordered by the MPR in x (that is, $F_{W|X^\theta}(\cdot|x)$ is ordered*

by $\succ_{PR-\omega_0}$ in x for each ω_0). Then $\frac{\partial}{\partial \theta} V^*(\theta; g) \geq 0$ for all $g \in U^{R^{SCQ}}$, if and only, if we let $T(x; \omega, \theta) \equiv F_{X^\theta}(x|W \leq \omega)$,

For all $\omega_0 > \inf \Omega$ and $y \in [0, 1]$, $F_W(\cdot | X^\theta \geq T^{-1}(y; \omega_0, \theta))$ is ordered by $\succ_{PR-\omega_0}$ in θ .
(MIO-scq)

Part (ii) of this proposition is originally due to Lehmann (1988); it was first stated in terms of single crossing orders over average posteriors by AL. Part (i) is due to AL, while part (iii) is new here (although the derivation follows directly from AL). (MIO-sc) has been applied in economics in the area of principal-agent problems (Kim, 1995; Jewitt, 1997) and auctions (Persico, 2000). AL apply (MIO-spm) to a number of economic problems, including adverse selection problems and oligopoly games.

To interpret the three conditions, consider $\theta_H > \theta_L$ and the corresponding signals, X^H and X^L . Notice that all three conditions are comparisons across information structures of the average over “high” posteriors (that is, posteriors corresponding to high signal realizations), where posteriors are ordered by the relevant stochastic single crossing order. Each monotone information order requires that the “high” posteriors for the good signal, X^H , are “higher” than the “high” posteriors for the bad signal, X^L . Recall that when the agent uses a monotone policy, higher signal realizations correspond to higher investments. For example, consider $U^{R^{SCQ}}$, let $\mathcal{A} = \{0, 1\}$, and consider an agent whose return to investment changes direction at ω_0 . If this agent observes signal X^L , she chooses a policy where, for some y , she invests when $X^L \geq F_{X^L}^{-1}(y|W \leq \omega_0)$. Then, (MIO-scq) implies that, if the agent instead observes signal X^H and invests exactly when $X^H \geq F_{X^H}^{-1}(y|W \leq \omega_0)$, the average over the agent’s posteriors conditional on investment are higher (according to $\succ_{PR-\omega_0}$).

Since the average of the posteriors is the prior, the conditions of Proposition 7 likewise require that the “low” posteriors for the good signal are “lower” than the “low” posteriors for the bad signal. Thus, another way to interpret the conditions is that they require the posteriors of an agent given a good signal to be more “spread out” in the appropriate stochastic order.

The three monotone information orders differ not only in the relevant order over posteriors; they also differ in that different “normalizations,” T , of the signals are used. The normalizations are tailored to the class of payoff functions, and they are derived in the proof of the proposition.

Observe that in part (ii) (and similarly, in part (iii)), we cannot simplify the statement of (MIO-sc) by requiring that $F_W(\cdot | X^\theta \geq T^{-1}(y; \omega, \theta))$ is ordered by the MLR in θ for all ω . To see why, observe that ω_0 appears twice in the statement of (MIO-sc): first, we condition on ω_0 in the normalization T , and second, we check that the posteriors are ordered by $\succ_{LR-\omega_0}$. It is important that the crossing point specified for the stochastic order is the same as the crossing point used in

the normalization T .

Finally, we note that we could also modify parts (ii) and (iii) to consider the sets $U^{R^{SC}(\omega_0)}$ and $U^{R^{SCQ}(\omega_0)}$. The orders over posteriors would simply be checked only for the relevant ω_0 .

AL establish that (MIO-spm) holds for all priors, if and only if (MIO-sc) holds (this is analogous to Lemma 2 (ii)). To verify that (MIO-scq) is stronger than (MIO-spm), simply let $\omega_0 = \sup \Omega$ in the definition of (MIO-scq). It can further be shown that, given a prior H , (MIO-scq) holds, if and only if (MIO-spm) holds for all priors of the form $G(\omega) \equiv H(\omega|W \leq a)$ (analogous to Lemma 2 (iii)).

3.1.1 Characterizations of Monotone Information Orders

This section provides three alternative characterizations of (MIO-sc) and (MIO-scq). The characterizations may be independently useful. We begin with a characterization that states the monotone information orders in a way that relates more closely to the \succ_{LR} and \succ_{PR} orders. It is straightforward to show that if each of Q and P has a differentiable density, $Q \succ_{LR} P$ if and only if for all $\omega \in \Omega$,

$$\frac{q'(\omega)}{q(\omega)} \geq \frac{p'(\omega)}{p(\omega)}.$$

Similarly, if Q and P have densities, $Q \succ_{PR} P$ if and only if for all $\omega \in \Omega$,

$$\frac{q(\omega)}{Q(\omega)} \geq \frac{p(\omega)}{P(\omega)}.$$

These characterizations are frequently useful in analyzing applied problems, such as auction games. The following characterization expresses the monotone information orders over average posteriors in an analogous way.

Proposition 8 (i) *Suppose that for all θ and all $x \in \mathcal{X}^\theta$, the posterior belief has a density that is differentiable. Then $F_{X^\theta, W}$ is ordered by (MIO-sc), if and only if, for all $\omega_0 \in \Omega$ and all $z \in [0, 1]$,*

$$\frac{f'_W(\omega_0 | X^\theta \geq F_{X^\theta|W}^{-1}(z|\omega_0))}{f_W(\omega_0 | X^\theta \geq F_{X^\theta|W}^{-1}(z|\omega_0))} \text{ is nondecreasing in } \theta. \quad (2)$$

(ii) *Suppose that for all θ and all $x \in \mathcal{X}^\theta$, the posterior belief has a density. Then $F_{X^\theta, W}$ is ordered by (MIO-scq), if and only if, for all $\omega_0 \in \Omega$ and all $z \in [0, 1]$,*

$$\frac{f_W(\omega_0 | X^\theta \geq F_{X^\theta}^{-1}(z|W \leq \omega_0))}{F_W(\omega_0 | X^\theta \geq F_{X^\theta}^{-1}(z|W \leq \omega_0))} \text{ is nondecreasing in } \theta. \quad (3)$$

When applying Proposition 8, it is important to note that ω_0 appears in two places in each of the formulas, once as the argument of the distribution or density, and a second time in normalizing the signals.

The proof of Proposition 8 builds on the following Lemma. Observe that the characterizations in this and the following Lemma are stated in terms of the distribution (or inverse distribution) of the signal conditional on a state or an interval of states, rather than in terms of posterior distribution.

Lemma 9 (i) *The following four conditions are equivalent: (a) $F_{X^\theta, W}$ is ordered by (MIO-sc); (b) For all $(y, \omega_0) \in [0, 1] \times \Omega$, and all $v > \omega_0$, $F_{X^\theta}(F_{X^\theta}^{-1}(y|W = \omega_0) | W = v)$ is nonincreasing in θ ; (c) For all $(y, \omega_0) \in [0, 1] \times \Omega$, and all $v < \omega_0$, $F_{X^\theta}(F_{X^\theta}^{-1}(y|W = \omega_0) | W = v)$ is nondecreasing in θ ; (d) for all $\theta_H > \theta_L$, $F_{X^H}^{-1}(F_{X^L}(x|W = \omega_0) | W = \omega_0)$ is nondecreasing in ω_0 for all $x \in \mathcal{X}^L$.*

(ii) *The following four conditions are equivalent: (a) $F_{X^\theta, W}$ is ordered by (MIO-scq); (b) For all $(y, \omega_0) \in [0, 1] \times \Omega$, and all $v > \omega_0$, $F_{X^\theta}(F_{X^\theta}^{-1}(y|W \leq \omega_0) | W \leq v)$ is nonincreasing in θ ; (c) For all $(y, \omega_0) \in [0, 1] \times \Omega$, and all $v < \omega_0$, $F_{X^\theta}(F_{X^\theta}^{-1}(y|W \leq \omega_0) | W \leq v)$ is nondecreasing in θ ; (d) for all $\theta_H > \theta_L$, $F_{X^H}^{-1}(F_{X^L}(x|W \leq \omega_0) | W \leq \omega_0)$ is nondecreasing in ω_0 for all $x \in \mathcal{X}^L$.*

Condition (i)(d) of this Lemma is the statement of (MIO-sc) given by Lehmann (1988), and AL show that (MIO-sc) is equivalent to (i)(d); part (ii) is new here. First, consider the interpretation of part (i)(b). Suppose that we seek to test the null hypothesis that $W \geq \omega_0$ against the alternative $W < \omega_0$, and we want a test that rejects the null hypothesis with probability no greater than y when the null hypothesis is true, that is, a test of size y . Observe that (given posteriors that are ordered by the MLR) we are most likely to reject the null incorrectly when $W = \omega_0$. With signal X^L , if the posteriors are ordered by MLR, to implement a test of size y we reject the null when $F_{X^L}(X^L|W = \omega_0) \leq y$. Now suppose we gain access to another signal, X^H , and consider implementing a hypothesis test of the same size, that is, a test that rejects the null if $F_{X^H}(X^H|W = \omega_0) \leq y$. We say that X^H is more informative than X^L if, given the size of test y , if the true state of the world is $v > \omega_0$, the probability of rejecting the null is smaller with X^H than X^L . Part (i)(c) requires that the probability of rejecting the null hypothesis goes up when the true state is $v < \omega_0$. In the terminology of classical statistics, for a fixed size y , X^H provides a uniformly more powerful hypothesis test than X^L .

Now consider the interpretation of part (ii)(c). Suppose that the agent's marginal return to investment, r , is in $R^{SCQ(\omega_0)}$. In words, r is nondecreasing on $\omega \leq \omega_0$, and remains nonnegative thereafter. Suppose that $\mathcal{A} = \{0, 1\}$. Then, for each θ , consider the policy where, for some y , the agent chooses to invest if the realization of X^θ is greater than $F_{X^\theta}^{-1}(y|W \leq \omega_0)$. Part (ii)(c) requires that if the true state of the world is in the set $\{\omega : \omega \leq v\}$, where $v < \omega_0$, a "better" signal makes it less likely that the agent chooses to invest. In contrast, by part (ii)(b), when $v > \omega_0$ a "better"

signal makes it more likely that the agent chooses to invest. Given the agent’s preferences, these changes increase ex ante expected utility.

Notice that, unlike (MIO-sc), (MIO-scq) places requirements on the “average” signal distributions. That is, in order to compute $F_{X^L}(\cdot | W \leq \omega_0)$, we need to know the prior distribution over W . Thus, while (MIO-sc) can be interpreted in terms of “classical” hypothesis testing in statistics, giving conditions for signals to be ranked for all possible priors, (MIO-scq) requires a Bayesian perspective, in particular, a pre-specified prior distribution. As AL discuss in more detail, in order to make use of the additional structure on the set of payoff functions R^{SCQ} (rather than R^{SC}), we must be willing to specify a restricted class of prior distributions. Notice, however, that given a prior H , the MPR order and (MIO-scq) are unchanged for a given signal family $F_{X^\theta}(\cdot | \omega)$ if we allow for an alternative prior of the form $\tilde{H}(\omega) \equiv H(\omega | W \leq v)$.

The next lemma builds on the characterizations from parts (i)(b) and (ii)(b), stating the conditions directly in terms of (average) signal distributions rather than the inverse distributions. The characterizations can be understood in terms of the transformation function, T . Suppose that for all θ , an investor follows the policy of investing when the realization of the normalized signal, $T(X^\theta; \theta, \omega_0)$, is greater than y . The following Lemma places conditions on how the “cutoff” level of x that implements this policy changes in θ and ω_0 .⁹

Lemma 10 (i) Define $\tilde{x}(y; \theta, \omega_0) \equiv F_{X^\theta|W}^{-1}(y | W = \omega_0)$. Then $F_{X^\theta, W}$ is ordered by (MIO-sc) if and only if $\tilde{x}(y; \theta, \omega_0)$ is supermodular in (θ, ω_0) , that is,

$$-\frac{\frac{\partial}{\partial \theta} F_{X^\theta}(x | W = \omega_0)}{f_{X^\theta}(x | W = \omega_0)} \text{ is nondecreasing in } \omega_0 \text{ for all } x. \quad (4)$$

(ii) Define $\check{x}(y; \theta, \omega_0) \equiv F_{X^\theta}^{-1}(y | W \leq \omega_0)$. Then if $F_{X^\theta, W}$ is ordered by (MIO-scq), $\check{x}(y; \theta, \omega_0)$ is supermodular in (θ, ω_0) , that is,

$$-\frac{\frac{\partial}{\partial \theta} F_{X^\theta}(x | W \leq \omega_0)}{f_{X^\theta}(x | W \leq \omega_0)} \text{ is nondecreasing in } \omega_0 \text{ for all } x. \quad (5)$$

To interpret the result of Lemma 10 (ii), $\frac{\partial}{\partial \omega_0} \check{x}(y; \theta, \omega_0)$ is a measure of the positive dependence between X^θ and W , where positive dependence is defined to increase expected payoffs for the relevant class of decision-makers (in this case, the MPR). Condition (5) requires that θ increases the positive dependence (in terms of the MPR) between X^θ and W .¹⁰

⁹Lemma 10 can also be used to establish directly the relationship between MIO-sc and MIO-scq. To see this, we can define $h(\omega_0; \eta)$ for $\eta \in \{\eta^L, \eta^H\}$ as follows: $h(\omega_0; \eta^L) \equiv f^\theta(x | \omega_0)$, $h(\omega_0; \eta^H) \equiv -\frac{\partial}{\partial \theta} F_{X^\theta|W}(x | \omega_0) \geq 0$. Then, (4) is nondecreasing in ω_0 exactly when $h(\omega_0; \eta)$ is log-supermodular. Since integrals of log-supermodular functions are log-supermodular (Karlin, 1968), (4) nondecreasing in ω_0 implies (5) is nondecreasing in ω_0 .

¹⁰Persico (2000) provides an analogous interpretation of MIO-sc. See also Jewitt (1997).

Finally, consider briefly an analogous characterization of (MIO-spm). In particular, notice that r is nondecreasing if and only if it is in $R^{SCQ(\bar{\omega})}$. Now, consider Lemma 9, part (ii)(b), when $\omega_0 = \bar{\omega}$. The requirement reduces to $\frac{\partial}{\partial \theta} \Pr(X^\theta \leq F_{X^\theta}^{-1}(y) \mid W \leq v) \geq 0$ for all v, y . This is equivalent to (MIO-spm). Equivalently, $\frac{\partial}{\partial \theta} F_{W, X^\theta}(v, F_{X^\theta}^{-1}(y)) \geq 0$ for all v, y .¹¹

3.2 Information Orders for a Risk-Averse Investor

For the problem of a risk-averse investor, payoffs satisfy single crossing of incremental returns, and so we can apply (MIO-sc). However, (MIO-sc) is potentially a very strong condition. Its restrictions must hold conditional on any pair of states. In contrast, (MIO-scq) allows some “averaging” over states.

Thus, we would like to relax (MIO-sc). The next result establishes that (MIO-scq) is the relevant information order for risk-averse investors.

Proposition 11 *Assume A1-A5. Assume that for each θ , $F_{W|X^\theta}(\cdot|x)$ is ordered by the MPR in x . Then $\frac{\partial}{\partial \theta} V^*(\theta; u, \pi) \geq 0$ for all u nondecreasing and concave and all π supermodular, if and only if (MIO-scq) holds.*

The proof of this result is in the Appendix. It builds on the proofs of Propositions 5 and 7, and it uses the characterization of (MIO-scq) derived in Lemma 10.

4 The Information Acquisition Problem

The information acquisition problem is stated as follows:

$$\theta^*(\rho) \equiv \arg \max_{\theta \in \Theta} \int_x \int_\omega u(\pi(\alpha^*(x; \theta, \rho), \omega); \rho) dF_{W, X^\theta}(\omega, x) - c(\theta),$$

where we have introduced a new parameter, ρ , which describes the agent’s risk preferences. Consider the following question about the comparative statics on information gathering: When does the amount of information acquired by the firm increase in ρ ?

To begin, we present a result that applies to the class of payoffs $U^{R^{SCQ}}$.

Proposition 12 *Let H be a prior probability distribution over Ω , let $F_{X^\theta|W}$ be a smoothly parameterized signal distribution, and let F_{W, X^θ} be the corresponding family of information structures. Assume A1-A2. Suppose that θ orders F_{W, X^θ} by (MIO-scq). Define*

$$w(x, \omega) = g(\alpha^*(x; \theta, \rho_H), \omega; \rho_H) - g(\alpha^*(x; \theta, \rho_L), \omega; \rho_L).$$

¹¹See AL for further characterizations of (MIO-spm) in terms of “marginal-preserving spreads.”

If for all $x_H > x_L$, $w(x_H, \omega) - w(x_L, \omega) \in R^{SCQ}$, then $\frac{\partial}{\partial \theta} V(\theta, u) \geq \frac{\partial}{\partial \theta} V(\theta, v)$.

When the relevant functions are differentiable, the requirement of the proposition is that $\frac{\partial^2}{\partial \rho \partial x} g(\alpha^*(x; \theta, \rho), \omega; \rho) \in R^{SCQ}$. In other words, ρ makes the incremental returns to investment “more R^{SCQ} .” For example, this condition would be satisfied if $g(\alpha^*(x; \theta, \rho), \omega; \rho) \equiv x \rho r(\omega)$, for $r \in R^{SCQ}$. This result is analogous to results established for R^{SC} in Persico (2000) and R^{ND} in AL.

Now consider applying this to a risk-averse investor. Let us restrict attention to the portfolio problem, where $\pi(a, \omega) \equiv a\omega + (1 - a)\omega_0$ (this is not necessary, but it simplifies an already complicated analysis). Further, we add the following assumption:

A6 There exists a $c > 0$ such that for each θ , $\alpha^*(x; \theta, \rho) > c$ for all realizations of the signal.

Further, the investor is eventually satiated: $u_1(c \cdot \bar{\omega}; \rho) = 0$.

The satiation condition is clearly undesirable, but it greatly simplifies the argument.

Consider small changes in ρ , and define

$$\begin{aligned} \Psi(x, \omega; \theta, \rho) &\equiv \frac{d}{d\rho} [\alpha^*(x; \theta, \rho) u_{11}(\pi(\alpha^*(x; \theta, \rho), \omega); \rho)] \\ &= \begin{aligned} &u_{112}(\pi(\alpha^*(x; \theta, \rho), \omega); \rho) \alpha^*(x; \theta, \rho) \\ &- \alpha_\rho^*(x; \theta, \rho) \omega_0 u_{111}(\pi(\alpha^*(x; \theta, \rho), \omega); \rho) \\ &+ \alpha_\rho^*(x; \theta, \rho) \left[\begin{array}{c} u_{11}(\pi(\alpha^*(x; \theta, \rho), \omega); \rho) \\ + u_{111}(\pi(\alpha^*(x; \theta, \rho), \omega); \rho) \pi(\alpha^*(x; \theta, \rho), \omega) \end{array} \right] \end{aligned} \end{aligned}$$

Then, we have the following comparative statics result for information acquisition.

Proposition 13 *Assume A1-A5, let $\mathcal{A} = [0, 1]$, and suppose $\pi(a, \omega) \equiv a\omega + (1 - a)\omega_0$. Assume (without loss of generality) that each X^θ is normalized so that $\frac{\partial}{\partial \theta} F_{X^\theta}(x) = 0$. Suppose further that for all θ , $F_{W|X^\theta}(\cdot|x)$ is ordered in the MPR in x , and θ orders the signal family F_{W, X^θ} according to (MIO-scq). Then, whenever $\Psi(x, \omega; \theta, \rho) \leq 0$ and α^* is supermodular in (x, ρ) , $\theta^*(\rho)$ is nondecreasing.*

To interpret these conditions, observe that if α^* is supermodular in (x, ρ) , higher values of ρ make the agent more sensitive to signal realizations. Typically, this effect would be likely to hold if ρ reduces risk aversion. The requirement that $\Psi \leq 0$ is more complicated. If α_ρ^* is close enough to zero, only the first term is important. Then, $\Psi \leq 0$ if $u_{112} \leq 0$. In words, higher values of ρ make the utility function more concave, which typically corresponds to greater risk aversion. This

captures the idea that for a given decision policy, greater risk aversion makes information more valuable.

The other terms in capture the effects that arise because ρ changes the optimal decision policy. If $\alpha_\rho^* \geq 0$ (i.e., higher values of ρ make the agent invest more), then ρ can be interpreted most naturally as a decrease in risk aversion. If, in addition, $\omega_0 > 0$ and $u_{111}(w; \rho) > 0$, then the second term is negative. Under the same conditions, the third term is negative if the agent's relative prudence, $-\frac{u_{111}(w; \rho)}{u_{11}(w; \rho)}w$, is less than 1. Since an agent's incentive to save in anticipation of a future (multiplicatively separable) risk is increasing in relative prudence, this requirement can be understood as a restriction on the extent to which an agent is willing to pay to avoid exposure to a multiplicative risk.

What more can we say about preferences for which $\Psi \leq 0$ and α^* is supermodular in (x, ρ) ? Consider first an agent with CARA preferences, with coefficient of risk aversion $-\rho$.¹² If the posterior beliefs have a normal distribution with mean x and variance independent of x ,¹³ an increase in ρ (a reduction in risk aversion) makes the agent more responsive to signal realizations, i.e., α^* is supermodular in (x, ρ) . For more general functional forms for posterior beliefs, additional assumptions are required to obtain this result. Now, consider our restriction on Ψ . It can be verified that sufficient conditions for $\Psi \leq 0$ are that, for some $\gamma > 0$, $1 \geq -\rho \geq \gamma$, and ω_0 and $\inf \Omega$ are both greater than $2/\gamma$.

Now consider another specific functional form, a quadratic utility function, where it is possible to provide a definitive prediction. Let $\omega_0 = 0$, $\mathcal{A} = [0, 1]$, $\Omega = [\underline{\omega}, \bar{\omega}]$ with $\bar{\omega} > 0 > \underline{\omega}$, and consider $\rho > \bar{\omega}$. Then define $u : \Omega \rightarrow \mathbb{R}$ by

$$u(m) \equiv \rho m - \frac{1}{2}m^2.$$

Notice that $u(a\omega)$ is not necessarily supermodular. Suppose that $\mathbb{E}_W[W|X^\theta = x]/\mathbb{E}_W[W^2|X^\theta = x] \in (0, 1)$ for all x . The first-order conditions for optimality entail that

$$\alpha^*(x, \rho) = \rho \frac{\mathbb{E}_W[W|X^\theta = x]}{\mathbb{E}_W[W^2|X^\theta = x]}.$$

If $F_{W|X^\theta}(\cdot|x)$ is ordered by the MPR, the optimal policy is nondecreasing in x . This in turn implies that the optimal policy is supermodular in (x, ρ) . Observe further that

$$\alpha^*(x; \theta, \rho)u_{11}(\pi(\alpha^*(x; \theta, \rho), \omega), \rho) = -\rho \frac{\mathbb{E}_W[W|X^\theta = x]}{\mathbb{E}_W[W^2|X^\theta = x]}.$$

¹²Note that this utility function does not satisfy our satiation requirement, so to apply our proposition directly we would need to extend the arguments of Proposition 13.

¹³Persico (1996) also analyzes this case, using Lehmann's (1988) information order.

Since the optimal policy is non-negative, the expression is nonincreasing in ρ (so that $\Psi(x, \omega; \theta, \rho) \leq 0$). Notice that

$$\frac{-u''}{u'}(m; \rho) = \frac{1}{\rho - m} \text{ and } \frac{\partial}{\partial \rho} \left(\frac{-u''}{u'}(m; \rho) \right) = \frac{-1}{(\rho - m)^2}.$$

So, risk aversion decreases with ρ . In this case, the result says that as the agent becomes less risk averse, the policy becomes more sensitive to information, and information becomes more valuable.

To apply these results, suppose that an investor can choose among the members of a family of signals, $\{X^\theta : \theta \in \Theta\}$, where the posteriors from each signal are totally according to MPR. The different signals might correspond to different opportunities for information gathering, such as annual reports of a firm, a full auditor's report, or being a manager of the firm. The investor's willingness to pay for signals in this family will be increasing in θ , if each member of the family is R^{SCQ} -ordered, and θ orders the family by (MIO-scq). Then, the results in this section can be used to provide conditions under which less risk averse agents purchase signals that more informative according to (MIO-scq).

Finally, observe that if we are willing to impose the more stringent requirements of MLR and (MIO-sc) in place of MPR and (MIO-scq), the requirements on risk preferences can be relaxed. See Persico (1996) or AL for details.

5 Conclusions

This paper derives sufficient conditions for comparative statics on investment and preference orderings over information structures for the problem faced by a risk averse firm, where the firm must make an investment decision on the basis of a noisy signal about the investment returns. The set of payoff functions generated by class of decision problems is larger than the set of supermodular functions, but smaller than the set of payoffs whose marginal returns are single crossing. Since the orders over posteriors and information structures induced by the set of single crossing payoffs are fairly restrictive, the weaker orders derived in this paper are potentially more useful in applications.

Using these orderings, we analyze how the value of information changes with risk preferences. We show that for a fixed policy, making a utility function more concave increases the demand for information. However, a decrease in risk aversion often makes an agent more responsive to an informative signal, and for some utility functions (such as quadratic), this effect dominates. Then, agents who are less risk averse purchase more information.

6 Appendix

Proof of Lemma 4: The fact that (1) is sufficient follows from checking the definitions. For necessity, observe that $R^{SCQ(\omega_0)}$ is a closed convex cone. But, for any closed convex cone R , $Q \succ_R P$ if and only if there exists a $\rho \geq 0$ such that $\int_{\Omega} rdQ \geq \rho \int_{\Omega} rdP$ for all $r \in R$.¹⁴ In the case of $R^{SCQ(\omega_0)}$, checking the latter inequality for $r(\omega) \equiv \mathbf{1}_{\{\omega \leq \omega_0\}}(\omega)$ and $r(\omega) \equiv -\mathbf{1}_{\{\omega \leq \omega_0\}}(\omega)$ implies that the inequality holds only if $\rho = Q(\omega_0)/P(\omega_0)$. Thus, (1) holds for all $r \in R^{SCQ(\omega_0)}$, if and only if $Q \succ_R P$ for $R = R^{SCQ(\omega_0)}$, as desired.

To prove the result more directly, observe that by linearity of the integral, (1) holds for all $r \in R^{SCQ(\omega_0)}$ if and only if it holds for all $r \in E^{SCQ(\omega_0)}$, where $E^{SCQ(\omega_0)}$ is a set of extreme points of $R^{SCQ(\omega_0)}$. That is, $E^{SCQ(\omega_0)}$ is a set of functions such that any $r \in R^{SCQ(\omega_0)}$ can be formed using convex combinations, positive scalar multiples, and limits of functions in $E^{SCQ(\omega_0)}$. The set $E^{SCQ(\omega_0)}$ is the union of the following sets: $\{-\mathbf{1}_{\{\omega \leq \omega_0\}}(\omega)\}$, $\{r : \text{for } a \leq \omega_0, r(\omega) \equiv \mathbf{1}_{\{a \leq \omega \leq \omega_0\}}(\omega)\}$, and $\{r : \text{for } a > \omega_0, r(\omega) \equiv \mathbf{1}_{\{\omega_0 \leq \omega \leq a\}}(\omega)\}$. Then, it is straightforward to verify that (1) is equivalent to $Q \succ_{PR-\omega_0} P$. ■

Proof of Proposition 6: Consider first necessity. The counter-examples required for necessity in part (iii) of Lemma 3 can be constructed using functions of the form $\mathbf{1}_{\{\omega \leq \omega_0\}}(\omega) \cdot r(\omega)$. Now, observe that given r nondecreasing and ω_0 , we can also find a u concave, a π supermodular, and an a such that $u'(\pi(a, \omega)) \frac{\partial}{\partial a} \pi(a, \omega)$ is approximately equal to $\mathbf{1}_{\{\omega \leq \omega_0\}}(\omega) \cdot r(\omega)$. To see this, choose π and $a > 0$ so that $\pi(a, \omega) = a \cdot r(\omega)$ for all ω , and then given this π , choose u so that $u' \approx 1$ when $\omega < \omega_0$ and $u' \approx 0$ for $\omega > \omega_0$. This u is concave.

It remains to establish sufficiency. Before beginning, we need a lemma. For variants on this result, see Athey (forthcoming).

Lemma 14 *Consider $\eta_H > \eta_L$, and suppose that for either $\eta = \eta_L$ or $\eta = \eta_H$, $g(\cdot, \eta) \in R^{SC(\omega_0)}$. Suppose further that $g(\omega, \eta_H) \geq g(\omega, \eta_L)$ for all $\omega \in \Omega$. Suppose that $k(\cdot; \eta_L)$ and $k(\cdot; \eta_H)$ are positive and continuous, and suppose that $k(\cdot; \eta_H)/k(\omega_0; \eta_H) - k(\cdot; \eta_L)/k(\omega_0; \eta_L)$ is weak single crossing at ω_0 . Then $\int g(\omega, \eta)k(\omega, \eta)d\omega$ satisfies weak single crossing in η . Further, for any $\eta_H > \eta_L$, let $l(\omega_0) = k(\omega_0, \eta_H)/k(\omega_0, \eta_L)$. Then,*

$$\int g(\omega, \eta_H)k(\omega, \eta_H)d\omega \geq l(\omega_0) \int g(\omega, \eta_L)k(\omega, \eta_L)d\omega. \quad (6)$$

¹⁴See Jewitt (1986), Gollier and Kimball (1995), and Athey (1999) for alternative proofs.

Proof of Lemma 14: Suppose that $g(\cdot, \eta_L) \in R^{SC(\omega_0)}$. Then,

$$\begin{aligned} \int g(\omega, \eta_H)k(\omega, \eta_H)d\omega &\geq \int g(\omega, \eta_L)k(\omega, \eta_H)d\omega \\ &= \int g(\omega, \eta_L)k(\omega, \eta_L)\frac{k(\omega, \eta_H)}{k(\omega, \eta_L)}d\omega \\ &\geq \frac{k(\omega_0, \eta_H)}{k(\omega_0, \eta_L)} \int g(\omega, \eta_L)k(\omega, \eta_L)d\omega \end{aligned}$$

The first inequality follows by our assumptions on g , since k is positive. The second inequality follows because $\frac{k(\omega, \eta_H)}{k(\omega, \eta_L)} \leq (\geq) \frac{k(\omega_0, \eta_H)}{k(\omega_0, \eta_L)}$ for ω where $g(\omega, \eta_L) \leq (\geq) 0$. If instead, it is $g(\cdot, \eta_H)$ that is in $R^{SC(\omega_0)}$, the equality and the second inequality hold replacing $g(\omega, \eta_L)$ with $g(\omega, \eta_H)$; then, because $\frac{k(\omega_0, \eta_H)}{k(\omega_0, \eta_L)} > 0$,

$$\frac{k(\omega_0, \eta_H)}{k(\omega_0, \eta_L)} \int g(\omega, \eta_H)k(\omega, \eta_L)d\omega \geq \frac{k(\omega_0, \eta_H)}{k(\omega_0, \eta_L)} \int g(\omega, \eta_L)k(\omega, \eta_L)d\omega$$

since g is nondecreasing in η . ■

Fix u and π , and let $\phi(a, x)$ represent the marginal returns to a given x , which are set equal to zero for an interior optimum:

$$\phi(a, x) \equiv \int u'(\pi(a, \omega))\pi_a(a, \omega)dF_{W|X}(\omega|x). \quad (\text{FOC})$$

Integrating by parts, we have

$$\begin{aligned} \phi(a, x) &= u'(\pi(a, \bar{\omega})) \int \pi_a(a, \omega)dF_{W|X}(\omega|x) - \int u''(\pi(a, \omega)) \int_{-\infty}^{\omega} \pi_a(a, t)dF_{W|X}(t|x)d\omega \\ &= \xi(a, x) + \psi(a, x), \end{aligned}$$

where we define

$$\xi(a, x) \equiv u'(\pi(a, \bar{\omega})) \int \pi_a(a, \omega)dF_{W|X}(\omega|x) \quad (7)$$

$$\psi(a, x) \equiv - \int u''(\pi(a, \omega)) \mathbb{E}_W[\pi_a(a, W)|x, W \leq \omega] F_{W|X}(\omega|x)d\omega. \quad (8)$$

Fix $x_H > x_L$, fix an a , and let $l(\omega_0) = \frac{F_{W|X}(\omega_0|x_H)}{F_{W|X}(\omega_0|x_L)}$. Since the MPR implies FOSD, $l(\omega_0) \leq 1$.

(a) Since u is nondecreasing, $\pi_a(a, \omega)$ is weak single crossing at ω_0 , and x orders $F_{W|X}(\cdot|x)$ by the MPR, Lemma 14 implies that $\xi(a, x_H) \geq \xi(a, x_L)$. Since $\xi(a, x_L) \geq 0$ and $l(\omega_0) \leq 1$, this implies that $\xi(a, x_H) \geq l(\omega_0)\xi(a, x_L)$.

(b) Now consider the second term, (8). By Lemma 2, $\mathbb{E}_W[\pi_a(a, W)|x, W \leq \omega]$ is nondecreasing in x , by the MPR and since $\pi_a(a, \omega)$ is nondecreasing in ω . Further, $\pi_a(a, \omega)$ nondecreasing in ω

implies that $\mathbb{E}_W[\pi_a(a, W)|x, W \leq \omega]$ is single crossing in ω . Since $u'' \leq 0$ by assumption, these facts imply that $-u''(\pi(a, \omega))\mathbb{E}_W[\pi_a(a, W)|x, W \leq \omega]$ satisfies single crossing in ω and is nondecreasing in x .

Fix $x_H > x_L$. Then, by Lemma 14, $\psi(a, x_H) \geq l(\omega_0)\psi(a, x_L)$. Combined with (a), that implies $\phi(a, x_H) \geq l(\omega_0)\phi(a, x_L)$. Thus, we have established that $\phi(a, x)$ satisfies weak single crossing in x , and the comparative statics conclusion follows. ■

Proof of Proposition 7: This result can be proved as an application of the results in AL for arbitrary sets R ; however, for completeness, we include a proof here that is tailored to the specific set of payoff functions of interest, R^{SCQ} . Before beginning, we recall that because the expected value of the posteriors is the prior, an equivalent condition to (MIO-scq) is that for all $\omega_0 > \inf \Omega$, if we let $T(x; \omega_0, \theta) \equiv F_{X^\theta}(X^\theta|W \leq \omega_0)$,

$$F_W(\cdot | X^\theta \leq T^{-1}(y; \omega_0, \theta)) \text{ is ordered by } \succ_{PR-\omega_0} \text{ in } -\theta \text{ for all } y \in [0, 1]. \quad (\text{MIO-scq}')$$

For the moment, we do not specify T ; it will be determined in the course of our analysis. Define

$$\tilde{V}(\theta; g, \beta, T) = \int_{\mathcal{X}} \int_{\Omega} g(\beta(y), \omega) dF_{W, X^\theta}(\omega, T^{-1}(y; \theta))$$

Let $L(U^R, \beta, \theta_H, \theta_L)$ be the minimized value of the following program:

$$\inf_{g \in U^R} \tilde{V}(\theta_H; g, \beta, T) - \tilde{V}(\theta_L; g, \beta, T)$$

$$\text{subject to: for all } y, \beta(y) \in \arg \max \int_{\Omega} g(a, \omega) dF_{W|X^\theta}(\omega | X^\theta = T^{-1}(y; \theta_L)).$$

If we consider small changes in θ , the envelope theorem implies that $\frac{\partial}{\partial \theta} V^*(\theta; g) \equiv \frac{\partial}{\partial \eta} \tilde{V}(\eta; g, \beta^*(\cdot; \theta), T)|_{\eta=\theta}$. In contrast, if θ_H is larger than θ_L by a discrete increment, $\inf_{g \in U^R} [V^*(\theta_H; g) - V^*(\theta_L; g)] \geq \sup_{\beta} L(U^R, \beta, \theta_H, \theta_L)$, since the decision-maker might like to choose a different (and better) policy under θ_H , while the program that determines L constrains the agent with signal θ_H to use the policy $\beta^*(\cdot; \theta_L)$. Therefore, a finding that $L(U^R, \beta, \theta_H, \theta_L)$ is non-negative for all β nondecreasing provides a sufficient condition for the conclusion that X^H is more informative than X^L for decision-makers in the class U^R ; for differential changes in θ , the condition is necessary and sufficient.

To proceed, we begin by showing that (MIO-scq) is necessary and sufficient for informativeness comparisons for small changes in θ , when we restrict attention to a subset of agents in U^R . We then argue that the informativeness rankings are sufficient to compare informativeness for all agents in U^R . The special class of decision problems allows only a binary choice, $a \in \{0, 1\}$, and all admissible utility functions take the form $g(a, \omega) \equiv a \cdot r(\omega)$, where $r(\cdot)$ is restricted to lie in R .

In such a problem, a monotone decision policy takes the following form: for some $y \in [0, 1]$, let $a = 1$ if $X^\theta \geq T^{-1}(y; \theta)$, and 0 otherwise. Fix a particular $\hat{y} \in [0, 1]$, and denote this policy by $\beta^{\hat{y}}$. Substituting in $\beta = \beta^{\hat{y}}$, and noting that the optimal policy will specify that $a = 1$ as soon as the expected returns are positive, the minimization problem can be rewritten as follows:

$$\inf_{r \in R} \int_{\Omega} r(\omega) d\omega \left[\begin{array}{l} F_W(\omega | X^H \geq T^{-1}(\hat{y}; \omega_0, \theta_H)) \Pr(X^H \geq T^{-1}(\hat{y}; \omega_0, \theta_H)) \\ - F_W(\omega | X^L \geq T^{-1}(\hat{y}; \omega_0, \theta_L)) \Pr(X^L \geq T^{-1}(\hat{y}; \omega_0, \theta_L)) \end{array} \right]$$

$$\text{subject to: } \int_{\Omega} r(\omega) dF_W(\omega | T(X^L; \omega_0, \theta_L) = \hat{y}) = 0.$$

The constraint is equivalent to the requirement that, given \hat{y} , we are only concerned with $r \in R$ where \hat{y} corresponds to the optimal policy for this r , when the agent has access to X^L .¹⁵ Then, $L(U^R, \beta^{\hat{y}}, \theta_H, \theta_L)$ will be non-negative if there exists a multiplier $\lambda(\hat{y})$ such that, for all $r \in R$,¹⁶

$$\begin{aligned} \int_{\Omega} r(\omega) d\omega [F_{W, X^L}(\omega, T^{-1}(\hat{y}; \omega_0, \theta_L)) - F_{W, X^H}(\omega, T^{-1}(\hat{y}; \omega_0, \theta_H))] \\ - \lambda(\hat{y}) \int_{\Omega} r(\omega) dF_W(\omega | T(X^L; \omega_0, \theta_L) = \hat{y}) \geq 0, \end{aligned} \quad (9)$$

where we used Bayes' rule (together with the fact that the expected value of the posteriors is the prior) to simplify the expression in brackets.

Focus now on the case where $R = R^{SCQ(\omega_0)}$ in particular. First, checking (9) when $r(\omega) \equiv \mathbf{1}_{\{\omega \leq \omega_0\}}(\omega)$ and when $r(\omega) \equiv -\mathbf{1}_{\{\omega \leq \omega_0\}}(\omega)$ requires that

$$F_{W, X^L}(\omega_0, T^{-1}(\hat{y}; \omega_0, \theta_L)) - F_{W, X^H}(\omega_0, T^{-1}(\hat{y}; \omega_0, \theta_H)) = \lambda(\hat{y}) F_W(\omega_0 | X^L = T^{-1}(\hat{y}; \omega_0, \theta_L)).$$

Letting $T(x; \omega_0, \theta) \equiv F_{X^\theta}(x | W \leq \omega_0)$, Bayes' rule implies that $\lambda(\hat{y}) = \hat{y} F_W(\omega_0) - \hat{y} F_W(\omega_0) = 0$. In words, when the policy function is written as a function of the signal using the ‘‘correct’’ normalization, then optimality of the policy function does not pose a binding constraint when minimizing the difference between expected payoffs at signals θ_H and θ_L .

When $\lambda(\hat{y}) = 0$, by Bayes' rule, (9) holds if and only if:

$$\begin{aligned} \int_{\Omega} r(\omega) dF_W(\omega | X^L \leq T^{-1}(\hat{y}; \omega_0, \theta_L)) \geq \\ \frac{\Pr(X^H \leq T^{-1}(\hat{y}; \omega_0, \theta_H))}{\Pr(X^L \leq T^{-1}(\hat{y}; \omega_0, \theta_L))} \int_{\Omega} r(\omega) dF_W(\omega | X^H \leq T^{-1}(\hat{y}; \omega_0, \theta_H)). \end{aligned} \quad (10)$$

¹⁵We have ignored the possibility that $\int_{\Omega} r(\omega) dF_W(\omega | T(X^L; \theta_L) = y)$ might be discontinuous in y . Allowing for this possibility by making the constraint an inequality does not affect the analysis.

¹⁶The approach we take here (looking for a multiplier, λ , that makes (9) positive) can alternatively be phrased in terms of Gollier and Kimball's (1995) ‘‘Diffidence Theorem’’; see also Jewitt (1986) or Athey (1999) for a discussion of this approach using tools from functional analysis.

Then, using Bayes' rule, we have

$$\frac{\Pr(X^H \leq T^{-1}(\hat{y}; \omega_0, \theta_H))}{\Pr(X^L \leq T^{-1}(\hat{y}; \omega_0, \theta_L))} = \frac{F_{W, X^H}(\omega_0, T^{-1}(\hat{y}; \omega_0, \theta_H)) / F_W(\omega_0 | X^H \leq T^{-1}(\hat{y}; \omega_0, \theta_H))}{F_{W, X^L}(\omega_0, T^{-1}(\hat{y}; \omega_0, \theta_L)) / F_W(\omega_0 | X^L \leq T^{-1}(\hat{y}; \omega_0, \theta_L))}. \quad (11)$$

Using Bayes' rule again and substituting in the definition of T , we have

$$\frac{F_{W, X^H}(\omega_0, T^{-1}(\hat{y}; \omega_0, \theta_H))}{F_{W, X^L}(\omega_0, T^{-1}(\hat{y}; \omega_0, \theta_L))} = \frac{F_{X^H}(F_{X^H}^{-1}(\hat{y} | W \leq \omega_0) | W \leq \omega_0)}{F_{X^L}(F_{X^L}^{-1}(\hat{y} | W \leq \omega_0) | W \leq \omega_0)} = 1. \quad (12)$$

Substituting (12) into (11), and then in turn substituting (11) into (10) yields

$$\int_{\Omega} r(\omega) dF_W(\omega | X^L \leq F_{X^L}^{-1}(\hat{y} | W \leq \omega_0)) \geq \frac{F_W(\omega_0 | X^L \leq F_{X^L}^{-1}(\hat{y} | W \leq \omega_0))}{F_W(\omega_0 | X^H \leq F_{X^H}^{-1}(\hat{y} | W \leq \omega_0))} \int_{\Omega} r(\omega) dF_W(\omega | X^H \leq F_{X^H}^{-1}(\hat{y} | W \leq \omega_0)) \quad (13)$$

Finally, Lemma 4 implies that (13) holds for all \hat{y} and all $r \in R^{SCQ(\omega_0)}$, if and only if (MIO-scq') holds. Thus, we have established that (MIO-scq') is equivalent to the result that $L(U^{R^{SCQ(\omega_0)}}, \beta^{\hat{y}}, \theta_H, \theta_L) \geq 0$ for all \hat{y} , which in turn implies that X^H is more informative than X^L for all agents in this restricted class of decision-makers (with utility functions given by $ar(\omega)$, for $r \in R^{SCQ(\omega_0)}$).

Consider a more direct approach to proving that (MIO-scq) is a necessary condition (as opposed to applying Lemma 4) for the informativeness comparison. We start from the result that given the appropriate T , (13) is satisfied if and only if an agent with payoff $a \cdot r$ has greater ex ante expected utility with X^H than with X^L , when she uses the policy $\beta^{\hat{y}}$. Suppose that (13) fails for some $r \in R^{SCQ(\omega_0)}$, \hat{y} , and ω_0 . Since the policy $\beta^{\hat{y}}$ is not necessarily optimal for the agent, there is no immediate contradiction. We now construct a new payoff function, for which $\beta^{\hat{y}}$ is optimal when the agent observes X^L . Observe that, given any constant K , adding $K \cdot \mathbf{1}_{\{\omega \leq \omega_0\}}(\omega)$ to r does not affect whether (13) holds; indeed, the transformation T was chosen for just this reason. Then, define $\tilde{r}(\omega) \equiv r(\omega) - \tilde{K} \cdot \mathbf{1}_{\{\omega \leq \omega_0\}}(\omega)$, where

$$\tilde{K} = \int_{\Omega} r(\omega) dF_W(\omega | X^L = F_{X^L}^{-1}(\hat{y} | W \leq \omega_0)) \Big/ F_W(\omega_0 | X^L = F_{X^L}^{-1}(\hat{y} | W \leq \omega_0)),$$

so that investing when $X^L \geq F_{X^L}^{-1}(\hat{y} | W \leq \omega_0)$ (that is, using the policy $\beta^{\hat{y}}$) is optimal policy for the agent with payoffs $a\tilde{r}$, with signal X^L . Observe that $\tilde{r} \in R^{SCQ(\omega_0)}$. Then, ex ante expected utility for an agent with payoffs $a\tilde{r}$ are lower with X^H than with X^L using the policy $\beta^{\hat{y}}$, because (13) fails. If $\theta_H = \theta_L + d\theta$ (that is, we consider a small change in θ), the envelope theorem implies that increasing θ decreases ex ante expected payoffs for the agent with payoff function $a\tilde{r}$.

Thus far, we have established informativeness orderings for a special class of decision problems, where utility functions take the form $g(a, \omega) \equiv a \cdot r(\omega)$. It is straightforward (see AL) to show that

the monotone information orders are further sufficient for larger classes of decision problems. To see this most easily, observe that when (A1-A4) hold, we may integrate by parts to yield $\tilde{V}(\theta; g, \beta, T) =$

$$- \int_{[0,1]} \int_{\Omega} \frac{\partial}{\partial y} g(\beta(y), \omega) d_{\omega} F_{W, X^{\theta}}(\omega, T^{-1}(y; \theta)) dy. \quad (14)$$

(To see an approach without using integration by parts, see the proof of Proposition 13 below.) Then, the inner integral of (14) is nonincreasing in θ whenever (9) holds for $r(\omega) \equiv \frac{\partial}{\partial y} g(\beta(y), \omega)$, $\hat{y} = y$, and $T(x; \omega_0, \theta) \equiv F_{X^{\theta}}(x|W \leq \omega_0)$ (so that $\lambda(\hat{y}) = 0$). But we have shown that for this T , if $r \in R^{SCQ(\omega_0)}$, (MIO-scq) implies that (9) holds. ■

Proof of Proposition 8: We prove part (ii); part (i) is analogous. Fix $\theta^H > \theta^L$. From Lemma 9 (ii)(d), we have that (MIO-scq) is equivalent to, for all $x \in \mathcal{X}^L$,

$$F_{X^H}^{-1}(F_{X^L}(x|W \leq \omega_0)|W \leq \omega_0) \text{ is nondecreasing in } \omega_0.$$

Take the derivative. First, by definition, $F_X^{-1}(F_X(x|W \leq \omega_0) | W \leq \omega_0) = x$, so that

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

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

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

Then, letting $z = F_{X^L}(x|W \leq \omega_0)$, we have

$$\begin{aligned} & \frac{\partial}{\partial \omega_0} F_{X^H}^{-1}(F_{X^L}(x|W \leq \omega_0)|W \leq \omega_0). \\ &= - \frac{\frac{\partial}{\partial \omega} F_{X^H}(F_{X^H}^{-1}(z|W \leq \omega_0) | W \leq \omega) \Big|_{\omega=\omega_0}}{f_{X^H}(F_{X^H}^{-1}(z|W \leq \omega_0) | W \leq \omega_0)} + \frac{\frac{\partial}{\partial \omega} F_{X^L}(F_{X^L}^{-1}(z|W \leq \omega_0) | W \leq \omega) \Big|_{\omega=\omega_0}}{f_{X^H}(F_{X^H}^{-1}(z|W \leq \omega_0) | W \leq \omega_0)}. \end{aligned}$$

This is nonnegative if:

$$\frac{\partial}{\partial \omega} F_{X^L}(F_{X^L}^{-1}(z|W \leq \omega_0) | W \leq \omega) \Big|_{\omega=\omega_0} \geq \frac{\partial}{\partial \omega} F_{X^H}(F_{X^H}^{-1}(z|W \leq \omega_0) | W \leq \omega) \Big|_{\omega=\omega_0},$$

or,

$$\frac{\partial}{\partial \omega} \left(\frac{F_{X^L}(F_{X^L}^{-1}(z|W \leq \omega_0), \omega)}{H(\omega)} \right) \Big|_{\omega=\omega_0} \geq \frac{\partial}{\partial \omega} \left(\frac{F_{X^H}(F_{X^H}^{-1}(z|W \leq \omega_0), \omega)}{H(\omega)} \right) \Big|_{\omega=\omega_0}.$$

Simplifying, we have

$$\begin{aligned} & -\frac{F_{W,X^L}(\omega_0, F_{X^L}^{-1}(z|W \leq \omega_0))}{H(\omega_0)} \frac{h(\omega_0)}{H(\omega_0)} + \frac{\frac{\partial}{\partial \omega} F_{W,X^L}(\omega, F_{X^L}^{-1}(z|W \leq \omega_0))|_{\omega=\omega_0}}{H(\omega_0)} \\ \geq & -\frac{F_{W,X^H}(\omega_0, F_{X^H}^{-1}(z|W \leq \omega_0))}{H(\omega_0)} \frac{h(\omega_0)}{H(\omega_0)} + \frac{\frac{\partial}{\partial \omega} F_{W,X^H}(\omega, F_{X^H}^{-1}(z|W \leq \omega_0))|_{\omega=\omega_0}}{H(\omega_0)} \end{aligned}$$

or, multiplying through by $H(\omega)$, factoring out a term, and using Bayes' rule, this is:

$$\begin{aligned} & F_{W,X^L}(\omega_0, F_{X^L}^{-1}(z|W \leq \omega_0)) \left[\frac{\frac{\partial}{\partial \omega} F_{W,X^L}(\omega, F_{X^L}^{-1}(z|W \leq \omega_0))|_{\omega=\omega_0}}{F_{W,X^L}(\omega_0, F_{X^L}^{-1}(z|W \leq \omega_0))} - \frac{h(\omega_0)}{H(\omega_0)} \right] \\ \geq & F_{W,X^H}(\omega_0, F_{X^H}^{-1}(z|W \leq \omega_0)) \left[\frac{\frac{\partial}{\partial \omega} F_{W,X^H}(\omega, F_{X^H}^{-1}(z|W \leq \omega_0))|_{\omega=\omega_0}}{F_{W,X^H}(\omega_0, F_{X^H}^{-1}(z|W \leq \omega_0))} - \frac{h(\omega_0)}{H(\omega_0)} \right]. \end{aligned} \quad (15)$$

But, using (12), $F_{W,X^L}(\omega_0, F_{X^L}^{-1}(z|W \leq \omega_0)) = F_{W,X^H}(\omega_0, F_{X^H}^{-1}(z|W \leq \omega_0)) = zH(\omega_0)$, and so (15) becomes (after applying Bayes' rule again)

$$\frac{f_W(\omega_0 | X^L \leq F_{X^L}^{-1}(z|W \leq \omega_0))}{F_W(\omega_0 | X^L \leq F_{X^L}^{-1}(z|W \leq \omega_0))} \geq \frac{f_W(\omega_0 | X^H \leq F_{X^H}^{-1}(z|W \leq \omega_0))}{F_W(\omega_0 | X^H \leq F_{X^H}^{-1}(z|W \leq \omega_0))}.$$

Finally, using Bayes' rule and the fact that the expectation of the posteriors must equal the prior, the latter inequality is equivalent to (3). ■

Proof of Lemma 9: Consider part (ii) (part (i) is analogous). First, consider the relationship between (b) and (d). Suppose that (b) holds. This is equivalent to requiring that, for all $\theta_H > \theta_L$ and $v > \omega_0$,

$$F_{X^H}(F_{X^H}^{-1}(y|W \leq \omega_0)|W \leq v) \leq F_{X^L}(F_{X^L}^{-1}(y|W \leq \omega_0)|W \leq v).$$

But then, apply $F_{X^H}^{-1}(\cdot|W \leq v)$ to both sides, and letting $x_y = F_{X^L}^{-1}(y|W \leq \omega_0)$, we have

$$F_{X^H}^{-1}(F_{X^L}(x_y|W \leq \omega_0)|W \leq \omega_0) \leq F_{X^H}^{-1}(F_{X^L}(x_y|W \leq v)|W \leq v),$$

as desired for (d). Finally, apply $F_{X^H}(\cdot|W \leq \omega_0)$ to both sides, and let $y' = F_{X^L}(x_y|W \leq v)$. This reduces to

$$F_{X^L}(F_{X^L}^{-1}(y'|W \leq v)|W \leq \omega_0) \leq F_{X^H}(F_{X^H}^{-1}(y'|W \leq v)|W \leq \omega_0). \quad (16)$$

Since $\omega_0 < v$, we can exchange the roles of ω_0 and v , so that (16) holds for all $\theta_H > \theta_L$, if and only if (c) holds. Thus, (b) and (c) are equivalent.

Now consider the relationship between (a) and (b). First we show that (a) implies (b), building on the proof of Proposition 7. In particular, fix $v > \omega_0$, and recall that, using the transformation $F_{X^\theta}(\cdot|W \leq \omega_0)$, (MIO-scq') implies that (10) holds for all $r \in R^{SCQ(\omega_0)}$. Then, check (10) for $r_v(\omega) \equiv \mathbf{1}_{\{\omega \leq v\}}(\omega)$. Finally, use Bayes' rule and the fact that $\Pr(\omega \leq v)$ is constant in θ to write the condition in terms of conditional probabilities, as in (b).

Now, we show that (b) implies (a). Fix $v > \omega_0$. Then, we have just argued that $F_{X^\theta}(F_{X^\theta}^{-1}(y|W \leq \omega_0)|W \leq v)$ is nonincreasing in θ , if and only if (10) holds for r_v . The proof of Proposition 7 establishes that (10) is equivalent to (13). In turn, (13) for r_v is equivalent to

$$\frac{F_W(v | X^L \leq F_{X^L}^{-1}(\hat{y}|W \leq \omega_0))}{F_W(\omega_0 | X^L \leq F_{X^L}^{-1}(\hat{y}|W \leq \omega_0))} - \frac{F_W(v | X^H \leq F_{X^H}^{-1}(\hat{y}|W \leq \omega_0))}{F_W(\omega_0 | X^H \leq F_{X^H}^{-1}(\hat{y}|W \leq \omega_0))} \geq 0 \quad (17)$$

Finally, consider the case where $v < \omega_0$, and recall that (b) implies (c). Using Bayes' rule, condition (c) requires that (10) fails for r_v . This is equivalent to requiring that (13) fails. Finally, that implies that (17) fails.

Then, we have established that if (b) holds, then $F_W(\cdot | X^\theta \leq F_{X^\theta}^{-1}(\hat{y}|W \leq \omega_0))$ is ordered by $\succ_{PR-\omega_0}$ in θ , as desired for (MIO-scq'), which is equivalent to (a). ■

Proof of Lemma 10: We prove part (ii); part (i) is analogous. Recall the characterization of (MIO-scq) from Lemma 9, part (ii)(b). Notice that, if we let $x_y = \check{x}(y; \theta, \omega_0) = F_{X^\theta}^{-1}(y|W \leq \omega_0)$,

$$\begin{aligned} & \frac{\partial}{\partial \theta} \Pr(X^\theta \leq F_{X^\theta}^{-1}(y|W \leq \omega_0) | W \leq v) \\ &= \frac{\partial}{\partial \theta} F_{X^\theta}(x_y | W \leq v) + f_{X^\theta}(x_y | W \leq v) \frac{\partial}{\partial \theta} F_{X^\theta}^{-1}(y | W \leq \omega_0) \\ &= \frac{\partial}{\partial \theta} F_{X^\theta}(x_y | W \leq v) - f_{X^\theta}(x_y | W \leq v) \frac{\frac{\partial}{\partial \theta} F_{X^\theta}(x_y | W \leq \omega_0)}{f_{X^\theta}(x_y | W \leq \omega_0)}. \end{aligned}$$

Thus, $\frac{\partial}{\partial \theta} \Pr(X^\theta \leq F_{X^\theta}^{-1}(y|W \leq \omega_0) | W \leq v) \leq 0$ for all $v > \omega_0$, if and only if (5) holds for all x . ■

Proof of Proposition 11: Necessity follows from Proposition 7 because, as argued above, we can approximate any payoff $g \in U^{R^{SCQ}}$ with an appropriate u and π . Now consider sufficiency. Before beginning, let us assume without loss of generality that X^θ is normalized so that $\frac{\partial}{\partial \theta} F_{X^\theta}(x) = 0$ (this does not affect the information content of X^θ , so it also does not affect whether MPR or (MIO-scq) hold). If we use the fact that $\Pr(W \leq v)$ does not vary with θ , Lemma 10 implies that (MIO-scq) holds if and only if

$$-\frac{\frac{\partial}{\partial \theta} F_{W, X^\theta}(v, x)}{F_{W|X^\theta}(v|x)} \text{ is nondecreasing in } v \text{ for all } x. \quad (18)$$

Now, observe that by the envelope theorem, it follows that

$$\frac{\partial}{\partial \theta} V^*(\theta; u \circ \pi) = \int_x \int_\omega u(\pi(\alpha^*(x; \theta), \omega)) d \left[\frac{\partial}{\partial \theta} F_{W, X^\theta}(\omega, x) \right].$$

Integrating by parts (which is valid by (A3)) and letting subscripts denote partial derivatives, we have:

$$- \int_x \alpha_x^*(x; \theta) \int_\omega u'(\pi(\alpha^*(x; \theta), \omega)) \pi_a(\alpha^*(x; \theta), \omega) d\omega \left[\frac{\partial}{\partial \theta} F_{W, X^\theta}(\omega, x) \right] dx. \quad (19)$$

Taking the inner integral and integrating it by parts again (grouping together the last two terms of the integrand, analogous to the proof of Proposition 6), we have the following condition which we will refer to as (II):

$$-u'(\pi(\alpha^*(x; \theta), \bar{\omega})) \int \pi_a(\alpha^*(x; \theta), \omega) d\omega \left[\frac{\partial}{\partial \theta} F_{W, X^\theta}(\omega, x) \right] \quad (\text{II-a})$$

$$+ \int_\omega \left[\begin{array}{c} u''(\pi(\alpha^*(x; \theta), \omega)) \pi_\omega(\alpha^*(x; \theta), \omega) \\ \cdot \left[\int_{-\infty}^\omega \pi_a(\alpha^*(x; \theta), t) dt \frac{\frac{\partial}{\partial \theta} F_{W, X^\theta}(t, x)}{\frac{\partial}{\partial \theta} F_{W, X^\theta}(\omega, x)} \right] \frac{\partial}{\partial \theta} F_{W, X^\theta}(\omega, x) \end{array} \right] d\omega \quad (\text{II-b})$$

To ensure that $\frac{\partial}{\partial \theta} V^*(\theta; u \circ \pi) \geq 0$, it suffices to establish that this expression is positive whenever the first order conditions (FOC) hold. Those, too, can be integrated by parts (as in the proof of the comparative statics proposition). Multiplying by -1 , the marginal returns to a are given by:

$$-u'(\pi(\alpha^*(x; \theta), \bar{\omega})) \int \pi_a(\alpha^*(x; \theta), \omega) d\omega F_{W|X^\theta}(\omega|x) \quad (\text{FOC-a})$$

$$+ \int_\omega \left\{ \begin{array}{c} u''(\pi(\alpha^*(x; \theta), \omega)) \pi_\omega(\alpha^*(x; \theta), \omega) \\ \cdot \left[\int_{-\infty}^\omega \pi_a(\alpha^*(x; \theta), t) dt \frac{F_{W|X^\theta}(t|x)}{F_{W|X^\theta}(\omega|x)} \right] \end{array} \right\} F_{W|X^\theta}(\omega|x) d\omega \quad (\text{FOC-b})$$

Notice that the first term of this expression (FOC-a) is non-positive, since u is nondecreasing and the expected marginal returns to a are positive by assumption. Thus, a necessary condition for the sum of the two terms to equal zero is that the second term (FOC-b) is non-negative.

Our approach to the remainder of the proof is as follows. Since we have established that (FOC) holds only if (FOC-b) ≥ 0 , it suffices to show that (FOC-b) ≥ 0 implies that (II) ≥ 0 . Let us proceed by showing first that (II-a) ≥ 0 always under our assumptions, and second showing that (FOC-b) ≥ 0 implies (II-b) ≥ 0 .

(a) Consider (II-a). Integrating by parts again yields

$$-u'(\pi(\alpha^*(x; \theta), \bar{\omega})) \left[\pi_a(\alpha^*(x; \theta), \bar{\omega}) \frac{\partial}{\partial \theta} F_{X^\theta}(x) - \int \pi_{a\omega}(\alpha^*(x; \theta), \omega) \frac{\partial}{\partial \theta} F_{W, X^\theta}(\omega, x) d\omega \right] \geq 0, \quad (20)$$

where the sign follows since $\frac{\partial}{\partial \theta} F_{X^\theta}(x) = 0$ by our normalization of x , by supermodularity of π , and because $\frac{\partial}{\partial \theta} F_{W, X^\theta}(\omega, x) \geq 0$ (as discussed above, the latter inequality is equivalent to (MIO-spm) under our normalization of x , $\frac{\partial}{\partial \theta} F_{X^\theta}(x) = 0$).

(b) Now consider the second term. Define $K_L(\omega) \equiv K(\omega|\kappa_L) \equiv F_{W|X^\theta}(\omega|x)$, and define $K_H(\omega) \equiv K(\omega|\kappa_H) \equiv \frac{\partial}{\partial \theta} F_{W, X^\theta}(\omega, x)$. Notice that $\frac{K_H(\cdot)}{K_L(\cdot)}$ is non-negative and nonincreasing (by (18)). Let $\varphi(\omega|\kappa) \equiv \int_{-\infty}^{\omega} -\pi_a(\alpha^*(x; \theta), t) dt \frac{K(t|\kappa)}{K(\omega|\kappa)}$. First, observe that $\varphi(\omega|\kappa_H) \geq \varphi(\omega|\kappa_L)$ (this follows because $\frac{K_H(t)}{K_H(\omega)} \geq \frac{K_L(t)}{K_L(\omega)}$ for $t < \omega$ and $-\pi_a(a, t)$ is nonincreasing in t). Second, note that $\varphi(\omega|\kappa_L)$ is weak single crossing in $-\omega$, since $-\pi_a(a, t)$ is nonincreasing in t . Next we apply Lemma 14. Let $g(\omega, \kappa) \equiv -u''(\pi(\alpha^*(x; \theta), \omega)) \cdot \pi_\omega(\alpha^*(x; \theta), \omega) \cdot \varphi(\omega|\kappa)$, where $g(\omega, \kappa_L)$ is weak single crossing in $-\omega$ and $g(\omega, \kappa)$ is nondecreasing in κ by our previous arguments, and since $u'' \leq 0$ and $\pi_\omega \geq 0$. Then, by Lemma 14, since $\frac{K_H(\omega)}{K_L(\omega)}$ is nondecreasing in ω , $\int_{\omega} g(\omega, \kappa) K(\omega|\kappa) d\omega$ satisfies weak single crossing in κ . In other words, (FOC-b) ≥ 0 implies (II-b) ≥ 0 , as desired.

Thus, we have established that (II) ≥ 0 , which implies that $\frac{\partial}{\partial \theta} V^*(\theta; u \circ \pi) \geq 0$. ■

Proof of Proposition 12: Fix $\omega_0 \in \Omega$, and define $Y^\theta = F_{X^\theta}(X^\theta|W \leq \omega_0)$. Let G_{W, Y^θ} be the joint distribution of W and Y^θ . Let $\beta^*(\cdot; \theta, \rho)$ be the optimal policy for an agent with utility parameter ρ , who observes the signal Y^θ . Further, define $w(y, \omega) = g(\beta^*(y; \theta, \rho_H), \omega; \rho_H) - g(\beta^*(y; \theta, \rho_L), \omega; \rho_L)$. By the Envelope Theorem,

$$\frac{\partial}{\partial \theta} V^*(\theta; g(\cdot; \rho_H)) - \frac{\partial}{\partial \theta} V^*(\theta; g(\cdot; \rho_L)) = \int_{\Omega} \int_{[0,1]} w(y, \omega) d \frac{\partial}{\partial \theta} \{G_{W, Y^\theta}(\omega, y)\}.$$

Assume \mathcal{A} is finite. Then there is some series of cut-points $\{y_i\}_{i=1}^n$, with $y_1 \leq y_2 \leq \dots \leq y_{N+1}$, such that $w(y, \omega) = w(y_i, \omega)$ on $[y_i, y_{i+1})$. Let $r_i(\omega) = w(y_i, \omega) - w(y_{i-1}, \omega)$, and note that $r_i \in R^{SCQ(\omega_0)}$ for all $i = 2, \dots, N$. Then, $\frac{\partial}{\partial \theta} V(\theta, g(\cdot; \rho_H)) - \frac{\partial}{\partial \theta} V(\theta, g(\cdot; \rho_L)) \geq 0$ if

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

which, we showed in the proof of Proposition 7, holds for all $r_i \in R^{SCQ(\omega_0)}$ and all y_i , if and only if (MIO-scq) holds (simply use the fact that the expectation of the posteriors must equal the prior to rearrange (12), (11), and (13) so that in each case we condition on the probability that the signal is high). The case where \mathcal{A} is convex follows via a limiting argument. ■

Proof of Proposition 13: Use subscripts to denote partial derivatives. For each ρ , the first-order conditions for optimality define $\alpha^*(x; \theta, \rho)$. Building on our integration by parts ((FOC-a) and (FOC-b) from above), and using our assumptions about the upper limit on the marginal utility,

the first-order conditions can be written as follows for each ρ :

$$\int_{\omega} u_{11}(\pi(\alpha^*(x; \theta, \rho), \omega); \rho) \alpha^*(x; \theta, \rho) \left[\int_{-\infty}^{\omega} (t - \omega_0) d_t \frac{F_{W|X^\theta}(t|x)}{F_{W|X^\theta}(\omega|x)} \right] F_{W|X^\theta}(\omega|x) d\omega = 0.$$

Since the first order conditions hold as an identity, we can differentiate with respect to ρ , yielding:

$$\int_{\omega} \Psi(x, \omega; \theta, \rho) \left[\int_{-\infty}^{\omega} (t - \omega_0) d_t \frac{F_{W|X^\theta}(t|x)}{F_{W|X^\theta}(\omega|x)} \right] F_{W|X^\theta}(\omega|x) d\omega = 0. \quad (21)$$

Likewise, we can simplify (II-a) and (II-b), and use the fact that $\frac{\partial}{\partial \theta} F_{X^\theta}(x) = 0$, to show that the increasing information condition (II) is increasing in ρ if

$$\int_{\omega} \Psi(x, \omega; \theta, \rho) \left[\int_{-\infty}^{\omega} (t - \omega_0) d_t \frac{\frac{\partial}{\partial \theta} F_{W, X^\theta}(t, x)}{\frac{\partial}{\partial \theta} F_{W, X^\theta}(\omega, x)} \right] \frac{\partial}{\partial \theta} F_{W, X^\theta}(\omega, x) d\omega \geq 0 \quad (22)$$

The proof of Proposition 11 (starting with part (b)) showed that when an expression analogous to (21) holds, then an expression analogous to (22) holds. In particular, for the special case of the portfolio problem, (FOC-b) and (II-b) are identical to (21) and (22) when $u''(\pi(\alpha^*(x; \theta, \rho), \omega)) \cdot \pi_{\omega}(\alpha^*(x; \theta, \rho), \omega)$ is replaced with $\Psi(x; \theta, \rho)$. Thus, if we simply establish that $\Psi(x, \omega; \theta, \rho) \leq 0$, we will have (21) implies (22).

However, there is one more effect to consider. (II) is just the inner integral of (19), which represents the change in ex ante expected payoffs from an increase in θ . Our arguments thus far provided conditions under which (II) is nondecreasing in ρ . When (MIO-scq) holds, (II) is positive, and if the posteriors are ordered by MPR, $\alpha_x^* \geq 0$. So (19) is nondecreasing in ρ if $\Psi(x, \omega; \theta, \rho) \leq 0$ and $\alpha_{x\rho}^* \geq 0$. ■

7 References

- ATHEY, SUSAN (1999): "Characterizing Properties of Stochastic Objective Functions," MIT Working Paper 96-1R.
- ATHEY, SUSAN (forthcoming): "Monotone Comparative Statics under Uncertainty," *Quarterly Journal of Economics*.
- ATHEY, SUSAN AND JONATHAN LEVIN (1998): "The Value of Information in Monotone Decision Problems," MIT Working Paper 98-24.
- BLACKWELL, DAVID (1951): "Comparisons of Experiments," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics*,
- BLACKWELL, DAVID (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

- BOYER, M. AND R. E. KIHLMSTROM, ED. (1984): *Bayesian Models in Economic Theory*. New York: North-Holland.
- ECKHOUDT, LOUIS AND CHRISTIAN GOLLIER (1995): "Demand for Risky Assets and the Monotone Probability Ratio Order," *Journal of Risk and Uncertainty* 11: 113-122.
- GOLLIER, CHRISTIAN (1995): "The Comparative Statics of Changes in Risk Revisited." *Journal of Economic Theory* 66 (2): 522-35.
- GOLLIER, C. AND M. KIMBALL (1995): "Toward a Systematic Approach to the Economic Effects of Uncertainty I: Comparing Risks," Mimeo, University of Toulouse.
- JEWITT, IAN (1986): "A Note on Comparative Statics and Stochastic Dominance," *Journal of Mathematical Economics*, 15, 249-254.
- JEWITT, IAN (1997): "Information and Principal Agent Problems," mimeo, University of Bristol.
- KARLIN, SAMUEL (1968): *Total Positivity: Volume I*, Stanford University Press.
- KIHLMSTROM, RICHARD (1974): "A General Theory of Demand for Information about Product Quality," *Journal of Economic Theory* 8 (4), pp. 413-39.
- KIM, SON KU (1995): "Efficiency of an Information System in an Agency Model," *Econometrica*, 63, 89-102.
- 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.
- MILGROM, PAUL (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 380-391.
- MILGROM, PAUL (1994): "Comparing Optima: Do Simplifying Assumptions Affect Conclusions?" *Journal of Political Economy*, 3 (102), 609-615.
- ORMISTON, M. AND E. SCHLEE (1992): "Necessary Conditions for Comparative Statics under Uncertainty," *Economic Letters* 40(4): 429-34.
- ORMISTON, M. AND E. SCHLEE (1993): "Comparative Statics Under Uncertainty for a Class of Economic Agents," *Journal of Economic Theory* 61: 412-422.
- PERSICO, NICOLA (1996): "Information Acquisition in Affiliated Decision Problems," mimeo, UCLA.
- PERSICO, NICOLA (2000): "Information Acquisition in Auctions," *Econometrica*.
- SANDMO, AGNAR (1971): "On the Theory of the Competitive Firm Under Price Uncertainty" *American Economic Review*, 61: 65-73.
- SHANNON, CHRIS (1995): "Weak and Strong Monotone Comparative Statics," *Economic Theory* 5 (2), March, pp. 209-27.