Discussion Paper No. 1149

INFORMATION ACQUISITION IN AFFILIATED DECISION PROBLEMS

by

Nicola Persico*

First version: September 1995

This version: February 1996

---

* Department of Economics, Northwestern University, Evanston, IL 60208-2600. *email:* npersico@casbah.acns.nwu.edu. I wish to thank Steve Matthews, Wolfgang Pesendorfer, Juuso Valimaki and Asher Wolinsky. A Sloan Foundation Fellowship is gratefully acknowledged.

Information Acquisition in Affiliated Decision Problems

Nicola Persico*

First version: September 1995.
This version, February 1996

Abstract

This paper investigates information acquisition in decision problems. We introduce a new notion of "better information", Accuracy-order (A-order), defined on continuous families of signals. Accuracy formalizes the idea that "a signal that is more correlated with the unknown random variable is better". This concept is indigenous to an economically interesting subset of all decision problems, those where signals are affiliated and the payoff function satisfies the single-crossing property. On this subset, this notion is found to be "tight", in the sense that A-order is an if-and-only-if condition for better information. Thus, a Blackwell-type result is obtained. On the subset, it is shown that Blackwell's Sufficiency is a special case of Accuracy. Finally, a comparative statics result is obtained, about which decision problem will induce more information acquisition.

1 Introduction

The present paper addresses the question of information acquisition in decision problems. The idea is that a signal is acquired before taking a decision, and the decision maker can choose its accuracy at a cost. We provide a unified

*Department of Economics, Northwestern University, Evanston IL 60208. e-mail npersico@casbah.acns.nwu.edu. I wish to thank Steve Matthews, Wolfgang Pesendorfer, Juuso Valimaki and Asher Wolinsky. A Sloan Foundation Fellowship is gratefully acknowledged.

model to address questions such as "what is useful information in a decision problem?" and "given two decision problems, which one will induce more information acquisition?".

Besides the theoretical interest, large amounts of resources are devoted to information acquisition. Consider a monopolist facing an unknown demand function; most likely he will acquire information about it, for example through a market research. But marketing departments in organizations are just one example: pre-auction exploration of oil tracts and preparation of biddings in procurement auctions are others. Many of these situations have a game-theoretic flavour. The framework developed here is applicable to games of incomplete information; a companion paper uses this framework to ask the question "which auction form (1st or 2nd price) will induce more information acquisition?", and draws conclusions that may (and in some circumstances do) overturn revenue-ranking. Endogeneizing information acquisition is bound to modify our perspective on the basic results of information economics.

The main contributions of this paper are:

a) offer a general framework to study information acquisition using differential methods. This allows us to reason in familiar terms of "marginal revenue" and "marginal cost" of information.

b) introduce Accuracy, a new concept of "better information", which refines Blackwell's notion of Sufficiency on a subset of decision problems. On this subset, Accuracy is a "tight" (if-and-only-if) notion.

c) spell out conditions under which one decision problem induces "more information acquisition" than another.

We posit that a signal X_η about an unobservable random variable V is chosen at a cost before taking an action a; any one signal in the family $\{X_\eta\}_{\eta\in E}$ may be chosen, at a cost $C(\eta)$. The goal is to maximize the expected value of a payoff function, $u(v,a)$, which depends on the realization of V and on the action. The more informative X_η is, the higher the expected payoff. Let $R(\eta)$ denote the expected value of the payoff function, given an optimal choice of a with signal X_η. We can formalize our information acquisition problem as

$$\max_{\eta\in E} R(\eta) - C(\eta).$$

We assume that E is an interval of the real line.

We give conditions under which $R(\eta)$ increases with η. In other words, we assess when increasing η to θ produces "better information", in the sense of raising the expected payoff to the decision maker. The answer to this question depends, in principle, on the characteristics of the payoff functions and on the statistical structure of the problem (the family of signals and the prior). Blackwell's well known answer is independent of all these primitives, and furthermore is a necessary and sufficient condition. It defines the notion of Sufficiency, and the answer to the question then is "for any decision problem, if and only if signal X_θ is sufficient for X_η". Our approach here is to place constraints on the payoff function and on the family of signals, that enable us to answer the question in the following way "for any decision problem in a restricted class, if and only if the family of signals is A-ordered". This is a Blackwell-type result. The next two paragraphs intuitively discuss the restriction on the payoff function, and the new concept of "better information". We then put the two together to provide intuition for Theorem 1: increasing Accuracy increases the payoff.

Roughly speaking, in a problem where the optimal strategy is increasing in the signal, our restriction on the payoff function is equivalent to the single-crossing property in $(a; r)$. The single-crossing property has fundamental relevance in the theory of monotone comparative statics (see Milgrom and Shannon [9]), and is satisfied in a number of economic situations. Loosely speaking, it can be phrased as "for each fixed a, $\frac{\partial}{\partial a}u(r,a)$ changes sign at most once, from negative to positive, as r increases". Consider for example the monopolist's case discussed in the beginning: here a is the quantity to produce, and let the demand function be $P(a) = r - a$, where r is the unknown parameter. Marginal costs are 0. Then $u(r,a) = aP(a) = ar - a^2$, hence $\frac{\partial}{\partial a}u(r,a)$ is just the marginal revenue: it is increasing in r, whereby the single-crossing property is satisfied.

The concept of A-order (Accuracy-order) is the new notion of "better information". Consider the transformation $T_{\eta,\theta,x}(\cdot)$ which, given two signals X_η and X_θ is uniquely defined by $T_{\eta,\theta,x}(X_\eta \mid r) \sim X_\theta \mid r$. This is the function that transforms a signal into another (possibly better) signal. We say that X_θ is more accurate than X_η when $\frac{\partial T_{\eta,\theta,x}(x)}{\partial r} > 0$. This means that the T transformation is positively correlated with r: when r is high, $T(x)$ will be a bit higher than x would have been, and a bit lower when r is low. In this sense, we can see the action of $T_{\eta,\theta,x}(\cdot)$ as "adding correlation with r" to X_η.

Now, we are ready to describe the first result of this work: why increasing Accuracy ("adding correlation") is useful when the decision payoff satisfies the single-crossing condition. The first-order conditions for the decision problem are the average of $\frac{\partial}{\partial a}u(r,a)$ with respect to r. By single-crossing, this quantity is negative for low r's and positive for high r's: this means that we would like to lower a when r is low, and increase a when r is high. But then this is exactly what $a(T_{\eta,\theta,x}(x))$ does, when X_θ is more accurate than X_η. In the monopolist's example, increasing Accuracy allows to produce more when the marginal revenue is higher, and less when it is low. This is the intuition behind Theorem 1.

We now want to understand which decision problem induces more information acquisition. We can loosely interpret the "single-crossingness" of a payoff as the degree to which it varies with the unknown v, i.e. its "variability". Then if a payoff function is "very single-crossing", i.e. $\frac{\partial}{\partial a}u(v,a)$ is highly sensitive to changes in v, it is reasonable that passing to a more accurate signal should give a higher increase in expected revenue than if u were "less single-crossing". We exploit this property in Theorem 3 to answer the question "which decision problem will induce more information acquisition".

Consider the set of decision problems where the family of signals is affiliated and the payoff is single-crossing. We have seen that increasing Accuracy is sufficient to guarantee increase in payoff: but is it necessary? Indeed, yes: we prove that on that set, Accuracy is a "tight" notion: if A-order fails, there exists a single-crossing payoff such that choosing a higher signal is detrimental. Thus we can't beat Accuracy. This is why this work is in the spirit of Blackwell's, only restricted to a subset of decision problems: those with affiliated signals and single-crossing payoff. Of course, we need to compare this new notion of "better information" with Blackwell's Sufficiency. Increasing Accuracy increases expected payoff on a *subset* of all decision problems, while Sufficiency works for *all* decision problems. Thus, given a signal, we may expect that the set of signals that represent an increase in Accuracy is *larger* than the sufficient ones: actually, we are able to demonstrate strict inclusion. Consider a family $\{X_\eta\}$ of signals affiliated with V: if $X_{\eta''}$ is sufficient for $X_{\eta'}$ then $X_{\eta''}$ is more accurate than $X_{\eta'}$, but the converse is not true. In this sense, we are able to prove that Accuracy refines Sufficiency on our subset of decision problems.

This paper makes two points: one more theoretical, the other more practical. The theoretical message is that, by restricting to an economically interesting subset of decision problems (those with affiliated signals and single-crossing payoff), we can refine Blackwell's notion of Sufficiency: this work introduces Accuracy, a tight (if-and-only-if) notion. The practical byproduct is a comparative statics result: the more the payoff is "sensitive to" (single-crossing in) the unknown variable, the more information will be acquired. As a special case, this holds when acquiring information in Blackwell's sense.

We start off by restricting the class of decision problems: some literature has investigated the opposite question, of how Blackwell's notion performs when the set of preferences is expanded to include non-expected-utility preferences (see Safra and Sulganik [12]). Blackwell's Sufficiency is found inadequate (not restrictive enough) in such settings. It is also worth pointing out that we analyze decision problems, and not games, like for example principal-agent games (e.g. Kim [7]). In a principal-agent model, information is acquired by the principal only in order better to monitor the agent. It follows that new information is useful only insofar as the agent believes the principal has acquired it, and adjusts his (the agent's) strategy accordingly. This is obviously a different case than our, where the principal (decision maker) plays against nature, and nature's behaviour does not change with the principal's information. Our theory can be used to analyze game-theoretic situations, for example auctions (see Persico [11]); however, the question of whether the information acquisition process is observable by the opponents is crucial in such an analysis.

The rest of the paper is organized as follows:

Section 2 spells out the model. Subsection 2.1 enumerates hypotheses on the payoff function, and 2.2 those on the signal structure.

Section 3 illustrates the definition of the T transformation, with the aid of an example (subsection 3.1); subsection 3.2 presents several other examples of A-ordered families of signals.

Section 4 contains the analysis. After two technical lemmas (subsection 4.1), subsection 4.2 contains the theorems that motivate this paper: Theorem 1 establishing the notion of Accuracy, and Theorem 2 showing that this notion is "tight" on the set of affiliated decision problems satisfying a monotonicity requirement on the payoff functions. Subsection 4.3 discusses this restriction on the payoff functions, under which Theorems 1 and 2 apply. Subsection 4.4 explains why, under affiliation, Accuracy refines Sufficiency; this is formalized in Proposition 5.

Section 5 investigates the information acquisition problem introduced in the beginning. It contains Theorem 3, a comparative statics result on which problem will induce more information acquisition.

Section 6 presents three applications: a monopolist faces an uncertain demand function, and can acquire information about it. The choice of a doctor under incomplete information about their skill. A take-it-or-leave-it offer with incomplete information about the value of the object.

Section 7 concludes.

2 The model

The model is the following:

Let us define a **payoff function** as a function

$$u(v, a) : \mathcal{V} \times \mathcal{A} \rightarrow \Re.$$

Here, a represents an action, and v is an unknown parameter, seen as the realization of a random variable V. Let $g(v)$ be a prior for V, with c.d.f. $G(v)$.

The decision maker cannot observe V, but can observe a **signal**, a random variable X_η with conditional density $f^\eta(x \mid v)$. The associate c.d.f. is denoted by $F^\eta(x \mid v)$. This signal will be chosen prior to observing its realization from a **family of signals** $\{X_\eta\}_{\eta \in E}$, where E is an interval of the real line. We will sometimes refer to a **statistical structure** as a prior for V together with a family of signals.

Script letters are reserved for supports of random variables and domains of functions. Thus, we denote with $\mathcal{V}$ the support of V, with $\mathcal{V}_x^\eta$ the support of $V \mid X_\eta = x$ and with $\mathcal{V}_x^{\eta C}$ its complement. Let $\mathcal{X}_\eta$ be the support of X_η, and $\mathcal{X} := \cup_{\eta \in E} \mathcal{X}_\eta$.

A payoff function together with a signal X_η and a prior for V constitute a **decision problem**, the problem being

$$\max_{a \in \mathcal{A}} \int_{\mathcal{V}} u(v, a) dG^\eta(v \mid x).$$

Define $a^\eta(x)$ as

$$a^\eta(x) \in \operatorname{argmax}_{a \in \mathcal{A}} \int_{\mathcal{V}} u(v, a) dG^\eta(v \mid x)$$

and let

$$u^\eta(v, x) := u(v, a^\eta(x)).$$

and

$$R(\eta) := \int_{\mathcal{V}} \int_{\mathcal{X}} u(v, a^\eta(x)) dF^\eta(x \mid v) dG(v)$$

be the expected revenue to the decision maker with signal X_η.

We will consider choosing one's signal X_η in a decision problem, at a cost $C(\eta)$. Given a payoff function, a prior for V, a family of signals $\{X_\eta\}_{\eta \in E}$ and a cost function $C(\eta)$, the following optimization problem

$$\max_{\eta \in E} R(\eta) - C(\eta)$$

will be referred to as the **information acquisition problem.**

Let

$$MR(\eta) := \frac{d}{d\theta} \int_{\mathcal{V}} \int_{\mathcal{X}} u(v, a^{\theta}(x)) dF^{\theta}(x \mid v) dG(v) \Big|_{\theta=\eta}.$$

denote the marginal revenue to the decision maker from increasing η, i.e. choosing a slightly higher signal.

Troughout this paper, smoothness assumptions are used in various places, and the flavour of results is often that of "marginal reasoning". The main reason for this is that Theorem 1 requires the use of an envelope condition on the optimal action. This point is expanded upon in the discussion following Theorem 1.

2.1 About payoff

Here we spell out some notation, and the restrictions we are going to place on the payoff function.

Let $\mathcal{I}$ be an interval of the real line. Let $\underline{0}$ denote the real function that is identically 0. We present now a definition due to Karamardian and Schaible [6].

Definition 1 *We say that a real function $H(v)$ is* **quasi-monotone on** $\mathcal{I}$ *if*

(QM)$\mathcal{I}$ $\qquad H(v) > 0 \Rightarrow H(v') \geq 0 \quad$ *for all $v' > v$, with v, v' in $\mathcal{I}$.*

We shall say that a function $H(\cdot)$ satisfies **(DQM)**$\mathcal{I}$ *whenever $-H(\cdot)$ satisfies* **(QM)**$\mathcal{I}$. *We say that a real function $H(\cdot)$ is* **strictly quasi-monotone on** $\mathcal{I}$ *or* **(SQM)**$\mathcal{I}$ *if H satisfies* **(QM)**$\mathcal{I}$ *and is almost everywhere different from 0 on $\mathcal{I}$. We define* **(SDQM)**$\mathcal{I}$ *analogously.*

Condition **(QM)**$\mathcal{I}$ is a condition of single-upward-crossing. It states that moving rightward on the v-axis inside $\mathcal{I}$ a function cannot become negative again once it has assumed positive values.

Definition 2 *Given two real functions H and L, we say that* $H \overset{(\mathbf{QM})}{\geq} L$ **on** $\mathcal{I}$ *if $H - L$ satisfies* **(QM)**$\mathcal{I}$.

We can loosely interpret the relation $H \overset{(\mathbf{QM})}{\geq} L$ as "H is more **(QM)** than L". A notion that is closely related to **(QM)** is that of weak single-crossing (see Milgrom and Shannon [9]):

Definition 3 *A real function $H(v,a)$ has the* **weak single crossing property in $(a;v)$ on $\mathcal{I}$** *if, for any fixed pair $a' > a$, we have*

$$\textbf{(WSCP)}\mathcal{I} \qquad D_{a,a'}(v) := H(v,a') - H(v,a) \textit{ satisfies } \textbf{(QM)}\mathcal{I}.$$

Remark 1 Whenever $H(v,a)$ is differentiable with respect to a and satisfies **(WSCP)**$\mathcal{I}$, then $\frac{\partial}{\partial a}H(v,a)$ satisfies **(QM)**$\mathcal{I}$. ◇

2.2 About signals

In the following we list several definitions pertaining to the signal structure of our model.

Definition 4 *We say that two random variables X and Y with joint distribution $f(x,y)$ are* **affiliated** *when*

$$\textbf{(AFF)} \qquad x' > x, y' > y \Rightarrow f(x',y')f(x,y) \geq f(x,y')f(x',y)$$

When X and Y are affiliated, we use the terms **affiliated signals** and **affiliated decision problem** (see Milgrom and Weber [10]).

Remark 2 It is easy to see that, if X and Y are affiliated, then $x' > x,\ y' > y \Rightarrow \frac{f(x|y')}{f(x'|y')} \leq \frac{f(x|y)}{f(x'|y)}$ whenever the denominators are different from 0. This is called **monotone likelihood ratio property**, because when y, y' are in the support of $f(y \mid x')$, increasing the value of y to y' monotonically reduces the ratio $\frac{f(x|\cdot)}{f(x'|\cdot)}$. Moreover, if X and Y are affiliated the ratio $\frac{F(x|y)}{f(x|y)}$ is nonincreasing in y (see Milgrom and Weber [10]). ◇

Definition 5 *We say that random variable X is* **non-weakly affiliated** *with random variable Y when X and Y are affiliated, and both of the following conditions hold:*

- *for all x' in the support of X and for all y in the support of $f(\cdot \mid x')$, there is an $x < x'$ such that y is a point of decrease for the ratio $\frac{f(x|\cdot)}{f(x'|\cdot)}$.*
- *for all $x < x' \in X$ such that $\mathcal{V}_x \cap \mathcal{V}_{x'} \neq \emptyset$, there is a y in the support of $f(\cdot \mid x')$ such that y is a point of decrease for the ratio $\frac{f(x|\cdot)}{f(x'|\cdot)}$.*

Remark 3 The concept of non-weak affiliation is weaker than that of *strict affiliation*, which requires the affiliation inequality to hold strictly inside the support of X and Y. For example, consider $f(x,y) = h(y)g(x \mid y)$ where $g(x \mid y)$ is a uniform on $[y-a, y+a]$ for some $a \in \Re_+$. This joint distribution implies that X is non-weakly affiliated with Y, but not strictly affiliated. $\diamond$

Remark 4 The first condition in the definition of non-weak affiliation implies that the ratio $\frac{F(x|\cdot)}{f(x|\cdot)}$ is strictly decreasing as a function of y. $\diamond$

Players can choose their signal from a family of random variables $\{X_\eta\}_{\eta \in E}$, each of which has a conditional c.d.f. $F^\eta(x \mid v)$. In what follows we shall make extensive use of an increasing function $T_{\eta,\theta,v}(\cdot)$ which has the following property: given $\eta < \theta$ and $X_\eta \sim F^\eta(x \mid v)$, then $T_{\eta,\theta,v}(X_\eta)$ has distribution $F^\theta(x \mid v)$. In other words,

Definition 6 *Given a family $\{X_\eta\}_{\eta \in E}$, the transformation $T_{\eta,\theta,v}(\cdot)$ is defined by*

$$T_{\eta,\theta,v} : \Re \to \Re \text{ is increasing} \quad \textit{and} \quad T_{\eta,\theta,v}(X_\eta \mid v) \sim X_\theta \mid v.$$

Definition 7 *Consider a family of signals $\{X_\eta\}_{\eta \in E}$. We say that this family of signals is* **A-ordered** *by η if, given any $\eta < \theta$ in E, we have $\frac{\partial}{\partial\theta} T_{\eta,\theta,v}(x)\Big|_{\theta=\eta}$ is nondecreasing in v, for all v and x in the support of $X_\eta \mid v$. If this quantity is strictly increasing in v, we say that the family is* **strictly A-ordered**.

The term "A-ordered" stands for "Accuracy-ordered": a more accurate signal corresponds to a higher parameter value. We will also use the term **A-ordered decision problem**, with the obvious meaning. Given an A-ordered family $\{X_\eta\}_{\eta \in E}$ we say that X_θ is **more accurate** than X_η if $\eta, \theta \in E$ and $\eta < \theta$.

3 The $T_{\eta,\theta,v}(\cdot)$ transformation

This paper introduces the transformation $T_{\eta,\theta,v}(\cdot)$. This is a novel concept. For given X_η, X_θ, v, this transformation is defined by

$$T_{\eta,\theta,v} : \Re \to \Re \text{ is increasing} \quad \text{and} \quad T_{\eta,\theta,v}(X_\eta \mid v) \sim X_\theta \mid v.$$

Our definition of "better information", or greater Accuracy, Definition 7, is based on this object. It is necessary to investigate the existence of such a transformation, before proceeding further. The next subsection explores this concept, while subsection 3.2 presents several examples.

3.1 A discussion of the T transformation

The next proposition asserts that the T transformation is well-defined. This transformation is a natural concept from statistics: it is the function needed to transform a conditional distribution (signal) into another (a possibly better signal). It is clear that the transformation must depend on v, since it has to work for all v's. This dependence makes it possible for T to "add correlation" to its argument.

Proposition 1 *Given a family of random variables $\{X_\theta\}$ with continuous density $f^\theta(x \mid v)$ with convex support, there is a unique real valued function $T_{\eta,\theta,v}(\cdot)$, strictly increasing on the support of $X_\eta \mid v$, such that*

$$T_{\eta,\theta,v}(X_\eta \mid v) \sim X_\theta \mid v.$$

Proof: For any triplet (η, θ, v) the function $T_{\eta,\theta,v}(\cdot)$ must solve the differential equation

$$f^\theta(T(x) \mid v)T'(x) = f^\eta(x \mid v)$$

(the change-of-variable identity), coupled with a suitable initial condition. A family of initial conditions that will work is

$$T_{\eta,\theta,v}(\text{ q-th quantile of } X_\eta \mid v) = \text{ q-th quantile of } X_\theta \mid v.$$

Then a standard theorem for differential equations gives existence and uniqueness of the T function. □

Definition 7 suggest that increasing the η index of a family of random variables represent an increase in the informative content of the signal whenever $\frac{\partial}{\partial\theta}T_{\eta,\theta,v}(x)\big|_{\theta=\eta}$ is nondecreasing in v. Let us comment on this notion of "useful information".

In this work the transformation $T_{\eta,\theta,v}(\cdot)$ will increase the informative content of a signal; thus, a more informative signal is defined starting from a less informative one. This is in contrast with the standard definition of a "more informative signal", given by Blackwell. There, a less informative signal is defined starting from a more informative one. We will see more about this in the following sections.

Like Blackwell's, our definition is independent of the prior on v, as is appropriate for a notion of accuracy of signals.

To get some intuition for the condition that $\frac{\partial}{\partial\theta}T_{\eta,\theta,v}(x)\big|_{\theta=\eta}$ be nondecreasing in v, we can profit from the following simple

Proposition 2 $\frac{\partial^2}{\partial\theta\partial v}T_{\eta,\theta,v}(x)\Big|_{\theta=\eta} > 0 \Leftrightarrow \frac{\partial}{\partial v}T_{\eta,\theta,v}(x) > 0$ *for* η, θ *sufficiently close.*

Proof: We can write

$$0 < \frac{\partial^2}{\partial\theta\partial v}T_{\eta,\theta,v}(x)\Big|_{\eta=\theta} = \lim_{\theta\downarrow\eta}\frac{\frac{\partial}{\partial v}T_{\eta,\theta,v}(x) - \frac{\partial}{\partial v}T_{\eta,\eta,v}(x)}{\theta-\eta}.$$

But $T_{\eta,\eta,v}(x) = x$, whereby the second term in the numerator of the fraction is 0. Thus, our condition is equivalent to

$$\frac{\partial}{\partial v}T_{\eta,\theta,v}(x) > 0$$

for η, θ sufficiently close. $\square$

Thus, our condition that the family of signals be A-ordered means that $\frac{\partial}{\partial v}T_{\eta,\theta,v}(x) > 0$, i.e. the transformation $T_{\eta,\theta,v}(\cdot)$ varies together with v. The new signal obtained applying this transformation, is a bit higher than the old signal when v is higher, and a bit lower when v is low. We see how our notion of a "more accurate signal" can be interpreted as one of a signal that is more correlated with the true value of the object, v. The $T(\cdot)$ transformation adds this additional correlation to the original signal.

Let us see the definition at work, in an example.

Example 1 Let $\eta \in (0, \infty)$, and V be distributed according to any c.d.f.. Let X_η be distributed according to a uniform on $[v - 1/\eta, v + 1/\eta]$.

$$f^\eta(x \mid v) = \frac{\eta}{2} \quad \text{on } [v - 1/\eta, v + 1/\eta].$$

Using the definition of affiliation we can check that X, V are affiliated. Furthermore, we can compute that

$$T_{\eta,\theta,v}(x) = \frac{\eta}{\theta}(x - v) + v.$$

Indeed, $T_{\eta,\theta,v}(x)$ solves the change-of-variable equation

$$\frac{\theta}{2} = \frac{\eta}{2}(T^{-1}(x))',$$

together with the relevant boundary conditions. $T_{\eta,\theta,v}(\cdot)$ is an increasing function and

$$\frac{\partial^2}{\partial\theta\partial v}T_{\eta,\theta,v}(x) = \frac{\eta}{\theta^2} > 0.$$

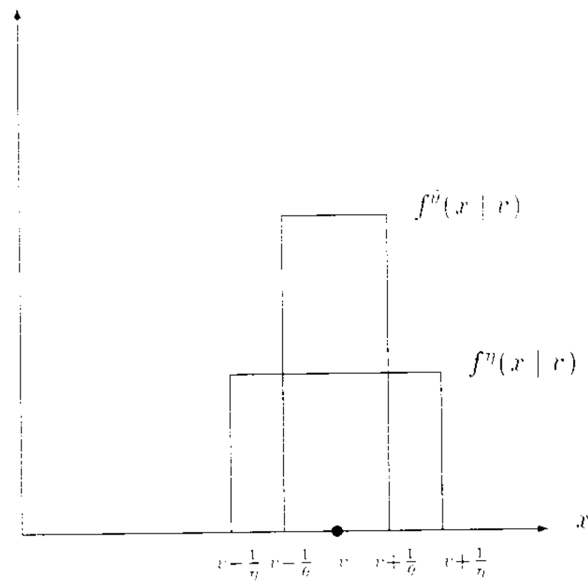

Figure 1: X_θ is more accurate than X_η

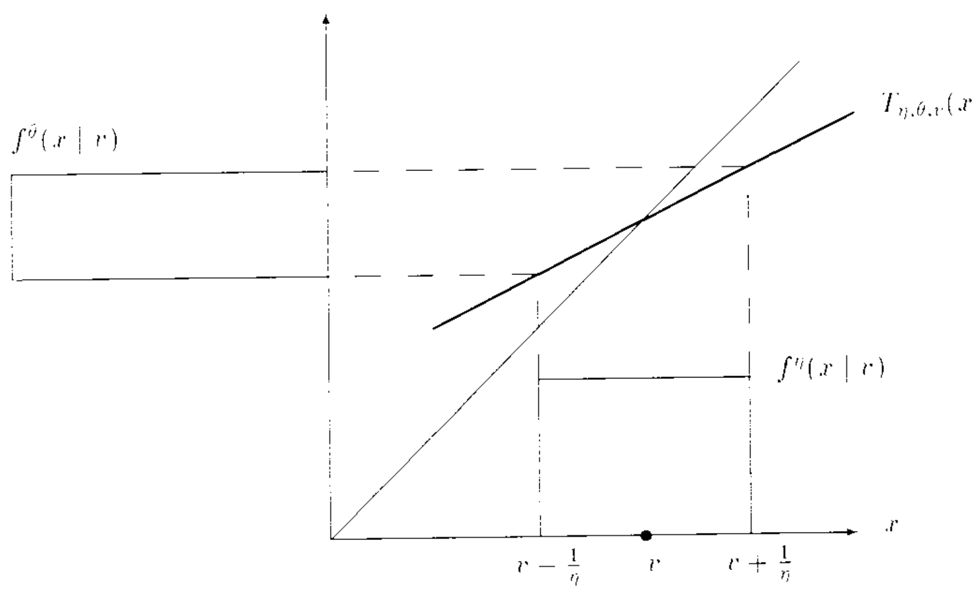

Figure 2: The $T(\cdot)$ transformation in action

Thus, the family $\{X_\eta\}_{\eta \in (0,\infty)}$ is strictly A-ordered by η. If the support of V is bounded, then $\frac{\partial}{\partial \theta} T_{\eta,\theta,v}(x)$ is bounded (as a function of v) uniformly in x. ◇

It is perhaps helpful to understand the geometric meaning of the condition that $\frac{\partial}{\partial v} T_{\eta,\theta,v} > 0$, with the aid of Figure 2. The first thing to notice is that $T_{\eta,\theta,v}(\cdot)$ has slope less than 1, and crosses the diagonal at v. This means that T "contracts" mass around v, the parameter we want to infer.

Second, observe that in this example $T_{\eta,\theta,v}(\cdot)$ is a straight line, depending on v only because it passes through (v, v). Thus, increasing v to v' produces a parallel shift of the T function, upwards to the line passing through a higher point (v', v'). Now, fix a value for x: such a shift produces an increase in $T(x)$ only if the slope of $T(\cdot)$ is less than 1. If the slope were higher than 1, increasing v would actually *reduce* $T(x)$. Hence, we see how in this example the property $\frac{\partial}{\partial v} T_{\eta,\theta,v} > 0$ ($T(x)$ increases with v) can be translated into a "contraction property" (the slope of $T(\cdot)$ is less than one).

3.2 More examples of A-ordered families

This subsection illustrates several examples of affiliated A-ordered families.

In the next two examples, η appears as an exponent of a c.d.f.: thus increasing η causes a first-order stochastic shift in the distribution of X conditional on v, without changing its support. Since v is also the upper bound of the support of $X \mid v$, this is the same as

$$\forall x \in \Re, \eta < \theta \text{ implies } P(X_\theta \in [x, v]) \geq P(X_\eta \in [x, v]).$$

In other words, the distribution of X_θ is more concentrated around the true value v than that of X_η. Thus it seems reasonable that an increase in η should increase the precision of a signal.

Examples 3 and 4 are of particular interest because there our definition of "more useful signal" coincides with Blackwell's notion of Sufficiency. For the first example this has been shown by Matthews [8], while the second example is the textbook example of Blackwell's theory (see for instance DeGroot [4]). We will indeed show later on that Sufficiency, in our "affiliated" framework, always implies Accuracy. Besides, example 4 treats the normal case, one that is of obvious interest in information theory.

Example 2 Let $\eta \in (0, \infty)$, and V be distributed according to any c.d.f.. Set

$$F^\eta(x \mid v) := \frac{x^{\eta(v+1)}}{v^{\eta(v+1)}} \text{ on } [0, v] \tag{1}$$

The corresponding density function is

$$f^{\eta}(x \mid v) = \eta(v+1)\frac{x^{\eta(v+1)-1}}{v^{\eta(v+1)}} \text{ on } [0, v]$$

It is readily verified that

$$\frac{\partial^2}{\partial x \partial v} \log f^{\eta}(x \mid v) = \frac{\eta}{x} > 0$$

whereby signals are affiliated for any η.

Let θ be some number greater than η; the corresponding c.d.f. is

$$F^{\theta}(x \mid v) := \frac{x^{\theta(v+1)}}{v^{\theta(v+1)}}. \tag{2}$$

Let

$$T_{\eta,\theta,v}(x) := x^{\eta/\theta} v^{1-\frac{\eta}{\theta}}.$$

Using the change-of-variable formula, one may verify that, when X is distributed according to (1), then $T_{\eta,\theta,v}(X)$ is distributed according to (2). Indeed, we may check that $T_{\eta,\theta,v}(\cdot)$ verifies the change-of-variable equality

$$\theta(v+1)\frac{y^{\theta(v+1)-1}}{v^{\theta(v+1)}} = \eta(v+1)\frac{[T^{-1}(y)]^{\eta(v+1)-1}}{v^{\eta(v+1)}}[T^{-1}(y)]'.$$

We observe that $T_{\eta,\theta,v}(x)$ is increasing in x, and can compute

$$\frac{\partial^2}{\partial \theta \partial v} T_{\eta,\theta,v}(x)|_{\theta=\eta} = \frac{x}{\eta v} > 0.$$

Hence, the family $\{X_\eta\}$ is strictly A-ordered. When the support of V is a finite strictly positive interval, $\frac{\partial}{\partial \theta} T_{\eta,\theta,\cdot}(x)$ is bounded (as a function of v) uniformly on x.

◇

Example 3 Let $\eta \in [0, \infty)$, and V be distributed according to any c.d.f.. Set

$$F^{\eta}(x \mid v) := \left[\frac{x}{v}\right]^{\eta} \text{ on } [0, v] \tag{3}$$

The corresponding density function is

$$f^{\eta}(x \mid v) = \eta \frac{x^{\eta-1}}{v^{\eta}} \text{ on } [0, v]$$

It is readily verified that

$$\frac{\partial^2}{\partial x \partial v} \log f^{\eta}(x \mid v) = 0$$

whereby signals are affiliated for all η.

Let θ be some number greater than η: the corresponding c.d.f. is

$$F^{\theta}(x \mid v) := \left[\frac{x}{v}\right]^{\theta}. \tag{4}$$

Let

$$T_{\eta,\theta,v}(x) := x^{\eta/\theta} v^{1-\frac{\eta}{\theta}}.$$

Using the change-of-variable formula, one may verify that, when X is distributed according to (3), then $T_{\eta,\theta,v}(X)$ is distributed according to (4). Indeed, we may check that $T_{\eta,\theta,v}(\cdot)$ verifies the change-of-variable equality

$$\theta \frac{y^{\theta-1}}{v^{\theta}} = \eta \frac{[T^{-1}(y)]^{\eta-1}}{v^{\eta}} [T^{-1}(y)]'.$$

Since this is the same $T(\cdot)$ as in the previous example, we can conclude that the family $\{X_\eta\}$ is strictly Λ-ordered. When the support of V is a finite strictly positive interval, $\frac{\partial}{\partial \theta} T_{\eta,\theta,\cdot}(x)$ is bounded (as a function of v) uniformly on x.

◇

Example 4

Let $\eta \in [0, \infty)$, and V be distributed according to any c.d.f.. Suppose X_η has a normal conditional distribution with mean v and variance $1/\eta^2$. It is readily verified that

$$\frac{\partial^2}{\partial x \partial v} \log f^{\eta}(x \mid v) = \eta^2 > 0,$$

whence all the random variables are affiliated. Moreover, basic notions regarding the normal distribution tell us that when

$$T_{\eta,\theta,v}(x) = \frac{\eta}{\theta}(x - v) + v,$$

we have that

$$T_{\eta,\theta,v}(X_\eta) \sim X_\theta.$$

Furthermore, $T_{\eta,\theta,v}(x)$ is increasing in x and we can compute

$$\frac{\partial^2}{\partial \theta \partial v} T_{\eta,\theta,v}(x) = \frac{\eta}{\theta^2} > 0.$$

Hence, the family $\{X_\eta\}$ is strictly Λ-ordered.

◇

4 The concept of Accuracy

In the first subsection we present a couple of Lemmas we shall need further on. The second subsection contains the main theorems involving Accuracy. The third discusses the import of an assumption of our theory, that of quasi-monotonicity. Subsection 4.4 discusses the relationship between Accuracy and Sufficiency.

4.1 Two technical lemmas

This subsection presents two lemmas, which will be of use in the following. First is a change-of-variable lemma:

Lemma 1 *Let X be a random variable, $X \sim F(\cdot)$, and let $h(\cdot)$ be any real function. If $T(\cdot)$ is an increasing real function such that $T(X) \sim G(\cdot)$, then*

$$\int_{-\infty}^{+\infty} h(x')dG(x') = \int_{-\infty}^{+\infty} h(T(x))dF(x).$$

Proof: This is just a restatement of the change-of-variable principle. □

The next lemma, like many theorems later on, is divided into a weak (W) and a strong (S) part.

Lemma 2 *Let (c, d) be an interval of the real line, $J(\cdot)$ a nondecreasing real function, $H(\cdot)$ a real function satisfying* $\mathbf{(QM)}(\boldsymbol{c}, \boldsymbol{d})$. *Assume that for some measure μ on $\Re$ we have*

$$\int_c^d H(r)d\mu(r) = 0. \tag{5}$$

Then
W)

$$\int_c^d H(r)J(r)d\mu(r) \geq 0.$$

S1) *if in addition the support of μ contains (c, d), $H(\cdot) \neq \underline{0}$ on (c, d) and $J(\cdot)$ is strictly increasing over (c, d), we have*

$$\int_c^d H(r)J(r)d\mu(r) > 0.$$

S2) *if in addition the support of μ contains (c, d), $H(\cdot)$ satisfies* **(SQM)**$(\boldsymbol{c}, \boldsymbol{d})$, *and $J(\cdot)$ has a point of increase on (c, d), we have*

$$\int_c^d H(r)J(r)d\mu(r) > 0.$$

Proof:

Part W): Because of (5), toghether with **(QM)**, we know that there must be a $r_0 \in [c, d]$, before which H is nonpositive, and after which it is nonnegative. Let $\tilde{J}(r) := J(r) - J(r_0)$; because $J(\cdot)$ is nondecreasing, $\tilde{J}(r)$ is nonpositive before r_0, and nonnegative after. Then we can write

$$\int_c^d H(r)J(r)d\mu(r) = \int_c^d H(r)\tilde{J}(r)d\mu(r) =$$

$$= \int_c^{r_0} H(r)\tilde{J}(r)d\mu(r) + \int_{r_0}^d H(r)\tilde{J}(r)d\mu(r) \geq 0$$

where the first equality uses (5) and the inequality follows from H and $\tilde{J}$ having the same sign on (c, r_0) and (r_0, d).

Parts S1), S2) In this cases, strict inequality will hold for at least one of the two addends after the second equality above. □

Remark 5 Suppose that, in the above Lemma, $c = -\infty$ and $d = +\infty$. Then, (5) reads $E_\mu(H) = 0$ and thus the conclusions of Lemma 2 can be restated as $Cov_\mu(H, J) \geq 0$. ◇

Remark 6 If $J(r)$ is decreasing, then the inequalities of part W) and S) are reversed. To see this, it is enough to observe that in this case, $-J(r)$ is increasing, and then apply Lemma 2. ◇

Remark 7 It is immediate to show that Lemma 2 holds whenever $H(\cdot)$ satisfies **(DQM)** *and* $J(\cdot)$ is decreasing. ◇

4.2 Accuracy

This is the section containing the results motivating this work. Theorem 1 can be read as follows: given a family of signals and a problem where $\frac{\partial}{\partial x}u^a$ satisfies **(QM)**, choosing a signal with a slightly higher index is beneficial if the family is Λ-ordered.

When the family of signals is non-weakly affiliated, Theorem 2 gives us the converse result: a failure of A-order implies that there is a problem where $\frac{\partial}{\partial x}u^{\eta}$ satisfies **(QM)** but choosing a signal with a higher index is detrimental.

Theorem 1 and 2 together show that, under non-weak affiliation of signals, the A-order condition is necessary and sufficient for an increase in payoff in the class of problems where $\frac{\partial}{\partial x}u^{\eta}$ satisfies **(QM)**. Therefore, A-order is a tight notion of "more informative signals" on a subclass of decision problems, those where a restriction of affiliation is placed on the signals, and one of monotonicity is placed on the payoff function.

The intuition underlying Theorem 1 is best understood when $\frac{\partial}{\partial x}u^{\eta}(v,x)$ satisfies **(QM)**. Consider the first-order conditions for a decision problem with Accuracy η, in terms of "type reports": the average of $\frac{\partial}{\partial x}u^{\eta}(v,x)$ with respect to v must equal 0. Condition **(QM)** implies that for low values of v this quantity is negative, while it is positive for high values of v. Thereby, in terms of first-order conditions, it would be desirable to report a slightly lower x when v is low, and a slightly higher x when v is high. But this is exactly what $T(\cdot)$ allows you to do when the family of signals is A-ordered. Indeed, introducing the T transformation sways the type report in the direction the decision maker would have, had he known the true value of v.

Theorem 1 *Consider an A-ordered family of signals and assume that, for all x, $\frac{\partial}{\partial x}u^{\eta}(v,x)$ satisfies* **(QM)**$\mathcal{V}_x^{\eta}$. *Then*

W) $MR(\eta) \geq 0$

S) *if in addition $\frac{\partial}{\partial x}u^{\eta}(v,x) \neq \underline{0}$ on $\mathcal{V}_x^{\eta}$ and the signal family is strictly A-ordered, we have $MR(\eta) > 0$.*

Proof:

Part W): We wish to show that $MR(\eta) \geq 0$, that is

$$\frac{d}{d\theta}\int_{\mathcal{V}}\int_{X} u(v,a^{\theta}(x))dF^{\theta}(x \mid v)dG(v)\bigg|_{\theta=\eta} \geq 0.$$

To apply Lemma 1, let us identify $F^{\theta}(\cdot \mid v)$ as $G(\cdot)$, $F^{\eta}(\cdot \mid v)$ as $F(\cdot)$ and $T_{\eta,\theta,v}(\cdot)$ as $T(\cdot)$. Then we can rewrite the above LHS as

$$\frac{d}{d\theta}\int_{\mathcal{V}}\int_{X} u(v,a^{\theta}(T_{\eta,\theta,v}(x)))dF^{\eta}(x \mid v)dG(v)\bigg|_{\theta=\eta} =$$

$$\frac{d}{d\theta}\int_{X}\int_{\mathcal{V}} u(v,a^{\theta}(T_{\eta,\theta,v}(x)))dG^{\eta}(v \mid x)dF^{\eta}(x)\bigg|_{\theta=\eta} =$$

$$\int_{\mathcal{X}} \frac{d}{d\theta} \int_{\mathcal{V}} u(v, a^{\theta}(T_{\eta,\theta,x}(x))) dG^{\eta}(v \mid x) dF^{\eta}(x) \Big|_{\theta=\eta}. \tag{6}$$

Taking into account an envelope result that allows us to omit differentiating the optimal strategy with respect to θ, the inner integral can be computed as

$$\int_{\mathcal{V}} \left[\frac{\partial}{\partial a} u(v, a^{\eta}(x))\right] a^{\eta\prime}(x) \frac{\partial}{\partial\theta} T_{\eta,\theta,x}(x))\Big|_{\theta=\eta} dG^{\eta}(v \mid x) =$$

$$\int_{\mathcal{V}} \frac{\partial}{\partial x} u^{\eta}(v, x) \frac{\partial}{\partial\theta} T_{\eta,\theta,x}(x))\Big|_{\theta=\eta} dG^{\eta}(v \mid x). \tag{7}$$

But, first-order conditions tell us that

$$\int_{\mathcal{V}} \frac{\partial}{\partial x} u^{\eta}(v, x) dG^{\eta}(v \mid x) = 0,$$

and since $\frac{\partial}{\partial x} u^{\eta}(v, x)$ satisfies **(QM)** by assumption, we can apply Lemma 2 to conclude that (7), and hence (6), are nonnegative.

Part S): follows from Lemma 2 part S1). □

A-order, our notion of "better information", is defined on a continuous family of signals, rather than through the comparison of any two signals. This is because Theorem 1 is a marginal result. Two things in this theorem require changes in Accuracy to be infinitesimal:

1. an envelope condition (allowing to disregard the change of the optimal action with small changes of Accuracy), and
2. a first-order (optimality) condition, needed to apply Lemma 2.

A discrete analogue of Theorem 1 is not straightforward, since it would presumably entail assumptions on how the optimal strategy changes when the signal changes. Furthermore, Lemma 2, the crucial tool in the proof of Theorem 1, fails in the absence of the first-order (zero-mean) condition which to be satisfied requires infinitesimal changes in Accuracy. Naturally, in an A-ordered family it is possible to rank the payoff of any two signals that are not infinitesimally close, since then Theorem 1 would hold for each signal in the family: moving from any signal towards a higher one in the family would monotonically increase payoff. It must be noticed however that this result relies on the existence of a family spanning the two signals we wish to compare.

The role of the first-order (optimality) condition in the proof of Theorem 1 suggests that this result is *not* in the spirit of the literature on stochastic dominance: there, optimality is not required, and the conclusions are different (and weaker).

The next theorem provides the converse to Theorem 1, under the assumption of non-weak affiliation of signals: if a family of signals is not A-ordered there exists a decision problem where increasing η is detrimental.

Theorem 2 *Consider a family $\{X_\eta\}_{\eta\in E}$ of signals non-weakly affiliated with V. If this family is not A-ordered, there exist payoff function $u(v,a)$, a prior $g(v)$ and an $\eta \in E$ such that $\frac{\partial}{\partial x}u^\eta(v,x)$ satisfies* $(\mathbf{QM})\mathcal{V}_x^\eta$ *for all x in the support of X_η, but $MR(\eta) < 0$.*

Proof: We will show that if $\{X_\eta\}$ is not A-ordered, i.e. for some (η, v^*, x^*) we have that $\frac{\partial}{\partial\theta}T_{\eta,\theta,x}(x^*)\Big|_{\theta=\eta}$ is strictly decreasing in v at v^*, then we can construct a decision problem in which slightly increasing η reduces the payoff.

First, we may assume, using continuity, that $\frac{\partial}{\partial\theta}T_{\eta,\theta,x}(x)$ is decreasing in v in some square neighbourhood N around (v^*, x^*). Furthermore, we choose a prior $g(v)$ that is all concentrated on N.

This done, consider the following decision problem (and revelation game):

$$u(v,a) := \int_{-\infty}^{a} [t_2(s)f(s \mid v) + t_1(s)F(s \mid v)]\, ds.$$

Here a is the action, and may be seen as the revelation of one's type. $t_1(\cdot)$ is an arbitrary negative function, and $t_2(\cdot)$ is some positive function. When $t_2(\cdot)$ is a bounded function, $u(v,a)$ is bounded from above. It is readily verified that $\frac{\partial}{\partial a}u(v,a)$ satisfies $(\mathbf{QM})$.

We first of all want to construct a payoff for which truthful revelation is the unique optimal action. This property has to hold for fixed η, hence to lighten notation we omit reference to η.

Part i: truthful revelation is optimal

The objective function for a decision maker observing signal $X = z$ (a "type z") is

$$\int_{\mathcal{V}} \int_{-\infty}^{a} [t_2(s)f(s \mid v) + t_1(s)F(s \mid v)]\, ds\, dG(v \mid z). \tag{8}$$

The first-order conditions are

$$\int_{\mathcal{V}} [t_2(a)f(a \mid v) + t_1(a)F(a \mid v)]\, dG(v \mid z) = 0. \tag{9}$$

Let $t_2(\cdot)$ be defined by

$$\int_{\mathcal{V}} \left[t_2(z) + t_1(z)\frac{F(z \mid v)}{f(z \mid v)} \right] f(z \mid v) dG(v \mid z) = 0 \text{ for each } z. \tag{10}$$

For every $t_1(\cdot)$, eq. (10) defines a function $t_2(\cdot)$ having the property that truthtelling is a stationary point for the decision problem. To show that it is indeed the unique maximum, suppose that a type z declares $\zeta > z$. First, it is suboptimal to declare a $\zeta > z$ such that $\mathcal{V}_\zeta \cap \mathcal{V}_z \neq \emptyset$, since reducing the action would improve the payoff at rate $t_1(a)$.

Then the first-order conditions are

$$\int_{\mathcal{V}} I_{\mathcal{V}_\zeta}(v) I_{\mathcal{V}_z}(v) \left[t_2(\zeta) + t_1(\zeta)\frac{F(\zeta \mid v)}{f(\zeta \mid v)} \right] f(\zeta \mid v) dG(v \mid z) +$$

$$+ \int_{\mathcal{V}} I_{\mathcal{V}_\zeta^C}(v) I_{\mathcal{V}_z}(v) \left[t_2(\zeta) f(\zeta \mid v) + t_1(\zeta) F(\zeta \mid v) \right] dG(v \mid z) =$$

$$\int_{\mathcal{V}} I_{\mathcal{V}_\zeta}(v) I_{\mathcal{V}_z}(v) \left[t_2(\zeta) + t_1(\zeta)\frac{F(\zeta \mid v)}{f(\zeta \mid v)} \right] f(\zeta \mid v) \frac{g(v) f(z \mid v)}{f(z)} dv +$$

$$+ \int_{\mathcal{V}} I_{\mathcal{V}_\zeta^C}(v) I_{\mathcal{V}_z}(v) \left[t_2(\zeta) f(\zeta \mid v) + t_1(\zeta) F(\zeta \mid v) \right] dG(v \mid z) =$$

$$\frac{f(\zeta)}{f(z)} \int_{\mathcal{V}} I_{\mathcal{V}_\zeta}(v) I_{\mathcal{V}_z}(v) \left[t_2(\zeta) + t_1(\zeta)\frac{F(\zeta \mid v)}{f(\zeta \mid v)} \right] f(\zeta \mid v) \left[\frac{f(z \mid v)}{f(\zeta \mid v)} \right] dG(v \mid \zeta) +$$

$$+ \int_{\mathcal{V}} I_{\mathcal{V}_\zeta^C}(v) I_{\mathcal{V}_z}(v) \left[t_2(\zeta) f(\zeta \mid v) + t_1(\zeta) F(\zeta \mid v) \right] dG(v \mid z) =$$

$$\frac{f(\zeta)}{f(z)} \int_{\mathcal{V}} I_{\mathcal{V}_\zeta}(v) \left[t_2(\zeta) + t_1(\zeta)\frac{F(\zeta \mid v)}{f(\zeta \mid v)} \right] f(\zeta \mid v) \left[\frac{f(z \mid v)}{f(\zeta \mid v)} \right] dG(v \mid \zeta) -$$

$$- \frac{f(\zeta)}{f(z)} \int_{\mathcal{V}} I_{\mathcal{V}_\zeta}(v) I_{\mathcal{V}_z^C}(v) \left[t_2(\zeta) + t_1(\zeta)\frac{F(\zeta \mid v)}{f(\zeta \mid v)} \right] f(\zeta \mid v) \left[\frac{f(z \mid v)}{f(\zeta \mid v)} \right] dG(v \mid \zeta) +$$

$$+ \int_{\mathcal{V}} I_{\mathcal{V}_\zeta^C}(v) I_{\mathcal{V}_z}(v) \left[t_2(\zeta) f(\zeta \mid v) + t_1(\zeta) F(\zeta \mid v) \right] dG(v \mid z) < 0$$

To justify the inequality, let us discuss the sign of each term in the above expression.

- The first term is nonzero, since we restrict attention to ζ's such that $\mathcal{V}_\zeta \cap \mathcal{V}_z \neq \emptyset$. Let us show that it is negative: we want to make use of Remark 6 applied to Lemma 2 S2). We can then identify $J(v)$ with

$\frac{f(z|r)}{f(\zeta|r)}$, since it is nonincreasing in r and has a point of decrease because of non-weak affiliation. The term

$$\left[t_2(\zeta) - t_1(\zeta)\frac{F(\zeta \mid r)}{f(\zeta \mid r)}\right] f(\zeta \mid r)$$

satisfies $(\mathbf{SQM})\mathcal{V}_\zeta$ (non-weak affiliation again), and because of eq. (10) may be identified with $H(r)$. We can therefore use Remark 6 applied to Lemma 2 S2) to conclude that the first term is negative.

- The second term vanishes because of the combined action of $I_{\mathcal{V}^c_\zeta}(r)$ and $f(z \mid r)$.
- The third term is nonpositive, because the indicators indicate a region of r's below the *inf* of $\mathcal{V}_\zeta$ (using affiliation). Thereby, $f(\zeta \mid r) = 0$ and $F(\zeta \mid r) = 1$.

The inequality shows that it is profitable to decrease one's bid. The converse holds for declarations below one's true type, and hence we have proved that truthtelling is the only optimal action.

We now construct the decision problem that will yield the result.

Part ii: increasing η lowers the payoff

For any given negative function $t_1^\eta(\cdot)$, pick the associate $t_2^\eta(\cdot)$ function that renders truthtelling optimal with signal X_η (Part i above ensures such a $t_2^\eta(\cdot)$ exists).

Let us define the payoff of our decision problem when the agent is observing X_θ as

$$\int_{\mathcal{X}}\int_{\mathcal{V}}\int_{-\infty}^{a^\theta(x)} [t_2^\eta(s)f^\eta(s \mid r) + t_1^\eta(s)F^\eta(s \mid r)]\, ds dG^\theta(r \mid x) dF^\theta(x) \qquad (11)$$

where $a^\theta(\cdot)$ is the optimal decision function with signal X_θ in problem (11). We can rewrite this expression, after the algebra of Theorem 1, as

$$\int_{\mathcal{X}}\int_{\mathcal{V}}\int_{-\infty}^{a^\eta(T_{\eta,\theta,x}(x))} [t_2^\eta(s)f^\eta(s \mid r) + t_1^\eta(s)F^\eta(s \mid r)]\, ds dG^\eta(r \mid x) dF^\eta(x)$$

In view of Part i above, $a^\eta(x) = x$ and thus the derivative with respect to θ evaluated at $\theta = \eta$ is

$$\int_{\mathcal{X}}\int_{\mathcal{V}} \left[t_2^\eta(x) + t_1^\eta(x)\frac{F^\eta(x \mid r)}{f^\eta(x \mid r)}\right] f^\eta(x \mid r) \left.\frac{\partial T_{\eta,\theta,x}(x)}{\partial\theta}\right|_{\theta=\eta} dG^\eta(r \mid x) dF^\eta(x). \qquad (12)$$

We will show that the inner integral, namely

$$\int_{\mathcal{V}} \left[t_2^{\eta}(x) + t_1^{\eta}(x) \frac{F^{\eta}(x \mid r)}{f^{\eta}(x \mid r)} \right] f^{\eta}(x \mid r) \left. \frac{\partial T_{\eta,\theta,x}(x)}{\partial \theta} \right|_{\theta=\eta} dG^{\eta}(r \mid x) \tag{13}$$

is negative for $x \in X$, and may be rendered arbitrarily close to 0 for $x \notin X$. This will allow us to conclude that (12) is negative.

To see that expression (13) is negative for $x \in X$, it is sufficient to recall that in this case $\left.\frac{\partial T_{\eta,\theta,x}(x)}{\partial \theta}\right|_{\theta=\eta}$ is strictly decreasing in r (recall $r \in X$ by construction), while the term in brackets is **(SQM)** becuse of non-weak affiliation. Then Remark 6 applied to Lemma 2 S2) will yield the result.

To make expression (13) arbitrarily small for $x \notin X$, it is sufficient to choose $t_1(x)$ very close to 0 for $x \notin X$. □

4.3 The (QM) assumption

Theorems 1 and 2, the theorems that motivate this paper, rely on $\frac{\partial}{\partial x}u^{\eta}(r,x)$ satisfying **(QM)**. This condition does not apply to primitives of the problem, but to optimal quantities (given a fixed Accuracy). It is then appropriate to provide a sufficient condition on primitives for $\frac{\partial}{\partial x}u^{\eta}(r,x)$ to satisfy **(QM)**.

Proposition 3 *Consider an affiliated decision problem where $u(r,a)$ has the single-crossing property in $(a;r)$ on $\mathcal{V}$. Then $\frac{\partial}{\partial x}u^{\eta}(r,x)$ satisfies* **(QM)**$\mathcal{V}$.

Proof: By definition,

$$\frac{\partial}{\partial x}u^{\eta}(r,x) = \left[\frac{\partial}{\partial a}u(r,a^{\eta}(x))\right] a^{\eta\prime}(x).$$

The first term in the product satisfies **(QM)**$\mathcal{V}$ since $u(r,a)$ satisfies **(WSCP)**$\mathcal{V}$ (see Remark 1). It is therefore sufficient to show that $a^{\eta\prime}(x)$ is nonnegative. But this follows from Theorem 5.1 of Athey [1] □

In addition to the above proposition, we are able to provide a characterization of the quantity $\frac{\partial}{\partial x}u^{\eta}(r,x)$ in a general affiliated decision problem (that is, without imposing **(WSCP)**). This is of special interest given the results of the next subsection. Roughly, these say that in an affiliated decision problem, "increasing information" in Blackwell's sense can be represented as increasing the Accuracy parameter in a family of signals. But the proof of Theorem 1 implies that increasing Accuracy, while "good" in a problem satisfying **(QM)** is "bad" in a problem where $\frac{\partial}{\partial x}u^{\eta}(r,x)$ satisfies **(SDQM)**. This suggests that in an affiliated problem $\frac{\partial}{\partial x}u^\eta(v,x)$ cannot satisfy **(SDQM)**, which is confirmed by the next proposition. The proof is carried out using second-order conditions.

Proposition 4 *In a non-weakly affiliated decision problem, given any x, $\frac{\partial}{\partial x}u^\eta(v,x)$ cannot satisfy* **(SDQM)**$\mathcal{V}$.

Proof: By contradiction, suppose $\frac{\partial}{\partial x}u^\eta(v,x)$ satisfies **(SDQM)**. Consider the first-order conditions for any type $z < x$ who plays like an x type:

$$\int_{\mathcal{V}} \left[\frac{\partial}{\partial x}u^\eta(v,x)\right] dG^\eta(v \mid z) =$$

$$\int_{\mathcal{V}} I_{\mathcal{V}_z}(v)\left(I_{\mathcal{V}_x}(v) + I_{\mathcal{V}_x^C}(v)\right)\left[\frac{\partial}{\partial x}u^\eta(v,x)\right] g(v)\frac{f^\eta(z \mid v)}{f^\eta(z)}dv =$$

$$\int_{\mathcal{V}} I_{\mathcal{V}_z}(v) I_{\mathcal{V}_x}(v)\left[\frac{\partial}{\partial x}u^\eta(v,x)\right]\frac{f^\eta(z \mid v)}{f^\eta(z)}\frac{f^\eta(x)}{f^\eta(x \mid v)}g(v)\frac{f^\eta(x \mid v)}{f^\eta(x)}dv +$$

$$+\int_{\mathcal{V}} I_{\mathcal{V}_z}(v) I_{\mathcal{V}_x^C}(v)\left[\frac{\partial}{\partial x}u^\eta(v,x)\right] dG^\eta(v \mid z) =$$

$$\frac{f^\eta(x)}{f^\eta(z)}\int_{\mathcal{V}} I_{\mathcal{V}_x}(v)\left[\frac{\partial}{\partial x}u^\eta(v,x)\right]\frac{f^\eta(z \mid v)}{f^\eta(x \mid v)}dG^\eta(v \mid x) -$$

$$-\frac{f^\eta(x)}{f^\eta(z)}\int_{\mathcal{V}} I_{\mathcal{V}_x}(v) I_{\mathcal{V}_z^C}(v)\left[\frac{\partial}{\partial x}u^\eta(v,x)\right]\frac{f^\eta(z \mid v)}{f^\eta(x \mid v)}dG^\eta(v \mid x) +$$

$$\int_{\mathcal{V}} I_{\mathcal{V}_z}(v) I_{\mathcal{V}_x^C}(v)\left[\frac{\partial}{\partial x}u^\eta(v,x)\right] dG^\eta(v \mid z) \tag{14}$$

Recalling now that first-order conditions for the x-type are

$$\int_{\mathcal{V}} \left[\frac{\partial}{\partial x}u^\eta(v,x)\right] dG(v \mid x) = 0,$$

that $\frac{f^\eta(z|v)}{f^\eta(x|v)}$ has a point of decrease at some v by non-weak affiliation, and that $\frac{\partial}{\partial x}u^\eta(v,x)$ satisfies **(SDQM)** by assumption, we can conclude that the first addend of (14) is positive by virtue of Remark 7 applied to Lemma 2 partS2). The second term is 0 because of the combined role of $I_{\mathcal{V}_z^C}$ and $f(z \mid v)$. The third term is nonnegative also, because of **(SDQM)**, since affiliation implies that the set of v's such that the indicators are nonzero lies in the area where $\frac{\partial}{\partial x}u^\eta(v,x)$ is nonnegative. Thus, the whole expression (14) is positive. But this contradicts the fact that playing $a^\eta(z)$ is optimal for type z. □

4.4 The relationship between Accuracy and Blackwell's Sufficiency

It must be noticed that our notion of Accuracy is indigenous to a particular class of decision problems, those where $\frac{\partial}{\partial x}u^{0}(v,x)$ satisfies **(QM)**. A notion that, instead, characterizes "better information" in *any* decision problem is Blackwell's definition of Sufficiency. Let us report it here, as presented in DeGroot [4], and explore its connection with Accuracy.

Let X and Y be random variables, or signals, with support S_X and S_Y respectively. It is said that the signal Y is **sufficient** for the signal X if there exists a nonnegative function h on the product space $S_X \times S_Y$ for which the following three relations are satisfied:

$$f_X(x \mid v) = \int_{S_Y} h(x,y) f_Y(y \mid v)dy \text{ for } v \in V \text{ and } x \in S_X$$

$$\int_{S_X} h(x,y)dx = 1 \text{ for } y \in S_Y \tag{15}$$

and

$$0 < \int_{S_Y} h(x,y)dy < \infty \text{ for } x \in S_X \ .$$

A nonnegative function that satisfies equation (15) is called a **stochastic transformation** from Y to X. Roughly, a "bad" signal is obtained subjecting a "good signal" to an additional randomization *independent of the true value*. It is intuitively appealing that, if the "good" signal was positively correlated with v, then the "bad" one will be less correlated, since it reflects an additional noise element independent of v. Thus, also Blackwell's notion informally predicts that a better signal will be more correlated with the true value. In some sense we can see the action of our $T(\cdot)$ transformation as the opposite process of Blackwell's stochastic transformation: the $T(\cdot)$ transformation produces "better" information by taking X and adding to it a little correlation with V; conversely, Blackwell's notion implies that "worse" information is obtained by taking X and adding a little noise to it, thereby reducing its correlation with V.

Blackwell's theorem asserts that if the decision maker can act upon observing X or Y, and seeks to maximize the expected value of u, then for every prior $g(v)$ and for every payoff function $u(v,a)$ he is better off observing signal Y if and only if Y is sufficient for X.

In the light of the results of Section 4.2, it is reasonable that Accuracy be a more general definition of "more informative" than Sufficiency, on the restricted set of decision problems to which the former applies. This is because Sufficiency is a necessary and sufficient characterization of "more useful information" for *all* decision problem. Section 4.2 shows that Accuracy is also necessary and sufficient, but of course on a *subset* of all decision problems. In other words, let D represent the set of all decision problems, and $D' \subset D$ represent the ones with a non-weakly affiliated signal structure and where $\frac{\partial}{\partial x} u^{\eta}$ satisfies **(QM)**. It must be that the signals which increase the revenue for all problems in D (Sufficiency) are fewer than those which increase the revenue for all problems in D'(Accuracy).

A more formal way to approach this issue is a restatement of Theorem 2. Consider a family $\{X_\eta\}$ non-weakly affiliated with V: taking into account Blackwell's theorem we can rewrite the statement of the theorem as

$$\{\text{The family } \{X_\eta\} \text{ is not A-ordered }\} \Rightarrow \{\ X_\theta \text{ is not sufficient for } X_\eta\ \}$$

The counternominal of this sentence shows that Sufficiency implies Accuracy in a non-weakly affiliated family, and thus in this setting Accuracy is more general (less restrictive a condition) than Sufficiency.

Even more formally, define a family of signals to be **S-ordered** (Sufficiency-ordered) when $\eta' < \eta''$ implies $X_{\eta''}$ is sufficient for $X_{\eta'}$. In other words, in an S-ordered family a higher signal is sufficient for a lower one. Then

Proposition 5 *If a family of signals is non-weakly affiliated and S-ordered, then it is A-ordered.*

Proof: The statement of Theorem 2 may be written as

$$\{\text{The family } \{X_\eta\} \text{ is not A-ordered }\} \Rightarrow \left\{\begin{matrix}\exists \eta\ :\ \forall \epsilon > 0, \exists \theta \in (\eta, \eta+\epsilon) \\ \text{such that } X_\theta \text{ is not sufficient} \\ \text{for } X_\eta\end{matrix}\right\}$$

Negating the above sentence gives

$$\left\{\begin{matrix}\forall \eta\ \exists \epsilon > 0 \text{ such that } \theta \in (\eta, \eta+ \\ \epsilon) \text{ implies } X_\theta \text{ is sufficient for} \\ X_\eta\end{matrix}\right\} \Rightarrow \{\text{ The family } \{X_\eta\} \text{ is A-ordered }\} \tag{16}$$

Since the sentence { The family $\{X_\eta\}$ is S-ordered } implies the left-hand sentence in (16), the proposition is proved. □

This means that all results of the form "increasing Accuracy implies something" hold, *a fortiori*, in the form "increasing informativeness in the Blackwell's sense implies something". This is of interest for results of the type of Theorem 3.

The next proposition shows that, while Sufficiency implies Accuracy, the converse is not true. Here we have a non-weakly affiliated family where increasing the parameter increases Accuracy (see Example 1), but does not increase informativeness in Blackwell's sense. To show this, we construct a payoff function for which increasing the Accuracy parameter results in a reduction of revenue; this, in view of Blackwell's theorem, shows that increasing Accuracy cannot produce a sufficient signal. Of course, to find such an example, we have to resort to a payoff where $\frac{\partial}{\partial x}u^{\xi}$ is *not* **(QM)**, otherwise Theorem 1 would doom the effort.

Proposition 6 *Let $b < c$ be two real numbers, and consider signals $X_b \sim U[v-b, v+b]$ and $X_c \sim U[v-c, v+c]$. If $c < 2b$ then X_b is* not *sufficient for X_c.*

Proof: In view of Blackwell's theorem, to prove our claim it is sufficient to find a decision problem and a prior for V such that X_b yields lower expected utility than X_c. Let us do this.

As a prior for V let us choose a very diffuse distribution (a Uniform on $[-1000b, +1000b]$ will do). Consider the following payoff function:

$$u(v,a) = \begin{cases} 0 \text{ if } a \in (v-b, v+b) \\ K_1 > 0 \text{ if } a \in (v-c, v-b) \cup (v+b, v+c) \\ K_2 < 0 \text{ elsewhere} \end{cases}$$

where a is an action. See fig 3

Consider the decision problem with signal X_b: the decision maker wants to play in $(v-c, v-b) \cup (v+b, v+c)$. However he does not know v, but only x, the realization of his signal. He can thus decide for an a lower or higher than x. Because of the payoff structure, any deviation of size less than $c-b$ is advisable, because at worse it produces no effect, and at best it puts him in the K_1 region. Any deviation of size larger than $c-b$, however, carries with it the risk of landing in the K_2 region. By choosing K_2 low enough, we can therefore make sure that the optimal strategy, $a^b(x)$, lies within $c-b+\epsilon$ of x. For definiteness, assume that at x the optimal strategy prescribes a *negative* deviation of size $c-b+\epsilon$.

Let us now calculate the optimal expected payoff to a decision maker with signal X_b and realization x. Given his signal, the posterior for V is

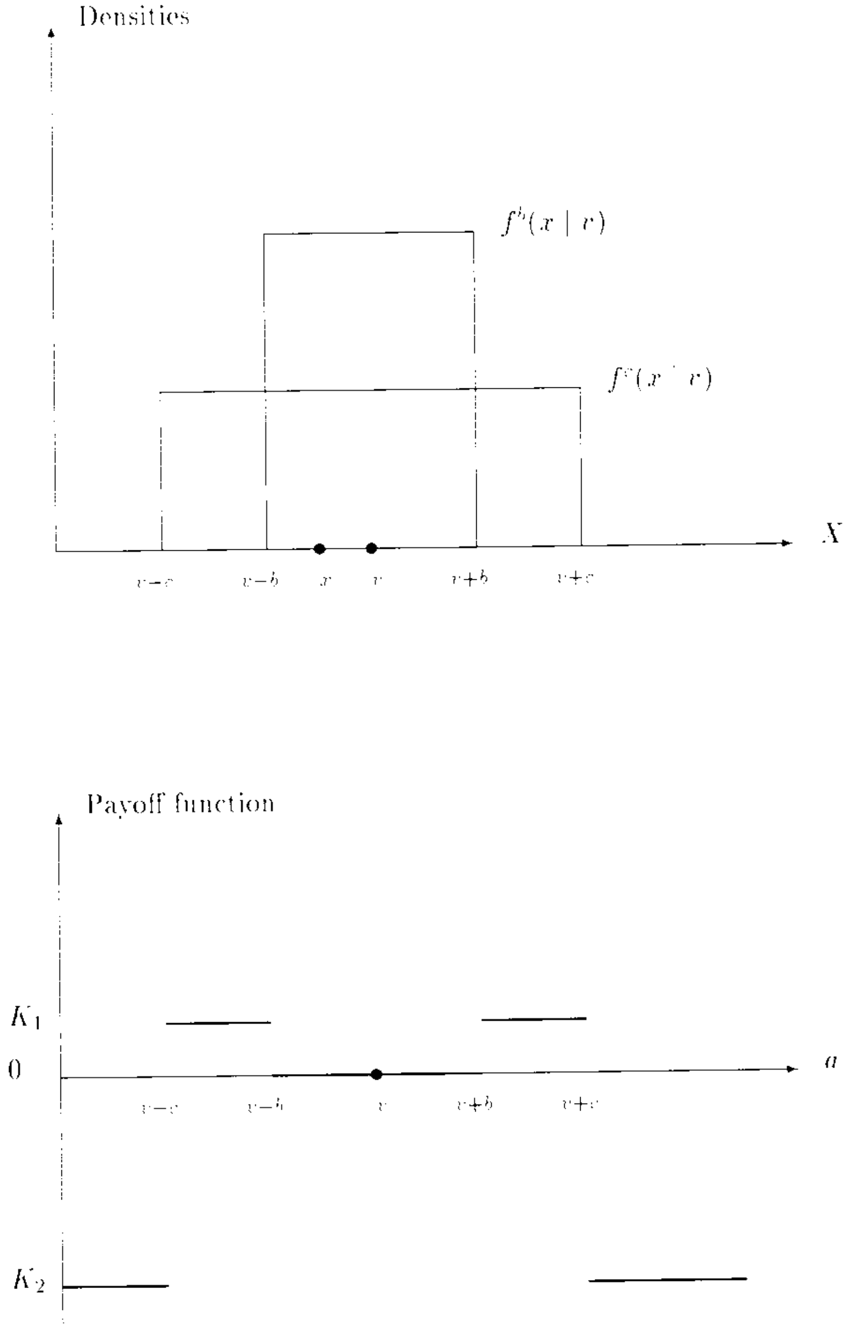

Figure 3: X_b is not sufficient for X_c

concentrated on $[x - b, x + b]$. His strategy will be successful whenever he lands on the K_1 region; because his action is lower than his signal, for this to happen it must be that

$$x - (c - b + \epsilon) < v - b. \tag{17}$$

Whenever the opposite inequality holds, he will have landed on the "0 region".

Because the prior is very diffuse, we may approximate the posterior distribution for V with a uniform on $[x - b, x + b]$, and the set of v's where the strategy is successful is (looking at eq. (17)) at most $[x + b - (c - b + \epsilon), \infty)$, whereby the posterior probability of success is at most $(c - b + \epsilon)/2b$. We can thus compute that given any signal x (except possibly those around $-1000b$ and $+1000b$)the expected payoff is at most

$$\frac{(c - b + \epsilon)}{2b} K_1 + 0.$$

Observe now that a decision maker endowed with signal X_c can, by just setting $a(x) = x$, obtain an expected payoff of

$$\frac{2(c - b)}{2c} K_1 + 0.$$

By choosing ϵ small enough, it is seen that signal X_c will yield a larger payoff than signal X_b when $c < 2b$, for nearly all values of x. Integrating over x concludes the proof. □

5 The information acquisition problem

In this section we discuss the information acquisition problem we presented in the introduction. The core is a comparative statics result, on which problem will induce more information acquisition, when acquiring information means increasing Accuracy. Since the previous section showed that under non-weak affiliation a sufficient signal is more accurate, we have that Theorem 3 holds, *a fortiori*, when acquiring information is meant in the Blackwell sense.

Proposition 7 and Corollary 1 give sufficient conditions for the information acquisition problem to admit finite solution. Theorem 3 parts 2) is a comparative statics result: it tells which of two payoff functions will yield greater information acquisition in equilibrium. In order to do so, and given a family of signals, Theorem 3 defines (in its parts 1)) a partial ordering on the set of payoff functions, according to the incentives they give to increase the Accuracy of one's signal.

The intuition underlying Proposition 7 below is quite simple. Suppose that when Accuracy goes to infinity the signal becomes totally accurate: then asymptotically the returns to increasing one's Accuracy must go to 0, since knowing almost everything already gets you very close to the "full information" optimum, which is the best one can do in the decision problem. In other words, we can think of the function $R(\eta)$ as being bounded from above, and that this bound is actually approached as $\eta \to \infty$: it therefore makes sense that its derivative with respect to η goes to 0 asymptotically.

Proposition 7 *Given an A-ordered information acquisition problem with $E = [\underline{\eta}, \infty)$, assume that $\lim_{\eta\to\infty} X_\eta = v$ almost surely. Suppose $a^\eta(x)$ has bounded derivative, that the support of V and of the family of signals is bounded, and finally that $\frac{\partial}{\partial\theta} T_{\eta,\theta,v}(x)\Big|_{\theta=\eta}$ is bounded (as a function of v) uniformly in x as $\eta \to \infty$. Then $\lim_{\eta\to\infty} MR(\eta) = 0$.*

Proof: Because $\lim_{\eta\to\infty} X_\eta = v$, we will have that

$$\lim_{\eta\to\infty} a^\eta(x) = \text{argmax}_a u(v, a).$$

Therefore, since $u(v, a)$ is continuously differentiable and $a^{\eta\prime}$ is bounded for all η,

$$\lim_{\eta\to\infty} \left[\frac{\partial}{\partial a} u(v, a^\eta(x))\right] a^{\eta\prime}(x) = 0a^{\eta\prime}(x) = 0.$$

But then equation (7) converges to 0 because $\frac{\partial}{\partial\theta} T_{\eta,\theta,v}(x)\Big|_{\theta=\eta}$ is bounded as $\eta \to \infty$, and this concludes the proof. □

Corollary 1 *If an information acquisition problem is as in Proposition 7 and if $\lim_{\eta\to\infty} MC(\eta) > 0$, an optimal Accuracy exists and is finite.*

Proof: Straightforward. □

We now turn to the main result of this subsection. The intuition behind Theorem 3 parts 1) is as follows. We have seen that Theorem 1 gets its kick from $\frac{\partial}{\partial x} u^\eta$ being **(QM)**. It is therefore reasonable that, the "more **(QM)**" $\frac{\partial}{\partial x} u^\eta$ is, the higher the increase in revenue from increasing η. The "more **(QM)**" relation is what we defined as $\overset{(\mathbf{QM})}{\geq}$. Once this is established, Parts 2) follow quite easily from marginal reasonings.

Theorem 3 *Suppose two payoff functions, say $u_I(v,a)$ and $u_{II}(v,a)$ are associated to the same A-ordered statistical structure, giving rise to two information acquisition problems. If for all x and η we have $\frac{\partial}{\partial x}u_I^\eta(v,x) \overset{(QM)}{\geq} \frac{\partial}{\partial x}u_{II}^\eta(v,x)$ on $\mathcal{V}_x^\eta$, then*

W)

1) for all η, $MR_I(\eta) \geq MR_{II}(\eta)$.

2) the set of optimal accuracies in problem I is higher (in the strong set order [1]) than that of problem II.

S)

suppose in addition that for all x, η we have $\frac{\partial}{\partial x}u_I \neq \frac{\partial}{\partial x}u_{II}$ on $\mathcal{V}_x^\eta$ and the family of signals is strictly A-ordered. We have

1) for all η, $MR_I(\eta) > MR_{II}(\eta)$.

2) the minimal optimal Accuracy in problem I is greater or equal than the maximal optimal Accuracy in problem II.

Proof:

Part W): The proof of 1) follows directly from the definition of the relation $\overset{(QM)}{\geq}$, coupled with Theorem 1 part W).

Part 2) follows from the fact that the function $\Pi_m(\eta) := R_m(\eta) - C(\eta)$ (where $m = I, II$ and $II < I$) is seen to have the single-crossing property in $(\eta; m)$. This is easily proved using part 1) above. Then Theorem 4 by Milgrom and Shannon [9] on monotone comparative statics yields the result.

Part S): 1) follows from Theorem 1 part S1).

The proof of part 2) is straightforward. □

6 Applications

6.1 A monopolist's information acquisition.

This subsection is meant to provide a simple application of the theory developed above to the problem of a monopolist acquiring information about an unknown parameter of his demand function. Throughout this section we will implicitly assume that the quantities involved are finite, in order to avoid trivial problems.

[1] Given A and B subsets of the real line, we say that A **is higher than** B **in the strong set order** when, for every $a \in A$ and $b \in B$ we have $\max\{a, b\} \in A$ and $\min\{a, b\} \in B$

Suppose that a monopolist has 0 marginal cost of producing quantity q of a good, for which the demand function is

$$P_{a,b}(q) = av - bq.$$

Here a, b are some known parameters, and v is unknown, with a prior distribution $g(v)$.

The monopolist observes a signal X_η, distributed according to $f^\eta(x \mid v)$, which conveys information about v. We assume that X and V are non-weakly affiliated, and that the family of signals $\{X_\eta\}$ is A-ordered.

The choice of η, the signal's Accuracy, is made prior to observing the signal, at a cost $C(\eta)$. We posit $\eta \in [\underline{\eta}, +\infty)$, and $\lim_{\eta \to \infty} C'(\eta) > 0$.

The monopolist's problem is, for a given x and η,

$$\max_q \int_V q P_{a,b}(q) dG^\eta(v \mid x).$$

Solving the first-order conditions it is seen that the optimal production $q^\eta_{a,b}(x)$ is

$$q^\eta_{a,b}(x) = \frac{a}{2b} E(V \mid X^\eta = x) = \frac{a}{b} q^\eta_{1,1}(x).$$

In view of non-weak affiliation, this quantity is increasing in x.

We may thus calculate the expression

$$\frac{\partial}{\partial x} u^\eta_{a,b}(v,x) = \frac{a^2}{b} q^{\eta\prime}_{1,1}(x)[v - 2q^\eta_{1,1}(x)] = \frac{a^2}{b} \frac{\partial}{\partial x} u^\eta_{1,1}(v,x) \tag{18}$$

and notice that it satisfies **(SQM)** (indeed, it is increasing in v). Thus, we can apply the theory developed above, and conclude that $MR_{a,b}(\eta) > 0$. The optimal Accuracy, let us call it $\eta^*_{a,b}$, will lie at a point where MR crosses C' from below.

As an application of Proposition 1, we have

Proposition 8 *When $\frac{\partial}{\partial \theta} T_{\eta,\theta,\cdot}(x)$ is bounded (as a function of v) uniformly on x, the optimal Accuracy $\eta^*_{a,b}$ is finite.*

It is of interest to conduct some comparative statics results on a and b, as they vary separately or together.

Proposition 9 *Consider two pairs of parameter values, (a,b) and (a',b'). Then $\eta^*_{a',b'} > \eta^*_{a,b}$ if and only if $\frac{a'^2}{b'} > \frac{a^2}{b}$.*

Proof: Using equation (18), it is immediate to show that

$$\left\{ \frac{\partial}{\partial x} u^{\eta}_{a',b'} \overset{\textbf{(QM)}}{\geq} \frac{\partial}{\partial x} u^{\eta}_{a,b} \text{ and } \frac{\partial}{\partial x} u^{\eta}_{a',b'} \neq \frac{\partial}{\partial x} u^{\eta}_{a,b} \right\} \Leftrightarrow \frac{a'^2}{b'} > \frac{a^2}{b}.$$

But then Theorem 3 part S2) gives the result. □

Corollary 2 *Suppose the monopolist had a constant marginal cost of production c. Then the higher c, the lower the equilibrium Accuracy.*

6.2 Market for doctors

Freixas and Kihlstrom in [5] examine the demand for information in the market for doctors.

The set of doctors is $(-\infty, +\infty)$, the patient's health after visiting doctor a is denoted by $u(r, a)$ where

$$u(r, a) = ra - ba^2.$$

Here b is a known positive constant, but r can take on any real value. Later on, they consider the alternative utility function

$$U(r, a) = -e^{-\gamma u(r,a)} K$$

where γ and K are known positive constants. In their work, they consider the case where the prior distibution on r is normal with mean μ and variance $1/v$, with v a known constant. In addition, the agent can observe a signal X_θ, with mean r and variance $1/\theta$. The parameter θ can be chosen by the agent at a cost. The agent's decision problem is, after having chosen a θ, to use it to infer the right a in order to maximize his expected utility.

It is straightforward to verify that both $u(r, a)$ and $U(r, a)$ satisfy **(WSCP)**. We are able to conclude that in this framework any affiliated A-ordered family of signals would have served the purpose, and it is not necessary to restrict to the normal environment.

6.3 A take-it-or-leave-it offer

A seller owns an object which he values r, and knows its value. A buyer has utility $u(r)$ for the object, where $u(r) > r$. The buyer does not know r, but can observe a signal X_η, at cost $C(\eta)$. The game is as follows: the buyer makes an offer $a(x)$, conditional on observing $X_\eta = x$. The seller either accepts or rejects. If the seller rejects, the buyer gets utility 0.

In any subgame perfect equilibrium, the seller accepts if and only if $a(x) > v$. Thus, we can express the payoff to the buyer as

$$\int_{-\infty}^{+\infty} [u(v) - a] I_{v<a} dG(v \mid x).$$

In the notation of this paper, we have

$$u(v, a) = [u(v) - a] I_{v<a}.$$

This is readily verified to be weakly single-crossing, and thus Theorem 5.1 of Athey [1] gives us monotonicity of the optimal strategy, under affiliation of signals.

Using an appropriate notion of derivative, we can write

$$\frac{\partial}{\partial a} u(v, a) = -I_{v<a} + [u(a) - a] g(a \mid x).$$

This is **(QM)**, and since $a'(x) > 0$, we conclude that $\frac{\partial}{\partial x} u^\eta(v, x) = \left[\frac{\partial}{\partial a} u(v, a(x))\right] a'(x)$ is **(QM)** also.

7 Conclusions

This paper deals with decision problems, where an agent has a payoff depending on his action and on an unobserved parameter. The agent maximizes his expected payoff conditional on the realization of a signal. The novelty is that the agent can choose the informational content of the signal, at a cost.

In this work we have introduced "A-order", a novel concept of "better information", suited to a subset of economically interesting decision problems: those where the payoff function exhibits the single-crossing property, and the optimal strategy is monotone in the signal.

The concept of A-order is defined for continuous families of signals: a family of signals is A-ordered when a slightly higher signal is slightly more correlated with the unknown random variable. This notion captures the intuitive idea that "more correlation with the unknown parameter is better". We show that, whenever a family of signals is non-weakly affiliated and is Sufficiency-ordered (higher signals are sufficient for lower ones), then it is A-ordered, while the converse is not true. In this sense the concept of A-order is more general (less restrictive) than Blackwell's Sufficiency, on the set of non-weakly affiliated families of signals.

Further, we show that the property of A-order is a "tight" notion on the class of decision problems with single-crossing payoffs and non-weakly affiliated signals: whenever A-order fails, we can find a single-crossing payoff where a higher signal gives lower payoff. In other words, by restricting to decision problems with single-crossing payoffs and non-weakly affiliated signals, we are able to give an if-and-only-if characterization of "better information". In this sense, this work is similar in spirit to Blackwell's, only restricted to a subset of economically interesting decision problems.

As a more practical byproduct of the analysis, a comparative statics result is developed, ordering decision problems by the amount of information acquired at the optimum. The flavour of the result is that the more single-crossing the payoff, the higher the optimal Accuracy. The intuition is that "single-crossingness" should be seen as "amount to which the payoff varies with the unknown parameter". Hence, when Accuracy is acquired at a cost, a problem where the payoff is highly dependent on the unknown variable (hence highly uncertain) will display a higher optimal choice of Accuracy. This result obviously holds when a higher signal is more informative in Blackwell's sense, since this is just a special case (under non-weak affiliation) of increasing Accuracy.

A-order, our notion of better information, is defined on continuous families of signals rather than through the comparison of two given signals: this is because Theorem 1 is a marginal result. The question of whether the restriction to continuous families of signals is inessential, or it carries with it mathematical substance, is an interesting one. This is a matter for future research.

References

[1] S. Athey, "Monotone Comparative Statics in Stochastic Optimization Problems", MIMEO January 1995.

[2] D. Blackwell, "The comparison of experiments". In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 93–102. Berkeley: University of California Press, 1951.

[3] D. Blackwell, "Equivalent comparison of experiments", *Annals of Mathematical Statistics*, pp. 265–272, June 1953.

[4] M. DeGroot, Optimal Statistical Decisions, New York, Mc Graw Hill, 1970.

[5] X. Freixas and R. Kihlstrom, "Risk aversion and information demand" In M. Boyer and R.E. Kihlstrom, eds.*Bayesian Models in Economic Theory*, pp. 93–104, Elsevier Science Publishers, 1984.

[6] S. Karamardian and S. Schaible, "Seven Kinds of Monotone Maps", *Journal of Optimization Theory and Applications*, **66** (1990) pp.37–46.

[7] S. K. Kim, "Efficiency of an Information System in an Agency Model", *Econometrica*, **63** (1996), pp. 89–102.

[8] S. Matthews, "Information Acquisition in Discriminatory Auctions", In M. Boyer and R.E. Kihlstrom, eds.*Bayesian Models in Economic Theory*, pp. 181–207, Elsevier Science Publishers, 1984.

[9] P. Milgrom and C. Shannon, "Monotone Comparative Statics", *Econometrica* **62** (1994), 157–180.

[10] P. Milgrom and R. Weber, "A Theory of Auctions and Competitive Bidding", *Econometrica* **50** (1982), 1089–1122.

[11] N. Persico, "Information Acquisition in Auctions", MIMEO Northwestern University, February 1996.

[12] Z. Safra and E. Sulganik, "On the Nonexistence of Blackwell's Theorem-Type Results with General Preference Relations", *Journal of Risk and Uncertainty*, **10** (1995), pp. 187–201.