Stochastic Orders and Decision under Risk
IMS Lecture Notes – Monograph Series (1991)

APPLICATIONS OF LIKELIHOOD RATIO ORDERINGS IN ECONOMICS

BY IAN JEWITT

University of Bristol

The likelihood ratio ordering has recently been used in economic theory. This paper examines some of its applications in portfolio theory, auction theory, and agency theory.

0. Introduction. A variety of familiar stochastic orderings are induced by classes of real valued functions: the random variable X_2 is said to stochastically dominate the random variable X_1 with respect to the class of functions C if

$$Eu(X_2) \geq Eu(X_1) \text{ for all } u \in C. \tag{0.1}$$

In what follows we shall be concerned with sufficient conditions expressed in terms of densities, so letting X_1 and X_2 have densities f_1 and f_2 respectively, (0.1) becomes

$$\int u(x)f_2(x)dx \geq \int u(x)f_1(x)dx \text{ for all } u \in C. \tag{0.2}$$

It is possible to consider inequality (0.2) a property of a linear transformation from the function space containing C to functions of the form $Tu : \{1,2\} \to \mathbb{R}$. It is required that Tu be increasing, where

$$Tu(i) = \int u(x)f_i(x)dx \qquad i = 1, 2.$$

Widening the enquiry somewhat, it is often of interest to know when a transformation T maps some known class of functions C into some other known class C'. So, we want conditions on f such that $TC \subset C'$ where

$$Tu(y) = \int u(x)f(x \mid y)dx.$$

AMS 1980 Subject Classification: Primary 90A05; Secondary 90A09.

Key words and phrases: Total positivity, risk aversion, auction theory, portfolio theory, principal-agent problem.

It is often useful to be able to verify the desired property after conditioning on certain sets. Conditioning on a set S corresponds to an operation on the density of the following form: $f \to f^*$, where

$$f^*(x \mid y) = \frac{f(x \mid y)g(x)}{\int f(x \mid y)g(x)dx} \tag{0.3}$$

and g is the indicator function of the set S. In some economic applications, a transformation of the form (0.3) occurs but in which g is not the indicator function of a set, but is some other function – for instance, a marginal utility.

The following simple problem illustrates this. The economic interpretation is that it is a portfolio problem with one safe and one risky asset. The problem is to choose s to maximize

$$\int u(sx + (1-s)r)f(x \mid y)dx,$$

where x denotes the risky return and r the safe return.

We want to find conditions on f which imply that the optimal choice of s increases with the parameter y for all increasing concave u. Given sufficient regularity at the optimum,

$$\int u'(sx + (1-s)r)(x-r)f(x \mid y)dx = 0. \tag{0.4}$$

By concavity, the left-hand side of (0.4) is decreasing in s therefore if at some s we have

$$\int u'(sx + (1-s)r)(x-r)f(x \mid y)dx > 0,$$

then it would be necessary to increase s in order to restore the first-order condition. In other words, if

$$\int (x-r)f^*(x \mid y)dx \ \text{ is increasing in } \ y$$

where

$$f^*(x \mid y) = \frac{f(x \mid y)u'(sx + (1-s)r)}{\int f(x \mid y)u'(sx + (1-s)r)dx}$$

then the optimal s is increasing in y. Since the likelihood ratio is not affected by the transformation $f \to f^*$, if it is possible to find conditions on the likelihood ratio which ensure that the distributions are ranked by their means then the desired result will obtain. It is of course well known that if f has a monotone likelihood ratio, then the condition is satisfied.

Milgrom (1981) characterizes the monotone likelihood ratio property and displays its usefulness in a variety of economic contexts including this simple portfolio problem. Landsberger and Meilijson (1989) also give a discussion of it.

It is therefore convenient to be able to deal with a class of densities which not only have the desired "stochastic ordering" properties, in that $TC \subset C'$ for the classes of functions C and C' of interest, but which is also invariant with respect to transformations of the form (0.3).

If x and y are both scalars, then the situation is relatively straightforward, any property defined by the sign of the determinants

$$\begin{vmatrix} f(x_1 \mid y_1) & \cdots & f(x_1 \mid y_n) \\ \vdots & & \vdots \\ f(x_n \mid y_1) & \cdots & f(x_n \mid y_n) \end{vmatrix}$$

will be invariant to transformations of the form (0.3).

1. Total Positivity. Karlin (1968) is the major work on this topic, and the reader is referred to it for further details. We give only a brief outline in what follows.

DEFINITION. Let X and Y be subsets of the real line $\mathbb{R}$, the function $K : X \times Y \to \mathbb{R}$ is totally positive of order n (TP_n) if $x_1 < \cdots < x_n$ and $y_1 < \cdots < y_n$ imply that

$$\begin{vmatrix} K(x_1, y_1) & \cdots & K(x_1, y_m) \\ \vdots & & \vdots \\ K(x_m, y_1) & \cdots & K(x_m, y_m) \end{vmatrix} \geq 0$$

for each $m = 1, \cdots, n$.

1.1. The Variation Diminishing Property. Let K be TP_n on $X \times Y$ and μ be a σ-finite measure on X. If the function $g : \mathbb{R} \to \mathbb{R}$ has at most $k \leq n-1$ sign changes, then the function

$$g^*(y) = \int g(x)K(x,y)d\mu(x), \qquad y \in Y,$$

has at most k sign changes. Moreover, if g^* has exactly k sign changes then g and g^* have the same pattern of sign changes.

We shall give two simple examples of the use of this property. Suppose $f(x \mid y)$ is a density with respect to μ for each value of the parameter y, so $\int f(x \mid y)\mu(dx) = 1$ for all $y \in Y$ and let

$$g^*(y) = \int g(x)f(x \mid y)d\mu(x), \qquad y \in Y \tag{1.1}$$

It follows directly from the variation diminishing property that

(a) if g is increasing and f is TP_2, then g^* is increasing.

(b) if g is quasi-concave and f is TP_3, then g^* is quasi-concave.

By way of illustration, we shall prove (b). Quasi-concavity of g can be characterized as follows: For each constant c, the function $g(x) - c$ has at most two sign changes, and moreover if two sign changes occur the first is from negative to positive. By the variation diminishing property,

$$\begin{aligned}\int (g(x) - c)f(x \mid y)d\mu(x) &= \int g(x)f(x \mid y)d\mu(x) - c \\ &= g^*(x) - c\end{aligned}$$

has the same sign change property, hence g^* is quasi-concave.

2. Application.

2.1. Preservation of the Arrow-Pratt Risk Aversion Ordering. The increasing utility u is said to be more risk averse than the increasing utility v if u is a concave transformation of v (Arrow (1970), Pratt (1964)). That is, u is more risk averse than v if for each α, β

$$u(x) - (\alpha + \beta v(x))$$

has at most two sign changes and if two sign changes actually occur then the first is from negative to positive. In other words, $u(x) - \beta v(x)$ is quasi-concave for each choice of β. It follows that the transformation $g \to g^*$ in (1.1) preserves the ordering of functions by risk aversion. Some economic applications and extensions are given in Jewitt (1987).

2.2. Multivariate Total Positivity. There is a multivariate generalization of TP_2. Standard references are Karlin and Rinott (1980) and Fortuin, Ginibre and Kasteleyn (1971). Let x and y be vectors from $\mathbb{R}^n$ and let $\vee$ and $\wedge$ be the usual lattice operations; $x \vee y$ is the vector with components $\max\{x_i, y_i\}$ and $x \wedge y$ is the vector with components $\min\{x_i, y_i\}$.

DEFINITION. A function $K : \mathbb{R}^n \to \mathbb{R}$ is said to be multivariate totally positive of order two (MTP_2) if

$$K(x \wedge y)K(x \vee y) \geq K(x)K(y) \ \text{ for all } \ x, y \in \mathbb{R}^n.$$

A *strictly* positive function is MTP_2 if and only if it is TP_2 in each pair of its arguments taken separately (see Karlin and Rinott (1980) or Eaton (1987)).

Much of the convenience of MTP_2 functions stems from the following facts:

(a) The monotonocity preserving property holds for MTP_2 densities. Let $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ and let $f(x \mid y)$ be a density with respect to Lebesgue measure. If $f(x \mid y)$ is MTP_2 in (x, y) then

$$g^*(y) = \int g(x) f(x \mid y) dx, \qquad y \in \mathbb{R}^m,$$

is increasing whenever g is increasing.

(b) If $K : \mathbb{R}^n \to \mathbb{R}$ is MTP_2, then so is $\overline{K} : \mathbb{R}^{n+m} \to \mathbb{R}$ where $\overline{K}$ is the extension of K defined as

$$\overline{K}(x, y) = K(x), \qquad (x, y) \in \mathbb{R}^{n+m}.$$

(c) If $K : \mathbb{R}^{n+m} \to \mathbb{R}$ is MTP_2, then so is the "marginal" function $L : \mathbb{R}^n \to \mathbb{R}$ defined by

$$L(x) = \int_{\mathbb{R}^m} K(x, y) dy, \qquad x \in \mathbb{R}^n.$$

(d) If $K : \mathbb{R}^n \to \mathbb{R}$ and $L : \mathbb{R}^n \to \mathbb{R}$ are MTP_2, then so is their product M,

$$M(x) = K(x) L(x) \qquad x \in \mathbb{R}^n.$$

(e) Let 1_+ be the indicator function of $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$, then $K : \mathbb{R}^2 \to \mathbb{R}$ defined by $K(x, y) = 1_+(x - y)$ is MTP_2 on $\mathbb{R}^2$.

A useful consequence of properties (a)–(e) is the following (e.g. Karlin and Rinott (1980), Milgrom and Weber (1982)). Let $(X_1, \cdots, X_n)$ be a random vector with MTP_2 density, then for any increasing function $\Psi : \mathbb{R}^n \to \mathbb{R}$,

$$E[\Psi(X_1, \cdots, X_n) \mid a_i \leq X_i \leq b_i;\ i = 1, \cdots n] \text{ is increasing in } a, b. \tag{2.2.1}$$

2.3. Application to Portfolio Theory. Given a concave increasing (utility) function $v : \mathbb{R}^n \to \mathbb{R}$ and an endowment $\bar{a} \in \mathbb{R}^n$ we can ask what *prices* $p \in \mathbb{R}^n$ support $\bar{a}$ as a *demand.* The prices determine a hyperplane which separate the a's which are better from those which are cheaper. We want to find p such that

$$\bar{a} \text{ maximizes } v(a) \text{ on } \{a \mid pa \leq p\bar{a}\}.$$

Many models of asset pricing in economics are essentially of this form. We suppose the economy to be inhabited by a single consumer who nevertheless sets himself prices at which he can trade freely. Equilibrium prices are then determined where supply equals demand for all commodities.

The assumption of a single representative consumer is not as capricious as it might first appear but can be regarded as a corollary of a very old idea which is that individuals acting in their own self interest leads to the common good. The name of Adam Smith is most closely associated with this but Hirschman (1977) traces it back further to Montesquieu albeit in a slightly different context (Montesquieu thought that having individuals seeking glory for themselves was for the good of the state).

Economists have now put the old idea into a more precise form. This is the statement that a competitive economy achieves an efficient allocation of resources. "Efficient" here is used in the sense attributed to Vilfredo Pareto, that it is not possible to make anybody better off without making someone else worse off. It follows from this that in a market equilibrium there exists a "representative consumer" in the sense that prices support the aggregate consumption as best in this consumer's preferences. Constantinides (1982) takes this perspective in an intertemporal asset pricing model.

Suppose that there are n risky assets with asset i yielding a random liquidation value of X_i, $i = 1, \cdots, n$, and one safe asset yielding $X_o = r$. For notational convenience, we normalize so that $r = 1$. The assets are divided into shares, and equilibrium prices are those for which it is optimal for the representative consumer to have a shareholding of 1 in each asset. That is, we look for prices p such that $s_i = 1$, $i = 0, \cdots, n$ solves the following maximization problem:

$$\text{maximize } Eu(\Sigma s_i X_i) \text{ subject to } \Sigma p_i s_i = \Sigma p_i.$$

Since the set of feasible s_i's is not altered by, say, a doubling of all prices, equilibrium prices are only determined up to a scale factor and only relative prices matter. It is convenient to normalize so that the price of the safe asset is equal to 1 and thereby measures the prices of the risky assets in terms of that of the safe asset.

Given concavity of the utility function u, the maximum will be characterized by the first order conditions, it follows that the prices p satisfy

$$\frac{Eu'(\Sigma X_j)X_i}{Eu'(\Sigma X_j)} = p_i, \qquad i = 1, \cdots, n.$$

It is of interest therefore to explore conditions under which

$$\frac{Eu'(\Sigma X_j)X_i}{Eu'(\Sigma X_j)} \geq \frac{Eu'(\Sigma Y_j)Y_i}{Eu'(\Sigma Y_j)}, \qquad i = 1, \cdots, n. \tag{2.3.1}$$

Let $(X_1, \cdots, X_n)$ have the density f_1 , and $(Y_1, \cdots, Y_n)$ have the density f_2, then inequality (2.3.1) can be written

$$\int x_i \bar{f}_1(x)dx \geq \int x_i \bar{f}_2(x)dx, \qquad i = 1, \cdots, n,$$

where

$$\bar{f}_i(x) = \frac{u'(\Sigma x_j) f_i(x)}{\int u'(\Sigma x_j) f_i(x) dx}, \qquad i = 1, 2.$$

If $u'(\Sigma x_j)$ is MTP_2 in x, and f_i is MTP_2 in (x, i), then it follows that $\bar{f}_i(x)$ is MTP_2 in (x, i) and the desired result follows from the monotonicity preserving property of MTP_2 densities.

The condition that $u'(\Sigma x_j)$ be MTP_2 is equivalent to $u'(x)$ being log convex. Since u' is positive by assumption, MTP_2 is equivalent to $u'(\Sigma x_j)$ being TP_2 in each pair of x's taken separately and therefore it suffices to consider the case $n = 2$. The condition reduces to $u'(x_1 + y_1)\, u'(x_2 + y_2) \geq u'(x_1 + y_2)\, u'(x_2 + y_1)$ for $x_1 < x_2$, $y_1 < y_2$. Evidently, $((x_1 + y_1), (x_2 + y_2))$ majorizes $((x_1 + y_2), (x_2 + y_1))$, hence the equivalence. The condition is actually a very natural one in this context, it corresponds to the utility function u being *decreasing absolute risk averse.* Let $A(y)$ be the set of gambles, X, that would be acceptable to an individual with utility u, and initial wealth $y \in \mathbb{R}$, so

$$A(y) = \{X \mid Eu(X + y) \geq u(y)\}$$

then $A(y)$ being increasing in the sense of set inclusion is equivalent to u' being log convex, Arrow (1970), Pratt (1964), Dubins and Savage (1965). Hence, log convexity of u' corresponds to reduced risk aversion at higher levels of initial wealth.

The conditions are not necessary ones and other convenient sufficient conditions have been derived for the case where there is only one risky asset. In this case, the question is equivalent to identifying the distribution changes of the risky asset which will encourage investors to hold more of the risky asset, which we touched on briefly in the introduction. Other conditions are derived by Landsberger and Meilijson (1989) and Black and Bulkley (1989). One is that the two distributions have the same means and a unimodal likelihood ratio. Whitt (1985) also discusses this relation but in a different context.

2.4. Application to Auction Theory. Different institutional arrangements often exist for selling different objects such as works of art, boxes of fish, cut flowers, mineral rights, houses, etc.

Dutch auctions are commonly executed with the aid of a clock, the single hand of which sweeps down from higher to lower prices until one of the participants stops the clock and is awarded the object at the price shown. (This is a common method of selling cut flowers in the Netherlands.) Vickrey (1961) had the insight that this way of disposing of an object is essentially equivalent to a sealed bid mechanism whereby the potential buyers all write their bids on pieces of paper and submit them to the auctioneer under the following rules:

the person who bids highest wins the object and pays the amount of his bid, nobody else makes any payment.

In an *English auction* increasing bids are solicited by the auctioneer until no one is willing to bid any further. The last person to bid is awarded the object at the price bid. This also has a sealed bid version which is conducted with written bids as before except now the highest bidder wins the object but only pays the amount of the *second* highest bid.

Evidently it would be possible in theory to design any number of sealed bid mechanisms in which the winner is charged some other statistic of the bids, e.g. the third highest bid, or the mean, or the median. Another variant has actually been used in practice in nineteenth century England. Each bidder submits two numbers, the bidder whose numbers contain the largest of those submitted wins the object but only pays the lowest of his numbers which is greater than all the numbers submitted by his competitors. One could try selling an object by a similar method but in which participants submit, say, four numbers. Another fairly common way of selling an object is by lottery. One well known race track gambler sold his Irish mansion by lottery and reputedly got a very good price, but some form of auction is a more common method.

It is of interest to analyze the performance of different auction mechanisms, in what follows, we shall concentrate on the criterion of expected revenue to the seller and we shall restrict attention to the sealed bid versions of the English (second price) and Dutch (first price) auctions.

In order to analyze the problem, it is necessary to cast it in game theoretic terms and adopt some notion of equilibrium. The Nash noncooperative equilibrium seems to be the most acceptable solution concept. In a Nash equilibrium each player's strategy maximizes his payoff *given* the strategies of the other players.

From a certain viewpoint, the game is one of incomplete information – for instance, if all the bidders know their own valuations for a work of art but do not know the valuations of the competing bidders. Vickrey (1961) made the important suggestion that the game of incomplete information should be treated as a game of complete information in which the first move is nature's choice of individual valuations. Vickrey likened the situation to a parlor game in which the game commenced by each player drawing his valuation from a "pack of cards".

2.5. The General Symmetric Model (Milgrom and Weber (1982)). There are $(n+1)$ random variables $Y_1, \cdots, Y_n, S$ with the following interpretation. Bidder i observes Y_i and this comprises the information on which he must formulate his bid. S is another random variable which might contribute to the value of the object to the bidders. In Vickrey's independent private values models, the Y_i are simply individual valuations. In the common model of Wilson, S is the common value of the object to the individuals. A concrete example would be that of an auction for mineral rights. S is the actual size of the deposit and the Y_i are the results of geological surveys. The value of the object to the n bidders is given by

$$V_i = u(Y_i, Y_{-i}, S), \qquad i = 1, \cdots, n,$$

where Y_{-i} is the vector $(Y_1, \cdots, Y_n)$ with the ith element deleted, and u is an increasing function. The analysis will rest heavily on symmetry so it is assumed that for each bidder $i = 1, \cdots, n$, $u(y_i, y_{-i}, s)$ is symmetric in y_{-i}. The random vector $(Y_1, \ldots, Y_n, S)$ has a density $f(y_1, \cdots, y_n, s)$ which is assumed to be symmetric in $(y_1, \cdots, y_n)$. It is natural to think of the y_i's as estimated valuations of the object, so we want $E[V_i \mid Y_i = y]$ to be increasing. This is ensured by the assumption that $(Y_1, \cdots, Y_n, S)$ are *affiliated*, to use the terminology introduced in Milgrom and Weber. A collection of random variables is said to be affiliated if their joint density is MTP_2. Moreover, if $(Y_1, \cdots, Y_n)$ are affiliated, then for any increasing Ψ

$$E[\Psi(Y_1, \cdots, Y_n) \mid \max\{Y_2, \cdots, Y_n\} \leq x, Y_1 = y]$$

is increasing in y. Also, since the indicator function for the set $\max\{y_2, \cdots, y_n\} \leq x$ is decreasing in $(y_2, \ldots, y_n)$,

$$\text{Prob}[\max\{Y_2, \cdots, Y_n\} \leq x \mid Y_1 = y]$$

is decreasing in y.

At the time of bidding, the only information that bidder i has is his estimate y_i, a realization of Y_i. The amount he bids in a given auction form will therefore, in general, depend on the y_i observed. Hence restricting attention to pure strategies, bidder i's strategy will be a function b_i mapping from the support of Y_i which we take to be $[0, \infty)$ to $[0, \infty)$.

Since the situation is symmetric between the bidders, it is natural to suppose that in equilibrium they all adopt the same strategy. So we look for a function b which is a best response when played against itself.

Take the first bidder as representative and let $Z = \max\{Y_2, \cdots, Y_n\}$. If the function $b^D : \mathbb{R}_+ \to \mathbb{R}_+$ is an equilibrium strategy for all the bidders in the first price auction (D stands for Dutch), then when bidder one has a valuation y_1, it is optimal for him to bid $b^D(y_1)$, and it is *not* generally optimal to bid $b^D(x)$ for $x \neq y$. Hence,

$$\begin{aligned}&E[(V_1 - b^D(x))1\{b^D(Z) \leq b^D(x)\} \mid Y_1 = y_1] \\ &\quad \leq E[(V_1 - b^D(y_1))1\{b^D(Z) \leq b^D(y_1)\} \mid Y_1 = y_1]\end{aligned}$$

for all $x, y_1 \in \mathbb{R}_+$. It is natural to suppose that in an equilibrium higher estimates lead to higher bids and given affiliation, this can indeed by proved (Milgrom and Weber (1982)). Hence, $1\{b^D(z) \leq b^D(x)\} = 1\{z \leq x\}$ and therefore in a (symmetric) equilibrium it must be the case that for each $y_1 \in [0, \infty)$, choosing $x = y_1$, maximizes

$$u^D(x, y_1) \equiv E[V_1 1\{Z < x\} \mid Y_1 = y_1] - E[b^D(Z)1\{Z < x\} \mid Y_1 = y_i].$$

$u^D(x, y_1)$ is the payoff that player 1 would get if everybody plays their equilibrium strategies but player 1 bids $b^D(x)$ when his estimate is y_1 rather than $b^D(y_1)$ i.e. if player 1 "pretends" to have a valuation x instead of y_1. $u^D(x, y_1)$ comprises two terms, one corresponds to the anticipated valuation of the object $A(x, y_1)$, and the other is simply the expected payment $C_u^D(x, y_1)$.

In the sealed bid version of the English auction the payoff to bidder 1 of bidding β when he has an estimate y_1 and everybody else adopts the strategy b^E is

$$E[(V_1 - b^E(Z))1\{b^E(Z) < \beta\} \mid Y_1 = y_1].$$

Taking for granted that the equilibrium strategy is increasing, it follows as before that for b^E to be an equilibrium strategy it must be the case that for each $y \in \mathbb{R}_+$, choosing $x = y$ maximizes

$$u^E(x, y_1) \equiv E[V_1 1\{Z < x\} \mid Y_1 = y_1] - E[b^E(Z)1\{Z < x\} \mid Y_1 = y_1].$$

Since for any y_1, both u^D and u^E are maximized at $x = y_1$, the partial derivatives vanish there,

$$\begin{aligned}A_1(y_1, y_1) - C_1^D(y_1, y_1) &\equiv 0 \\ A_1(y_1, y_1) - C_1^E(y_1, y_1) &\equiv 0,\end{aligned}$$

the 1 subscript represents differentiation with respect to the first argument of the function.

Hence,

$$C_1^D(y, y) = C_1^E(y, y) \text{ for all } y \in \mathbb{R}_+. \tag{2.5.1}$$

The identity (2.5.1) is useful in that it considerably simplifies analysis of the derivative of $C^D(y, y) - C^E(y, y)$.

Now

$$\begin{aligned} C^D(x,y) - C^E(x,y) \\ &= E[b^D(x) - b^E(Z) \mid Z \le x, Y_1 = y] \operatorname{Prob}[Z \le x \mid Y_1 = y] \end{aligned} \tag{2.5.2}$$

It follows from affiliation that as a function of y, the expression in (2.5.2) is the product of a decreasing positive function and a decreasing function, it follows that it is decreasing whenever it is positive. That is,

$$C^D(x,y) \ge C^E(x,y) \Rightarrow C_2^D(x,y) \le C_2^E(x,y).$$

Since $C_1^D(y,y) \equiv C_i^E(y,y)$, this implies that

$$C^D(y,y) \ge C^E(y,y) \Rightarrow \frac{d}{dy} C^D(y,y) \le \frac{d}{dy} C^E(y,y).$$

Hence, $C^D(y,y) - C^E(y,y)$ is decreasing wherever it is positive, since $C^D(0,0) = C^E(0,0) = 0$, it follows that $C^D(y,y) \le C^E(y,y)$ for all $y \in \mathbb{R}_+$. If he has an estimate $Y_1 = y_1$, the expected cost to player 1 in the first price auction is $C^D(y_1, y_1)$, so his expected payment is $EC^D(Y_1, Y_1)$. Since expected payments by bidders are expected revenues for the seller, it follows that the second price auction yields a higher expected revenue.

2.6. Application to Agency Theory. A franchise contract, say for a fast food restaurant, could take a variety of forms. At one extreme the franchisee would take all the profits and pay a fixed sum for the operating license. At the other extreme, the franchisor could take all the profits and pay the franchisee a fixed wage. The disadvantage of the first system is that the franchisee bears all the risks of the enterprise. This is not efficient if the franchisee is risk averse. The disadvantage of the second arrangement is that, since the franchisee gets paid the same whatever the profits are, he has no incentive to work hard. There is therefore a conflict between risk sharing and the provision of incentives. It is likely that similar conflict occurs rather widely. Crop-sharing arrangements in rural India are presumably an attempt to strike some sort of balance between incentives and risk sharing. The phenomenon has long been recognized in insurance circles where it is known as moral hazard: supplying people with insurance may mean that they take less care so the occurrence of the insured event becomes more frequent than an actuarial study of frequencies would have suggested prior to the introduction of insurance. A remarkable example arose when Japanese insurance companies started providing coverage against the eventuality of scoring "a hole in one" at golf (the custom is to stand a substantial round of drinks, plant a tree, and generally become out of pocket as a result of the good luck). Not surprisingly, the insurers noticed a remarkable increase in the frequency of reported "holes in one". A benign government faces something of a similar problem in designing a progressive tax system – there is a conflict between redistributing income to the poor and maintaining incentives for work.

The simplest version of the moral hazard problem can be formulated as a maximization problem. The *principal* cannot observe the action a taken by the agent but can observe the "output" x resulting from it. Given an action a, revenue is a random variable with (known) density f. The principal chooses a wage schedule $s : \mathbb{R} \to \mathbb{R}$ and an action a for the agent to maximize expected profit subject to the *incentive compatibility constraint* that the level of effort is optimal for the agent given s, and the *participation constraint*, that the expected utility attained by the agent is at least as great as some reservation utility – what the agent could get in an alternative occupation.

The problem is,

$$\underset{s,a}{\text{maximize}} \int (x - s(x))f(x,a)dx$$

subject to

$$\int u(s(x))f(x,a)dx - c(a) \geq \int u(s(x))f(x,a')dx - c(a') \text{ for all } a' \tag{IC}$$

$$\int u(s(x))f(x,a)dx - c(a) \geq R. \tag{PC}$$

This is an awkward problem to analyze under general conditions. A considerable simplification of the problem would occur if it were admissible to replace the constraint (IC) with the weaker one that the agent's expected utility be at a stationary point in effort. The problem is that there may be more stationary points than global maxima and so replacing the (IC) constraint in this way might well enlarge the principal's choice set in a way which affects the solution to the problem. *If* it is admissible to replace (IC) with

$$\int u(s(x))f_a(x,a)dx = c'(a) \tag{IC$'$}$$

where the subscript a denotes the partial derivative with respect to a, then a standard variational argument leads to

$$\frac{1}{u'(s(x))} = \lambda + \mu \frac{f_a(x,a)}{f(x,a)}. \tag{2.6.1}$$

Given (IC'),

$$\int u(s(x))\frac{f_a(x,a)}{f(x,a)}f(x,a)dx = c'(a) > 0.$$

It follows from this that s cannot be a decreasing transformation of f_a/f for then we would have

$$\int u\frac{f_a}{f}fdx \leq \int ufdx \cdot \int \frac{f_a}{f}fdx = 0 < c'(a).$$

The first term above is the covariance of u and f_a/f, if s is a decreasing transformation of f_a/f, then the covariance must be negative, but by (IC') it is equal to $c'(a)$ which is positive by assumption. Hence, $\mu > 0$ in (2.6.1). Given $\mu > 0$ we see (2.6.1) implies that the optimal payment schedule is increasing if $f_a(x,a)/f(x,a)$ is increasing in x. Given the assumed differentiability, this is equivalent to $f(x,a)$ having the monotone likelihood ratio property. In which case, we have the result that the monotone likelihood ratio property implies that the agents remuneration is increasing in output.

In order to justify the replacement of (IC) by (IC'), it suffices to show that at the choice of s solving the problem with constraint (IC'),

$$U(a) = \int u(s(x))f(x,a)dx - c(a)$$

is quasi-concave. To establish this for arbitrary convex increasing c, it is necessary and sufficient that $\int u(s(x))f(x,a)dx$ be concave.

Given the assumed monotone likelihood ratio property $u(s(x))$ is known to be increasing at the solution to the problem with the relaxed constraint (IC') so the desired conclusion will follow if the transformation $\varphi \rightarrow \varphi^*$ defined by:

$$\varphi^*(a) = \int \varphi(x)f(x,a)dx$$

maps the class of increasing functions into concave functions. A necessary and sufficient condition for this is that the distribution function $F(x,a)$ be convex in a for each value of x. This condition, together with the monotone likelihood ratio condition, is due to Mirrlees (1976). This is perhaps too strong an assumption to make, and so it may be preferable to weaken it. It is possible to make assumptions on u and f which ensure that (2.6.1) implies that $u(s(x))$ is concave (Jewitt (1988)). Hence it becomes of interest to establish conditions on f such that the transformation $\varphi \rightarrow \varphi^*$ defined above maps concave functions into concave functions. It is sufficient that $f(x,a)$ be TP_3 and $\int xf(x,a)dx$ be concave. Since TP_3 densities preserve the risk aversion relation, it follows that assuming $f(x,a)$ to be TP_3, and in addition that $\int xf(x,a)dx$ is concave suffices for $\int u(s(x))f(x,a)dx$ to be concave whenever $u(s(x))$ is concave. The assumption that $\int xf(x,a)dx$ be concave is a fairly natural assumption of decreasing returns.

The assumption that $f(x,a)$ be TP_3 can be utilized in this problem more directly.

Rather than set the problem up with continuous variables we follow Grossman and Hart (1983) and formulate a discrete version. There are m possible levels of effort: $a_1, \cdots, a_m$ and n possible outcomes $x_1, \cdots, x_n$. The probability of outcome x_i occurring when effort level a_j is adopted is $\pi_{ij} > 0$. If it is optimal for the principal to implement effort level k, say, then it is optimal to do so as cheaply as possible and so the optimal payment schedule: pay s_i if x_i occurs must minimize $\Sigma \pi_{ik} s_i$ subject to

$$\sum_i u(s_i)\pi_{ik} - c(a_k) \geq \sum_i u(s_i)\pi_{ij} - c(a_j) \quad j = 1, \cdots, n, j \neq k \tag{IC}$$

and

$$\sum_i u(s_i)\pi_{ik} - c(a_k) \geq R. \tag{PC}$$

This is a standard programming problem for which the Kuhn-Tucker conditions hold at the optimum. Hence, at the solution there are non-negative numbers λ and $\mu_1, \cdots, \mu_{n-1}$ such that

$$\frac{1}{u'(s_i)} = \lambda - \sum_j \frac{\pi_{ij}}{\pi_{ik}} \mu_j.$$

Consider the function $i \to \sum_j \frac{\pi_{ij}}{\pi_{ik}}(\mu_j - c\delta_{jk})$ where $\delta_{jk} = 1$ if $j = k$ and 0 otherwise.

For non-negative μ and fixed k, $\mu_j - c\delta_{jk}$ has at most two sign changes and when two occur, the first is from positive to negative. Since π_{ij}/π_{ik} is TP_3 in (i,j) if π_{ij} is, the variation diminishing property ensures that

$$\sum_j \frac{\pi_{ij}}{\pi_{ik}} \mu_j - c$$

inherits the sign change property and therefore $\sum_j \frac{\pi_{ij}}{\pi_{ik}} \mu_j$ is quasi-convex. Since $1/u'(s)$ is an increasing function, it follows that the mapping $i \to s_i$ is quasi-concave. Hence, if the density is TP_3, then the optimal payment schedule is quasi-concave. This is perhaps rather a weak result, but general results have proved hard to come by in this model, and our main intention is to illustrate the applicability of total positivity arguments to economic problems.

REFERENCES

ARROW, K. J. (1970). The theory of risk aversion, in Arrow, K. J. *Essays in the Theory of Risk Bearing*. North-Holland, Amsterdam.

BLACK, J. M. and BULKLEY, G. (1989). A ratio criterion for signing the effects of an increase in uncertainty. *Internat. Econom. Rev.* **30**, 119–130.

CONSTANTINIDES, G. M. (1982). Intertemporal asset pricing with heterogeneous consumers and without demand aggregation. *J. Business* **55**, 253–267.

DUBINS, L. and SAVAGE, L. J. (1965). *Inequalities for Stochastic Processes: How to Gamble if You Must.* Dover Publications, New York.

EATON, M. L. (1987). *Lectures on Topics in Probability Inequalities.* Centrum voor Wiskunde en Informatica, Amsterdam.

FORTUIN, C. M., GINIBRE, G. and KASTELEYN, P. N. (1971). Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.* **22**, 89–103.

GROSSMAN, S. and HART, O. (1983). An analysis of the principal-agent problem. *Econometrica* **51**, 7–345.

HIRSCHMAN, A. O. (1977). *The Passions and the Interests.* Princeton University Press.

JEWITT, I. (1987). Risk aversion and the choice between risky prospects: The preservation of comparative statics results. *Rev. Econom. Stud.* **54**, 73–85.

JEWITT, I. (1988). Justifying the first-order approach to principal-agent problems. *Econometrica* **56**, 1177–1190.

KARLIN, S. (1968). *Total Positivity.* Stanford University Press.

KARLIN, S. and RINOTT, Y. (1980). Classes of orderings of measures and related correlation inequalities, I. Multivariate totally positive distributions. *J. Multivariate Anal.* **10**, 467–498.

LANDSBERGER, M. and MEILIJSON, I. (1989). Mean preserving portfolio dominance. Mimeo, University of Haifa.

MILGROM, P. R. (1981) Good news and bad news: Representation theorems and applications. *Bell J. Econom.* **12**, 380–391.

MILGROM, P. R. and WEBER, R. J. (1982). A theory of auctions and competitive bidding. *Econometrica* **50**, 1089–1122.

MIRRLEES, J. (1976). The optimal structure of authority and incentives within an organization. *Bell J. Econom.* **7**, 105–131.

PRATT, J. W. (1964). Risk aversion in the small and in the large. *Econometrica* **32**, 122–136.

VICKREY, W. (1961). Counter-speculation, auctions and competitive sealed tenders. *J. Finance* **16**, 305–321.

Whitt, W. (1985). Uniform conditional variability ordering of probability distributions. *J. Appl. Probab.* **22**, 619–633.

Department of Economics
University of Bristol
40 Berkeley Square
Bristol, BS8 1HY, United Kingdom