Co-monotone allocations, Bickel–Lehmann dispersion and the Arrow–Pratt measure of risk aversion

Michael Landsberger

Department of Economics, University of Haifa, Haifa, Israel

Isaac Meilijson

Raymond and Beverly Sackler Faculty of Exact Sciences, School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel

For every integrable allocation $(X_1, X_2, \ldots, X_n)$ of a random endowment $Y = \sum_{i=1}^n X_i$ among n agents, there is another allocation $(X_1^*, X_2^*, \ldots, X_n^*)$ such that for every $1 \le i \le n$, X_i^* is a nondecreasing function of Y (or, $(X_1^*, X_2^*, \ldots, X_n^*)$ are *co-monotone*) and X_i^* dominates X_i by Second Degree Dominance.

If $(X_1^*, X_2^*, \ldots, X_n^*)$ is a co-monotone allocation of $Y = \sum_{i=1}^n X_i^*$, then for every $1 \le i \le n$, Y is more dispersed than X_i^* in the sense of the Bickel and Lehmann stochastic order.

To illustrate the potential use of this concept in economics, consider insurance markets. It follows that unless the uninsured position is Bickel and Lehmann more dispersed than the insured position, the existing contract can be improved so as to raise the expected utility of both parties, regardless of their (concave) utility functions.

Keywords: Co-monotonicity, Bickel–Lehmann dispersion.

1. Introduction

It is well known that if an allocation $X = (X_1, X_2, \ldots, X_n)$ of a random endowment Y among n agents with prespecified nondecreasing and concave utility functions is Pareto optimal, then X is *co-monotone*, i.e., each X_i is a nondecreasing function of $Y = \sum_{i=1}^n X_i$ (see Borch [5] and Wilson [21]). It follows that for every allocation which is not co-monotone and every n-tuple $(U_1, U_2, \ldots, U_n)$ of concave utilities there is a co-monotone allocation X^* of $Y = \sum X_i = \sum X_i^*$ such that $EU_i(X_i^*) \ge EU_i(X_i)$ for all $1 \le i \le n$.

The shortcoming of this important result, applied by many authors in the context of risk sharing programs, is the dependence of X^* on the specific utilities $(U_i)_{i=1}^n$, since in practice utilities are hardly ever known. Insurance companies (see Raviv [14]), other entrepreneurs (see Pratt and Zeckhauser [13]) or designers of financial assets who wish to reach wide markets with a small number of financial instruments may find little comfort in knowing how to design securities or other risk sharing contracts for agents with prespecified utility functions. It is only natural to investigate whether there exist allocations which are preferred to other allocations by *all* risk averse agents.

Being able to extend the properties of efficient allocations without the need to restrict attention to prespecified utility functions should be of significant importance. For example, financial instruments such as contingent claims may be contingent on primitives inducing income streams which are not co-monotone. Should such a situation be discerned, "co-monotonizing" contingent claims should be desired by all risk averse agents who hold the primitive assets. In real life, it may not be easy to identify such situations. However, if these arbitrage opportunities are identified, we should expect that income streams held by agents be co-monotone. This conclusion does not follow from "utility specific" co-monotonicity.

We prove by a constructive algorithm (proposition 1(i)) that for every allocation X which is not co-monotone there exists a co-monotone one, X^*, such that $EU(X_i^*) \geq EU(X_i)$ for all $1 \leq i \leq n$ and all concave, nondecreasing U. We should add that for prespecified utilities not all co-monotone allocations are Pareto optimal but, as we further prove, there is no co-monotone allocation to which some other allocation is preferred by all risk averters (proposition 1(ii)).

The concept of *co-monotonicity* has been applied earlier. Schmeidler [18, 19] established the concept of *co-monotonic independence* which he applied in the context of decision making under uncertainty. Yaari [22] applied it in the context of designing the Dual Theory. Wakker [20] used it to characterize the concepts of Pessimism and Optimism. The concept of co-monotonicity applied in this paper is very similar to the one used by Yaari. The fact that variants of what is basically the same concept were successfully applied in providing characterizations of different theoretical problems can serve as an indication of its potential use in other, as yet unexplored, problems.

Co-monotone allocations induce an interesting stochastic dominance relation between the total amount Y to be allocated and the amount X_i allocated to the ith agent: the latter is less dispersed in the sense of the Bickel–Lehmann (Bickel and Lehmann [3, 4]) dispersion order. This order, well known to statisticians, seems not to have been applied yet in economics. As an illustration of a possible application of this order note that as long as the marginal income tax is between 0 and 1, net income and tax revenues are co-monotone. Under these conditions, net income is Bickel–Lehmann less dispersed than gross income. This order is stronger than Second Degree Dominance, whose relation with the elasticity of income after tax with respect to income before tax has been investigated by Jacobsson [8].

We note that the Bickel–Lehmann dispersion order is also useful in analyzing behavior under risk. Ross [15] raised the issue that in environments where full insurance is not available the Arrow–Pratt measure or risk aversion (Arrow [1], Pratt [12]) does not necessarily yield the nice and intuitively appealing results rendered in situations where all risk can be eliminated. The reductions in risk considered by Ross and others have the Rothschild–Stiglitz structure (Mean Preserving Increases in Risk, Rothschild and Stiglitz [16]). In an insurance setting, this structure is induced by allocations between insurer and insured in which the amount allocated to the insurer is "noise" with respect to the amount allocated to the insured. We prove elsewhere (see Landsberger and Meilijson [11]) that the Bickel–Lehmann reductions in risk are the ones compatible with the Arrow–Pratt index in the class of all nondecreasing utilities.

This result extends results by Jewitt [9] who developed an order, weaker than Bickel–Lehmann Dispersion and stronger than Mean Preserving Increase in Risk, compatible with the Arrow–Pratt index in the class of all concave nondecreasing utilities.

2. Co-monotone allocations

We use throughout the paper the terms *allocation*, *dominance*, *co-monotonicity*. We define these terms formally and summarize some of their basic properties.

An *allocation* of an integrable random variable Y (defined on some probability space (Ω, B, P)) between n agents is an n-tuple $X = (X_1, X_2, \dots, X_n)$ of integrable random variables on that space, such that $\sum_{i=1}^{n} X_i = Y$ almost surely.

The allocation $X^* = (X_1^*, X_2^*, \dots, X_n^*)$ of Y *dominates* X if every risk averse agent (weakly) prefers X_i^* to X_i, for all $1 \le i \le n$. It is easy to see that if X^* dominates X, then $E(X_i^*) = E(X_i)$ for all $1 \le i \le n$.

Random variables $(X_1, X_2, \dots, X_n)$ are *co-monotone* if any of the following equivalent conditions hold. For proofs, see Dellacherie [6].

(i) For all $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$,

$$P(X_1 \le x_1, X_2 \le x_2, \dots, X_n \le x_n) = \min_{1 \le i \le n} P(X_i \le x_i).$$

(ii) There exist non-decreasing functions $f_i : \mathbb{R} \to \mathbb{R}$, $1 \le i \le n$, and a random variable W such that for all $1 \le i \le n$, $X_i = f_i(W)$ almost surely.

(iii) There exist non-decreasing functions $f_i : \mathbb{R} \to \mathbb{R}$, $1 \le i \le n$, such that for all $1 \le i \le n$, $X_i = f_i(\sum_{j=1}^{n} X_j)$ almost surely.

Condition (i) implies that there is one and only one co-monotone joint distribution with given marginal distributions $F_1, F_2, \dots, F_n$. To build co-monotone random variables with these marginals, let W be uniformly distributed in the unit interval and define, for $1 \le i \le n$, $X_i = F_i^{-1}(W)$. This canonical construction of co-monotone random variables with given marginals is one of those postulated by condition (ii). Condition (iii) stipulates that a co-monotone allocation assigns to each agent an amount which is a nondecreasing function of the total amount to be allocated.

As mentioned in the introduction, Pareto optimal allocations between agents with specific concave utilities are well known to be co-monotone. Thus, for every allocation X and every n-tuple $\{U_i\}_{i=1}^n$ of concave utility functions, there is a co-monotone allocation X^* (typically dependent on $\{U_i\}$) which every agent prefers. This does not imply the existence of a co-monotone X^* which dominates X. The following proposition claims that such an X^* exists and its proof shows how to construct it.

PROPOSITION 1

(i) Every allocation is dominated by some co-monotone allocation.

(ii) Given two co-monotone allocations (X_1, X_2) and (X'_1, X'_2) of $Y = X_1 + X_2 = X'_1 + X'_2$, if every risk averter prefers X_1 to X'_1 then every risk averter prefers X'_2 to X_2.

Proof

(i) We first show that every allocation X of a random asset Y is dominated by another, which allocates to each agent an amount determined by $Y = \sum_{j=1}^n X_j$. Plainly, the dominating allocation may be taken to be $X' = (X'_1, X'_2, \ldots, X'_n)$, with $X'_i = E(X_i | Y)$: X' is clearly an allocation, and every risk averter gladly gives up any integrable random variable in favor of its conditional expectation given any other random variable. Hence, without loss of generality, we may assume that all coordinates of the allocation X are functions of Y. In what follows we restrict the proof to the case $n = 2$; the generalization to $n > 2$ is straightforward. We further restrict the proof to Y supported by a finite set. The limiting argument for the general case is omitted. Let $y_1 > y_2 > \ldots > y_m$, with probabilities $p_1 > 0$, $p_2 > 0, \ldots, p_m > 0$, be the values of Y, and let $(x_1, z_1), (x_2, z_2), \ldots, (x_m, z_m)$, with $x_i + z_i = y_i$ for all $1 \le i \le m$ be the allocation X. If X is not co-monotone then for some $1 \le i \le m$, $x_1 \ge x_2 \ge \ldots \ge x_i$ and $z_1 \ge z_2 \ge \ldots \ge z_i$ but the point (x_{i+1}, z_{i+1}) does not belong to the Southwestern quadrant $\{(x, z) | x < x_i, z < z_i\}$ (see figure 1). The line with slope -1 through (x_k, z_k) is "fiber" k. Since (x_{i+1}, z_{i+1}) is not in the above quadrant, it can be in sections I or III of fiber $i + 1$. Without loss of generality, let it be in section I. There is a minimal $1 \le j \le i$ such that $z_{i+1} > z_j$. Slide (x_{i+1}, z_{i+1}) by an amount α in the South-eastern direction and slide $(x_i, z_i), (x_{i-1}, z_{i-1}), \ldots, (x_j, z_j)$ by an amount β in the Northwestern direction (with α and β related so as to preserve the expectations of both coordinates of the allocation) until the points in fibers $i + 1$ and j have the same z-value. The new allocation dominates the old one because each of the two coordinates has experienced a decrease in risk of the Diamond and Stiglitz [7] single-crossing type. It should now be clear that this algorithm will

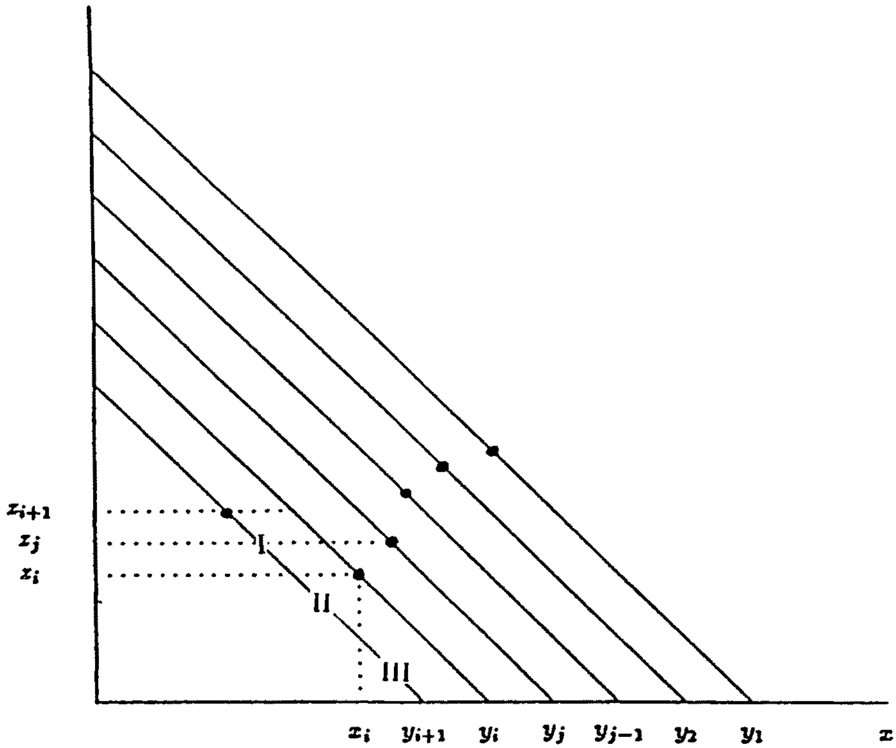

Fig. 1.

terminate with a co-monotone allocation in at most $m(m-1)/2$ sliding operations of this kind. For the sake of completeness and programming, in the general step described above, x_{i+1} moves to the right by $(z_{i+1}-z_j)\sum_j^i p_k/\sum_j^{i+1} p_k$ and each x_k, $j \le k \le i$, moves to the left by $(z_{i+1}-z_j)p_{i+1}/\sum_j^{i+1} p_k$.

(ii) Assume that X_1 dominates X'_1 by Second Degree Stochastic Dominance.

The proof relies on an alternative characterization of SSD, in terms of Lorenz curves (see Atkinson [2]) according to which F dominates G if and only if for all $u \in (0,1)$, $\int_0^u G^{-1}(v)\,\mathrm{d}v \le \int_0^u F^{-1}(v)\,\mathrm{d}v$. To see this, observe that $(1/u)\int_0^u F^{-1}(v)\,\mathrm{d}v$, the conditional expectation of an F-distributed random variable given that it belongs to the lower u-bracket of F, may be geometrically identified as the intercept on the x-axis of the linear function with slope u which is tangent to $\varphi_F(\cdot) = \int_{-\infty}^{\cdot} F(y)\,\mathrm{d}y$. Clearly, pointwise inequality between these intercepts is equivalent to pointwise inequality between the φ-functions, $\varphi_G(t) \ge \varphi_F(t)$, $\forall t$, and the latter is the well-known integral characterization of SSD.

Since the joint distribution of co-monotone random variables (X_1, X_2) is equal to that of $F_{X_1}^{-1}(U)$, $F_{X_2}^{-1}(U)$ for $U \sim U[0,1]$, if $Y = X_1 + X_2$ then

$F_Y^{-1} \equiv F_{X_1}^{-1} + F_{X_2}^{-1}$ because the equality $F_Y^{-1}(U) = F_{X_1}^{-1}(U) + F_{X_2}^{-1}(U)$ holds almost everywhere and these three inverse functions are right-continuous. We thus get that $F_{X_1}^{-1} + F_{X_2}^{-1} = F_Y^{-1} = F_{X_1'}^{-1} + F_{X_2'}^{-1}$, from which it follows that

$$\int_0^u F_{X_2}^{-1}(v)\,\mathrm{d}v - \int_0^u F_{X_2'}^{-1}(v)\,\mathrm{d}v = -\left[\int_0^u F_{X_1}^{-1}(v)\,\mathrm{d}v - \int_0^u F_{X_1'}^{-1}(v)\,\mathrm{d}v\right] \geq 0,$$

which completes the proof. □

As far as the economic literature is concerned, Atkinson [2] was the first to notice the equivalence between ranking distributions by SSD and the Lorenz criterion. Atkinson elaborated more on the case where the distributions have equal means but it is clear, upon reading his paper (see pp. 246–47), that he was aware that the equivalence holds in the more general case when means are not necessarily equal.

Proposition 1(ii) claims that no co-monotone allocation may be improved for the benefit of all risk averse agents. This does not preclude the existence of co-monotone allocations which are preferred to other co-monotone allocations by some risk averters. Not only are there co-monotone allocations which are not Pareto optimal but in fact for *every* co-monotone allocation there exist $\{U_i\}_{i=1}^n$ for which the allocation is not Pareto optimal. This fact may be inferred directly from the Wilson [21] characterization of Pareto optimality. We provide a direct proof.

LEMMA 1

Let $X = (X_1, X_2, \ldots, X_n)$ be a co-monotone allocation of Y such that some X_i is not constant. Then there exist $\{U_i\}_{i=1}^n$ concave, nondecreasing, and a co-monotone allocation X^* such that $EU_i(X_i^*) > EU_i(X_i)$ for all $1 \leq i \leq n$.

Proof

Suppose X_1 is not constant. Let U_1 be any strictly concave utility. Then the first agent strictly prefers $(E(X_1) + X_1)/2$ to X_1. Hence, for some $\epsilon > 0$, s/he strictly prefers $X_1^* = -\epsilon + (E(X_1) + X_1)/2$ to X_1. Split the excess $X_1 - X_1^*$ (whose mean is $\epsilon > 0$) evenly between all other agents. If these are risk neutral or nearly so, they will strictly prefer the new amounts $X_i^* = X_i + (X_1 - X_1^*)/(n-1)$ allocated to them. □

The following example contrasts co-monotone allocations with those which eliminate "noise".

Table 1

Situation	Stock pays $2 house intact	Stock pays zero house intact	Stock pays $2 house – total loss	Stock pays zero house – total loss
Uninsured wealth position Y	13	11	3	1
Probability	0.45	0.45	0.05	0.05

EXAMPLE 1

Consider a risk averse agent who owns a house worth $10. The house is susceptible to a total loss with probability 1/10. The agent also owns $1 in cash and a source of random income (stock) which pays $2 or nothing with probabilities 1/2 each. Uninsured wealth positions are summarized in table 1.

Assume that fair insurance of the house is available but uncertainty associated with capital (stock) income is retained. Under fair insurance, the wealth position X of the insured is $10 or $12 with equal probabilities and the insurer gets Z which takes the values -9 with probability 1/10 and 1 with probability 9/10. X and Z are independent and $X+Z=Y$. Risk averse agents welcome such a program, which eliminates one source of risk.

Wealth positions of the insured and the insurer are given in table 2.

Table 2

Situation	Stock pays $2 house intact	Stock pays zero house intact	Stock pays $2 house – total loss	Stock pays zero house – total loss
Insured's wealth position X	12	10	12	10
Insurer's wealth position Z	1	1	−9	−9
Probability	0.45	0.45	0.05	0.05

Table 3

Situation	Stock pays $2 house intact	Stock pays zero house intact	Stock pays $2 house – total loss	Stock pays zero house – total loss
Insured's wealth position X^*	12	10.2	10.2	10
Insurer's wealth position Z^*	1	0.80	−7.2	−9
Probability	0.45	0.45	0.05	0.05

The allocation (X, Z) is obviously not co-monotone; therefore, it can be improved to the benefit of all risk averse agents. To achieve a co-monotone reduction in risk proceed as in the proof of proposition 1(i). The new insurance program given in table 3, is co-monotone and dominates the one presented in table 2.

3. The Bickel–Lehmann dispersion order

DEFINITION (Bickel and Lehmann [3, 4])

The distribution F is less dispersed than the distribution G if for every $0 < u < v < 1$

$$F^{-1}(v) - F^{-1}(u) \leq G^{-1}(v) - G^{-1}(u). \tag{1}$$

This inequality says that the interval between the u-quantile and the v-quantile of G is at least as long as the corresponding interval for F. Inequality (1) can be rewritten as

$$G^{-1}(v) - F^{-1}(v) \geq G^{-1}(u) - F^{-1}(u). \tag{2}$$

which means that F is less dispersed than G if and only if $G^{-1} - F^{-1}$ is a non-decreasing function on (0, 1).

A useful implication of the above is that for every real c, $F(x-c)$ and $G(x)$ cross at most once and if they cross, $F(x-c)$ lies below G to the left of the crossing. To see this, note that the horizontal distance between $F(\cdot - c)$ and $G(\cdot)$ differs from the horizontal distance $G^{-1} - F^{-1}$ between F and G, by the constant c. Hence, horizontal distances are monotone for all c or for none.

In particular, if $-\infty < \int x \, \mathrm{d}G(x) \leq \int x \, \mathrm{d}F(x) < \infty$ and F is less dispersed than G, then F dominates G by Second Degree Stochastic Dominance, or, every risk averter prefers F to G. This is so, because the single-crossing property between mean-ordered distributions implies second degree dominance, as proved by Diamond and Stiglitz [7]. The following proposition, which somewhat generalizes theorem 1 in Bickel and Lehmann [4], identifies as dispersion the relation between the distribution of a random variable co-monotonically allocated to a number of agents, and the distribution of the amount allocated to any one of them.

PROPOSITION 2

A distribution F is less dispersed than a distribution G if and only if there exist on some probability space two co-monotone random variables X and Z such that $X \sim F$ and $X + Z \sim G$.

Proof

Suppose that F is less dispersed than G. Then, by virtue of (2), $G^{-1} - F^{-1}$ is nondecreasing, as is F^{-1}. Hence, $X = F^{-1}(U)$ and $Z = G^{-1}(U) - F^{-1}(U)$ are co-monotone. To prove the other direction let X and Z be co-monotone, $X \sim F$ and $Z \sim H$, and let $Y = X + Z$. Since co-monotonicity uniquely determines the joint distribution of X and Z which in turn uniquely determines the distribution of $X + Z$, we may assume that X and Z are realized as $X = F^{-1}(U)$ and $Z = H^{-1}(U)$ for $U \sim U[0, 1]$. Now, letting $Y = X + Z = F^{-1}(U) + H^{-1}(U)$, if we denote $F^{-1}(U) + H^{-1}(U)$ by $T(U)$, then $T(U) \sim G$. But the essentially only nondecreasing function on $(0, 1)$ with a prescribed distribution is the inverse of the distribution. Hence, $T = G^{-1}$ a.e. and $G^{-1} - F^{-1} = T - F^{-1} = H^{-1}$ is non-decreasing, which, by (2), implies that F is less dispersed. □

Proposition 2 characterizes Bickel–Lehmann dispersions as the addition of a co-monotone variable. This is close in spirit to the characterization of Rothschild–Stiglitz Increase in Risk as the addition of "noise". We provided in the introduction a conceptual illustration of Bickel–Lehmann dispersion, the relation between gross and net income. We now provide a technical illustration, which appears in the economic literature as an illustration of Rothschild–Stiglitz Increase in Risk.

EXAMPLE 2: "Stretches" of distributions are more dispersed.

G is a stretch of F (Arrow [1], Sandmo [17]) if there exist $\alpha > 1$ and x such that for $X \sim F$, $x + \alpha(X - x) \sim G$. Since $x + \alpha(X - x) = X + (\alpha - 1)(X - x)$ is a sum of two co-monotone random variables one of which is X, G is more dispersed than F.

In particular, normal, exponential and uniform families of distributions are dispersion-ordered by their variances.

We remark that the proof of the lemma in section 2 was based on a stretch.

Another important use of the Bickel–Lehmann order is in characterizing behavior under risk. As noted by Ross [15] and Kihlstrom et al. [10], under partial insurance the Arrow–Pratt measure of risk aversion does not assure that more risk averse agents are willing to pay more for the elimination of some risk. The authors proved in Landsberger and Meilijson [11] that this shortcoming of the Arrow–Pratt measure does not apply to Bickel–Lehmann risk reductions, namely, in the class of all agents with nondecreasing utility functions, more-risk-averse agents pay a higher (not necessarily non-negative) risk premium for a reduction in risk if and only if the less risky distribution is Bickel–Lehmann less dispersed. Jewitt [9] established a risk concept, *Location Independent Risk*, and proved that agents with concave, nondecreasing utility functions who are more risk averse, are willing to pay a higher premium for a reduction in this risk.

Acknowledgements

We are grateful to D. Kreps, M. Machina, D. Schmeidler and a referee for helpful comments and constructive suggestions.

References

[1] K. Arrow, The theory of risk aversion, in: *Essays in the Theory of Risk Bearing*, ed. K. Arrow (North-Holland, Amsterdam, 1970).
[2] A.B. Atkinson, On the measurement of inequality, J. Econ. Theory (1970) 244–263.
[3] P.J. Bickel and E.L. Lehmann, Descriptive statistics for nonparametric models, III. Dispersion, Ann. Statist. 4 (1976) 1139–1158.
[4] P.J. Bickel and E.L. Lehmann, Descriptive statistics for nonparametric models, IV. Spread. in: *Contributions to Statistics*, ed. Jureckova (Reidel, 1979).
[5] K. Borch, Equilibrium in reinsurance markets, Econometrica 30 (1962) 424–444.
[6] C. Dellacherie, Quelques commentaires sur les prolongements de capacités, *Séminaire de Probabilités 1969/70*, Strasbourg, L.N. in Math. 191 (Springer, 1970) pp. 77–81.
[7] P.A. Diamond and J.E. Stiglitz, Increases in risk and in risk aversion, J. Econ. Theory, 9 (1974) 337–360.
[8] M. Jacobsson, On the measurement of the degree of progression, J. Public Econ. 5 (1976) 161–168.
[9] I. Jewitt, Choosing between risky prospects: the characterization of comparative statics results, and location independent risk, Manag. Sci. 35 (1989) 60–70.
[10] R.E. Kihlstrom, D. Romer and S. Williams, Risk aversion with random initial wealth, Econometrica 49 (1981) 911–920.
[11] M. Landsberger and I. Meilijson, The generating process and an extension of Jewitt's location independent risk concept, Manag. Sci. (1993), to appear.
[12] J.W. Pratt, Risk aversion in the small and in the large, Econometrica 32 (1964) 122–136.
[13] J.W. Pratt and R. Zeckhauser, The impact of risk sharing on efficient decision, J. Risk and Uncertainty 2 (1989) 219–223.
[14] A. Raviv, The design of an optimal insurance policy, Amer. Econ. Rev. 69 (1979) 84–96.
[15] S.A. Ross, Some stronger measures of risk aversion in the small and the large with applications, Econometrica 49 (1981) 621–663.
[16] M. Rothschild and J.E. Stiglitz, Increasing risk I: A definition, J. Econ. Theory 2 (1970) 337–360.
[17] A. Sandmo, Competitive firm under price uncertainty, Amer. Econ. Rev. 61 (1971) 65–73.
[18] D. Schmeidler, Integral representation without additivity, Proc. AMS 97 (1982) 255–261.
[19] D. Schmeidler, Subjective probability and expected utility without additivity, Econometrica 57 (1989) 571–587.
[20] P. Wakker, Characterizing optimism and pessimism directly through comonotonicity, J. Econ. Theory 52 (1990) 453–463.
[21] R. Wilson, The theory of syndicates, Econometrica 36 (1968) 119–132.
[22] M.E. Yaari, The dual theory of choice under risk, Econometrica 55 (1985) 95–115.