



Four notions of mean preserving increase in risk, risk attitudes and applications to the Rank-Dependent Expected Utility model

Alain Chateauneuf, CERMSEM, Université de Paris I,

Michèle Cohen, EUREQua, Université de Paris I,

Isaac Meilijson, School of Mathematical Sciences, Tel Aviv University

Abstract

Abstract : This article presents various notions of risk generated by the intuitively appealing single-crossing operations between distribution functions. These stochastic orders, Bickel & Lehmann dispersion or (its equal-mean version) Quiggin's monotone mean-preserving increase in risk and Jewitt's location-independent risk, have proved to be useful in the study of Pareto allocations, ordering of insurance premia and other applications in the Expected Utility setup. These notions of risk are also relevant to the Quiggin-Yaari Rank-dependent Expected Utility (RDEU) model of choice among lotteries. Risk aversion is modeled in the vNM Expected Utility model by Rothschild & Stiglitz's Mean Preserving Increase in Risk (MPIR). Realizing that in the broader rank-dependent set-up this order is too weak to classify choice, Quiggin developed the stronger monotone MPIR for this purpose.

This paper reviews four notions of mean-preserving increase in risk - MPIR, monotone MPIR and two versions of location-independent risk (renamed here left and right monotone MPIR) - and shows which choice questions are consistently modeled by each of these four orders.

Keywords : Location-independent risk, monotone increase in risk, rank-dependent expected utility.

JEL classification: D81

Acknowledgements : We would like to thank Jan Werner for his helpful comments.

1 Introduction

Second degree Stochastic Dominance (SSD) and its equal-mean version Mean-Preserving Increase in Risk (MPIR) play a leading role in the modeling of increasing risk in Economics, just as First degree Stochastic Dominance (FSD) plays a leading role in modeling increasing random wealth. None of these orders provides clear-cut answers to all questions of interest that have been posed, but they do provide a general framework in which these questions can be studied by developing more stringent types of FSD or SSD.

For example (see Rothschild & Stiglitz [36]), if a portfolio based on one safe and one risky asset is subjected to a MPIR of the risky asset, expected utility of the optimal portfolio decreases, but the fraction invested optimally in the risky asset does not necessarily decrease, as intuition would dictate. There is a vast literature dedicated to the discovery of special types of MPIR or special types of concave utility that coherently rank these fractions.

As a second example (see Ross [34]), an expected utility maximizer will always be willing to buy full insurance accepted by a less risk averse one (this is in fact the Pratt postulate for *less risk averse*), but this is not necessarily so for partial insurance that reduces risk in the MPIR sense, as, again, intuition would suggest. Some stronger forms of MPIR (Bickel & Lehmann [2, 3], Jewitt [19], Landsberger & Meilijson [24, 25]) perform coherently on this problem, as sketched a few paragraphs below.

These stochastic orders, some of their properties and some of their roles in the Expected Utility (EU) and in the Rank-dependent Expected Utility (RDEU) setups, are the subject matter of this paper.

The common link of the orders under study is that they are generated as successive applications of mean-preserving or mean-reducing *single crossing* operations between distribution functions: as shown by Diamond & Stiglitz [13], SSD is the *uniformly integrable transitive closure* of this simple, intuitive operation. Why intuitive? Because FSD is more intuitive than SSD, and single crossing means that for some $0 < p < 1$, the restrictions of F and G to a common lower (upper) p -quantile make F dominate G (G dominate F) by FSD: mass on the left (right) is transferred further to the left (right).

- **Single crossing.** G *single crosses* F if for some $x_0 \in [-\infty, \infty]$, $G(x) \geq F(x)$ for $x < x_0$ and $G(x) \leq F(x)$ for $x > x_0$. Equivalently, if the horizontal distance $G^{-1}(v) - F^{-1}(v)$ between F and G , as a function on $(0, 1)$, is non-positive below some $v_0 \in [0, 1]$ and non-negative thereafter. As shown by Diamond & Stiglitz [13] and Machina & Pratt [28], if applied in a mean-preserving way, this type of stretch generates mean-preserving increase in risk (MPIR). If applied in a mean-preserving or mean-reducing way, it generates second degree dominance.
- **Monotone single crossing.** G is a monotone single crossing of F if the horizontal distance $G^{-1}(v) - F^{-1}(v)$ between F and G is a non-decreasing function on the interval $(0, 1)$. Equivalently, if G single crosses every horizontal shift of F . This is the notion

of dispersion introduced by Bickel & Lehmann [2, 3] and the notion of monotone increase in risk (MIR) introduced by Quiggin [32] for the mean-preserving case.

- **Left-monotone single crossing.** G is a left-monotone single crossing of F if G is a single crossing of F such that $G^{-1}(v) - F^{-1}(v)$ is non-decreasing in every interval (of the type $(0, v_0)$) where it is non-positive. Equivalently, if G single crosses every horizontal shift of F towards the left. As shown by Landsberger & Meilijson [25], if applied in a mean-preserving way this type of stretch generates left-monotone increase in risk (LMIR), that Jewitt defined under the name of *location-independent risk*, without the requirement of mean preservation.
- **Right-monotone single crossing.** G is a right-monotone single crossing of F if G is a single crossing of F such that $G^{-1}(v) - F^{-1}(v)$ is non-decreasing in every interval (of the type $(v_0, 1)$) where it is non-negative. Equivalently, if G single crosses every horizontal shift of F towards the right. This mirror image of Jewitt's order generator gives rise to a similar stochastic order, dual location-independent risk or (under equal means) right-monotone increase in risk. (Y is a right-monotone increase in risk (RMIR) with respect to X iff $-Y$ is a left-monotone increase in risk with respect to $-X$.)

Marshall & Olkin [29] and Moyes [30] have developed the Relative Differentials Ordering (RDO) as a relative measure of income inequality: the logarithms of RDO-ordered random variables are Bickel & Lehmann ordered (one is a monotone single crossing of the other). RDO is a special case of dual location-independent risk (Chateauneuf & Moyes [9]). More generally, we prove in Theorem 1 that if Y is a right-monotone (left-monotone) single crossing of X then $g(Y)$ is a right-monotone (left-monotone) single crossing of $g(X)$ for every g that is increasing and convex (concave).

Monotone single crossing is a transitive relation, so it defines directly a stochastic order. The other three are not, so they generate stochastic orders by uniformly integrable transitive closure. What this means is that if F dominates G (by SSD, MPIR, MIR, LMIR or RMIR) then G is the limit of a uniformly integrable sequence of distributions starting with F and converging to G , such that each is obtained from the previous one by the corresponding single-crossing operation (applied mean preservedly or mean reducingly in the case of SSD and mean preservedly for each of the other four orders), and conversely. Note: uniform integrability is automatic if F , G and all interim distributions are supported by a finite interval.

Single crossings of horizontal shifts can play a decisive role in Insurance, as they order premia for partial insurance by the Arrow-Pratt index: let the uninsured position Y be replaced by an insured position $X - C$. If their distributions single cross, so will those of $u(Y)$ and $u(X - C)$ for non-decreasing u . If the insurance contract is welcome under u , second order stochastic dominance yields the same preference inequality between

$E\eta(u(Y))$ and $E\eta(u(X - C))$ for η non-decreasing and concave, that is, for a utility function $v(\cdot) = \eta(u(\cdot))$ Arrow-Pratt more risk averse than u .

As pointed out earlier (Ross [34]), regular single crossing is not enough: even if G single crosses F , it need not single cross all horizontal shifts of F . Thus, unlike MPIR, Bickel & Lehmann dispersion is well suited to the EU study of insurance premia under general non-decreasing utilities (where the premium C may be positive or negative), while Jewitt's location-independent risk fits the needs under concave utilities, since only non-negative premia C are extracted from EU risk averters.

So far we have been working within the expected utility (EU) model, under which a decision maker is characterized by a utility function u . This non-decreasing function, often assumed to be continuous, becomes also concave when the decision maker is postulated to be averse to risk. Under the rank-dependent expected utility (RDEU) model, a decision maker is characterized by two functions, a utility function u (that plays the role of *utility under certainty*) in conjunction with a *probability-transformation* function $f : [0, 1] \rightarrow [0, 1]$ that is non-decreasing and satisfies $f(0) = 0, f(1) = 1$. Such a decision maker prefers a random variable X to another, Y , if and only if $V(X) > V(Y)$, where the *Quiggin-Yaari functional* V (see Quiggin [32], Yaari [42]) is given by

$$V(Z) = V_{u,f}(Z) = - \int_{-\infty}^{\infty} u(x) df(P(Z > x)) \quad (1)$$

A winning probability p is “perceived” by the decision maker as $f(p)$, in the sense that expected utility is computed as if the winning probability was $f(p)$ rather than p . If the transformation function f is the identity function $f(p) \equiv p$, then $V(Z) = V_{u,I}(Z)$ is the expected utility $E[u(Z)]$ of the random variable. The *Yaari functional* (see Yaari [42]) is the special case $V(Z) = V_{I,f}(Z)$, in which the utility under certainty is the identity function $u(x) \equiv x$. If both transformation and utility are identity functions, then $V(Z) = V_{I,I}(Z)$ is simply the expected value $E[Z]$ of the random variable.

In the EU model, aversion to risk is a rather clear-cut notion: The decision maker always prefers the expectation of a random variable to the random variable itself. This property is fully characterized by concavity of the utility function u . The question can now be reversed as to what are the pairs (X, Y) of random variables with the same finite mean for which every risk averse decision maker prefers X to Y , leading to the characterization of MPIR.

The situation in the RDEU model is less clear-cut due to the interaction of the utility function and the probability transformation function in implying risk aversion. The universal preference of expected value to the random variable itself for all concave utilities u is modeled by transformation functions f that display *pessimism*, i.e., are majorized by the identity function (Yaari [42], Quiggin [32, 33]). However, the universal preference (under concave utilities) for mean-preserving decreases in risk requires the stricter condition that f display *increasing pessimism*, i.e., that f be convex (Chew, Karni & Safra [10]). Quiggin [33] showed that the (stronger) Bickel & Lehmann [2, 3] dispersion order, or equivalently Quiggin's [32, 33] monotone increase in risk, is the weakest order that coherently ranks

the preferences (between pairs of random variables with the same finite mean) of decision makers with concave utility function and probability transformation function majorized by the identity function.

We will rederive these two classes of probability transformation function, pessimism and increasing pessimism, ordered by inclusion, and derive two intermediate classes, the *star-shaped* probability transformation functions (Lemma 4), by postulating aversion to different types of risk. These two classes display *increasing relative pessimism* with respect to a sure loss or with respect to a sure gain.

Diverse attitudes to different types of MPIR lie at the core of the departure of the RDEU model from the EU model. According to the von-Neumann & Morgenstern Independence Axiom (see [39]), a decision maker that prefers a random variable X^* to a random variable Y^* will also prefer X (a mixture of X^* and another random variable W by the flip of some coin) to Y (the mixture of Y^* and W by the flip of the same coin), whatever be the distributions of the outcome of the coin and of the random variable W .

Consider two equal-mean dichotomous random variables X^* , with values $x_1 < x_2$, and Y^* , with values $y_1 \leq x_1$ and $y_2 \geq x_2$. The (risk averse) preference of X^* to Y^* is quite natural. The question of (von-Neumann & Morgenstern) Independence axiom consistency is whether this preference is so sharp that it is preserved even when the spread pair (x_1, x_2) is only part of the distribution - the other part, left unchanged, being represented by W . In the absence of this powerful axiom, it need not be. The attitude of RDEU decision makers may vary according to the relative location of the other atoms of W in relation to the values (x_1, x_2) . As will be seen in Lemma 4, if W is supported by the interval $[x_1, x_2]$ then RDEU preference for X over Y (i.e., aversion to spread of extremes) only requires f to be majorized by the identity function while if W is allowed to have an arbitrary atom on each of the two rays $(-\infty, x_1)$ and (x_2, ∞) , the global preference of X to Y (aversion to middle spread) requires f to be convex. The presence of an arbitrary atom of W in $(-\infty, x_1)$ forces f to be *star shaped at 0*, the model for aversion to spread of gains, while its presence in (x_2, ∞) forces f to be star-shaped at 1, modeling aversion to spread of losses. (g is said to be star-shaped at μ if $(g(x) - g(\mu))/(x - \mu)$ is monotone (non-decreasing, in this application) on $\mathcal{R} - \{\mu\}$).

We thus have four conditions on the probability transformation function that model aversion to particular types of mean-preserving spread in the RDEU set-up, by DM's with concave utility function. These types of increase in risk, so far illustrated only by some simple canonical scenarios, correspond to the four types of single-crossing operation introduced earlier (Theorems 2, 3, 4 and 5.)

The first part of this paper (Section 2) is dedicated to the four stochastic orders under study. The second part (composed of sections 3 and 4) focuses on the probability transformation function of RDEU DM's, postulating their utility functions to be concave. The authors study in [7] a departure from concavity of u , and characterize the pairs (u, f) of RDEU agents averse to monotone increase in risk: as long as f displays strong enough

pessimism in some sense, u can even be convex.

To stay at the conceptual level, we choose to avoid heavy measure-theoretic machinery: formal treatment will be restricted to random variables with compact support and to continuous probability transformation functions.

2 An introduction to the four relevant stochastic orders

This section summarizes properties of mean-preserving increase in risk MPIR (see Hardy, Littlewood & Polya [17], Strassen [37], Rothschild & Stiglitz [35], Diamond & Stiglitz [13], Lorenz [27]), monotone increase in risk (see Bickel & Lehmann [2, 3], Quiggin [32], Cohen [12]), left-monotone increase in risk (or location-independent risk, see Jewitt [19], Landsberger & Meilijson [25]) and a dual form of the latter, right-monotone increase in risk. For each stochastic order we present, following the summary of its relevant basic properties, an indication of its role as a generalization of one of the canonical spreads (17)-(20).

The focus on attitude to increase in risk by RDEU DM's should leave out of this study distribution changes that do not preserve the mean. However, many of the other applications of the stochastic orders under study benefit from relaxing the assumption of equal means. We will try to accomodate both.

Given a distribution function H , one of the most common ways to define a probability space and on it a random variable Z distributed H , is to let the probability space be the interval $(0, 1)$ endowed with its Borel σ -field and on it a uniformly distributed random variable U . Defining on the unit interval $[0, 1)$ the *quantile function* or generalized inverse H^{-1} of the (not necessarily invertible) distribution function H by

$$H^{-1}(p) = \inf\{x | H(x) > p\} \tag{2}$$

completed as $H^{-1}(1) = \sup\{x | H(x) < 1\}$, let $Z = H^{-1}(U)$. The elementary fact that Z is distributed H will be used repeatedly. A less elementary fact is the uniqueness of H^{-1} for this purpose: if $g : (0, 1) \rightarrow \mathcal{R}$ is non-decreasing and $g(U) \sim H$, then $g(U) = H^{-1}(U)$ a.s. This construction of a random variable with given distribution function can be extended to the construction of a pair of comonotone random variables (i.e., non-decreasing functions of some other random variable) with given marginal distributions F and G as $F^{-1}(U), G^{-1}(U)$, and to that of a corresponding pair of anti-thetic random variables (hedged against each other) as $F^{-1}(U), G^{-1}(1 - U)$.

Define two integrated distribution functions

$$\Psi_H(x) = \int_x^\infty (1 - H(t))dt = E[\max(Z - x, 0)] \tag{3}$$

$$\Phi_H(x) = \int_{-\infty}^x H(t)dt = \Psi_H(x) + x - E[Z] . \tag{4}$$

Each of these functions is convex and non-negative, with slope never steeper than 1. Φ increases from zero to become asymptotic at ∞ to the linear function $x - E[Z]$ and Ψ decreases to zero from being asymptotic towards $-\infty$ to $E[Z] - x$.

2.1 Second-degree dominance and mean-preserving increase in risk

This order, first studied by Hardy, Littlewood and Polya [17] and generalized by Strassen [37], has been well known in Economics since Rothschild & Stiglitz [35]. It was earlier studied by Hadar & Russel [15] and Hanoch & Levy [16] as Second Degree Dominance, the case of possibly unequal means.

Given two distributions F and G with equal finite means, G is said to represent a Mean-preserving Increase in Risk (MPIR) with respect to F if $\Phi_G(x) \geq \Phi_F(x)$ for all $x \in \mathcal{R}$, or equivalently, if $\Psi_G(x) \geq \Psi_F(x)$ for all $x \in \mathcal{R}$. While these two conditions are equivalent in the case of equal means, Second Degree Dominance of G by F is defined by the first condition and Convex Dominance of F by G (common in the Reliability Engineering literature) is defined by the second condition.

Let $p \in (0, 1)$ and consider the linear function with slope p tangent to the convex function Φ . This linear function can be seen to intersect the x -axis at the conditional expectation

$$L_H(p) = E[H^{-1}(U)|U < p] = \frac{\int_0^p H^{-1}(q) dq}{p} \quad (5)$$

of Z on its lower p -quantile. For the case where H is the distribution of a non-negative random variable (income), the function $\frac{pL_H(p)}{E[Z]}$ was introduced by Lorenz [27] as a relative measure of inequality. It is geometrically obvious that Φ_G majorizes Φ_F if and only if L_F majorizes L_G . This equivalent characterization of Second Degree Dominance is the well known Lorenz-function characterization of MPIR.

Besides these integral characterizations, MPIR has also been characterized in terms of a *noise* or Martingale condition (the existence of a probability space and on it two random variables X and Y with respective marginal distributions F and G , such that $E(Y|X) = X$ almost surely) and by its relation to (EU) risk aversion (the weak preference of F to G by all EU decision makers with concave utility functions).

It follows from the work of Diamond & Stiglitz [13] and Machina & Pratt [28] that for two distribution functions F and G with bounded support and equal means, G is a MPIR of F if and only if there is a sequence (F_i) of distribution functions with uniformly bounded supports such that (i) $\lim F_i(x) = G(x)$ at all continuity points x of G and (ii) for each $i \geq 1$, F_i is a single crossing spread of F_{i-1} , where $F_0 = F$.

2.2 Dispersion and monotone increase in risk

The dispersive order, introduced and first studied by Bickel & Lehmann [2, 3] in the context of Non-parametric methods in Statistics, without requiring the two distributions to have equal nor even existing means, was re-defined by Quiggin in its mean-preserving version, and

applied to choice between lotteries in the RDEU model. Landsberger & Meilijson [24, 25] applied it to the design and study of efficient insurance contracts in the EU model.

(i) G is *more dispersed than* F (F less dispersed than G) if for every $0 < p < q < 1$, the length $G^{-1}(q) - G^{-1}(p)$ of the interval between the p - and the q - quantiles of G (weakly) exceeds the corresponding length for F . Equivalently, as defined in the Introduction,

(ii) G is more dispersed than F if the horizontal distance

$$D(p) = F^{-1}(p) - G^{-1}(p) \tag{6}$$

between F and G is a non-increasing function on the interval $(0, 1)$. Further equivalently,

(iii) G is more dispersed than F if on some probability space there are two comonotone random variables X and Z such that $X \sim F$ and $X + Z \sim G$.

If G is more dispersed than F , and if G and F have the same means, G will be said to be a monotone increase in risk with respect to F .

Dispersion is location-free, i. e., it is preserved if each random variable is shifted by the addition of an arbitrary constant. So, if F and G (with finite means and G more dispersed than F) are shifted so as to equate their means, G represents a monotone increase in risk with respect to F .

The middle definition implies the single-crossing property between F and G (in fact, a stronger property holds: any horizontal shift of F single crosses G). Hence, if the mean of G does not exceed that of F , F dominates G by SSD, according to Diamond & Stiglitz [13]. Landsberger & Meilijson [24] proved in the EU context that if the non-decreasing utility function v displays Arrow-Pratt more risk aversion than the non-decreasing utility function u , then a decision maker with utility function v is willing to buy any partial insurance contract that makes the insured position *less dispersed* than the uninsured position, as soon as a decision maker with utility function u is willing to buy it. As is well known, Pratt [31] introduced the notion of *more risk averse* by precisely this requirement, applied to *full* insurance. Ross [34] showed that mean-preserving increase in risk is not strong enough to yield this result. Monotone reduction in risk is not a strong assumption to make on insurance contracts: Landsberger and Meilijson [24] proved (again, in the EU context) that by the third definition of dispersion and the statement of the next sub-section, whenever an insurance contract does not reduce risk in the monotone way, there is another contract, preferred to the first by both parties, that does. Thus, monotone reduction in risk is a consequence of Pareto optimality:

2.2.1 The role of comonotonicity in Pareto-optimal allocations

Landsberger & Meilijson [24] showed that for any two integrable random variables X and Z , there are two other random variables X^* and Z^* defined on the same probability space,

such that (i) $X^* + Z^* = X + Z$ almost surely, (ii) X is a MPIR with respect to X^* , (iii) Z is a MPIR with respect to Z^* and (iv) X^* and Z^* are comonotone.

This statement means that every non-comonotone way in which two risk averse partners split a random endowment Y as $Y = X + Z$ (or *syndicate*, see Wilson [40]), can be replaced by a different contingent claim, weakly preferred by each, that is comonotone, a desirable property: if the partnership does better, no one of the partners does worse. Insurance contracts are examples of such partnerships, between insured and insurer.

2.2.2 The role of comonotonicity and monotone increase in risk on ordering premia

As pointed out earlier, if (in fact, only if) Y is monotonically riskier than X , then for every (negative, zero or positive premium) C , the distribution of Y single crosses the distribution of $X - C$. But then the distribution of $u(Y)$ single crosses that of $u(X - C)$ for all non-decreasing u . From this it follows that if an EU maximizer with utility function u is willing to pay C for the replacement of the uninsured position Y by the insured position X then so does another with utility function v , that is Arrow-Pratt more risk averse than the first. Observe that the value of C to achieve the inequality $Eu(Y) \leq Eu(X - C)$ for a non-decreasing function u can be positive or negative, since such functions cover risk averters as well as risk seekers.

However, if attention is restricted to risk averse DM's, then only positive C are of interest, since it is not needed that G single cross *every* horizontal shift of F ; it is enough that G single cross all horizontal shifts of F *towards the left*. This is why the Jewitt location-independent risk order, studied in the next section, is strong enough to be respected by premia paid by risk averse EU maximizers.

2.3 Location-independent risk and its dual version, left-monotone and right-monotone increase in risk

Jewitt [19] defined as *location-independent riskier* the following inequality between the integrated cumulative distribution functions, without the requirement of equal means.

Definition 1 G is *location-independent riskier* than F if $\Phi_G(G^{-1}(q)) \geq \Phi_F(F^{-1}(q))$ for every $q \in (0, 1)$.

If G is *location-independent riskier* than F and the two have the same finite mean, G will be said to be *left-monotone riskier* than F .

Analogously, Definition 2 introduces a mirror-image order via the corresponding inequality between the integrated de-cumulative distribution functions.

Definition 2 G is *dual location-independent riskier* than F if $\Psi_G(G^{-1}(q)) \geq \Psi_F(F^{-1}(q))$ for every $q \in (0, 1)$.

If G is dual location-independent riskier than F and the two have the same finite mean, G will be said to be right-monotone riskier than F .

Observe that Φ_H is a convex function, supported at the point $H^{-1}(q)$ by a linear tangent with slope q . $\Phi_G(G^{-1}(q)) \geq \Phi_F(F^{-1}(q))$ means that at whichever pair of points Φ_G and Φ_F have the same slope, Φ_G is higher. This is the same as saying that at whichever pair of points Φ_G and Φ_F have the same height, Φ_G has smaller slope. Since the same holds for Ψ ,

Lemma 1 (Landsberger & Meilijson [25]).

G is location-independent riskier than F if and only if the horizontal distance $\Phi_G^{-1}(x) - \Phi_F^{-1}(x)$ between the integrated cumulative distribution functions is non-decreasing in $x \in \mathcal{R}^+$.

G is dual location-independent riskier than F if and only if the horizontal distance $\Psi_G^{-1}(x) - \Psi_F^{-1}(x)$ between the integrated de-cumulative distribution functions is non-increasing in $x \in \mathcal{R}^+$.

Of course, under equality of means, the same conditions characterize left and right monotone increase in risk.

2.3.1 The role of left monotonicity in ordering premia

Location-independent risk, weaker than Bickel & Lehmann dispersion, was introduced by Jewitt [19] in an EU context of Insurance. Just as the insurance premium for a monotone reduction in risk is ordered by the Arrow-Pratt index of risk aversion in the class of non-decreasing utility functions, the insurance premium for a left-monotone reduction in risk is ordered by the index of risk aversion in the class of concave non-decreasing utility functions. In other words, as proved by Jewitt [19], if (the uninsured position) G is left-monotone riskier than (the insured position) F and some risk averse EU maximizer is willing to pay a given premium for the replacement of G by F , so is any such decision maker that is Arrow-Pratt more risk averse than the first. Furthermore, (Jewitt [19]) left-monotone increase in risk is the weakest order with this property, just as (Landsberger & Meilijson [24]) monotone increase in risk is the weakest order with this property in the class of all non-decreasing utility functions. Since it has already been explained above that left-monotone single crossings order premia as claimed, the picture will be complete if left-monotone increase in risk is identified as the uniformly integrable transitive closure of left-monotone single crossings. That is, phrasing somewhat informally a result in Landsberger & Meilijson [25],

Lemma 2 G is location-independent (dual location-independent, left-monotone, right-monotone) riskier than F if it is possible to obtain G as the limit of a sequence of distributions F_i such that $F_0 = F$ and for each $i \geq 1$, F_i left (right, mean-preservedly left, mean-preservedly right) monotonically single crosses F_{i-1} .

Proofs can be found in Landsberger & Meilijson [25]. We only show here that if G is a left-monotone single crossing of F then G is left-monotone riskier than F , by expressing

$$\Phi_F(F^{-1}(p)) = \int_{-\infty}^{F^{-1}(p)} F(t)dt = \int_0^p qdF^{-1}(q) = pF^{-1}(p) - \int_0^p F^{-1}(q)dq \quad (7)$$

to obtain

$$\Phi_G(G^{-1}(p)) - \Phi_F(F^{-1}(p)) = \int_0^p [(G^{-1}(p) - F^{-1}(p)) - (G^{-1}(q) - F^{-1}(q))]dq \quad (8)$$

$$= p[G^{-1}(p) - F^{-1}(p) + L_F(p) - L_G(p)], \quad (9)$$

that must be shown to be non-negative. Now observe that for p below the single crossing the integrand in the RHS of (8) is non-negative, while for p above the single crossing, (9) is non-negative because it is the sum of two non-negative terms - recall the Lorenz characterization of second degree dominance and the fact that single crossing (whether left-monotone or not), implies second degree dominance.

Since the sum $\Phi_H^{-1}(x) - \Phi_F^{-1}(x) = (\Phi_H^{-1}(x) - \Phi_G^{-1}(x)) + (\Phi_G^{-1}(x) - \Phi_F^{-1}(x))$ of non-decreasing functions is non-decreasing, left-monotone increase in risk is transitive. This shows that any G generated from F by a sequence of left-monotone single crossings (with Φ -functions converging from below, a property related to uniform integrability), is left-monotone riskier than F .

2.3.2 Monotone transformations of random variables ordered by the various notions of dispersion

It is clear that the single crossing property is preserved by increasing transformations. It is just as clear that concave increasing transformations preserve second-degree dominance and convex increasing transformations preserve convex dominance.

Let Y be Bickel & Lehmann more dispersed than X . We will see that if g is increasing and concave, then $g(Y)$ is location-independent riskier than $g(X)$, and if g is increasing and convex, then $g(Y)$ is riskier than $g(X)$ in the sense of the mirror image of this order. In particular, if a constant C is adjusted in each of the two cases so that $Eg(Y) = Eg(X - C)$, then $g(Y)$ is accordingly a left or right monotone increase in risk with respect to $g(X - C)$. More generally,

Theorem 1 (i) *If Y is location-independent riskier than X and $g : \mathcal{R} \rightarrow \mathcal{R}$ is increasing and concave, then $g(Y)$ is location-independent riskier than $g(X)$.*

(ii) *If Y is a dual location-independent riskier than X and $g : \mathcal{R} \rightarrow \mathcal{R}$ is increasing and convex, then $g(Y)$ is dual location-independent riskier than $g(X)$.*

Proof. By Lemma 2, it is enough to study the effect of these monotone transformations on left or right monotone single crossings. To see (i), observe that the definition of left-monotone single crossing implies that on $U \sim U(0, 1)$ as probability space, it is possible to

express $Y = F_Y^{-1}(U)$ as a sum of $X = F_X^{-1}(U)$ and left comonotone noise as $F_Y^{-1}(U) = F_X^{-1}(U) + Z(U)$, where Z single crosses zero and is non-decreasing while negative. Since $g(F_Y^{-1}(U))$ and $F_{g(Y)}^{-1}(U)$ are identically distributed non-decreasing functions of U , they must be a.s. equal. Consider $g(Y) = g(X) + [g(X + Z) - g(X)]$ and let $u < v$. Then as long as $Z(u) \leq 0$, $g(X(v) + Z(v)) - g(X(v)) \geq g(X(v) + Z(u)) - g(X(v))$ because g and Z are non-decreasing. Furthermore, on this range $g(X(v) + Z(u)) - g(X(v)) \geq g(X(u) + Z(u)) - g(X(u))$ because g is concave. Hence, $F_{g(Y)}^{-1}(U) - F_{g(X)}^{-1}(U) = Z^*(U)$ is non-decreasing while negative, because the equality $Z^*(U) = g(X(U) + Z(U)) - g(X(U))$ holds while negative. The behavior beyond the crossing is covered by the fact that the single crossing property is preserved under increasing transformations. Thus, Y is a left-monotone single crossing of X . The proof of (ii) is analogous.

Corollary 1 (Chateauneuf & Moyes [9]) *Let F and G be distributions of positive random variables (with the same mean), and let G display more inequality in the RDO sense (Marshall & Olkin [29], Moyes [30]). Then G is dual location-independent (respectively, right-monotone) riskier than F .*

Proof. Let $X \sim F$ and $Y \sim G$. The notion of RDO means that for any $0 < p < q < 1$, the relative $\frac{p}{q}$ -quantile $\frac{F^{-1}(p)}{F^{-1}(q)}$ of F within its q -quantile weakly exceeds the corresponding expression for G . This means that $\log(Y)$ is Bickel & Lehmann more dispersed than $\log(X)$. Thus, X and Y are the result of applying the increasing convex transformation $g(x) = \exp(x)$ to two random variables ordered by dual location-independent risk.

The converse to Corollary 1 is wrong: since the property of being a left or right monotone single crossing is not transitive, RDO does not coincide with dual location-independent risk.

2.3.3 Deductible insurance left monotonically dominates any other

Arrow [1] proved that for a given premium (based on the expected value of the loss), insurance contracts with deductible are optimal for EU risk averse agents.

This result was generalized by Gollier & Schlesinger [14] to agents averse to MPIR, independently of their decision model under risk.

Vergnaud [38] extended this result further: the insured position induced by any contract with the given premium is left-monotone riskier than the insured position induced by the corresponding contract with deductible.

Furthermore, in a result complementing Jewitt's claim on the behavior of the Arrow-Pratt index, Chateauneuf, Cohen and Vergnaud [8] showed that aversion to left-monotone risk is the weakest notion of risk aversion under which deductible insurance contracts are optimal.

2.3.4 The role of left and right monotone increase in risk in a classical search problem

The Chow & Robbins search problem [11] (see also Karni & Schwartz [20, 21]) considers a seller of one good that is given by re-callable potential buyers successive offers X_1, X_2, \dots , assumed to be independent and identically F -distributed non-negative random variables. Each such offer incurs a cost C for the seller so selling after n offers will yield a net revenue $NR_{F,C,n} = \max_{1 \leq i \leq n} X_i - nC$. The optimal stopping time, maximizing $E[NR_{F,C,\tau}]$ over all stopping times τ , is known to be of a control limit form, that dictates selling at the first offer not below the threshold $b = \Psi_F^{-1}(C)$ solving $\Psi_F(b) = C$ (see (3)). The optimal threshold $\Psi_F^{-1}(C)$ turns out to be the optimal expected revenue as well. Thus, the expected overpayment over the threshold exactly pays for sampling. Indeed, restricting attention to such control-limit stopping rules τ_b , $E[NR_{F,C,\tau_b}] = E[X|X \geq b] - C/(1 - F(b-))$, the first term being the expected price received and the second term being the expected sampling cost, expressed in terms of the expectation $1/(1 - F(b-)) = 1/P(X > b)$ of the (geometrically distributed) number of offers sampled. Now,

$$\begin{aligned} E[NR_{F,C,\tau}] &= E[X|X \geq b] - \frac{C}{P(X \geq b)} = \frac{E[XI_{X \geq b}] - C}{1 - F(b-)} \\ &= b + \frac{E[\max(X - b, 0)] - C}{1 - F(b-)} = b + \frac{\Psi(b) - C}{1 - F(b-)} \end{aligned} \quad (10)$$

from which it is not hard to get the result

$$\sup_{\tau} E[NR_{F,C,\tau}] = E[X|X \geq \Psi_F^{-1}(C)] - \frac{C}{1 - F(\Psi_F^{-1}(C)-)} = \Psi_F^{-1}(C) \quad (11)$$

The Chow & Robbins seller's search problem [11] can be turned over into the problem of a shopper that gets independent F -distributed price quotes Y_i at a cost C each, and waits optimally for a price not exceeding the control limit $d = \Phi_F^{-1}(C)$ that satisfies $\Phi_F(d) = C$ (see (4)). The optimal total cost is

$$\inf_{\tau} E[TC_{F,C,\tau}] = E[Y|Y \leq \Phi_F^{-1}(C)] + \frac{C}{F(\Phi_F^{-1}(C))} = \Phi_F^{-1}(C). \quad (12)$$

Suppose that the price quotes the buyer gets are distributed G rather than F . The G 's under which the buyer is better off than under F for every $C > 0$ are those such that for every $C > 0$, the point $\Phi_G^{-1}(C)$ where Φ_G reaches height C is to the left of the corresponding point $\Phi_F^{-1}(C)$. In other words, $\Phi_G \geq \Phi_F$, or F must dominate G by SSD. If F and G have the same expectation, this means that G must be a MPIR of F .

This makes sense: dispersion helps the patient buyer to benefit from the heavier lower tail. Intuition further dictates the rationale that in a more dispersed market, it pays to the buyer to sample longer in expectance of lower offers. After all, the least dispersed case, the "distribution" of a fixed price, dictates buying at the first quote. This rationale would mean that for all $C > 0$, the expected sample size $1/(1 - G(\Phi_G^{-1}(C)-))$ under G should not

be less than the corresponding expected sample size under F . I.e., the slope $1 - G(\Phi_G^{-1}(C))$ should be less than or equal to the slope $1 - F(\Phi_F^{-1}(C))$ for all $C > 0$, or, the horizontal distance function $\Phi_G^{-1}(C) - \Phi_F^{-1}(C)$ should be non-decreasing. While SSD or MPIR are not strong enough to yield this inequality, by Lemma 1 this condition says precisely that G is location-independent riskier than F . Summarizing for the case of equal means,

- If the distribution G of the asking prices is a MPIR of F , the buyer's optimal expected total cost is lower under G , for all $C > 0$.
- If G is a left-monotone increase in risk with respect to F , then under G the buyer's optimal expected total cost is lower while the expected sampling cost is higher than under F , for all $C > 0$.

This means that in a left-monotone more dispersed market, the sellers will get a much lower revenue, since besides making the buyer better off, they have to subsidize the buyer's higher sampling cost.

Going back to the seller's problem, it should be clear that it admits such an analysis too. Summarizing the results for the case of equal means,

- If the distribution G of the price offers is a MPIR of F , the seller's optimal expected revenue is higher under G , for all $C > 0$.
- If G is a right-monotone increase in risk with respect to F , the seller's optimal expected revenue and sampling cost are higher under G , for all $C > 0$.

This means that in a right-monotone more dispersed market, the buyer will pay a much higher price, since besides making the seller better off, the buyer has to subsidize the seller's higher sampling costs.

3 Attitudes to increase in risk in the RDEU model

For convenience of exposition, distributions $D(X) = (x_1, p_1; x_2, p_2; \dots; x_{k-1}, p_{k-1}; x_k, p_k)$ of random variables X with finite support, tacitly assumed to satisfy $x_1 < x_2 < \dots < x_k$, will be represented in the equivalent form

$$D(X) = (x_1, 1 - q_1; x_2, q_1 - q_2; \dots; x_{k-1}, q_{k-2} - q_{k-1}; x_k, q_{k-1}) . \quad (13)$$

The original expression is in terms of the probabilities $P(X = x_i) = p_i$ of the atoms while (13) is a representation in terms of the right-tail probabilities $P(X > x_i) = q_i$, with $p_i = q_{i-1} - q_i$ under the convention $q_0 = 1, q_k = 0$. For such a discrete distribution, the expected utility and the Quiggin-Yaari functional (see (1)) may be expressed as

$$\begin{aligned} E[u(X)] = & - \int u(x) dF^*(x) = u(x_1) + q_1[u(x_2) - u(x_1)] \\ & + q_2[u(x_3) - u(x_2)] + \dots + q_{k-1}[u(x_k) - u(x_{k-1})] \end{aligned} \quad (14)$$

and

$$V(X) = - \int u(x)df(F^*(x)) = u(x_1) + f(q_1)[u(x_2) - u(x_1)] \\ + f(q_2)[u(x_3) - u(x_2)] + \cdots + f(q_{k-1})[u(x_k) - u(x_{k-1})] \quad (15)$$

respectively, where $F^*(x) = P(X > x)$ is the *survival function* of the random variable X . The function $F(x) = 1 - F^*(x) = P(X \leq x)$ is its *distribution function*.

If the distribution of X is perturbed by mean-preservedly spreading out two atoms strictly preserving the order of all atoms, the resulting distribution (of a random variable Y) may be expressed in the form (compare with (13))

$$D(Y) = (x_1, 1 - q_1; \dots; x_{i-1}, q_{i-2} - q_{i-1}; x_i - \epsilon[q_{j-1} - q_j], q_{i-1} - q_i; x_{i+1}, q_i - q_{i+1}; \\ \dots; x_{j-1}, q_{j-2} - q_{j-1}; x_j + \epsilon[q_{i-1} - q_i], q_{j-1} - q_j; x_{j+1}, q_j - q_{j+1}; \dots; x_k, q_{k-1}) . \quad (16)$$

This type of spread keeps all probabilities unchanged. It also keeps all except two of the atoms of the distribution of X unchanged. The two exceptional atoms ($x_i < x_j$) are set further apart by the precise multiples of a sufficiently small arbitrary ϵ that will keep the expected value constant, without disrupting the order of the atoms. Such a spread can be conveniently parameterized by specifying ϵ, i and j . The analysis, to be performed in Section 3.1, will be restricted to the four canonical examples (see (17)-(20)), that are respective special cases of the four types of single-crossing operations defined in the Introduction.

The fact that each of the stochastic orders under consideration is generated by a single-crossing operation between distribution functions, has far reaching implications for analysis. As formalized by the following Lemma, in order to study the effect of each of these four types of increase in risk on the Quiggin-Yaari functional under general concave utility functions u and general transformation functions f , it is enough to study their effect on the *Yaari* functional, i.e., the Quiggin-Yaari functional with linear utility under certainty $u(x) = x$.

Lemma 3 *If the distribution G of a random variable Y is a mean-preserving or mean-reducing single crossing spread of the distribution F of a random variable X and f is a probability transformation function such that $V_{I,f}(X) \geq V_{I,f}(Y)$, then $V_{u,f}(X) \geq V_{u,f}(Y)$ for all concave non-decreasing utility-on-certainty functions u .*

Proof: If the distribution G single crosses the distribution F , then the distribution $1 - f(G^*)$ single crosses the distribution $1 - f(F^*)$. Hence, if $1 - f(F^*)$ has expectation bigger or equal to that of $1 - f(G^*)$ (i.e., if the Yaari functionals satisfy the corresponding inequality), then the former dominates the latter by second degree dominance (see Diamond & Stiglitz [13]). But then, the preference of F by decision makers with linear utility extends automatically to its preference by decision makers with concave utility.

In other words, since the Quiggin-Yaari functional is in essence Expected Utility with respect to the transformation-function-transformed distributions, the introduction of concave

utilities under certainty will automatically preserve the corresponding inequalities satisfied by linear utilities.

Lemma 3 and all the discussion so far deal with different types of transformation function, asking which type models some specific kind of aversion to risk, by DM's with concave utility functions. The authors have characterized in [7] the pairs u, f with possibly non-concave u for which the DM is averse to monotone increase in risk. The relaxation of concavity under the Jewitt orders is the subject of our current research.

3.1 RDEU decision makers' aversion to middle spread, spread of extremes, spread of losses and spread of gains

Returning to the four canonical examples of mean-preserving spread announced earlier and motivated by Lemma 3, each is now analyzed in terms of the Yaari functional. As an example, $k = 4, i = 2$ and $j = 3$ means that the distribution is supported by four atoms, of which the middle two are subjected to a mean-preserving spread. For each canonical example, the difference $V_{I,f}(Y) - V_{I,f}(X)$ is displayed. Aversion to a particular type of mean-preserving spread means that $V(Y) - V(X)$ is non-positive for all choices of the distribution, in the corresponding canonical class. It is always assumed that $\epsilon > 0$ is small enough, so that the order of the atoms is preserved.

1. **Canonical middle spread:** $k = 4, i = 2$ and $j = 3$. Then

$$V_{I,f}(Y) - V_{I,f}(X) = \epsilon(q_1 - q_3)[f(q_2) - \frac{q_2 - q_3}{q_1 - q_3}f(q_1) - \frac{q_1 - q_2}{q_1 - q_3}f(q_3)] . \quad (17)$$

2. **Canonical extreme spread:** $k = 3, i = 1$ and $j = 3$. Then

$$V_{I,f}(Y) - V_{I,f}(X) = \epsilon(1 - q_1)q_2[\frac{f(q_2)}{q_2} - \frac{1 - f(q_1)}{1 - q_1}] . \quad (18)$$

3. **Canonical spread of losses:** $k = 3, i = 1$ and $j = 2$. Then

$$V_{I,f}(Y) - V_{I,f}(X) = \epsilon(1 - q_1)(1 - q_2)[\frac{1 - f(q_2)}{1 - q_2} - \frac{1 - f(q_1)}{1 - q_1}] . \quad (19)$$

4. **Canonical spread of gains:** $k = 3, i = 2$ and $j = 3$. Then

$$V_{I,f}(Y) - V_{I,f}(X) = \epsilon q_1 q_2[\frac{f(q_2)}{q_2} - \frac{f(q_1)}{q_1}] . \quad (20)$$

Thus,

Lemma 4 (i) *The Yaari functional displays aversion to canonical middle spread (17) if and only if the transformation function f is convex.*

(ii) *The Yaari functional displays aversion to canonical extreme spread (18) if and only if the transformation function f is majorized by the identity function.*

(iii) The Yaari functional displays aversion to canonical spread of losses (19) if and only if the transformation function f is majorized by the identity function and star-shaped at 1.

(iv) The Yaari functional displays aversion to canonical spread of gains (20) if and only if the transformation function f is majorized by the identity function and star-shaped at 0.

As defined in the Introduction,

Definition 3 A function g defined on some interval I is said to be star-shaped at $\mu \in I$ if $(g(x) - g(\mu))/(x - \mu)$ is monotone on $I - \{\mu\}$.

For a probability transformation function g majorized by the identity function and star-shaped at $\mu = 0$ or 1 , $(g(x) - g(\mu))/(x - \mu)$ is monotone non-decreasing.

Proof of Lemma 4: In each of the four cases we must characterize the class of transformation functions f for which $V_{I,f}(Y) - V_{I,f}(X)$ is non-positive for all choices of the distribution parameters. Equivalently, the f 's for which the expression in brackets in the corresponding RHS is non-positive. In expression (17) this is readily seen to be the class of convex functions. In expression (18), one direction is clear: if f is majorized by the diagonal then $\frac{f(q)}{q} \leq 1 \leq \frac{1-f(q)}{1-q}$. As for the other direction, due to the monotonicity of f , if f is not majorized by the diagonal, then there are necessarily $0 < q_2 < q_1 < 1$ with $f(q_i) > q_i$. Since $\frac{f(q_2)}{q_2} > 1 > \frac{1-f(q_1)}{1-q_1}$, this decision maker strictly prefers Y to X . The remaining two expressions in brackets in (19) and (20) are non-positive throughout if and only if the corresponding f is star-shaped at the corresponding point, by definition.

3.2 Star-shaped probability transformation functions: EU and RDEU, lotteries and insurance

Star-shapedness of f at 0 and at 1 can be given a related but different role, that sharpens the edge between EU and RDEU. Let a risk averse EU-maximizer with utility function u be offered full insurance for a single loss L that occurs with probability $1 - p$ to her house worth W . The fair premium is $P_f = (1 - p)L$ and the supremal premium that can be extracted from the agent is $P_s = W - u^{-1}(u(W-L) (1-p) + u(W) p)$ so the *load factor* the agent can withstand is

$$\begin{aligned}
 LF_{EU} &= \frac{P_s}{P_f} = \frac{W - u^{-1}(u(W) - (1 - p)(u(W) - u(W - L)))}{(1 - p)L} \\
 &= \frac{u(W) - u(W - L)}{L} \frac{u^{-1}(u(W)) - u^{-1}(u(W) - (1 - p)(u(W) - u(W - L)))}{(1 - p)(u(W) - u(W - L))} \\
 &= \frac{u(W) - u(W - L)}{L} \frac{u^{-1}(v) - u^{-1}(v - \Delta)}{\Delta}, \tag{21}
 \end{aligned}$$

where $v = u(W)$ and $\Delta = (1 - p)(u(W) - u(W - L))$. Since u is concave, u^{-1} is convex and the load factor is a non-decreasing function of p . Although a change in p is a first

rather than a second degree change, EU risk aversion has a clear cut effect: risk averse EU maximizers are willing to pay relatively more for a rarer loss event.

After studying this issue purely in terms of the utility u , we now study it purely in terms of the transformation function f , by its effect on the Yaari index. Skipping straightforward details, the load factor is now $LF_{Yaari} = \frac{1-f(p)}{1-p}$. Thus, the load factor is a non-decreasing function of p if and only if f is star-shaped at 1. Among the RDEU agents with linear utility, those with transformation function star-shaped at 1 are precisely the ones that display the unanimous attitude of EU risk averters on this issue.

A related question (but notably less important) involving lotteries instead of insurance is the attitude of EU risk averters to a lottery with price C and probability p of winning a single prize K . The fair value of K is of course $K_f = \frac{C}{p}$, and the infimal prize for participation of the EU agent is $K_i = C + u^{-1}\left(\frac{u(0)-u(-C)(1-p)}{p}\right)$.

Similarly to the analysis for insurance, the top-prize load factor $\frac{K_i}{K_f}$ can be seen to be a non-increasing function of p : EU risk averters unanimously demand relatively higher prizes as compensation for lower winning probabilities.

The corresponding top-prize load factor under the Yaari index is $\frac{p}{f(p)}$, that is non-increasing if and only if the transformation function f is star-shaped at 0.

Combining these diverse attitudes to lottery loading or spread of gains and premium loading or spread of losses, we can see that although RDEU pessimism is more flexible than EU risk aversion, this flexibility is distinctly normative.

Some further remarks

Landsberger & Meilijson [23] introduced (in the EU framework) a notion of *selective* risk aversion that is modeled by star-shaped utility functions. Instead of the property of decreasing *marginal* utility displayed by concavity, star-shape yields decreasing *average* utility, as measured from the point of reference μ .

It may be of interest to notice the geometrical fact that the transformation functions that are star-shaped at *both* 0 *and* 1 are those expressible as pointwise minima of convex transformation functions. Since this class strictly contains the convex functions, aversion to both left-monotone and right-monotone increase in risk does not imply aversion to MPIR. Perhaps there is an interesting stochastic order hidden behind this fifth class.

In the framework of inequality measurement Chateauneuf [4] proved that the Yaari [43] index of inequality measurement respects RDO if and only if f is majorized by the identity function and is star-shaped at 0. This is consistent with Corollary 1. In this context, for every income random variable X and any f star shaped at 0, the Yaari index can be interpreted as the average level of income after (rank-dependent) progressive income taxation.

3.3 Formulation of the main results

Once the four special classes above of transformation functions have been obtained from basic principles and justified intuitively, it becomes a natural question of interest to study the stochastic orders that model the choices of decision makers with such transformation functions. In other words, the task is to discover the different restrictive notions of mean-preserving increase in risk, of which the four canonical mean-preserving spreads (17)-(20) are special cases. The main results are formulated in the four theorems below. Their proofs appear in the next section.

The rationale behind the three different statements in each of the following theorems, is that the first statement identifies the type of transformation function of a given DM with linear utility, from the DM's attitude to a specific type of mean-preserving increase in risk; the second statement qualifies the attitude to the various kinds of mean-preserving increase in risk of a DM with concave utility function, in terms of the DM's type of transformation function; the third statement identifies the type of dominance relating two given distributions with equal finite means, from unanimous preference pattern between the two by all DM's with linear utility function and a certain type of transformation function.

Thus, in each theorem, statements (i) and (ii) deal with an individual DM with concave utility and characterize the DM's type in terms of aversion to the various kinds of increase in risk, while statements (ii) and (iii) deal with a particular pair of distributions with equal means and characterize the type of increase in risk connecting these two distributions, in terms of the unanimous preference for one over the other by all DM's of a given type.

Theorem 2 Mean-preserving increase in risk and convex transformation function

(i) (Chew, Karni & Safra [10], Yaari [42]) *Let a RDEU decision maker with linear utility function be averse to mean-preserving increase in risk. Then the decision maker's transformation function f is convex.*

(ii) (Chew, Karni & Safra [10]) *Let the transformation function be convex and the utility function be concave (e.g., linear). Then the RDEU decision maker is averse to mean-preserving increase in risk.*

(iii) (Chateauneuf [5], Yaari [41]) *Let F and G be two distributions with the same finite mean, such that all RDEU decision makers with linear utility function and convex probability transformation function prefer F to G . Then G represents a mean-preserving increase in risk with respect to F .*

Theorem 3 Monotone increase in risk and transformation function below the identity function

(i) (Quiggin [33]) *Let a RDEU decision maker with linear utility function be averse to monotone increase in risk. Then the decision maker's transformation function f is bounded*

from above by the identity function.

(ii) (Quiggin [33]) Let the transformation function be bounded from above by the identity function, and the utility function be concave (e.g., linear). Then the RDEU decision maker is averse to monotone increase in risk.

(iii) Let F and G be two distributions with the same finite mean, such that all RDEU decision makers with linear utility function and probability transformation function bounded from above by the identity function prefer F to G . Then G represents a monotone increase in risk with respect to F .

Theorem 4 Left-monotone increase in risk and transformation function below the identity function, star-shaped at 1

(i) Let a RDEU decision maker with linear utility function be averse to left-monotone increase in risk. Then the decision maker's transformation function f is bounded from above by the identity function and is star-shaped at 1.

(ii) Let the transformation function be bounded from above by the identity function and star-shaped at 1, and the utility function be concave (e.g., linear). Then the RDEU decision maker is averse to left-monotone increase in risk.

(iii) Let F and G be two distributions with the same finite mean, such that all RDEU decision makers with linear utility function and probability transformation function bounded from above by the identity function and star-shaped at 1 prefer F to G . Then G represents a left-monotone increase in risk with respect to F .

Theorem 5 Right-monotone increase in risk and transformation function below the identity function, star-shaped at 0

(i) Let a RDEU decision maker with linear utility function be averse to right-monotone increase in risk. Then the RDEU decision maker's transformation function f is bounded from above by the identity function and is star-shaped at 0.

(ii) Let the transformation function be bounded from above by the identity function and star-shaped at 0, and let the utility function be concave (e.g., linear). Then the decision maker is averse to right-monotone increase in risk.

(iii) Let F and G be two distributions with the same finite mean, such that all RDEU decision makers with linear utility function and probability transformation function bounded from above by the identity function and star-shaped at 0 prefer F to G . Then G represents a right-monotone increase in risk with respect to F .

4 Proof of Theorems 2, 3, 4 and 5.

The four statements (i) of Theorems 2, 3, 4 and 5 are corollaries of Lemma 4, since the four canonical examples are special cases of the corresponding stochastic orders.

The four statements (ii) of Theorems 2, 3, 4 and 5 follow directly from Lemma 3 for the case of concave utility functions, provided these statements hold for linear utility functions, the case to be proven in the sequel.

Since the four classes of transformation function were constructed as necessary conditions for the proper preferences of decision makers when subjected to each of the canonical spreads (17)-(20), each stochastic order turns out to provide a *characterization* of choice by RDEU decision makers with linear utility and transformation function in the corresponding class, as claimed by Theorems 2, 3, 4 and 5.

4.1 Mean-preserving increase in risk

In order to prove Theorem 2 (ii), consider the following representations (22) and (23) of the Yaari functional (substitute $u(z) = z$ in (1)) of a H -distributed random variable Z . Representation (22) follows easily from first principles, the definition of the Stieltjes integral as a limit of the integral of discretizations of the integrand. Representation (23) is meaningful for convex f . While the formula of integration by parts holds for all p_0 , (23) is obtained by substituting $p_0 = 1$.

$$\begin{aligned}
V_{I,f}(Z) &= - \int_{-\infty}^{\infty} x df(H^*(x)) = \int_0^1 (H^*)^{-1}(p) df(p) & (22) \\
&= f'(p) \int_{p_0}^p (H^*)^{-1}(q) dq \Big|_0^1 - \int_0^1 \left[\int_{p_0}^p (H^*)^{-1}(q) dq \right] df'(p) \\
&= f'(p) \int_{p_0}^p (H^*)^{-1}(q) dq \Big|_0^1 + \int_0^1 \left[\int_{1-p_0}^{1-p} H^{-1}(r) dr \right] df'(p) \\
&= f'(0)E[Z] + \int_0^1 L_H(1-p)(1-p) df'(p) . & (23)
\end{aligned}$$

It follows from (23) that if f is convex and $Y \sim G$ is a MPIR with respect to $X \sim F$, then $V_{I,f}(X) \geq V_{I,f}(Y)$. This is so because $E[X] = E[Y]$, $(1-p)df'(p)$ is a non-negative measure and the integrands are pointwise ordered in the direction needed, by the Lorenz-type characterization of SSD following (5). This proves Theorem 2 (ii).

To prove Theorem 2 (iii), suppose that G does not represent a mean-preserving increase in risk with respect to F . A convex transformation function f must be displayed under which the Yaari functional shows strict preference for G . Under the assumption above, the distance $L_F(\cdot) - L_G(\cdot)$ between the L-functions (see 5) fails to be non-negative. Let $N \subseteq (0, 1)$ be the set of p on which $L_F(1-p) - L_G(1-p)$ is strictly negative. Since N is a non-empty open set, it has positive Lebesgue measure. Motivated by representation (23), we want to build a convex transformation function f whose derivative f' increases only on N . With this in mind, let $f : [0, 1] \rightarrow [0, 1]$ be defined by

$$f(0) = 0 ; f'(q) = \text{Leb}(N \cap [0, q]) / \int_0^1 \text{Leb}(N \cap [0, p]) dp , q \in (0, 1) . \quad (24)$$

where Leb stands for Lebesgue measure. The derivative of f measures the local density of the set N , so it increases only on N . Since f' is a properly normalized non-negative and non-decreasing function, f is a convex probability transformation function. By (23),

$$\begin{aligned} V_{I,f}(X) - V_{I,f}(Y) &= \int_0^1 [L_F(1-p) - L_G(1-p)](1-p)df'(p) \\ &= \int_N [L_F(1-p) - L_G(1-p)](1-p)df'(p) < 0 . \end{aligned} \quad (25)$$

Hence, G is strictly preferred to F . This proves Theorem 2 (iii).

Remark. The simplifying assumption of continuity of f can be seen at play here, since a discontinuous f would render the RHS of (22) unstably dependent on which version of $(H^*)^{-1}$ is used - left continuous, right continuous or otherwise.

4.2 Monotone increase in risk

Consider the representation (22) of the Yaari functional of a H -distributed random variable and the notation (6) for the horizontal distance between distribution functions. If $Y \sim G$ represents a monotone increase in risk with respect to $X \sim F$ then $V_{I,f}(X) - V_{I,f}(Y) = \int_0^1 D(1-p)df(p)$ is the expectation of a non-decreasing function of a f -distributed random variable. Since f is assumed to be majorized by the identity function, this “distribution” $df(p)$ dominates the uniform distribution dp on $(0, 1)$ by first degree stochastic dominance. Hence, an EU decision maker with utility function D prefers a f -distributed random variable to a uniformly distributed one. In other words,

$$\begin{aligned} V_{I,f}(X) - V_{I,f}(Y) &= \int_0^1 D(1-p)df(p) \geq \int_0^1 D(1-p)dp \\ &= \int_0^1 [F^{-1}(1-p) - G^{-1}(1-p)]dp = E[X] - E[Y] = 0 , \end{aligned} \quad (26)$$

thus proving Theorem 3 (ii).

To prove Theorem 3 (iii), suppose that G does not represent a monotone increase in risk with respect to F . We have to display a transformation function f supported from above by the identity function, under which the Yaari functional shows strict preference for G . Under the current assumption, the function D is not non-decreasing. Keeping in mind (26), there is a distribution function f supported by the unit interval, that first-degree dominates the uniform distribution (i.e., $f(q) \leq q$ for all $q \in (0, 1)$), such that a EU-maximizer decision maker with utility function D strictly prefers the uniform distribution to the distribution f . This f can be taken to be continuous and strictly increasing, to conform with the properties of probability transformation functions. Since this provides a strict reversal of the inequality in (26), a transformation function majorized by the identity has been constructed, under which G is strictly preferred to F . This proves Theorem 3 (iii).

4.3 Left-monotone increase in risk

The proof of Theorem 4 is entirely analogous to the proof of Theorem 5 for right-monotone increase in risk. Since the latter has never been dealt with before, it is chosen as the setup for the proof.

4.4 Right-monotone increase in risk and transformation functions star-shaped at zero

Clearly, this stochastic order is characterized by pointwise inequality between distances of Ψ_G and Ψ_F , just as left-monotone increase in risk is characterized by the corresponding inequality for the Φ functions (see (8)).

If the transformation function f is below the diagonal and star-shaped at 0, then $\frac{f(p)}{p}$ is non-decreasing. Since it equals 1 at $p = 1$, consider the distribution function ν of a random variable U with values in the interval $[0, 1]$ given by $P(U \leq p) = \nu(p) = \frac{f(p)}{p}$ for $p \in (0, 1]$, with $P(U = 0) = \nu(0) = \lim_{p \downarrow 0} \frac{f(p)}{p}$. Under this notation,

Lemma 5 *For a bounded random variable Z distributed H ,*

$$V_{I,f}(Z) = E(Z) - \int_0^1 \Psi_H(H^{-1}(1 - q)) d\nu(q) \quad (27)$$

Remark: It should be intuitively clear why is a representation like this useful: Since (see (8)) right-monotone increase in risk is characterized by pointwise inequality between functions like $\Psi_H(H^{-1}(1 - q))$, integration with respect to a positive measure will preserve the inequality, i.e., it will transfer it to the Yaari index.

Although we discovered representation (27) of $V_{I,f}$ by a number of steps of integration by parts, we opt for proving it otherwise.

Proof of Lemma 5: The proof will be divided into two steps. In the second step the Lemma will be proved for random variables supported by finite sets. In the first step it will be shown that for arbitrary $\epsilon > 0$, if $P(|Z - \tilde{Z}| < \epsilon) = 1$, then each of the three functionals in (27) changes by at most 2ϵ when Z is replaced by \tilde{Z} . This is enough because every random variable (say, Z) with bounded support can be uniformly ϵ -approximated by a random variable (call it \tilde{Z}) with finite support. Simply, let \tilde{Z}/ϵ be the integer part of Z/ϵ .

Step 1. The fact that $|V_{I,f}(Z) - V_{I,f}(\tilde{Z})|$ and $|E(Z) - E(\tilde{Z})|$ are less than ϵ are immediate consequences of each of $V_{I,f}$ and E being (i) monotone with respect to First Degree Dominance and (ii) linear with respect to addition of constants. Details are skipped, except the proof for the integral term. Since $|Z - \tilde{Z}| < \epsilon$ a.s.,

$$|H^{-1}(1 - q) - \tilde{H}^{-1}(1 - q)| < \epsilon \text{ for each } q \in (0, 1) . \quad (28)$$

Since $\Psi_H(x) = E(Z - x)^+ = E(\max(Z, x)) - x$, subtract x from each term in

$$\begin{aligned} E(\max(\tilde{Z}, x + \epsilon)) - \epsilon &= E(\max(\tilde{Z} - \epsilon, x)) \leq E \max(Z, x) \\ &\leq E(\max(\tilde{Z} + \epsilon, x)) = E(\max(\tilde{Z}, x - \epsilon)) + \epsilon \end{aligned} \quad (29)$$

to obtain that for all x , $\Psi_{\tilde{H}}(x + \epsilon) \leq \Psi_H(x) \leq \Psi_{\tilde{H}}(x - \epsilon)$. The third and last preparatory step needed is inferring from the representation $\Psi_H(x) = \int_x^\infty (1 - H(t))dt$ that the derivative of Ψ is in $(-1, 0)$. Indeed,

$$\begin{aligned} \Psi_{\tilde{H}}(\tilde{H}^{-1}(1 - q)) - 2\epsilon &\leq \Psi_{\tilde{H}}(\tilde{H}^{-1}(1 - q) + 2\epsilon) \leq \Psi_H(\tilde{H}^{-1}(1 - q) + \epsilon) \\ &\leq \Psi_H(H^{-1}(1 - q)) \leq \Psi_H(\tilde{H}^{-1}(1 - q) - \epsilon) \\ &\leq \Psi_{\tilde{H}}(\tilde{H}^{-1}(1 - q) - 2\epsilon) \leq \Psi_{\tilde{H}}(\tilde{H}^{-1}(1 - q)) + 2\epsilon, \end{aligned} \quad (30)$$

finishing Step 1.

Step 2. Proof for the discrete case, by induction on the number of atoms of the distribution.

If the distribution has two atoms, express it as $(x_1, p_1; x_2, p_2)$. Then $E(Z) - V_{I,f}(Z) = (x_1 + p_2(x_2 - x_1)) - (x_1 + f(p_2)(x_2 - x_1)) = (p_2 - f(p_2))(x_2 - x_1)$. Evaluation of the integral yields

$$\begin{aligned} \int_0^1 \Psi_H(H^{-1}(1 - q))d\nu(q) &= \Psi(x_2)\nu(0, p_2) + \Psi(x_1)\nu(p_2, 1) \\ &= \Psi(x_1)\nu(p_2, 1) = (x_2 - x_1)p_2\nu(p_2, 1). \end{aligned} \quad (31)$$

The two are equal if and only if $\nu(p_2, 1) = 1 - f(p_2)/p_2$, or $\nu(0, p_2) = f(p_2)/p_2$, that holds by definition of the probability measure ν .

The inductive step consists of replacing the distribution $H = (x_1, p_1; x_2, p_2; \dots; x_k, p_k)$ of Z (with $x_1 < x_2 < \dots < x_k$ and $k \geq 3$) by \tilde{H} of \tilde{Z} obtained from H by replacing x_1 by x_2 . In other words, $P(\tilde{Z} = x_i) = P(Z = x_i) = p_i$ for $3 \leq i \leq k$ and $P(\tilde{Z} = x_2) = p_1 + p_2$. Since the induction hypothesis postulates that (27) holds for \tilde{Z} , it remains to show that the difference

$$\begin{aligned} (E(Z) - V_{I,f}(Z)) - (E(\tilde{Z}) - V_{I,f}(\tilde{Z})) &= (E(Z) - E(\tilde{Z})) - (V_{I,f}(Z) - V_{I,f}(\tilde{Z})) \\ &= -(x_2 - x_1)p_1 + (x_2 - x_1)(1 - f(1 - p_1)) = (x_2 - x_1)(1 - p_1 - f(1 - p_1)) \end{aligned} \quad (32)$$

is equal to the difference

$$\begin{aligned} &\int_0^1 \Psi_H(H^{-1}(1 - q))d\nu(q) - \int_0^1 \Psi_{\tilde{H}}(\tilde{H}^{-1}(1 - q))d\nu(q) \\ &= \int_{1-p_1-p_2}^{1-p_1} \Psi_H(H^{-1}(1 - q))d\nu(q) + \int_{1-p_1}^1 \Psi_H(H^{-1}(1 - q))d\nu(q) \\ &\quad - \int_{1-p_1-p_2}^1 \Psi_{\tilde{H}}(\tilde{H}^{-1}(1 - q))d\nu(q) \\ &= \Psi(x_2)\nu(1 - p_1 - p_2, 1 - p_1] + \Psi(x_1)\nu(1 - p_1, 1] - \Psi(x_2)\nu(1 - p_1 - p_2, 1] \\ &= (\Psi(x_1) - \Psi(x_2))\nu(1 - p_1, 1] = (x_2 - x_1)(1 - p_1)\nu(1 - p_1, 1]. \end{aligned} \quad (33)$$

The two are equal if and only if $\nu(1 - p_1, 1] = 1 - f(1 - p_1)/(1 - p_1)$, or $\nu[0, 1 - p_1) = f(1 - p_1)/(1 - p_1)$. Again, this follows from the definition of the probability measure ν .

This finished the proof of Lemma 5. Returning now to the proof of Theorem 5, it follows from (27) that

$$V_{I,f}(X) - V_{I,f}(Y) = E[X] - E[Y] + \int_0^1 [\Psi_G(G^{-1}(1 - q)) - \Psi_F(F^{-1}(1 - q))] d\nu(q). \quad (34)$$

Hence, if the expression in square brackets in (34) is non-negative, then $V_{I,f}(X) \geq V_{I,f}(Y)$ as soon as $E[X] \geq E[Y]$. But, as pointed out earlier, non-negativity of the integrand $\Psi_G(G^{-1}(1 - q)) - \Psi_F(F^{-1}(1 - q))$ for all q is one of the equivalent definitions of right-monotone increase in risk. In other words, decision makers with probability transformation function star-shaped at 0 are averse to (mean-reducing or mean-preserving) left-monotone increase in risk. This proves statement (ii) of Theorems 4 and 5.

To prove Theorem 5 (iii), suppose that G does not represent a right-monotone increase in risk with respect to F . We have to display a transformation function f supported from above by the identity function and star-shaped at 0, under which the Yaari functional shows strict preference for G . Under the current assumption, the clue to the construction of a transformation function under which G is strictly preferred is provided by expression (34). The set N of the previous case (see (24)) is replaced by the set $M \subseteq (0, 1)$ where $\Psi_G(G^{-1}(1 - q)) - \Psi_F(F^{-1}(1 - q))$ is strictly negative, a non-empty open set. In analogy to the previous case, let the measure ν in (34) be defined by

$$\nu(A) = \lambda I_A(0) + (1 - \lambda) \frac{\text{Leb}(A \cap M)}{\text{Leb}(M)},$$

for Borel subsets A of the unit interval, where $\lambda \in [0, 1)$ is a desired value of $\nu(\{0\}) = \lim_{x \downarrow 0} \frac{f(x)}{x}$ and I_A is the indicator function of the set A .

References

- [1] Arrow, K., 1971. *Essays in the theory of risk bearing*. North Holland Publishing Co: Amsterdam–London.
- [2] Bickel, P. J. and E. L. Lehmann, 1976. Descriptive statistics for non-parametric models, III. Dispersion. *Annals of Statistics*, **4**, 1139–1158.
- [3] Bickel, P. J. and E. L. Lehmann, 1979. Descriptive statistics for non-parametric models, IV. Spread. In *Contributions to Statistics*, Jureckova (Ed.), Reidel.
- [4] Chateauneuf, A., 1996. Decreasing inequality : an approach through non-additive models, *Cahiers Eco & Maths*, Paris I, 96.58.
- [5] Chateauneuf, A., 1991. On the use of capacities in modeling uncertainty aversion and risk aversion. *Journal of Mathematical Economics*, **20**, 343–369.

- [6] Chateauneuf, A., M. Cohen and R. Kast, 1997. A review of some results related to comonotonicity. *Cahiers Eco & Maths*, **97.32**, Université Paris I.
- [7] Chateauneuf, A., M. Cohen and I. Meilijson, 1997. More pessimism than greediness : a characterization of monotone risk aversion in the Rank Dependent Expected Utility model, *Cahiers Eco & Maths*, **97.53**, Université Paris I.
- [8] Chateauneuf, A., M. Cohen and Vergnaud, J. C., 2001. Left-monotone reduction of risk, deductible and call. Mimeo.
- [9] Chateauneuf, A. and P. Moyes, 2000. Inequality measurement and the weakening of the transfer principle. Working paper, CERMSEM and GRAPE.
- [10] Chew, S., E. Karni and Z. Safra, 1987. Risk aversion in the theory of expected utility with Rank Dependent preferences. *Journal of Economic Theory*, **42**, 370-381.
- [11] Chow, Y. S. and H. Robbins, 1963. On optimal stopping rules.
Z. Wahrscheinlichkeitstheorie, **2**, 33-49.
- [12] Cohen, M., 1995. Risk aversion concepts in expected- and non-expected-utility models. *The Geneva Papers on Risk and Insurance Theory*, **20**, 73-91.
- [13] Diamond, P. and J. Stiglitz, 1974. Increases in risk and risk aversion. *Journal of Economic Theory*, **8**, 337-360.
- [14] Gollier, C. and H. Schlesinger, 1996. Arrow's theorem on the optimality of deductibles: a stochastic dominance approach. *Economic Theory*, **7**, 359-364.
- [15] Hadar, J. and W. Russel, 1969. Rules for ordering uncertain prospects. *American Economic Review*, **59**, 25-34.
- [16] Hanoch, G. and H. Levy, 1969. The efficiency analysis of choices involving risk. *Review of Economic Studies*, **36**, 335-346.
- [17] Hardy, G. H., J. E. Littlewood and G. Polya, 1934. *Inequalities*. Cambridge University Press, (2nd edition, 1952).
- [18] Jewitt, I., 1987. Risk aversion and the choice between risky prospects: the preservation of comparative statics results. *Review of Economic Studies*, **54**, 73-85.
- [19] Jewitt, I., 1989. Choosing between risky prospects: the characterisation of comparative static results and location independent risk. *Management Science*, **35**, 60-70.
- [20] Karni, E. and A. Schwartz, 1977. Search Theory: the case of search with uncertain recall. *Journal of Economic Theory*, **16**, 38-52.

- [21] Karni, E. and A. Schwartz, 1977. Two theorems on optimal stopping with backward solicitation. *Journal of Applied Probability*, **14**, 869–875.
- [22] Landsberger, M. and I. Meilijson, 1990. Demand for Risky Financial Assets: A Portfolio Analysis. *Journal of Economic Theory*, **50**, 204–213.
- [23] Landsberger, M. and I. Meilijson, 1992. Lotteries, Insurance and Star-shaped utility functions. *Journal of Economic Theory*, **52**, 1–17.
- [24] Landsberger, M. and I. Meilijson, 1994. Comonotone allocations, Bickel-Lehmann dispersion and the Arrow-Pratt measure of risk aversion. *Annals of Operations Research*, **52**, 97–106.
- [25] Landsberger, M. and I. Meilijson, 1994. The generating process and an extension of Jewitt's location independent risk concept. *Management Science*, **40**, 662–669.
- [26] Lehmann, E., 1959. *Testing Statistical Hypotheses*. Wiley: New York.
- [27] Lorenz, M. O., 1905. Methods of measuring concentration of wealth, *J. of American Stat. Association*, **9**, 209–219.
- [28] Machina, M. J. and J. W. Pratt, 1997. Increasing risk: some direct constructions. *Journal of Risk and Uncertainty*, **14–15**, 103–127.
- [29] Marshall, A. W. and I. Olkin, 1979. *Inequalities: theory of majorization and its applications*. Academic Press: New York.
- [30] Moyes, P., 1994. Inequality reducing and inequality preserving transformations of incomes: symmetric and individualistic transformations. *Journal of Economic Theory*, **63**, 271–298.
- [31] Pratt, J. W., 1964. Risk Aversion in the small and large, *Econometrica*, **32**, 122–136.
- [32] Quiggin, J., 1982. A theory of anticipated utility, *Journal of Economic Behavior and Organization*, **3**, 323–343.
- [33] Quiggin, J., 1992. Increasing risk : another definition, in Chikan (Ed), *Progress in Decision Utility and Risk Theory*. Kluwer: Dordrecht.
- [34] Ross, S. 1981. Some stronger measures of risk aversion in the small and the large with applications, *Econometrica*, **49**, **3**, 621–638.
- [35] Rothschild, M. and J. Stiglitz, 1970. Increasing Risk: I. A definition. *Journal of Economic Theory*, **2**, 225–243.
- [36] Rothschild, M. and J. Stiglitz, 1970. Increasing Risk: II. Its economic consequences. *Journal of Economic Theory*, **3**, 66–84.

- [37] Strassen, V. 1965. The existence of probability measures with given marginals, *Ann. Math. Stat.*, **36**, 423-439.
- [38] Vergnaud, J. C. 1997. Analysis of risk in a non-expected utility framework and application to the optimality of the deductible. *Finance*, **18-1**, 155-167.
- [39] von Neumann, J. and O. Morgenstern, 1947. *Theory of games and economic behavior*, Princeton University Press, (3rd edition, 1954).
- [40] Wilson, 1968. The theory of syndicates, *Econometrica*, **36**, 119-138.
- [41] Yaari, M. E. 1984. Risk aversion without diminishing marginal utility. International Center for Economics and Related Disciplines, London.
- [42] Yaari, M. E. 1987. The dual theory of choice under risk. *Econometrica*, **55**, 95-115.
- [43] Yaari, M. E. 1988. A controversial proposal concerning inequality measurement. *Journal of Economic Theory*, **44**, 381-397.