"On Risk And Price: Stochastic Orderings And Measures "

Werner Hürlimann
Switzerland

Summary

Following the axiomatic approach to measures of statistical quantities initiated by van Zwet(1964) and developed by several other authors, we present a general axiomatic system for the measure of the quantities risk and price. We argue that risk and insurance price are closely related through the notion of risk loading, viewed as function of the measure of risk, and that risk should be closely related to the measures of scale, skewness and kurtosis. We consider "universal" measures of scale and risk, which can be adjusted for skewness and kurtosis. Concerning the measure of price, the distortion pricing principle introduced by Denneberg(1990), studied further by Wang(1996a/b), and justified axiomatically as insurance price in a competitive market setting by Wang et al.(1997), is a measure of price for our more general axiomatic system. Our presentation includes numerous examples, some of which have so far not been encountered in actuarial science.

Key words : axiomatic approach, measure of risk, measure of price, distortion pricing, stop-loss order, relative inverse convex order, scale, skewness, kurtosis

"Risque et Prix: Ordres Stochastiques et Mesures"

Werner Hürlimann
Switzerland

Résumé.

Suivant l'approche axiomatique pour les mesures de quantités statistiques initiée par van Zwet(1964) et développée par divers autres auteurs, nous présentons un système axiomatique général pour la mesure des quantités risque et prix. Nous supposons que le risque et le prix de l'assurance sont étroitement liés par la notion de surcharge de risque, conçue comme fonction de la mesure de risque, et que le risque est étroitement lié aux mesures d'échelle, d'asymétrie et d'élongation. Nous considérons des mesures universelles d'échelle et de risque, qui peuvent être ajustées pour l'asymétrie et l'élongation. En ce qui concerne la mesure de prix, le principe de distortion introduit par Denneberg(1990), qui a été étudié par Wang(1996a/b) et justifié axiomatiquement comme prix d'assurance dans un marché compétitif par Wang et al.(1997), est une mesure de prix pour notre système axiomatique plus général. Notre présentation inclut de nombreux examples, en partie inédits en sciences actuarielles.

<u>Mots clés</u> : approche axiomatique, mesure de risque, mesure de prix, principe de distortion, ordre stop-loss, ordre convexe relatif inverse, échelle, asymétrie, élongation

1. The formal approach based on orderings and measures.

It is often believed that science progresses through the identification of important concepts, which are first classified (e.g. risk), compared (e.g. riskier) and then quantified (e.g. measure of risk), where these three stages of development rarely follows the given temporal order (see e.g. Oja(81) who refers to Carnap(1962)).

Descriptive Statistics offers the theory and methods to identify *statistical quantities* associated with such concepts. In general, there is only a vague concept defining the statistical quantity, which can be formalized in many ways for practical use. This observation is typical for the important statistical properties of skewness and kurtosis associated to a distribution function (e.g. Balanda and MacGillivray(1988) and Groeneveld(1991)), and the same can be said about risk and price, for which a vague concept only begins to emerge in actuarial science and finance (e.g. Ramsay(1993)). In view of the different possible formalizations of risk and price, it seems more reasonable to accept "vague concepts" and develop a coherent structure of such formalizations (so-called axiomatic approach) rather than to concentrate on specific measures of risk and price.

The historical development of actuarial science reveals that insurance risks, and their associated premiums or prices, are closely related with the concepts of location, scale, skewness and kurtosis. Therefore, a "minimal" theory of actuarial risks and prices should be based on a thorough understanding of these four statistical quantities (cf. the comments by Tompkins, p.549, and Clarkson, p.597, in Howison et al.(1994), made in a finance context, but also valid in actuarial science). In our view, a coherent structure of risk and price should simultaneously discuss measures of all the relevant concepts as well as their interplay through common and diverging properties.

The axiomatic approach to measures of statistical quantities involves the following main steps (e.g. van Zwet(1964), Oja(1981), MacGillivray(1986), Balanda and MacGillivray(1988/90), among others) :

(i) Define *stochastic (partial) orders* on random variables or distribution functions, which allow for comparisons of the given statistical quantity.

(ii) Identify *measures of a statistical quantitiy* by considering functionals of distributions that preserve each of the plausible partial orders, and use only such measures in practical work.

(iii) Make only comparisons for *classes* or *families* of distributions, which are *totally ordered* with respect to the selected stochastic order.

(iv) Display the *hierarchy of* (possible) *stochastic orders* for a given statistical quantity. In particular, identify the weakest (resp. strongest) order(s), which should cover a largest (resp. smallest) possible class or family of distributions.

Following this fundamental philosophy, let us fix ideas by specifying the different axiomatic systems, which so far best fit the vague concepts behind the considered statistical quantities. Denote by $\leq$ a selected ordering of a given statistical quantity, and let F be a family of random variables X such that if $X, Y \in F$, then either $X \leq Y$ or $Y \leq X$.

Concerning the structure of location, scale, skewness and kurtosis, the basic "minimal" axiomatic framework proposed by Oja(1981) has been widely recognized in the statistical literature. In some applications, further axioms or/and requirements may be considered (e.g. Ruppert(1987), Groeneveld(1991)). In what follows, it is understood that the selected ordering $\le$ and family F will change and may vary for and within each definition.

Definition 1.1. The function $L: F \to R$ is a *measure of location* in F if

(Lo1) $L[aX+b]=aL[X]+b$ for all $a,b\in R$, X, $aX+b\in F$,

(Lo2) $L[X]\le L[Y]$ if $X,Y\in F$ and $X\le Y$.

Definition 1.2. The function $S: F \to R$ is a *measure of scale* in F if

(Sc1) $S[aX+b]=|a|\cdot S[X]$ for all $a,b\in R$, X, $aX+b\in F$,

(Sc2) $S[X]\le S[Y]$ if $X,Y\in F$ and $X\le Y$.

Definition 1.3. The function $\gamma: F \to R$ is a *measure of skewness* in F if

(Sk1) $\gamma[X]=0$ for all symmetric $X\in F$,

(Sk2) $\gamma[aX+b]=\gamma[X]$ for all $a,b\in R$, $a>0$, X, $aX+b\in F$,

(Sk3) $\gamma[-X]=-\gamma[X]$ for all $X,-X\in F$,

(Sk4) $\gamma[X]\le\gamma[Y]$ if $X,Y\in F$ and $X\le Y$.

Definition 1.4. The function $\gamma_2: F \to R$ is a *measure of kurtosis* in F if

(Ku1) $\gamma_2[aX+b]=\gamma_2[X]$ for all $a,b\in R$, $a\neq 0$, X, $aX+b\in F$,

(Ku2) $\gamma_2[X]\le\gamma_2[Y]$ if $X,Y\in F$ and $X\le Y$.

One notes that the defined measures of scale, skewness and kurtosis are all *location free* (or *location invariant*), that is invariant when one replaces X by $X+b$ for all $b\in R$. During the course of the development following Oja(1981), the need for *location dependent* measures has been advocated (e.g. MacGillivray(1986), Arnold and Groeneveld(1995)). These location dependent measures are defined as follows, where the functional $\theta_X = L[X]$, $X\in F$, represents some fixed *location parameter* θ (e.g. mean, median, mode, etc.), the value taken by some measure of location.

Definition 1.2'. The function $S: F \to R$ is a *measure of scale with respect to* θ in F if

(Sc1') $S[aX;\theta_{aX}]=|a|\cdot S[X;\theta_X]$ for all $a,b\in R$, X, $aX\in F$,

(Sc2') $S[X;\theta_X]\le S[Y;\theta_Y]$ if $X,Y\in F$ and $X\le Y$.

Definition 1.3'. The function $\gamma: F \to R$ is a *measure of skewness with respect to* θ in F if

(Sk1') $\gamma[X;\theta_X]=0$ for all symmetric $X\in F$,

(Sk2') $\gamma[aX;\theta_{aX}]=\gamma[X;\theta_X]$ for all $a,b\in R$, $a>0$, X, $aX\in F$,

(Sk3') $\gamma[-X;\theta_{-X}]=-\gamma[X;\theta_X]$ for all $X,-X\in F$, (Sk4') $\gamma[X;\theta_X] \le \gamma[Y;\theta_Y]$ if $X, Y \in \mathcal{F}$ and $X \le Y$.

Definition 1.4'. The function $\gamma_2 : \mathcal{F} \to R$ is a *measure of kurtosis with respect to* θ in $\mathcal{F}$ if

(Ku1') $\gamma_2[aX;\theta_{aX}] = \gamma_2[X;\theta_X]$ for all $a,b \in R$, $a \ne 0$, $X, aX \in \mathcal{F}$,

(Ku2') $\gamma_2[X;\theta_X] \le \gamma_2[Y;\theta_Y]$ if $X, Y \in \mathcal{F}$ and $X \le Y$.

Since there is in general no agreement on what risk measures are, and how prices are built, universally accepted axiomatic systems for the structures of risk and price do not yet seem to exist in actuarial science and finance. It appears reasonable to start from "minimal" sets of axioms that are believed most important, and then refine or replace non-appropriate axioms by others if inconsistencies are revealed either by theoretical or practical work.

For the structure of risk, we distinguish between "absolute" and "relative" risk (Garrido(1993), Hürlimann(1992/95)). We assume that $\mathcal{F}$ contains only non-negative random variables. A precise measure of (actuarial) risk formalizes the intuitive notion of *actuarial risk*, understood as variation and uncertainty in potential future insurance losses (e.g. Ramsay(1993)). This notion must not be confounded with the intuitive notion of *financial risk*, which is interpreted as the capital investment required to render a future value of a position acceptable, and which can be formalized similarly by defining a precise measure of (financial) risk in the sense of Arztner et al.(1997a/b).

Definition 1.5. The function $R : \mathcal{F} \to R$ is a *measure of (absolute) risk* in $\mathcal{F}$ if

(aR1) $R[X] = 0$ if $X \in \mathcal{F}$ is riskless, that is $\Pr(X = \mu) = 1$ for some constant μ

(aR2) $R[X] \ge 0$ for all $X \in \mathcal{F}$,

(aR3) $R[X+Y] \le R[X] + R[Y]$ for all $X, Y \in \mathcal{F}$ such that X, Y are independent random variables and $X+Y \in \mathcal{F}$,

(aR4) $R[aX+b] = a \cdot R[X]$ for all $a,b \in R$, $a > 0$, and X, $aX + b \in \mathcal{F}$,

(aR5) $R[X] \le R[Y]$ if $X, Y \in \mathcal{F}$ and $X \le Y$.

Though we will not use it in the present work, it seems that a corresponding structure of "relative" risk should satisfy the following axioms (cf. Garrido(1993) for the property (rR4)).

Definition 1.6. The function $r : \mathcal{F} \to R$ is a *measure of relative risk* in $\mathcal{F}$ if

(rR1) $r[X] = 0$ if $X \in \mathcal{F}$ is riskless,

(rR2) $r[X] \ge 0$ for all $X \in \mathcal{F}$,

(rR3) $r[X+Y] \le r[X] + r[Y]$ for all $X, Y \in \mathcal{F}$ such that X, Y are independent random variables and $X+Y \in \mathcal{F}$,

(rR4) $r[aX] = r[X]$ for all $a > 0$, X, $aX \in \mathcal{F}$,

(rR5) $r[X] \le r[Y]$ if $X, Y \in \mathcal{F}$ and $X \le Y$.

The structure of price for absolute or/and relative risks is presumably quite complex because it should take into consideration all of the preceding structures. For simplicity, we will just focus on price for absolute risks. A precise definition of a measure of price formalizes the intuitive notion of price, which is understood as a certainty equivalent of risk. To be applicable in actuarial science, a "minimal" set of axioms for a coherent structure of price should (at least) contain the following plausible requirements (e.g. Denneberg(1990)).

Definition 1.7. The function $P: \mathcal{F} \to R$ is a *measure of price* in $\mathcal{F}$ if

(P1) $P[X] \geq E[X]$ if $X \in \mathcal{F}$ is non-negative,

(P2) $P[X] \leq \sup[X]$ if $X \in \mathcal{F}$ is non-negative,

(P3) $P[aX+b] = a \cdot P[X] + b$ for all $a, b \in R$, $a > 0$, $X, aX+b \in \mathcal{F}$,

(P4) $P[X+Y] \leq P[X] + P[Y]$ for all $X, Y \in \mathcal{F}$ such that $X+Y \in \mathcal{F}$,

(P5) $P[X] \leq P[Y]$ if $X, Y \in \mathcal{F}$ and $X \leq Y$.

Let us follow these definitions by a brief outline of the content of our study together with some significant actuarial motivation. In Section 2, a review of the class of stop-loss transform orders is given. These orders are of basic importance in risk and price theory because they essentially provide the possible ordering of risks, under which measures of risk and price should be preserved. Then we describe in Section 3 the class of relative inverse convex orders of arbitrary degree, whose lower degree orders are essential for the definition of coherent measures of location, scale, skewness and kurtosis. Section 4 is a summary about some main facts concerning the measures of scale, skewness and kurtosis, including some signigicant examples. The actuarial main part of the subject is an illustrative review about measures of risk (Section 5) and price (Section 6). It is argued that actuaries describe insurance price roughly as "expected cost plus a safety loading in form of a function of the measure of risk", and that risk should be closely related to scale, skewness and kurtosis. In particular, it is of interest to analyze when a measure of scale, which preserves an ordering of scale, also preserves an ordering of risk. Some counterexamples are analyzed in Section 5.1, and "universal" examples of simultaneous measures of scale and risk are displayed in Section 5.2. As a novelty, we show in Section 5.3 how to adjust a measure of scale and risk for positive skewness. As a concrete example, we adjust the median absolute deviation measure of risk for the modified Yule measure of skewness. Section 6 contains a short survey of the main pricing principles, which define measures of price in the sense of Definition 1.7. We recall that the distortion pricing principle introduced by Denneberg(1990), studied further by Wang(1996a/b), and justified axiomatically as insurance price in a competitive market setting by Wang et al.(1997), is a measure of price in our more general axiomatic approach to pricing theory (see also Hürlimann(1998c) for complements on this approach). As a recent new example of distortion pricing measure, we mention an "entropy" measure of price, which has been motivated from considerations in option pricing theory (see also Hürlimann(1997c/98d)). Finally, we conclude with a measure of price derived from the measure of risk adjusted for positive skewness, which has been described in Example 5.8.

2. The class of stop-loss transform orders.

For simplicity, let *X, Y* be random variables with absolutely continuous and strictly increasing distributions $F_X(x), F_Y(x)$, which are defined on the supports S_F, S_G. For each n=0,1,2,..., the *degree* n *stop-loss transform* of X consists of the collection of partial moments of order n given by $\pi_X^n(x) = E\left[(X-x)_+^n\right]$, $x \in R$. The convention is made that

$\pi_X^0(x) = \overline{F}(x) = 1 - F(x)$ is the survival function of X. For n=1 the function $\pi_X^1(x)$ is the usual stop-loss transform $\pi_X(x)$, written without upper index.

Recall the class of higher degree stop-loss orders (e.g. van Heerwaarden(1991), Section 5, or Kaas et al.(1994), chap. V).

Definition 2.1. The random variable X precedes Y in the *degree n stop-loss order*, written $X \le_{sl}^{(n)} Y$, if the moments of order n are finite and the following conditions are satisfied :

(2.1) $$E\left[X^k\right] \le E\left[Y^k\right], \quad k = 1,\dots,n-1,$$

(2.2) $$\pi_X^n(x) \le \pi_Y^n(x), \text{ uniformly for all } x \in R.$$

For n=0 the relation $\le_{sl}^{(0)}$ is identical to $\le_{st}$. The order $\le_{sl}^{(1)}$ coincides with $\le_{sl}$ or $\le_{icx}$. The restriction (2.1) is required for the characterization as common preferences of a group of decision makers with increasingly regular utility functions (e.g. Kaas et al.(1994), Theorem V.2.1). For this reason, the class of stop-loss transform orders is highly significant in actuarial science, finance and economics, and plays a fundamental role in the definition of coherent structures of risk and price. Also, by equal means and variances, and possibly higher order moments, the usual stochastic and stop-loss order comparisons of two random variables do not apply. In this situation, a higher degree stop-loss order comparison can be useful (e.g. Hürlimann(1995), Section 6, Kaas et al.(1995)). In applications, to establish stop-loss order comparison properties, one requires some fundamental facts and equivalent characterizations, which are described below (see also Hürlimann(1998b)).

First of all, the following well-known elementary equivalent statements hold :

(SL1) $X \le_{sl} Y$

(SL2) $E[\varphi(X)] \le E[\varphi(Y)]$ for all increasing convex functions $\varphi(x)$

(SL3) $E[\max(x,X)] \le E[\max(x,Y)]$ uniformly for all $x \in R$

Furthermore, recall the following famous and widely known sufficient condition.

Lemma 2.1. (Karlin-Novikoff(1963), Ohlin(1969)). Let X and Y be random variables with finite means such that $\mu_X \le \mu_Y$, and there exists c such that $F_X(x) \le F_Y(x)$, for $x \le c$, and $F_X(x) \ge F_Y(x)$, for $x > c$. Then X precedes Y in *dangerousness order*, written $X \le_D Y$, and this implies the stop-loss order $X \le_{sl} Y$.

Through application of appropriate limiting arguments, it is often possible to restrict the attention to random variables, which belong to the large set S, which consists of all non-negative random variables with finite means, such that the distribution functions of any two of them cross finitely many times (*finite crossing condition*).

A generalized version of Lemma 2.1 is the sign-change characterization of the stop-loss order (Taylor(1983), Stoyan(1977)).

Theorem 2.1. Let $X, Y \in S$ such that the distributions cross n≥1 times in the crossing points $t_1 < t_2 < .. < t_n$. Then one has $X \le_{sl} Y$ if, and only if, one of the following is fulfilled :

Case 1 : The first sign change of the difference $F_Y(x)-F_X(x)$ occurs from $-$ to $+$, there is an even number of crossing points n=2m, and one has the inequalities

(2.3) $$\pi_X(t_{2j-1}) \le \pi_Y(t_{2j-1}), j=1,...,m$$

Case 2 : The first sign change of the difference $F_Y(x)-F_X(x)$ occurs from $+$ to $-$, there is an odd number of crossing points n=2m+1, and one has the inequalities

(2.4) $$\mu_X \le \mu_Y, \quad \pi_X(t_{2j}) \le \pi_Y(t_{2j}), \quad j=1,...,m$$

Proof. This is shown in Hürlimann(1998b). ◊

The condition of Lemma 2.1 is not a transitive relation. The *transitive (stop-loss)-closure* of the order $\le_D$, denoted by $\le_{D*}$, which is defined as the smallest partial order containing all pairs (X,Y) with $X \le_D Y$ as a subset, identifies with the stop-loss order (Kaas and Heerwaarden(1992), Müller(1996)). For finitely many sign changes one has the result.

Theorem 2.2. Let $X, Y \in S$ such that $X \le_{sl} Y$. Then there exists a finite sequence of random variables $Z_1, Z_2,..., Z_n$ such that $X = Z_1, Y = Z_n$ and $Z_i \le_D Z_{i+1}$ for all i=1,...,n–1.

Proof. This is Kaas et al.(1994), Theorem III.1.3. Alternatively, the ordered sequences (2.8) and (2.12) in Hürlimann(1998b) yield a more detailed constructive proof of this result. ◊

The stop-loss order separates as follows.

Theorem 2.3. If $X \le_{sl} Y$, then there exists a random variable Z such that $X \le_{st} Z \le_{sl,=} Y$.

Proof. Proofs are given by Kaas et al.(1994), Theorem IV.2.1, Makowski(1994), Shaked and Shanthikumar(1994), Theorem 3.A.3, and Müller(1996), Theorem 3.7. ◊

Other characterizations of the stop-loss order can be obtained. For a random variable X with finite mean and quantile function $F_X^{-1}(u)$, the *Hardy-Littlewood transform* X^H of X is defined by its quantile function on $[0,1]$ through the formula

(2.5) $$(F_X^H)^{-1}(u) = \begin{cases} \frac{1}{1-u}\int_u^1 F_X^{-1}(v)dv, & u<1, \\ F_X^{-1}(u), & u=1. \end{cases}$$

Its name stems from the Hardy-Littlewood(1930) maximal function. The random variable X^H is the least majorant with respect to $\le_{st}$ among all random variables $Y \le_{sl} X$ (e.g. Meilijson and Nàdas(1979)). Its great importance in applied probability and related fields has been noticed by several further authors, among others Blackwell and Dubins(1963), Dubins and Gilat(1978), Rüschendorf(1991), and Kertz and Rösler(1990/92/93). Recent actuarial applications are found in Hürlimann(1997a/c).

Theorem 2.4. For i=1,2, let $X_i \in S$ with finite means. Then one has $X_1 \le_{sl} X_2$ if, and only if, one has $X_1^H \le_{st} X_2^H$.

Proof. This is Lemma 1.8 in Kertz and Rösler(1992) (see also Hürlimann(1998b)). ◊

By existence of a common mean, the characterization of the convex order $X_1 \le_{cx} X_2 \quad \Leftrightarrow \quad X_1^H \le_{st} X_2^H$ is found in van der Vecht(1986), p.69. For any non-negative random variable X with fixed mean μ, consider the *integrated tail* transform X^I of X with survival function defined by

$$(2.6) \qquad \overline{F}_X^I(x) = \frac{\pi_X(x)}{\mu}, \quad x \ge 0,$$

which plays an important role in actuarial ruin models (e.g. Embrechts et al.(1997)). In renewal theory (2.6) is called *stationary renewal distribution*.

Theorem 2.5. For i=1,2, let X_i be non-negative random variables with common finite mean. Then $X_1 \le_{sl,=}^{(n)} X_2 \quad \Leftrightarrow \quad X_1^I \le_{sl,=}^{(n-1)} X_2^I$ for all $n = 1,2,3,\ldots$.

Proof. Consult for example van Heerwaarden(1991), p.69. ◊

An alternative characterization goes back to Blackwell(1953). A bivariate real function $T : R^2 \to [0,1]$ is called *Markov kernel* if for each $x \in R$ the function $T(x,y)$ of y defines a distribution function. The function T is called *mean preserving Markov kernel* if additionally the mean value of $T(x,y)$ by fixed x is preserved, that is

$$(2.7) \qquad \int_{-\infty}^{\infty} y\,dT(x,y) = x, \text{ for all } x \in R.$$

The function T associates to X the transformed random variable $Y = T(X)$ with

$$(2.8) \qquad F_Y(x) = \int_{-\infty}^{\infty} T(x,y)\,dF_X(x) = E[T(X,y)].$$

Theorem 2.6. Let X and Y with a common finite mean. Then one has $X \le_{cx} Y$ if, and only if, there exists a mean preserving Markov kernel $T(x,y)$ such that $Y = T(X)$.

Proof. The sufficient part is immediate. The necessary condition is easily verified for discrete distributions (see Szekli(1995)) while a general proof is beyond elementary mathematics (see Alfsen(1971)). ◊

A more recent characterization of the convex order relies on the mathematical notion of *fusion* for probability measures as studied by Elton and Hill(1992). For more details on this, the interested reader should consult Szekli(1995), Theorem 1.3D.

3. The class of relative inverse convex orders.

The class of higher degree stop-loss transform orders is of fundamental importance in risk and price theory. There is a second class of increasingly more complex orders, which is of similar importance in the modern theory of descriptive statistics.

For simplicity, let X, Y be random variables with absolutely continuous and strictly increasing distributions $F(x)$, $G(x)$, densities $f(x)$, $g(x)$, which are defined on the supports S_F, S_G. The *relative inverse* function of F with respect to G (G with respect to F) is defined by $R_{X,Y}(x) = G^{-1}F(x), x \in S_F$ ($S_{X,Y}(x) = R_{X,Y}^{-1}(x) = F^{-1}G(x), x \in S_G$), and is viewed as a real functional of the pair (X,Y). It is useful to set $D_{X,Y} = R_{X,Y}(x) - x$ and $E_{X,Y} = S_{X,Y}(x) - x$ (which correspond to $\Delta(x)$ and $\Delta^*(x)$ in Oja(1981)).

Recall the notion of *higher degree convexity* (e.g. Karlin(1968), p.23, and Karlin and Ziegler(1976)).

Definition 3.1. For a real interval I, a function $f : I \to R$ is called *convex (concave) of degree* 0 if

$$(\mathrm{C}_0) \qquad f(x) \le (\ge) 0 \quad \text{for all } x \in I$$

It is called *convex (concave) of degree* 1 if the determinant

$$(\mathrm{C}_1) \qquad \begin{vmatrix} 1 & 1 \\ f(x_1) & f(x_2) \end{vmatrix} \ge (\le) 0 \quad \text{for all } x_1 < x_2, x_1, x_2 \in I,$$

and *convex (concave) of degree* n≥2 if the determinant

$$(\mathrm{C}_n) \qquad \begin{vmatrix} 1 & 1 & & 1 \\ x_1 & x_2 & & x_{n+1} \\ & & & \\ x_1^{n-1} & x_2^{n-1} & & x_{n+1}^{n-1} \\ f(x_1) & f(x_2) & & f(x_{n+1}) \end{vmatrix} \ge (\le) 0 \quad \text{for all } x_1 < x_2 < \ldots < x_{n+1}, \quad x_1, x_2, \ldots, x_{n+1} \in I.$$

For the lower degrees, one notes that the statement f is convex (concave) of degree 0 says that f is *non-negative (non-positive)*, f is convex (concave) of degree 1 means that f is *non-decreasing (non-increasing)* and f is convex (concave) of degree 2 exactly when f is *convex (concave)* (Karlin(1968), p.280-83). As a special case, note that if the n-th derivative exists, then f is convex (concave) of degree n if and only if $f^{(n)}(x) \ge (\le) 0$.

The class of higher degree relative inverse convex orders has been introduced by Oja(1981) as follows.

Definition 3.2. The random variable X precedes Y in the *degree* n *relative inverse convex order*, written $X \le_n Y$, if the function $D_{X,Y}(x)$ is convex of degree n.

The significance of the lower degree orders for $n = 0,1,2$ lies in the fundamental role they take in clarifying coherent structures of location, scale, skewness and kurtosis (see Section 4). In the following, we present some useful alternative characterizations and relate these orders to alternative order concepts and other notions. The given facts will be exploited later throughout.

The order $\le_0$ is identical to $\le_{st}$ while $\le_1$, introduced by Doksum(1969), coincides with the dispersion order $\le^{disp}$ introduced by Bickel and Lehmann(1979).

Definition 3.3. The random variable X precedes Y in the *dispersion order*, written $X \le^{disp} Y$, if $F^{-1}(\beta) - F^{-1}(\alpha) \le G^{-1}(\beta) - G^{-1}(\alpha)$ for all $0 < \alpha < \beta < 1$.

This order can be characterized through sign change properties, where the number of sign changes is described by the function $S^-(\cdot)$. The notation $f_c(x) = f(x+c)$ is used.

Theorem 3.1. Let X and Y be random variables with absolutely continuous and strictly increasing distributions. The following statements are equivalent :

(a) $X \le_1 Y$

(b) $X \le^{disp} Y$

(c) $S^-(G_c - F) \le 1$ for all $c \in R$, with a sign sequence $(+,-)$ if a sign change occurs

(d) $R'_{X,Y}(x) \ge 1$ for all $x \in S_F$

(e) The density-quantile functions satisfy the inequality $f\left[F^{-1}(u)\right] \ge g\left[G^{-1}(u)\right]$, $u \in (0,1)$

(f) The distribution-density-inverse functions satisfy the inequality $F\left[f^{-1}(x)\right] \ge G\left[g^{-1}(x)\right]$ provided f and g are decreasing functions

(g) $(X - F^{-1}(u))_+ \le_{st} (Y - G^{-1}(u))_+$ for all $u \in (0,1)$

(h) If X and Y are non-negative random variables, (a)-(g) are equivalent to $X \le_{st} Y$ and $S^-(G_c - F) \le 1$ for all $c \in R_+$, with a sign sequence $(+,-)$ if a sign change occurs

Proof. The equivalence of (a) and (b) is shown in Deshpande and Kochar(1983). Similarly simple is the equivalence of (a) and (d). The criteria (e) and (f) are restatements of (d). The equivalence of (b) and (c) is in Shaked(1982) (also Szekli(1995), p.29). The equivalence of (b) and (g) is in Muñoz-Pérez(1990). The equivalence of (b) and (h) is in Shaked(1982). ◊

The following sufficient (but not necessary) criterion is often useful.

Theorem 3.2. Let X and Y be absolutely continuous random variables with densities f and g. If $S^-(g_c - f) \le 2$ for all $c \in R$, with a sign sequence $(+,-,+)$ in case of equality, then one has $X \le^{disp} Y$.

Proof. A proof is given in Szekli(1995), p.30. ◊

An elementary general fact, which is useful in both risk and price theory involves comonotone random variables, a notion introduced by Schmeidler(1986) and Yaari(1987).

Definition 3.4. Two random variables X and Y are called *comonotone* if there exists Z and weakly increasing functions u and v such that $X = u(Z), \quad Y = v(Z)$.

Theorem 3.3. If X and Y are comonotone, then one has $X, Y \le^{disp} X + Y$.

Proof. This follows from Definition 3.3 using the well-known fact $F_{X+Y}^{-1}(u) = F_X^{-1}(u) + F_Y^{-1}(u)$ (e.g. Denneberg(1990/94), Landsberger and Meilijson(1994)). ◊

There exist a number of interesting results relating the dispersion order with some aging classes (e.g. Bartoszewicz(1985/87), Bagai and Kochar(1986)).

Theorem 3.4. If X and Y are non-negative random variables, one has the statements :

(a) If $X \le_{hr} Y$ and X or Y is of class DFR, then $X \le^{disp} Y$.

(b) If $X \le^{disp} Y$ and X or Y is of class IFR, then $X \le_{hr} Y$

(c) Let X be of class NBU (new better than used) and Y of class NWU (new worse than used). Then one has $X \le^{disp} Y$ if and only if $X \le_{hr} Y$

Proof. Bartoszewicz(1987), Prop. 1, 2, Bagai/Kochar(1986), Theorem 2.1, Corollary 2.1. ◊

The order $\le_2$ coincides with "convex order" introduced by van Zwet(1964) denoted by $\le_c$. Setting $G_{a,b}(x) = G(ax+b)$, an equivalent condition is that $S^-(G_{a,b} - F) \le 2$ for all $a, b \in R$ with a sign sequence $(-,+,-)$ if two sign changes occur.

Besides its use as ordering of skewness (see Section 4), the stochastic order $\le_c$ is of importance in reliability theory and related to the starshaped and superadditive orders. Let X and Y be non-negative random variables with distributions $F(x)$ and $G(x)$ such that $F(0) = G(0) = 0$, which have finite means μ_X, μ_Y. For the class of functions $\varphi : R_+ \to R_+$, which are continuous such that $\varphi(0) = 0$, one says that φ is *starshaped* if for each $\alpha \in [0,1]$, and all x, one has $\varphi(\alpha x) \le \alpha \varphi(x)$. The function φ is *superadditive* if $\varphi(x+y) \ge \varphi(x) + \varphi(y)$ for all x and y. When applied to the relative inverse function $\varphi(x) = R_{X,Y}(x) = G^{-1}F(x)$, one obtains the following stochastic orders. As already stated X precedes Y in the "convex" order, written $X \le_c Y$, if $R_{X,Y}(x)$ is convex. Similarly X precedes Y in the *starshaped order*, written $X \le_* Y$, if $R_{X,Y}(x)$ is starshaped, and X precedes Y in the *superadditive order*, written $X \le_{su} Y$, if $R_{X,Y}(x)$ is superadditive. Since the convex property implies the starshaped property, which implies itself the superadditive property (e.g. Bruckner and Ostrow(1962)), these stochastic orders are increasingly weaker in the sense that $\le_c \Rightarrow \le_* \Rightarrow \le_{su}$. It is worthwile to mention the following result.

Theorem 3.5. Let X and Y be non-negative random variables with finite means $\mu_X \le \mu_Y$. If $X \le_* Y$ (or stronger $X \le_c Y$) then $X \le_{sl} Y$.

Proof. Consult Dharmadhikari and Joag-Dev(1988), Theorem 9.3, and Szekli(1995), p.23. ◊

4. Measures of scale, skewness and kurtosis.

Measures of statistical quantities based on divers stochastic orders are relatively long known. The usual stochastic order $\leq_{st} \equiv \leq_0$ has been introduced by Mann and Whitney(1947) and studied further by Lehmann(1955). As an ordering of location, it has been used by Bickel and Lehmann(1975) to define measures of location for asymmetric random variables. These authors(1976/79) also introduced $\leq^{disp} \equiv \leq_1$ as an ordering of scale. Van Zwet's(1964) ordering $\leq_c \equiv \leq_2$ is used as an ordering of skewness, and a modification of it is used as an ordering of kurtosis for symmetric random variables. Oja(1981) unified the previous work and introduced some weaker stochastic orders of scale, skewness and kurtosis. The subsequent papers by Groeneveld and Meeden(1984), MacGillivray(1986), Balanda and MacGillivray(1988/90), Groeneveld(1991) and Arnold and Groeneveld(1995) discuss how to refine and obtain "complete" structures of scale, skewness and kurtosis. These results are of interest for defining measures of risk and price, as exemplified in Section 5 and 6.

4.1. Measures of scale.

For the ordering of scale $\leq_1 \equiv \leq^{disp}$, the location parameter may be arbitrary, a fact which is reflected in the property (c) of Theorem 3.1. Fixing the particular location parameter around which scale is taken necessarily weakens the order $\leq_1$. Though there is a huge number of possible location parameters (e.g. Andrews et al.(1972) study 58 different location estimators), let us restrict our attention to three of the most important ones, namely the mean μ_X, the median m_X and the mode M_X of a given random variable X (cf. Hutchinson(1993)). In all definitions, we assume that these quantities are well-defined, and consider only random variables for which this is the case.

Definition 4.1. (Oja(1981)) The random variable X precedes Y in the *order of mean-scale*, written $X \leq_1^{\mu} Y$, if one of the following equivalent statements holds :

(μS1) $S^-(G_{\mu_Y} - F_{\mu_X}) = 1$ with a sign sequence $(+,-)$

(μS2) $X - \mu_X \leq_{D,=} Y - \mu_Y$

(μS3) There exists c such that $D_{X,Y}(x) \leq (\geq) \mu_Y - \mu_X$ for all $x \leq (\geq) c$

(μS4) There exists $\xi \in (0,1)$ such that $G^{-1}(u) - F^{-1}(u) \leq (\geq) \mu_Y - \mu_X$ for all $u \in (0,1)$ such that $u \leq (\geq) \xi$

Definition 4.2. (MacGillivray(1986)) The random variable X precedes Y in the *order of median-scale*, written $X \leq_1^m Y$, if one of the following holds :

(mS1) $D_{X,Y}(x) \leq (\geq) m_Y - m_X$ for all $x \leq (\geq) m_X$

(mS2) $G^{-1}(u) - F^{-1}(u) \leq (\geq) m_Y - m_X$ for all $u \in (0,1)$ such that $u \leq (\geq) \frac{1}{2}$

Note that in contrast to Definition 4.1, the crossing point is necessarily m_X (resp. $\frac{1}{2}$) because if $X \leq^{disp} Y$ then $G^{-1}(u) - F^{-1}(u)$ is necessarily increasing in u.

Definition 4.3. The random variable X precedes Y in the *order of mode-scale*, written $X \le_1^M Y$, if one of the following holds :

(MS1) There exists c such that $D_{X,Y}(x) \le (\ge) M_Y - M_X$ for all $x \le (\ge) c$

(MS2) There exists $\xi \in (0,1)$ such that $G^{-1}(u) - F^{-1}(u) \le (\ge) M_Y - M_X$ for all $u \in (0,1)$ such that $u \le (\ge) \xi$

(MS3) $S^-(G_{M_Y} - F_{M_X}) = 1$ with a sign sequence $(+,-)$

The equivalent characterizations of $\le_1$ given in Theorem 3.1 yield the following result.

Theorem 4.1. (*Hierarchy of orderings of scale*) Each of the partial orders $X \le_1^\mu Y$, $X \le_1^m Y$ and $X \le_1^M Y$ is implied by the location free order of scale $X \le_1 Y$.

Remark 4.1. The order $\le_1^{**}$ of Oja(1981) weakens further $\le_1^\mu$ but seems less significant.

For each of the orderings of scale, it is possible to describe measures of scale together with totally ordered families of random variables for which these measures apply. Only a small list will suffice for our purpose. In view of Theorem 4.1, each measure of scale with respect to one of $\le_1^\mu, \le_1^m$ and $\le_1^M$ is also a measure of scale for the stronger order $\le_1$.

Examples 4.1 : location free and location dependent measures of scale

(1) location free measures

(1a) The Gini measure $Gini[X] = \frac{1}{2} E[|X - Y|]$, where X and Y are independent and identically distributed (see Oja(1981)).

(1b) The class of density-quantile measures $DQ_\varphi[X] = E[-(\varphi \circ f)(X)]$, where $f = F'$ is the density and $\varphi(x)$ is a monotone increasing function defined on $f(S_F)$ (use criterion (e) of Theorem 3.1). In the special case $\varphi(x) = \ln(x)$ one recovers the measure of entropy $En[X] = E[-\ln\{f(X)\}]$ mentioned in Oja(1981).

(1c) The class of survival-density-inverse measures $SD_\varphi[X] = E[-\frac{d}{dx}(\varphi \circ \overline{F})(X)]$, where $\varphi(x)$ is a monotone increasing function defined on $(0,1)$, and X is a non-negative random variable with decreasing density f (use (f) of Theorem 3.1). In the special case $\varphi(x) = \ln(x)$ one obtains the measure of hazard $Eh[X] = E[h(X)]$, where $h(x)$ is the hazard rate.

(2) mean dependent measures

By Definition 4.1, statement (µS2), the relation $X \le_1^\mu Y$ implies $X - \mu_X \le_{sl,=} Y - \mu_Y$ (by the Karlin-Novikoff(1963) cut-criterion or dangerousness order relation). It follows that any functional $S_{\mu,\varphi}[X] = E[\varphi(X - \mu_X)]$, where $\varphi(x)$ is a convex function, preserves the ordering of scale $\le_1^{\mu}$. The choice of $\varphi(x)$ is further restricted by the scaling axiom (Sc1') in Definition 1.2'. As special cases one has :

(2a) The standard deviation $\sigma_{\mu}[X]=\sqrt{E[(X-\mu_X)^2]}$, which is Corollary 4.2 in Oja(1981).

(2b) The "stop-loss at mean" measure or "one half of the mean absolute deviation" measure $\pi_{\mu}[X]=E[(X-\mu_X)_+]=\frac{1}{2}E[|X-\mu_X|]$.

(3) <u>median dependent measures</u>

Appropriate functionals, which preserve $\le_1^{m}$ are of the form $S_{m,\varphi}[X]=\int_0^1 \varphi\{|F^{-1}(u)-m_X|\}du$, where $\varphi(x)$ is a monotone increasing function on $(0,1)$. As an example one has the absolute deviation from the median defined by $\sigma_m[X]=E[|X-m_X|]$.

(4) <u>mode dependent measures</u>

By Definition 4.3, statement (MS3), the dangerousness order relation $X-M_X \le_D Y-M_Y$, which implies $X-M_X \le_{sl} Y-M_Y$, is fulfilled provided $M_Y-M_X \le \mu_Y-\mu_X$. Under the latter mean-mode condition, any functional $S_{M,\varphi}[X]=E[\varphi(X-M_X)]$, where $\varphi(x)$ is an increasing convex function, preserves the ordering $\le_1^{M}$. Measures of this type include :

(4a) The square root of the quadratic deviation from the mode $\sigma_M[X]=\sqrt{E[(X-M_X)^2]}$.

(4b) The "stop-loss at mode" measure $\pi_M[X]=E[(X-M_X)_+]$.

4.2. Measures of skewness.

For weakening the ordering of skewness $\le_c \equiv \le_2$, the location parameter may be chosen arbitrarily, and there is also much freedom in the choice of the remaining scale parameter. Due to its complexity, an exhaustive classification has not yet been given in the statistical literature. For this reason, we restrict our attention to some main situations.

Definition 4.4. (MacGillivray(1986), refined ordering $\le_2^*$ of Oja(1981)) The random variable X precedes Y in the *order of mean-skewness*, written $X \le_2^{\mu} Y$, if the standardized distribution difference $d_{X,Y}(x)=G(\mu_Y+\sigma_Y x)-F(\mu_X+\sigma_X x)$ is either identically zero or changes sign twice with a sign sequence $(-,+,-)$.

The next four orderings of skewness, in other nomenclature, were all proposed by MacGillivray(1986).

Definition 4.5. The random variable X precedes Y in the *order of median-star-skewness*, written $X \le_{2,*}^{m} Y$, if $\left(R_{X,Y}(x)-m_Y\right)/\left(x-m_X\right)$ is non-decreasing in S_F, or equivalently $\left(G^{-1}(u)-m_Y\right)/\left(F^{-1}(u)-m_X\right)$ is nondecreasing in $(0,1)$.

Definition 4.6. The random variable X precedes Y in the *order of median-Doksum-skewness*, written $X \le_{2,D}^{m} Y$, if $R_{X,Y}(x)-\frac{f(m_X)}{g(m_Y)}x$ is nonincreasing (nondecreasing) for all $x \in S_F$ such that $x<(\ge) m_X$, or equivalently $G^{-1}(u)-\frac{f(m_X)}{g(m_Y)}F^{-1}(u)$ is nonincreasing (nondecreasing) for all $u \in (0,1)$ such that $u<(\ge)\frac{1}{2}$.

Definition 4.7. The random variable X precedes Y in the *order of median-skewness*, written $X \le_{2}^{m} Y$, if $\left(R_{X,Y}(x)-m_Y\right)\cdot g(m_Y) \ge \left(x-m_X\right)\cdot f(m_X)$ for all $x \in S_F$ or equivalently $\left(G^{-1}(u)-m_Y\right)\cdot g(m_Y) \ge \left(F^{-1}(u)-m_X\right)\cdot f(m_X)$ for all $u \in (0,1)$.

Definition 4.8. The random variable X precedes Y in the *order of median-quantile-skewness*, written $X \le_{2,\gamma}^{m} Y$, if $\frac{G^{-1}\left[\overline{F}(x)\right]-m_Y}{F^{-1}(1-u)-m_X} \ge \frac{R_{X,Y}(x)-m_Y}{x-m_X}$ for all $x \le m_X$, or equivalently $\frac{G^{-1}\left[1-u\right]-m_Y}{F^{-1}(1-u)-m_X} \ge \frac{G^{-1}\left[u\right]-m_Y}{F^{-1}(u)-m_X}$ for all $u \in (0,\frac{1}{2})$.

Though a measure of skewness with respect to the mode has been recently proposed by Arnold and Groeneveld(1995), corresponding orderings with respect to the mode, which are weaker than van Zwet's ordering $\le_2$, are still missing.

Theorem 4.2. (*Hierarchy of orderings of skewness*) Between the defined orderings of skewness, the following implications hold :

(i) $X \le_2 Y \;\Rightarrow\; X \le_2^{\mu} Y$

(ii) $X \le_2 Y \;\Rightarrow\; X \le_{2,*}^{m} Y \;\Rightarrow\; X \le_{2,D}^{m} Y \;\Rightarrow\; X \le_2^{m} Y \;\Rightarrow\; X \le_{2,\gamma}^{m} Y$

Proof. For details, consult MacGillivray(1986). ◊

There are four classical measures of skewness, which were known by 1920 :

$\gamma_K\left[X\right]=\frac{\mu_X-M_X}{\sigma_X}$ (Karl Pearson(1895))

$\gamma\left[X\right]=\frac{\mu_{3,X}}{\sigma_X^3}$ (Edgeworth(1904), Charlier(1905), Doodson(1917), Haldane(1942))

$\gamma_{0.25}\left[X\right]=\frac{F^{-1}(0.75)-F^{-1}(0.25)-2m_X}{F^{-1}(0.75)-F^{-1}(0.25)}$ (Bowley(1901), David and Johnson(1954/56))

$\gamma_m^0\left[X\right]=\frac{\mu_X-m_X}{\sigma_X}$ (Yule(1911), Hotelling/Solomons(1932), Garver(1932), Majindar(1962))

By 1964, when van Zwet's ordering of skewness revolutionized the subject, only the last one did not pass the ordering axiom (Sk4), Definition 1.3. In fact, failing is solely due to the wrong choice of the measure of scale, and this measure can be replaced by

$$\gamma_m[X] = \frac{\mu_X - m_X}{E[|X - m_X|]} \quad \text{(see Groeneveld and Meeden(1984))}$$

There are many different formalizations of the concept of skewness, which according to Groeneveld(1991) should best be viewed as "a location- and scale-free deformation of the probability mass of a symmetric distribution". Further information about measures of skewness can be extracted from the influence function of Hampel(1968/74). We list various measures of skewness, which preserve one of the possible orderings of skewness. It seems that the generalized Bowley quantile measure $\gamma_u[X]$ with respect to the weakest ordering $\le_{2,\gamma}^m$ (suggested first by David and Johnson(1956)) is valuable in discussing both skewness and asymmetry (see MacGillivray(1986)). As a result in robust statistics, it is constantly affected by contamination in the tails of the distribution (see Groeneveld(1991)).

Examples 4.2 : location free and location dependent measures of skewness

(1) location free measures

Oja(1981) gives $\gamma_O[X] = E\left[\frac{X_{(3)} - X_{(2)}}{X_{(3)} - X_{(1)}}\right]$, with $X_{(1)}, X_{(2)}, X_{(3)}$ an ordered sample from X.

(2) mean dependent measures

The usual coefficient of skewness $\gamma[X] = \frac{\mu_{3,X}}{\sigma_X^3}$ is justified as measure of skewness with respect to the mean by the following result.

Theorem 4.3. (MacGillivray(1985/86)) If $X \le_2^\mu Y$ then one has the inequalities

$$\frac{\mu_{2n+1,X}}{\sigma_X^{2n+1}} \le \frac{\mu_{2n+1,Y}}{\sigma_Y^{2n+1}}, \quad n = 1,2,3,\dots .$$

If equality holds for any n, then the standardized distributions of X and Y coincide.

(3) median dependent measures

Details about the following results are found in MacGillivray(1986).

(3a) The generalized Bowley quantile measure of skewness defined by

$$\gamma_u[X] = \frac{F^{-1}(1-u) - F^{-1}(u) - 2m_X}{F^{-1}(1-u) - F^{-1}(u)}, \quad u \in (0, \tfrac{1}{2}),$$

preserves the weakest ordering $\leq_{2,\gamma}^{m}$.

(3b) The modified Yule measure of skewness $\gamma_m[X]=\dfrac{\mu_X - m_X}{E[|X-m_X|]}$ by Groeneveld and Meeden(1984) preserves the ordering $\leq_2^m$.

(3c) The symmetrically weighted quantile averages based measure of skewness

$$\gamma_m^K[X]=(\mu_K[X]-m_X)\cdot f(m_X)\text{, with } \mu_K[X]=\int_0^1 F^{-1}(u)dK(u),$$

where $K(u)$ is a distribution function on $(0,1)$ with symmetry center $\frac{1}{2}$ preserves $\leq_2^m$.

(3d) The median-density scaled measure of skewness

$$\gamma_m^d[X]=E\left[(X-m_X)^3\right]\cdot f(m_X)$$

preserves $\leq_2^m$. In fact, all the higher order measures $E\left[(X-m_X)^{2n+1}\right]\cdot f(m_X), \quad n=1,2,3,...$, are preserved by $\leq_2^m$.

(3e) The pure quantile measure of skewness

$$\gamma_{u,v}[X]=\frac{F^{-1}(1-u)+F^{-1}(u)-F^{-1}(1-v)-F^{-1}(v)}{F^{-1}(1-v)-F^{-1}(v)}, \quad 0\leq u<v\leq \tfrac{1}{2},$$

preserves $\leq_{2,*}^m$.

(4) <u>mode dependent measures</u>

Let $\mathcal{F}$ be a class of random variables, which is totally ordered with respect to $\leq_2$, such that each member X in $\mathcal{F}$ with distribution $F(x)$ has a continuously differentiable density $f(x)>0$ defined on an interval $S_F=(a,b), \quad -\infty\leq a<b\leq\infty$. Assume that $f'(x)>0$ for $x<M_X$, $f'(x)<0$ for $x>M_X$, and $f'(M_X)=0$. In particular, M_X is the unique mode of X. Then the measure of skewness with respect to the mode

$$\gamma_M[X]=1-2\cdot F(M_X)$$

preserves the ordering of skewness $\leq_2$. For details, consult Arnold and Groeneveld(1995).

4.3. <u>Measures of kurtosis</u>.

According to Balanda and MacGillivray(1988/90) *increasing kurtosis* should best be defined vaguely as "the location- and scale-free movement of probability mass from the "shoulders" of a distribution into its centre and tails" and formalize this movement by a coherent structure of orderings and measures.

For the class of symmetric distributions, a suitable ordering of kurtosis has been introduced by van Zwet(1964) (see also Oja(1981)).

Definition 4.9. If X and Y are symmetric, then X precedes Y in the *s-order*, written $X \leq_s Y$, if $R_{X,Y}(x) = G^{-1}F(x)$ is convex for $x > m_X$ and concave for $x < m_X$.

It is only quite recently that orderings of kurtosis, which apply in asymmetrical models and which imply $\leq_s$ in symmetrical models, have been constructed. Using quantile-based and moment-based spread functions, Balanda and MacGillivray(1990) have proposed two such coherent structures. The strongest orderings in these structures are defined as follows.

Definition 4.10. The random variable X precedes Y in the *s-order*, written $X \leq_s Y$, if $s_Y\left[s_X^{-1}(x)\right]$ is convex for $x \geq 0$, where $s_X(u)$ (resp. $s_Y(u)$) is the *quantile-based spread* function of X (resp. Y) defined by $s_X(u) = F^{-1}\left(\frac{1}{2}+u\right) - F^{-1}\left(\frac{1}{2}-u\right)$, $0 \leq u < \frac{1}{2}$.

Definition 4.11. The random variable X precedes Y in the *ζ-order*, written $X \leq_\zeta Y$, if $\zeta_Y\left[\zeta_X^{-1}(x)\right]$ is convex for $x \geq 0$, where $\zeta_X(u)$ (resp. $\zeta_Y(u)$) is the *moment-based spread* function of X (resp. Y) defined by $\zeta_X(u) = F(\mu_X + x) - F(\mu_X - x)$, $x \geq 0$.

For kurtosis comparisons within larger classes of random variables than those possible with $\leq_s$ *and* $\leq_\zeta$, it is necessary to weaken successively the convexity conditions defining these orders. Let us describe the hierarchy of kurtosis orderings obtained by Balanda and MacGillivray(1990).

For a specific scale-matching of kurtosis, one needs the following measure of scale, which preserves the dispersion order $\leq_1 \equiv \leq^{disp}$.

Definition 4.12. Let X be a random variable with distribution $F(x)$ and let $0 \leq \alpha, \beta < \frac{1}{2}$. The average change of spread over the interval (α, β) defines the measure of scale

$$\sigma_{\alpha,\beta}[X] = \frac{s_X(\beta) - s_X(\alpha)}{\beta - \alpha},$$

which reduces to $s'_X(\alpha)$ in case $\beta = \alpha$.

In the following, suppose that $0 \leq \gamma < \frac{1}{2}$ and $0 \leq \alpha, \beta < \frac{1}{2}$.

Definition 4.13. The random variable X precedes Y in the *star-shaped order of kurtosis*, written $X \leq_*^\gamma Y$, if

$$\frac{s_Y(u) - s_Y(\gamma)}{s_X(u) - s_X(\gamma)}$$

is increasing for $0 \leq u < \frac{1}{2}$.

Definition 4.14. The random variable X precedes Y in the *Doksum order of kurtosis*, written $X \le_{D,\beta}^{\alpha} Y$, if

(a) $$\frac{s_Y(u)-s_Y(\alpha)}{\sigma_{\alpha,\beta}[Y]}-\frac{s_X(u)-s_X(\alpha)}{\sigma_{\alpha,\beta}[X]} \text{ is decreasing for } 0\le u<\alpha$$

(b) $$\frac{\sigma_{\alpha,\beta}[Y]}{s_Y(u)-s_Y(\alpha)}-\frac{\sigma_{\alpha,\beta}[X]}{s_X(u)-s_X(\alpha)} \text{ is decreasing for } \alpha\le u\le\beta$$

(c) $$\frac{s_Y(u)-s_Y(\alpha)}{\sigma_{\alpha,\beta}[Y]}-\frac{s_X(u)-s_X(\alpha)}{\sigma_{\alpha,\beta}[X]} \text{ is increasing for } \beta\le u<\tfrac{1}{2}.$$

Definition 4.15. The random variable X precedes Y in the *weak order of kurtosis*, written $X \le_{\beta}^{\alpha} Y$, if

$$\frac{s_Y(u)-s_Y(\alpha)}{\sigma_{\alpha,\beta}[Y]}\begin{Bmatrix}\ge\\ \le\\ \ge\end{Bmatrix}\frac{s_X(u)-s_X(\alpha)}{\sigma_{\alpha,\beta}[X]} \text{ for } \begin{Bmatrix}0<u<\alpha\\ \alpha\le u<\beta\\ \beta\le u<\tfrac{1}{2}\end{Bmatrix}.$$

For each pair (α,β), one has the following ordering of kurtosis structure.

Theorem 4.4. (*Hierarchy of orderings of kurtosis*)

(a) $X \le_s Y \iff X \le_*^{\gamma} Y$ for all $\gamma\in[0,\tfrac{1}{2})$

(b) $X \le_*^{\gamma} Y \iff X \le_{D,\beta}^{\alpha} Y$ for all $\delta\in[\gamma,\tfrac{1}{2})$

$\iff X \le_{\delta}^{\gamma} Y$ for all $\delta\in[\gamma,\tfrac{1}{2})$

(c) $X \le_s Y \Rightarrow X \le_*^{\alpha} Y \Rightarrow X \le_{D,\beta}^{\alpha} Y \Rightarrow X \le_{\beta}^{\alpha} Y$.

Proof. For details, consult Balanda and MacGillivray(1990). ◊

Similarly to the generalized Bowley quantile measure of skewness, which preserves the weakest ordering of skewness as given in (3a) of the Examples 4.2, there exists a measure of kurtosis, which preserves the weakest ordering of kurtosis.

Example 4.3 : measures of kurtosis and tailweight

For each $0\le\alpha,\beta<\tfrac{1}{2}$, $0\le u<\tfrac{1}{2}$, the quantile-based functional

$$\gamma_2^{\alpha,\beta,u}[X]=\begin{cases}\dfrac{s_X(u)-s_X(\alpha)}{\sigma_{\alpha,\beta}[X]}, & 0<u<\alpha, \quad or \quad \beta<u<\tfrac{1}{2},\\[2ex] \dfrac{\sigma_{\alpha,\beta}[X]}{s_X(u)-s_X(\alpha)}, & \alpha<u<\beta,\end{cases}$$

preserves $\le_{\beta}^{\alpha}$ and all partial orderings preceding it in the hierarchy of Theorem 4.1(c). The special case $0=\alpha<\beta$ yields a large class of tailweight measures

$$\gamma_2^{\alpha,\delta}[X]=\frac{F^{-1}\left(\frac{1}{2}+\gamma\right)-F^{-1}\left(\frac{1}{2}-\gamma\right)}{F^{-1}\left(\frac{1}{2}+\delta\right)-F^{-1}\left(\frac{1}{2}-\delta\right)},\quad 0<\delta<\gamma<\tfrac{1}{2},$$

which is well-known in robust statistics literature. Balanda and MacGillivray(1990) have noticed that the quantile-based functional

$$\gamma_2^{\alpha}[X]=\frac{F^{-1}\left(\frac{3}{4}+\alpha\right)+F^{-1}\left(\frac{3}{4}-\alpha\right)-2\cdot F^{-1}\left(\frac{3}{4}\right)}{F^{-1}\left(\frac{3}{4}+\alpha\right)-F^{-1}\left(\frac{3}{4}-\alpha\right)},\quad 0<\alpha<\tfrac{1}{4},$$

which has been suggested as a measure of kurtosis for symmetric random variables by Groeneveld and Meeden(1984), does not preserve van Zwet's extension $\leq_s$ to the asymmetric case. Further details are found in the papers by Balanda and MacGillivray.

5. <u>Measures of "absolute" risk</u>.

Historically, the concept of "risk" has been used in many different senses, and has been defined according to its context of application in economics, statistics, insurance and finance (e.g. Borch(1967)).

One of the first notion of "risk" is attributed to Tetens(1786), who defined risk as "expected loss to the company, if the insurance contract leads to a loss". Greatly simplified, this yields risk as "one half of the mean absolute deviation". In modern terms, this first measure of risk is simultaneously our measure of scale (2b) in Example 4.1, that is

(5.1) $$Tetens[X]=E[(X-\mu_X)_+]=\tfrac{1}{2}E[|X-\mu_X|].$$

Though obsolete from a dynamic point of view (e.g. Borch(1967)), renewed interest, mainly in "static risk theory", can be found in many quite recent papers (consult Hürlimann(1998a) and references). Since it is usually more convenient to work with the standard deviation, the latter measure of scale has also been advocated as measure of risk (e.g. Hausdorff(1897)).

Concerning the practical use of measures of risk, a vague concept of risk emerged through the concept of (insurance) price defined roughly as "expected cost plus a safety loading in form of a function of the measure of risk", first suggested by Wold(1936), reconsidered in Ramsay(1993), and which can be formalized in many ways. For example, the concept of "safety loading" can be reconciled with expected utility theory (e.g. Nolfi(1957), Borch(1990) among others).

The above historical sketch reveals a certain connection between the concepts of scale and risk intuitively perceived by each actuary. In fact, the formal analogy in the axiomatic foundation of these notions (see Definitions 1.2 and 1.5) suggest that any measure of scale is potentially a candidate for a measure of risk. However, one requires more from the latter and there is also a main difference, which has been made precise only quite recently through the development of the theory of ordering of actuarial risks (a monograph entirely devoted to this topic is Kaas et al.(1994)). Indeed, the selected ordering of scale (say out of $\leq_1$, $\leq_1^{\mu}$, $\leq_1^{m}$, $\leq_1^{M}$, etc.) usually differs from the selected ordering of risk (say out of $\leq_{st}$, $\leq_{sl,=}$, $\leq_{sl}$, etc.). Though there exist relationships between these orderings, it is not always guaranteed that a measure of scale, which preserves an ordering of scale, also preserves an ordering of risk. In Section 5.1, some counterexamples are analyzed. In particular, we show that entropy is not a measure of risk, a fact which may be of interest in connection with the controverse raised by Sundt(1982). In contrast to this, "universal" examples, which are simultaneously measures of scale and risk, are displayed in Section 5.2.

On the other side, there is an indisputable need in actuarial science and finance to develop measures of risk, which explicitely use measures of skewness and kurtosis. As an example, Ramsay(1993), Section 4.3, derived from the normal power (NP) approximation to a moderately positively skewed random variable (e.g. Beard et al.(1984), Section 3.11) the "measure of risk" $R[X]=\sqrt{Var[X]\cdot(1+\frac{1}{18}\gamma[X]^2)}$. However, as shown by Promislow(1993), the Rothschild-Stiglitz ordering of risk $\le_{cx}\equiv\le_{sl,=}$ is not preserved by this (NP) measure of risk, rejecting it as a genuine measure of risk. In Section 5.3, we show how to adjust a measure of scale and risk for a positive measure of skewness. As illustration, an authentic skewness based measure of risk is discussed in Example 5.8. Similarly, using measures of kurtosis as constructed in Section 4.3, it should be possible to adjust a measure of scale and risk for kurtosis or for both skewness and kurtosis. However, this is a widely open area of investigation, which is beyond the scope of the present work and requires further research.

5.1. Some counterexamples.

To define a measure of risk, one has to specify a family F of non-negative random variables, which is totally ordered with respect to the selected ordering of risk, and verify if the axioms (aR1)-(aR4) in Section 1 are fulfilled for the proposed measure of risk. If a proposal passes this test for "big" classes of possible families F, it will be called a *universal* measure of risk. Otherwise, there exists a family F in such a "big" class, for which one axiom is not fulfilled, and the corresponding measure of risk is a counterexample.

Example 5.1 : standard deviation with respect to the usual stochastic order

The ordering of risk $\le_{st}$ is often assumed. Let F be the totally ordered family of diatomic random variables X with support $\{a,b\}, 0\le a<b$, b fixed, and probabilities $\Pr(X=a)=\Pr(X=b)=\frac{1}{2}$. For $X=\{0,b\}$, $Y=\{a,b\}$, $0<a<b$, one has $X\le_{st}Y$, but the standard deviations satisfy $\sigma_X=\frac{1}{2}b>\sigma_Y=\frac{1}{2}(b-a)$.

Example 5.2 : entropy with respect to the stop-loss order

From Examples 4.1, (1b), we know that entropy, defined by $En[X]=E[-\ln\{f(X)\}]$ with $f(x)$ the density of the random variable X, is a measure of scale. Consider the family of Pareto risks $X\sim Par(a,\gamma)$ with survival function $\overline{F}(x)=\left(\frac{x}{a}\right)^{-\gamma}$, $x\ge a>0$, $\gamma>1$. Since $\gamma>1$, the mean exists and equals $\mu=a\cdot\left(\frac{\gamma}{\gamma-1}\right)$. In the parametrization (μ,γ), one has the stop-loss order comparison rules (e.g. van Heerwaarden(1991), Section 6.4) :

$$Par(\mu_1,\gamma_1)\le_{sl}Par(\mu_2,\gamma_2) \quad \Leftrightarrow \quad 0<\mu_1\le\mu_2 \text{ and } \gamma_1\ge\gamma_2>1.$$

As totally ordered family F, choose only those $X\sim Par(\mu,\gamma)$ with parameters (μ,γ) in the region defined by the right-hand side in this comparison result. From calculation, the entropy of a member of this family is given by $En[X;\mu,\gamma]=\left(\frac{1+\gamma}{\gamma}\right)\cdot\left(2\cdot\ln\left\{\frac{\gamma}{\gamma-1}\right\}-\ln\{\mu\}\right)$. Now, if $\mu_1<\mu_2$, $\gamma_1=\gamma_2=\gamma>1$, corresponding Pareto risks $X_1\le_{sl}X_2$ are stop-loss ordered, but the measures of entropy satisfy $En[X_1;\mu_1,\gamma]>En[X_2;\mu_2,\gamma]$.

Example 5.3 : hazard with respect to the stop-loss order by equal means

Examples 4.1, (1c), tells us that the expected hazard $Eh[X]=E[h(X)]$ for non-negative random variables with decreasing densities is a measure of scale. Consider the one-parameter family of Pareto risks $X\sim Par(\mu,\gamma)$ with fixed mean μ and $\gamma>1$. The hazard rate is $h(x)=\frac{\gamma}{x}$ and its expected value equals $Eh[X]=E\left[\frac{\gamma}{X}\right]=\frac{1}{\mu}\cdot\left(\frac{\gamma^3}{\gamma^2-1}\right)$. Let $X_i\sim Par(\mu,\gamma_i)$, $i=1,2$, with $\gamma_1>\gamma_2>1$, hence $X_1\le_{sl,=}X_2$. It is easy to see that

$$Eh[X_1]<Eh[X_2],\text{ if } 1<\gamma_2<\gamma_1<\sqrt{3}\text{, and } Eh[X_1]>Eh[X_2],\text{ if } \gamma_1>\gamma_2\ge\sqrt{3}.$$

In particular, the expected hazard is not a measure of risk with respect to $\le_{sl,=}$ for the Pareto family with fixed mean and index $\gamma\ge\sqrt{3}$.

5.2. Universal measures of scale and risk.

Among the measures of scale mentioned in the Examples 4.1, the following ones, may be viewed as measures of risk with respect to some ordering of risk.

Example 5.4 : standard deviation with respect to the Rothschild-Stiglitz order $\le_{cx}\equiv\le_{sl,=}$

It is well-known that $X\le_{sl,=}Y$ implies $Var[X]\le Var[Y]$, hence axiom (aR5) is fulfilled. Since the remaining axioms are satisfied, standard deviation is a measure of risk with respect to $\le_{sl,=}$. By unequal means, this is not a measure of risk, as follows from Example 5.1 ($\le_{st}$ implies $\le_{sl}$). Moreover, as in case of the variance, its application has hitherto been justified only in situations for which a normal distribution or a quadratic utility can be assumed (see e.g. the comments and references in Ramsay(1993)).

Example 5.5 : Teten's measure with respect to the Rothschild-Stiglitz order

It is remarkable that the oldest proposed measure (5.1) is a genuine measure of risk with respect to the stop-loss ordering of risk by equal means. For several families of random variables it preserves also the weaker stop-loss order by unequal means. Moreover, it satisfies a lot of other interesting and important statistical, economical and actuarial properties. For some references, consult Hürlimann(1998a).

Example 5.6 : a class of distortion measures with respect to the Rothschild-Stiglitz order

A close look at the common structure of the median absolute deviation and the Gini measure suggests the following generalized measure.

Theorem 5.1. (*Distortion measures of scale and risk*) Let $\varphi(x)$ be a differentiable increasing concave function on $(0,\frac{1}{2})$ such that $\varphi(0)=0$, $\varphi(\frac{1}{2})\in(0,1)$, $0\le\varphi'(x)\le 1$ for $x\in\left[0,\frac{1}{2}\right)$, and set

$$h(x)=\begin{cases}\varphi(x), & x\in\left[0,\frac{1}{2}\right),\\ \varphi(1-x), & x\in\left[\frac{1}{2},1\right].\end{cases}$$

For the class of non-negative random variables with survival function $\overline{F}_X(x)$, the functional

(5.2) $$R[X]=\int_0^\infty h\left[\overline{F}_X(x)\right]dx$$

is a measure of scale for $\le_1$ and a measure of risk for $\le_{sl,=}$.

Proof. Making the substitution $F(x)=u$, one obtains

(5.3) $$R[X]=\int_0^\infty h\left[\overline{F}_X(x)\right]dx=\int_0^1 F_X^{-1}(1-u)h'(u)du=\int_0^{\frac{1}{2}}\left\{F_X^{-1}(1-u)-F_X^{-1}(u)\right\}\varphi'(u)du\ .$$

This quantile integral representation shows that $R[X]$ preserves the ordering $\le_1\equiv\le^{disp}$, and since $R[aX]=|a|\cdot R[X]$, it is a measure of scale. To verify axiom (aR5) for $\le_{sl,=}$, observe that the associated "measure of price" in the sense of Definition 1.7 defined by

$$P[X]=E[X]+\theta\cdot R[X]=\int_0^\infty g\left[\overline{F}_X(x)\right]dx,\quad g(x)=x+\theta\cdot h(x),\quad 0<\theta\le 1,$$

preserves $\le_{sl}$ by Theorem 6.1. By equal means, it follows that $R[X]$ preserves $\le_{sl,=}$. It remains to verify (aR3). From Wang(1996), Theorem 2, applied to the measure of price $P[X]$, one obtains the stronger subadditive property $R[X+Y]\le R[X]+R[Y]$ for all X and Y regardless of dependence. ◊

As straightforward examples, one recovers the mean absolute deviation from median measure $\sigma_m[X]=E\left[|X-m_X|\right]$ for $\varphi(x)=x$, and the Gini measure $Gini[X]=\int_0^\infty F(x)\overline{F}(x)dx$ for $\varphi(x)=x(1-x)$. Among recent examples, one notes $\varphi(x)=-x\cdot\ln(x)$, which has been considered in relation with option pricing theory (see Hürlimann(1997c/98d)) and $\varphi(x)=\sqrt{x}$ used by Wang(1998) to define an actuarial index of the right-tail risk. In general, it is not clear when the functional (5.2) preserves the usual stochastic order $\le_{st}$, from which it would follow through application of the so-called separation theorem (e.g. Kaas et al.(1994), Theorem IV.2.1) that it is a measure with respect to the weaker ordering $\le_{sl}$. In any case, this is true for many families or random variables or/and choices of $\varphi(x)$. It is also interesting to mention that (5.3) is a measure of scale with respect to the (new) ordering of scale $\le_1^q$, which is weaker than $\le^{disp}$ and defined by

$$X\le_1^q Y\quad\Leftrightarrow\quad F_X^{-1}(1-u)-F_X^{-1}(u)\le F_Y^{-1}(1-u)-F_Y^{-1}(u)\quad\text{for all }\ u\in(0,\tfrac{1}{2}).$$

Example 5.7 : stop-loss at mode with respect to the stop-loss order

Assume F is a totally ordered family of non-negative unimodal random variables with respect to $\le_{sl}$ such that $M_Y \le M_X$ if $X \le_{sl} Y$ and $M_{X+Y} \ge M_X + M_Y$ for independent $X, Y \in F$ such that $X+Y \in F$. To be sure that $X+Y$ is unimodal for independent X and Y, one can assume that either X or Y is strongly unimodal, a notion introduced by Ibragimov(1956) (see the book by Dharmadhikari and Joag-dev(1988)). Then the stop-loss at mode functional $\pi_M[X] = E[(X - M_X)_+]$ is a measure of scale for $\le_1$ and a measure of risk for $\le_{sl}$. By assumption, one has $M_Y \le M_X$ in case $X \le_{sl} Y$, hence the stop-loss order is preserved. Since $\le_1$ implies $\le_{st}$ (and $\le_{sl}$) (condition (g) in Theorem 3.1), the ordering of scale is also preserved. Finally, subadditivity for independent X and Y follows from the assumption $M_{X+Y} \ge M_X + M_Y$ and the inequalities

$$E[(X+Y-M_{X+Y})_+] \le E[(X+Y-M_X-M_Y)_+] \le E[(X-M_X)_+] + E[(Y-M_Y)_+].$$

5.3. Adjustment of measures of scale and risk for positive skewness.

It is well-known that the standard deviation measure of risk (with respect to $\le_{cx} \equiv \le_{sl,=}$) is usually accepted as measure of risk only if random variables are (approximately) normally distributed. There does not seem to exist in the actuarial and financial literature measures of risk, which adjust a "universal" measure of scale and risk (as standard deviation) for positive skewness (and kurtosis) risk, and preserves an accepted ordering of risk (as the Rothschild-Stiglitz measure). Indeed, the most recent proposed special normal power measure of risk

$$R[X] = \sigma[X] \cdot \left(1 + \tfrac{1}{18}\gamma[X]^2\right), \quad \gamma[X] = \frac{\mu_{3,X}}{\sigma_X^3},$$

does not always preserve $\le_{cx}$ (comment by Promislow(1993)). A general method to adjust a measure of scale and risk $S[X]$ for a positive measure of skewness $\gamma[X] \ge 0$ is to set

(5.4) $$R[X] = S[X] \cdot (1 + c \cdot \gamma[X]), \quad c \ge 0,$$

such that the required axioms for a measure of risk are satisfied. A single example suffices to illustrate what is meant.

Example 5.8 : the median absolute deviation measure of risk adjusted for the modified Yule measure of skewness

If in (5.4) one sets $S[X] = E[|X - m_X|]$ (see Example 5.6), $\gamma[X] = \gamma_m[X] = \dfrac{\mu_X - m_X}{E[|X - m_X|]}$ (which is (3b) from Example 4.2), and $c = 1 + \varepsilon, \varepsilon \ge 0$, one obtains

(5.5) $$R[X] = E[|X - m_X|] + c \cdot (\mu_X - m_X) = \varepsilon \cdot (\mu_X - m_X) + 2 \cdot E[(X - m_X)_+].$$

There are many families of random variables for which (5.5) is a measure of risk with respect to $\leq_{cx}$. In particular, when $\varepsilon = 0$, replacing the mean by the median in Teten's measure of risk (Example 5.5) yields a measure of risk, which adjusts for positive skewness. An example often used to model financial assets under a risk-neutral valuation assumption is the lognormal $\ln N(\mu_X, \sigma_X)$ with constant mean $r = \exp\{\mu_X + \frac{1}{2}\sigma_X^2\}$ and median $m_X = \exp\{\mu_X\} = r \cdot \exp\{-\frac{1}{2}\sigma_X^2\}$. It is well-known that lognormal distributions with equal mean increase in stop-loss order with increasing volatility parameter (e.g. Hürlimann(1995), Lemma 4.1). It follows that $\mu_X - m_X = r \cdot (1 - \exp\{-\frac{1}{2}\sigma_X^2\})$ and (5.5) are preserved with respect to $\leq_{cx}$. A similar popular example, used to model large claims in reinsurance, is the Pareto already considered in Example 5.2. For the reparametrized version in terms of (μ, γ), recall the stop-loss order comparison rule :

(5.6) $$Par(\mu_X, \gamma_X) \leq_{sl} Par(\mu_Y, \gamma_Y) \quad \Leftrightarrow \quad 0 < \mu_X \leq \mu_Y \text{ and } \gamma_Y \geq \gamma_X > 1.$$

A calculation yields $m_X = \mu_X \cdot (\gamma_X - 1) \cdot (2^{\frac{1}{\gamma_X}} - 1)$, $E[(X - m_X)_+] = \frac{1}{2}\mu_X 2^{\frac{1}{\gamma_X}}$, and thus

(5.7) $$R[X] = \mu_X \cdot \left\{ \varepsilon \left[1 - (\gamma_X - 1) \cdot (2^{\frac{1}{\gamma_X}} - 1) \right] + 2^{\frac{1}{\gamma_X}} \right\}.$$

Taking the derivative with respect to γ_X, one sees that if $0 \leq \varepsilon \leq \ln(4) = 1.3863$, then (5.7) preserves the stop-loss order.

6. <u>Measures of price</u>.

Since Bühlmann(1970) the functional approach to premium calculation has seen a tremendous development. Monographs of risk theory containing accounts of this approach include Gerber(1979), Goovaerts et al.(1984), Heilmann(1987) and Kaas et al.(1994). Though the stop-loss ordering preserving property of the Swiss family of premium calculation principles has been known since its consideration in Bühlmann et al.(1977), the recognition of $\leq_{sl}$ as a sound ordering of risk seems more recent. For example, the order preserving axiom (P5) is considered in Heilmann(1987) but without mention of a specific partial order, which could be used as selected ordering of risk. Furthermore, the absolute deviation principle and the Gini principle, introduced by Denneberg(1985/90), and which satisfy axioms (P1)-(P4), have been shown to satisfy (P5) for $\leq_{sl}$ only very recently (Theorem 6.1 and its comments). For this reason, it seems useful to present a short chronological review of some main non-trivial pricing functionals, which preserve $\leq_{sl}$, and inspect whether the remaining axioms (P1)-(P4) are satisfied. We restrict our attention to non-negative random variables.

The Swiss family is positively homogeneous if, and only if, it is the net principle (see Schmidt(1989), simpler proof by Hürlimann(1997b), Example 4.1 (continued), p.9). The first genuine "measures of price", which satisfy (P1)-(P5), are the absolute deviation principle $P[X] = E[X] + \theta \cdot E[|X - m_X|]$, $0 \leq \theta \leq 1$ (Denneberg(1985/90)) and the Gini principle $P[X] = E[X] + \theta \cdot Gini[X]$, $0 \leq \theta \leq 1$ (Denneberg(1990)). These functionals are special cases of the class of distortion pricing principles

(6.1) $$P[X]=\int_0^{\infty} g[\overline{F}_X(x)]dx=\int_0^1 F_X^{-1}(1-u)dg(u),$$

where $g(x)$ is an increasing concave function such that $g(0)=0, g(1)=1$, and F_X^{-1} is a generalized inverse of F_X. The right-hand side representation has been introduced by Denneberg(1990) and its equivalence with the first integral has been used by Wang(1996a) and Wang et al.(1997).

The following main result is in a great extent accessible from an elementary perspective, as shown in Hürlimann(1998b) (see also Goovaerts and Dhaene(1998)).

Theorem 6.1. (*Distortion measure of price with respect to stop-loss order*) Let F be the family of non-negative random variables X with survival functions $\overline{F}_X(x)$ and quantile functions $F_X^{-1}(u)$, and let $g(x)$ be a differentiable increasing concave function on $[0,1]$ such that $g(0)=0, g(1)=1$. Then the functional (6.1) satisfies the axioms (P1)-(P5) of a measure of price.

Theorem 6.1 yields the first general rather elementary method to generate valuable measures of price. Another attractive special case is the PH-transform principle studied by Wang(1995a/95b/96a/96b), and a new "entropy" principle generated by the distortion function $g(x)=x-\theta\cdot x\cdot\ln(x)$, which is related to the new measure of risk mentioned after Theorem 5.1. Previously to the last examples had appeared the Dutch principle (see van Heerwaarden(1991a/91b), van Heerwaarden and Kaas(1992), Kaas et al.(1994)) and a slight generalization of it (see Hürlimann(1994/95a/95b)). A pricing principle from the Dutch family satisfies (P1)-(P5) if, and only if, it is of the form

(6.2) $$P[X]=E[X]+\theta\cdot E[(X-E[X])_+],\quad 0\le\theta\le 1,$$

and is directly related with Teten's measure of risk (see Example 5.5). The Dutch family is a special case of the class of so-called "quasi-mean value principles" considered recently by the author. However, only sporadic members of this class define feasible measures of price, satisfying (P1)-(P5), of which one may mention the interesting Example 11.1 in Hürlimann(1997b).

A generalization of the class of distortion pricing principles is the class of Choquet pricing principles in Chateauneuf et al.(1996), which is based on the theory of capacities and non-additive measures (exposed in Denneberg(1994)), and breaks with the traditional probabilitistic foundations of actuarial science and finance. Finally, let us mention that one misses still feasible "measures of price" along the economic approach initiated by Bühlmann(1980/84) (see the critical comments by Lemaire(1988)).

As follows from Section 5.2, there should be a close relationship between measures of risk and price. Indeed, if $P[X]$ is a measure of price preserving $\le_{sl}$, then the safety loading (insurance terminology) or risk premium (finance terminology) given by $(P[X]-E[X])$ defines a measure of risk with respect to the stronger ordering of risk $\le_{sl,=}$. Reciprocally, not every measure of risk $R[X]$ with respect to $\le_{sl,=}$ yields a measure of price $P[X]=E[X]+\theta\cdot R[X]$ with respect to $\le_{sl}$. The most straightforward counterexample is the standard deviation principle $P[X]=E[X]+\theta\cdot\sqrt{Var[X]}$, which does not preserve $\le_{st}$, and a fortiori $\leq_{sl}$. Despite the lack of theoretical justification, this pricing principle is still encountered in actuarial practice.

To conclude with a main message, an important merit of the formal approach based on orderings and measures is the possibility to construct explicit measures of price, which take into account non-neglible skewness and kurtosis effects. For example, the measure of risk presented in Example 5.8 generates a pricing functional

$$P[X] = E[X] + \alpha \cdot (E[X] - m[X]) + \beta \cdot E[(X - m[X])_+], \quad (6.3)$$

which for specific values of $\alpha, \beta \geq 0$ and families of random variables (e.g. the Pareto family) can be viewed as a measure of price with respect to the stop-loss ordering of risk.

References.

Alfsen, E. (1971). Compact Convex Sets and Boundary Integrals. Springer. New York.

Andrews, D.F., Bickel, P.J., Hampel, F.R., Huber, P.J., Rogers, W.H. and J.W. Tukey (1972). Robust Estimates of Location : Surveys and Advances. Princeton University Press, Princeton.

Arnold, B.C. and R.A. Groeneveld (1995). Measuring skewness with respect to the mode. The American Statistician 49(1), 34-38.

Artzner, P., Delbaen, F., Eber, J.-M. and D. Heath (1997a). Definition of coherent measures of risk. Risk Management Symposium of the European Finance Association Meeting, Vienna, August 27-30, 1997.

Artzner, P., Delbaen, F., Eber, J.-M. and D. Heath (1997a). Thinking coherently. RISK 10, no. 11 (November 1997), 68-70.

Bagai, I. and S.C. Kochar (1986). On tail-ordering and comparison of failure rates. Communications in Statistics - Theory and Methods 15(4), 1377-88.

Balanda, K.P. and H.L. MacGillivray (1988). Kurtosis : a critical review. The American Statistician 42, 111-19.

Balanda, K.P. and H.L. MacGillivray (1990). Kurtosis and spread. The Canadian Journal of Statistics 18, 17-30.

Bartoszewicz, J. (1985). Dispersive ordering and monotone failure rate distributions. Advances in Applied Probability 17, 472-474.

Bartoszewicz, J. (1987). A note on dispersive ordering defined by hazard functions. Statistics and Probability Letters 6, 13-16.

Beard, R.E., Pentikäinen, T. and E. Pesonen (1984). Risk Theory. The Stochastic Basis of Insurance. Chapman and Hall(1984).

Bickel, P.J. and E.L. Lehmann (1975). Descriptive statistics for nonparametric models I. Introduction. II. Location. Annals of Statistics 3, 1038-1069.

Bickel, P.J. and E.L. Lehmann (1976). Descriptive statistics for nonparametric models III. Dispersion. Annals of Statistics 4, 1139-1158.

Bickel, P.J. and E.L. Lehmann (1979). Descriptive statistics for nonparametric models IV. Spread. In : J. Jurechova (Ed.). Contributions to Statistics (Hàjek Memorial Volume). Reidel : Amsterdam.
Blackwell, D. (1953). Equivalent comparisons of expirements. Annals of Mathematical Statistics 24, 265-272.
Blackwell, D. and L.E. Dubins (1963). A converse to the dominated convergence theorem. Illinois Journal of Mathematics 7, 508-514.
Borch, K.H. (1967). The Theory of Risk. Journal of the Royal Statistical Society, Series B 29(3), 432-67. Reprinted in Borch(1990), 105-34.
Borch, K.H. (1990). Economics of Insurance. Advanced Textbooks in Economics 29. North-Holland.
Bowley, A.L. (1901). Elements of Statistics (6th ed., 1937). Staples Press Ltd., London.
Bruckner, A.M. and E. Ostrow (1962). Some function classes related to the class of convex functions. Pacific Journal of Mathematics 12, 1203-1215.
Bühlmann, H. (1970). Mathematical Methods in Risk Theory. Springer-Verlag.
Bühlmann, H. (1980). An economic premium principle. ASTIN Bulletin 11, 52-60.
Bühlmann, H. (1984). The general economic premium principle. ASTIN Bulletin 14, 13-21.
Bühlmann, H., Gagliardi, B., Gerber, H.U. and E. Straub (1977). Some inequalities for stop-loss premiums. ASTIN Bulletin 9, 75-83.
Charlier, C.V.L. (1905). Über das Fehlergesetz. Arkiv fir Math., Astronomi och Fysik 2, 8.
Chateauneuf, A., Kast, R. and A. Lapied (1996). Choquet pricing for financial markets with frictions. Mathematical Finance 6(3), 323-30.
David, F.N. and N.L. Johnson (1954). A test for skewness with ordered variables. Annals Eugenics London 18, 351-353.
David, F.N. and N.L. Johnson (1956). Some tests of significance with ordered variables. Journal Royal Statistical Society, Series B 18, 1-20.
Denneberg, D. (1985). Valuation of first moment risk for decision purposes in Finance and Insurance. In : Goppel, H., Henn, R. (ed.) 3. Tagung Geld, Banken und Versicherungen, Karlsruhe, 855-869. Verlag Versicherungswirtschaft, Karlsruhe.
Denneberg, D. (1990). Premium calculation : why standard deviation should be replaced by absolute deviation. ASTIN Bulletin 20, 181-190.
Denneberg, D. (1994). Non-Additive Measure and Integral. Theory and Decision Library, Series B, vol. 27. Kluwer Academic Publilshers.
Deshpande, J. and S. Kochar (1983). Dispersive ordering is the same as tail-ordering. Advances in Applied Probability 15, 686-87.
Dharmadhikari, S.W. and K. Joag-dev (1988). Unimodality, Convexity, and Applications. Academic Press, Boston.
Doksum, K. (1969). Star-shaped transformations and the power of rank tests. Annals of Mathematical Statistics 40, 1167-76.
Doodson, A.T. (1917). Relation of the mode, median and mean in frequency curves. Biometrika 11, 425-429.
Dubins, L.E. and D. Gilat (1978). On the distribution of the maxima of martingales. Transactions of the American Mathematical Society 68, 337-38.
Edgeworth, F.Y. (1904). The law of error. Transactions of the Cambridge Phil. Soc. 20, 36-65 and 113-114.
Elton, J. and T.P. Hill (1992). Fusions of a probability distribution. The Annals of Probability 20(1), 421-54.

Embrechts, P., Klüppelberg, C. and Th. Mikosch (1997). Modelling Extremal Events for Insurance and Finance. Applications of Mathematics - Stochastic Modelling and Applied Probability, vol. 33.
Garrido, J. (1993). Discussion of the paper by Ramsay(1993).
Garver, R. (1932). Concerning the limits of a measure of skewness. Annals of Mathematical Statistics 3, 358-360.
Gerber, H.U. (1979). An Introduction to Mathematical Risk Theory. Huebner Foundation Monograph 8. R.D. Irwin, Homewood, Illinois.
Goovaerts, M.J., DeVylder, F., Haezendonck, J. (1984). Insurance Premiums. North-Holland.
Goovaerts, M.J. and J. Dhaene (1998). On the characterization of Wang's class of premium principles. Transactions 26th International Congress of Actuaries, vol. 4, 121-34.
Groeneveld, R.A. (1991). An influence function approach to describing the skewness of a distribution. The American Statistician 45, 97-102.
Groeneveld, R.A. and G. Meeden (1984). Measuring skewness and kurtosis. The Statistician 33, 391-399.
Haldane, J.B.S. (1942). The mode and median of a nearly normal distribution with given cumulants. Biometrika 32, 294-299.
Hampel, F.R. (1968). Contributions to the theory of robust estimation. Ph. D. Dissertation. University of California, Berkeley.
Hampel, F.R. (1974). The influence curve and its role in robust estimation. Journal of the American Statistical Association 69, 383-393.
Hardy, G.H. and J.E. Littlewood (1930). A maximal theorem with function-theoretic applications. Acta Mathematica 54, 81-116.
Hausdorff, F. (1897). Das Risico bei Zufallsspielen. Leipziger Berichte 49, 497-548.
Heerwaarden, van A.E. (1991a). Analysis of premium calculation principles; design of a new principle. Tinbergen Institute Research Bulletin 3, 101-113.
Heerwaarden, van A.E. (1991b). Ordering of risks : theory and actuarial applications. Ph.D. Thesis, Tinbergen Research Series no. 20, Amsterdam.
Heerwaarden, van A.E. and R. Kaas (1992). The Dutch premium principle. Insurance : Mathematics and Economics 11, 129-133.
Heilmann, W.-R. (1987). Grundbegriffe der Risikotheorie. Verlag Versicherungswirtschaft, Karlsrune. (English translation (1988). Fundamentals of Risk Theory)
Hotelling, H. and L.M. Solomons (1932). The limits of a measure of skewness. Annals of Mathematical Statistics 3, 141-142.
Howison, S.D., Kelly, F.P. and P. Willmott (1994) (Editors). Mathematical models in finance. Phil. Transactions Royal Society London, Series A 347, 449-598.
Hürlimann, W. (1992). Ordering of risks through loss ratios. Insurance : Mathematics and Economics 11, 49-54.
Hürlimann, W. (1994). A note on experience rating, reinsurance and premium principles. Insurance : Mathematics and Economics 14, 197-204.
Hürlimann, W. (1995a). Links between premium principles and reinsurance. International Congress of Actuaries, Brussels, vol. 2, 141-167.
Hürlimann, W. (1995b). Transforming, ordering and rating risks. Bulletin of the Swiss Association of Actuaries, 213-236.
Hürlimann, W. (1997a). On distribution-free safe layer-additive pricing. Insurance : Mathematics and Economics 22 (1998), 277-285.
Hürlimann, W. (1997b). On quasi-mean value principles. Blätter der Deutschen Gesellschaft für Versicherungsmathematik XXIII, Heft 1, 1-16.

Hürlimann, W. (1997c). Inequalities for lookback option strategies and exchange risk modelling. Presented at the first Euro-Japanese Workshop on Stochastic Modelling in Insurance, Finance, Production and Reliability, Brussels, September 7-10, 1998.

Hürlimann, W. (1998a). Distribution-free excess-of-loss reserves for some actuarial protection models. Trans. 26th International Congress of Actuaries, vol. 4, 291-317.

Hürlimann, W. (1998b). On stop-loss order and the distortion pricing principle. ASTIN Bulletin 28(1), 119-134.

Hürlimann, W. (1998c). Truncation transforms, stochastic orders and layer pricing. Transactions of the 26th International Congress of Actuaries, vol. 4, 135-151.

Hürlimann, W. (1998d). Einführung in die theoretische Prämienkalkulation. Manuscript (available from the author).

Hutchinson, T.P. (1993). Essentials of Statistical Methods, in 41 Pages. Rumsby Scientific Publishing. Adelaide, South Australia.

Ibragimov, I.A. (1956). On the composition of unimodal distributions. Theory of Probability and its Applications 1, 255-266.

Kaas, R. and A.E. van Heerwaarden (1992). Stop-loss order, unequal means, and more dangerous distributions. Insurance : Mathematics and Economics 11, 71-77.

Kaas, R., Heerwaarden, van A.E. and M.J. Goovaerts (1994). Ordering of Actuarial Risks. CAIRE Education Series 1, Brussels.

Kaas, R., and O. Hesselager (1995). Ordering claim size distributions and mixed Poisson probabilities. Insurance : Mathematics and Economics 17, 193-201.

Karlin, S. (1968). Total Positivity. Stanford University Press, Stanford.

Karlin, S. and A. Novikoff (1963). Generalized convex inequalities. Pacific Journal of Mathematics 13, 1251-1279.

Karlin, S. and Z. Ziegler (1976). Some applications to inequalities of the method of generalized convexity. Journal Anal. Math. 30, 281-303.

Kertz, R.P. and U. Rösler (1990). Martingales with given maxima and terminal distributions. Israel Journal of Mathematics 69, 173-192.

Kertz, R.P. and U. Rösler (1992). Stochastic and convex orders and lattices of probability measures, with a martingale interpretation. Israel Journal of Mathematics 77, 129-164.

Kertz, R.P. and U. Rösler (1993). Hyperbolic-concave functions and Hardy-Littlewood maximal functions. In : Stochastic Inequalities. IMS Lecture Notes - Monograph Series, vol. 22.

Landers, D. and L. Rogge (1983). Consistent estimation of the natural median. Statistics and Decisions 1, 269-284.

Landsberger, M. and I. Meilijson (1994). Co-monotone allocations, Bickel-Lehmann dispersion and the Arrow-Pratt measure of risk aversion. Annals of Operations Research 52, 97-106.

Lehmann, E.L. (1955). Ordered families of distributions. Annals of Mathematical Statistics 26, 399-419.

Lemaire, J. (1988). Actuarial Challenges of Reinsurance. 24-th International Congress of Actuaries, Helsinki.

MacGillivray, H.L. (1985). A crossing theorem for distributions and their moments. Bulletin Australian Mathematical Society 31, 413-419.

MacGillivray, H.L. (1986). Skewness and asymmetry : measures and orderings. The Annals of Statistics 14, 994-1011.

Majindar, K.N. (1962). Improved bounds on a measure of skewness. Annals of Mathematical Statistics 33, 1192-1194.

Makowski, A.M. (1994). On an elementary characterization of the increasing convex ordering,

with an application. Journal of Applied Probability 31, 834-40.
Mann, H.B. and D.R. Whitney (1947). On a test of whether one of two random variables is stochastically larger than the other. Annals of Mathematical Statistics 18, 50-60.
Meilijson, I. and A. Nàdas (1979). Convex majorization with an application to the length of critical paths. Journal of Applied Probability 16, 671-77.
Muñoz-Pérez, J. (1990). Dispersive ordering by the spread function. Statistics and Probability Letters 10, 407-10.
Nolfi, P. (1957). Zur mathematischen Darstellung des Nutzens in der Versicherung. Bulletin of the Swiss Association of Actuaries, 395-407.
Ohlin, J. (1969). On a class of measures of dispersion with application to optimal reinsurance. ASTIN Bulletin 5, 249-66.
Oja, H. (1981). On location, scale, skewness and kurtosis of univariate distributions. Scandinavian Journal of Statistics 8, 154-168.
Pearson, K. (1895). Contributions to the mathematical theory of evolution II. Skew variation in homogeneous material. Phil. Trans. Royal Society London, Series A 186, 343-414.
Promislow, S.D. (1993). Discussion of the paper by Ramsay(1993).
Ramsay, C.M. (1993). Loading gross premiums for risk without using utility theory. Transactions of the Society of Actuaries XLV, 305-306. Discussion, 337-349.
Ruppert, D. (1987). What is kurtosis ? An influence function approach. The American Statistician 41(1), 1-5.
Rüschendorf, L. (1991). On conditional stochastic ordering of distributions. Advances in Applied Probability 23, 46-63.
Schmeidler, D. (1986). Integral representation without additivity. Proceedings of the American Mathematical Society 97, 255-261.
Schmidt, K.D. (1989). Positive homogeneity and multiplicativity of premium principles on positive risks. Insurance : Mathematics and Economics 8, 315-319.
Shaked, M. (1982). Dispersive ordering of distributions. Journal of Applied Probability 19, 310-320.
Stoyan, D. (1977). Qualitative Eigenschaften und Abschätzungen Stochastischer Modelle. Akademie-Verlag, Berlin. (English version (1983). Comparison Methods for Queues and Other Stochastic Models. J. Wiley, New York.)
Sundt, B. (1982). Minimum entropy in Risk Theory. Bulletin of the Swiss Association of Actuaries, 296-299.
Szekli, R. (1995). Stochastic Ordering and Dependence in Applied Probability. Lecture Notes in Statistics 97. Springer-Verlag.
Taylor, J.M. (1983). Comparisons of certain distribution functions. Math. Operationsforschung und Statistik, Ser. Stat. 14(3), 397-408.
Tetens, J.N. (1786). Einleitung zur Berechnung der Leibrenten und Anwartschaften. Leipzig.
van der Vecht, D.P. (1986). Inequalities for stopped Brownian motion. C.W.I. Tracts 21, Mathematisch Centrum, Amsterdam.
van Zwet, W.R. (1964). Convex Transformations of Random Variables. Mathematics Centre Tract 7, Mathematisch Centrum, Amsterdam.
Wang, S. (1995a). Insurance pricing and increased limits ratemaking by proportional hazards transforms. Insurance : Mathematics and Economics 17, 43-54.
Wang, S. (1995b). Insurance pricing and increased limits ratemaking by proportional hazards transforms. Proc. XXV. Int. Congress of Actuaries Brussels, vol. 2, 293-323.
Wang, S. (1995c). Risk loads in life/non-life insurance : a unified approach. XXVI. ASTIN Colloquium, Leuven.

Wang, S. (1996a). Premium calculation by transforming the layer premium density. ASTIN Bulletin 26, 71-92.
Wang, S. (1996b). Ordering of risks under PH-transforms. Insurance : Mathematics and Economics 18, 109-114.
Wang, S. (1998). An actuarial index of the right-tail risk. North American Actuarial Journal 2(2), 88-101.
Wang, S., Young, V.R. and H.H. Panjer (1997). Axiomatic characterization of insurance prices. Insurance : Mathematics and Economics 21, 173-183.
Wold, H. (1936). Landsbygdens Brandförsäkringsbolags Maximaler och Aterförsäkring. Stockholm.
Yaari, M.E. (1987). The dual theory of choice under risk. Econometrica 55, 95-115.
Yule, G.U. (1911). Introduction to the Theory of Statistics (13th ed., 1944). Griffin, London.