Westfälische Wilhelms-
Universität Münster

Institut für Kreditwesen
Direktor: Prof. Dr. Andreas Pfingsten

Institut für
Kreditwesen

Diskussionsbeitrag 99-03

Learning from Poverty Measurement: An Axiomatic Approach to Measure Downside Risk

Dipl.-Kfm. Carsten Breitmeyer

Dipl.-Math. Hendrik Hakenes

Prof. Dr. Andreas Pfingsten

Dipl.-Wi.-Inf. Christoph Rechtien

ISSN 0949-6610

August 2000

Abstract

It has been long known that a close relation between the measurement of inequality and the measurement of risk exists. We demonstrate a similar relation between measures of poverty and downside risk respectively. Among others, a checklist for reasonable measures of downside risk emerges. The widely used "Value at Risk" fails quite a few tests, whereas some other indices do much better.

Key Words: Value at Risk, Lower Partial Moments, Downside Risk, Shortfall Risk Measurement, Poverty Measurement.

An earlier version of this paper was presented at the Third International Stockholm Seminar on Risk Behaviour and Risk Management in 1999, the 1999 European Meeting of the Econometric Society in Santiago de Compostela, the Symposium on Operations Research in Magdeburg (SOR '99) and the 1999 Meeting of the German Finance Association in Aachen. We would like to thank the participants for helpful comments, but remain sole responsibility for all remaining errors and omissions.

Westfälische Wilhelms-Universität Münster
Institut für Kreditwesen
Universitätsstraße 14-16
D-48143 Münster
Telefon: +49 251 83-22881
Telefax: +49 251 83-22882
E-Mail: `21anpf@wiwi.uni-muenster.de`

Contents

1 Introduction

Decision-making under uncertainty is required in many situations. In particular in portfolio management this is indeed the key issue. But how should it be done?

A number of approaches have been put forward over the last decades, all of them suggesting more or less appealing principles which ought to be observed. It seems fair to say that expected utility theory, based on the ingenious work of von Neumann and Morgenstern (1947), has gained most support of all as far as theoretical papers are being concerned. At the same time it is worth noting that a number of paradoxes are known, where human decisions deviate from what is suggested by expected utility theory (cf. Eisenführ and Weber (1999, p. 327)). This has lead to alternative theories, e.g. the prospect theory by Kahnemann and Tversky (1979), which will not be discussed here.

Given expected utility theory is applied, the alternative with the highest expected utility is chosen and risk does not matter as a separate feature. This is not true for risk-value models (cf. Sarin and Weber (1993)). There, an alternative is evaluated taking only its value or worth (e.g. its expected return) and its riskiness (e.g. measured by its variance of returns) into account. Obviously the mean-variance approach widely applied in finance belongs to this class of models. Like other risk-value models it is only compatible with expected utility theory under certain restrictions.

In the present paper, we will concentrate on risk-value models and we will do so in the context of portfolio management. Until quite recently it was common in that area to use the variance or the standard deviation of returns as risk measures (the term volatility being used for return variations over time). Both evaluate deviations from the mean symmetrical, i.e. shortfalls (sometimes labeled risk in a narrow sense) and excess returns (sometimes labeled chance) are both included.

Spectacular crashes caused by traders at Barings and other investment banks caused in particular the supervisors' interest to focus on losses. They were concerned only with outcomes below certain targets which were, for example, given by initial investments. In the sequel this triggered applications of a measure termed "Value at Risk" (VaR). It measures that shortfall from a target that, within a certain holding period, is only exceeded with (at most) a given probability (e.g. 1 %). This measure is now widely applied in financial institutions. However, its normative foundation was not examined carefully and by now a number of papers have shown serious shortcomings (cf. Artzner et al. (1997), Guthoff et al. (1997), Johanning (1998)). This calls for a second thought on shortfall risk and an axiomatic approach as initiated by Artzner et al. (1996). It is generally possible, and also desirable, to determine utility functions which are compatible with risk-value models using specific risk measures. Such an exercise (cf. Sarin and Weber (1993)) reveals value judgements implicit in the applied risk measure. In the present paper we will yet refrain from such an analysis, concentrating on "statistical" risk measures as opposed to "economic" risk measures.[1]

There are numerous definitions of risk available in the literature. One of the more influential papers is Rothschild and Stiglitz (1970). They present three equivalent definitions, each of them intuitively appealing in a different way. Later it was shown that their notion of one distribution being riskier than another is also equivalent to a concept used to compare the inequality of income distributions (cf. Atkinson (1970), Dasgupta et al. (1973), Rothschild and Stiglitz (1973)). Nermuth (1993) and Eichhorn and Vogt (1990) present surveys of the relation between risk measures and inequality measures.

We have stressed earlier that many risk measures take the whole distribution into account. This is also true for the inequality measures. Shortfall risk, or downside risk as we will call it henceforth, by contrast only evaluates a lower part of the distribution (to be defined in precise terms later). But there is an analogy for this in welfare theory as well: Poverty measures only care for poor people (to be defined later, too) and therefore also ignore the upper part of the (income) distribution. To the best of our knowledge, and certainly to our surprise, this analogy does not seem to have been exploited yet.

Probably the most important conceptual contribution of our paper, apart from presenting a useful checklist for measures of downside risk, will be the transfer of axioms from poverty measurement to the measurement of downside risk. This is not always a trivial exercise. Without going into many details, we comment on two of the conceptual differences which mainly make the analysis somewhat difficult.

Firstly, poverty measures are often calculated for a population of $N \in \mathbb{N}$ individuals (or households). This facilitates the intuitive explanation of axioms considerably. In portfolio management we could identify events with people, but at least for practical applications one would like to work with a continuous distribution function instead of

[1] Such a distinction, where "economic" represents the explicit use of utility functions has some history in price measurement and inequality measurement (cf. Diewert and Montmarquette (1983), Bossert and Pfingsten (1990)).

a discrete number of events. We will formally allow for both which will generate a few problems.

Secondly, the incomes even of the poor are usually assumed to be positive (or at least nonnegative). This restriction would be inappropriate in portfolio management. Thus, we will drop it. When discussing specific poverty indices this creates problems because, e.g., the mean income of bad events could be zero or negative and hence may not be suitable as a divisor in an index number. Therefore a suitable re-definition of a known poverty index might be necessary to allow the application as a downside risk measure (cf. Eggers et al. (1999)).

Poverty measures, concentrating on the income situation of the poor, need a basic definition who is called poor. There are two competing views. The first one sees poverty as an absolute concept and assumes that a poverty line z, defining as poor all those with incomes less than z, is exogenously given. Empirical poverty comparisons may be influenced by the choice of z. Therefore research has identified cases, i.e. restrictions on distribution functions, where the ranking of societies or groups according to poverty is independent of z (cf. Foster and Shorrocks (1988)). The second view sees poverty as a relative concept, deriving the poverty line from the income distribution under inspection. It is sometimes suggested, for example, to consider those as poor who have an income below 50% of the mean.

In poverty measurement the choice of a poverty line z is a difficult ethical problem. Measuring downside risk seems to be a little easier. Losses are naturally defined as negative deviations from initial investments or from the (expected) final wealth of a (riskless) portfolio. However, in addition to these two possible benchmarks one could also take a value including information on reserves to cover losses. At any rate, fixing a critical line, which divides losses from gains, relative to the return distribution does not seem to make much sense: Why should an actual return of 2% be interpreted as a loss when the expected return of a portfolio is 5% but not a loss when the latter is 3%?

Throughout our paper we will assume that the distribution functions and poverty lines, respectively critical lines, are all given. We are intentionally abstracting from the measurement issues involved with these data in order to concentrate on the main purpose of our paper.

We will start our analysis of analogies between poverty and downside risk measurement, respectively, in Section 2 with a number of basic concepts and definitions. They are fairly technical but will simplify the axiomatic approach suggested in Section 3. There we present a number of properties which are more or less reasonable requirements for measures of downside risk. The axioms are either transferred from the literature on poverty measurement (cf. the most recent survey of Zheng (1997)) or added as appealing according to our own judgement. Still the list is far from complete. All axioms transferred from poverty measurement will be explained informally in the poverty context first. Since it is sometimes easier to understand the idea of formal axioms by an example, we will check Lower Partial Moments (see below) with respect to the properties. Section 4 concludes.

It was already mentioned that drawing upon the analogy between poverty measurement and the measurement of downside risk is one contribution of our paper. We will not check poverty measures systematically for all the properties suggested. This was done elsewhere (cf. Zheng (1997)) and the one-to-one transfer to downside risk measures is beyond the scope of this paper. However, we think that the axioms as such provide a worthwhile checklist for anybody who is in search for a reasonable downside risk measure.

Axiomatic approaches often lead to impossibility results, Arrow's famous theorem being only the one best known, because usually there are sets of axioms which cannot be satisfied simultaneously. In this sense, our readers are encouraged to make up their own minds about a hierarchy of the axioms presented. It may well be the case that a regulator, e.g., favours other properties than a wealthy investor.

2 Basics

Before risk measures can be introduced, it is useful to get acquainted with the similarities of poverty and downside risk. Basic definitions and concepts are presented as a basis for the understanding of the following sections.

2.1 Poverty and Downside Risk

It has been known for some time (and was discussed quite a bit) that close formal ties between risk and inequality exist. We call income inequality the fact that not all people in a society earn the same income in a given period. Similarly, a distribution of returns is called risky if there are events where returns are different. We feel that the resemblance of poverty and downside risk also is striking: Both have their focus mainly on the lower part of the distribution of payments, concentrating on the income of the poor and the bad outcomes, respectively. After having defined the basic notions of poverty measurement, we want to trace back the measurement of poverty to the measurement of downside risk.

Be the set $\Pi := \{i \in \mathbb{N} | 0 < i \leq N\}$ a population of $N \in \mathbb{N}$ individuals. Let $e_i > 0$ be the income of person i. If z is the poverty line and

$$F(y) := \frac{|\{i \in \Pi | e_i \leq y\}|}{N}$$

the income distribution function,[2] then all necessary information is given to calculate a poverty index $P(F, z)$ (cf. Sen (1976)). $F(y)$ is the share of individuals earning less than or equal to y. F is a monotone right-continuous function with discontinuities of multiples of $1/N$. A poverty index must certainly not be dependent on the names and on the order of the individuals. Thus it is sufficient to use the distribution function F and the poverty line z as the only determinants of the poverty index.

Now let the future yield or value of a given portfolio be characterized by a random variable Y. Let Ω be the set of events, with $Y : \Omega \to \mathbb{R}$. Let P be a probability

[2] $|\Omega| \leq \infty$ denotes the cardinality of a set Ω.

measure on Ω.[3] The corresponding probability distribution function[4] is defined by[5]

$$F(y) := P((-\infty, y]) = P(\{Y \leq y\}).$$

If a probability density function exists, let it be denoted by f with $F(y) = \int_{-\infty}^{y} f(t)\,dt$.

An event $\omega \in \Omega$ defines a probability $P(\omega)$ and a corresponding yield or value of the portfolio $Y(\omega)$ and corresponds to one individual in the concept of poverty. However, a natural relation between the measurement of poverty and the measurement of downside risk exists. If one writes the incomes of the population on N balls, puts them into a box and draws one of them, then the probability distribution function is identical to the income distribution function. The number of individuals N corresponds to the set Ω, the income e_i to the portfolio value $Y(\omega_i)$, and the weight $1/N$ of each person corresponds to the probability of an event $P(\omega_i)$. The poverty line z which divides the poor (earning less than or equal to z) from the rich (earning more than z) corresponds to the critical line that divides critical events with portfolio values less than or equal to z from uncritical events with portfolio values greater than z.

2.2 Transfer of Concepts

In this section, we will introduce some concepts for the modification of distribution functions. This is a vital step in order to establish the system of axioms. The concepts are related to concepts used in the discussion of poverty measurement.[6]

A relative change in the measurement of poverty[7] is defined by the multiplication of the income of each individual as well as the poverty line by the same positive number λ. The analogon in the measurement of downside risk is the multiplication of the outcomes of all events as well as the critical line z with the positive number λ.

Definition 1 (Relative Change) *A relative change is the transition* $(Y, z) \to (\lambda\, Y, \lambda\, z)$ *for* $\lambda \in \mathbb{R}_{++}$.[8] *A relative change can also be characterized by the transition* $(F_Y, z) \to (F_{\lambda Y}, \lambda\, z)$, *where the distribution function of* $\lambda\, Y$ *is defined by* $F_{\lambda Y}(y) = F_Y(y/\lambda)$.

An equal absolute change in the measurement of poverty[9] is defined by the addition of the same positive number τ to the income of each individual as well as to the poverty line. The translation in the measurement of risk is done analogously but, as negative outcomes and a negative critical line are not excluded, τ is allowed to take on negative values in this case.

---

[3] To be precise, P is defined on a σ-algebra on Ω, and further properties of measurability must be given. For a detailed introduction cf. Bauer (1991).

[4] Where misconceptions are impossible, we will say distribution function instead of probability distribution function as well as density function instead of probability density function.

[5] In the following $Y, Y_1, Y_a, \tilde{Y}, Z, \ldots$ denote different random variables. Let $F, F_1, F_a, \tilde{F}, G, \ldots$ be the corresponding probability distribution functions and $f, f_1, f_a, \tilde{f}, g, \ldots$ the corresponding probability density functions.

[6] Cf. Zheng (1997, p. 128).

[7] Cf. Zheng (1997, p. 129).

[8] $\mathbb{R}_{++} := \{y \in \mathbb{R} : y > 0\}$ and $\mathbb{R}_{+} := \mathbb{R}_{++} \cup \{0\}$.

[9] Cf. Zheng (1997, p. 129).

Definition 2 (Absolute Change) *An absolute change is the transition* $(Y, z) \to (Y + \tau, z + \tau)$ *for* $\tau \in \mathbb{R}$*. An absolute change can also be characterized by the transition* $(F_Y, z) \to (F_{Y+\tau}, z + \tau)$*, where the distribution function of* $Y + \tau$ *is defined by* $F_{Y+\tau}(y) = F_Y(y - \tau)$*.*

The increment τ for all events and the critical line z results in a shift of the whole distribution function to the right by τ. This shift leaves the original shape of the curvature unchanged. However, the multiplication of all events by the factor λ leads to a stretching (crushing) of the distribution function. This induces a flatter (steeper) slope of the function in case of $\lambda > 1$ ($\lambda < 1$).

In the following, the set of distribution functions is denoted by $\mathcal{V} := \{F : \mathbb{R} \to [0,1] | F \text{ increasing}, \lim_{y\to-\infty} F(y) = 0, \lim_{y\to\infty} F(y) = 1, F \text{ right-continuous}\}$.

A simple increment in the context of poverty[10] is defined as an additional payment to one person. In risk measurement the analogous concept means that the value in at least one event will be increased. This modifies the density function by shifting "probability mass" to the right, because less (more) profitable events will be less (more) probable. The distribution function is thus caused to move downwards in the relevant part. The shift increases the first moment, i.e. the mean, of the corresponding random variable. It will therefore be called 1^{st} degree bonus.

Definition 3 (1^{st} Degree Bonus) *A distribution function* F_2 *is achieved by a* 1^{st} *degree bonus from* F_1 *(or a random variable* Y_2 *by a* 1^{st} *degree bonus from* Y_1*), if and only if*[11]

$$\exists\, \underline{y}, \overline{y} \in \mathbb{R} : \underline{y} < \overline{y} \text{ and}$$
$$F_1(y) - F_2(y) \begin{cases} = 0 : -\infty < y < \underline{y} \\ > 0 : \quad \underline{y} < y < \overline{y} \\ = 0 : \quad \overline{y} \leq y < \infty \end{cases}.$$

Outside the interval $[\underline{y}, \overline{y}]$, F_1 and F_2 are equal, inside the interval $(\underline{y}, \overline{y})$, F_2 is smaller than F_1. Hence F_2 dominates F_1 according to first-degree stochastic dominance. In particular, the mean of Y_2 is greater than that of Y_1.

A 1^{st} degree malus is defined symmetrically to the 1^{st} degree bonus: F_2 is achieved by a 1^{st} degree malus from F_1 if and only if F_1 is achieved by a 1^{st} degree bonus from F_2.

Definition 4 (2^{nd} Degree Bonus) $F_3 \in \mathcal{V}$ *is achieved by a* 2^{nd} *degree bonus from* $F_1 \in \mathcal{V}$*, if and only if a* $F_2 \in \mathcal{V}$ *exists in the following form:*

- F_2 *is derived from* F_1 *by a* 1^{st} *degree bonus,* $G_1 := F_2 - F_1$*,*
- F_3 *is derived from* F_2 *by a* 1^{st} *degree malus,* $G_2 := F_3 - F_2$*,*
- $\exists\, c \in \mathbb{R}_+ : G_1(y) = -G_2(y + c)\ \forall\, y \in \mathbb{R}$*,*
- $\{\, y \in \mathbb{R} : G_1(y) \neq 0\} \cap \{\, y \in \mathbb{R} : G_2(y) \neq 0\} = \emptyset$.[12]

[10] Cf. Zheng (1997, p. 129).

[11] Note that there are no additional restrictions that could be claimed for the difference $F_1 - F_2$ at the point $\underline{y}$. Because F_1 and F_2 are right-continuous, $F_1 - F_2$ must also be right-continuous, and therefore $F_1(\underline{y}) - F_2(\underline{y})$ is determined by its right neighborhood.

[12] The closure of the set $\{\, y \in \mathbb{R} : G(y) \neq 0\}$ is called the support of G.

A 2^{nd} degree bonus denotes a composition of a 1^{st} degree bonus and a 1^{st} degree malus in such a way that a less profitable event gets a higher payment (a 1^{st} degree bonus) and a more profitable event gets a lower payment (a 1^{st} degree malus), but in total with an unchanged mean.

Remark 1 *If $F_3 \in \mathcal{V}$ is achieved by a 2^{nd} degree bonus from $F_1 \in \mathcal{V}$, then $E[F_3] = E[F_1]$.*

Proof:[13] Let $G_1(y)$ and $G_2(y)$ be defined as in Definition 4. Let $G(y) := G_1(y)$ with $G_1(y) \leq 0$, thus $G_2(y) = -G(y-c)$ and $F_3(y) - F_1(y) = G(y) - G(y-c)$. For reasons of comprehensibility, we prove the remark only for differentiable distribution functions.

$$\begin{aligned}
F_3(y) =& F_1(y) + G(y) - G(y-c) \\
f_3(y) =& f_1(y) + g(y) - g(y-c) \\
E[F_3] =& \int_{-\infty}^{\infty} y\,(f_1(y) + g(y) - g(y-c))\,dy \\
=& E[F_1] + \int_{-\infty}^{\infty} y\,g(y) - y\,g(y-c))\,dy \\
=& E[F_1] + \int_{\underline{y}}^{\overline{y}} y\,g(y)\,dy - \int_{\underline{y}+c}^{\overline{y}+c} y\,g(y-c)\,dy \\
=& E[F_1] + \int_{\underline{y}}^{\overline{y}} y\,g(y)\,dy - \int_{\underline{y}}^{\overline{y}} (y+c)\,g(y)\,dy \\
=& E[F_1] + \int_{\underline{y}}^{\overline{y}} c\,g(y)\,dy \\
=& E[F_1] + c\,(G(\overline{y}) - G(\underline{y})) = E[F_1] + c\,(0-0) = E[F_1].
\end{aligned}$$

For non-differentiable distribution functions, the proof can be completed by a limit argument. ■

The second moment, i.e. the variance, of the corresponding random variable will decrease by a 2^{nd} degree bonus.

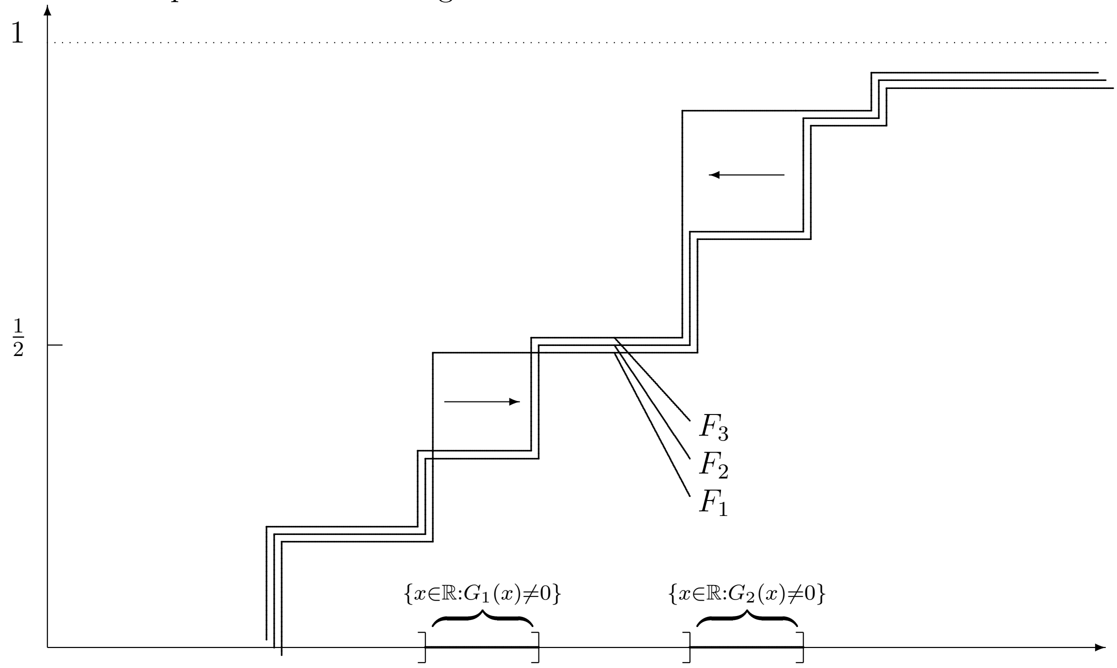
The opposite of the 2^{nd} degree bonus is the 2^{nd} degree malus.[14] Thus the 2^{nd} degree malus requires that F_2 in the definition above is built by a 1^{st} degree malus and F_3 by a 1^{st} degree bonus. Thereby, again, the mean will stay unchanged and the variance will increase which is called a mean preserving spread. In the next step we can combine a 2^{nd} degree bonus and a 2^{nd} degree malus in such a way, that the first and the second moments of the random variable do not change. If the 2^{nd} degree bonus occurs below (above) the 2^{nd} degree malus, we get a 3^{rd} degree bonus (malus).

This process of combining a bonus and a malus of the same degree can be repeated. In general, one can combine a $(n-1)^{th}$ degree bonus and a $(n-1)^{th}$ degree malus

[13] The proof that an n^{th} degree bonus leaves the $(n-1)^{th}$ stochastic moment unchanged can be done in the same way.

[14] The mean preserving spread (for a definition cf. Rothschild and Stiglitz (1970) or Guthoff et al. (1997)) is more general than the 2^{nd} degree malus and used frequently in the literature. Nevertheless, the concept of the 2^{nd} degree malus is sufficient for our purposes.

Figure 1: Example for a Second Degree Bonus on a Discrete Distribution Function



in such a way, that the first $n-1$ stochastic moments of the random variable do not change. If the $(n-1)^{th}$ degree bonus occurs below (above) the $(n-1)^{th}$ degree malus, we receive a n^{th} degree bonus (malus).

Again, a n^{th} degree malus is defined symmetrically to the n^{th} degree bonus: F_3 is obtained by a n^{th} degree malus of F_1 if and only if F_1 is obtained by a n^{th} degree bonus of F_3.

3 Axioms

In the beginning of this chapter, we define the notion of a downside risk measure, for which we introduce axioms in the following. The axioms have to be understood as not only possible, but in addition (more or less) desirable requirements for downside risk measures. Their desirability depends on the context in which risk is to be measured. They are not statements that need to be true for every risk measure.

If one looks for the "best" risk measure in a certain context, one should proceed in the following way: Choose all axioms from the catalogue[15] of axioms in this section which seem to be useful in the special context. If the chosen axioms lead to a contradiction, the choice must be revised. A risk measure is dominated by another if it satisfies only a subset of axioms of the other. It is possible that there is a set of risk measures left in which one does not dominate another. Then a choice can be made by examining further properties.

The axioms of the following sections are required for every choice of a probability distribution and a critical line. This will not be mentioned explicitly in the axioms.

[15] Note that this catalogue does not claim to be complete.

3.1 Basic Axioms and the Definition of Risk Measures

Definition 5 (Downside Risk Measure) *Let $\mathcal{V}$ be defined as in Definition 3, and let $z \in \mathbb{R}$ be called critical line.*[16] *A downside risk measure D is then a function $D : \mathcal{V} \times \mathbb{R} \to \mathbb{R}$ that satisfies the following axioms 1, 2, 3 and 4.*

Axiom 1 (Focus) *If $\tilde{F}$ is obtained from F by a 1^{st} degree bonus with $\underline{y} > z$, then $D(\tilde{F}, z) = D(F, z)$.*

Axiom 2 (Normalization) *If $F(z) = 0$ then $D(F, z) = 0$.*

Axiom 3 (Non-Negativity) *$D(F, z) \geq 0$.*

Axiom 4 (Weak Monotonicity) *If $\tilde{F}$ is obtained from F by a 1^{st} degree bonus, then $D(\tilde{F}, z) \leq D(F, z)$.*

Let us for the sake of comprehensibility start with comments on the definition and each of the axioms. The definition implies that for a given z two different portfolios with the same distribution function lead to the same risk measure. The definition is based on distribution functions F rather than on random variables Y. A random variable Y determines its distribution function F, but not vice versa.[17]

In financial theories like Markowitz's portfolio theory (cf. Markowitz (1952)) it is common to define risk as the dispersion of possible outcomes. However, risk measures like the now commonly used Value at Risk (*VaR*[18]) disregard the positive deviation of the outcome from the reference value (e.g. the target) and concentrate on the lower part of the distribution function. Focusing on the lower part of the distribution function is also the notion used in poverty measurement as formulated in Zheng (1997). The Focus Axiom (Axiom 1) itself can be spelled: Changes of the distribution above the critical line z do not affect the downside risk measure D, which is independent of payments above z.

Normalization (Axiom 2) is more or less a convention or a calibration of the risk measure. According to Zheng (1997, p. 140) the poverty measure must vanish if there are no poor people. We likewise claim the downside risk measure to vanish if the probability of payments below the critical line equals zero. A random variable with no realizations below the critical line is "riskless" in the sense that the downside risk measure of the random variable is zero.

Axioms 3 and 4 claim that a downside risk measure always is a nonnegative number, and that a first degree bonus never leads to an increase of risk. This implies that the measure reacts in the "right" direction: increasing payments will never increase (downside) risk.

[16] The idea is that z divides the elementary events into critical events (with values below or equal to the critical line) and uncritical events (with payments above).

[17] Different portfolios can lead to the same distribution function, if they are affected by the possible events in such a way that the possible outcomes and the according probabilities are equal.

[18] A precise definition and analysis will be presented below.

Remark 2 *Suppose D_1, D_2 are downside risk measures, $\lambda_1, \lambda_2 > 0$ are real numbers, $\phi : \mathbb{R} \to \mathbb{R}$ is a monotone increasing function with $\phi(0) = 0$. Then*

1. $\lambda_1 D_1 + \lambda_2 D_2$ *and* $D_1 D_2$ *are downside risk measures,*
2. $\max(D_1, D_2)$ *and* $\min(D_1, D_2)$ *are downside risk measures,*
3. $\phi(D_1)$ *is a downside risk measure.*

Let us now define three specific examples for downside risk measures.

Definition 6 (Lower Partial Moment Zero) *The Lower Partial Moment Zero (LPM_0) is defined as*

$$LPM_0(F, z) := \int_{-\infty}^{z} f(y)\, dy = P(Y \leq z) = F(z).$$

Focus Axiom[19], Normalization Axiom[20] Non-Negativity Axiom[21] and Weak Monotonicity Axiom[22] are satisfied by LPM_0. It is therefore a downside risk measure in the sense of Definition 5.

Definition 7 (Lower Partial Moment One) *The Lower Partial Moment One (LPM_1) is defined as*

$$LPM_1(F, z) := \int_{-\infty}^{z} (z - y)\, f(y)\, dy.$$

The LPM_1 fulfills axioms 1 to 4, as can be shown similarly as for the LPM_0.

Definition 8 (Value at Risk) *Let α be a fixed number called confidence level with $0 < \alpha < 1$. The Value at Risk (VaR) relating to α is defined as*[23]

$$VaR(F, z) := \max\{0, z - F^{-1}(1 - \alpha)\}.$$

The VaR fulfills Focus Axiom[24], Normalization Axiom[25], Non-Negativity-Axiom[26] and the Weak Monotonicity Axiom[27] and is therefore a downside risk measure, too. VaR and LPM_0 are more or less the "inverse" functions of one another.[28]

We use the VaR and LPM_0 as examples in the following. Each of the following axioms will be checked for whether VaR and LPM_0 satisfy or violate it. When the LPM_0 fails, we additionally use the LPM_1 as an example.

---

19 $LPM_0(F, z) = F(z)$ by definition, thus $LPM_0(F, z)$ depends on F only at z.

20 $LPM_0(F, z) = F(z)$, thus $LPM_0(F, z) = 0 \iff F(z) = 0$.

21 $LPM_0(F, z) = F(z)$, thus $LPM_0(F, z) \geq 0$.

22 $\tilde{F}(z) \leq F(z)$ for all z if $\tilde{F}$ is a 1^{st} degree bonus of F.

23 Cf. Jorion (1997, pp. 87) and Hartmann-Wendels et al. (2000, pp. 330). The definition is not quite correct, because the inverse of F does not necessarily exist. A precise definition would include $\inf\{y \in \mathbb{R} | F(y) \leq 1 - \alpha\}$ instead of the possibly illdefined $F^{-1}(1 - \alpha)$.

24 If $VaR(F, z) = 0$ then $z \leq F^{-1}(1 - \alpha)$ which implies $z \leq \tilde{F}^{-1}(1 - \alpha)$ and therefore $VaR(\tilde{F}, z) = 0$. If $VaR(F, z) > 0$ then $z > F^{-1}(1 - \alpha)$ which implies $\tilde{F}^{-1}(1 - \alpha) = F^{-1}(1 - \alpha)$ and therefore $VaR(\tilde{F}, z) = VaR(F, z)$.

25 $F(z) = 0 \implies F^{-1}(1 - \alpha) > z \implies z - F^{-1}(1 - \alpha) < 0$, which implies $VaR(F, z) = \max\{0, z - F^{-1}(1 - \alpha)\} = 0$.

26 $\max\{0, z - F^{-1}(1 - \alpha)\} \geq 0$ trivially.

27 $\tilde{F} \leq F(z) \implies \tilde{F}^{-1}(1 - \alpha) \geq F^{-1}(1 - \alpha) \implies VaR(\tilde{F}, z) \leq VaR(F, z)$.

28 For a given α, $VaR(F, z) = \max\{0, z - F^{-1}(1 - \alpha)\} = \max\{0, z - [LPM_0(F, \cdot)]^{-1}(1 - \alpha)\}$.

3.2 Continuity Axioms

Before formalizing axioms of continuity, it is necessary to say why they are useful. We want to give two reasons:

1. Assume an investor is interested in the risk of his portfolio. However, he does not know the precise probability distribution of payments but only some estimate. When using a risk measure for portfolio optimization, he wants to be sure that, provided that his estimate is "close" to the actual distribution, the risk measure of the estimate is close to the actual risk, too.

2. Assume the trader at a bank has to report the probability distribution of his trades' yields to the head office, which is interested in riskiness and expected yield. The trader has the possibility to report distorted data. The head office can discover larger deceit and expect the trader to distort only little. It wants therefore a risk measure which does not react too strongly when the data are manipulated only "a little".

Thus whenever the basis for the calculation of the risk measure deviates for whatever reason "a little" from the actual distribution function, the calculated risk must not deviate too much from the actual risk either. This requirement is formalized in the Continuity Axiom and the Lipschitz Continuity Axiom below.

But what does "a little" mean? In the case of real numbers, the answer is clear: the distance between two real numbers is the absolute of the difference. In case of random variables, the answer is more difficult. We choose

Definition 9 (Distance of Random Variables) *Let*

$$d(Y_1, Y_2) := \|F_1 - F_2\|_1 = \int |F_1(y) - F_2(y)|\, dy$$

define the distance between two random variables.

There is a number of justifications for this definition:

- If and only if two random variables lead to the same distribution function, their distance is zero and they are therefore regarded as equal, as we are not interested in the actual random variable, but in the probabilities of payments.

- Let $Y, \tilde{Y} : \Omega \to \mathbb{R}$ with $Y \equiv \tilde{Y}$ on $\Omega \setminus \{\omega_0\}$, then $d(Y, \tilde{Y}) = P(\{\omega_0\})\, |Y(\omega_0) - \tilde{Y}(\omega_0)|$. Thus the impact of a misestimation (manipulation) is proportional to the misestimation (manipulation) itself, if the probability of the manipulated event $P(\omega_0)$ is greater than zero.[29]

The above definition of a distance leads directly to a concept of continuity that corresponds to the concept used in Zheng (1997, p. 131).

[29] $P(\{\omega_0\}) > 0$ is not possible if F is continuous.

Axiom 5 (Continuity) *$D(F,z)$ is continuous in F for a fixed z. Thus if $(Y_i)_{i\in\mathbb{N}}$ and Y are random variables, $(F_i)_{i\in\mathbb{N}}$ and F the corresponding distributions with*

$$\lim_{i\to\infty} \|F_i - F\|_1 = 0,$$

then

$$\lim_{i\to\infty} D(F_i, z) = D(F, z)$$

must hold.

In words: As the estimate approaches the actual distribution function, the risk measure of the estimate approaches the actual risk measure, too.

Remark 3 *The LPM_0 does not fulfill Continuity.*

Proof: Consider a critical line $z := 0$ and a sure payment of variable size:

$$F(y) := \begin{cases} 0 & : \quad y < 0 \\ 1 & : \quad y \geq 0 \end{cases} \qquad \text{and} \quad F_i(y) := \begin{cases} 0 & : \quad y < 1/i \\ 1 & : \quad y \geq 1/i \end{cases}.$$

Then

$$\lim_{i\to\infty} \|F - F_i\|_1 = \lim_{i\to\infty} 1/i = 0, \qquad \text{but}$$
$$\lim_{i\to\infty} LPM_0(F_i, 0) = \lim_{i\to\infty} F_i(0) = 0 \neq 1 = F(0) = LPM_0(F, 0).$$

Thus LPM_0 is not continuous in the first parameter.[30] ■

Neither is the VaR continuous.[31]

Remark 4 *The LPM_1 fulfills Continuity.*

Proof: It is sufficient to prove $|LPM_1(F,z)| \leq \|F\|_1$ for all F, z.[32] We have

$$\begin{aligned} |LPM_1(F,z)| &= \int_{-\infty}^{z} (z-y)\, f(y)\, dy \\ &= \int_{-\infty}^{z} z\, f(y)\, dy - \int_{-\infty}^{z} y\, f(y)\, dy \\ &= z\, F(z) - \left[y\, F(y)\Big|_{-\infty}^{z} - \int_{-\infty}^{z} F(y)\, dy \right] \\ &= \int_{-\infty}^{z} F(y)\, dy \leq \int_{-\infty}^{\infty} F(y)\, dy = \|F\|_1. \end{aligned}$$

■

If Continuity is proven for a risk measure, it is of interest to know whether there is an upper bound for the error in the risk measure given that an upper bound for the error in the distribution function is known.

---

30 Note that one can also prove the result using continuous functions F and F_i.

31 This can be proven similarly for all $0 < \alpha < 1$ by setting $F(y) := \begin{cases} 0 & : \quad y < z-1 \\ 1 & : \quad y \geq z-1 \end{cases}$ and $F_i(y) := \begin{cases} 0 & : \quad y < z-2 \\ \alpha\,(1-1/i) & : \quad z-2 \leq y < z-1 \\ 1 & : \quad y \geq z-1 \end{cases}$.

32 Cf. Rudin (1991, p. 24).

Axiom 6 (Lipschitz Continuity) *$D(F, z)$ is Lipschitz-continuous in F for fixed z, i.e. there is an $L \in \mathbb{R}_+$ such that for all random variables Y_1, Y_2*

$$|D(F_1, z) - D(F_2, z)| \leq L \, \|F_1 - F_2\|_1 \tag{1}$$

holds.

Note that Axiom 6 implies Axiom 5.[33] Therefore VaR and LPM_0, already failing to fulfill Axiom 5, can neither fulfill Axiom 6. LPM_1 is Lipschitz continuous with $L = 1$.[34]

It is also reasonable to claim that a risk measure depends continuously on the critical line z. Otherwise a slight change in z could greatly influence the risk assessment.

Axiom 7 (Critical Line Continuity) *$D(F, z)$ is continuous in z for a fixed F, i.e., if F is a distribution function and $z_i, z \in \mathbb{R}$ with*

$$\lim_{i \to \infty} |z_i - z| = 0,$$

then

$$\lim_{i \to \infty} D(F, z_i) = D(F, z)$$

must hold.

With an argument similar to the one above, it can be shown that Axiom 7 is not met[35] by the LPM_0, but it is met by the VaR.[36]

One could, of course, formulate many more axioms, like e.g. a Lipschitz Axiom for the critical line or even axioms on differentiability with respect to the distribution or the critical line. We feel, however, that these would most likely be of less intuitive appeal.

3.3 Invariance and Gauge Axioms

In this section we will deal with two more groups of axioms, invariance and gauge axioms. The former require downside risk measures to be invariant under certain types of joint changes of the distribution function and the critical line. The latter pose restrictions on the range of the measures.

At first glance, the two groups of axioms seem to be so different that the analysis under one heading looks odd. However, it turns out that some gauge axioms do not make sense (and in fact cannot be satisfied) in conjunction with some of the invariance axioms.

[33] To see this, look at (1) when L is fixed.
[34] See the proof of Remark 4.
[35] Choose F as in the proof of Remark 3 and $z := 0$, $z_i := -1/i$.
[36] As a function of z only, $VaR(z) = \max\{0, z - c\}$, which is clearly continuous for any $c \in \mathbb{R}$.

3.3.1 Invariance Axioms

Having said that invariance axioms require constancy of a measure under specific changes in its arguments, the natural question arises what kinds of changes are such that we would be prepared to make normative statements. The Focus Axiom is a basic invariance axiom which we have introduced right at the beginning. Building upon the literature on inequality and poverty measurement once again, relative changes and absolute changes are further natural candidates.

The Scale Invariance Axiom is derived from Zheng (1997, p. 138). The interpretation of scale invariance in the risk context is similar to the one in the poverty context. If scale invariance is fulfilled, then the poverty measure and also the downside risk measure are unaffected by the unit or the currency in which income or payments are measured. Therefore it does not matter whether the payments are in deutschmarks, euro or dollar: The downside risk measure is unit-free.

Axiom 8 (Scale Invariance)

$$\forall\, \lambda \in \mathbb{R}_{++} : D(F_{\lambda Y}, \lambda\, z) = D(F_Y, z).$$

The LPM_0,[37] in contrast to the VaR,[38] fulfills this axiom. One must choose carefully whether to require Scale Invariance or not. Imagine tossing a coin: in one situation, you can win or lose 1 deutschmark, in another, 1000 deutschmarks are at stake. Intuitively, one would say the second situations is riskier. However, if a downside risk measure satisfies Scale Invariance, its value is the same in both situations[39], if $z = 0$. The reader should notice that the Scale Invariance Axiom requires in particular that the distributions $F_{\lambda y}$ and F_y are assigned the same downside risk if evaluated with respect to the critical line $z = 0$. In this case, a change in the currency would look just like a change in the "size" of the lottery mentioned above. This conceptual problem appears in the literature on inequality measurement as well as in the price index literature in connection with the commensurability axiom.[40]

The Homogeneity Axiom demands the downside risk measure to react linearly on stretching and crushing of the distribution function. If e.g. the random payments of a portfolio are measured once in deutschmarks and once in euro, then the downside risk measures must differ by the exchange rate[41]. The homogeneity of risk measures is also required by Artzner et al. (1996, p. 3, Property 2.3).

Axiom 9 (Homogeneity)

$$\forall\, \lambda \in \mathbb{R}_{++} : \frac{D(F_{\lambda Y}, \lambda\, z)}{\lambda} = D(F_Y, z).$$

[37] $F_{\lambda Y}(z) = F_Y(z/\lambda) \implies LPM_0(F_{\lambda Y}, \lambda\, z) = F_{\lambda Y}(\lambda\, z) = F_Y(\lambda\, z/\lambda) = F_Y(z)$.

[38] Cf. footnote 43.

[39] $F_{1000Y}(y) = F_Y(y/1000)$ implying $D(F_{1000y}, 1000z) = D(F_Y, z)$ if Scale Invariance holds.

[40] Cf. Eichhorn (1988) and Fuchs-Seliger et al. (1986).

[41] Thus the measure has the same dimension (e.g. DM, hfl, £, ...) as the random variable.

This axiom directly contradicts[42] the Scale Invariance Axiom. It is met by the *VaR*.[43] The LPM_0 fulfills the Scale Invariance Axiom, hence it cannot fulfill the Homogeneity Axiom. LPM_1 is homogeneous.[44]

The next axiom also defines the reaction of measures to a class of changes. Formulated in Zheng (1997, p. 138) the Translation Invariance Axiom requires that the poverty measure remains unchanged under an absolute change. In the context of risk, this can be reformulated as follows. If a bank receives a sure additional payment and the critical line z rises by the same amount, then the risk measure should show an unchanged risk. The risk measure is then translation invariant.

Axiom 10 (Translation Invariance)

$$\forall\, \tau \in \mathbb{R} : D(F_{\tau+Y}, \tau + z) = D(F_Y, z).$$

VaR[45] as well as LPM_0[46] fulfill translation invariance.

The invariance axioms require further inspection with respect to their normative content. Among others, it needs to be checked how sensible the axioms are when applied to distributions of different variables, e.g. cash flows, losses (measured in money terms), or yields. Comparable requirements can be found in Artzner et al. (1997, p. 3, 4, Properties 2.3, 2.5).

3.3.2 Gauge Axioms

Gauge axioms normalize the values a downside risk measure may obtain in different ways. The reader may notice that our basic axioms 2 and 3, Normalization and Non-Negativity, are gauge axioms, too.

If the risk measure is a relative measure, the image can be standardized to the unit interval.

Axiom 11 (Unit Interval) $D(F, z) \leq 1$.

As the LPM_0 is a percentage, it fulfills[47] axiom 11. The *VaR* does not, because it satisfies Homogeneity (axiom 9).

---

42 To see this, set $\lambda = 2$. The only downside risk measure that satisfies both of the axioms is the trivial measure $D \equiv 0$. This does also imply that no downside risk measure can satisfy every axiom.

43 $F_{\lambda Y}(y) = 1 - \alpha = F_Y(y/\lambda)$. Taking the inverse of each of the equations yields $y = F^{-1}_{\lambda Y}(1-\alpha)$ and $F_Y^{-1}(1-\alpha) = y/\lambda$. This implies $\lambda\, F_Y^{-1}(1-\alpha) = F^{-1}_{\lambda Y}(1-\alpha)$, and therefore $\lambda\, VaR(F_Y, z) = VaR(F_{\lambda Y}, \lambda\, z)$. The *VaR* is therefore not scale invariant.

44 The substitution rule of integration (with $\phi(y) = y/\lambda$) yields $LPM_1(F_{\lambda Y}, \lambda\, z) = \int_{-\infty}^{\lambda z} (\lambda\, z - y)\, f(y/\lambda)/\lambda\, dy = \int_{-\infty}^{z} (\lambda\, z - \lambda\, y) f(y)\, \lambda/\lambda\, dy = \lambda\, LPM_1(F_Y, z)$.

45 This can be proven in the same way as the homogeneity of the *VaR*. $F_{Y+\tau}(y) = 1 - \alpha = F_Y(y - \tau)$. Taking the inverse of each of the equations yields $F^{-1}_{Y+\tau}(1 - \alpha) = y$ and $F_Y^{-1}(1 - \alpha) = y - \tau$. This implies $F_Y^{-1}(1 - \alpha) - \tau = F^{-1}_{Y+\tau}(1 - \alpha)$, and therefore $VaR(F_Y, z) = VaR(F_{Y+\tau}, z + \tau)$.

46 $F_{Y+\tau}(y) = F_Y(y - \tau) \implies LPM_0(F_{Y+\tau}, z+\tau) = F_{Y+\tau}(z+\tau) = F_Y(z+\tau-\tau) = F_Y(z) = LPM_0(F_Y, z)$.

47 $LPM_0(F, z) = F(z) \implies 0 \leq LPM_0(F, z) \leq 1$.

If a downside risk measure has the dimension of a currency, one can require the Limitedness Axiom: the value of the downside risk measure should not be greater than the maximal potential loss. This requirement has already been formulated by Artzner et al. (1997, p. 2, Property 2.1).

Axiom 12 (Limitedness) *Be* $\underline{y}_F := \inf\{y|F(y) > 0\}$*. Then* $\underline{y}_F < z$ *implies*

$$D(F, z) < z - \underline{y}_F.$$

The LPM_0 fulfills Scale Invariance (axiom 8) and therefore cannot fulfill Limitedness.[48] The VaR is limited.[49]

Table 1: The Contradictions between Gauge and Invariance Axioms. Crosses (×) stand for contradictions. Examples are put forward when there are no contradictions.

	A 8	A 9	A 10	A 11	A 12
(A 8) Scale Invariance	—	×	LPM_0	LPM_0	×
(A 9) Homogeneity		—	VaR	×	VaR
(A 10) Translation Invariance			—	LPM_0	VaR
(A 11) Unit Interval				—	$\min\{LPM_0, VaR\}$
(A 12) Limitedness					—

Remark 5 *The Unit Interval Axiom (axiom 11) is independent of the Scale Invariance Axiom (axiom 8).*

Proof: One can convert any downside risk measure D into a downside risk measure D_1 fulfilling axiom 11 by defining $D_1 := \frac{D}{D+1}$. Then D_1 satisfies axioms 11, and it satisfies the basic axioms because D does. ■

Let us now prove two incompatibilities between the introduced invariance and gauge axioms.

Remark 6 *The only risk measure* D *that satisfies Homogeneity (Axiom 9) and Unit Interval (Axiom 11) is* $D(F, z) \equiv 0$.

Proof: If $D \not\equiv 0$, then there exist $F_Y \in \mathcal{V}$ and $z \in \mathbb{R}$ with $D(F_Y, z) = d > 0$. Then $D(F_{2Y/d}, z) = 2\,d/d > 1$, which is a contradiction to Unit Interval. ■

Remark 7 *The only risk measure* D *that satisfies Scale Invariance (Axiom 8) and Limitedness (Axiom 12) is* $D(F, z) \equiv 0$.

The remark can be proved symmetrically to Remark 6. ■

48 Cf. Remark 7 on page 16.

49 If $VaR(F, z) > 0$ then $F^{-1}(1 - \alpha) > \underline{y}_F$, and therefore $VaR(F, z) \leq z - \underline{y}_F$.

Excluding the trivial index $D \equiv 0$, we have thus contradictions between Scale Invariance and Homogeneity, between Scale Invariance and Limitedness and between Unit Interval and Homogeneity. There are downside risk measures that satisfy Scale Invariance and Unit Interval (e.g. the LPM_0), and downside risk measures that satisfy Homogeneity and Limitedness (e.g. the LPM_1), although there are implications in neither direction. Note that there is no contradiction between Unit Interval and Limitedness.[50]

3.4 Impact Axioms

In contrast to section 3.3.1, the axioms of this section do not require downside risk measures to be invariant under certain classes of changes, but to change in a specific way under certain other classes of changes. These changes can affect either the distribution function or the critical line.

3.4.1 Impact of Changes in the Distribution Function

The following axioms are ordered by two criteria. Firstly, axioms are divided into reagibility axioms and sensitivity axioms. Reagibility axioms require the downside risk measure to react in a specific way under certain changes, sensitivity axioms compare the impacts of different changes of the same type. Secondly, both types of axioms come in different degrees. This provides a twodimensional pattern of axioms. The higher degree axioms are more subtle, there is no point claiming higher degree axioms for a downside risk measure that already fails fulfilling lower degree axioms (cf. Figure 2).

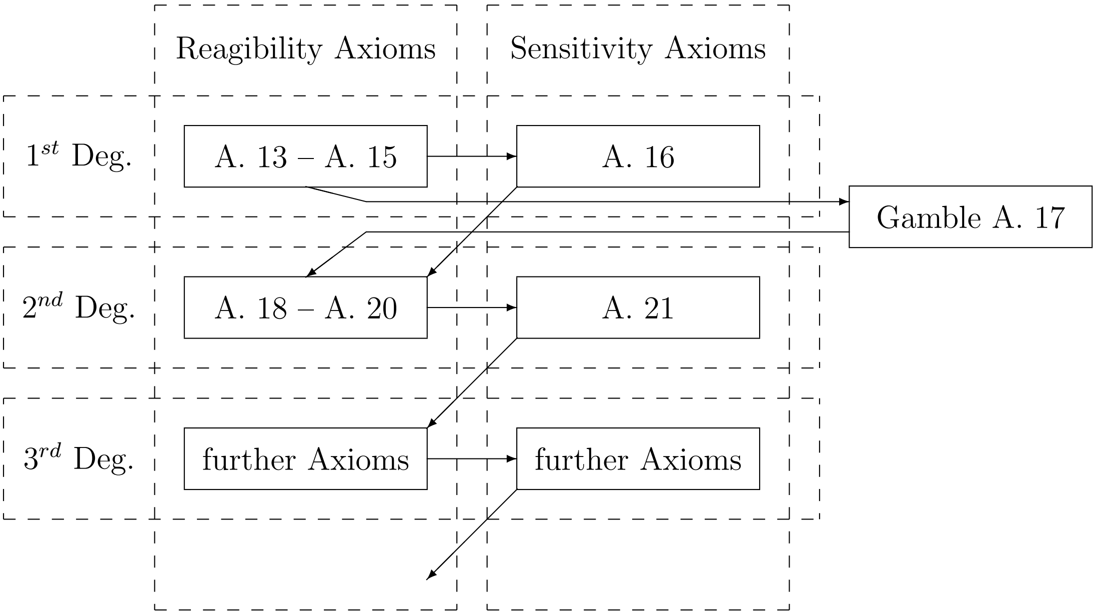
First Degree Axioms

First Degree Distribution Reagibility Axioms The following axiom is derived from the context of poverty measures (cf. Zheng (1997, p. 131)). It originally expresses in essence: *If money is taken from a poor person, the poverty measure should increase.* We require that a downside risk measure should increase if the payment in a critical event is diminished, which seems natural for a non-satiated investor. In the discrete case, the decrease of a payment in one event implies a 1^{st} degree malus. That is the reason why we, deriving downside risk measures from distribution functions rather than from payments in different events, base the following axioms on 1^{st} degree bonuses or maluses. This makes the axioms also applicable to the continuous case. As 1^{st} degree bonuses and maluses define two particular numbers ($\underline{y}$ and $\overline{y}$), we can distinguish the three cases $z < \underline{y}$, $\underline{y} < z < \overline{y}$ and $\overline{y} < z$. Because of the Focus Axiom (axiom 1), the first case does not affect the value of the poverty measure, and we receive two semi-strong monotonicity axioms.

Axiom 13 (Semi-Strong Monotonicity 1) *If F_2 is obtained from F_1 by a 1^{st} degree malus below z (i.e. with $\overline{y} < z$), then $D(F_2, z) > D(F_1, z)$.*

[50] To see this, let D_U be a downside risk measure satisfying Unit Interval, D_L a downside risk measure satisfying Limitedness. Then $D := \min(D_U, D_L)$ satisfies both of the axioms.

Figure 2: The Interrelation between the Impact Axioms.
Arrows stand for necessary conditions.



We want to stress the fact that this inequality is strict. It does not hold for the LPM_0.

Remark 8 *The* LPM_0 *does not satisfy Semi-Strong Monotonicity 1.*

Proof: Let

$$F_1(y) := \begin{cases} 0 & : \quad y < -1 \\ 0.1 & : \quad -1 \leq y < 10 \\ 1 & : \quad 10 \leq y \end{cases} \quad \text{and} \quad F_2(y) := \begin{cases} 0 & : \quad y < -2 \\ 0.1 & : \quad -2 \leq y < 10 \\ 1 & : \quad 10 \leq y \end{cases}.$$

Then F_2 *is* a 1^{st} degree malus of F_1 in a critical event, but $D(F_1, 0) = D(F_2, 0) = 10\%$. This proves that the Semi-Strong Monotonicity Axiom 1 is not met by the LPM_0. ■

At least, the LPM_0 is never increased by a 1^{st} degree malus.) Nor does it for the *VaR*.[51]

Axiom 14 (Semi-Strong Monotonicity 2) *If* F_2 *is obtained from* F_1 *by a* 1^{st} *degree malus around* z *(i.e. with* $\underline{y} < z < \overline{y}$*), then* $D(F_2, z) > D(F_1, z)$.

This axiom holds for the LPM_0,[52] but it does not for the *VaR*.[53]

The application of the following axiom to poverty measurement (cf. Zheng (1997)) is in essence: *If money is given to a poor person, the poverty measure is due to fall.*

Axiom 15 (Strong Monotonicity) *If* F_2 *is obtained from* F_1 *by a* 1^{st} *degree malus in a critical event (therefore* $\underline{y} < z$*), then* $D(F_2, z) > D(F_1, z)$.

[51] Choose $F^{-1}(1-\alpha) > \overline{y}$.
[52] $\underline{y} < z < \overline{y} \implies F_2(z) < F_1(z) \implies LPM_0(F_2, z) < LPM_0(F_1, z)$ by definition.
[53] Choose $F^{-1}(1-\alpha) < \underline{y}$.

The Strong Monotonicity implies both of the Semi-Strong Monotonicity Axioms.[54] LPM_1 satisfies Strong Monotonicity.[55]

Although the LPM_0 satisfies none of the following axioms, they are still useful to assess the quality of downside risk measures with better properties than the LPM_0, e.g. the LPM_1.

First Degree Distribution Sensitivity Axioms The following axiom is again transferred from the context of poverty (cf. Zheng (1997)), where it essentially expresses: *If a fixed amount of money is taken from a poor person, the poverty measure should react the stronger, the poorer the person is.* In terms of risk, the measure should react the stronger on a payment reduction in an event the lower the original payment in the initial event is. This reads in the continuous case:

Axiom 16 (Monotonicity Sensitivity) *If F_1 and F_2 are obtained from F by a 1^{st} degree malus below the critical line respecting $G_1 = F_1 - F\ (\geq 0)$, $G_2 = F_2 - F\ (\geq 0)$ and*

$$\exists\, c \in \mathbb{R}_{++}\ : G_1(y) = G_2(y + c)\ \forall\, y \in \mathbb{R},$$

then $D(F_1, z) > D(F_2, z)$.

LPM_1 satisfies Monotonicity Sensitivity.[56] LPM_0 and VaR do not (cf. Figure 2).

Second Degree Axioms

The following axiom is different from all the other distribution sensitivity axioms. It is not, like the others, transferred from the concepts of poverty (cf. Zheng (1997)), but a new suggestion. It is quite intuitive, we think.

If a risk-averse investor is offered — in addition with his portfolio — an uncorrelated gamble with mean zero, he should refuse.[57] Therefore, the value of the downside risk measure of the distribution function with the additional gamble should be higher than that of the downside risk measure of the original distribution.

Axiom 17 (Additional Gamble)

$$D(\frac{F_{Y-\varepsilon} + F_{Y+\varepsilon}}{2}, z) < D(F, z) \textit{ for all } \varepsilon > 0 \textit{ if } F(z) > 0. \tag{2}$$

The LPM_0 does not fulfill the axiom. However, it does fulfill inequality (2) for small ε and fixed (twice differentiable) F if and only if $F''(z) > 0$. The inequality holds for the LPM_0 and many natural choices of Y and z, e.g. if f_Y has only got one local maximum at some y^* (e.g. if Y is normally distributed) and $z < y^*$. The VaR does not fulfill Axiom 17.[58]

[54] Thus the LPM_0 cannot meet Strong Monotonicity.

[55] To be proved similarly to Remark 1.

[56] To be proved similarly to Remark 1.

[57] The idea is that adding noise to the distribution increases risk (cf. Rothschild and Stiglitz (1973)).

[58] For z sufficiently large, $VaR(F_{Y-\varepsilon}, z) - \varepsilon = VaR(F_{Y-\varepsilon}, z) = VaR(F_{Y+\varepsilon}, z) + \varepsilon$, and therefore $VaR((F_{Y-\varepsilon} + F_{Y+\varepsilon})/2, z) = VaR(F, z)$. $F_{Y+\varepsilon}$ is a transfer of F_Y, cf. Definition 2.

Second Degree Distribution Reagibility Axioms The following axioms are very similar, each of them is translated into the context of risk measurement from Zheng (1997), where they read roughly: *If a richer person gives to a poorer person and somewhere else on the income scale a poorer person gives the same amount to a richer person, then the effect of the money transfer in the poorer part of the population should dominate the other.*

As 2^{nd} degree bonuses/maluses define four particular numbers (lower and upper bound of the support of G_1 and lower and upper bound of the support of G_2), we can distinguish five cases. Two of the cases are already covered by first degree axioms, one is covered by the Focus Axiom (axiom 1). Thus we receive two new 2^{nd} degree distribution sensitivity axioms.

Axiom 18 (Semi-Strong 2^{nd} Degree Distribution Reagibility 1) *If F_2 is obtained from F_1 by a 2^{nd} degree malus with* $\inf\{y \in \mathbb{R} : G_2(y) \neq 0\} < z < \sup\{y \in \mathbb{R} : G_2(y) \neq 0\}$*, then $D(F_2, z) > D(F_1, z)$.*

Axiom 19 (Semi-Strong 2^{nd} Degree Distribution Reagibility 2) *With notation as in Definition 4, if F_3 is obtained from F_1 by a 2^{nd} degree malus with* $\sup\{y \in \mathbb{R} : G_2(y) \neq 0\} < z$*, then $D(F_3, z) > D(F_1, z)$.*

As already mentioned, the LPM_0 does not fulfill any of these axioms. A 2^{nd} degree bonus can even lead to an increase of the VaR.[59]

Axiom 20 (Strong 2^{nd} Degree Distribution Reagibility) *With notation as in Definition 4, if F_3 is obtained from F_1 by a 2^{nd} degree malus with* $\inf\{y \in \mathbb{R} : G_1(y) \neq 0\} < z$*, then $D(F_1, z) > D(F_3, z)$.*

Second Degree Distribution Sensitivity Axioms

Axiom 21 (2^{nd} Degree Distribution Sensitivity) *If F_1 and F_2 are derived from F by a 2^{nd} degree malus below the critical line with $G_1 = F_1 - F$, $G_2 = F_2 - F$ and*

$$\exists\, c \in \mathbb{R}_{++}\ \forall\, y \in \mathbb{R} : G_1(y) = G_2(y + c),$$

then $D(F_1, z) < D(F_2, z)$.

n^{th} Degree Axioms

One can now come to other more or less desirable properties of downside risk measures by recursion, following the principle: "If a change in the distribution is 'good' for the downside risk measure, then the impact should be the greater, the lower the payments are where the change takes place", which leads to sensitivity axioms, as well as "If the same change is carried out on the same distribution function simultaneously at two different places, the effect of the change taking place at the lower part of the distribution should dominate", which leads to new reagibility axioms.

59 Cf. Guthoff et al. (1997).

3.4.2 Impact of Shifting the Critical Line

A poverty measure should depend positively on the poverty line (cf. Zheng (1997)). In other words, if the poverty line is risen, more (at least not less) people become poor by definition and also the distance between the income of the already poor and the poverty line increases. Therefore the poverty measure should increase as well.

Similarly, downside risk measures must also depend positively on the critical line. We distinguish between two gradings of critical line axioms; the strong form (axiom 23) again implies the semi-strong form (axiom 22). The Strong Increasing Critical Line Axiom claims that the downside risk measure rises whenever the critical line increases. The Semi-Strong Increasing Critical Line Axiom is weaker and claims only that the downside risk measure must rise at least if the probability of payments between the old and the new critical line is positive.

Axiom 22 (Semi-Strong Increasing Critical Line) *If* $z_1 < z_2$ *and* $F(z_2) - F(z_1) > 0$ *then* $D(F, z_1) < D(F, z_2)$.

Axiom 23 (Strong Increasing Critical Line) *If* $z_1 < z_2$ *and* $F(z_2) > 0$ *then* $D(F, z_1) < D(F, z_2)$.

Trivially[60], the LPM_0 fulfills the Semi-Strong, but not the Strong Increasing Critical Line Axiom. The VaR meets none of the axioms.[61]

3.5 Subgroup Axioms

As mentioned in the Introduction, there exist concepts of absolute and relative poverty measurement in the literature. Assume a population is divided into two subgroups. Then, following the relative concept of poverty, one could imagine the value of the poverty measure of the whole population to rise even if the poverty measures in each of the subgroups decrease. This would happen e.g. if the wealthier of the subgroups becomes even richer, the poorer subgroups stays nearly unchanged, so that the gap between the subgroups widens. When using the absolute concept of poverty, the poverty change of the entire population should be consistent with the poverty changes of subgroups of the population.

As described in section 2.1, the persons of the population are "translated" into elementary events in the context of risk. Therefore, instead of subpopulations, events (that consist of several elementary events) must be examined. As the set of subgroups sums up to the whole population again, the events are supposed to be a partition of the event space Ω.

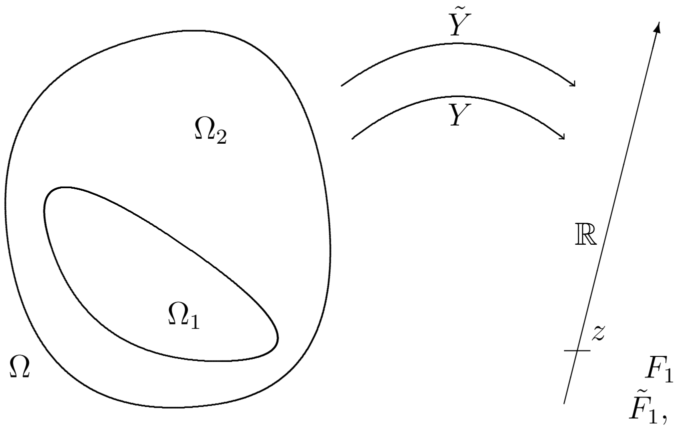
The first subgroup axiom expresses in the context of poverty: *Assume that a population is split into several subgroups. If poverty in one subgroup decreases and remains unchanged in the other groups, then poverty of the whole population is due to decrease, too.*[62]

[60] $z_1 < z_2 \Longrightarrow \forall F \in \mathcal{V} : [LPM_0\,(F, z_1) < LPM_0\,(F, z_2) \iff F(z_2) - F(z_1) > 0]$.

[61] To prove this, choose small z_1 and z_2.

[62] Cf. Zheng (1997, p. 135).

Figure 3: Risk Measurement on Disjunct Events



$\Omega_2 = \Omega \setminus \Omega_1$
$Y, \tilde{Y} : \Omega \to \mathbb{R}$
F_1, F_2 distrib. of $Y|_{\Omega_1}, Y|_{\Omega_2}$
$\tilde{F}_1, \tilde{F}_2$ distrib. of $\tilde{Y}|_{\Omega_1}, \tilde{Y}|_{\Omega_2}$

This requirement makes economic sense in the risk context, too. Suppose, there is an important possible future random event[63] E_1 (e.g. the bancruptcy of a big multinational). Then the set of possible events is split into two subgroups: one subgroup on which E occurs, and its complement. If an investor can then reduce the risk of his portfolio given the case E_1 occurs, leaving the risk unchanged given E_1 does not occur, then the aggregated risk measure should decrease as well.

Axiom 24 (Subgroup Consistency) *Let* $F = \lambda\, F_1 + (1-\lambda)\, F_2$ *($\tilde{F}$ analogous) with* $0 < \lambda < 1$. *If* $D(F_1, z) < D(\tilde{F}_1, z)$ *and* $D(F_2, z) = D(\tilde{F}_2, z)$, *then*

$$D(F, z) < D(\tilde{F}, z).$$

Here, λ denotes the probability of the event in which the risk is reduced ($F_1 \to \tilde{F}_1$), accordingly $1-\lambda$ denotes the probability of the unchanged complement ($F_2 = \tilde{F}_2$). The LPM_0 fulfills the axiom,[64] the VaR does not.[65]

Suppose now that all events are partitioned in two sets with different (or equal) downside risk. Then we require that the aggregate downside risk measure lies between

[63] $\Omega_1 := \{\omega \in \Omega | E_1\}, \Omega_2 := \Omega \setminus \Omega_1$.

[64] This is easy to see:

$$\begin{aligned}
LPM_0(F_{X_1}, z) = F_{X_1}(z) &< F_{Y_1}(z) = LPM_0(F_{Y_1}, z),\\
LPM_0(F_{X_2}, z) = F_{X_2}(z) &= F_{Y_2}(z) = LPM_0(F_{Y_2}, z) \implies\\
\lambda\, F_{X_1}(z) + (1-\lambda)\, F_{X_2}(z) &< \lambda\, F_{X_1}(z) + (1-\lambda)\, F_{X_2}(z) \implies\\
LPM_0(F_X, z) = F_X(z) &< F_Y(z) = LPM_0(F_Y, z).
\end{aligned}$$

[65] For a counterexample, choose $\alpha := 97.5\%, z := 0, \lambda := 1/2$,

$$Y_1 := \{-1 \text{ with } p = 3\%, 1 \text{ with } p) = 97\%\}, Y_2 := 1,$$
$\tilde{Y}_1 := \{-1 \text{ with } p = 2\%, 1 \text{ with } p = 98\%\}, \tilde{Y}_2 := 1$. This implies $D(F, z) = D(\tilde{F}, z) (= 0)$.

the values for the two sets. In the context of poverty measurement, the poverty of an entire population should lie between the poverties of the two subpopulations.

Axiom 25 (Mean) *Let* $F := \lambda F_1 + (1-\lambda) F_2$. *If* $D(F_1, z) \le D(F_2, z)$, *then*

$$D(F_1, z) \le D(F, z) \le D(F_2, z)$$

This axiom is met by LPM_0[66] and VaR.[67]

The following decomposibility axiom does not only provide bounds for the aggregated poverty measure, but even a formula for its calculation. It implies linearity for the downside risk measure. Thus, if decomposibility is given, one can apply the tools of linear functional analysis (cf. Rudin (1991)) on risk measurement. In poverty literature a decomposability axiom is formulated (cf. Zheng (1997, p. 136)) as a special case of the subgroup consistency axiom. It says that the poverty value of a population is equal to the weighted sum of the poverty values of the subgroups, where the weights are equal to the sizes of the subgroups relative to the whole group.

The decomposibility axiom is on the one hand restrictive, on the other hand it is useful: If Ω is finite, then the downside risk measure can be calculated as the weighted sum of the downside risk measures of the elementary events.

Axiom 26 (Decomposability) *Using the above notation,*

$$D(F, z) = \lambda D(F_1, z) + (1-\lambda) D(F_2, z).^{68}$$

The LPM_0 is a decomposable downside risk measure,[69] the VaR is not.[70]

The two following axioms are derived from Zheng (1997, p. 137): If a poor person vanishes from the population or a rich person appears, then the poverty measure should decrease. The translation of this is neither obvious, nor is it clear whether it is desirable in the context of risk at all. The counterparts of persons are elementary events, but events cannot "appear" or "vanish". That is why the appearance of persons is interpreted as mergers of the population with one-inhabitant-populations (which are poor or not poor, respectively). Therefore, the axioms are classified as subgroup axioms.

[66] $F_1(z) < \lambda F_1(z) + (1-\lambda) F_2(z) < F_2(z) \iff F_1(z) < F_2(z)$.

[67] Assume $0 < VaR(F_1, z) \le VaR(F_2, z)$, and that inverse functions exist. Then $F_2^{-1}(1-\alpha) \le F_1^{-1}(1-\alpha)$, and in the interval $[F_2^{-1}(1-\alpha), F_1^{-1}(1-\alpha)]$, we have $F_1 \le F \le F_2$. Therefore, $F(F_2^{-1}(1-\alpha)) \le 1-\alpha$ and $1-\alpha \le F(F_1^{-1}(1-\alpha))$, thus $F_2^{-1}(1-\alpha) \le F^{-1}(1-\alpha) \le F_1^{-1}(1-\alpha)$, which completes the proof.

[68] This axiom was first proposed for risk measurement by Huang (1971). Weber and Bottom (1989) demonstrate some empirical evidence for its violation in measurement of perceived risk by individuals.

[69] $LPM_0(F, z) = F(z) = \lambda F_1(z) + (1-\lambda) F_2(z) = \lambda LPM_0(F_1, z) + (1-\lambda) LPM_0(F_2, z)$.

[70] Counterexample: distribution $\tilde{Y}$ in footnote 65.

Axiom 27 (Growth of Safety) *With notation as above, if* $F_2(z) = 0$ *then*

$$D(F, z) \leq D(F_1, z).$$

This axiom is met by LPM_0[71] and VaR.[72]

Axiom 28 (Growth of Risk) *With notation as above, if* $F_2(z) = 1$ *then*

$$D(F, z) \geq D(F_1, z).$$

The LPM_0 fulfills this axiom trivially,[73] the VaR does not.[74]

4 Conclusion and Outlook

In our paper, we have demonstrated the interrelation between the measurement of poverty in a society and the measurement of downside risk. Because of the close relation, we were able to transfer the concepts of poverty measurement in the form of axioms to a context of downside risk. We have added further axioms and modified others, because the interrelation is not one to one. We have shown how to apply the axioms to risk measures, in our case to the LPM_0, and we have found that the LPM_0 (and likewise the Value at Risk to which it is closely related) does not fulfill many of the axioms. Thus there is scope for improvements based on our detailed catalogue of desirable properties.

The most obvious application is to examine whether several downside risk measures, either genuine risk measures or measures transferred from the poverty context, fulfill the axioms. By doing so, one will be able to create a hierarchical ordering[75] for downside risk measures. From the set of "efficient" downside risk measures[76], one can then choose the measure fitting to the very application. Eggers et al. (1999) derive a downside risk measure from poverty measurement and then examine its properties.

Furthermore, one can examine implications and inconsistencies between the axioms. A focus of future research may lie on downside risk measures that fulfill the Decomposability Axiom (axiom 26). Such risk measures are related to additively separable poverty measures (Atkinson (1987)) and possess interesting properties related to utility theory. Furthermore, the axioms formulated in our paper have to be compared with different axioms of rational decision making (e.g. von Neumann and Morgenstern (1947)). Last but not least, empirical tests will have to show how sensitive the assessment of the downside risk of portfolios is with respect to the choice of measures.

---

[71] $F_2(z) = 0 \implies F_2(z) \leq F_1(z) \implies F(z) \leq F_1(z)$.

[72] The claim follows from $VaR(F_2, z) = 0$ and the Mean axiom.

[73] $F_2(z) = 1 \implies F_2(z) \geq F_1(z) \implies F(z) \geq F_1(z)$.

[74] Counterexample: Choose $F_2(z) = 1$, but $F_1^{-1}(1-\alpha) < F_2^{-1}(1-\alpha)$.

[75] One downside risk measure dominates another if it meets every axiom the other meets and at least one more. Therefore, the ordering will in general not be complete, but only partial.

[76] In this sense, "efficient" risk measures are measures that are not dominated.

Table 2: The Properties of the presented DRM
Here, 'Y' indicates that a property is met, '$\times$' that it is not met.

	VaR	LPM_0	LPM_1
(A 5) Continuity	$\times$	$\times$	Y
(A 6) Lipschitz Continuity	$\times$	$\times$	Y
(A 7) Critical Line Continuity	Y	$\times$	
(A 8) Scale Invariance	$\times$	Y	
(A 9) Homogeneity	Y	$\times$	Y
(A 10) Translation Invariance	Y	Y	Y
(A 11) Unit Intervall	$\times$	Y	$\times$
(A 12) Limitedness	Y	$\times$	Y
(A 13) Semi-Strong Monotonicity 1	$\times$	$\times$	Y
(A 14) Semi-Strong Monotonicity 2	$\times$	Y	Y
(A 15) Strong Monotonicity	$\times$	$\times$	Y
(A 16) Monotonicity Sensitivity	$\times$	$\times$	Y
(A 17) Additional Gamble	$\times$	$\times$	
(A 18) Semi-Strong 2^{nd} Degree Distribution Reagibility 1	$\times$	$\times$	
(A 19) Semi-Strong 2^{nd} Degree Distribution Reagibility 2	$\times$	$\times$	
(A 20) Strong 2^{nd} Degree Distribution Reagibility	$\times$	$\times$	
(A 21) 2^{nd} Degree Distribution Sensitivity	$\times$	$\times$	
(A 22) Semi-Strong Increasing Critical Line	$\times$	Y	
(A 23) Strong Increasing Critical Line	$\times$	$\times$	
(A 24) Subgroup Consistency	$\times$	Y	Y
(A 25) Mean	Y	Y	Y
(A 26) Decomposability	$\times$	Y	Y
(A 27) Growth of Safety	Y	Y	Y
(A 28) Growth of Risk	$\times$	Y	$\times$

References

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber and David Heath (1996), A Characterization of Measures of Risk, Technical Report 1186, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York.

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber and David Heath (1997), Thinking Coherently, *Risk*, vol. 10, pp. 68–71.

Atkinson, Anthony B. (1970), On the Measurement of Inequality, *Journal of Economic Theory*, vol. 2, pp. 244–263.

Atkinson, Anthony B. (1987), On the Measurement of Poverty, *Econometrica*, vol. 55, pp. 749–764.

Bauer, Heinz (1991), Wahrscheinlichkeitstheorie, de Gruyter, Berlin.

Bossert, Walter and Andreas Pfingsten (1990), Intermediate Inequality: Concepts, Indices, and Welfare Implications, *Mathematical Social Sciences*, vol. 19, pp. 117–134.

Dasgupta, Partha, Amartya Sen and David Starrett (1973), Notes on the Measurement of Inequality, *Journal of Economic Theory*, vol. 6, pp. 180–187.

Diewert, W. Erwin and Claude Montmarquette (eds.) (1983), Price Level Measurement, Statistics Canada, Ottawa, Ontario.

Eggers, Frank, Andreas Pfingsten and Sven Rieso (1999), Three Dimensions of Shortfall Risk: Sen's Poverty Index Revisited, Technical Report 99-04, Institut für Kreditwesen, Westfälische Wilhelms-Universität Münster.

Eichhorn, Wolfgang (1988), On a class of Inequality Measures, in: Wulf Gaertner and Prasanta K. Pattanaik (eds.), *Distributive Justice and Inequality*, Springer, Berlin, pp. 83–89.

Eichhorn, Wolfgang and Arthur Vogt (1990), Gemeinsames bei der Messung von Ungleichheit, Streuung, Risiko und Information, *Mitteilungen der Schweizerischen Vereinigung der Versicherungsmathematiker*, pp. 53–73.

Eisenführ, Franz and Martin Weber (1999), Rationales Entscheiden, 3. ed., Springer, Berlin.

Foster, James E. and Anthony F. Shorrocks (1988), Poverty Orderings, *Econometrica*, vol. 56, pp. 173–177.

Fuchs-Seliger, Susanne, Ulrich Niemeyer and Andreas Pfingsten (1986), An Analysis of Statistically-Motivated Tests in Economic Price Index Theory, in: Martin J. Beckmann, Karl-Walter Gaede, Klaus Ritter and Hans Schneeweiss (eds.), *Methods of Operations Research (Vol. 54)*, Anton Hain, Königstein/Ts., pp. 233–244.

Guthoff, Anja, Andreas Pfingsten and Juliane Wolf (1997), Effects on Risk Taking Resulting from Limiting the Value at Risk or the Lower Partial Moment One, in: Institute of Actuaries of Australia (ed.), *7th International AFIR Colloquium Proceedings*, Southwood Press, Sydney, pp. 355–378.

Hartmann-Wendels, Thomas, Andreas Pfingsten and Martin Weber (2000), Bankbetriebslehre, 2nd ed., Springer, Berlin.

Huang, L.C. (1971), The expected risk function, Technical Report 71-6, Michigan Mathematical Psychology Program Report.

Johanning, Lutz (1998), Value-at-Risk zur Marktrisikosteuerung und Eigenkapitalallokation, Uhlenbruch, München.

Jorion, Philippe (1997), Value at Risk, Irwin, Chicago.

KAHNEMANN, Daniel and Amos TVERSKY (1979), Prospect theory: An analysis of decision under risk, *Econometrica*, vol. 47, pp. 263–291.

MARKOWITZ, Harry M. (1952), Portfolio Selection, *Journal of Finance*, vol. 7, pp. 77–91.

NERMUTH, Manfred (1993), Different Economic Theories with the Same Formal Structure: Risk, Income Inequality, Information Structures, etc., in: W. Erwin DIEWERT, Klaus SPREMANN and Frank STEHLING (eds.), *Mathematical Modelling in Economics*, Springer, Berlin, pp. 271–277.

ROTHSCHILD, Michael and Joseph E. STIGLITZ (1970), Increasing Risk I: A Definition, *Journal of Economic Theory*, vol. 2, pp. 225–243.

ROTHSCHILD, Michael and Joseph E. STIGLITZ (1973), Some Further Results on the Measurement of Inequality, *Journal of Economic Theory*, vol. 6, pp. 188–204.

RUDIN, Walter (1991), Functional Analysis, Mc Graw-Hill, New York.

SARIN, Rakesh K. and Martin WEBER (1993), Risk-Value Models, *European Journal of Operational Research*, vol. 70, pp. 135–149.

SEN, Amartya (1976), Poverty: An Ordinal Approach to Measurement, *Econometrica*, vol. 44, pp. 219–231.

von NEUMANN, John and Oskar MORGENSTERN (1947), Theory of Games and Economic Behavior, Princeton University Press, New York.

WEBER, Elke U. and William P. BOTTOM (1989), Axiomatic Measures of Perceived Risk: Some Tests and Extensions, *Journal of Behavioral Decision Making*, vol. 2, pp. 113–131.

ZHENG, Buhong (1997), Aggregate Poverty Measures, *Journal of Economic Surveys*, vol. 11, pp. 123–162.