Westfälische Wilhelms-Universität Münster

Institut für Kreditwesen

Direktor: Prof. Dr. Andreas Pfingsten

Institut für Kreditwesen

Diskussionsbeitrag 01-01

The Properties of Downside Risk Measures

Dipl.-Kfm. Carsten Breitmeyer

Dipl.-Math. Hendrik Hakenes

Prof. Dr. Andreas Pfingsten

ISSN 0949-6610

März 2001

The Properties of Downside Risk Measures

Carsten Breitmeyer†
Hendrik Hakenes†
Andreas Pfingsten†

May 28, 2001

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

Abstract

A close formal relationship between poverty measurement and the measurement of downside risk was established by Breitmeyer et al. (1999). Moreover, they show that well-known requirements for poverty measures can motivate reasonable properties of downside risk measures, too. We build on their work by transforming a number of poverty measures into (new) downside risk measures. Also, we test these measures (as well as some previously suggested ones) with respect to their axioms. This sometimes is a tedious task, but is simplified considerably for decomposable measures for which we prove some fairly general features. A table summarizes the tests and can serve as a useful tool for anybody who is concerned with downside risk, e.g. risk managers, investors or regulators.

JEL classification: D81, G11, G20, I32

Key Words: risk measures, downside risk, poverty measurement, shortfall risk measurement, decomposable measures, Value at Risk, Lower Partial Moments

Contents

1 Introduction

The theory of finance over the past 50 years has passed a number of important milestones. Among the most significant contributions are the introduction of modern portfolio theory by Markowitz (1952) and the creation of the now standard capital asset pricing model in a series of independent papers by Sharpe (1964), Lintner (1965), and Mossin (1966). In all of these μ - σ - approaches, risk is being represented by the variance of returns, and it is probably fair to say that at least until quite recently many people in the profession would almost have taken risk and variance as being synonyms. However, it has also been pointed out clearly that this identity does not hold. For instance, it is now well established that there exist distributions X and Y with X having a higher mean and a lower variance, but where Y will be chosen by some risk-averse, perfectly rational, expected utility maximizing individual. Also, from a behavioral finance point of view, Brachinger and Weber (1997) have shown by examining empirical studies that the variance insufficiently represents perceived risk.

Rothschild and Stiglitz (1970) provided a number of appealing definitions of risk, which are closely related to second-order stochastic dominance and turned out to be equivalent (cf. Atkinson (1970), Dasgupta et al. (1973), and Rothschild and Stiglitz (1973)). These approaches, while being able to explain the "variance paradox" mentioned above, are still restricting risk measurement to be a symmetrical evaluation in the sense that both upside and downside deviations from the mean matter.

This is different in a line of research linked to the seminal paper of Roy (1952). He not only called the minimum-variance portfolio on a Markowitz efficiency frontier the safety-first portfolio, but also introduced the notion of shortfall risk (using the term disaster). In this concept only shortfalls matter, i.e. negative deviations from a target. If investments are chosen such that the probability of a shortfall is minimized then this is called a safety-first strategy (cf. Elton and Gruber (1995) for a good account of the theory and further references).

It is usually a matter of personal preferences whether an investor will stick to the variance notion of risk or rather use a shortfall concept. Banks, however, have to care for downside risk almost by assumption. Given the widespread view that they have to be regulated and supervised in order to protect depositors in particular and the financial system in general, a view we will neither defend nor challenge here, authorities try to induce them to be managed as to avoid failures whenever possible. Consequently, regulators are focused on the downside part of the distribution, an attitude that has become quite visible when Value at Risk was accepted as a risk measure for market risk (cf. Basle Committee on Banking Supervision (1996)).

Value at Risk, i.e. that shortfall below a target that is only exceeded with a given probability (1 %, say), has soon become popular (cf. Jorion (1997)). In our opinion it is not unlikely that this has happened mainly because Value at Risk is easy to understand and fairly easy to calculate, in particular with the support provided by RiskMetrics™ introduced by J.P. Morgan. While studies on the sensitivity of Value at Risk with respect to different ways to calculate it and efficient computation methods became issues early (e.g., Beder (1995), Hendricks (1996), and Linsmeier and Pearson (1996)), critical comments were not available for some time. This has changed by now. In particular Artzner et al. (1999), in an article which had been available as a preliminary working paper since 1996, have provided a rigorous analysis. Starting from first principles they conclude that Value at Risk should not be regarded as a very appealing measure and is inferior to their class of coherent risk measures.

We appreciate the axiomatic approach, but will follow a different track. Breitmeyer et al. (1999) have investigated similarities between the measurement of poverty and downside risk measurement. They adapt the conceptual framework of the former to produce requirements for the latter. The general idea is that for particular changes in distribution functions we are prepared to make normative statements how this should affect downside risk. In this sense, their approach carries same relation to the Rothschild and Stiglitz (1970) approach to define increasing risk, yet it concentrates on the downside part of distributions. In their paper, Breitmeyer et al. (1999) have tested for only a few downside risk measures (Value at Risk, Lower Partial Moment 0, and partly for Lower Partial Moment 1) whether or not they fulfill a long list of potentially desirable axioms. The first contribution of our paper therefore is a far more extensive coverage of indices.

In the process of transferring axioms from poverty measurement to the measurement of downside risk some obstacles are in the way. Firstly, in poverty measurement a discrete number of individuals is commonly assumed whereas in portfolio management one would like to work with a continuous distribution function (although in practice, e.g. in the Monte Carlo models some banks apply, a continuous distribution is often only approximated by a finite number of a few thousand observations). Thus the axioms, and more importantly the measures, have to be partly reformulated to be applicable for continuous distributions. Secondly, in poverty measurement incomes are usually assumed to be positive (or at least nonnegative, but certainly with a positive mean). The yields or the changes of the value of a portfolio can be negative. In fact we have seen real world examples where even the expected value of the daily change in the portfolio value was negative for a large number of days. This causes some problems when poverty indices are to be transformed into downside risk measures. Some of the initial measures are not defined for negative values, others are not even defined for zero incomes. In particular whenever some expression is divided by the mean (what is often done in poverty measurement for reasons that will become clear later) this is obviously problematic. The second contribution of our paper, therefore, is the presentation of solutions (or at least suggestions) for those of the mentioned problems that were still open issues.

When transferring indices from poverty measurement to the measurement of downside risk we cannot be exhaustive (cf. Zheng (1997) for a wealth of measures). We will however suggest a procedure how to do this transformation for further measures. That is, some new downside risk measures as well as a way to generate even more constitute our third contribution.

We build on the work of Breitmeyer et al. (1999) and refer to it repeatedly. In order to make our paper self-contained without replicating theirs, we have decided to collect important definitions, concepts and axioms in an appendix. Our own analysis starts in section 2 with our definition of a downside risk measure. After that we introduce some genuine downside risk measures conforming to this very definition. After a motivation of how to transform a poverty measure into a downside risk measure, we are able to write down the definitions of some poverty measures and their corresponding downside risk measures. The formulation and classification of the poverty measures we use is based on Zheng (1997). section 3 is devoted to the analysis which of the proposed axioms decomposable downside risk measures fulfill. We find that in this case the analysis can be simplified considerably by deriving some general observations for elements of this class. In section 4 we check how some important non-decomposable measures perform with respect to the axioms. section 5 winds up with a table which summarizes our results, i.e. it shows which measures satisfy which axioms, and section 6 concludes.

2 A Selection of Downside Risk Measures

In this chapter, the notion of a downside risk measure is defined, and a number of downside risk measures is introduced. A downside risk measure is a function that aggregates the distribution of a random variable, e.g. of a future portfolio value or of a lottery outcome, into a real number. This real number is then supposed to indicate the riskiness of the random variable. The risk measure is called downside risk measure, because it takes into account only the part of the distribution which is below a critical line z. 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 values above).

Definition 1 (Downside Risk Measure) *Let $I \subseteq \mathbb{R}$ be an interval on $\mathbb{R}$, and be $\mathcal{V}$ the set of probability distributions on I:*

$$\mathcal{V} := \{F : I \to [0,1] | F \textit{ increasing}, \inf_{y \in I} F(y) = 0, \sup_{y \in I} F(y) = 1, F \textit{ right-continuous}\}.$$

Let $z \in I$ be called critical line. A downside risk measure D is then a function $D : \mathcal{V} \times I \to \mathbb{R}$ 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 $D(\tilde{F}, z) \leq D(F, z)$.*

The definition of a 1^{st} degree bonus, like other concepts and axioms, is provided in the appendix. Its notion is that the random variable is increased in a sub-interval of I, where $\underline{y}$ is the lower bound of this sub-interval (cf. definition 6 on page 27). The four axioms are called basic, because they are a prerequisite for all sensible downside risk measures. They exclude obscure mappings D that have nothing in common with the intuitive notion of downside risk.
Still, infinitely many different downside risk measures satisfying the above definition exist. Therefore a selection must be made. We make a representative choice: some measures are genuine risk measures (chapter 2.1), others are derived from poverty measures (chapter 2.2).

2.1 Genuine Downside Risk Measures

The distinction between genuine downside risk measures and those transferred from poverty measurement is not unambiguous. Some genuine downside risk measures (e.g. the LPM_0) do also exist in an equivalent form in poverty measurement (in this case the headcount ratio). The relationships between them will also be shown in the following. Definition 2 collects some selected genuine downside risk measures which will be discussed below.

Definition 2 (Some Genuine Downside Risk Measures)

$$VaR_\alpha(F,z) := \max\{0, z - F^{-1}(1-\alpha)\} \quad \textit{for } \alpha \in [0,1] \quad \textit{Value at Risk}, \tag{1}$$

$$LPM_n(F,z) := \int_{-\infty}^{z} (z-y)^n f(y)\,dy \quad \textit{for } n \in [0,\infty] \quad \textit{Lower Partial Moment } n. \tag{2}$$

The Shortfall-Value of Risk, $S\text{-}VaR(F,z) := z - LPM_1(F,z)/F(z)$ fails to fulfill Axiom 4, therefore it is not a downside risk measure in the sense of definition 1. The measures $LPM_0(F,z) = \int_0^z f(y)\,dy = F(z)$ (Lower Partial Moment 0) and $LPM_1(F,z) = \int_0^z (z-y)\,f(y)\,dy$ (Lower Partial Moment 1) are special cases of the general LPM_n. For $z = E(y)$, the Lower Partial Moment 2 becomes the semi-variance. The LPM_n is commonly defined for natural numbers $n \in \mathbb{N}$ only, but can as well be calculated for nonnegative reals α instead of n.

2.2 The Transformation of Poverty Measures into Downside Risk Measures

Zheng (1997) provides a collection and classification of poverty measures. His classification will be adopted because it gives a good hint how to characterize and how to examine the different measures. We select some measures from each class of his collection and transform them into downside risk measures, some of which are already known and have been previously applied in downside risk measurement. Each measure will be written down in its original form in the context of poverty measurement and as a downside measure. The transformation used will, of course, be also given where it is not evident.

We need to define what we mean by a transformation of a measure because otherwise any downside risk measure could be called a transfer of a poverty measure. The intuition is the following: A downside risk measure is called a transformation of a poverty measure, if the poverty measure for a population is equal to a downside risk measure for a lottery in which the payoffs are the incomes of the individuals of the population, with uniform probability (after suitable normalization). Formally:

Definition 3 (Tranfer of a Poverty Measure) *Suppose a poverty measure P has the form that it maps a vector x of incomes of a population (with $x \in \bigcup_{n=0}^{\infty} \mathbb{R}_+^n$, n the total population, therefore $x = \{x_1, \ldots, x_n\}$) and a poverty line z (with $z \in \mathbb{R}_+$) into the real numbers $\mathbb{R}$. A downside risk measure $\widetilde{D}$ is then called transformation of a poverty measure D if for all $x \in \bigcup_{n=0}^{\infty} \mathbb{R}_+^n$ and for all $z \in \mathbb{R}_+$*

$$D(x;z) = \widetilde{D}(F_x, z) \quad \textit{with} \quad F_x(y) := \sum_{x_i \le y} \frac{\chi_{x_i}(y)}{n}, \tag{3}$$

where $\chi_{x_i}(y) := \{0 \text{ for } x_i < y; 1 \text{ else}\}$ *is the unit step function.*

In the following, the symbol '$\rightsquigarrow$' will be used for the transformation of a poverty measure (e.g. $W(x; z)$) into a downside risk measure (e.g. $\widetilde{W}(F, z)$), where a weaker symbol than '$\Rightarrow$' is used because the transfer is not unambiguous. One important source of ambiguity is the fact that poverty measures are defined on positive numbers only, whereas the portfolio values in downside risk measures might also become negative, e.g. when the portfolio contains short positions. Therefore, the extension of a poverty measure on the negative numbers is not trivial. If the poverty measure has a singularity in zero, then it cannot be continuously extended on the negative numbers at all. However, it is sometimes possible to define a modified transformation of a poverty measure that allows for negative portfolio values, especially if the possible portfolio values are limited downwards. For example, the Watts measure (see below) could be shifted by a positive constant c (which is an absolute change; cf. definition 5), with c e.g. equal to the lowest possible portfolio value. Any modification, however, usually changes some of the measure's properties. We will not discuss such modifications in this paper, as the proofs of the properties (though not the results) would usually be similar to those for the original measures.
Other than poverty measures, downside risk measures are usually defined for continuous distribution functions. In practice, however, downside risk measures are often calculated from a finite set of data, because the actual distribution function is not known, but can only be estimated. For example, in some banks a historical simulation or a Monte Carlo simulation with a fixed number of scenarios are used (cf. Beder (1995)). Based on this observation, one might be tempted to argue that a discrete formulation of a downside risk measure is already sufficient. But for two reasons this is *not* the case. *Firstly*, in order to receive an unbiased estimator for a statistic out of a sample, the statistical formulas would have to be adjusted, just like the formula for the variance has to be adjusted if it is to be inferred from the variance of a sample. Thus, the formula for a poverty measure is *not* the same as the (also discrete) formula for the estimator of a downside risk measure calculated from a sample. *Secondly*, other banks use variance-covariance approaches which are based on continuous distributions (e.g. multivariate normals) and a general downside risk measure should be applicable in these cases, too. *Thirdly*, there are approaches that assume that the relevant probability distribution can be approximated by continuous distributions out of a class of functions with a finite number of parameters. The parameters are then estimated from the data of a sample, and the (downside) risk measure is calculated from the approximating continuous function with the estimated parameters (cf. e.g. the proposals of Sortino and Forsey (1994) and Sortino and Forsey (1996)).

2.2.1 Class 0: Distribution-Insensitive Poverty Measures

Class 0 consists of three simple poverty measures. They do not change when certain changes (e.g., 2^{nd} degree bonus and malus) of the distribution function below the poverty line z occur and are therefore called distribution-insensitive. Their transformation into downside risk measures varies, but is always obvious (and sometimes even already known).

The headcount ratio, which has been widely used in poverty measurement, is defined as the ratio between the number of the poor and the population size. For downside risk measurement this is the probability that an event occurs which has a value below the target. In other words, this is the relative frequency of shortfalls (shortfall probability), which is also known as the LPM_0. In the following, $q(x;z)$ (or simply q) denotes the number of the poor, and $n(x)$ (resp. n) the total number of people.

Measure 1 (Headcount Ratio)

$$\begin{aligned} H(x;z) &= \frac{q(x;z)}{n(x)} \\ \rightsquigarrow \widetilde{H}(F,z) &= F(z) \end{aligned}$$

The transfer of the headcount ratio allows for negative portfolio values and for negative z. With a glance at their definitions it is clear that the headcount ratio is equivalent to the LPM_0.

The next simple poverty measure is the income gap ratio. It is defined as the percentage of the average income shortfall of the poor from the poverty line. By $\mu_p(x;z)$ (or simply μ_p) we denote the average income of the poor. In downside risk measurement the index describes the average shortfall relative to the critical line, i.e. the average severity of shortfalls.

Measure 2 (Income Gap Ratio)

$$\begin{aligned} I(x;z) &= 1 - \frac{\mu_p(x;z)}{z} \\ \rightsquigarrow \widetilde{I}(F,z) &= 1 - \frac{1}{z\,F(z)} \int_0^z y\,f(y)\,dy \end{aligned}$$

Because of division by z, the Income Gap Ratio is not defined for $z = 0$ and cannot be extended on negative portfolio values.

The product of the headcount ratio and the income gap ratio is the poverty measure called poverty gap ratio. This measure was first mentioned by Sen (1976). It will be shown below that this measure satisfies some axioms which both headcount ratio and income gap ratio violate.

Measure 3 (Poverty Gap Ratio)

$$\begin{aligned} HI(x;z) &= H(x;z) \cdot I(x;z) \\ \rightsquigarrow \widetilde{HI}(F,z) &= \widetilde{H}(F,z) \cdot \widetilde{I}(F,z) \end{aligned}$$

There is a relationship between the LPM_1 and the poverty gap ratio. After a few transformations it can be seen that the LPM_1 is equivalent to the poverty gap ratio times the critical line z.

$$
\begin{aligned}
LPM_1(F,z) &= \int_0^z (z-y) f(y) dy \\
&= \int_0^z z\, f(y) dy - \int_0^z y\, f(y) dy \\
&= z\, F(z) - \int_0^z y\, f(y) dy = z\, F(z)\, \left(1 - \frac{1}{z} \int_0^\infty y\, \frac{f(y)}{F(z)}\, dy\right) \\
&= F(z)\, (z - \mu_p) \qquad (4)\\
&= z\, F(z)\, \widetilde{I}(F,z) = z\, \widetilde{H}(F,z)\, \widetilde{I}(F,z) \\
&= z\, \widetilde{HI}(F,z). \qquad (5)
\end{aligned}
$$

Note that, although $\widetilde{H}$ and LPM_1 are defined for negative and positive portfolio values, the domain of $\widetilde{HI}$ is restricted to the positive numbers. The reason is that, according to the above definition (but letting integrals run from $-\infty$ instead of 0) with a multiplication by negative z, $\widetilde{HI}$ would violate the Non-Negativity Axiom and the Weak Monotonicity Axiom. The Poverty Gap Ratio is identical to an index called Normalized Deficit, which has been developed by Watts (1968).

2.2.2 Class 1: Sen-Type Poverty Measures

Sen (1976) demanded that a poverty measure, other than the indices of class 0, be sensitive to changes in the distribution function below z and constructed a poverty measure that has this property. Based on it, Zheng (1997, pp. 144) built up a class of measures. The indices in this class are proposed by Sen himself or are modifications of his indices. Therefore this class is called “Sen-type”. The modifications of Thon (1979), Kakwani (1980a), and Takayama (1979) cannot be transformed into downside risk measures, because they are not replication invariant. This means that if two countries with identical income distributions are united, the resulting poverty differs from the initial poverty. Thus the same income distribution leads to different degrees of poverty, depending on the population size. This implies that only replication invariant poverty measures can be tranferred according to definition 3, as a downside risk measure may depend only on the critical line z and the distribution F. Consequently, we only choose one modification of the classical Sen measure from this class, namely the one suggested in Sen (1976), which is transferable into a downside risk measure.

Measure 4 (Sen)

$$
S'(x;z) = H(x;z)\,[I(x;z) + (1 - I(x;z))G(x;z)] \quad \textit{with} \qquad (6)
$$

$$
G = \frac{1}{2\,q^2\,\mu_p} \sum_{i=1}^{q} \sum_{j=1}^{q} |y_i - y_j| = \frac{1}{2\,\mu_p} \sum_{i=1}^{q} \sum_{j=1}^{q} \frac{1}{q^2}\, |y_i - y_j|
$$

$$
\rightsquigarrow \widetilde{S}(F,z) = F(z)\, \left[\widetilde{I}(F,z) + (1 - \widetilde{I}(F,z))\, \widetilde{G}(F,z)\right] \quad \textit{with}
$$

$$
\widetilde{G}(F,z) = \frac{1}{2}\, \frac{F(z)}{z\, F(z) - LPM_1} \int_0^z \int_0^z |x - y|\, \frac{f(x)}{F(z)}\, \frac{f(y)}{F(z)}\, dx\, dy.
$$

We take a slight modification S' of the original Sen-index,

$$S(x;z) = H(x;z)\left[I(x;z) + (1 - I(x;z))G(x;z)\frac{q}{q+1}\right], \tag{7}$$

and transform it into the downside risk measure $\widetilde{S}$. The step from (7) to (6), the omission of the factor $q/q+1$, has already been proposed by Sen[1] himself, in order to make his measure replication invariant. Because of equation (4) μ_p has been replaced by $(z\,F(z) - LPM_1)/F(z)$. Further replacing and reducing terms yields

$$\begin{aligned}
\widetilde{S}(F,z) &= F(z)\left[\frac{LPM_1}{z\,F(z)} + \frac{z\,F(z) - LPM_1}{z\,F(z)}\,\frac{1}{2}\,\frac{F(z)}{z\,F(z) - LPM_1}\int_0^z\int_0^z |x-y|\ldots\right] \\
&= \frac{F(z)}{z}\left[\frac{LPM_1}{F(z)} + \int_0^z\int_0^y (y-x)\ldots\right] \\
&= \frac{F(z)}{z}\left[\int_0^z \left((z-y) + \int_0^y (y-x)\,\frac{f(x)}{F(z)}\,dx\right)\frac{f(y)}{F(z)}\,dy\right] \qquad (8) \\
&= \frac{F(z)}{z}\left[\int_0^z \left((z-y) + \frac{LPM_1(F,y)}{F(z)}\right)\frac{f(y)}{F(z)}\,dy\right] \\
&= \frac{1}{z}\left[LPM_1(F,z) + \int_0^z LPM_1(F,y)\frac{f(y)}{F(z)}\,dy\right]. \qquad (9)
\end{aligned}$$

While all of the latter formulations are equivalent, it is not obvious which one is the most favorable for computations and proofs. In (9), e.g., the expression is reduced to a fairly simple integral over LPM_1's.
In poverty measurement, as well as in the straightforward transformation for downside risk measurement, the Sen-index is restricted to positive incomes or positive portfolio values. One can see from equation (9) that (a further generalization of) the Sen-index can even be calculated if negative portfolio values are possible. The integrals then have to start from $-\infty$ instead of 0. As the LPM_1's are always well-defined and positive, this does not seem to cause formal difficulties.[2] Thus, the Sen-index can be continued on negative portfolio values, but not for nonpositive z.

2.2.3 Class 2: Ethical Poverty Measures

The Sen-type indices can often be interpreted in terms of social welfare. Along this line, the measures in this class are modifications with respect to the measure of income inequality used or the social welfare function applied. That is the reason why they are labeled as ethical poverty measures. They are all quite different and therefore it is difficult to provide a joint characterization. Generally speaking, their structure and properties depend upon the specific social welfare function used.

[1] Cf. Sen (1973, p. 31).

[2] However, only the Sen-index defined on positive numbers exclusively is going to be discussed below, as a modification changes properties of the index (e.g. the Unit Interval Axiom (A11)).

We have deliberately chosen to present just one of these measures, namely the fairly well-known index proposed by Clark et al. (1981). They had found that the Sen-index, or rather one of the variants that come under this heading, can be written as a combination of the poverty gap ratio and the Gini coefficient. They replaced the Gini coefficient with another inequality measure — the Atkinson measure.

Measure 5 (Clark, Hemming, Ulph 1)

$$\begin{aligned} C_{1,\alpha}(x;z) &= \frac{q}{n\,z}\left[\frac{1}{q}\sum_{i=1}^{q}(z-x_i)^{\alpha}\right]^{\frac{1}{\alpha}} \qquad \textit{with } \alpha \geq 1 \\ \rightsquigarrow \widetilde{C}_{1,\alpha}(F,z) &= \frac{F(z)}{z}\left[\int_0^z (z-y)^{\alpha}\frac{f(y)}{F(z)}\right]^{\frac{1}{\alpha}} \\ &= \frac{F(z)^{\frac{\alpha-1}{\alpha}}}{z}\sqrt[\alpha]{LPM_{\alpha}(F,z)}. \end{aligned}$$

Like the LPM_{α}, this downside risk measure allows for negative portfolio values, as long as z remains positive.

Although we do not want to transform them now, we still want to give brief descriptions of other indices within this class for a reason that will become apparent. The poverty measure of Blackorby and Donaldson (1980) is also derived from one of the Sen measures, noting that it can be interpreted as the product of the headcount ratio and the percentage shortfall of the representative income of the poor from the poverty line. The representative income of the poor, analog to the equally distributed equivalent income known in the inequality literature, and hence the index and its properties depend on the social welfare function. This makes a general transformation into a downside risk measure difficult, but also opens an interesting route for further research. Some of our readers might also want to see not only statistical indices, which look more or less appealing, but might want to know the normative significance of using specific measures. In other words, what is the utility function implicitly used when subscribing to one of these indices? Ethical poverty indices could in the future provide a sensible starting point here because from their outset they are each based on welfare concepts about which sometimes a lot is known at least in the poverty context.

Another future choice could be one of the many measures which were formulated by Chakravarty (1983b). He uses a specific social welfare function. Social welfare functions are also directly used in the work of Vaughan (1987). By contrast, the measure suggested by Kakwani (1980b) is not useful for our purpose. He generalized the Sen-type measures by dropping, among others, the focus axiom. His measure, which has counterintuitive properties in poverty measurement, too, therefore would certainly violate our basic axioms if transformed into a downside risk measure.

2.2.4 Class 3: Distribution-Sensitive, Subgroup Consistent Poverty Measures

This last class proposed by Zheng (1997) contains poverty measures which all are distribution-sensitive and fulfill the Axiom of Subgroup Consistency (cf. Axiom 24 in the appendix). This property requires that the value of the measure is bounded

by the two values resulting if the whole population, respectively the set of all events, is partitioned into two sets. Also, all of these measures are decomposable "or can be expressed as increasing transformations of some decomposable poverty measures" (Zheng (1997, p. 150)). Decomposability (cf. Axiom 26 in the appendix) means that in fact the value for the whole population, respectively for the set of all events, need not only be bounded by the two values for the subsets, but must even be a convex combination with weights equal to the relative sizes of the subsets. We will transform all measures of this class into downside risk measures.
When examining these measures with respect to the axioms (cf. section 3), this turns out to be relatively easy. It will be shown that any decomposable measure fulfills some axioms automatically if some simple conditions hold. Hence it can be shown that all decomposable measures fulfill a whole group of axioms. Only a few axioms remain which need to be checked individually.
The first distribution-sensitive (and decomposable) poverty measure was proposed by Watts (1968). An axiomatic characterization for the Watts measure has not been provided by Watts himself but by Zheng (1993).

Measure 6 (Watts)

$$\begin{aligned} W(x;z) &= \frac{1}{n}\sum_{i=1}^{q}(\ln z - \ln x_i) \\ \rightsquigarrow \widetilde{W}(F,z) &= \int_0^z (\ln z - \ln y)\ f(y)\,dy. \end{aligned}$$

As $\widetilde{W}$ has singularities in $y = 0$ and $z = 0$, i.e. neither the portfolio values y nor z may be negative, and $\widetilde{W}$ cannot be extended on the negative numbers.
Another poverty measure proposed by Clark et al. (1981) is subgroup consistent but not decomposable. The measure can be expressed as a monotone increasing function of a decomposable poverty measure, namely the Chakravarty index $\widetilde{Ch}_e$ presented right below (cf. Zheng (1997, p. 151)).

Measure 7 (Clark, Hemming, Ulph 2)

$$\begin{aligned} C_{2,\beta}(x;z) &= 1 - \frac{1}{z}\left[\frac{1}{n}\sum_{i=1}^{n}(\min\{x_i,z\})^{\beta}\right]^{\frac{1}{\beta}} \qquad \textit{for } \beta < 1 \\ \rightsquigarrow \widetilde{C}_{2,\beta}(F,z) &= 1 - \frac{1}{z}\left[\int_0^{\infty}(\min\{y,z\})^{\beta}\, f(y)\,dy\right]^{\frac{1}{\beta}} \\ &= 1 - \sqrt[\beta]{1 - \widetilde{Ch}_{\beta}(F,z)}. \end{aligned}$$

Thus $\widetilde{C}_{2,\beta}$ can be derived from $\widetilde{Ch}_e$, setting $e = \beta$. The measure by Chakravarty (1983a)[3] is a transformation of the one above, and vice versa. Chakravarty derived his decomposable poverty measure by following Sen's axiomatic approach, solving a functional equation arising from Sen's three basic axioms.

[3] In table 3.1. on p. 143 in Zheng (1997) the Chakravarty (1983a) poverty measure Ch_e is by mistake sorted into class 2, but in the text it is sorted into class 3.

Measure 8 (Chakravarty)

$$Ch_e(x; z) = \frac{1}{n}\sum_{i=1}^{q}\left[1 - \left(\frac{x_i}{z}\right)^e\right] \qquad \textit{for } 0 < e < 1$$
$$\rightsquigarrow \widetilde{Ch}_e(F, z) = \int_0^z \left[1 - \left(\frac{y}{z}\right)^e\right] f(y)\, dy.$$

Another class of poverty measures which includes only decomposable measures has been developed by Foster et al. (1984). The measures use a power of the income gap as a weighting.

Measure 9 (Foster, Greer, Thorbecke)

$$FGT_\alpha(x; z) = \frac{1}{n}\sum_{i=1}^{q}\left(1 - \frac{x_i}{z}\right)^\alpha, \quad \alpha \geq 0$$
$$\rightsquigarrow \widetilde{FGT}_\alpha(F, z) = \int_0^z \left(1 - \frac{y}{z}\right)^\alpha f(y)\, dy.$$

Because of multiplications or divisions by z, neither $\widetilde{C}_{2,\beta}$, $\widetilde{Ch}_e$ nor $\widetilde{FGT}_\alpha$ can be continued for negative critical lines z. Negative portofolio values are, however, possible. Hagenaars (1987) proposed poverty measures based on the assumption of utilitarian social welfare functions, which therefore bear some resemblance to the indices of the previous class. The exact properties of these measures depend on the characteristics of the individual utility functions U. In Hagenaars (1987) and in Zheng (1997, p. 152) some relationships between these utility function and the axioms can be found. Zheng (1997, pp. 152) mentioned also the relationship between the measures of Hagenaars, Dalton (therefore the symbol *HD*), and Chakravarty. The general form of these measures is

Measure 10 (Hagenaars 1)

$$HD_U(x; z) = 1 - \frac{1}{n\,U(z)}\sum_{i=1}^{n}\min\{U(x_i), U(z)\}$$
$$= \frac{1}{n}\sum_{i=1}^{q}\left(1 - \frac{U(x_i)}{U(z)}\right) \tag{10}$$
$$\rightsquigarrow \widetilde{HD}_U(F, z) = 1 - \frac{1}{U(z)}\int_0^\infty \min\{U(y), U(z)\}\, f(y)\, dy$$
$$\textit{with} \quad U'(y) > 0, U''(y) < 0\ \forall\, y \in \mathbb{R}_+. \tag{11}$$

Setting $U(y) := \ln y$ yields the Hagenaars Measure $\widetilde{Ha}$ as a special case, which concludes the selection of poverty and downside risk measures.

Measure 11 (Hagenaars 2)

$$Ha(x; z) = \frac{1}{n}\sum_{i=1}^{q}\left(1 - \frac{\ln x_i}{\ln z}\right)$$
$$\rightsquigarrow \widetilde{Ha}(F, z) = \int_0^z \left(1 - \frac{\ln y}{\ln z}\right) f(y)\, dy.$$

In the following two sections we will establish a connection between the downside risk measures introduced above and the axioms proposed by Breitmeyer et al. (1999). The names of these properties are sometimes self-explanatory, but for additional reference we have collected brief descriptions in the appendix.

3 Decomposable Downside Risk Measures

In this section we show that it is possible to derive some general results, stating which axioms are satisfied, for the decomposable downside risk measures.

Proposition 1 (Characterization of Decomposable Downside Risk Measures) *If ϕ is piecewise continuous and D is a functional of the form*

$$D(F, z) = \int_{-\infty}^{\infty} \phi(y, z)\, f(y)\, dy, \tag{12}$$

then D is a downside risk measure (and fulfills the four basic axioms) if and only if

$$\phi(y, z) = 0 \quad \forall y, z \in \mathbb{R}, y > z, \tag{13}$$

$$\phi(y, z) \geq 0 \quad \forall y, z \in \mathbb{R}, \tag{14}$$

$$\phi(y_1, z) \geq \phi(y_2, z) \quad \forall y_1, y_2, z \in \mathbb{R}, y_2 > y_1. \tag{15}$$

Note that the integral in (12) covers the complete real numbers. Of course, if ϕ is only defined on an interval $I \subseteq \mathbb{R}$, then only the defined part of $\phi(y, z)\, f(z)$ needs to be integrated, and the measure D is only meaningful for density functions f with supp $f \subseteq I$.

Proof: The proof can be lead independently for each of the four basic axioms.[4]

A1 Focus $\Longrightarrow$ $\phi(y, z) = \tilde{\phi}(z)$ for all $y > z$, because: Let $y_2 > y_1 > z$, $F_1 := \chi_{y_1}$ and $F_2 := \chi_{y_2}$, where χ is a characteristic function as defined in definition 3. Then F_2 is derived by a 1^{st} degree bonus from F_1, and $f_i = F_i'$ is a Dirac Distribution[5] with $\int \phi_(y, z)\, f_i(y)\, dy = \phi(y_i)$. This implies

$$\phi(y_1, z) = \int \phi(y, z)\, f_1(y)\, dy = D(F_1, z) = D(F_2, z)$$
$$= \int \phi(y, z)\, f_2(y)\, dy = \phi(y_2, z),$$

thus $\phi(y_1, z) = \phi(y_2, z) =: \tilde{\phi}(z)$. The proof of the other direction, that ϕ constant above z implies the Focus Axiom, is straightforward.

A2 It is therefore obvious that Focus and Normalization hold if and only if $\tilde{\phi}(z) \equiv 0$, i.e. if $\phi(y, z)$ vanishes for all y above z.

---

[4] The numbering of the statements is adjusted to the numbering of the axioms in the appendix and in Breitmeyer et al. (1999).

[5] For literature on the theory of distributions, cf. Rudin (1991, pp. 149). Distributions will be used if necessary in order to simplify proofs.

A3 Positivity implies $\phi(y,z) \geq 0$ for all $y, z \in \mathbb{R}$. Suppose $\phi(\tilde{y}, z) < 0$ for some $\tilde{y}$ and z, then choose $F := \chi_{\tilde{y}}$ with χ as defined in definition 3. Then $D(F,z) = \phi(\tilde{y}, z) < 0$ which is a contradiction to Positivity. The other direction of the implication is obvious.

A4 Weak Monotonicity implies $\phi(y_1, z) \geq \phi(y_2, z)$ for $y_2 > y_1$. To see this, choose F_1 and F_2 as in the proof of A1. Let us now prove Weak Monotonicity, assuming $\phi(y_1, z) \geq \phi(y_2, z)$ for all $y_2 > y_1$. Be F_2 a 1^{st} degree bonus of F_1, thus $F_1 - F_2 \geq 0$ everywhere. If ϕ is differentiable, then $\partial\phi(y,z)/\partial y \geq 0$, and integration by parts yields

$$\begin{aligned} D(F_1, z) &= \int_{-\infty}^{z} \phi(y,z)\, f_1(y)\, dy \\ &= \phi(y,z)\, F_1(y)\Big|_{-\infty}^{z} - \int_{-\infty}^{z} \frac{\partial\phi(y,z)}{\partial y} F_1(y)\, dy \\ &= \phi(z,z)\, F_1(z) - \int_{-\infty}^{z} \frac{\partial\phi(y,z)}{\partial y} F_1(y)\, dy \\ &\geq \phi(z,z)\, F_2(z) - \int_{-\infty}^{z} \frac{\partial\phi(y,z)}{\partial y} F_2(y)\, dy \\ &= \int_{-\infty}^{z} \phi(y,z)\, f_2(y)\, dy = D(F_2, z). \end{aligned}$$

If ϕ contains jumps, then $\partial\phi/\partial y$ is a distribution, but the proof is led the same way. ■

Note that decomposable downside risk measures are linear[6] in the first variable F (or f), thus some instruments of functional analysis can be applied.

Proposition 2 (Properties of Decomposable Downside Risk Measures) *Let D be a decomposable downside risk measure, ϕ as in (12). Then D fulfills*

A5 Continuity and A6 Lipschitz Continuity if and only if $|\partial\phi(y,z)/\partial y| \leq C(z)$ for some function $C(z)$.

A7 Critical Line Continuity if and only if $\phi(y,z)$ is continuous in the second variable and either $\phi(y,z)$ is a monotone function in z, or $\phi(y,z) \leq \tilde{\phi}(y)$, ϕ is dominated by some integrable function $\tilde{\phi}$.

A8 Scale Invariance if and only if $\phi(\lambda\, y, \lambda\, z) = \phi(y,z) \quad \forall \lambda > 0$, i.e. if ϕ is homogeneous of degree 0.

A9 Homogeneity if and only if $\phi(\lambda\, y, \lambda\, z) = \lambda\, \phi(y,z) \quad \forall \lambda > 0$, i.e. if ϕ is homogeneous of degree 1.

A10 Translation Invariance if and only if $\phi(y,z) = \tilde{\phi}(y - z)$.

[6] Actually, they can be embedded in a linear functional. They are not linear themselves, e.g. the condition $D(F_1) + D(F_2) = D(F_1 + F_2)$ does not make sense as D is defined on probability distributions only, and $F_1 + F_2$ is not a distribution if F_1 and F_2 are.

A11 Unit Intervall if and only if $0 \leq \phi(y,z) \leq 1$.

A12 Limitedness if and only if $\phi(y,z) \leq z - y$.

A13 Semi-Strong Monotonicity 1 if and only if $\phi(y,z)$ *is a strongly monotone decreasing function for* $y \in (-\infty, z)$.

A14 Semi-Strong Monotonicity 2 if and only if $\phi(y,z) > 0$ *for all* $y < z$.

A15 Strong Monotonicity if and only if $\phi(y,z)$ *is a strongly monotone decreasing function for* $y \in (-\infty, z]$.

A16 Monotonicity Sensitivity, A17 Additional Gamble and A19 Semi-Strong 2^{nd} *Degree Distribution Reagibility 2 if and only if* $\phi(y,z)$ *is a convex function for* $y \in (-\infty, z)$, *thus if* $\phi(y) < (\phi(y_2,z) - \phi(y_1,z))/(y_2 - y_1)$ *for* $y_1 < y < y_2 < z$, *or if* $\partial^2\phi(y,z)/\partial y^2 > 0$ *for* $y < z$. *These axioms are thus equivalent for decomposable Downside Risk Measures.*

A18 Semi-Strong 2^{nd} *Degree Distribution Reagibility 1 and A20 Strong* 2^{nd} *Degree Distribution Reagibility if and only if* $\phi(y,z)$ *is convex for* $y \in (-\infty, z]$. *These axioms are equivalent for decomposable Downside Risk Measures.*

A21 2^{nd} *Degree Distribution Sensitivity if and only if* $\partial\phi(y,z)/\partial y$ *is a convex function for* $y \in (-\infty, z)$.

A22 Semi-Strong Increasing Critical Line if and only if $\phi(y,z_2) \geq \phi(y,z_1)$ *for* $z_2 > z_1$ *and* $\phi(y,z) > 0$ *for all* $y < z$.

A23 Strong Increasing Critical Line if and only if $\phi(y,z_2) > \phi(y,z_1)$ *for* $z_2 > z_1$.

A24 Subgroup Consistency, A25 Mean and A27 Growth of Safety independently of any further preconditions.

A28 Growth of Risk if and only if $\phi(y,z) = const.\ \forall y < z$.

Proof: Because many of the proofs follow comparable patterns, some of them are skipped in order not to bore the reader.[7]

A5 and A6: Because of the linear form of D, it is continuous whenever it is bounded[8], i.e. if

$$|D(F_1,z) - D(F_2,z)| \leq C(z) \cdot \|F_1 - F_2\|_1 = C(z) \cdot \int |F_1(y) - F_2(y)|dy. \quad (16)$$

Obviously, Continuity and Lipschitz Continuity are identical for linear functions with $L = C$. For one direction of the proof, assume that $\partial\phi/\partial y$ is bounded

[7] Those proofs are available from the authors upon request.

[8] Cf. Rudin (1987, pp. 95).

($|\partial\phi(y,z)/\partial y| \le C(z)$). Integration by parts yields

$$
\begin{aligned}
&|D(F_1,z) - D(F_2,z)| \\
&= \left| \int_{-\infty}^{\infty} \phi(y,z)\ (f_1(y) - f_2(y))\ dy \right| \\
&= \left| \phi(y,z)\,(F_1(y) - F_2(y)) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{\partial\phi(y,z)}{\partial z}\ (F_1(y) - F_2(y))\ dy \right| \\
&\overset{(\dagger)}{=} \left| \int_{-\infty}^{\infty} \frac{\partial\phi(y,z)}{\partial z}\ (F_1(y) - F_2(y))\ dy \right| \\
&\overset{(*)}{\le} C(z) \left| \int F_1(y) - F_2(y)\, dy \right| \\
&\le C(z) \int |F_1(y) - F_2(y)|\ dy = C(z)\|F_1 - F_2\|_1,
\end{aligned}
$$

where ($\dagger$) holds because $\lim_{y\to-\infty}(F_1(y) - F_2(y)) = 0 - 0 = 0$ and $\lim_{y\to\infty}(F_1(y) - F_2(y)) = 1 - 1 = 0$, and ($*$) holds if $|\partial\phi(y,z)/\partial z|$ is bounded by some function $C(z)$. For the proof of the opposite direction, assume that $\partial\phi/\partial y$ is unbounded and construct a contradiction by choosing series y_n, h_n with $|h_n| \ge 1/n$, $h_n \to 0$ and

$$\left| \frac{\phi(y_n + h_n, z) - \phi(y_n, z)}{h_n} \right| \ge n.$$

Choosing $F_{1,n} = \chi_{y_n}$ and $F_{2,n} = \chi_{y_n + h_n}$, one receives $\lim_{n\to\infty} \|F_{1,n} - F_{2,n}\|_1 = \lim_{n\to\infty} h_n = 0$, but $\lim_{n\to\infty} D(F_{1,n}, z) - D(F_{2,n}, z) = \lim_{n\to\infty} \phi(y_n, z) - \phi(y_n + h_n, z) \ge \lim_{n\to\infty} n\,h_n \ge 1 \not\to 0$, which provides the desired contradiction. Whenever ϕ contains jumps, $\partial\phi/\partial y$ is an unbounded distribution, and D is not continuous.

A7 Assume Critical Line Continuity holds, and $\lim_{i\to\infty} |z_i - z| = 0$ and choose $F_i(y) = \chi_{z_i}(y)$, with χ as defined in definition 3. Then $\phi(y, z_i) = \lim_{i\to\infty} D(F, z_i) = D(F, z) = \phi(y, z)$. To prove the other direction, assume continuity in the second variable ($\lim_{i\to\infty} \phi(y, z_i) = \phi(y, z)$ for $z_i \to z$). Then[9]

$$
\begin{aligned}
|D(F, z_i) - D(F, z)| &= \left| \int_{-\infty}^{z_i} \phi(y, z_i) f(y) dy - \int_{-\infty}^{z} \phi(y, z) f(y) dy \right| \\
&= \left| \int_{-\infty}^{z_i} \phi(y, z_i) - \int_{-\infty}^{z_i} \phi(y, z) + \int_{-\infty}^{z_i} \phi(y, z) - \int_{-\infty}^{z} \phi(y, z) \right| \\
&\le \left| \int_{-\infty}^{z_i} \phi(y, z_i) - \int_{-\infty}^{z_i} \phi(y, z) \right| + \left| \int_{-\infty}^{z_i} \phi(y, z) - \int_{-\infty}^{z} \phi(y, z) \right| \\
&= \left| \int_{-\infty}^{z_i} \phi(y, z_i) - \phi(y, z) \right| + \left| \int_{z}^{z_i} \phi(y, z) \right| \\
&\to 0 \quad \text{for} \quad i \to \infty.
\end{aligned}
$$

[9] In the integrals, $f(y)\,dy$ is omitted for reasons of readability.

The first integral vanishes (if ϕ is continuous in z) either because of Lebesgue's Dominated Convergence Theorem[10] or because of Lebesgue's Monotone Convergence Theorem[11]. The second integral vanishes for any integrable ϕ.

A8, A9, A10 and A11: All of these can be proven similarly, thus we show only the proof of A10:

$$\begin{aligned} D(F_{Y+\tau}, z+\tau) &= D(F_Y, z) \quad \text{for all } \tau \text{ and } F \\ \int_{-\infty}^{z} \phi(y,z)\, f(y)\, dy &= \int_{-\infty}^{z+\tau} \phi(y, z+\tau)\, f(y-\tau)\, dy \\ &= \int_{-\infty}^{z} \phi(y+\tau, z+\tau)\, f(y)\, dy \quad \text{for all } \tau \text{ and } f \\ \overset{(\dagger)}{\Longleftrightarrow}\ \phi(y+\tau, z+\tau) &= \phi(y,z) \quad \text{for all } \tau \\ \Longleftrightarrow\ \phi(y-z, 0) &= \phi(y,z). \end{aligned}$$

Of (†), '$\Longleftarrow$' is obvious, and '$\Longrightarrow$' can again be proven with the aid of characteristic functions χ as defined in definition 3. Setting $\tilde{\phi}(y) = \phi(y, 0)$ completes the proof.

A12 '$\Longleftarrow$' is straightforward to prove. One can again use the characteristic function χ to prove '$\Longrightarrow$': Choose $f(y) := \chi_{y_{\min}}(y)$, then $\underline{Y}_F = y_{\min}$ and

$$D(F,z) = \int_{-\infty}^{z} \phi(y,z)\, \delta_{y_{\min}}(y)\, dy = \phi(y_{\min}, z)$$

which must be less or equal $z - \underline{y}_F$. This can only be achieved if $\phi(y,z) \le z - y$.

A13, A14 and A15: Can be proven similarly to the equivalence of Weak Monotonicity and (15). Note that in the case of decomposable downside risk measure, Semi-Strong Monotonicity 1 implies Semi-Strong Monotonicity 2.

A16 Let $\mathcal{F}_i$ be the primitive functions of F_i. If both F_1 and F_2 are 1^{st} degree mali of the same F, but the malus of F_1 lies below the one of F_2, then $\mathcal{F}_1 \ge \mathcal{F}_2$. Integrating by parts twice yields

$$\begin{aligned} D(F_1, z) &= \int_{-\infty}^{z} \phi(y,z)\, f_1(y)\, dy \\ &= \phi(z,z)\, F_1(z) - \int_{-\infty}^{z} \frac{\partial \phi(y,z)}{\partial y} F_1(y)\, dy \\ &= \underbrace{\phi(z,z)}_{=0} F_1(z) - \underbrace{\frac{\partial \phi}{\partial y}(z,z)}_{\le 0} \underbrace{\mathcal{F}_1(z)}_{>\mathcal{F}_2} + \int_{-\infty}^{z} \underbrace{\frac{\partial^2 \phi(y,z)}{\partial y^2}}_{?} \underbrace{\mathcal{F}_1(y)}_{\overset{(\dagger)}{>}\mathcal{F}_2} dy \\ &> D(F_2, z) \end{aligned}$$

[10] The theorem applies if the functions $\phi(y, z_i)$ are dominated by an integrable function $\tilde{\phi}$, cf. Rudin (1987, p. 26).

[11] The theorem applies for monotone $\phi(y, z_i)$, cf. Rudin (1987, p. 21).

if and only if $\partial^2\phi(y,z)/\partial y^2 > 0$ for $y < z$. Again, $\partial^2\phi/\partial y^2$ can be a distribution. $\phi(z,z) = 0$ because convex functions are continuous. (†) is supposed to mean that at $\mathcal{F}_1 > \mathcal{F}_2$ strictly at least in one point below z.

A17 Note that $F(y-\varepsilon)$ is a 1^{st} degree bonus of $F(y)$, and $F(y+\varepsilon)$ is a 1^{st} degree malus of $F(y)$, only that the malus is applied by ε above the bonus. An additional gamble is therefore a special case of a 2^{nd} degree malus, and the proof can be lead similarly to the proofs of Monotonicity Sensitivity (A16) or 2^{nd} degree Reagibility Axioms.

A18, A19, A20 and A21: are all to be proven similarly, thus only one proof (A20) will be carried out. Because of the Decomposability of D, results on Monotonicity Sensitivity can be transferred to 2^{nd} degree Reagibility.[12] Monotonicity Sensitivity yields $D(F+G_1,z) > D(F+G_2,z)$ if $G_1(y) \equiv G_2(y+c)$ for some $c > 0$. Because of the linearity of D,

$$\begin{aligned} D(F,z) + D(G_1,z) &> D(F,z) + D(G_2,z) \\ D(G_1,z) &> D(G_2,z). \end{aligned}$$

Thus if F_3 is a 2^{nd} degree malus of F_1 and therefore $F_3 \equiv F_1 - G_1 + G_2$,

$$D(F_3,z) = D(F_1,z) + \underbrace{D(G_1,z) - D(G_2,z)}_{>0} > D(F_1).$$

A22 For $z_2 > z_1$ and $F(z_2) > F(z_1)$

$$D(F,z_2) = \int_{-\infty}^{z_1} f(y)\,\phi(y,z_2)\,dy + \int_{z_1}^{z_2} f(y)\,\phi(y,z_2)\,dy \tag{17}$$

which is greater than $D(F,z_1)$ if and only if $\phi(y,z_2) \geq \phi(y,z_1)$ and $\phi(y,z) > 0$ for all z.

A23 Here, (17) holds if and only if $\phi(y,z_2) > \phi(y,z_1)$.

A24 Subgroup Consistency is easy to prove:

$$\begin{aligned} D(F,z) &= \lambda\, D(F_1,z) + (1-\lambda)\, D(F_2,z) \\ &< \lambda\, D(\tilde{F}_1,z) + (1-\lambda)\, D(\tilde{F}_2,z) = D(\tilde{F},z). \end{aligned}$$

A25 Mean and A27 Growth of Safety are just as straightforward to prove.

A28 '$\Leftarrow$' is straightforward to prove. For the proof of '$\Rightarrow$', assume that $D(F,z) = \lambda\, D(F_1,z) + (1-\lambda)\, D(F_2,z) \geq D(F_1,z)$ whenever $F_2(z) = 1$, independently of the shape of $F_1(z)$ below z. This implies $D(F_2,z) \geq D(F_1,z)$ whenever $F_2(z) = 1$. This is only possible if $D(F_2,z) = D(F_1,z)$ holds for $F_1(z) = 1$, thus $D(F,z) = const.$ whenever $F(z) = 1$. Therefore, $\phi(y,z) = const.$ ■

[12] For decomposable D, this statement applies also for higher degree sensitivity and reagibility axioms, because the integral always falls into pieces.

With this result, it is sufficient to examine not the decomposable downside risk measures $D(F, z)$ themselves, but only their characteristic functions $\phi_D(y, z)$. They are

$$\begin{aligned} \phi_{LPM_n}(y,z) &= (z-y)^n, & \phi_{HI}(y,z) &= \frac{z-y}{z}, \\ \phi_W(y,z) &= \ln z - \ln y, & \phi_{Ch_e}(y,z) &= 1-\left(\frac{y}{z}\right)^e, \\ \phi_{FGT_\alpha}(y,z) &= \left(1-\frac{y}{z}\right)^\alpha, & \phi_{HD_U}(y,z) &= 1-\frac{U(y)}{U(z)} \quad \text{and} \\ \phi_{Ha}(y,z) &= 1-\frac{\ln y}{\ln z}. \end{aligned}$$

In order to prove the axioms for these decomposable downside risk measures, it remains to check the properties of the corresponding ϕ-functions, e.g. whether their first derivatives are bounded (for A5), whether they are homogeneous (for A8 and A9), whether they are bounded (for A11 and A12) and so forth. The results of these checks are collected in table 1 on page 26.

4 Non-Decomposable Downside Risk Measures

In this section, we perform a similar exercise as in the previous one, but this time for the few non-decomposable downside risk measures. Also, in Breitmeyer et al. (1999), the axioms are already checked for the VaR, $LPM_0$and LPM_1. The results can directly be used for table 1.

Of the non-decomposable downside risk measures, some are closely related to decomposable ones. If this is the case, results sometimes are easily derived from the properties of the related decomposable downside risk measures. This is shown in section 4.2. Of the remaining downside risk measures, that are neither decomposable nor closely related to decomposable downside risk measures, the most prominent is the Sen-index. Its properties are not always easy to check. We illustrate this by analyzing some axioms for the Sen-index to start with.

4.1 The Sen-Index

In Eggers et al. (1999), it is shown that the Sen-Index, if transferred from poverty measurement to downside risk measurement, has problems because of the negative portfolio values that might appear. This problem can be solved partly if there exists a minimum portfolio value $y_{\min}$ (with possibly $y_{\min} < 0$) by shifting the distribution function F and the critical line z by $y_{\min}$. Let us now prove some properties of S, restricted to positive portfolio values. We do not prove all properties, but hope to have chosen a representative selection.

Proposition 3 (Properties of the Sen-Index) *The Sen-Index does*

A5 not fulfill Continuity,

A8 fulfill Scale Invariance,

A10 not fulfill Translation Invariance, *A11 fulfill Unit Interval,*

A13 fulfill Semi-Strong Monotonicity 1,

A15 fulfill Strong Monotonicity,

A16 fulfill Monotonicity Sensitivity,

A21 not fulfill 2^{nd} *Degree Distribution Sensitivity.*

The proof starts with A13, because one of the provisional results can be used for the proofs of other axioms.

A13 In (8), the Sen-index is written as

$$\begin{aligned}
\widetilde{S}(F,z) &= \frac{F(z)}{z}\left[\int_0^z \left((z-y)+\int_0^y (y-x)\,\frac{f(x)}{F(z)}\,dx\right)\frac{f(y)}{F(z)}\,dy\right], \quad \text{which is} \\
&= \frac{1}{z\,F(z)}\left[\int_0^z \left((z-y)\,F(z)+\int_0^y (y-x)\,f(x)\,dx\right)\,f(y)\,dy\right] \\
&\overset{(*)}{=} \frac{1}{z\,F(z)}\int_0^z (2F(z)-F(y))\;F(y)\,dy \qquad (18) \\
&\overset{(\dagger)}{\geq} \frac{1}{z\,F(z)}\int_0^z \left(2F(z)-\tilde{F}(y)\right)\,\tilde{F}(y)\,dy = \widetilde{S}(\tilde{F},z), \qquad (19)
\end{aligned}$$

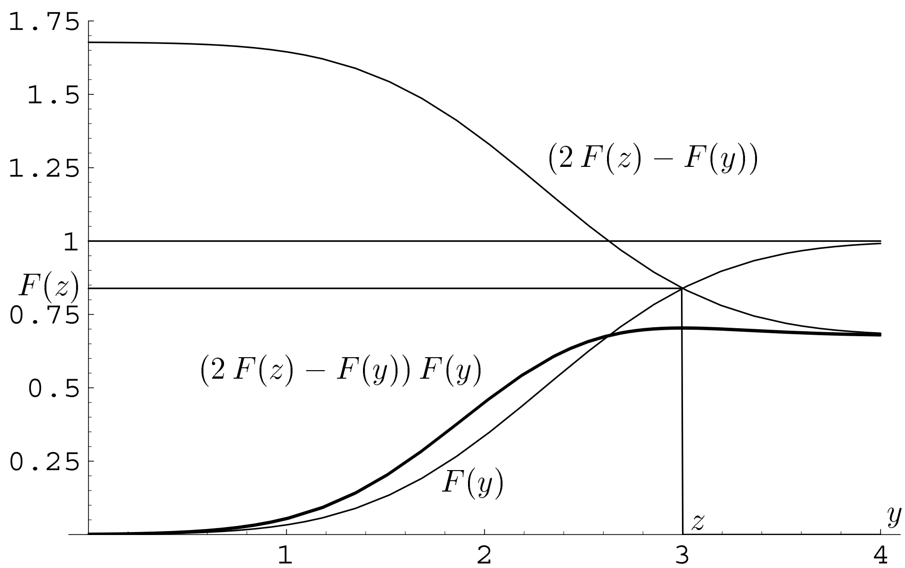
where $(*)$ holds because several integrations by parts yield

$$\begin{aligned}
&\int_0^z \left((z-y)\,F(z)+\int_0^y (y-x)\,f(x)\,dx\right)\,f(y)\,dy \\
&= \int_0^z (z-y)\,F(z)+\left[\underbrace{(y-x)\,F(x)\Big|_0^y}_{=0}-\int_0^y (-1)F(x)\,dx\right]\,f(y)\,dy \\
&= \int_0^z (z-y)\,F(z)+\int_0^y F(x)\,dx\,f(y)\,dy \\
&= \int_0^z \left[(z-y)\,F(z)+\mathcal{F}(y)\right]\,f(y)\,dy \\
&= \underbrace{\left[(z-y)\,F(z)+\mathcal{F}(y)\right]\,F(y)\Big|_0^z}_{\text{with } F(0)=\mathcal{F}(0)=0}-\int_0^z (F(z)-F(y))\,F(y)\,dy \\
&= \mathcal{F}(z)\,F(z)+\int_0^z (F(z)-F(y))\,F(y)\,dy \\
&= \int_0^z 2F(z)\,F(y)-F(y)^2\,dy = \int_0^z (2F(z)-F(y))\;F(y)\,dy,
\end{aligned}$$

where $\mathcal{F}$ is the primary function of F. $(\dagger)$ holds whenever $\tilde{F}$ is a 1^{st} degree bonus of F below z, thus $\tilde{F}<F$ somewhere below z and $\tilde{F}(z)=F(z)$, because the

integral in (19) is monotone increasing as a function[13] of F. Equation (18) can be used to calculate or illustrate the Sen-index. Note that also the customary form of the Sen-index (equation (6), especially the Gini-index, can be illustrated as a part of a surface (cf. Eggers et al. (1999)). However, as figure 1 illustrates the whole Sen-index instead of one of its components, it can help to elucidate properties of the Sen-index, e.g. A5, A11 and A13. In figure 1, the Sen-index can be read as the proportion between the surface under the graph $(2\,F(z) - F(y))\,F(y)$ (between 0 and z) and the surface of the rectangle $[0, F(z)] \times [0, z]$.

Figure 1: Graphic Determination of the Sen-Index



A5 In equation (9), all terms apart from $F(z) = LPM_0(F, z)$ are continuous, thus the discontinuity of LPM_0 is inherited by $\widetilde{S}$.

A8 In the writing of equation (18), we have

$$\begin{aligned}\widetilde{S}(F_{\lambda Y}, \lambda\, z) &= \frac{1}{\lambda\, z\, F(z)} \int_0^{\lambda\, z} (2F(z) - F(y/\lambda))\, F(y/\lambda)\, dz \\ &= \frac{1}{\lambda\, z\, F(z)} \int_0^{z} \lambda\, (2F(z) - F(y))\, F(y)\, dz = \widetilde{S}(F, z).\end{aligned}$$

Because A8 stands in contradiction to A9 and A12, the proof implies already that $\widetilde{S}$ cannot fulfill A9 and A12.

[13] To see this, note that

$$\frac{\partial (2F(z) - F(y))F(y)}{\partial F(y)} = 2F(z) - 2F(y) > 0$$

if $F(z) > F(y)$, which is the case because F is an increasing function.

A10 Nearly any arbitrary example can be taken as a counterexample.

A11 $(2\,F(z) - F(y))\,F(y) \leq F(z)$ for $0 \leq y \leq z$ implies $\int_0^z (2\,F(z) - F(y))\,F(y)\,dy \leq z\,F(z)$, thus one can see that $\widetilde{S}(F, z) \leq 1$ in the spelling of equation (18).

A15 If $\tilde{F}$ is a 1^{st} degree bonus of F, then the bonus can be split into two parts: With

$$\hat{F}(y) := \begin{cases} \tilde{F}(y) : y < z \\ F(y) : \text{ else} \end{cases},$$

$\hat{F}$ is a 1^{st} degree bonus of F below z, and $\tilde{F}$ is a 1^{st} degree bonus of F, where probability density is only 'pushed' over the critical line z, i.e. $\hat{F} = \tilde{F}$ below z. Then $\widetilde{S}(\tilde{F}, z) \leq \widetilde{S}(\hat{F}, z) \leq \widetilde{S}(F, z)$, where at least one of the inequalities is strict. The first inequality holds because of A13, and the second because the Sen-index in equation (19) is monotone increasing with $F(z)$.

A16 Let $\tilde{F}(y) = F(y) + G(y + c)$ with $G(y + c) > 0$. Let us assume that the support of $G(y + c)$ always lies below the critical line z, thus $\tilde{F}(z) = F(z)$. Then

$$\begin{aligned} &z\,\tilde{F}(z)\,\widetilde{S}(\tilde{F}, z) \\ &= \int_0^z (2\,F(z) - \tilde{F}(y))\,\tilde{F}(y)\,dy \\ &= \int_0^z (2\,F(z) - F(y) - G(y + c))\,(F(y) + G(y + c))\,dy \\ &= \underbrace{z\,F(z)\,\widetilde{S}(F, z)}_{(\dagger)} + 2\underbrace{\int_0^z (F(z) - F(y))\,G(y + c))\,dy}_{(*)} - \underbrace{\int_0^z G(y + c)^2\,dy}_{(\ddagger)}. \end{aligned}$$

Now $(\dagger)$ and $(\ddagger)$ are independent of c. An increase of c shifts the malus G downwards. Because $(F(z) - F(y))$ is a decreasing function, this leads to an increase of $(*)$. This increase is strict, because $F(y)$ must increase strictly in the relevant part, otherwise the malus G would not be applicable to F.

A21 For a counterexample, choose $z = 5$, and let $P(y = 1) = P(y = 2) = P(y = 3) = P(y = 4) = 0.25$. Now there are different ways to calculate the Sen-index. A discrete calculation (like in equation (6)) yields $F(z) = 1$, $\widetilde{I}(F, z) = 0.5$, and $\widetilde{G}(F, z) = 0.25$, thus $\widetilde{S}(F, z) = 0.625$. The continuous spelling (like in equation (18)) yields the same result, another indication for the validity of equation (18): $\int_0^5 (2 - F(y))\,F(y)\,dy = \int_0^1 0 \cdot 2 + \int_1^2 0.25 \cdot 1.75 + \int_2^3 0.5 \cdot 1.5 + \int_3^4 0.75 \cdot 1.25 + \int_4^5 1 \cdot 1 = 3.125$, thus again $\widetilde{S}(F, z) = 0.625$. Now taking a 1^{st} degree malus of the distribution at different places below z and calculating the Sen-index, e.g. choosing once $P(y = 0.5) = P(y = 2.5) = P(y = 3) = P(y = 4) = 0.25 \implies \tilde{S}(\tilde{F}, z) = 0.6375$ and then $P(y = 1) = P(y = 2) = P(y = 2.5) = P(y = 4.5) = 0.25 \implies \tilde{S}(\tilde{F}, z) = 0.6375$ yields the same Sen-index, which is a contradiction to A16. Similar counterexamples can be chosen to prove that the Sen-index does not fulfill A17 – A21. ■

4.2 Other Indices

Many downside risk measures, and all of the measures defined in this paper but not discussed yet, are based on decomposable downside risk measures. For this reason, the properties of those measures depend strongly on the properties of the measures they are derived from.
The measure of Clark, Hemming and Ulph (1981) is an example for a measure that is subgroup consistent, but not decomposable. It has the form

$$\begin{aligned} C_{2,\beta}(F,z) &= 1 - \frac{1}{z}\left[\int_0^z (\min\{y,z\})^\beta \, f(y)\, dy\right]^{\frac{1}{\beta}}, \\ &= \Phi\left(z, \int_{-\infty}^z \phi(y,z)\, f(y)\, dy\right) \end{aligned} \tag{20}$$

with $\phi(y,z) = 1 - (y/z)^\beta$ and $\Phi(z,x) = 1 - \sqrt[\beta]{1-x}/z$. According to Foster and Shorrocks (1991, p. 696), subgroup consistent measures can typically be written as in equation (20) with suitable ϕ and Φ.
If a downside risk measure D has the form (20), then many of its properties can easily be derived from the properties of the accompanying decomposable measure ϕ, which have been discussed in section 3. As an example, all Distribution Reagibility and Distribution Sensitivity Axioms hold for the subgroup consistent measure D if and only if they hold for the accompanying decomposable measure (with kernel ϕ). If the decomposable measure (with kernel ϕ) and Φ are continuous in the sense of A5, then so is D. The same applies for A6 Lipschitz Continuity. Even Invariance and Gauge Axioms can be transferred: As $\phi(y,z) = 1 - (y/z)^\beta$ lies in $[0,1]$ for $0 \le y \le z$ and $0 < \beta < 1$, and Φ maps $[0,1]$ on $[0,1]$, $C_{2,\beta}$ fulfills A11 Unit Interval.
One more downside risk measure that behaves completely different from any decomposable measure is the *VaR*. The reason is that the definition of the *VaR* contains inverse functions. The *VaR* is not discussed here, because its properties have already been checked in Breitmeyer et al. (1999).
Hopefully, it has become obvious that most of the properties of the common downside risk measures as well as of measures transferred from poverty measurement are easy to analyze. The examinations becomes less straightforward as the definition of measures becomes more complex, such as the Sen-index.

5 Conclusion

After a motivation why downside risk measures are useful, in particular for the application in banks, we have shown how the close relation between downside risk and poverty can be exploited: Many poverty measures, although not all (e.g. measures which are not replication invariant), can be transformed into downside risk measures via the straightforward mechanism (definition 3). This supplies us with downside risk measures that have not yet been used for this purpose. Unfortunately some of these new downside risk measures can only be applied for positive values, i.e. for distributions of portfolio values but not for profit and loss distributions.

With an increased set of downside risk measures available, the question which one(s) is/are reasonable (or even best) gains importance. To answer it, we have checked a number of indices with respect to an extensive list of requirements ("axioms"). The results, including findings of other papers and some obvious properties, are collected in table 1 on page 26. One important observation emerging from it is the fact that the widely used Value at Risk performs quite poorly in comparison to other indices. In particular the decomposable measures, for which some general results were derived, do much better. Their violation of invariance properties and range restrictions (cf. axioms A8 through A12) is not an extremely serious shortcoming because these properties are somewhat debatable in the downside risk context.

From a theoretical point of view, the transformation and thorough examination of further indices might be desirable. However, we see two different tasks ahead that seem to be even more important:

Firstly, for a number of specific applications, i.e. given data sets, risk preferences etc., the results of table 1 should be used to select "the best" downside risk measure(s). Portfolio managers, e.g., might make a choice different from that of regulators.

Secondly, but probably not finally, we intend to test empirically whether or not the violations of axioms (and hence the suggested inferiority of an index) indeed shows up for real world data. Also we want to find out how different the risk rankings are which are generated by various indices.

Apparently, much more work is required. At the very least we hope to have convinced the reader that it is worth being done.

Table 1: The Properties of Selected Downside Risk Measures

	Axiom ↓ / DRM →	VaR	LPM_0	LPM_1	LPM_2	$\widetilde{I}$	$\widetilde{HI}$	$\widetilde{S}$	$\widetilde{C}_{1,\alpha}$	$\widetilde{Ch}_e$	$\widetilde{FGT}_\alpha$	$\widetilde{W}$	$\widetilde{C}_{2,\beta}$	$\widetilde{Ha}$	$\widetilde{HD}_U$
(A5)	Continuity	–	–	✓	–	–	✓	–	–	–	–	–	–	–	–
(A6)	Lipschitz Continuity	–	–	✓	✓	–	✓	–	–	–	–	–	–	–	–
(A7)	Critical Line Continuity	✓	–	✓	✓	–	✓	–	✓	✓	✓	✓	✓	✓	✓
(A8)	Scale Invariance	–	✓	–	–	✓	✓	✓	–	–	–	–	–	–	–
(A9)	Homogeneity	✓	–	✓	–	–	–	–	–	–	–	–	–	–	–
(A10)	Translation Invariance	✓	✓	✓	✓	–	–	–	✓	–	–	–	–	–	–
(A11)	Unit Intervall	–	✓	–	–	✓	✓	✓	–	–	–	–	–	–	–
(A12)	Limitedness	✓	–	✓	–	–	–	–	–	–	–	–	–	–	–
(A13)	Semi-Strong Monotonicity 1	–	–	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
(A14)	Semi-Strong Monotonicity 2	–	✓	✓	✓	–	✓	✓	✓	✓	✓	✓	✓	✓	✓
(A15)	Strong Monotonicity	–	–	✓	✓	–	✓	✓	✓	✓	✓	✓	✓	✓	✓
(A16)	Monotonicity S.	–	–	✓	✓	–	✓	✓	✓	✓	✓	✓	✓	✓	✓
(A17)	Additional Gamble	–	–	–	✓	–	–	✓	✓	✓	✓	✓	✓	✓	✓
(A18)	Semi-Strong 2^{nd} Deg. D.R. 1	–	–	–	✓	–	–	✓	✓	✓	✓	✓	✓	✓	✓
(A19)	Semi-Strong 2^{nd} Deg. D.R. 2	–	–	–	✓	–	–	✓	✓	✓	✓	✓	✓	✓	✓
(A20)	Strong 2^{nd} Deg. D.R.	–	–	–	✓	–	–	✓	✓	✓	✓	✓	✓	✓	✓
(A21)	2^{nd} Deg. D.S.	–	–	–	✓	–	–	–	✓	✓	✓	✓	✓	✓	✓
(A22)	Semi-Strong Inc. C.L.	–	✓	✓	✓	–	–	–	✓	✓	✓	✓	✓	✓	✓
(A23)	Strong Inc. C.L.	–	–	✓	✓	–	–	–	✓	✓	✓	✓	✓	✓	✓
(A24)	Subgroup Consistency	–	✓	✓	✓	✓	✓	–	✓	✓	✓	✓	✓	✓	✓
(A25)	Mean	✓	✓	✓	✓	✓	✓	–	✓	✓	✓	✓	✓	✓	✓
(A26)	Decomposability	–	✓	✓	✓	–	✓	–	✓	✓	✓	✓	–	✓	✓
(A27)	Growth of Safety	✓	✓	✓	✓	–	✓	✓	✓	–	✓	✓	✓	✓	✓
(A28)	Growth of Risk	–	✓	–	–	–	–	–	–	–	–	–	–	–	–

Note: '✓' indicates that a property is met, '–' that it is not met.

A Preliminary Definitions and Notions

This appendix provides a condensed version of definitions in Breitmeyer et al. (1999). Due to limitation of space, we have to leave out many of the comments on the appropriateness and economic intuition of definitions and axioms.
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 measure on Ω. The corresponding probability distribution function is defined by $F(y) := P((-\infty, y]) = P(\{Y \leq y\})$. If a probability density function exists, let it be denoted by f with $F(y) = \int_{-\infty}^{y} f(t)\,dt$.

A.1 Concepts

In this section of the appendix, we formulate concepts for the modification of distribution functions. The concepts are related to concepts used in the discussion of poverty measurement.[14]
A relative change is the multiplication of the outcomes of all events as well as the critical line z with the positive number λ.

Definition 4 (Relative Change) *A relative change is the transition* $(Y, z) \to (\lambda Y, \lambda z)$ *for* $\lambda \in \mathbb{R}_{++}$.[15] *A relative change can also be characterized by the transition* $(F_Y, z) \to (F_{\lambda Y}, \lambda z)$*, where the distribution function of* λY *is defined by* $F_{\lambda Y}(y) = F_Y(y/\lambda)$.

An equal absolute change is defined by the addition of the same number τ to the outcomes of all events as well as the critical line z. τ is allowed to take on negative values.

Definition 5 (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* $F_{Y+\tau}(y) = F_Y(y - \tau)$.

A 1^{st} degree bonus 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 implies first order stochastic dominance (cf. Bawa (1982) oder Bawa (1975) oder Hanoch and Levy (1969) oder Levy (1992)).

Definition 6 (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* $\exists\, \underline{y}, \overline{y} \in \mathbb{R} : \underline{y} < \overline{y}$ *and* $F_1(y) - F_2(y) \begin{cases} = 0 : & -\infty < y < \underline{y} \\ > 0 : & \underline{y} < y < \overline{y} \\ = 0 : & \overline{y} \leq y < \infty \end{cases}$.

[14] Cf. Zheng (1997, p. 128).
[15] $\mathbb{R}_{++} := \{y \in \mathbb{R} : y > 0\}$ and $\mathbb{R}_{+} := \mathbb{R}_{++} \cup \{0\}$.

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. 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. The second moment, i.e. the variance, of the corresponding random variable will decrease by a 2^{nd} degree bonus, which implies second order stochastic dominance (cf. Hanoch and Levy (1969)).

Definition 7 (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, $G_2 := F_3 - F_2$,*
- $\exists\, c \in \mathbb{R}_+ \; : G_1(y) = -G_2(y + c) \;\forall\, y \in \mathbb{R}$,
- $\{\, y \in \mathbb{R} : G_1(y) \neq 0\} \cap \{\, y \in \mathbb{R} : G_2(y) \neq 0\} = \emptyset$.

The opposite of the 2^{nd} degree bonus is the 2^{nd} degree malus. 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 of the same size and type.

A.2 Axioms

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.

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.

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 $|D(F_1, z) - D(F_2, z)| \leq L\,\|F_1 - F_2\|_1$ holds.*

Note that Axiom 6 implies Axiom 5.
It is also reasonable to require 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.*

The Scale Invariance Axiom can be motivated as follows: If scale invariance is fulfilled, then the downside risk measure is unaffected by the unit or the currency in which portfolio values are measured. Notice that this axiom is problematic for $z = 0$ because then, e.g., the "doubling of a game", i.e. $\lambda = 2$, would not increase downside risk.

Axiom 8 (Scale Invariance) $\forall\, \lambda \in \mathbb{R}_{++} : D(F_{\lambda Y}, \lambda z) = D(F_Y, z)$.

The Homogeneity Axiom demands the downside risk measure to react linearly on stretching and crushing of the distribution function. In this case the above example would imply a doubling of the downside risk, too.

Axiom 9 (Homogeneity) $\forall\, \lambda \in \mathbb{R}_{++} : \frac{D(F_{\lambda Y}, \lambda z)}{\lambda} = D(F_Y, z)$.

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. This property is formulated in the Translation Invariance Axiom.

Axiom 10 (Translation Invariance) $\forall\, \tau \in \mathbb{R} : D(F_{Y+\tau}, z + \tau) = D(F_Y, z)$.

If the downside risk measure is a relative measure, i.e. it satisfies scale invariance, the image can be standardized to the unit interval.

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

If a downside risk measure has the dimension of a currency, e.g. when it satisfies Homogeneity, one can require the Limitedness Axiom: the value of the downside risk measure should not be greater than the maximal potential loss.

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

Weak Monotonicity is already required as a basic axiom. Additionally, a downside risk measure should indicate higher risk if payments in critical events are diminished.

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

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

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

The Strong Monotonicity implies both of the Semi-Strong Monotonicity Axioms.
The downside risk measure should react the stronger on a payment reduction in an event the lower the original payment in the initial event is.

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

The value of a downside risk measure of a distribution function with an additional gamble should be higher than that of the downside risk measure of the original distribution. This is related to the Rothschild and Stiglitz (1970) notion of increasing risk by adding noise.

Axiom 17 (Additional Gamble) *$D(\frac{F_{Y-\varepsilon}+F_{Y+\varepsilon}}{2}, z) > D(F, z)$ for all $\varepsilon > 0$ if $F(z) > 0$.*

We get two new 2^{nd} degree distribution sensitivity axioms.

Axiom 18 (Semi-Strong 2^{nd} Degree Distribution Reagibility 1) *With notation as in Definition 7, if F_3 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 $D(F_3, z) > D(F_1, z)$.*

Axiom 19 (Semi-Strong 2^{nd} Degree Distribution Reagibility 2) *With notation as in Definition 7, 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 $D(F_3, z) > D(F_1, z)$.*

The following axiom implies both of the Semi-Strong 2^{nd} Degree Distribution Reagibility axioms.

Axiom 20 (Strong 2^{nd} Degree Distribution Reagibility) *With notation as in Definition 7, 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 $D(F_1, z) > D(F_3, z)$.*

The downside risk measure should react the stronger, the lower the portfolio values are where the 2^{nd} degree malus happens.

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

Downside risk measures must 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).

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 $D(F, z_1) < D(F, z_2)$.*

Suppose the set of possible events is split into two subgroups: one subgroup on which an event E_1 occurs, and its complement E_2. 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 E_1 in which the risk is reduced $(F_1 \to \tilde{F}_1)$, accordingly $1 - \lambda$ denotes the probability of the complement, where downside risk has not changed.
Suppose now that again all events are partitioned in two sets. Then we require that the value of the aggregate downside risk measure lies between the values for the two sets.

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

The following decomposability 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. The decomposability 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. This property is very helpful when proving other axioms.

Axiom 26 (Decomposability) *Using the above notation, $D(F, z) = \lambda D(F_1, z) + (1 - \lambda) D(F_2, z)$.*

The following two axioms describe how a downside risk measure should react if an elementary event has the probability 0 or 1.

Axiom 27 (Growth of Safety) *With notation as above, if $F_2(z) = 0$ then $D(F, z) \leq D(F_1, z)$.*

Axiom 28 (Growth of Risk) *With notation as above, if $F_2(z) = 1$ then $D(F, z) \geq D(F_1, z)$.*

References

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber and David Heath (1999), Coherent Measures of Risk, *Mathematical Finance*, vol. 9, pp. 203–228.

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

Basle Committee on Banking Supervision (1996), Amendment to the Capital Accord to Incorporate Market Risks, Technical report, Basle, January.

Bawa, Vijay S. (1975), Optimal Rules for Ordering Uncertain Prospects, *Journal of Financial Economics*, pp. 95–121.

Bawa, Vijay S. (1982), Stochastic Dominance: A Research Bibliography, *Management Science*, vol. 28, pp. 698–712.

Beder, Tanya S. (1995), VAR: Seductive but Dangerous, *Financial Analysts Journal*, pp. 12–24.

Blackorby, Charles and David Donaldson (1980), Ethical Indices for the Measurement of Poverty, *Econometrica*, vol. 48, pp. 1053–1060.

Brachinger, Hans W. and Martin Weber (1997), Risk as a Primitive: A Survey of Measures of Perceived Risk, *OR Spektrum*, vol. 19, pp. 235–250.

Breitmeyer, Carsten, Hendrik Hakenes, Andreas Pfingsten and Christoph Rechtien (1999), Learning From Poverty Measurement: An Axiomatic Approach to Measure Downside Risk, Technical Report 99-03, Institut für Kreditwesen, Westfälische Wilhelms-Universität Münster.

Chakravarty, Satya R. (1983a), A New Index of Poverty, *Mathematical Social Sciences*, vol. 6, pp. 307–313.

Chakravarty, Satya R. (1983b), Ethically Flexible Measures of Poverty, *Canadian Journal of Economics*, vol. 16, pp. 74–85.

Clark, Stephen, Richard Hemming and David Ulph (1981), On Indices for the Measurement of Poverty, *Economic Journal*, vol. 91, pp. 515–526.

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

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

Elton, Edwin J. and Martin J. Gruber (1995), Modern Portfolio Theory and Investment Analysis, John Wiley & Sons, New York.

Foster, James, Joel Greer and Erik Thorbecke (1984), Notes and Comments: A Class of Decomposable Poverty Measures, *Econometrica*, vol. 52, No. 3, pp. 761–767.

Foster, James E. and Anthony F. Shorrocks (1991), Subgroup Consistent Poverty Indices, *Econometrica*, vol. 59, pp. 687–709.

Hagenaars, Aldi (1987), A Class of Poverty Indices, *International Economic Review*, vol. 28, pp. 583–607.

Hanoch, Giora and Haim Levy (1969), The Efficiency Analysis of Choices Involving Risk, *Review of Economic Studies*, vol. 36, pp. 335–346.

Hendricks, Darryll (1996), Evaluation of Value-at-Risk Models Using Historical Data, *Economic Policy Review (Federal Reserve Bank of New York)*, pp. 39–69.

Jorion, Philippe (1997), Value at Risk: The New Benchmark for Controlling Market Risk, 2nd ed., McGraw Hill, New York et. al.

Kakwani, Nanak C. (1980a), On a Class of Poverty Measures, *Econometrica*, vol. 48, pp. 437–446.

Kakwani, Nanak C. (1980b), Income Inequality and Poverty: Methods of Estimation and Policy Applications, Oxford University Press, New York.

Levy, Haim (1992), Stochastic Dominance and Expected Utility: Survey and Analysis, *Management Science*, vol. 38, pp. 555–593.

Linsmeier, Thomas J. and Neil D. Pearson (1996), Risk Measurement: An Introduction to Value at Risk, Technical Report July, Department of Accountancy and Department of Finance, University of Illinois, Urbana-Champaign.

Lintner, John (1965), The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, *The Review of Economics and Statistics*, vol. 47, pp. 13–37.

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

Mossin, Jan (1966), Equilibrium in a Capital Asset Market, *Econometrica*, vol. 34, pp. 768–783.

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.

Roy, Andrew D. (1952), Safety First and the Holding of Assets, *Econometrica*, vol. 20, pp. 431–449.

Rudin, Walter (1987), Real and Complex Analysis, Mc Graw-Hill, New York.

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

Sen, Amartya (1973), On Economic Inequality, Clarendon Press, Oxford.

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

Sharpe, William F. (1964), Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *Journal of Finance*, vol. 19, pp. 425–442.

Sortino, F. and H. Forsey (1994), Performance Measurement in a Downside Risk Framework, *Journal of Investing*, vol. 3, pp. 59–64.

Sortino, F. and H. Forsey (1996), On the use and Misuse of Downside Risk, *Journal of Portfolio Management*, vol. 2, pp. 35–42.

Takayama, Noriyuki (1979), Poverty, Income Inequality, and Their Measures: Professor Sen's Axiomatic Approach Reconsidered, *Econometrica*, vol. 47, pp. 747–759.

Thon, Dominique (1979), On Measuring Poverty, *Review of Income and Wealth*, vol. 25, pp. 429–439.

VAUGHAN, Richard N. (1987), Welfare Approaches to the Measurement of Poverty, *Economic Journal*, vol. 97, pp. 160–170.

WATTS, Harold W. (1968), An Economic Definition of Poverty, in: Daniel P. MOYNIHAN (ed.), *On Understanding Poverty*, Basic Books, Inc., New York and London, pp. 316–329.

ZHENG, Buhong (1993), An Axiomatic Characterization of the Watts Poverty Measure, *Economics Letters*, vol. 42, pp. 81–86.

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