The Geneva Papers on Risk and Insurance Theory, 23: 89–117 (1998)

An Extended Family of Financial-Risk Measures

CHRISTIAN S. PEDERSEN AND STEPHEN E. SATCHELL
Trinity College, Cambridge University, Cambridge, U.K. CB2 ITQ

Abstract

Recalling the class of risk measures introduced by Stone [1973], the authors survey measures from different academic disciplines—including psychology, operations research, management science, economics, and finance—that have been introduced since 1973. We introduce a general class of risk measures that extends Stone's class to include these new measures. Finally, we give four axioms that describe necessary attributes of a good financial risk measure and show which of the measures surveyed satisfy these. We demonstrate that all measures that satisfy our axioms, as well as those that do not but are commonly used in finance, belong to our new generalized class.

Key words: inequality, risk, utility

1. Introduction

In 1973, Bernell K. Stone presented his class of risk measures, which included most of the measures used in both empirical and theoretical finance. Since then, there has been considerable development in risk measurement not only in economics and finance but also in other fields, such as psychology, experimental psychology, operational research, and management science. Additionally, related measures of income inequality have appeared in the economics literature. The purpose of this article is to extend Stone's class to take into account developments over the last twenty five years.

We survey inequality measures presented in economic literature and outline the axiomatic approaches of Fishburn [1980, 1981], Luce [1980], Sarin [1987], and Pollatsek and Tversky [1970] from the psychology literature. From the more recent work in finance, we draw on such measures as the generalized Gini in Yitzhak [1983], lower partial moment, and probability of making a loss (which forms the basis for value at risk [1993]), both presented in Bawa and Lindenberg [1977], as well as some related statistical measures of dispersion. Utility-based risk measures from the operational research and management science literature are also reviewed.

Adapting the axioms of Kijima and Ohnishi [1993], we give four desirable properties of a good financial-risk measure. We show that almost all measures satisfying these belong to a new extended family. In addition, this family also contains measures commonly used in theoretical finance or by practitioners today that fail one or more of these properties. Further possible generalizations are discussed, as are possible relaxations and tightenings of the axioms.

Much risk theory is concerned with notions of stochastic dominance (e.g., Rothschild and Stiglitz [1970] and Bawa et al. [1985]). This imposes a partial ordering on risky prospects but will typically leave many investment unranked. By contrast, a risk measure will allow comparison between all prospects. The relationship between our proposed risk measures and stochastic dominance is addressed in Section 3.

The article is set out as follows. In Section 2 we introduce the Stone family of risk measures and motivate our plan to generalize further by our survey. In Section 3, we present our new extended family of risk measures, listing its special cases, and introduce the desired axioms we wish financial-risk measures to satisfy. We examine all measures against these axioms. Proofs are relegated to an Appendix. Section 4 is reserved for our conclusions.

2. Survey of risk measures

We proceed with our survey. In order to keep a consistent notation throughout the article, random returns are denoted by $\tilde{y}$, its density by $f(y)$, and cumulative distribution function by $F(y)$. The mean of $\tilde{y}$ is μ_y. In some contexts, however, $\tilde{y}$ could be second-period wealth or the price. We first recall Stone's original class of risk measures, which motivated our interest in this topic.

2.1. Stone's family of risk measures

In the financial literature, the most extensive work on general measures of risk was based on Stone [1973]. The family of risk measures defined by him was

$$R_S[Y_0, k, A](f) = \left(\int_{-\infty}^{A} |y - Y_0|^k f(y)\, dy \right)^{1/k}, \tag{1}$$

where Y_0, k, and A are all real numbers and $k > 0$. Note that this includes several commonly used measures of risk and dispersion, including the standard deviation

$$R_{SD}(f) = R_S[\mu_y, 2, \infty](f) = \left(\int (y - \mu_y)^2 f(y)\, dy \right)^{1/2} \tag{2}$$

as well as the semistandard deviation

$$R_{SSD}(f) = R_S[\mu_y, 2, 0](f) = \left(\int_{-\infty}^{0} (y - \mu_y)^2 f(y)\, dy \right)^{1/2} \tag{3}$$

and the mean absolute deviation

$$R_{MAD}(f) = R_S[\mu_y, 1, \infty](f) = \int |y - \mu_y| f(y)\, dy. \tag{4}$$

Stone's measure is thus quite broad and, as is shown, invariably satisfies our basic properties (to be defined later). Note further that the variance is

$$R_{VM}(f) = R_S[\mu_y, 2, \infty]^2(f) = \int (y - \mu_y)^2 f(y)\, dy, \tag{5}$$

while the semivariance measure is

$$R_{SVM}(f) = R_S[\mu_y, 2, 0]^2(f) = \int_{-\infty}^{0} (y - \mu_y)^2 f(y)\, dy. \tag{6}$$

Several of these subcases have later been converted into equilibrium risk measures in asset pricing, including the semivariance and the lower partial moments, which were applied to derive alternative CAPM-models by Bawa and Lindenberg [1977] as well as the original Sharpe model (see [1964]), which is based on the variance.

Stone's class is also closely related to the popular $\alpha - t$ measures of Fishburn [1977], which have been used in applied economic decision making under uncertainty, especially in the financial references. The $\alpha - t$ measures are defined as

$$R_{\alpha-t}(f) = R_S[t, \alpha, t]^\alpha(f) = \int_{-\infty}^{t} (t - y)^\alpha f(y)\, dy \tag{7}$$

and thus include both the variance and semivariance ((5) and (6), respectively) as special cases.

2.2. *The inequality measures*

There exists a parallel literature on inequality measurement in economics. Measuring economic inequality essentially reduces to the ranking of alternative income distributions according to some measure of dispersion. Consequently, as risk is commonly assumed to increase with dispersion (see, for example, Rothschild and Stiglitz [1970]), the inequality measures are prime candidates for potential adaptation to risk measurement.

The income inequality literature offers many different types of measures, arising from various axiomatic approaches. Most seem to be intuitive measures of dispersion, and some are adapted from the fields of physics and engineering. Although these are separately introduced, they can all be found in the books by Cowell [1995] and Sen [1973].

The first work by Dalton [1920] was based on a utilitarian framework. Comparing actual levels of average utility to the level of utility arising under perfect equality, he suggests the following functional form:

$$R_D(f) = 1 - \frac{\int y^{1-e} f(y)\, dy - 1}{\mu_y^{1-e} - 1}, \tag{8}$$

where e is the associated nonnegative measure of inequality aversion, defined on page 37 of Cowell [1995]. This was adapted by Atkinson [1970], who suggested

$$R_{A_e}(f) = 1 - \left(\int y^{1-e} f(y)\, dy \right)^{\frac{1}{1-e}}. \tag{9}$$

Following his work, axiomatizations using a social-welfare function approach appeared in large numbers in the literature. This shift to more formal frameworks corresponded to the rigorous risk axiomatizations appearing in the mathematical psychology literature in the mid-1970s, which are surveyed in Section 2.5. The general entropy measure

$$R_{GEM_e}(f) = \frac{1}{e(e-1)} \left[\int \left(\frac{y}{\mu_y} \right)^e f(y)\, dy - 1 \right] \tag{10}$$

was derived from such an axiomatic approach (see Cowell [1995], p. 60, for details). A special case of this is cardinally equivalent to the independently introduced Herfindahl's index

$$R_H(f) = \frac{1}{\mu_y^2} \int y^2 f(y)\, dy, \tag{11}$$

where E denotes the expectation with respect of $f(y)$. Another suggested expression of this type is Theil's measure

$$R_T(f) = \frac{1}{\mu_y} \int y \log\left[\frac{y}{\mu_y} \right] f(y)\, dy, \tag{12}$$

derived from the notion of entropy in information theory (see Cowell [1995], pp. 56–59, or Theil [1967] for details). Note that this measure is not defined for negative values of y and so is not a suitable measure for data on returns but appropriate for wealth or prices. The more basic range measure is simply

$$R_R(f) = y_{\max} - y_{\min}. \tag{13}$$

We need also mention the coefficient of variation

$$R_{CV}(f) = \frac{R_{SD}(f)}{\mu_y}, \tag{14}$$

effectively a unit-free spread indicator, and the Gini coefficient

$$\begin{aligned} R_G(f) &= \int\int |y_i - y_j| f(y_i) f(y_j)\, dy_i dy_j = 2Cov[y, F(y)] \\ &= \int [1 - F(y)]\, dy - \int [1 - F(y)]^2\, dy, \end{aligned} \tag{15}$$

which is one of the inequality measures most quoted and used by researchers in this field. The Gini has been generalized by Yitzhaki [1983], who defines the extended Gini coefficients

$$R_{EG_V}(f) = \int [1 - F(y)]\,dy - \int [1 - F(y)]^V\,dy \tag{16}$$

for $V > 0$. From the derivation of the Lorenz curve (see, for example, Sen [1973]), we get two more inequality measures—the equal share coefficient

$$R_{ECS}(f) = F(\mu_y) \tag{17}$$

and the minimal majority coefficient

$$R_{MMC}(f) = F\left(G^{-1}\left[\frac{1}{2}\right]\right), \tag{18}$$

where $G(y) = \frac{1}{\mu_y}\int_0^y zf(z)\,dz$, and (F, G) define the Lorentz curve (for details see Sen [1973]).

The class of linear inequality measures we define following Mehran [1976] as

$$R_{LM}(f) = \int (y - \mu_y)W[F(y)]f(y)\,dy, \tag{19}$$

where $W[F(y)] = aF(y) + b$ for $a > 0$—that is, W is a linear function of the cumulative distribution function. Analogously, we define the piecewise linear measures, $R_{PLM}(f)$, which was the original class presented by Mehran, and the nonlinear measures, $R_{NLM}(f)$, which both have the same functional form as (19) but where W is, respectively, piecewise linear or nonlinear.

2.3. Commonly used financial-risk measures

From financial-risk analysis, we draw on a few measures that are used by practitioners as well as theorists today. The first is the probability of making a loss (whose inverse C is the value at risk), defined as

$$P_{PL}(f) = \Pr_f(y < C) \tag{20}$$

for some critical level C. The value at risk is currently used by regulators (see Basle Committee on Banking Supervision [1993]), and the probability of making a loss has been incorporated into safety-first rules by Roy [1952], which were applied to asset pricing by Bawa [1978]. The interplay between safety-first rules and expected utility maximization with regard to portfolio selection is discussed in Pyle and Turnovsky [1970]. Another measure used in applications is the generalized lower partial moment, presented in Bawa and Lindenberg [1977]:

$$R_{LPM}(f) = \left(\int_{-\infty}^{\mu_y} (\mu_y - y)^r f(y)\, dy \right)^{\frac{1}{r}}. \tag{21}$$

It is easy to verify that this measure is closely related to Fishburn's $\alpha - t$ measures (7) and Stone's original class (1). Several reformulations of this measure are possible, as seen in the proofs in the Appendix. More recently, Sinclair, Power, and Lonie [1995] use the results of Greenwood et al. [1979] and Hosking [1986] by adapting the probability weighted moments (PWM)

$$\beta_{r,s}(y) = E[yF(y)^r(1 - F(y))^s] = \int yF(y)^r(1 - F(y))^s f(y)\, dy \tag{22}$$

to define risk measures on a distribution. Note that in (22), varying the parameters r and s will change the emphasis put on the extreme tails of the distribution. Also, $r = 1, s = 0$ yields half the Gini coefficient plus $\frac{\mu_y}{2}$ (see (15)), while $\{\beta_{r,s}(y) : r + s = 2\}$ is a subset of the quadratic class, $\{\beta_{r,s}(y) : r + s = 3\}$ is in the cubic class, and so on (see (19)), whenever $\mu_y = 0$. Sinclair et al. [1995] show that the population counterparts to the important sample L-moments,

$$L_r = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \frac{(r-1)!}{k!(r-k-1)!} \hat{\beta}_{r-k,k}(y), \tag{23}$$

are in fact linear combinations of the sample statistics of (22), denoted $\hat{\beta}_{r-k,k}(y)$, for varying r and k. These provide an alternative method of describing location, dispersion, and skewness, which complement conventional techniques and the usual moments around the mean and as such are potentially good risk measures.

2.4. *Operational research and management science*

Risk has mainly been approached from a utility perspective in the fields of operational research and management science. As a result, risk measures defined explicitly over the cumulative distribution function have been extracted from utility functions that have been shown to be compatible with some type of mean-risk model (see Bell [1995]) or risk-value model (see Sarin and Weber [1993]).

Bell [1995] generalizes his earlier work [1988] and suggests three risk measures. As this article is primarily intended as a survey, we refer the reader to the original work for the motivation behind these. Consider

$$R_{B1}(f) = \frac{2}{c^2} \log\left[\int e^{-c(y-\mu_y)} f(y)\, dy \right], \tag{24}$$

where c is a constant. Equation (24) approaches the variance (5) as c gets closer to zero, while larger c gives more weight to large downside observations. He generalizes this to

$$R_{B2}(f) = \frac{2}{c-a}\left(\frac{1}{c}\log\left[\int e^{-c(y-\mu_y)} f(y)\,dy\right] - \frac{1}{a}\log\left[\int e^{-a(y-\mu_y)} f(y)\,dy\right]\right), \tag{25}$$

where c and a are constants. The third risk measure introduced in Bell in [1995] is

$$R_{B3}(f) = \int e^{-c(y-\mu_y)} f(y)\,dy - 1, \tag{26}$$

which resembles the Sarin measure [41] in its use of the exponential function.

In a recent work on mean-risk models, Kijima and Ohnishi [1993] introduce the following measure

$$R_{KO1}(f) = \left(\int |y-\mu_y|^k f(y)\,dy\right)^{1/k}, \tag{27}$$

where $k > 1$. This is in fact a subcase of Stone (1) since $R_{KO}(f) = R_S[\mu_y, k, \infty](f)$. They extend this to

$$R_{KO2}(f) = \left(c^+ \int_{\mu_y}^{\infty} (y-\mu_y)^k f(y)\,dy + c^- \int_{-\infty}^{\mu_y} (\mu_y - y)^k f(y)\,dy\right)^{1/k} \tag{28}$$

for positive constants c^+ and c^-. This introduces the concept of asymmetry in risk measurement since contributions from below-mean returns can be weighed differently from above-mean returns. We will see more of this type of measures in the next section. Note that for $c^+ = c^- = 1$, (28) reduces to (27).

We mention two works from this literature that deal with risk-value models. Sarin and Weber [1993] use measures from psychology (next section) to define the risk-value models. Their work was extended by Jia and Dyer [1996], who prove that the general form of a risk measure compatible with a suitable expected utility function $U(y)$ through a risk-value model is

$$R(y) = -E[U(y-\mu_y)], \tag{29}$$

which is, of course, expressed in terms of the utility function. As such, we can not directly compare (29) to any of our other measures without specializing further. Jia and Dyer highlight several potentially attractive utility functions and give the associated risk measures. Most of these have already appeared elsewhere in the literature (for example, (41), (26), (27), and (28)). In addition, they introduce

$$R_{JD}(f) = \int (y-\mu_y)^2 f(y)\,dy - \int (y-\mu_y)^3 f(y)\,dy, \tag{30}$$

which allows skewness to affect preferences positively in financial decisions based on the associated risk-value model.

2.5. *Psychology*

Since many readers will not be familiar with this literature, we present an extended survey of the psychology risk literature. The field of experimental psychology has produced interesting approaches to the ranking of risky prospects using concise mathematical axiomatizations, and the attempts at formal treatment of risk measurement undertaken by mathematical psychologists are briefly reviewed in this section. The initial approaches of Coombs and Huang [1970] on percieved risk were taken up by Pollatsek and Tversky [1970], who derive the first closed form risk measure from axioms on the space of density functions to be measured and the specific functional form of the risk function. This was at the same time as the inequality literature, following the result by Atkinson (see (9) in Section 2.2), moved in the direction of more formal axiomatizations. Their work, which is also set out in Chapter 3 of Krantz et al. [1971], was later generalized by Luce and Weber [1986], who separate the effects of the upside and downside effects on overall risk. This is looked at in parallel with the interesting work of Fishburn [1980, 1981], which is the most general work done in this area. The final part of this section is devoted to the work of Luce [1980] and Sarin [1987], who produce specific testable functional forms by considering the effects of addition and scalar multiplication on risk, when the measure takes one of two general forms. We start with the basic result of Pollatsek and Tversky (PT) in [1970], briefly discussing their assumptions and final measure. For the formal proof, we refer the reader to the original paper.

Like most formal derivations of risk measures, PT assume the existence of a binary relation of risk, $\preceq$, on a set of distribution functions, Ω, on the real line. Ω is assumed closed under convolutions. The double $\{\Omega, \preceq\}$ is a weak order—that is, complete and transitive. For three distinct distributions f, g, and h all in Ω, $f \succeq g \Leftrightarrow f \circ h \succeq g \circ h$, where $\circ$ denotes convolution—that is, the convolution of two distinct distributions with a third distribution, statistically independent of both, leaves the risk rankings of the original two unchanged. Under these assumptions and the further restriction that no distribution may take on infinite risk, (PT) prove the existence of a risk functional R, which preserves the ordering $\preceq$ and has the property that $R(f \circ g) = R(f) + R(g)$—that is, the risk of a convolution of two distributions is the sum of the risks of those individual distributions.

By imposing further restrictions on $\{\Omega, \preceq\}$ and the functional R, one can establish more specific functional forms of the risk measure. PT assume *completeness*—that is, for all f in Ω, the distributions f_α induced by multiplying all outcomes in f by a scalar α and sure-thing distributions (those yielding any number β with probability one), are all in Ω. Further, they assume *positivity* that is, convoluting a risky distribution with a sure-thing reduces risk) and *scalar monotonicity* (that is, if $E(f) = 0$ and $\alpha > 0$, then $R(f_\alpha) > R(f)$ as well as if $E(f) = E(g) = 0$ and $f \succeq g$, then $R(f_\alpha) > R(g_\alpha)$). Finally, a *continuity* assumption on the actual functional representation R requires that a series of distributions in Ω converging to say, f, will induce a convergent series of risk measures with limit $R(f)$.

Under these assumptions, PT prove that the risk functional takes the form

$$R_{PT}(f) = aVar(f) - (1-a)E(f). \tag{31}$$

The proof is given in detail in Pollatsek and Tversky [1970].

Observe that this is almost the negative of the expected quadratic utility function of mean-variance analysis, except the expected value does not enter quadratically but linearly in (31). Moreover, this analysis is based on pairwise independent random variables, which may be more applicable to psychology, where latent factors are thought of as independent via factor analysis. Although used in arbitrage pricing theory by Ross [1976], this assumption is not generally accepted when working with financial data.

Apart from this, (31) is undesirable for at least two other reasons. First, the risk function is linear in the expected value, and since we seek to separate the effects on decision-making of expectation and risk, (31) may not be a good risk measure as it captures both effects simultaneously. Second, the measure takes no account of higher-order moments, a major criticism that even the CAPM model suffers under, and one may justifiably conclude that the analysis of PT is geared toward individuals perception of the risk of relatively simple gambles rather than the financial equilibrium applications we have in mind.

We now turn to the work of Fishburn [1980, 1981], who draws on the work of Krantz et al. [1971] and derives measures using a separation of the distribution function. Fishburn had previously applied measures of downside risk to define decision rules in a financial context in [1977], which are shown to be congruent to some notion of expected utility.

His work is extensive and we include a brief discussion of his approach along with the measures he finally derives. As he used twenty four axioms in all, we refer the reader to the cited papers for a more rigorous presentation and the formal proofs of his theorems.

Let Θ be the set of all possible outcomes, Θ^- the set of outcomes less than a preset target (assumed without loss of generality to be zero), and similarly Θ^+ is the set of outcomes larger than zero. The inclusion of an arbitrary target to separate "gain" from "loss" looks appealing in a financial context since asset managers often seek to outperform some benchmark, not always identical to the riskless rate. Fishburn further defines P^- to be the set of probability measures that always yield a loss—that is, $p(\Theta^-) = 1$ for all $p \in P^-$ and P^+ is similarly defined. Both these sets are assumed closed under convex combinations. A general probability measure then lies in the set

$$\Gamma = \{(a, p, b, q) : a \geq 0, b \geq 0, a + b \leq 1, p \in P^-, q \in p^+\}$$

so that the measure assigns probability $ap(Y)$ to each $Y \in \Omega^-$ and $bq(Z)$ to each $Z \in \Omega^+$, leaving probability $(1 - a - b)$ for the zero outcome. Hence, a is a loss probability (set $Y = \Omega^-$), b is a gain probability (set $Z = \Omega^+$) and p, q are the conditional loss and gain distributions respectively. There is a unique mapping from a density function to a member of Γ, and we are able to separate the effects of upside and downside upon overall risk measurement.

Fishburn [1981] proceeds to propose alternative sets of axioms for ranking any two such members of Γ. In particular, he is able to generalize his earlier work [1980] on the same topic and so assumes different properties for upside and downside. Fishburn's first six basic axioms are much along the lines of those of PT. The ranking needs to be a weak order, risk increases in a and decreases in b if $a > 0$, and sure-thing distributions are ranked in order of what they guarantee. He assigns zero risk if and only if $a = 0$, which contradicts one of our basic properties later. Continuity and upper boundedness of risk are assumed, and the existence of a general functional form established.

Fishburn places more restrictions and proves six theorems presenting increasingly more specific functional forms. We present the final risk functions derived and leave the reader to examine the formal details in the cited paper. Consider first the risk functional

$$R_{F1}(f) = T(a, p)S(b, q), \tag{32}$$

where $T(a, p)$ and $S(b, q)$ are both continuous and unique up to exponential transformations, with $T(a, p)$ increasing in a and $S(b, q)$ decreasing in b. This is a very general way of separating the effects of upside and downside effects on overall risk using a multiplicative aggregation technique, referred to by Fishburn as *separability*. He next makes specific the forms of $T(a, p)$ and $S(b, q)$ to derive *separable expected loss risk*. Assuming lower boundedness and continuity of functions whose expectations we wish to examine, gives

$$R_{F2}(f) = S(b, q) \int_{\Theta^-} T(a, x)\, dp(x) \tag{33}$$

—that is, the downside component is an expectation with respect to the conditional downside distribution. Here, $T(a, x)$ is increasing in a, decreasing in x and unique up to a similarity transformation. By assuming a similar form for $S(b, q)$, we obtain

$$R_{F3}(f) = \int_{\Theta^+} S(b, y)\, dq(y) \int_{\Theta^-} T(a, x)\, dp(x). \tag{34}$$

The properties of $S(b, y)$ are the obvious analogues to those described for $T(a, x)$ above. Extending this further, Fishburn gives necessary axioms to get *separable loss and gain probabilities*—that is, the two-argument functions $T(a, x)$ and $S(b, y)$ are split into two separating functions ($T_1(a)$, $T_2(x)$, and $S_1(b)$, $S_2(y)$, respectively) in their respective arguments. This gives an opportunity to further differentiate upside and downside effects and yields

$$R_{F4}(f) = \left[T_1(a) \int_{\Theta^-} T_2(x)\, dp(x)\right]\left[1 - S_1(b) \int_{\Theta^+} S_2(y)\, dq(y)\right], \tag{35}$$

where the separate functions retain the induced properties from their parental functions. For details of even more specific constructions, all special cases of (35), we refer the reader to Fishburn ([1981], pp. 240–243). It is worth noting that by taking logs of the presented measures we could get an additive framework rather than a multiplicative one, some cases of which are examined in the later analysis by Luce [1980] and Sarin [1987]. It is also worth noting that the method of separating the upside and downside of a distribution has not only been used by psychologists but also by financial econometricians (e.g., see Knight, Satchell, and Tran [1995]). Indeed, this is a natural subdivision of financial data. One analyzes returns, and risk, by the upside and downside probabilities, and conditional on the sign of the realized return, by the magnitude.

In a separate study, Luce and Weber [1986] axiomatize a risk measure much in the same way as Pollatsek and Tversky, although they use mixture distributions and not convolutions as a means of aggregation. Like Fishburn, they also separate the distribution and derive the following measure, known as *conjoint expected risk*:

$$\begin{aligned} R_{LW}(f) &= \alpha_0 \Pr(y=0) + \Pr(y>0)(\alpha_+ + \beta_+ E(y^{k_+} \mid y>0)) \\ &\quad + \Pr(y<0)(\alpha_- + \beta_- E(y^{k_-} \mid y<0)), \end{aligned} \tag{36}$$

where $\alpha_0,\ \alpha_+,\ \alpha_-,\ \beta_+,\ \beta_-,\ k_+$, and k_- are all constants with $k_+ > 0$ and $k_- > 0$.

This is a significant improvement on Pollatsek and Tversky and considers loss with different weight than gain, which certainly should be desirable for anyone engaging in a monetary gamble, if non-risk-neutral at least. Two higher partial moments are also allowed in the risk measure. Finally, note that this measure in fact is a linear transformation of a logarithmic type of the Fishburn measures (32) to (35) discussed above.

Another approach to axiomatization of functional forms of risk measures was taken by Luce [1980] and Sarin [1987]. They considered the effects of addition and scalar multiplication on the desired-risk measure, allowing the actual risk function to take one of two different general forms. Each author produces several measures (two of Luce's were corrected a year later in Luce [1981] using similar procedures. We thus present their analysis simultaneously, giving the axioms presented by both authors, stating their theorems as one major result and illustrating how their proofs can be considerably shortened if one knows in advance the functional form of the measure.

Recall that $R(f)$ denotes the risk of distribution f, f_α denotes the distribution induced by multiplying all outcomes in f by the constant α, and f^β denotes the distribution induced by adding a constant β to all outcomes.

Axiom 1 (LIA): *For a strictly increasing function* $S(\alpha)$ *with* $S(1)=0$ *defined for* $\alpha > 0$

$$R(f_\alpha) = R(f) + S(\alpha).$$

Axiom 2 (L1M): *For an increasing function* $S(\alpha)$ *with* $S(1)=1$ *defined for* $\alpha > 0$

$$R(f_\alpha) = R(f)S(\alpha).$$

Axiom 3 (L2): *There exists a nonnegative function* T, *with* $T(0)=0$, *such that for all* f,

$$R(f) = \int T[f(x)]\,dx$$

(*that is*, *transform of the actual density function*).

Axiom 4 (L3): *There exists a function T such that for all densities f,*

$$R(f) = \int T(x) f(x)\,dx = E[T(x)].$$

Axiom 5 (S1): *For a strictly monotone function $S(\beta)$ defined for $\beta > 0$*

$$R(f^{\beta}) = R(f) + S(\beta).$$

Axiom 6 (S2): *There are functions $T(x)$ and $S(x)$ such that*

$$R(f) = \int T(x) f(x)\,dx + \frac{1}{2}\left(\int S(x) f(x)\,dx\right)^2.$$

The abbreviated expressions above in parentheses refer to the names of the axioms in the original papers, S denoting Sarin [1987] and L denoting Luce [1980]. As mentioned earlier, the proofs of Luce and Sarin can be considerably shortened if one knows the answer. We illustrate how for one of the cases of Luce, leaving the others available on request from the authors. We need the following preliminary lemmas:

Lemma 1 (Part 1 of proof of Luce [1980], Theorem 1): *If Axiom* 1 *holds, then* $S(\alpha) = A \log \alpha$ *for* $A > 0$.

Lemma 2 (Part 1 of proof of Luce [1980], Theorem 2): *If Axiom* 2 *holds, then for* $\theta > 0$, $S(\alpha) = \alpha^{\theta}$.

The proofs of these two Lemma's are given as the first part of Luce's major proofs and follow from Aczel [1966]. The following Lemma is our own, given to simplify the major proofs to follow.

Lemma 3: *If Axiom* 3 *holds, then*

$$\int_{R(y)} T[f_{\alpha}]\,dy = \alpha \int_{R(x)} T\left[\frac{1}{\alpha} f(x)\right] dx,$$

where y is induced by multiplying all outcomes x by the constant α.

Proof.

$$\int_{R(y)} T[f_{\alpha}]\,dy = \int_{R(\alpha x)} T[f_{\alpha}]\,d(\alpha x) = \int_{R(x)} T\left[\frac{1}{\alpha} f(x)\right] d(\alpha x) \text{ and } d(\alpha x) = \alpha\,dx.$$

□

Similar Lemmas exist for the other assumptions on the functional form, but these are omitted as they will not play a part in our demonstration proof. We not turn to the theorems as stated in Luce [1980], with corrections in Luce [1981] and Sarin [1987] proving one case only as the others follow an identical path.

Theorem 4: *The following results are taken directly from the original papers*:

- *Luce* ([1980], *Theorem* 1): *If Axiom* 1 *and Axiom* 3 *hold*, *then for constants* $A > 0$ *and* $B \geq 0$,

$$R_{L1}(f) = -A \int f(x) \log[f(x)]\, dx + B, \tag{37}$$

which is closely related to the entropy measure (10).

- *Luce* ([1980], *Theorem* 2): *If Axiom* 2 *and Axiom* 3 *hold*, *then for constants* $B, \theta > 0$ *and* $A > 0$,

$$R_{L2}(f) = A \int f(x)^{1-\theta}\, dx + B. \tag{38}$$

- *Luce* ([1981], *Theorem* 3 *corrected*): *If Axiom* 1 *and Axiom* 4 *hold*, *then for constants* $A > 0$, B *and* C,

$$R_{L3}(f) = B \int_0^{\infty} f(x)\, dx + C \int_{-\infty}^{0} f(x)\, dx + A \log(|\tilde{x}|). \tag{39}$$

- *Luce* ([1981], *Theorem* 4 *corrected*): *If Axiom* 2 *and Axiom* 4 *hold*, *then for constant* $\theta > 0$,

$$R_{L4}(f) = A \int_0^{\infty} x^{\theta} f(x)\, dx + B \int_{-\infty}^{0} |x|^{\theta} f(x)\, dx, \tag{40}$$

where $A = (\theta + 1) \int_0^1 T(x)\, dx$ *and* $A = (\theta + 1) \int_{-1}^{0} T(x)\, dx$.

- *Sarin* ([1987], *Theorem* 1): *If Axiom* 3 *and Axiom* 5 *hold*, *then for constants A and B*,

$$R_{S1}(f) = A \int_0^{\infty} e^{Bx} f(x)\, dx, \tag{41}$$

where $sign(A) = -sign(B)$.

- *Sarin* ([1987], *Theorem* 2): *If Axiom* 2 *and Axiom* 6 *hold*, *then for constants* $A > 0$, B *and* C,

$$R_{S2}(f) = B \int_0^{\infty} f(x)\, dx + C \int_{-\infty}^{0} f(x)\, dx + A E(\log|\tilde{x}|) - \frac{A^2}{2} Var(\log|\tilde{x}|). \tag{42}$$

- *Sarin* ([1987], *Theorem* 3): *If Axiom* 1 *and Axiom* 6 *hold*, *then for constants* $A > 0, \theta > 0$ *and* C,

$$R_{S3}(f) = C\int_0^\infty |\tilde{x}|^\theta f(x)\,dx + A\int_{-\infty}^0 |\tilde{x}|^\theta f(x)\,dx + \frac{\theta}{2(\theta-1)}E(|\tilde{x}|^{2\theta}) - \frac{1}{2}Var(|\tilde{x}|^\theta). \tag{43}$$

Proof. We present the Luce 1 case (37), as all other cases follow an identical pattern. This version of the proof relies on prior knowledge of the solution and is thus a much simplified version of the original proof. By Lemma 1, Lemma 3, and Axiom 1,

$$\alpha\int_{R(x)} T\left[\frac{1}{\alpha}f(x)\right]dx = \int_{R(x)} T[f(x)]\,dx + A\log\alpha$$

for $A > 0$. Substituting in $T(x) = x[B - \log x]$, this will be the correct solution iff the above equation holds for A and B both positive. We get

$$\alpha\int_{R(x)} \frac{1}{\alpha}f(x)\left(B - \log\left[\frac{1}{\alpha}f(x)\right]\right)dx = \int_{R(x)} f(x)(B - \log f(x))\,dx + A\log\alpha$$

for $A > 0$. Multiplying out and cancelling terms gives

$$\int_{R(x)} f(x)\log\alpha\,dx = A\log\alpha,$$

which trivially hold for any B and $A = 1$, since f(x) integrates to 1. Hence, $T(x) = x[B - \log x]$ is the correct for $T(x)$. Substituting this into Axiom 3 gives the result. □

Sarin argues that all of Luce's measures are too specialized to be applied to positive returns although they could be used for negative returns (details in Sarin [1987]). His own improvements to this are mainly through the generalized Axiom 6 (S2), which allows the introduction of a second moment term, but the overall benefit of this change has not yet been thoroughly empirically examined.

3. A new family of risk measures and isolation of "good" measures

We now present a new general class of risk measures and its special cases, which includes, or is closely related to, all of the previously presented measures from finance, Stone's class (1) and Mehran's work (19), as well as almost all inequality measures. We also show that this class covers most measures derived from the psychology approaches, except those of Fishburn (32) to (35), and is related to numerous others. Further possible generalizations are suggested and discussed. We subsequently present four axioms we believe describe necessary features of a “good” financial risk measure and test all measures surveyed against these.

3.1. A new class of risk measures

Based on Stone’s class of risk measures (1), we present the following extended family of risk measures.

Definition 5: *The risk measure* $R[A,\ b,\ \alpha,\ \theta,\ W(\cdot)]$ *is defined by*

$$R[A,\ b,\ \alpha,\ \theta,\ W(\cdot)] = \left[\int_{-\infty}^{A} |y-b|^{\alpha} W[F(y)] f(y)\, dy\right]^{\theta} \tag{44}$$

for some bounded function $W(\cdot)$ *and real numbers* $A, b, \alpha,$ *and* θ, *where* $\theta > 0$ *and* $\alpha > 0$.

This new family contains, or is related to, a large number of the risk measures surveyed in this article. The following are special cases of (44):

- Stone’s measure (1) is

$$R\left[A,\ b,\ \alpha,\ \frac{1}{\alpha},\ 1\right].$$

Consequently, we automatically have the following:

- Standard deviation (2) is

$$R\left[\infty,\ \mu_y,\ 2,\ \frac{1}{2},\ 1\right].$$

- Semistandard deviation (3) is

$$R\left[\mu_y,\ \mu_y,\ 2,\ \frac{1}{2},\ 1\right].$$

- Mean absolute deviation (4) is

$$R[\infty,\ \mu_y,\ 1,\ 1,\ 1].$$

- Kijima and Ohshini’s first measure (27) is

$$R\left[\infty,\ \mu_y,\ \alpha,\ \frac{1}{\alpha},\ 1\right].$$

Furthermore, to generalize Stone's measure, we can now also include the following:

- The variance (5) is

$$R[\infty,\ \mu_y,\ 2,\ 1,\ 1].$$

- The semivariance (6) is

$$R[\mu_y,\ \mu_y,\ 2,\ 1,\ 1].$$

- The probability of making a loss (20) is

$$R[0,\ 0,\ 0,\ 1,\ 1].$$

- Lower partial moments (21) are

$$R\left[\mu_y,\ \mu_y,\ \alpha,\ \frac{1}{\alpha},\ 1\right].$$

- Fishburns $\alpha - t$ measures (7) are

$$R[t,\ t,\ \alpha,\ 1,\ 1].$$

So far, we have not utilized the $W(\cdot)$-function in (44). By abstracting in this direction, we capture the following additional measures:

- The probability weighted moments (22) are

$$R[\infty,\ 0,\ 0,\ 1,\ F(y)^r(1-F(y))^s].$$

- The extended Gini coefficient (16) is

$$R[\infty,\ \mu_y,\ 1,\ 1,\ 2(1-F(y))^k \text{ for positive integer } k].$$

- The linear[1] measures (19) are

$$R[\infty,\ \mu_y,\ 1,\ 1,\ aF(y)+b \text{ for constants } a \text{ and } b, \text{ where } a>0].$$

- The Gini coefficient (15) is

$$R[\infty,\ \mu_y,\ 1,\ 1,\ 2F(y)].$$

In addition, further special cases of the extended Stone family describes several more measures up to a linear transformation. In particular,

- Dalton's measure (8) is

$$\left(1+\frac{1}{\mu_y^{1-e}-1}\right)-\left(\frac{1}{\mu_y^{1-e}-1}\right)R[\infty,\ 0,\ 1-e,\ 1,\ 1].$$

- Atkinson's measure (9) is

$$1-R\left[\infty,\ 0,\ 1-e,\ \frac{1}{(1-e)},\ 1\right].$$

- The general entropy measure (10) is

$$\left(-\frac{1}{e(e-1)}\right)+\left(\frac{1}{\mu_y^e e(e-1)}\right)R[\infty,\ 0,\ e,\ 1,\ 1].$$

- Herfindahl's index (11) is

$$\frac{1}{\mu_y^2}R[\infty,\ 0,\ 2,\ 1,\ 1].$$

- The coefficient of variation (14) is

$$\frac{1}{\mu_y}R\left[\infty,\ \mu_y,\ 2,\ \frac{1}{2},\ 1\right].$$

- The Pollatsek and Tversky (31) measure is

$$aR[\infty,\ \mu_y,\ 2,\ 1,\ 1]-(1-a)\mu_y.$$

One measure is the linear combination of two extended Stone measures:

- The Jia and Dyer measure (30) is

$$R[\infty,\ \mu_y,\ 2,\ 1,\ 1]-R[\infty,\ \mu_y 3,\ 1,\ 1].$$

Sarin's third measure (43), Luce and Weber (36), Luce's fourth measure (40) and Kijima and Ohnishi's second measure (28) all include terms that are in (44), but additional terms defined on ranges with finite lower limit. Since we elect to fix our lower limit at $-\infty$, in order to keep to a five-parameter family while preserving downside risk considerations, these can not be fully covered by our class.

The measures that are not directly related to (44) under this general class include the inequality measures derived from the Lorentz curve (that is, the equal shares coefficient (17) and minimal majority coefficient (18)), as well as some of the psychology measures. They include the general axiomatic work of Fishburn from the psychology literature (32), Luce's measures (37) to (39), since they employ functions of the density function $f(y)$ rather than the cumulative density function. Similarly, Sarin's (41) and (42), Bell's measures (24) to (26) and Theil's measure (12) involve expectations of either logarithmic or exponential functions of returns, in contrast to (44), which uses a polynomial. Finally, the range (13), really a sample measure, cannot be expressed as a case of (44).

The new family of risk measures hence incorporates many general classes and special cases from each academic discipline surveyed, including all the financial measures as well as Stone (1) and Mehran's (19) original work while maintaining an algebraically simple format. If the only aim had been to include as many special cases as possible, we could easily have chosen more general families. For instance, the 6 K-parameter family

$$R[K, A_k, c_k, b_k, \alpha_k, \theta_k, W_k(\cdot)] = \left[\sum_{k=1}^{K} \int_{-\infty}^{A_k} e^{-c_k y} |y - b_k|^{\alpha_k} W_k[F(y)] f(y)\, dy \right]^{\theta_k} \tag{45}$$

will capture significantly more cases even at $K = 2$, while the 11 K-parameter measure $R[K,\ A_k^U,\ A_k^D,\ d_k,\ f_k,\ c_k,\ b_k,\ \alpha_k,\ \theta_k,\ W_{1k}(\cdot),\ W_{2k}(\cdot)] =$

$$\left[\sum_{k=1}^{K} \int_{A_k^D}^{A_k^U} (\log d_k y)^{f_k} e^{-c_k y} |y - b_k|^{\alpha_k} W_{1k}[f(y)] W_{2k}[F(y)] f(y)\, dy \right]^{\theta_k} \tag{46}$$

captures all measures surveyed when $K = 5$. The algebraic complications that accompany (45) and (46) are, not surprisingly, far greater than the usefulness of having such a general form. Consequently, these two are included merely to indicate possible generalizations, and we do not pursue them in further detail.

In order to select between all the measures presented in this article so far, both those in the new family (44) and those left out, we now turn to suitable decision criteria.

3.2. *Desirable properties of a financial risk measure*

In order to specify axioms from which to separate out the more desirable financial risk measures, it is essential to fully understand exactly what it is we wish to measure. A *risk measure in the financial sense* will be a measure that, given a random variable $\tilde{y}$, returns a nonnegative number on the real line. Thus, we effectively seek a functional

$$R[\tilde{y}] : \Omega \mapsto \Re^+, \tag{47}$$

where we allow $\tilde{y}$ to vary in Ω, depending on the financial entity whose risk we wish to measure. We may elect to restrict wealth to be positive or arithmetic returns to be larger than -1 (that is, limited liability). We do not allow negative risk and the functional should be able to accommodate a binary relation of risk. In addition, we require the following properties to be satisfied.

Definition 6: *Let $\tilde{y}$ be a random variable. The basic properties (BP) of a financial risk measure functional $R[\tilde{y}]$ as defined in* (47) *are*

($BP1$) (*Nonnegativity*): $R[\tilde{y}] \geq 0$.
($BP2$) (*Homogeneity*): $R[\lambda \tilde{y}] = |\lambda|\, R[\tilde{y}]$ *for* $\lambda \geq 0$.
($BP3$) (*Subadditivity*): $R[\tilde{y}_1 + \tilde{y}_2] \leq R[\tilde{y}_1] + R[\tilde{y}_2]$.
($BP4$) (*Shift-invariance*): $R[\tilde{y} + \lambda] \leq R[\tilde{y}]$ *for all* λ.

These are adapted from the axioms employed by Kijima and Ohnishi [1993] in their definition of risk in risk-value models but are strengthened in that BP4 now has to hold for all λ rather than just $\lambda \geq 0$. Nonnegativity and shift invariance were also considered necessary properties of a risk measure by Bell [1995]. Taken together, BP2 and BP3 imply that constant distributions are assigned zero risk, in keeping with our wish of measuring risk as dispersion only, essential to fully explore the effects of location and dispersion on financial decision-making in mean-risk setups. Also, BP2 and BP4 can be used to show that our measure $R[\tilde{y}]$ is convex, which implies that diversification will reduce risk and one obtains separation in optimal portfolio selection. The advantage of convexity also follows from the definition of second-degree stochastic dominance, which we denote by SSD. It is well known that if $X_{ssd}Y$, then $Y = X + \epsilon$, where $E(\epsilon \mid X) = 0$ (see, for instance Rothschild and Stiglitz [1970]). Huang and Litzenberger [1988] then show that if $G(.)$ is a convex function, then

$$E[G(Y)] = E[G(X+\epsilon)] = E[E[G(X+\epsilon)] \mid X] \leq E[G(X)]$$

—that is, the risk measure assigns less risk to stochastically dominating investments when (ii) is satisfied. It may be possible to link other notions of stochastic dominance to our risk measures, but we do not pursue this in our article. Additionally, BP2 forces our measure to respond proportionally to scale changes, as homogeneity of degree one gives a constant risk-return tradeoff in portfolio selection.

BP4 makes the measure invariant to the addition of a constant to the random variable. This might seem to be counterintuitive at first, since we will in effect allow some acceptable measures to assign equal risk to two coin tosses, where one pays -1 for "heads," $+1$ for "tails," and the other $10^{100} - 1$ for "heads," $10^{100} + 1$ for "tails." However, this may be resolved by reiterating that our risk measures are defined in terms of dispersion only, so that when combined with some location measure in a mean-risk model, the latter gamble will undoubtedly be preferred to the former (indeed it stochastically dominates the former) due to the vast difference in their expected value. This is a classic example where, from a financial perspective, risk and return are not effectively separated.

We next present the main lemma and theorems of the article, establishing which risk measures surveyed satisfy our three basic properties. All proofs can be found in the Appendix.

Lemma 7: *A general member $R[A,\ b,\ \alpha,\ \theta,\ W(\cdot)]$ of the class* (44) *satisfies BP1, BP2, and BP4 if $b = \mu_y$, $\theta = \frac{1}{\alpha}$ and $W(.)$ is positive valued.*

Theorem 8: *The following measures fail* (*BP*):

(a) *Pollatsek and Tversky* (31)
(b) *Luce and Weber* (36)
(c) *Luce's measures* ((37) *to* (40)) *except trivial special cases*
(d) *Sarin's measures* ((41) *to* (43))
(e) *Fishburn's psychology measures* ((32) *to* (35))
(f) *Dalton's measure* (8)
(g) *Atkinson's measure* (9)
(h) *General entropy measure* (10)
(i) *Herfindahl's index* (11)
(j) *Theil's measure* (12)
(k) *Coefficient of variation* (14)
(l) *Equal shares coefficient* (17)
(m) *Minimal majority coefficient* (18)
(n) *General Fishburn* $\alpha - t$ *class* (7)
(o) *The probability of making a loss* (20)
(p) *L-moments* (23)
(q) *Variance* (5)
(r) *Semi variance* (6)
(s) *Bell's measures* ((24) *to* (26))
(t) *Jia and Dyer's measure* (30).

Theorem 9: *The following measures satisfy* (*BP*):

(a) *Stone's Class for* $k > 1$ *and* $Y_0 = \mu_y$ *and* $A = \mu_y$ *or* $A = \infty$ (1)
- *Standard deviation* (2)
- *Mean absolute deviation* (4)
- *Fishburn's* $\alpha - t$ *measures for* $t = \mu_y$ *raised to power* $\frac{1}{k}$ (7)
- *Semistandard deviation* (3)
- *The first Kijima-Ohnishi measure* (27)
- *Generalized lower partial moment* (21)

(b) *The range* (13)
(c) *The piecewise linear measures* (19)
- *The Gini coefficient* (15)
- *The L-moments for* $r + s < 2$ (23)

(d) *Kijima and Ohsniki's second measure* (28).

Although the number of measures failing our proposed axioms seems very large, a closer look at the list does illustrate that a majority of these are from fields other than finance. Some of the measures that seem to fail our axioms but are commonly used risk measures on the empirical side have a related measure satisfying all four axioms. For instance, the general Fishburn $\alpha - t$ class (7), which does not satisfy BP2 or BP4, does have a related special case deemed satisfactory. Likewise, the standard deviation and semistandard deviation satisfy our axioms, while the variance and semivariance fail property BP2, merely because squaring the former two makes the resulting measures homogeneous of degree two instead of one. Furthermore, the scaling of the standard deviation by the mean to get the coefficient of variation (14) results in homogeneity of degree zero and so again violation of BP2. The only measure used in popular financial decision models that fails three properties (BP2, BP3, and BP4) is the probability of making a loss (20), interestingly probably the most discussed risk measure in the practitioner literature at present due to its close links with the value at risk. Note that although the newly defined family of risk measures (44) does not include all the presented measures, it does have as special cases most the measures that satisfy all three axioms as well as those that fail one or more of the properties but are commonly used.

An earlier version of this article deals with a similar set of axioms, all of which are implied by the above except the additional requirement that zero risk should be assigned only if a distribution is constant. This new property could be added without changing the main results. Also, it is worth mentioning that one could automatically relax BP2 to

$$(BP2') \quad R[\lambda\tilde{y}] \begin{cases} = |\lambda|\, R[\tilde{y}] & |\lambda| \geq 1 \\ \leq |\lambda|\, R[\tilde{y}] & |\lambda| \leq 1, \end{cases}$$

which follows from BP2, BP3, and BP4. An anonymous referee has suggested relaxing BP2 further to

$$(BP2'') \quad R[\lambda\tilde{y}] \begin{cases} \geq |\lambda|\, R[\tilde{y}] & |\lambda| \geq 1 \\ \leq |\lambda|\, R[\tilde{y}] & |\lambda| \leq 1, \end{cases}$$

which would imply that doubling your investment would more than double risk, while halving your investment would result in risk reducing by a factor larger than 2. This is a considerable relaxation, and the question of what further measures can be accepted under these or similar slacker conditions merits a survey of possible axioms in itself. Consequently, this survey sticks with the axioms at hand and leaves to future research the investigations of stronger theorems than those presented.

4. Conclusion

We have presented an extended survey of risk measures from several academic disciplines. We have defined a new family of risk measures that accommodates most presently used measures of risk in finance and those from related fields satisfying four criteria we would expect any good financial risk measure to satisfy. The analysis of financial decision-making can thus be generalized substantially to account for axiomatically derived general risk families rather than the present use of the variance and other measures selected for their algebraic simplicity or by personal preference. The new family also contains the common statistical moments around the mean as well as the alternative L-moments.

Further research along the lines of Fishburn [1977] and Holthausen [1981] can be used to relate decision rules based on combinations of returns and risk measures to expected utility decision rules. This links our research to the work of Machina [1982] and Yaari [1987] where expected utility functions that are nonlinear in probabilities are examined. Our general measure (44) holds several risk measures that, when combined with the mean in some mean-risk decision rule, induce an expected utility function nonlinear in probabilities. This is certainly the case when $W[F(y)]$ is nonconstant, as illustrated for the mean-Gini model in Shalit and Yitzhaki [1984]. It has been shown by Newbury [1970] that the Gini coefficient is inconsistent with expected utility theory.

Contrary to the notion of stochastic dominance, which is common tool in risk comparisons but produces only a partial ranking or prospects, our risk measures rank all investment opportunities. Additionally, our axioms BP2 and BP4 ensure that acceptable risk measures are convex and so assign lower risk to stochastically dominating portfolios.

For the measures that are congruent to expected utility theory, one can take his one step further. Bell [1988] combines measures of risk and return to generate risk-return models, congruent to specific utility functions. Sarin and Weber [1993], who introduce the risk-value models, likewise link explicit measures of risk to decision models. In finance, one can now seek to explain observed anomalies in financial data that were fitted to models relying on specific risk measures by switching to alternative explicit risk representations. On a theoretical basis, asset pricing models can be constructed and equilibrium risk measures derived, which depend on the form of the risk measure implicit in the individual utility functions. These could complement the CAPM models of Shalit and Yitzhaki [1984], Bawa and Lindenberg [1977] and Harlow and Rao [1989] as well as the original Sharpe-Lintner CAPM.

Appendix

For notational convenience, we sometimes work with the measures in expectations form. Note that

$$\int g(x) f(x)\, dx = E[g(x)]$$

and further that the expectation of $g(x)$ is a linear operator (that is, $E[ag(x)+b] = aE[g(x)]+b$).

Proof of Lemma 7

Consider (44) with $b=\mu_y$, $\theta=\frac{1}{\alpha}$ and $W[F(y)]$ positive valued.

BP1 Since $W(.)$ is positive valued, everything under the integral is nonnegative. Result then follows on integration.

BP2

$$\begin{aligned} R(\lambda y) &= (E[|\lambda y - \lambda\mu_y|^\alpha W[F_{\lambda y}(\lambda y)]])^{\frac{1}{\alpha}} \\ &= (|\lambda|^\alpha E[|y-\mu_y|^\alpha W[F_y(y)]])^{\frac{1}{\alpha}} \end{aligned}$$

since $\lambda \geq 0$. This becomes

$$= |\lambda|(E[|y-\mu_y|^{\alpha} W[F_y(y)]])^{\frac{1}{\alpha}}$$
$$= |\lambda| R(y).$$

BP4

$$\begin{aligned} R(y+\lambda) &= (E[|(y+\lambda)-(\mu_y+\lambda)|^{\alpha} W[F_{y+\lambda}(y+\lambda)]])^{\frac{1}{\alpha}} \\ &= (E[|y-\mu_y|^{\alpha} W[F_y(y)]])^{\frac{1}{\alpha}} \\ &= R(y). \end{aligned}$$

□

Proof of Theorem 8

The theorem is proved by exhibiting the violation of one axiom for each proposed measure.

(a) **BP1** Let $a = \frac{1}{2}$ and $y \sim N(10, 1)$. Then $R_{PT}(y) = -4\frac{1}{2} < 0$.
(b) **BP2** Let $\alpha_0 = \alpha_+ = \alpha_- = 0$, $\beta_+ = \beta_- = 1$, $\Pr(y > 0) = \Pr(y < 0) = \frac{1}{2}$. Then,

$$\begin{aligned} R_{LW}(\lambda y) &= \frac{1}{2}(E_{\lambda y}[(\lambda y)^{k_+} \mid \lambda y > 0] + E_{\lambda y}[(\lambda y)^{k_-} \mid \lambda y < 0]) \\ &= \frac{1}{2}(E[(\lambda y)^{k_+} \mid y > 0] + E[(\lambda y)^{k_-} \mid y < 0]) \\ &= \frac{1}{2}(\lambda^{k_+} E[y^{k_+} \mid y > 0] + \lambda^{k_-} E[y^{k_-} \mid y < 0]). \end{aligned}$$

Now suppose $y \sim N(0, 1)$, $\lambda = 2$, $k_+ = 1$ and $k_- = 2$. Then

$$R_{LW}(\lambda y) = \frac{1}{2}\left[4\left(\frac{1}{2}\right)\right] = 1.$$

However,

$$|\lambda|\ R_{LW}(y) = 2\left(\frac{1}{2}\right)4\left(\frac{1}{2}\right) = 2.$$

(c) + (d) We start with R_{L1}, R_{L3}, and R_{L2}, which rely on an axiom of Luce—namely, (L1A). This is incompatible with BP2. To see this, suppose both are satisfied and consider two distinct r.v.'s $\tilde{x}$ and $\tilde{y}$, where $R[\tilde{x}] > R[\tilde{y}]$. By L1A,

$$R[\lambda\tilde{y}] = R[\tilde{y}] + S(\lambda) \text{ and } R[\lambda\tilde{x}] = R[\tilde{x}] + S(\lambda).$$

By BP2,

$$R[\lambda\tilde{y}] = |\lambda|\ R[\tilde{y}] \text{ and } R[\lambda\tilde{x}] = |\lambda|\ R[\tilde{x}].$$

Subtraction gives

$$R[\lambda\tilde{y}] - R[\lambda\tilde{x}] = R[\tilde{y}] - R[\tilde{x}] = |\lambda|\,(R[\tilde{y}] - R[\tilde{x}]),$$

which is not satisfied if $\lambda \neq 1$.

Next consider R_{L2} and R_{L4}, which are derived using axiom Luce (L1M). By a similar argument to the above, we can show that we need $S(a) = a$ in L1M for it be compatible with BP2. Following the proofs of Luce, this further implies that R_{L2} is a constant and so uninteresting. R_{L4} becomes

$$A\left[\int_{-\infty}^{0} xf(x)\,dx - \int_{0}^{\infty} |x|f(x)\,dx\right].$$

Clearly, if we take $A > 0$ and any negatively skewed distribution, $R_{L4} < 0$, and so BP1 is violated.

For R_{S1}, we consider axiom S2, which implies that

$$R[\tilde{y} + \lambda] = R[\tilde{y}] + S(\lambda),$$

where $S(\lambda)$ is monotonic. In order to be compatible with BP4, however, we would need $S(\lambda) = 0$, which is not monotonic.

Finally, for R_{S3}, suppose $C = -A = 1$ and $\theta = 2$. Consider $\tilde{y} \sim N(0,\ 1)$. Then

$$R_{S3}(\tilde{y}) = -\frac{1}{2}\mathrm{var}[|\tilde{y}|^2] < 0,$$

which violates BP1.

(e) Fishburn's axiom B4 implies that $R(\tilde{y}) = 0 \Leftrightarrow \Pr(\tilde{y} > 0) = 1$. Hence, if k is larger than the minimum value of $\tilde{y}$, this means that $R(\tilde{y} + (-k)) \neq 0$. But BP4 implies that $R(\tilde{y} + (-k)) = R(\tilde{y}) = 0$.

(f) **BP2** Note that

$$R_D(\lambda\tilde{y}) = \frac{\lambda^{1-e}(\mu_y - 1)}{\lambda^{1-e}\mu_y - 1}R(\tilde{y}).$$

So if, say, $e = -1$ and $\mu_y = 2$,

$$R_D(\lambda\tilde{y}) = \frac{\lambda^2}{2\lambda^2 - 1}R(\tilde{y})$$

and taking $\lambda = 2$ gives the desired contradiction.

(g) **BP2** Clearly,

$$\begin{aligned} R_{A_e}(\lambda\tilde{y}) &= 1 - [\lambda^{1-e} y^{1-e} f(\lambda y)\, d(\lambda y)]^{\frac{1}{1-e}} \\ &= 1 - \lambda[y^{1-e} f(y)\, d(y)]^{\frac{1}{1-e}} \\ &= 1 - \lambda R_{A_e}(\tilde{y}), \end{aligned}$$

so if, say, $R_{A_e}(\tilde{y}) = \frac{1}{2}$ and $\lambda = 2$, we get a contradiction.
(h), (i), (j), and (k) **BP2** Contradiction established as in (f) and (g).
(l) **BP2** Pick any $\lambda > 0$. Then $R_{ESC}(\lambda\tilde{y}) = F_{\lambda\tilde{y}}(\lambda\mu_y) = F(\mu_y) \neq |\lambda|\, R_{ESC}(\tilde{y})$.
(m) $G^{-1}(\frac{1}{2})$ is not defined if $\tilde{y}$ is a constant r.v.
(n) **BP2** Let $t = \mu_y$ and $\alpha > 1$. Then, for $\lambda > 1$,

$$R_{\alpha-t}(\lambda\tilde{y}) = \lambda^{\alpha} R(\tilde{y}) \neq |\lambda|\, R(\tilde{y}).$$

(o) **BP2** Pick any $\lambda > 0$. Then $R_{PL}(\lambda\tilde{y}) = \Pr(\lambda\tilde{y} < \lambda C) = \Pr(\tilde{y} < C) \neq |\lambda| R_{PL}(\tilde{y})$.
(p) **BP4**

$$\begin{aligned} \beta_{r,s}(\tilde{y}+\lambda) &= E((y+\lambda)\{F_{y+\lambda}(y+\lambda)\}^r\{1 - F_{y+\lambda}(y+\lambda)\}^s) \\ &= \beta_{r,s}(\tilde{y}) + \lambda E(\{F(y)\}^r\{1 - F(y)\}^s), \end{aligned}$$

so if, say, $\tilde{y} \sim N(0, 1)$,

$$\beta_{r,s}(\tilde{y}+\lambda) \neq \beta_{r,s}(\tilde{y}).$$

(q) **BP2** $\mathrm{Var}(\lambda\tilde{y}) = \lambda^2 \mathrm{Var}(\tilde{y}) \neq \lambda \mathrm{Var}(\lambda\tilde{y})$ for $\lambda \neq 1$.
(r) **BP2** Same as (q).
(s) **BP2** Let $c = 1$ and $f(y) = 1$ with probability $\frac{1}{2}$, 0 with probability $\frac{1}{2}$. Then

$$R_{B1}(\lambda\tilde{y}) = \log E\left[e^{-(\lambda y - \lambda\mu_y)}\right] = \log\left[e^{-\frac{\lambda}{2}} + e^{\frac{\lambda}{2}}\right],$$

but

$$|\lambda|\, R(\tilde{y}) = |\lambda| \log\left[e^{-\frac{1}{2}} + e^{\frac{1}{2}}\right].$$

These are clearly not always equal.

The other two measures of Bell—that is, R_{B2} and R_{B3}—follow in a similar fashion.
(t) **BP2** Here, suppose we consider a r.v. $\tilde{y}$ where $E[(y-\mu_y)^2] = E[(y-\mu_y)^3]$.
Then clearly $R_{JD}(\tilde{y}) = 0$ and

$$\begin{aligned} R_{JD}(\lambda\tilde{y}) &= \lambda^2 E[(y-\mu_y)^2] - \lambda^3 E[(y-\mu_y)^3] \\ &= (\lambda^2 - \lambda^3) E[(y-\mu_y)^2], \end{aligned}$$

which is nonzero for, say, $\lambda = 2$. □

Proof of Theorem 9

(a) By Lemma 7, we know that BP1, BP2, and BP4 hold in both cases. We thus need to show BP3 is satisfied.
Consider first the case where $k > 1$, $Y_0 = \mu_y$, and $A = \mu_y$. Note that then

$$R_S(\tilde{y}) = \left(\int_{-\infty}^{\mu_y} |y - \mu_y|^k f(y)\, dy \right)^{1/k}$$
$$= \left(\int \max[0, |y - \mu_y|]^k f(y)\, dy \right)^{1/k}.$$

Hence,

$$R_S(\tilde{y} + \tilde{x}) = \left(\int\int \max[0, |y + x - \mu_y\, \mu_x|]^k f(y,\, x)\, dydx \right)^{1/k}$$
$$\left(\int\int \max[0, |(y - \mu_y) + (x - \mu_x)|]^k f(y, x)\, dydx \right)^{1/k}$$
$$\leq \left(\int\int \max[0, |y - \mu_y|]^k f(y, x)\, dydx \right.$$
$$\left. + \int\int \max[0, |x - \mu_x|]^k f(y, x)\, dydx \right)^{1/k}$$

by the triangle inequality and the fact that $k > 0$. Applying Minkowski's inequality (see Rudin [1976]), implies that this quantity is less than or equal to

$$\left(\int\int \max[0, |y - \mu_y|]^k f(y, x)\, dydx \right)^{1/k}$$
$$+ \left(\int\int \max[0, |x - \mu_x|]^k f(y, x)\, dydx \right)^{1/k}.$$

Integrating out the marginal distributions, this is equal to

$$\left(\int \max[0, |y - \mu_y|]^k f(y)\, dy \right)^{1/k} + \left(\int \max[0, |x - \mu_x|]^k f(x)\, dx \right)^{1/k}$$
$$= R_S(\tilde{y}) + R_S(\tilde{x}).$$

The case for which $k > 1$, $Y_0 = \mu_y$, and $A = \infty$ can be proved similarly and in any case follows from the work of Kijima and Ohnishi [1993].

(b) For the range measure, we demonstrate that all four properties are satisfied:

- BP1 Clearly, $y_{\max} - y_{\min} \geq 0$.
- BP2

$$\begin{aligned} R_R(\lambda\tilde{y}) &= (\lambda\tilde{y})_{\max} - (\lambda\tilde{y})_{\min} \\ &= \begin{cases} \lambda(y_{\max} - y_{\min}) & \text{if } \lambda > 0 \\ \lambda(y_{\min} - y_{\max}) & \text{if } \lambda \leq 0 \end{cases} \\ &= |\lambda|(y_{\max} - y_{\min}) = |\lambda| R_R(\tilde{y}). \end{aligned}$$

- BP3

$$\begin{aligned} R_R(\tilde{y} + \tilde{x}) &= (x + y)_{\max} - (x + y)_{\min} \\ &\leq x_{\max} + y_{\max} - x_{\min} - y_{\max} \end{aligned}$$

by applying the triangle inequality twice. This is in turn equal to

$$\begin{aligned} &(x_{\max} - x_{\min}) + (y_{\max} - y_{\min}) \\ &\quad = R_R(\tilde{y}) + R_R(\tilde{x}). \end{aligned}$$

- BP4

$$\begin{aligned} R_R(\tilde{y} + \lambda) &= (y + \lambda)_{\max} - (y + \lambda)_{\min} \\ &= y_{\max} - y_{\max} = R_R(\tilde{y}). \end{aligned}$$

(c) First note that the linear measures are in fact constant multiples of the Gini coefficient. To see this, note that the linear measure is

$$cov[y - \mu_y,\ aF(y) + b],$$

where a and b are constants and $a > 0$. This simplifies to

$$acov[y,\ F(y)] = \frac{a}{2} R_G(\tilde{y}).$$

Hence, if the Gini coefficient satisfies the basic properties, so do the linear measures. Since the Gini belongs to our new class (44) (in fact, it is $R[\infty,\ \mu_y,\ 1,\ 1,\ 2F(y)]$), Lemma 7 applies. Consequently, BP1, BP2, and BP4 are all satisfied by the Gini coefficient. We thus show

$$\begin{aligned} &\text{BP3} R_G(\tilde{x} + \tilde{y}) \\ &= \int\!\!\int\!\!\int\!\!\int |(x_i + y_i) - (x_j + y_j)|\, f(x_i,\ x_j,\ y_i,\ y_j)\, dx_i dx_j dy_i dy_j \\ &\leq \int\!\!\int\!\!\int\!\!\int |x_i - x_j|\, f(x_i,\ x_j,\ y_i,\ y_j)\, dx_i dx_j dy_i dy_j \\ &\quad + \int\!\!\int\!\!\int\!\!\int |y_i - y_j|\, f(x_i,\ x_j,\ y_i,\ y_j)\, dx_i dx_j dy_i \end{aligned}$$

by the triangle inequality. Integrating out the marginals gives

$$\int\int |x_i - x_j| f(x_i,\ x_j)\, dx_i dx_j + \int\int |y_i - y_j| f(y_i,\ y_j)\, dy_i dy_j$$
$$= R_G(\tilde{y}) + R_G(\tilde{x})$$

Hence, the Gini and so the linear measures satisfy all properties.

To generalize this to the piecewise-linear measures, if one assumes that the $W[F(y)]$ is invertible, one may use the fact that the sum of convex functions is convex and refer to the linear case just proved.

(d) This result is given in Kijima and Ohnishi [1993]. □

Acknowledgments

The authors wish to thank Babak Eftekari, Professor Christian Gollier, and two anonymous referees for helpful comments. The first author has been supported by the Economic and Social Research Council, while the second author has been supported by Inquire and the Newton Trust.

Note

1. This can be extended to the piecewise linear measures (see (19)) simply by making $W[F(y)]$ piecewise linear in $F(y)$.

References

ACZEL, J. [1966]: *Lectures on Functional Equations and Their Applications*, Academic Press, New York.
ATKINSON, A. [1970]: "On the Measurement of Inequality," *Journal of Economic Theory*, 2, 244–257.
Basle Committee on Banking Supervision. [1993]: *Measurement of Banks Exposure to Interest Rate Risk.* Committee on Banking Supervision, Basle.
BAWA, V. [1978]: "Safety First, Stochastic Dominance and Optimal Portfolio Choice," *Journal of Financial and Quantitative Analysis*, 13(2) (June), 255–271.
BAWA, V., BODURTHA, J., RAO, M., and SURI, H. [1985]: "On Determination of Stochastic Dominance Optimal Sets," *Journal of Finance*, 40(2), 417–431.
BAWA, V., and LINDENBERG, E. [1977]: "Capital Market Equilibrium in a Mean-Lower Partial Moment Framework," *Journal of Financial Economics*, 5, 189–200.
BELL, D. [1988]: "One-Switch Utility Functions and a Measure of Risk," *Management Science*, 34(12) (December), 1416–1424.
BELL, D. [1995]: "Risk Return and Utility," *Management Science*, 41(1) (January), 23–30.
COOMBS, C., and HUANG, L. [1970]: "Tests of Portfolio Theory of Risk Preference," *Journal of Experimental Psychology*, 85(6), 23–29.
COWELL, F. [1995]: *Measuring Inequality*, Prentice-Hall/Harvester Wheatleaf, London.
DALTON, H. [1920]: "The measurement of the Inequality of Income," *Economic Journal*, 30, 348–361.
FISHBURN, P. [1977]: "Mean-Risk Analysis with Risk Associated with Below-Target Returns," *American Economic Review*, 67(2) (March), 116–126.
FISHBURN, P. [1980]: "Foundations of Risk Measurement: I. Risk as a Probability of Loss," *Management Science*, 30, 396–406.

FISHBURN, P. [1981]: "Foundations of Risk Measurement: I. Effects of Gains on Risk," *Journal of Mathematical Psychology*, 25, 226–242.

GREENWOOD, J., LANDWEHR, J., MATALAS, N., and WALLIS, J. [1979]: "Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressible in Inverse Form," *Water Resources Research*, 15, 1049–1054.

HARLOW, W., and RAO, R. [1989]: "Asset Pricing in a Generalised Mean-Lower Partial Moment Framework: Theory and Evidence," *Journal of Financial and Quantitative Analysis*, 24(3), 285–311.

HOLTHAUSEN, D. [1981]: "A Risk-Return Model with Risk and Return Measured as Deviations from a Target Return," *American Economic Review*, 71(1), 182–188.

HOSKING, J. [1986]: "The Theory of Probability Weighted Moments." Technical report, IBM Research Report, RC12210 (No. 54860), Thomas J. Watson Research Center, New York, 1986.

HUANG, C., and LITZENBERGER, R. [1988]: *Foundations for Financial Economics*. Elsevier Science, London.

JIA, J., and DYER, J.S. [1996]: "A standard Measure of Risk and Risk-Value Models," *Management Science*, 42(12) (December), 1691–1705.

KIJIMA, M., and OHNISHI, M. [1993]: "Mean-Risk Analysis of Risk Aversion and Wealth Effects on Optimal Portfolios with Multiple Investment Opportunities," *Annals of Operations Research*, 45, 147–163.

KNIGHT, J., SATCHELL, S., and TRAN, K. [1995]: "Statistical Modelling of Assymetric Risk in Asset Returns," *Applied Mathematical Finance*, 2, 155–172.

KRANTZ, D., LUCE, R., SUPPES, P., and TVERSKY, A. [1971]: *Foundations of Measurement I.* Academic Press, New York.

LUCE, R. [1980]: "Several Possible Measures of Risk," *Theory and Decision*, 12, 217–228.

LUCE, R. [1981]: "Correction to Several Possible Measures of Risk," *Theory and Decision*, 13, 381.

LUCE, R., and WEBER, E. [1986]: "An Axiomatic Theory of Conjoint Expected Risk," *Journal of Mathematical Psychology*, 30, 188–205.

MACHINA, M. [1982]: "Expected Utility Analysis Without the Independence Axiom," *Econometrica*, 50(2), 277–323.

MEHRAN, F. [1976]: "Linear Measures of Income Inequality," *Econometrica*, 44(4), 805–809.

NEWBURY, D. [1970]: "A Theorem on the Measurement of Inequality," *Journal of Economic Theory*, 2, 264–266.

POLLATSEK, A., and TVERSKY, A. [1970]: "A Theory of Risk," *Journal of Mathematical Psychology*, 7, 540–553.

PYLKE, D.H., and TURNOVSKY, S.J. [1970]: "Safety-First and Expected Utility Maximisation in Mean-Standard Deviation Portfolio Analysis," *Review of Economics and Statistics*, 52(1) (February), 75–81.

ROSS, S. [1976]: "The Arbitrage Pricing Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341–360.

ROTHSCHILD, M., and STIGLITZ, J. [1970]: "Increasing Risk: I. A definition," *Journal of Economic Theory*, 2(3), 225–243.

ROY, A. [1952]: "Safety First and the Holding of Risky Assets," *Econometrica*, 20(3), 431–449.

RUDIN, W. [1976]: *Principles of Mathematical Analysis*, 3rd. ed. McGraw-Hill, New York.

SARIN, R. [1987]: "Some Extensions of Luce's Measure of Risk," *Theory and Decision*, 22, 125–141.

SARIN, R.K., and WEBER, M. [1993]: "Risk-Value Models," *European Journal of Operational Research*, 70, 135–149.

SEN, A. [1973]: *On Economic Inequality*. Clarendon Press, Oxford.

SHALIT, H., and YITZHAKI, S. [1984]: "Mean-Gini, Portfolio Theory and the Pricing of Risky Assets," *Journal of Finance*, 29(5), 1449–1468.

SHARPE, W. [1964]: "Capital Asset Prices: A Theory of Capital Market Equilibrium Under Conditions of Risk," *Journal of Finance*, 19, 425–442.

SINCLAIR, C., POWER, D., and LONIE, A. [1995]: "Risk Triangles and Other Novel Measures of Risk and Return for Stocks or Portfolios." *Manuscript*.

STONE, B. [1973]: "A General Class of Three-Parameter Risk Measures," *Journal of Finance*, 28, 675–685.

THEIL, H. [1967]: *Economics and Information Theory*, North-Holland, Amsterdam.

YAARI, M. [1987]: "The Dual Theory of Choice Under Risk," *Econometrica*, 55(1), 95–115.

YITZHAKI, S. [1983]: "On an Extension of the Gini Inequality Index," *International Economic Review*, 24(3), 617–628.