

Generalized Systematic Risk

Ohad Kadan Fang Liu Suying Liu*

July 2015

Abstract

We generalize the concept of “systematic risk” to a broad class of risk measures potentially accounting for high distribution moments, downside risk, rare disasters, as well as other risk attributes. We offer two different approaches. First is an equilibrium framework generalizing the Capital Asset Pricing Model, two-fund separation, and the security market line. Second is an axiomatic approach resulting in a systematic risk measure as the unique solution to a risk allocation problem. Both approaches lead to similar results extending the traditional beta to capture multiple dimensions of risk. The results lend themselves naturally to empirical investigation.

Risk is a complex concept. The definition of risk and its implications have long been the subject of both academic and practical debate. This issue has gained even more prominence during the recent financial crisis, when markets and individual assets were hit by catastrophic events whose ex-ante probabilities were considered negligible. Indeed, these events demonstrate that “risk” accounts for much more than what is measured by the variance of the returns of an asset. High distribution moments, rare disasters, and downside risk are just some of the different aspects that may be of interest when measuring risk.

*Ohad Kadan: Olin Business School, Washington University in St. Louis, Campus Box 1133, One Brookings Dr., St. Louis, MO 63130 (e-mail: kadan@wustl.edu); Fang Liu: School of Hotel Administration, Cornell University, 435C Statler Hall, Ithaca, NY 14853 (e-mail: fl357@cornell.edu); Suying Liu: Olin Business School, Washington University in St. Louis, Campus Box 1133, One Brookings Dr., St. Louis, MO 63130 (e-mail: sliua@wustl.edu). We thank Jason Donaldson, Phil Dybvig, Brett Green, Sergiu Hart, Steve Ross, and the referees as well as seminar participants at UC Berkeley, Cornell University, Hebrew University, Hong Kong Polytechnic University, Indiana University, University of Pennsylvania, and Washington University in St. Louis for helpful comments and suggestions.

In this paper we allow “risk” to take a very general form. We then revisit the classic notion of “systematic risk,” which reflects the contribution of an asset to the risk of a portfolio. Traditional measures of systematic risk focus on a narrow set of risk attributes. In particular, the most well-known and widely used measure of systematic risk is the beta of the asset, which is the slope from regressing the asset returns on portfolio returns (Sharpe (1964), Lintner (1965a, b), and Mossin (1966)). Beta is the contribution of an asset to the risk of the portfolio as measured by the variance of its return. It sets the foundations for all risk-return analysis as part of the Capital Asset Pricing Model (CAPM). However, the traditional beta ignores all aspects of risk other than the variance, such as high distribution moments and rare disasters.

We offer two different approaches to generalizing systematic risk. First we study an equilibrium framework modifying the traditional CAPM to allow for a broad set of risk attributes. The equilibrium approach allows us to extend classic results such as the geometry of efficient portfolios, the two-fund separation theorem, the efficiency of the market portfolio, and the security market line. Second is an axiomatic approach in which we recast the issue as a risk allocation problem. We then specify desirable properties of systematic risk, leading to a unique solution. Both approaches yield similar results, generalizing the traditional beta to reflect a variety of risk attributes.

We begin with a broad definition of what would constitute a measure of risk. We define a risk measure as any mapping from random variables to real numbers. That is, a risk measure is simply a summary statistic that encapsulates the randomness using just one number. The variance (or standard deviation) is obviously the most commonly used risk measure. However, many other risk measures have been proposed and used. For example, high distribution moments can account for skewness and tail risk, downside risk accounts for the variation in losses, and value at risk is a popular measure of disaster risk. Recently, Aumann and Serrano (2008) and Foster and Hart (2009) offered two appealing risk measures that account for all distribution moments and for disaster risk.¹ All of these measures fall under our wide umbrella

¹See Hart (2011) for a unified treatment of these two measures and Kadan and Liu (2014) for an analysis of the moment properties of these measures.

of risk measures. Moreover, any linear combination of risk measures is itself a risk measure. Thus, one can easily create measures of risk that account for a number of dimensions of riskiness, assigning the required weight to each dimension.

Our first analysis generalizes the classic CAPM to allow for a broad set of risk measures. The idea is simple. In the classic CAPM setting investors are assumed to have mean-variance preferences. That is, their utility is increasing in the expected payoff and decreasing in the variance of their payoffs. In our generalized setting we assume that investors have mean-risk preferences, where the term “risk” stands for a host of potential risk measures. We provide mild sufficient conditions under which these preferences are locally consistent with expected utility in the sense of Machina (1982).

We consider an exchange economy with a finite number of risky assets, one risk-free asset, and a finite number of investors with mean-risk preferences. As usual, in equilibrium each investor chooses a portfolio of assets from the set of efficient portfolios, minimizing risk for a given expected return. However, due to the generality of the risk measure, the geometry of this set is more complicated than in the case where risk is measured by the variance. Nevertheless, we establish sufficient conditions on the risk measure under which the solution to each investor’s problem satisfies Tobin’s (1958) two-fund separation property. That is, each investor’s optimal portfolio of assets can be presented as a linear combination of the risk-free asset and a unique portfolio of risky assets. We demonstrate that a variety of risk measures satisfy these sufficient conditions, where the variance is just one special case. A consequence of two-fund separation is that the equilibrium market portfolio lies on the efficient frontier. Using this we establish a generalization of the classic security market line (SML) to a large class of risk measures. Specifically, in equilibrium, the expected return of each risky asset i satisfies

$$E(\tilde{z}_i) = r_f + \mathcal{B}_i^R (E(\tilde{z}^M) - r_f),$$

where \tilde{z}_i is the risky return of asset i , \tilde{z}^M is the risky return of the market portfolio, r_f is the risk-free rate, and \mathcal{B}_i^R is the systematic risk of asset i given the risk measure R . Moreover, \mathcal{B}_i^R is given in closed form as the marginal contribution of asset i to the market risk scaled by the weighted average of such marginal contributions

across all assets in the economy.

In the special case in which R is the variance, \mathcal{B}_i^R coincides with the traditional beta. More generally, we show that our equilibrium setting is versatile enough to allow for a variety of risk attributes such as tail risk, downside risk, and rare disasters, among others. Our setting can also readily account for risk measures that reflect several of these risk attributes, assigning different weights to each of them. We illustrate that in all these cases one can readily derive closed form solutions for the generalized betas. Typically, these betas reflect the covariation of the return of asset i with some function of the market return. In general, these betas do not take the form of a regression coefficient. Nevertheless, they can be estimated directly from return data and applied in a standard Fama-MacBeth (1973) cross-sectional analysis.

The CAPM equilibrium can be thought of as a special case of the more general problem of risk allocation. Indeed, the CAPM beta measures the contribution of one asset to the risk of the market portfolio. Many other problems of considerable economic import require estimating the contribution of one asset to some specific portfolio of assets (not necessarily the market portfolio). For example, the government is constantly interested in the contribution of particular banks and other financial institutions to the total market risk (known as systemic risk). Banks and other financial institutions may also find it useful to calculate the contribution of different assets on their balance sheet to the total risk of the institution, so that each asset or business unit could be “taxed” appropriately. All of these problems are essentially risk allocation problems in which total risk should be allocated among the constituents of a portfolio. We broaden the term “systematic risk” to designate solutions to such problems. That is, a systematic risk measure is a vector specifying the portion of the total portfolio risk allocated to each asset in the portfolio. The literature has not yet presented a general solution to this problem for a broad set of risk measures and for arbitrary portfolios.

In the second part of this paper we tackle this problem from an axiomatic point of view. We state desirable properties of systematic risk measures, which we call axioms, and we look for solutions that satisfy these properties. Unlike in the equilibrium setting, here we do not need to impose almost any structure on the risk

measure. Moreover, the portfolio of assets is arbitrary and is not limited to the market portfolio.

We state four economically plausible axioms that systematic risk measures are expected to satisfy. We then show that these four axioms imply a unique systematic risk measure which applies to all risk allocation problems. This measure is given by a scaled version of the Aumann-Shapley (1974) diagonal formula, which was developed as a solution concept in cooperative game theory. Essentially, this formula calculates for each asset the average of its marginal contributions to portfolios along a diagonal starting from the origin and ending at the portfolio of interest. In the common case in which the risk measure is homogeneous of some degree, the solution becomes very simple, and it coincides with the generalized beta obtained in the equilibrium setting above. In particular, it assigns to each asset its marginal contribution to total portfolio risk scaled by the weighted average of marginal contributions of all assets. Our proof of the axiomatization result relies on a mapping between risk allocation problems and cost allocation problems studied in Billera and Heath (1982).

The paper proceeds as follows. Section I discusses the related literature. In Section II we define the notion of risk measures. Section III studies the equilibrium setup and offers a generalization of the CAPM. In Section IV we present the axiomatic approach. Section V concludes. Proofs of the main theorems are in Appendix A, proofs of propositions and other derivations are in Appendix B, and other technical results are provided in an Online Appendix.

I. Related Literature

Our paper contributes to several strands of the literature. First, the paper adds to the growing literature on high distribution moments, disaster risk, and other risk attributes, as well as their effect on prices. Rubinstein (1973), Kraus and Litzenberger (1976), Jean (1971), Kane (1982), and Harvey and Siddique (2000) argue that investors favor right-skewness of returns, and demonstrate the cross-sectional implications of this effect. In addition, Barro (2006, 2009), Gabaix (2008, 2012), Gourio (2012), Chen, Joslin, and Tran (2012), and Wachter (2013) study the aver-

sion of investors to tail risk and rare disasters. Ang, Chen, and Xing (2006) and Lettau, Maggiori, and Weber (2014) show that downside risk is a good explanatory variable for returns in both equity and currency markets. Our paper adds to this literature by outlining a general approach to measuring systematic risk that can capture the contribution of an asset to a range of risk dimensions such as high distribution moments, downside risk, and rare disasters. Our framework is flexible and can account for either one risk aspect or a combination of several of them.

Our equilibrium approach follows a reduced form, where preferences are described through the aversion to broadly defined risk. Our main results are derived without the need to specify an exact form of the utility function. This is different from the approach in consumption-based asset pricing models (e.g., Bansal and Yaron (2004) and Campbell and Cochrane (1999)). These models rely on the specification of a particular utility function (such as Epstein and Zin (1989) preferences or preferences reflecting past habits). One advantage of our approach is that it provides a parsimonious and simple one-factor model that can capture different aspects of risk in a manner that may lend itself naturally to empirical investigation. Another feature of our approach is that, unlike consumption-based models, it resorts to prices directly. Thus, one can potentially test our model without relying on consumption data.

The paper also adds to the growing literature on risk measurement. This literature dates back to Hadar and Russell (1969), Hanoch and Levy (1969), and Rothschild and Stiglitz (1970) who extend the notion of riskiness beyond the “variance” framework by introducing stochastic dominance rules. Artzner et al. (1999) specify desirable properties of coherent risk measures, and Rockafellar, Uryasev, and Zabarankin (2006a) introduce the notion of generalized deviation measures. More recently, Aumann and Serrano (2008), Foster and Hart (2009, 2013), and Hart (2011) have come up with appealing risk measures that generalize conventional stochastic dominance rules. Notably, all the risk measures discussed in this literature are idiosyncratic in nature. Our paper contributes to this literature by specifying a method to calculate the systematic risk of an asset for any given risk measure. This in turn allows us to study the fundamental risk-return trade-off associated with a risk measure.

Our paper also adds to the recent literature on systemic risk, which is the risk that the entire economic system collapses. Adrian and Brunnermeier (2014) define the $\Delta CoVaR$ measure as the difference between the value at risk of the banking system conditional on the distress of a particular bank and the value at risk of the banking system given that the bank is solvent. Acharya et al. (2010) propose the Systemic Expected Shortfall measure, which estimates the exposure of a particular bank in terms of undercapitalization to a systemic crisis. Huang, Zhou, and Zhu (2009) measure the systemic risk of a financial institution by the price of insurance against financial distress. Our paper takes a general approach to the problem of estimating the contribution of one asset to the risk of a portfolio of assets. We provide an easy-to-calculate and intuitive measure that applies to a wide variety of risk measures, as well as in an array of contexts.

Our paper also contributes to the literature studying conditions for two-fund separation. The idea of two-fund separation was introduced by Tobin (1958). Since then the literature discussed different sufficient conditions in terms of either agents' utility (e.g., Cass and Stiglitz (1970) and Dybvig and Liu (2015)) or the distribution of returns (e.g., Ross (1978)). Here we take a somewhat different approach, as we specify sufficient conditions for two-fund separation in terms of properties of the risk measure. This approach is similar to the one taken in Rockafellar, Uryasev, and Zabarankin (2006b), who consider general deviation measures. Our restrictions on risk measures are weaker than theirs as we do not require homogeneity. All of these papers consider two-fund separation only and do not provide any generalization of the notion of systematic risk, which is the focus of our paper.

Finally, the paper adds to an extensive list of studies applying the Aumann-Shapley solution concept in different contexts, e.g., Billera, Heath, and Raanan (1978), Samet, Tauman, and Zang (1984), Powers (2007), and Billera, Heath, and Verrecchia (1981). Tarashev, Borio, and Tsatsaronis (2010) use the Shapley value (Shapley (1952), a discrete version of the Aumann-Shapley solution concept) to measure systemic risk. Our paper offers theoretical foundations for their practical approach.

II. Risk Measures and Their Properties

Let (Ω, \mathcal{F}, P) be a probability space, where Ω is the state space, \mathcal{F} is the σ -algebra of events, and $P(\cdot)$ is a probability measure. As usual, a random variable is a measurable function from Ω to the reals. In the context of investments, we typically consider random variables representing the payoffs or the returns of financial assets. Thus, we often refer to random variables as “investments” or “random returns.” We generically denote random variables by \tilde{z} , which is a shorthand notation for $\tilde{z}(\omega)$, $\forall \omega \in \Omega$. We restrict attention to random variables for which all moments exist. We denote the expected value of \tilde{z} by $E(\tilde{z})$ and its k^{th} central moment by $m_k(\tilde{z}) = E(\tilde{z} - E(\tilde{z}))^k$, where $k \geq 2$.

A risk measure is simply a function that assigns to each random variable a single number summarizing its riskiness. Formally,

Definition 1 *A risk measure is a function mapping random variables to the reals.*²

We generically denote risk measures by $R(\cdot)$. The simplest and most commonly used risk measure is the variance ($R(\tilde{z}) = m_2(\tilde{z})$). However, many other risk measures have been proposed in the literature, capturing higher distribution moments and other risk attributes. A risk measure $R(\cdot)$ is *homogeneous of degree k* , if for any random return \tilde{z} and positive number $\lambda > 0$,

$$R(\lambda\tilde{z}) = \lambda^k R(\tilde{z}).$$

A weaker requirement, which is sufficient for most of our results, is that the risk ranking between two investments does not depend on scaling. We say that $R(\cdot)$ is *scaling independent* if for all $\lambda > 0$ and any two random returns \tilde{z}_1 and \tilde{z}_2 , $R(\tilde{z}_1) > R(\tilde{z}_2)$ implies $R(\lambda\tilde{z}_1) > R(\lambda\tilde{z}_2)$.

The next property of risk measures which will become useful is convexity. Formally, we say that a risk measure $R(\cdot)$ is *convex* if for any two random returns \tilde{z}_1

²Strictly speaking, a risk measure is also a function of the underlying probability measure P . However, in our analysis we fix P throughout, and yet consider different random variables. Thus, it is convenient to think about risk measures as functions of the random variables, viewing the probability measure as a fixed parameter.

and \tilde{z}_2 , and for any $\lambda \in (0, 1)$, we have

$$R(\lambda\tilde{z}_1 + (1 - \lambda)\tilde{z}_2) \leq \lambda R(\tilde{z}_1) + (1 - \lambda)R(\tilde{z}_2),$$

with equality holding only when $\tilde{z}_1 = \tilde{z}_2$ with probability 1. Notice that $\lambda\tilde{z}_1 + (1 - \lambda)\tilde{z}_2$ can be considered as the return of a portfolio that assigns weights λ and $1 - \lambda$ to \tilde{z}_1 and \tilde{z}_2 , respectively. Then the convexity condition says that the risk of the portfolio should not be higher than the corresponding weighted average risk of the constituent investments. Thus, convexity of a risk measure captures the idea that diversifying among two investments lowers the total risk.

Next we would like to formalize a property dealing with the type of assets that are risk-free. We say that a risk measure $R(\cdot)$ has the *risk-free property*, if (i) $R(\tilde{z}) \geq 0$ for all \tilde{z} ; (ii) $R(\tilde{z}) = 0$ if and only if $P(\{\tilde{z} = c\}) = 1$ for some constant c ; and (iii) $R(\tilde{z}_1 + \tilde{z}_2) = R(\tilde{z}_1)$ whenever $R(\tilde{z}_2) = 0$. Namely, R has the risk-free property if the only assets with zero risk are those that pay a constant amount with probability 1, if all other assets have strictly positive risk, and if adding a zero-risk asset does not change risk. In what follows, we often refer to assets satisfying $R(\tilde{z}) = 0$ as risk-free.

Risk measures can be applied to individual random variables or to portfolios of random variables. Formally, assume there are n random variables represented by the vector $\tilde{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_n)$. A *portfolio* is a vector $\mathbf{x} = (x_1, \dots, x_n) \in \mathcal{R}^n$, where x_i is the dollar amount invested in \tilde{z}_i .³ Then, $\mathbf{x} \cdot \tilde{\mathbf{z}} = \sum_{i=1}^n x_i \tilde{z}_i$ is itself a random variable. We then say that the risk of portfolio \mathbf{x} is simply $R(\mathbf{x} \cdot \tilde{\mathbf{z}})$. When the vector of random variables is unambiguous, we often abuse notation and denote $R(\mathbf{x})$ as a shorthand for $R(\mathbf{x} \cdot \tilde{\mathbf{z}})$. We say that a risk measure is *smooth* if for any vector of random returns $\tilde{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_n)$ and for all portfolios $\mathbf{x} = (x_1, \dots, x_n)$ we have that $R(\mathbf{x} \cdot \tilde{\mathbf{z}})$ is continuously differentiable in x_i for $i = 1, \dots, n$. We then write $R_i(\mathbf{x})$ (or $R_i(\mathbf{x} \cdot \tilde{\mathbf{z}})$) for the partial derivative of $R(\cdot)$ with respect to the amount invested in the i^{th} asset evaluated at \mathbf{x} .⁴

³Throughout the paper we denote vectors using bold notation (for both numbers and random variables).

⁴Note that we use subscripts to denote both elements of a vector and partial derivatives. For example, x_i is the i^{th} element of the vector \mathbf{x} while $R_i(\cdot)$ is the partial derivative of $R(\cdot)$ considered as a function of portfolio amounts. This notation does not result in any ambiguity since the only case

When restricting attention to homogeneous risk measures, the properties discussed above are maintained when taking convex combinations of different risk measures with the same degree of homogeneity. Thus, we can easily create new risk measures satisfying these properties from existing homogeneous risk measures. That is, let s be a positive integer, let $R^1(\cdot), \dots, R^s(\cdot)$ be risk measures, and choose $\theta = (\theta_1, \dots, \theta_s)$ with $\theta_j > 0 \forall j$. We can then define a new risk measure by

$$R^\theta(\tilde{z}) = \sum_{j=1}^s \theta_j R^j(\tilde{z}),$$

where θ_j reflects the weight assigned to the risk attribute measured by R^j . We then have the following trivial but useful lemma.

Lemma 1 *Assume that each R^j is homogeneous of degree k , convex, smooth, and satisfies the risk-free property. Then, R^θ also satisfies all of these properties.*

A. Examples of Risk Measures

Below we present some popular examples of risk measures and discuss their properties. Each of these examples highlights a different aspect of risk that may be of interest in applications. These examples will be crucial later in the paper when we demonstrate how to apply our main results.

Example 1 *Even central moments and normalized even central moments.* For any integer $k \geq 2$ even, the central moment $R(\tilde{z}) = m_k(\tilde{z})$ is a risk measure which is homogeneous of degree k , convex, smooth and satisfies the risk-free property. The normalized central moment $w_k(\tilde{z}) = (m_k(\tilde{z}))^{1/k}$ is also a risk measure. For example, when $k = 2$, $w_k(\tilde{z})$ is the standard deviation of \tilde{z} . Normalized central moments satisfy all of the above properties as well (with homogeneity of degree 1). Indeed, homogeneity, smoothness, and the risk-free property are trivial in these cases. Convexity stems from the following result, which shows that $w_k(\tilde{z})$ is convex, and thus $m_k(\tilde{z})$ is a fortiori convex.

in which the subscript should be interpreted as a partial derivative is when applied to the risk measure considered as a function of portfolio amounts.

Proposition 1 For all $k \geq 2$ even, $R(\tilde{z}) = w_k(\tilde{z})$ is a convex risk measure.

Example 2 Odd central moments and normalized odd central moments. For any integer $k \geq 3$ odd, the central moment $R(\tilde{z}) = m_k(\tilde{z})$ is a risk measure which is homogeneous of degree k and smooth. Similarly, the normalized odd moments $w_k(\tilde{z})$ are homogeneous of degree 1 and smooth. In contrast to the even central moments, neither convexity nor the risk-free property holds in this case.⁵

Evidently, the feature of odd central moments that prevents them from satisfying convexity and the risk-free property is that they admit negative values. A natural way to fix this is to focus on just one side of the distribution. The next example follows this idea, allowing one to readily incorporate aspects of odd central moments (such as skewness) into risk measures that also satisfy convexity and the risk-free property.

Example 3 Downside risk. When considering risk, investors sometimes restrict attention to the lower outcomes of the distribution, in particular to those which fall below the mean. Such an approach is called downside risk. Formally, for any integer $k \geq 2$, define the downside risk of order k of \tilde{z} as

$$\text{DR}_k(\tilde{z}) = (-1)^k \left(\mathbb{E}([\tilde{z} - \mathbb{E}(\tilde{z})]^-)^k \right)^{\frac{1}{k}},$$

where $[t]^- = \min(t, 0)$ for $t \in \mathcal{R}$. Often, this measure is used in the special case of $k = 2$. More generally, for any $k \geq 2$, $\text{DR}_k(\tilde{z})$ is a risk measure which is homogeneous of degree 1, smooth, and satisfies the risk-free property. The next proposition establishes that this risk measure is also convex.

Proposition 2 For any $k \geq 2$, $\text{DR}_k(\tilde{z})$ is a convex risk measure.

⁵To see the former, consider the simple example of two random returns, \tilde{z}_1 and \tilde{z}_2 , which are independent and have negative third central moments $m_3(\cdot)$. Then, by independence and the homogeneity of central moments,

$$m_3\left(\frac{1}{2}\tilde{z}_1 + \frac{1}{2}\tilde{z}_2\right) = \left(\frac{1}{2}\right)^3 m_3(\tilde{z}_1) + \left(\frac{1}{2}\right)^3 m_3(\tilde{z}_2) > \frac{1}{2}m_3(\tilde{z}_1) + \frac{1}{2}m_3(\tilde{z}_2),$$

since $m_3(\tilde{z}_1) + m_3(\tilde{z}_2) < 0$. To see the latter, note that the third moment can be negative, violating the risk-free property.

Example 4 Value at risk. A risk measure widely used in financial risk management is the Value at Risk (VaR), designed to capture the risk associated with rare disasters. VaR measures the amount of loss not exceeded with a certain confidence level. Formally, given some confidence level $\delta \in (0, 1)$, for any random return \tilde{z} , the VaR measure is defined as the negative of the δ -quantile of \tilde{z} , i.e.,

$$(1) \quad \text{VaR}_\delta(\tilde{z}) = -\inf \{z \in \mathcal{R} : F(z) \geq \delta\},$$

where $F(\cdot)$ is the cumulative distribution function of \tilde{z} . Notice that we include the minus sign to reflect the fact that a larger loss indicates higher risk. If \tilde{z} is continuously distributed with a density function $f(\cdot)$, then (1) is implicitly determined by

$$(2) \quad \int_{-\infty}^{-\text{VaR}_\delta(\tilde{z})} f(z) dz = \delta.$$

This risk measure is homogeneous of degree 1 and smooth.⁶ For any risk-free return \tilde{z} with $P(\{\tilde{z} = c\}) = 1$, we have $\text{VaR}_\delta(\tilde{z}) = -c$, implying that the VaR of risk-free assets depends on the risk-free return. Hence, the risk-free property is not satisfied. In addition, it is not hard to find examples where convexity is violated for the VaR measure.

Example 5 Expected shortfall and demeaned expected shortfall.⁷ These measures capture the average loss from disastrous events, defined as those involving a loss larger than the VaR. Formally, assume that \tilde{z} can be represented by a density $f(\cdot)$. Given some confidence level $\delta \in (0, 1)$, for any random return \tilde{z} the Expected Shortfall (ES) is the negative of the conditional expected value of \tilde{z} below the δ -quantile. That is,

$$(3) \quad \text{ES}_\delta(\tilde{z}) = -\frac{1}{\delta} \int_{-\infty}^{-\text{VaR}_\delta(\tilde{z})} z f(z) dz.$$

Additionally, when $\tilde{z} = c$ (a constant) with probability 1 we set $\text{ES}_\delta(\tilde{z}) = -c$. Similar to VaR, ES is homogeneous of degree 1 and is smooth, but it does not satisfy

⁶Formally, smoothness follows if a joint density of the random returns in a portfolio exists. This is shown using an application of the implicit function theorem to (2). We omit the proof for brevity.

⁷Expected shortfall is sometimes termed “conditional VaR.”

the risk-free property. To ensure that the risk-free property is satisfied it is useful to consider the demeaned version of ES defined as

$$\text{DES}_\delta(\tilde{z}) = -\frac{1}{\delta} \int_{-\infty}^{-\text{VaR}_\delta(\tilde{z})} (z - \text{E}(\tilde{z})) f(z) dz = \text{ES}_\delta(\tilde{z}) + \text{E}(\tilde{z}).$$

Similar to ES, DES also captures the expected loss from a rare disaster. This risk measure is also homogeneous of degree 1, smooth, and it satisfies the risk-free property.⁸ Moreover, unlike VaR, both ES and DES satisfy the convexity property as shown in the next proposition.

Proposition 3 For any $\delta \in (0, 1)$, $R(\tilde{z}) = \text{ES}_\delta(\tilde{z})$ and $R(\tilde{z}) = \text{DES}_\delta(\tilde{z})$ are convex.

Example 6 The Aumann-Serrano and Foster-Hart risk measures. Two measures of riskiness have been proposed by Aumann and Serrano (2008, hereafter AS) and Foster and Hart (2009, hereafter FH). These measures generalize the notion of second-order stochastic dominance (SOSD). The AS measure $R^{AS}(\tilde{z})$ is given by the unique positive solution to the implicit equation

$$(4) \quad \text{E} \left[\exp \left(-\frac{\tilde{z}}{R^{AS}(\tilde{z})} \right) \right] = 1.$$

The FH measure $R^{FH}(\tilde{z})$ is given by the unique positive solution to the implicit equation

$$(5) \quad \text{E} \left[\log \left(1 + \frac{\tilde{z}}{R^{FH}(\tilde{z})} \right) \right] = 0.$$

Both these measures are homogeneous of degree 1 and smooth. These two risk measures also satisfy the convexity property.⁹ By contrast, these two measures do not satisfy the risk-free property.¹⁰

⁸The risk-free property follows since $\text{ES}_\delta(\tilde{z}) + \text{E}(\tilde{z}) \geq 0$ for all significance level $0 < \delta < 1$ with equality if and only if \tilde{z} is a constant with probability 1.

⁹This follows since these risk measures are subadditive and homogeneous of degree 1.

¹⁰To see this, note that for any constant $c > 0$, $\tilde{z} + c$ first-order stochastically dominates \tilde{z} . Since R^{AS} is consistent with first-order stochastic dominance, we have that $R^{AS}(\tilde{z} + c) < R^{AS}(\tilde{z})$. A similar argument applies to R^{FH} . Also, technically, these two risk measures are not defined for risk-free assets.

All of the risk measures discussed thus far are homogeneous of some degree. However, most of our results do not require homogeneity. The next set of examples illustrates how nonhomogeneous risk measures satisfying all of the other properties can be constructed.

Example 7 Nonhomogeneous risk measures. *Let R be a risk measure which is homogeneous of some degree k , convex, smooth, and satisfies the risk-free property, and let $h : [0, \infty) \rightarrow \mathcal{R}$ be a strictly increasing, strictly convex, and continuously differentiable function. Define a new risk measure \hat{R} by*

$$\hat{R}(\tilde{z}) = h(R(\tilde{z})) - h(0).$$

It is straightforward to verify that \hat{R} is scaling independent, convex, smooth, and satisfies the risk-free property. However, \hat{R} may fail to be homogeneous of any degree. For a concrete example, set $h(x) = e^x$, and let $R(\tilde{z}) = m_k(\tilde{z})$ for k even. Then, $\hat{R}(\tilde{z}) = e^{R(\tilde{z})} - 1$ is not homogeneous of any degree and yet it satisfies all of the other properties.

III. Systematic Risk in an Equilibrium Setting

Traditionally, systematic risk is derived from the CAPM equilibrium setting. We will now present a generalized version of this model. We first outline the setup of the model. We then study the geometry of solutions, and present a two-fund separation result implying the efficiency of the market portfolio. Finally, we derive a variant of the security market line, enabling us to obtain a generalization of the traditional beta as a measure of systematic risk.

A. Model Setup

Investors, Assets, and Timing. Assume a market with $n + 1$ assets $\{0, \dots, n\}$. Assets $1, \dots, n$ are risky and pay a random amount denoted by $(\tilde{y}_1, \dots, \tilde{y}_n)$. Asset 0 is risk-free, paying an amount \tilde{y}_0 which is equal to some constant $y_0 \neq 0$ with probability 1. Denote $\tilde{\mathbf{y}} = (\tilde{y}_0, \dots, \tilde{y}_n)$. There are ℓ investors in the market, all of whom agree on the parameters of the model. The choice set of each investor is

\mathcal{R}^{n+1} , where $\zeta^j \in \mathcal{R}^{n+1}$ represents the number of shares investor j chooses in each asset $i = 0, \dots, n$, i.e., ζ^j is a bundle of assets. Negative numbers represent short sales, and we impose no short-sale constraints. The initial endowment of investor j is a nonzero $e^j \in \mathcal{R}_+^{n+1}$. We assume that $\sum_{j=1}^{\ell} e_i^j > 0$ for $i = 1, \dots, n$. That is, risky assets are in positive net supply. An *allocation* is an ℓ -tuple $\mathcal{A} = (\zeta^1, \dots, \zeta^\ell)$ consisting of a bundle $\zeta^j \in \mathcal{R}^{n+1}$ for each investor. An allocation \mathcal{A} is *attainable* if $\sum_{j=1}^{\ell} \zeta^j = \sum_{j=1}^{\ell} e^j$, that is, if it clears the market. A *price system* is a vector $\mathbf{p} = (p_0, \dots, p_n)$ specifying a price for each asset. Similar to the standard CAPM setting, there are two dates. At Date 0, investors trade with each other and prices are set. At Date 1, all random variables are realized.

Risk and Preferences. The traditional approach features investors with mean-variance preferences, i.e., they prefer higher mean and lower variance of investments. Instead, we assume that investors have mean-risk preferences. Formally, fix a risk measure $R(\cdot)$. The utility that investor $j = 1, \dots, \ell$ assigns to a bundle $\zeta \in \mathcal{R}^{n+1}$ is given by

$$(6) \quad U^j(\zeta) = V^j(\mathbb{E}(\zeta \cdot \tilde{\mathbf{y}}), R(\zeta \cdot \tilde{\mathbf{y}})),$$

where V^j is continuous, strictly increasing in its first argument (expected payoff) and strictly decreasing in its second argument (risk of payoff), and quasi-concave.

Note that $U^j(\zeta)$ cannot be in general supported as a von Neumann-Morgenstern utility. Nevertheless, in the Online Appendix we show that if V^j is differentiable and if the risk measure is a differentiable function of a finite number of moments, then $U^j(\zeta)$ is a local expected utility function in the sense of Machina (1982). Namely, comparisons of “close by” investments are well approximated by expected utility. These conditions apply to a wide range of risk measures representing high distribution moments.

An implication of quasi-concavity of V^j is that when plotted in the mean-risk space, the upper contour of each indifference curve is convex. Similar to the standard mean-variance case, we will assume that a risk-free asset cannot be created synthetically from risky assets. That is, there is no redundant risky asset: for any $\zeta = (\zeta_0, \zeta_1, \dots, \zeta_n) \in \mathcal{R}^{n+1}$ we have $R(\zeta \cdot \tilde{\mathbf{y}}) \neq 0$ unless $(\zeta_1, \dots, \zeta_n) = (0, \dots, 0)$.¹¹

¹¹In the standard mean-variance case this condition corresponds to the variance-covariance matrix

Equilibrium. An *equilibrium* is a pair $(\mathbf{p}, \mathcal{A})$ where $\mathbf{p} \neq 0$ is a price system and $\mathcal{A} = (\zeta^1, \dots, \zeta^\ell)$ is an attainable allocation, such that for each $j \in \{1, \dots, \ell\}$, $\mathbf{p} \cdot \zeta^j = \mathbf{p} \cdot \mathbf{e}^j$, and if $\zeta \in \mathcal{R}^{n+1}$ and $U^j(\zeta) > U^j(\zeta^j)$ then $\mathbf{p} \cdot \zeta > \mathbf{p} \cdot \mathbf{e}^j$. In words, an equilibrium is a price system and an allocation that clear the market such that each investor optimizes subject to her budget constraint. The next theorem specifies conditions under which an equilibrium exists.

Theorem 1 *Suppose that $R(\cdot)$ is convex, smooth, and satisfies the risk-free property. Then, an equilibrium exists.*

It is well known that the CAPM setting can yield negative or zero prices (see for example Nielsen (1992)). The reason for this is that preferences are not necessarily monotone in the number of shares. Specifically, the expected payoff to an investor's bundle increases as she holds more shares of a risky asset, but so does the risk. It may well be that at some point, the additional expected payoff gained from adding more shares to the bundle is not sufficient to compensate for the increase in risk. If the equilibrium happens to fall in such a region then the asset becomes undesirable, rendering a negative price. For our following results we will need that prices are positive for all assets. The literature has suggested several ways to guarantee such an outcome. In the Online Appendix we provide one sufficient condition which follows Nielsen (1992). Other (and possibly weaker) sufficient conditions may be obtained, but are beyond the scope of this paper.

From now on we will only consider equilibria with positive prices. Given positivity of prices, naturally, each equilibrium induces a vector of random returns $\tilde{z}_i = \tilde{y}_i/p_i$, and we can talk about the expected returns and the risk of the returns in equilibrium, as in the usual CAPM setting. In particular, the equilibrium return from the risk-free asset \tilde{z}_0 is equal to some constant r_f with probability 1. We now study these returns.

of risky assets being positive-definite.

B. A Generalized CAPM

i. Geometry of Efficient Portfolios

Let $(\mathbf{p}, \mathcal{A})$ be an equilibrium. The equilibrium allocation $(\zeta^1, \dots, \zeta^\ell)$ naturally induces a portfolio for each investor j given by $\mathbf{x}^j = (x_0^j, \dots, x_n^j)$, where $x_i^j = p_i \zeta_i^j$ is the amount invested in asset i , and where the vector of portfolio weights of investor j is denoted by α^j and given by $\alpha_i^j = x_i^j / (\sum_{h=0}^n x_h^j)$. Let

$$\mu^j = \sum_{i=0}^n \alpha_i^j \mathbb{E}(\tilde{z}_i)$$

be the expected return obtained by investor j in equilibrium. The next theorem shows that the standard procedure of “minimizing risk for a given expected return” applies to the equilibrium setting. It relies on the scaling independence and convexity of the risk measure.

Theorem 2 *Suppose that $R(\cdot)$ is scaling independent and convex. Then, in an equilibrium with positive prices, for all investors $j \in \{1, \dots, \ell\}$, α^j is the unique solution to*

$$(7) \quad \begin{aligned} & \min_{\alpha \in \mathcal{R}^{n+1}} R(\alpha \cdot \tilde{\mathbf{z}}) \\ & \text{s.t.} \\ & \sum_{i=0}^n \alpha_i \mathbb{E}(\tilde{z}_i) = \mu^j. \\ & \sum_{i=0}^n \alpha_i = 1. \end{aligned}$$

Given this, we can now discuss the geometry of portfolios in the μ - R plane where the horizontal axis is the risk of the return of a portfolio (R) and the vertical axis is the expected return (μ). The locus of portfolios minimizing risk for any given expected return is the boundary of the portfolio opportunity set. This set is convex in the μ - R plane whenever $R(\cdot)$ is a convex risk measure. This follows simply because the expectation operator is linear, implying that the line connecting any two portfolios in the μ - R plane lies to the right of the set of portfolios representing convex combinations of these two portfolios. Figure 1 illustrates two curves. The

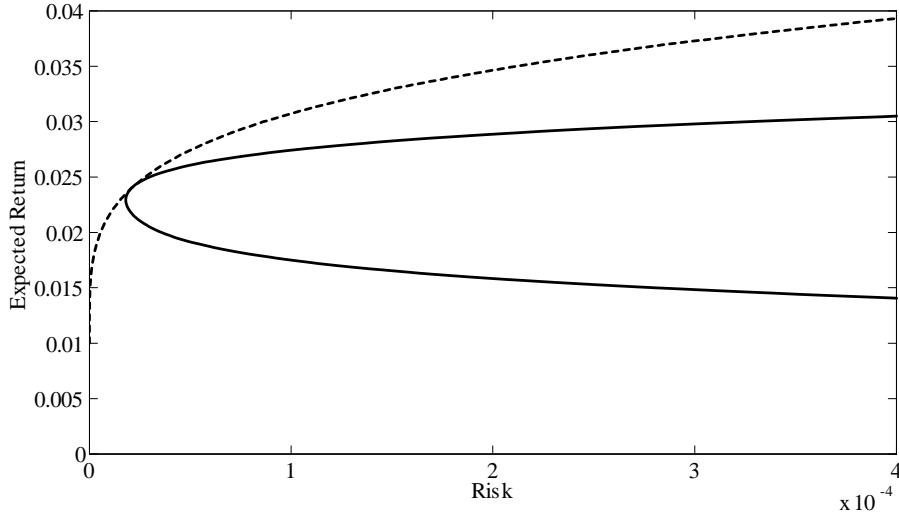


Figure 1: Portfolio Opportunity Set and Efficient Frontier

solid curve depicts the opportunity set of risky assets only. The dashed curve depicts portfolios minimizing risk for a given expected return, corresponding to Program (7). Both of these are defining convex sets. Unlike in the special case of the standard deviation, we do not, in general, obtain a straight line connecting the risk-free asset and risky portfolios. We say that a portfolio is *efficient* if it solves Program (7) for some $\mu^j \in \mathcal{R}$. Thus, the dashed curve in Figure 1 corresponds to the set of efficient portfolios.

ii. Two-Fund Separation

We say that two-fund separation holds if the equilibrium optimal portfolios for all investors can be spanned by the risk-free asset and a unique portfolio of risky assets. That is, there exists a unique portfolio with weights α^P such that $\alpha_0^P = 0$, and for all investors $j \in \{1, \dots, \ell\}$, the solution to Problem (7) is a linear combination of α^P and the risk-free asset.

Theorem 3 Consider an equilibrium with positive prices. Assume that $R(\cdot)$ is scaling independent, convex, and satisfies the risk-free property. Then, two-fund separation holds.

The proof is very intuitive, and we show it here. Let α^1 and α^2 be solutions to Problem (7) for investors $j_1 \neq j_2$, respectively, and without loss of generality assume $j_1 = 1$ and $j_2 = 2$. The case of interest is when both α^1 and α^2 have nonzero weights in some risky assets.¹² By the risk-free property and by the nonredundancy assumption, $R(\alpha^j \cdot \tilde{\mathbf{z}}) > 0$ for $j = 1, 2$. Hence, $\mu^j = E(\alpha^j \cdot \tilde{\mathbf{z}}) > r_f$ for $j = 1, 2$, since otherwise α^j would be mean-risk dominated by the risk-free asset, and thus would not be optimal.

Now, consider all the linear combinations of these two portfolios with the risk-free asset. Since $R(\cdot)$ is assumed convex, the resulting curves are concave in the μ - R plane as illustrated in Figure 2. Note that both α^1 and α^2 can be presented as a linear combination of the risk-free asset and some portfolios α^{P_1} and α^{P_2} of risky assets only (i.e., $\alpha_0^{P_1} = \alpha_0^{P_2} = 0$). To show two-fund separation we need to show that $\alpha^{P_1} = \alpha^{P_2}$. Suppose this is not the case. Then let $\hat{\alpha}^1$ be a linear combination of α^{P_2} and the risk-free asset such that $E(\hat{\alpha}^1 \cdot \tilde{\mathbf{z}}) = \mu^1$. Similarly, let $\hat{\alpha}^2$ be a portfolio of α^{P_1} and the risk-free asset such that $E(\hat{\alpha}^2 \cdot \tilde{\mathbf{z}}) = \mu^2$. By convexity of $R(\cdot)$, α^1 and α^2 are the unique solutions to Program (7) for $j = 1, 2$. Hence,

$$(8) \quad R(\hat{\alpha}^1 \cdot \tilde{\mathbf{z}}) > R(\alpha^1 \cdot \tilde{\mathbf{z}}) \text{ and } R(\hat{\alpha}^2 \cdot \tilde{\mathbf{z}}) > R(\alpha^2 \cdot \tilde{\mathbf{z}}).$$

Thus, as illustrated in Figure 2, the two curves must cross at least once. We will now show that such crossings are impossible. Indeed, by scaling independence (8) implies that for any $\lambda > 0$,

$$R(\lambda \alpha^1 \cdot \tilde{\mathbf{z}}) < R(\lambda \hat{\alpha}^1 \cdot \tilde{\mathbf{z}}),$$

which together with risk-free property implies

$$R(\lambda \alpha^1 \cdot \tilde{\mathbf{z}} + (1 - \lambda) r_f) < R(\lambda \hat{\alpha}^1 \cdot \tilde{\mathbf{z}} + (1 - \lambda) r_f).$$

This means that all linear combinations of α^1 with the risk-free asset (with positive λ) lie strictly to the left of all linear combinations of $\hat{\alpha}^1$ with the risk-free asset. In particular, $\hat{\alpha}^2$ can be obtained as a linear combination of α^1 with the risk-free asset by setting

$$\lambda = \frac{\mu^2 - r_f}{\mu^1 - r_f} > 0,$$

¹²If only one investor holds nonzero weights in risky assets then two-fund separation is trivial.

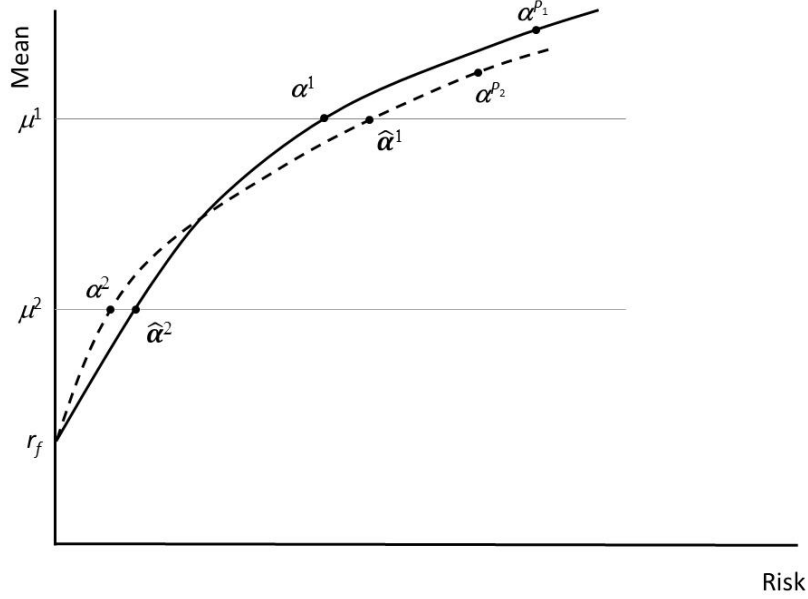


Figure 2: Graphical Illustration of the Proof of Theorem 3

where the inequality follows since $\mu^j > r_f$ for $j = 1, 2$. But, using this λ we obtain

$$R(\hat{\alpha}^2 \cdot \tilde{\mathbf{z}}) < R(\alpha^2 \cdot \tilde{\mathbf{z}}),$$

contradicting (8). Thus, two-fund separation must hold.

A corollary is that the unique portfolio α^P is efficient. Indeed, let $\mu^P = E(\alpha^P \cdot \tilde{\mathbf{z}})$. Since in equilibrium all investors hold a linear combination of the risk-free asset and α^P , and since $\mu^j = E(\alpha^j \cdot \tilde{\mathbf{z}}) \geq r_f$ for all j with strict inequality for some j , we have two cases:¹³ (i) all investors hold α^P with a nonnegative weight, and $\mu^P > r_f$; or (ii) all investors hold α^P with a nonpositive weight, and $\mu^P < r_f$. But, the second case is impossible since then the market cannot clear for at least one risky asset, which is held in positive weight in α^P . Thus, $\mu^P > r_f$.

Now, assume that $\alpha' \neq \alpha^P$ solves Problem (7) for $\mu^j = \mu^P$. Then, $R(\alpha' \cdot \tilde{\mathbf{z}}) < R(\alpha^P \cdot \tilde{\mathbf{z}})$, and so by the same argument as in the proof of Theorem 3, all linear combinations of α' with the risk-free asset would have strictly lower risk than the corresponding linear combinations of α^P with the risk-free asset. This contradicts

¹³If all investors choose the risk-free asset then the market for risky assets cannot clear.

that α^P and the risk-free asset span all optimal portfolios. We thus have:

Corollary 1 *Under the conditions of Theorem 3, the portfolio α^P is efficient. In particular, it solves Problem (7) for some $\mu^P > r_f$.*

Let $x_i^M = \sum_{j=1}^{\ell} x_i^j$ be the total amount invested in asset i in equilibrium. We call $\mathbf{x}^M = (x_1^M, \dots, x_n^M)$ the *market portfolio* (consisting of risky assets only). Let α^M be the corresponding portfolio weights. By Theorem 3, in equilibrium, the market portfolio is equal to α^P , the unique portfolio of risky assets that together with the risk-free asset spans all optimal portfolios.¹⁴ Moreover, by corollary 1, the market portfolio is efficient, and its expected return is strictly higher than r_f .

Corollary 2 *Under the conditions of Theorem 3, the market portfolio is efficient. In particular, it solves Problem (7) for some $\mu^M > r_f$.*

iii. A Generalized Security Market Line

In the traditional CAPM framework, the security market line describes the equilibrium relation between the expected returns of individual assets and the market expected return. Specifically, the expected return of any asset in excess of the risk-free rate is proportional to the excess market expected return, with the coefficient of proportionality being equal to the traditional beta. The following theorem provides sufficient conditions under which a similar relation holds for a broad set of risk measures.

Theorem 4 *Consider an equilibrium with positive prices and let α^M be the market portfolio. Assume that $R(\cdot)$ is scaling independent, convex, smooth, and satisfies the risk-free property. Then, for each asset $i = 1, \dots, n$,*

$$(9) \quad \mathbb{E}(\tilde{z}_i) = r_f + \mathcal{B}_i^R (\mathbb{E}(\alpha^M \cdot \tilde{\mathbf{z}}) - r_f),$$

where

$$(10) \quad \mathcal{B}_i^R = \frac{R_i(\alpha^M)}{\sum_{h=1}^n \alpha_h^M R_h(\alpha^M)}.$$

¹⁴Note that α^P is of dimension $n + 1$, but its first component is zero. By saying that $\alpha^P = \alpha^M$ we mean that $\alpha_i^P = \alpha_i^M$ for $i = 1, \dots, n$.

If $R(\cdot)$ is also homogeneous of some degree k , then (10) takes the form

$$\mathcal{B}_i^R = \frac{R_i(\boldsymbol{\alpha}^M)}{kR(\boldsymbol{\alpha}^M)}.$$

Thus, the security market line has the traditional form, with the generalized systematic risk measure (\mathcal{B}_i^R) given as the marginal contribution of asset i to the total risk of the market portfolio, scaled by the weighted average of marginal contributions of all assets. If R is furthermore homogeneous, it is simply given by the marginal contribution of asset i scaled by total risk multiplied by the degree of homogeneity.

To see the intuition for this result, start with an efficient portfolio $\boldsymbol{\alpha}^*$ and consider borrowing one dollar at the risk-free rate and investing this dollar in asset i . The effect of this exercise on the risk of the portfolio is (up to first-order approximation) $R_i(\boldsymbol{\alpha}^*) - R_0(\boldsymbol{\alpha}^*)$, which by the risk-free property is just $R_i(\boldsymbol{\alpha}^*)$. Since $\boldsymbol{\alpha}^*$ is efficient, the effect of this exercise on risk is equal to the shift in the expected return constraint times the shadow price of the constraint, ξ , i.e.,

$$(11) \quad R_i(\boldsymbol{\alpha}^*) = \xi (E(z_i) - r_f).$$

Taking the weighted average using the portfolio weights gives

$$(12) \quad \sum_{i=1}^n \alpha_i^* R_i(\boldsymbol{\alpha}^*) = \xi (E(\boldsymbol{\alpha}^* \cdot \tilde{\mathbf{z}}) - r_f).$$

Using (11) and (12) we obtain that for any efficient portfolio $\boldsymbol{\alpha}^*$,

$$\frac{R_i(\boldsymbol{\alpha}^*)}{\sum_{i=1}^n \alpha_i^* R_i(\boldsymbol{\alpha}^*)} = \frac{E(z_i) - r_f}{E(\boldsymbol{\alpha}^* \cdot \tilde{\mathbf{z}}) - r_f}.$$

Namely, in equilibrium, \mathcal{B}_i^R (as given in (10)) equals the ratio of the expected excess return of any asset i to the expected excess return of the efficient portfolio $\boldsymbol{\alpha}^*$. Finally, since $\boldsymbol{\alpha}^M$ has been shown to be efficient (Corollary 2) we obtain the result.

C. Applications and Empirical Implementation

We now provide several applications to illustrate the versatility and power of Theorem 4 and its potential empirical usefulness. We show how to use this theorem to

generalize the traditional CAPM to account for high distribution moments, downside risk, rare disasters, as well as combinations thereof. We also discuss the economic interpretation of systematic risk in these cases and explain how these results can be implemented empirically.

i. Applications

Application I: The standard CAPM. When the risk measure R is the variance, i.e., $R(\tilde{z}) = \text{Var}(\tilde{z})$, Theorem 4 coincides with the standard CAPM (see Appendix B for the derivation). Namely,

$$(13) \quad \mathcal{B}_i^R = \frac{\text{Cov}(\tilde{z}_i, \boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})}{\text{Var}(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})}.$$

Thus, in this case systematic risk is measured as the standard regression coefficient. The same result is obtained when $R(\tilde{z}) = w_2(\tilde{z})$, i.e., the standard deviation of returns.

Application II: A CAPM reflecting aversion to tail risk. The simplest generalization of the standard CAPM is to the case in which investors are averse to any moment of an even degree. That is, set $R(\tilde{z}) = m_k(\tilde{z}) = \text{E}(\tilde{z} - \text{E}(\tilde{z}))^k$, k even. This risk measure satisfies all of the conditions in Theorem 4 (see Example 1). In this case the systematic risk takes the form (see Appendix B for the derivation)

$$(14) \quad \mathcal{B}_i^R = \frac{\text{Cov}\left(\tilde{z}_i, \left(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}} - \boldsymbol{\alpha}^M \cdot \text{E}(\tilde{\mathbf{z}})\right)^{k-1}\right)}{m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})}.$$

That is, the systematic risk of asset i is proportional to the covariance of \tilde{z}_i with the $(k-1)^{th}$ power of the demeaned market return. In the special case of $k=2$ (variance), this reduces to (13) as expected. Another important special case is $k=4$, in which $R(\tilde{z})$ measures the tail risk of \tilde{z} . Then,

$$\mathcal{B}_i^R = \frac{\text{Cov}\left(\tilde{z}_i, \left(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}} - \boldsymbol{\alpha}^M \cdot \text{E}(\tilde{\mathbf{z}})\right)^3\right)}{m_4(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})}.$$

Namely, the systematic risk of asset i is proportional to the co-kurtosis of \tilde{z}_i with the demeaned market return. Similarly, when $R(\tilde{z}) = w_k(\tilde{z})$, the normalized k^{th} central moment, \mathcal{B}_i^R again takes the form (14).

Application III: A CAPM reflecting aversion to downside risk. Assume $R(\tilde{z}) = \text{DR}_k(\tilde{z})$ for $k \geq 2$. This risk measure satisfies all of the conditions in Theorem 4 (see Example 3). The systematic risk is then given by (see Appendix B for the derivation)

$$(15) \quad \mathcal{B}_i^R = (-1)^k \frac{\text{Cov} \left[\tilde{z}_i, \left([\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}} - \text{E}(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})]^- \right)^{k-1} \right]}{(\text{DR}_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}}))^k}.$$

That is, the systematic risk of asset i is proportional to the covariance of \tilde{z}_i with the $(k-1)^{\text{th}}$ power of the demeaned market return, censored at zero.

Application IV: A CAPM reflecting aversion to rare disasters. To account for rare disasters we can use the demeaned expected shortfall measure, which satisfies all the requirements in Theorem 4 (see Example 5). Assume then that $R(\tilde{z}) = \text{DES}_\delta(\tilde{z})$, where $0 < \delta < 1$ is some confidence level. The systematic risk in this case is given by (see Appendix B for the derivation)

$$(16) \quad \mathcal{B}_i^R = - \frac{\text{E} \left[\tilde{z}_i - \text{E}(\tilde{z}_i) \mid \boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}} \leq -\text{VaR}_\delta(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}}) \right]}{\text{DES}_\delta(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})}.$$

Thus, the systematic risk of asset i in this case equals (the negative of) the expected demeaned return of asset i conditional on the market being in a disaster, scaled by the market's demeaned expected shortfall. An equivalent expression is

$$\mathcal{B}_i^R = - \frac{\text{Cov} \left[\tilde{z}_i, \mathbb{1}_{\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}} \leq -\text{VaR}_\delta(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})} \right]}{\delta \cdot \text{DES}_\delta(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})},$$

showing that systematic risk in this case is proportional to the covariance of the asset return with an indicator equal to one when the market is in a disaster.

So far we have restricted our applications to cases in which investors are averse to just one risk aspect. In reality, it is likely that investors are averse to several risk attributes. Our framework allows for this by constructing risk measures that account for several risk characteristics using Lemma 1. The next application illustrates this point.

Application V: A CAPM reflecting aversion to variance, downside skewness, tail risk, and rare disasters. Consider the following family of risk measures

$$R(\tilde{z}) = \theta_1 w_2(\tilde{z}) + \theta_2 \text{DR}_3(\tilde{z}) + \theta_3 w_4(\tilde{z}) + \theta_4 \text{DES}_\delta(\tilde{z})$$

for some confidence level δ . Here $\theta_1, \dots, \theta_4$ are nonnegative weights accounting for the degree of aversion to variance, downside skewness, tail risk, and rare disasters, respectively.¹⁵ The case $\theta_1 = 1$ and $\theta_2 = \theta_3 = \theta_4 = 0$ corresponds to the traditional CAPM, whereas different values of the weights allow us to reflect different levels of aversion to the different risk attributes.

By Lemma 1 these risk measures satisfy all the conditions in Theorem 4 and so all the CAPM results above hold. The resulting systematic risk measure accounts for the contribution of asset i to all four risk attributes. It is simply given by a weighted average of the systematic risk measures as calculated in the above applications (see Appendix B). Namely,

$$(17) \quad \mathcal{B}_i^R = \frac{R^1(\boldsymbol{\alpha}^M)}{R(\boldsymbol{\alpha}^M)} \mathcal{B}_i^{R^1} + \frac{R^2(\boldsymbol{\alpha}^M)}{R(\boldsymbol{\alpha}^M)} \mathcal{B}_i^{R^2} + \frac{R^3(\boldsymbol{\alpha}^M)}{R(\boldsymbol{\alpha}^M)} \mathcal{B}_i^{R^3} + \frac{R^4(\boldsymbol{\alpha}^M)}{R(\boldsymbol{\alpha}^M)} \mathcal{B}_i^{R^4},$$

where $R^1(\cdot) = \theta_1 w_2(\cdot)$, $R^2(\cdot) = \theta_2 \text{DR}_3(\cdot)$, $R^3(\cdot) = \theta_3 w_4(\cdot)$, and $R^4(\cdot) = \theta_4 \text{DES}_\delta(\cdot)$, and where $\mathcal{B}_i^{R^1}$, $\mathcal{B}_i^{R^2}$, $\mathcal{B}_i^{R^3}$, and $\mathcal{B}_i^{R^4}$ are given by (13)–(16).

ii. Empirical Implementation

Similar to the classic CAPM, Theorem 4 and its applications lend themselves naturally to empirical investigation. The standard approach for testing and applying the CAPM follows Fama and MacBeth (1973) and Fama and French (1992). The first stage in their approach consists of estimating beta through time-series regressions, whereas the second stage consists of cross-sectional regressions of excess asset returns on estimated betas.

To apply this approach in our case, one needs to first take a stand on what the risk measure R is. Then, using Theorem 4 one can estimate \mathcal{B}_i^R from time-series data. For example, if R takes the form as in Application V above, then we need time

¹⁵Note that we are using here $w_2(\cdot)$ and $w_4(\cdot)$ (the normalized second and fourth moments) instead of $m_2(\cdot)$ and $m_4(\cdot)$. This is done to make sure that all of the components in $R(\cdot)$ are homogeneous of degree 1, and so $R(\cdot)$ is homogeneous.

series return data for asset i and for the market portfolio in order to estimate \mathcal{B}_i^R from (17). This will be a weighted average of the betas prescribed in Applications I–IV. Note that unlike in the classic CAPM, \mathcal{B}_i^R is in general not a regression coefficient. Nevertheless, it often takes the form of some scaled covariance of the asset returns and some function of the market returns (see Applications I–IV). Thus, \mathcal{B}_i^R can still be readily estimated from time-series return data. The cross-sectional part is then identical in nature to that in Fama and MacBeth (1973).

It is important to note that the model does not provide us with guidance as to what R is. Rather, for any given risk measure the model provides an expression for the associated systematic risk. In practice we believe that the data can guide us in finding what the “true” risk is, to which investors are averse. For example, consider Application V, which allows the risk measure to reflect aversion to variance, downside skewness, tail risk, and disaster risk. One still has a lot of flexibility in choosing the weights $\theta_1, \dots, \theta_4$, which determine the degree of aversion to each particular aspect of risk. The model can then allow the data to determine which set of weights obtains the most support. This flexibility is tantamount to the freedom provided by the Arbitrage Pricing Theory (Ross (1976)) in which the model suggests the existence of multiple systematic factors but does not provide guidance as to what these factors are.

D. Further Discussion

Note that Theorem 4 relies on the market portfolio being efficient, and that two-fund separation is a way to achieve this efficiency result. Our assumptions on the risk measure are sufficient for two-fund separation, but they are by no means necessary. Weaker conditions that guarantee two-fund separation may exist. Further, even when two-fund separation fails, it does not necessarily mean that market efficiency is rejected. The literature explores market efficiency from both theoretical (see, for example, Dybvig and Ross (1982)) and empirical (see, for example, Levy and Roll (2010)) views. Our generalized SML remains valid as long as we have evidence that the market portfolio is mean-risk efficient.

We should also mention that the classical notion of beta and its relation to expected returns go beyond the standard CAPM setup. Specifically, as long as there is

no arbitrage and so a stochastic discount factor exists, a beta representation of the form

$$E(\tilde{z}_i) = \gamma + \mathcal{B}_i \lambda$$

exists (see Hansen and Richard (1987) and Cochrane (2001) Ch. 6). This does not stand in conflict to the results in this section. Rather, our results essentially identify a class of stochastic discount factors driven by the mean-risk preferences being assumed.

IV. Systematic Risk as a Solution to a Risk Allocation Problem

The equilibrium approach presented in the previous section generalizes the classic CAPM, but it has two limitations. First, this approach allows us to calculate the contribution of an asset to the risk of the market portfolio, but not to arbitrary portfolios of risky assets. Second, to obtain the equilibrium results we imposed restrictions on the risk measures (scaling independence, convexity, and the risk-free property). These restrictions allow us to establish existence of equilibrium and efficiency of the market portfolio. However, some risk measures do not satisfy these conditions.

In this section we offer an alternative approach to developing a systematic risk measure. This approach applies to any portfolio of risky assets and to a broader class of risk measures. For example, if a bank would like to use the VaR measure to estimate the contributions of different assets on its balance sheet to the total VaR of the bank, then the results in this section can be applied. Importantly, when the risk measure is homogeneous, the two approaches lead to an identical result, generalizing the traditional beta.

Our approach is to consider this issue as a risk allocation problem, where the total risk of a given portfolio needs to be “fairly” allocated among its components. We offer four axioms that describe reasonable properties of solutions to risk allocation problems. We then show that these axioms determine a unique formula for the systematic risk of an asset, the contribution of the asset to the risk of the portfolio.

A. Axiomatic Characterization of Systematic Risk

A *risk allocation problem* of order $n \geq 1$ is a pair (R, \mathbf{x}) , where R is a risk measure and $\mathbf{x} \in \mathcal{R}_{++}^n$ is a portfolio specifying the dollar amount invested in each of n assets $\tilde{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_n)$, and $R(\mathbf{x} \cdot \tilde{\mathbf{z}}) \neq 0$. Denote the total dollar amount invested by $\bar{x} = \sum_{i=1}^n x_i$. Also, let α be the vector of corresponding portfolio weights, i.e., $\alpha_i = x_i/\bar{x}$. The only two requirements we impose on R in this section are that $R(0) = 0$ (i.e., zero investment entails no risk) and that $R(\cdot)$ is smooth.

A *systematic risk measure* is a function mapping any risk allocation problem of order n to a vector $\mathbf{B}^R(\mathbf{x}) = (\mathcal{B}_1^R(\mathbf{x}), \dots, \mathcal{B}_n^R(\mathbf{x}))$ in \mathcal{R}^n . Intuitively, one can think of $\mathcal{B}_i^R(\mathbf{x})$ as the contribution of asset i to the total risk of portfolio \mathbf{x} , which is $R(\mathbf{x} \cdot \tilde{\mathbf{z}})$. Note that a systematic risk measure applies to all possible pairs of risk measures and portfolios, rather than to a given pair.

We now state four axioms specifying desirable economic properties of systematic risk measures. The intuition for why these axioms make sense mostly comes from the traditional beta. Here we simply try to identify properties of beta and ask how these properties could be generalized to arbitrary risk measures. It is important to emphasize that these axioms do not impose any restriction on the risk measure. Rather, they impose structure on what would constitute a solution to the risk allocation problem.

The first axiom postulates that (as for the traditional beta) the weighted average of systematic risk values across all assets is normalized to 1.

Axiom 1 *Normalization:* $\sum_{i=1}^n \alpha_i \mathcal{B}_i^R(\mathbf{x}) = 1$.

The sum of any two risk measures is itself a risk measure. The next axiom requires that in such a case the systematic risk measure of the sum will be a risk-weighted average of systematic risk based on each of the two risk components.

Axiom 2 *Linearity:* If $R(\cdot) = R^1(\cdot) + R^2(\cdot)$, then

$$\mathcal{B}_i^R(\mathbf{x}) = \frac{R^1(\mathbf{x})}{R(\mathbf{x})} \mathcal{B}_i^{R^1}(\mathbf{x}) + \frac{R^2(\mathbf{x})}{R(\mathbf{x})} \mathcal{B}_i^{R^2}(\mathbf{x}) \text{ for all } i = 1, \dots, n.$$

When risk is measured using variance, the notion of systematic risk is closely tied to the concepts of correlation and covariance. It is not easy to generalize these concepts to arbitrary risk measures. However, two features can be easily generalized laying the foundations for the next two axioms.

First, while the concept of “correlation” is not easy to generalize, the idea of “perfect correlation” does lend itself to a natural generalization. The intuition is that if several assets are perfectly correlated, then essentially they can be thought of as the same asset. Thus, a portfolio of perfectly correlated assets can be viewed as one “big” asset. This intuition comes from the standard notion of correlation relating to risk being measured by the variance, but it can easily be generalized to arbitrary risk measures.

Formally, given a risk measure R , we say that assets $\tilde{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_n)$ are R -perfectly correlated if there exists a function $g(\cdot) : \mathcal{R} \mapsto \mathcal{R}$ and a nonzero vector $\mathbf{q} = (q_1, \dots, q_n) \in \mathcal{R}_+^n$, such that for any portfolio $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n) \in \mathcal{R}_+^n$ we have $R(\boldsymbol{\eta} \cdot \tilde{\mathbf{z}}) = g(\boldsymbol{\eta} \cdot \mathbf{q})$. That is, the n assets are R -perfectly correlated if the risk of any portfolio of these assets as measured by R only depends on some linear combination of their investment amounts. In essence, this means that the n assets can be aggregated into one “big” asset by assigning each asset a certain weight specified by the vector \mathbf{q} .¹⁶ Note that different risk measures correspond to different concepts of R -perfect correlation, which typically would not coincide with the standard notion of perfect correlation associated with the variance.¹⁷

The next axiom imposes that if the n assets are R -perfectly correlated, then their systematic risk measures are proportional to each other.

¹⁶To see the correspondence to the standard notion of perfect correlation, consider the following example. Assume risk is measured using variance and let $\tilde{\mathbf{z}} = (\tilde{z}_1, \tilde{z}_2, \tilde{z}_3)$ with $\tilde{z}_2 = 2\tilde{z}_1$ and $\tilde{z}_3 = 5\tilde{z}_1$. Then, all three assets are perfectly correlated and for any portfolio (η_1, η_2, η_3) we have

$$Var(\eta_1 \tilde{z}_1 + \eta_2 \tilde{z}_2 + \eta_3 \tilde{z}_3) = (\eta_1 + 2\eta_2 + 5\eta_3)^2 Var(\tilde{z}_1).$$

Thus, we can set $g(t) = t^2$ and the vector of weights is $\mathbf{q} = \sqrt{Var(\tilde{z}_1)}(1, 2, 5)$. More generally, it is easy to verify that when risk is measured using variance, the concept of R -perfect correlation coincides with the standard definition of perfect correlation.

¹⁷In the standard notion of perfect correlation, we differentiate between positive and negative perfect correlation. We could do the same here by allowing elements of \mathbf{q} to take negative values. However, this is not needed for our axiomatic characterization.

Axiom 3 Proportionality: If $\tilde{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_n)$ are R -perfectly correlated with weights $\mathbf{q} = (q_1, \dots, q_n)$, then

$$(18) \quad q_j \mathcal{B}_i^R(\mathbf{x}) = q_i \mathcal{B}_j^R(\mathbf{x}) \text{ for all } i, j = 1, \dots, n.$$

Next we turn to generalize the idea of “positive correlation.” Assume first that risk is measured using variance. Then, if two assets are positively correlated, adding additional units of an asset to any portfolio of the two always increases total variance. We can then rely on this feature to get a generalized notion of positive correlation. Specifically, given a risk measure R , we say that assets $\tilde{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_n)$ are R -positively correlated if $R_i(\boldsymbol{\eta} \cdot \tilde{\mathbf{z}}) \geq 0$ for all $\boldsymbol{\eta} \in \mathcal{R}_+^n$ and for all $i = 1, \dots, n$. Namely, the assets are R -positively correlated if adding one more unit of an asset to any portfolio with nonnegative weights can never reduce total risk. The key to this definition is that for the assets to be R -perfectly correlated it is not enough that adding one more unit of an asset would increase risk for a particular portfolio. Rather, this property has to hold for all possible portfolios of these assets.¹⁸ The next axiom requires that when the assets are R -positively correlated, the systematic risk of all assets is nonnegative.

Axiom 4 Monotonicity: If $\tilde{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_n)$ are R -positively correlated, then $\mathcal{B}_i^R(\mathbf{x}) \geq 0$ for all $i = 1, \dots, n$.

Our main result in this section follows. It states that Axioms 1–4 are sufficient to pin down a unique systematic risk measure, which takes on a very simple and intuitive form. Moreover, when the risk measure is homogeneous, the solution coincides with the equilibrium result in Theorem 4.

Theorem 5 *There exists a unique systematic risk measure satisfying Axioms 1–4. For each risk allocation problem (R, \mathbf{x}) of order n , it is given by*

$$(19) \quad \mathcal{B}_i^R(\mathbf{x}) = \frac{\bar{x} \int_0^1 R_i(tx_1, \dots, tx_n) dt}{R(x_1, \dots, x_n)} \text{ for } i = 1, \dots, n.$$

¹⁸It is easy to check that when risk is measured using variance, the assets are R -positively correlated if and only if the correlation between any two assets is nonnegative.

Furthermore, if R is homogeneous of some degree k , then (19) reduces to

$$(20) \quad \mathcal{B}_i^R(\mathbf{x}) = \frac{R_i(\boldsymbol{\alpha})}{\sum_{h=1}^n \alpha_h R_h(\boldsymbol{\alpha})}$$

$$(21) \quad = \frac{R_i(\boldsymbol{\alpha})}{kR(\boldsymbol{\alpha})}.$$

Thus, when R is homogeneous (which is a common case), the systematic risk of asset i is measured simply as the marginal contribution of asset i to the total risk of the portfolio, scaled by the weighted average of marginal contributions of all assets. This is identical to the result in Theorem 4 only with respect to an arbitrary portfolio rather than the market portfolio. When the risk measure is not homogeneous, the expression in (19) shows that systematic risk depends not only on marginal contributions at \mathbf{x} , but rather on marginal contributions along a diagonal between $(0, \dots, 0)$ and \mathbf{x} . This is a variation of the diagonal formula of Aumann and Shapley (1974). The integral can be interpreted as an average of marginal contributions of asset i to the risk of portfolios along the diagonal. Then, $\mathcal{B}_i^R(\mathbf{x})$ is simply a scaled version of the integral where the scaling ensures that Axiom 1 is satisfied.

Note that when the risk measure is homogeneous, $\mathcal{B}_i^R(\mathbf{x})$ depends only on the portfolio weights $\boldsymbol{\alpha}$ (and not on the dollar amounts invested in each asset). Indeed, in the homogeneous case $R_i(tx_1, \dots, tx_n)$ is proportional to $R_i(x_1, \dots, x_n)$ for all $t \in [0, 1]$, yielding the simple expression in (20). When the risk measure is not homogeneous, the actual investment amounts (not just the weights) are necessary for the calculation of systematic risk.

The uniqueness part of the proof of Theorem 5 is in Appendix A. It relies on the solutions to cost allocation problems established in Billera and Heath (1982).¹⁹ In this proof we draw a one-to-one mapping between risk allocation problems and cost allocation problems, and from systematic risk measures to solutions of cost allocation problems. Then, we show that given these mappings, our set of axioms is stronger than the set of conditions specified in Billera and Heath (1982). This in turn allows us to apply their result to obtain uniqueness.

¹⁹Billera and Heath (1982) define a cost allocation problem of order n as a pair (h, \mathbf{x}) where $h : \mathbb{R}_+^n \rightarrow \mathbb{R}$ is continuously differentiable and $h(\mathbf{0}) = 0$. They interpret \mathbf{x} as a vector of inputs and h as a cost function. The question they ask is how to allocate total cost among the different inputs. See Appendix I for more details on their model.

Existence is straightforward and we show it below by demonstrating that (19) satisfies Axioms 1–4. Suppose that $\mathcal{B}_i^R(\mathbf{x})$ is given by (19). Then,

$$\begin{aligned} \sum_{i=1}^n \alpha_i \mathcal{B}_i^R(\mathbf{x}) &= \sum_{i=1}^n \frac{x_i \bar{x}}{\bar{x}} \frac{\int_0^1 R_i(tx_1, \dots, tx_n) dt}{R(x_1, \dots, x_n)} \\ &= \frac{\int_0^1 \sum_{i=1}^n x_i R_i(tx_1, \dots, tx_n) dt}{R(x_1, \dots, x_n)} \\ &= \frac{\int_0^1 \frac{dR(tx_1, \dots, tx_n)}{dt} dt}{R(x_1, \dots, x_n)} = 1, \quad (\text{since } R(0) = 0) \end{aligned}$$

and so Axiom 1 holds. To see Axiom 2, suppose $R(\cdot) = R^1(\cdot) + R^2(\cdot)$. Then,

$$\begin{aligned} \mathcal{B}_i^R(\mathbf{x}) &= \frac{\bar{x} \int_0^1 R_i(tx_1, \dots, tx_n) dt}{R(x_1, \dots, x_n)} \\ &= \frac{\frac{\bar{x} \int_0^1 R_i^1(tx_1, \dots, tx_n) dt}{R^1(x_1, \dots, x_n)} R^1(x_1, \dots, x_n) + \frac{\bar{x} \int_0^1 R_i^2(tx_1, \dots, tx_n) dt}{R^2(x_1, \dots, x_n)} R^2(x_1, \dots, x_n)}{R(x_1, \dots, x_n)}, \end{aligned}$$

as required. Next, for Axiom 3, suppose that $\tilde{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_n)$ are R -perfectly correlated. Then, there exists $g(\cdot) : \mathcal{R} \mapsto \mathcal{R}$ and a nonzero vector $\mathbf{q} \in \mathcal{R}_+^n$ such that for all $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)$ we have $R(\boldsymbol{\eta}) = g(\boldsymbol{\eta} \cdot \mathbf{q})$. It follows that

$$R_i(\boldsymbol{\eta}) = q_i g'(\boldsymbol{\eta} \cdot \mathbf{q}) \text{ for all } i = 1, \dots, n.$$

Hence, for all $i = 1, \dots, n$,

$$\mathcal{B}_i^R(\mathbf{x}) = \frac{\bar{x} q_i \int_0^1 g'(t\mathbf{x} \cdot \mathbf{q}) dt}{R(x_1, \dots, x_n)},$$

which implies (18). Finally, given the definition of R -positive correlation, it is immediate that (19) satisfies Axiom 4.

B. Applying the Result

In Section III.C we have provided several applications and shown how to calculate systematic risk for different risk measures. All of these results apply to the approach presented in this section as well, but now they can be used with respect to arbitrary portfolios rather than just the market portfolio. The next example illustrates a case

of risk measures that do not satisfy the conditions in Section III, but for which Theorem 5 applies.

Recall the Aumann-Serrano and Foster-Hart risk measures in Example 6. These measures are homogeneous, convex, and smooth, but they do not satisfy the risk-free property.²⁰ Still, Theorem 5 allows us to calculate the systematic risk associated with these risk measures.

Using Theorem 5 and applying the implicit function theorem to (4) and (5) yields the systematic risk of individual assets associated with the AS and FH measures relative to any portfolio weights α as follows:

$$\mathcal{B}_i^{RAS}(\alpha) = \frac{\mathbb{E} \left[\exp \left(-\frac{\alpha \cdot \tilde{\mathbf{z}}}{R(\alpha)} \right) \tilde{z}_i \right]}{\mathbb{E} \left[\exp \left(-\frac{\alpha \cdot \tilde{\mathbf{z}}}{R(\alpha)} \right) \alpha \cdot \tilde{\mathbf{z}} \right]},$$

and

$$\mathcal{B}_i^{RFH}(\alpha) = \frac{\mathbb{E} \left[\frac{\tilde{z}_i}{R(\alpha) + \alpha \cdot \tilde{\mathbf{z}}} \right]}{\mathbb{E} \left[\frac{\alpha \cdot \tilde{\mathbf{z}}}{R(\alpha) + \alpha \cdot \tilde{\mathbf{z}}} \right]}.$$

C. Discussion

It is interesting to ask what would happen if we used (20) to define systematic risk when R is not homogeneous (instead of using (19)). In particular, this alternative measure only relies on the marginal contribution of asset i at α and not along the diagonal. In the absence of homogeneity these two alternative definitions yield different results. Thus, given Theorem 5, it must be that (20) violates at least one of our axioms. It is straightforward to check that the axiom being violated in this case is Axiom 2 while the other three axioms are satisfied.

²⁰Although $R^{AS}(0)$ and $R^{FH}(0)$ are not defined, they can be approximated using a limiting argument. Specifically, take any random return \tilde{z} satisfying $\mathbb{E}(\tilde{z}) > 0$ and $P(\{\tilde{z} < 0\}) > 0$. Then, for both $R(\cdot) = R^{AS}(\cdot)$ and $R(\cdot) = R^{FH}(\cdot)$, we can define $R(0)$ by

$$R(0) \equiv \lim_{t \rightarrow 0} R(t\tilde{z}) = 0,$$

where the equality follows since both the AS and the FH measures are homogeneous of degree 1.

Another alternative to measuring systematic risk might be to define

$$\mathcal{B}_i^R(\boldsymbol{\alpha}) = \frac{R_i(\boldsymbol{\alpha})}{R(\boldsymbol{\alpha})},$$

namely, the systematic risk of an asset is the marginal contribution of the asset to total risk, scaled by total risk. This measure satisfies Axioms 2, 3, and 4 but it fails Axiom 1, so it cannot be considered as a generalization of the traditional beta.

Finally, it is worth noting that (20) can also be written as

$$\mathcal{B}_i^R(\boldsymbol{\alpha}) = \frac{\left. \frac{d}{dt} \right|_{t=0} R(\boldsymbol{\alpha} + t\boldsymbol{\varepsilon}^i)}{\left. \frac{d}{dt} \right|_{t=0} R(\boldsymbol{\alpha} + t\boldsymbol{\alpha})},$$

where $\boldsymbol{\varepsilon}^i$ is an n -dimensional vector equal to 1 at the i^{th} dimension and zero elsewhere. Namely, when the risk measure is homogeneous, systematic risk of asset i can be thought of as the directional derivative of total risk along the i^{th} dimension scaled by the derivative along the diagonal in the direction of the portfolio itself.

V. Conclusion

In this paper we generalize the concept of systematic risk to account for a variety of risk characteristics. Our equilibrium approach shows that results attributed to the classic CAPM hold much more broadly. In particular, aspects of the geometry of efficient portfolios, two-fund separation, and the security market line are derived in a setting where risk can account for a variety of attributes. Our axiomatic approach specifies four economically meaningful conditions that pin down a unique measure of systematic risk. Both approaches lead to similar generalizations of the traditional beta.

When risk is confined to measure the variance of a distribution, our systematic risk measures coincide with the traditional beta, the slope from regressing asset returns on portfolio returns. More generally, systematic risk is not a regression coefficient. Our equilibrium setting leads to the conclusion that systematic risk is simply the marginal contribution of the asset to the risk of the portfolio of interest, scaled by the weighted average of all such marginal contributions. An identical result is obtained in the axiomatic approach for homogeneous risk measures. When the risk

measure is not homogeneous, the axiomatic approach gives rise to an expression for systematic risk that involves averaging marginal contributions of the asset along a diagonal from the origin to the portfolio of interest.

Our axiomatic approach applies to a wide variety of risk measures, requiring of them only smoothness and zero risk for zero investment. The equilibrium framework imposes additional conditions in the form of scaling independence, convexity, and the risk-free property. Nevertheless, even in the equilibrium framework we are still left with an extensive class of risk measures. Indeed, this class is sufficiently broad to potentially account for high distribution moments, downside risk, rare disasters, and other aspects of risk. A limitation of our framework is that we restrict all investors to use the same risk measure. Future research may direct at developing weaker conditions on the risk measures and introducing more heterogeneity to investor risk preferences.

Finally, our approach is agnostic regarding the choice of a particular risk measure. Indeed, which risk measures better capture the risk preferences of investors is ultimately an empirical question. Our framework therefore provides foundations for testing the appropriateness of risk measures and consequently selecting those that are supported by the data.

Appendix A: Proofs of Main Theorems

Proof of Theorem 1: Our setting is a special case of the setting in Nielsen (1989). To show the existence of equilibrium Nielsen requires that preferences satisfy the following three conditions: (i) each investor's choice set is closed and convex, and contains her initial endowment; (ii) The set of $\{\zeta \in \mathcal{R}^{n+1} : U^j(\zeta) \geq U^j(\zeta')\}$ is closed for all $\zeta' \in \mathcal{R}^{n+1}$ and for all $j = \{1, \dots, \ell\}$; (iii) If $\zeta, \zeta' \in \mathcal{R}^{n+1}$ and $U^j(\zeta') > U^j(\zeta)$, then $U^j(t\zeta' + (1-t)\zeta) > U^j(\zeta)$ for all t in $(0, 1)$.

Condition (i) is satisfied in our setting since the choice set of each investor is \mathcal{R}^{n+1} , which is closed and convex, and contains e^j for all j . Condition (ii) holds since V is assumed continuous and R is assumed smooth, and so their composition is continuous. Condition (iii) follows since $V(\cdot)$ is quasi-concave, strictly increasing in its first argument and strictly decreasing in its second argument, and $R(\cdot)$ is a convex risk measure.

Given these properties of the preferences, Nielsen (1989) establishes two conditions as sufficient for the existence of a quasi-equilibrium: (i) positive semiindependence of directions of improvement, and (ii) nonsatiation at Pareto attainable portfolios. Condition (i) follows in our setting as in Nielsen (1990, Proposition 1) since in our setting all investors agree on all parameters of the problem (in particular on the expected returns), and due to the nonredundancy of risky assets assumption. To see why condition (ii) holds in our setting note that we assume the existence of a risk-free asset paying a nonzero payoff with probability 1. Since $R(\cdot)$ satisfies the risk-free property, we have that $R(\tilde{z}_1 + \tilde{z}_2) \leq R(\tilde{z}_1)$ whenever \tilde{z}_2 is risk-free with $P(\{\tilde{z}_2 > 0\}) = 1$. Thus, adding a positive risk-free asset can only (weakly) reduce risk. It follows that we can always add this positive risk-free asset to any bundle ζ , strictly increasing the expected return while weakly decreasing risk. This implies that in our model there is no satiation globally. Thus, a quasi-equilibrium exists in our setting. Moreover, any quasi-equilibrium is, in fact, an equilibrium in our setting. This follows from the conditions in Nielsen (1989 p. 469). Indeed, in our setting each investor's choice set is convex and unbounded, and the set $\{\zeta \in \mathcal{R}^{n+1} : U^j(\zeta) > U^j(\zeta')\}$ is open for all j and $\zeta' \in \mathcal{R}^{n+1}$. ■

Proof of Theorem 2: Suppose that the equilibrium bundle of investor j is ζ^j . Let

$\bar{x}^j = \sum_{i=0}^n x_i^j = \mathbf{p} \cdot \boldsymbol{\zeta}^j$ be the total dollar amount of investment of investor j . Then,

$$\begin{aligned}
\text{(A1)} \quad U^j(\boldsymbol{\zeta}^j) &= V^j \left(\mathbb{E} \left(\sum_{i=0}^n \zeta_i^j \tilde{y}_i \right), R \left(\sum_{i=0}^n \zeta_i^j \tilde{y}_i \right) \right) \\
&= V^j \left(\bar{x}^j \mathbb{E} \left(\sum_{i=0}^n \frac{\zeta_i^j}{\bar{x}^j} \tilde{y}_i \right), R \left(\bar{x}^j \sum_{i=0}^n \frac{\zeta_i^j}{\bar{x}^j} \tilde{y}_i \right) \right) \\
&= V^j \left(\bar{x}^j \mathbb{E} \left(\sum_{i=0}^n \frac{\zeta_i^j p_i}{\bar{x}^j} \frac{\tilde{y}_i}{p_i} \right), R \left(\bar{x}^j \sum_{i=0}^n \frac{\zeta_i^j p_i}{\bar{x}^j} \frac{\tilde{y}_i}{p_i} \right) \right) \\
&= V^j \left(\bar{x}^j \mathbb{E} \left(\sum_{i=0}^n \frac{x_i^j}{\bar{x}^j} \tilde{z}_i \right), R \left(\bar{x}^j \sum_{i=0}^n \frac{x_i^j}{\bar{x}^j} \tilde{z}_i \right) \right) \\
&= V^j \left(\bar{x}^j \mathbb{E} \left(\sum_{i=0}^n \alpha_i^j \tilde{z}_i \right), R \left(\bar{x}^j \sum_{i=0}^n \alpha_i^j \tilde{z}_i \right) \right) \\
&= V^j \left(\bar{x}^j \mathbb{E} \left(\boldsymbol{\alpha}^j \cdot \tilde{\mathbf{z}} \right), R \left(\bar{x}^j \left(\boldsymbol{\alpha}^j \cdot \tilde{\mathbf{z}} \right) \right) \right).
\end{aligned}$$

From the definition of equilibrium, each investor chooses $\boldsymbol{\zeta}^j$ to maximize $U^j(\boldsymbol{\zeta}^j)$ subject to $\bar{x}^j \leq \mathbf{p} \cdot \mathbf{e}^j$, where by the positivity of prices $\bar{x}^j = \mathbf{p} \cdot \mathbf{e}^j > 0$ (using that $\mathbf{e}^j \in \mathcal{R}_+^{n+1}$ is not zero by assumption). From (A1) and since V^j is strictly increasing in the first argument and strictly decreasing in the second argument, we have that for any positive \bar{x}^j , $U^j(\boldsymbol{\zeta}^j)$ is strictly increasing in $\mathbb{E}(\boldsymbol{\alpha}^j \cdot \tilde{\mathbf{z}})$ and strictly decreasing in $R(\bar{x}^j(\boldsymbol{\alpha}^j \cdot \tilde{\mathbf{z}}))$. Therefore, in equilibrium, $\boldsymbol{\alpha}^j$ must minimize $R(\bar{x}^j(\boldsymbol{\alpha} \cdot \tilde{\mathbf{z}}))$ for a given level of expected return $\mathbb{E}(\boldsymbol{\alpha}^j \cdot \tilde{\mathbf{z}})$. By scaling independence, this is equivalent to minimizing $R(\boldsymbol{\alpha} \cdot \tilde{\mathbf{z}})$ for a given level of expected return, and thus, to solving Problem (7). The solution is unique since we assumed that $R(\cdot)$ is a convex risk measure, and so $R(\boldsymbol{\alpha} \cdot \tilde{\mathbf{z}})$ is convex as a function of $\boldsymbol{\alpha}$. ■

Proof of Theorem 4: By the smoothness of $R(\cdot)$ and by Theorem 1, the solution to Problem (7) for some $\mu^j = \mu$ is determined by the first order conditions. To solve this program, form the Lagrangian

$$\mathcal{L}(\boldsymbol{\alpha}) = R(\boldsymbol{\alpha}) - \xi \left(\sum_{i=1}^n \alpha_i \mathbb{E}(\tilde{z}_i) + \left(1 - \sum_{i=1}^n \alpha_i \right) r_f - \mu \right),$$

where ξ is a Lagrange multiplier. Equivalently,

$$\mathcal{L}(\boldsymbol{\alpha}) = R \left(1 - \sum_{i=1}^n \alpha_i, \alpha_1, \dots, \alpha_n \right) - \xi \left(\sum_{i=1}^n \alpha_i \mathbb{E}(\tilde{z}_i) + \left(1 - \sum_{i=1}^n \alpha_i \right) r_f - \mu \right).$$

The first order condition states that for all $i = 1, \dots, n$,

$$(A2) \quad -R_0(\boldsymbol{\alpha}^*) + R_i(\boldsymbol{\alpha}^*) - \xi(\mathbb{E}(\tilde{z}_i) - r_f) = 0,$$

where $\boldsymbol{\alpha}^*$ is any efficient portfolio (the market portfolio being a special case). By the risk-free property, $R_0(\boldsymbol{\alpha}^*) = 0$. Hence,

$$(A3) \quad R_i(\boldsymbol{\alpha}^*) = \xi(\mathbb{E}(\tilde{z}_i) - r_f).$$

It follows that

$$(A4) \quad \begin{aligned} \sum_{i=1}^n \alpha_i^* R_i(\boldsymbol{\alpha}^*) &= \xi \sum_{i=1}^n \alpha_i^* (\mathbb{E}(\tilde{z}_i) - r_f) \\ &= \xi(\mathbb{E}(\tilde{\mathbf{z}} \cdot \boldsymbol{\alpha}^*) - r_f). \end{aligned}$$

From (A3) and (A4) we obtain

$$\frac{R_i(\boldsymbol{\alpha}^*)}{\sum_{h=1}^n \alpha_h^* R_h(\boldsymbol{\alpha}^*)} = \frac{\mathbb{E}(\tilde{z}_i) - r_f}{\mathbb{E}(\tilde{\mathbf{z}} \cdot \boldsymbol{\alpha}^*) - r_f},$$

as required. If R is homogeneous of degree k , then by Euler's homogeneous function theorem and using the risk-free property this is also equivalent to

$$\frac{R_i(\boldsymbol{\alpha}^*)}{kR(\boldsymbol{\alpha}^*)} = \frac{\mathbb{E}(\tilde{z}_i) - r_f}{\mathbb{E}(\tilde{\mathbf{z}} \cdot \boldsymbol{\alpha}^*) - r_f}.$$

■

Proof of Theorem 5: The proof relies on a mapping between risk allocation problems as defined in Section IV.A and cost allocation problems as defined in Billera and Heath (1982, hereafter BH). Specifically, BH define a cost allocation problem of order n as a pair (h, \mathbf{x}) where $h : \mathcal{R}_+^n \rightarrow \mathcal{R}$ is continuously differentiable and $h(\mathbf{0}) = 0$. Since R is smooth and satisfies $R(\mathbf{0}) = 0$ we can view any risk allocation problem, (R, \mathbf{x}) , of order n , as a cost allocation problem as defined in BH by setting $h(\mathbf{x}) = R(\mathbf{x} \cdot \tilde{\mathbf{z}})$. Given this mapping we will use (R, \mathbf{x}) to denote both the risk allocation problem and its corresponding cost allocation problem. BH define a cost allocation procedure as a function assigning each cost allocation problem (R, \mathbf{x}) of order n a vector $\mathbf{c}(R, \mathbf{x}) \in \mathcal{R}^n$. That is, $\mathbf{c}(R, \mathbf{x})$ should be interpreted as the cost allocated to each of the n goods or services.

We can then consider a natural mapping between systematic risk measures as defined in Section IV.A and the BH cost allocation procedures as follows. If $\mathbf{B}^R(\mathbf{x})$ is a systematic risk measure of the risk allocation problem (R, \mathbf{x}) , then

$$(A5) \quad \mathbf{c}(R, \mathbf{x}) = \frac{\mathbf{B}^R(\mathbf{x}) R(\mathbf{x} \cdot \tilde{\mathbf{z}})}{\bar{x}}$$

is a cost allocation procedure for the corresponding cost allocation problem (R, \mathbf{x}) . Namely, risk allocation measures can be viewed as scaled versions of cost allocation procedures for the corresponding problems.

Lemma A1 *If a systematic risk measure $\mathbf{B}^R(\mathbf{x})$ satisfies Axioms 1-4, then the corresponding cost allocation procedure $\mathbf{c}(R, \mathbf{x})$ satisfies Conditions (2.1)-(2.4) in BH.*

It is important to note that Axioms 1-4 and Conditions (2.1)-(2.4) in BH are not equivalent to each other either as a group or individually. Rather, our four axioms as a set are stronger than their four conditions as a set. The proof of this lemma follows from the next four steps.

Step 1. Axiom 1 is satisfied if and only if Condition (2.1) in BH holds. Indeed, $\sum_{i=1}^n \alpha_i \mathcal{B}_i^R(\mathbf{x}) = 1$ is equivalent to $\sum_{i=1}^n (x_i R(\mathbf{x}) \mathcal{B}_i^R(\mathbf{x}) / \bar{x}) = R(\mathbf{x})$, which using (A5) is equivalent to $\sum_{i=1}^n x_i c_i(R, \mathbf{x}) = R(\mathbf{x})$. This is Condition (2.1).

Step 2. Axiom 2 is satisfied if and only if Condition (2.2) in BH holds. Indeed, suppose $R(\cdot) = R^1(\cdot) + R^2(\cdot)$ and

$$\mathcal{B}_i^R(\mathbf{x}) = \frac{R^1(\mathbf{x})}{R(\mathbf{x})} \mathcal{B}_i^{R^1}(\mathbf{x}) + \frac{R^2(\mathbf{x})}{R(\mathbf{x})} \mathcal{B}_i^{R^2}(\mathbf{x}).$$

Then

$$\frac{\mathcal{B}_i^R(\mathbf{x}) R(\mathbf{x})}{\bar{x}} = \frac{R^1(\mathbf{x}) \mathcal{B}_i^{R^1}(\mathbf{x})}{\bar{x}} + \frac{R^2(\mathbf{x}) \mathcal{B}_i^{R^2}(\mathbf{x})}{\bar{x}}.$$

That is,

$$c_i(R, \mathbf{x}) = c_i(R^1, \mathbf{x}) + c_i(R^2, \mathbf{x}),$$

which is Condition (2.2).

Step 3. Axioms 1 and 3 jointly imply Condition (2.3).

Assume that both Axioms 1 and 3 are satisfied and assume that for all $\boldsymbol{\eta} \in \mathcal{R}_+^n$,

$$(A6) \quad R(\boldsymbol{\eta} \cdot \tilde{\mathbf{z}}) = g(\boldsymbol{\eta} \cdot \mathbf{q})$$

for some function $g(\cdot)$ and a nonzero vector $\mathbf{q} \in \mathcal{R}_+^n$. Then, $(\tilde{z}_1, \dots, \tilde{z}_n)$ are R -perfectly correlated.

By Axiom 3 for all $i, j = 1, \dots, n$,

$$(A7) \quad q_j \mathcal{B}_i^R(\mathbf{x}) = q_i \mathcal{B}_j^R(\mathbf{x}),$$

and hence

$$\alpha_i q_j \mathcal{B}_i^R(\mathbf{x}) = \alpha_i q_i \mathcal{B}_j^R(\mathbf{x}).$$

Summing over $i = 1, \dots, n$ gives

$$(A8) \quad q_j \sum_{i=1}^n \alpha_i \mathcal{B}_i^R(\mathbf{x}) = (\boldsymbol{\alpha} \cdot \mathbf{q}) \mathcal{B}_j^R(\mathbf{x}).$$

By Axiom 1 we know that $\sum_{i=1}^n \alpha_i \mathcal{B}_i^R(\mathbf{x}) = 1$. Plugging this into (A8) we have

$$q_j = (\boldsymbol{\alpha} \cdot \mathbf{q}) \mathcal{B}_j^R(\mathbf{x}) \text{ for } j = 1, \dots, n.$$

By (A5), and recalling that $R(\mathbf{x}) \neq 0$,

$$(A9) \quad q_j = (\boldsymbol{\alpha} \cdot \mathbf{q}) \frac{c_j(R, \mathbf{x}) \bar{x}}{R(\mathbf{x})} = (\mathbf{x} \cdot \mathbf{q}) \frac{c_j(R, \mathbf{x})}{R(\mathbf{x})} \text{ for } j = 1, \dots, n.$$

If $\mathbf{x} \cdot \mathbf{q} = 0$ this implies that $q_j = 0$ for all j , contradicting that \mathbf{q} is a nonzero vector. Hence, $\mathbf{x} \cdot \mathbf{q}$ is not zero. We then have

$$(A10) \quad c_j(R, \mathbf{x}) = \frac{q_j R(\mathbf{x})}{(\mathbf{x} \cdot \mathbf{q})} \text{ for all } j = 1, \dots, n.$$

Consider an asset with return $\tilde{w} = (\mathbf{x} \cdot \tilde{\mathbf{z}}) / (\mathbf{x} \cdot \mathbf{q})$. Namely, investing $\mathbf{x} \cdot \mathbf{q}$ dollars in this asset yields the same return as of the portfolio \mathbf{x} . Then,

$$R((\mathbf{x} \cdot \mathbf{q}) \tilde{w}) = R(\mathbf{x} \cdot \tilde{\mathbf{z}}) = g(\mathbf{x} \cdot \mathbf{q}).$$

Consider now the risk allocation problem of order 1 with the single asset \tilde{w} held at the amount $\mathbf{x} \cdot \mathbf{q}$. By Axiom 1 the systematic risk measure of this asset must satisfy

$$\mathcal{B}^R(\mathbf{x} \cdot \mathbf{q}) = 1,$$

or equivalently using (A5),

$$c(g, \mathbf{x} \cdot \mathbf{q}) = \frac{R((\mathbf{x} \cdot \mathbf{q}) \tilde{w})}{\mathbf{x} \cdot \mathbf{q}} = \frac{g(\mathbf{x} \cdot \mathbf{q})}{\mathbf{x} \cdot \mathbf{q}}.$$

Plugging back into (A10) and using that $R(\mathbf{x}) = g(\mathbf{x} \cdot \mathbf{q})$ we have

$$c_j(R, \mathbf{x}) = c(g, \mathbf{x} \cdot \mathbf{q}) q_j.$$

This is exactly what Condition (2.3) in BH requires, restricting attention to the case that \mathbf{q} is a nonzero vector of nonnegative integers.

Step 4. Axiom 4 is satisfied if and only if Condition (2.4) holds. This follows directly from (A5) and the definition of R -positive correlation.

Having established Lemma A1 we now turn to completing the proof of the theorem. First, existence has been proved in the text by showing that (19) satisfies Axioms 1-4. To show uniqueness note that Lemma A1 implies that Axioms 1-4 are jointly stronger than Conditions (2.1)-(2.4) in BH. From BH's main result we know that there is a unique cost allocation procedure $\mathbf{c}(R, \mathbf{x})$ satisfying Conditions (2.1)-(2.4). It follows (using the mapping (A5)) that there is a unique systematic risk measure satisfying Axioms 1-4. Thus, the unique systematic risk measure is given by (19).

Finally, to see that (19) and (20) are equivalent when R is homogeneous of degree k , note first that in this case

$$\begin{aligned} \int_0^1 R_i(tx_1, \dots, tx_n) dt &= R_i(x_1, \dots, x_n) \int_0^1 t^{k-1} dt \\ &= \frac{R_i(x_1, \dots, x_n)}{k}, \end{aligned}$$

where the first equality follows since R_i is homogeneous of degree $k - 1$. It follows that

$$\begin{aligned}
 \mathcal{B}_i^R(\mathbf{x}) &= \frac{\bar{x} \int_0^1 R_i(tx_1, \dots, tx_n) dt}{R(x_1, \dots, x_n)} \\
 &= \frac{\bar{x} R_i(x_1, \dots, x_n)}{k R(x_1, \dots, x_n)} \\
 &= \frac{\bar{x} R_i(\bar{x}\alpha_1, \dots, \bar{x}\alpha_n)}{k R(\bar{x}\alpha_1, \dots, \bar{x}\alpha_n)} \\
 &= \frac{R_i(\alpha_1, \dots, \alpha_n)}{k R(\alpha_1, \dots, \alpha_n)} \\
 &= \frac{R_i(\alpha_1, \dots, \alpha_n)}{\sum_{h=1}^n \alpha_h R_h(\alpha_1, \dots, \alpha_n)},
 \end{aligned}$$

where the penultimate equality follows from the homogeneity of degrees k and $k - 1$ of R and R_i respectively, and the last equality follows from Euler's homogeneous function theorem. This completes the proof of Theorem 5. ■

Appendix B: Other Proofs and Derivations

Proofs of Propositions

Proof of Proposition 1: We need to show that for any random returns \tilde{z}_1 and \tilde{z}_2 , and any $0 < \lambda < 1$,

$$(A11) \quad w_k(\lambda \tilde{z}_1 + (1 - \lambda) \tilde{z}_2) \leq \lambda w_k(\tilde{z}_1) + (1 - \lambda) w_k(\tilde{z}_2).$$

Letting $\hat{z}_1 = \tilde{z}_1 - \mathbb{E}(\tilde{z}_1)$ and $\hat{z}_2 = \tilde{z}_2 - \mathbb{E}(\tilde{z}_2)$, (A11) can be rewritten as

$$(A12) \quad \left(\mathbb{E} \left[(\lambda \hat{z}_1 + (1 - \lambda) \hat{z}_2)^k \right] \right)^{\frac{1}{k}} \leq \lambda \left(\mathbb{E} [\hat{z}_1^k] \right)^{\frac{1}{k}} + (1 - \lambda) \left(\mathbb{E} [\hat{z}_2^k] \right)^{\frac{1}{k}}.$$

Applying the binomial formula to the LHS of (A12) implies that we need to show

$$\left(\sum_{i=0}^k \binom{k}{i} \lambda^{k-i} (1 - \lambda)^i \mathbb{E} (\hat{z}_1^{k-i} \hat{z}_2^i) \right)^{\frac{1}{k}} \leq \lambda \left(\mathbb{E} [\hat{z}_1^k] \right)^{\frac{1}{k}} + (1 - \lambda) \left(\mathbb{E} [\hat{z}_2^k] \right)^{\frac{1}{k}}.$$

Since k is even, replacing each \hat{z}_1 and \hat{z}_2 with $|\hat{z}_1|$ and $|\hat{z}_2|$ will not affect the RHS, but it might increase the LHS. So, it is sufficient to show that

$$\left(\sum_{i=0}^k \binom{k}{i} \lambda^{k-i} (1 - \lambda)^i \mathbb{E} (|\hat{z}_1^{k-i} \hat{z}_2^i|) \right)^{\frac{1}{k}} \leq \lambda \left(\mathbb{E} [|\hat{z}_1|^k] \right)^{\frac{1}{k}} + (1 - \lambda) \left(\mathbb{E} [|\hat{z}_2|^k] \right)^{\frac{1}{k}}.$$

Since both sides are positive we can raise both sides to the k^{th} power, maintaining the inequality. Thus, it would be sufficient to show

$$\sum_{i=0}^k \binom{k}{i} \lambda^{k-i} (1 - \lambda)^i \mathbb{E} (|\hat{z}_1^{k-i} \hat{z}_2^i|) \leq \left(\lambda \left(\mathbb{E} [|\hat{z}_1|^k] \right)^{\frac{1}{k}} + (1 - \lambda) \left(\mathbb{E} [|\hat{z}_2|^k] \right)^{\frac{1}{k}} \right)^k.$$

Applying the binomial formula to the RHS implies that it would be sufficient to show

$$\sum_{i=0}^k \binom{k}{i} \lambda^{k-i} (1 - \lambda)^i \mathbb{E} (|\hat{z}_1^{k-i} \hat{z}_2^i|) \leq \sum_{i=0}^k \binom{k}{i} \lambda^{k-i} (1 - \lambda)^i \left(\mathbb{E} [|\hat{z}_1|^k] \right)^{\frac{k-i}{k}} \left(\mathbb{E} [|\hat{z}_2|^k] \right)^{\frac{i}{k}}.$$

To establish this inequality we will show that it actually holds term by term. That is, it is sufficient to show that for each $i = 0, \dots, k$,

$$\mathbb{E} (|\hat{z}_1^{k-i} \hat{z}_2^i|) \leq \left(\mathbb{E} \left[|\hat{z}_1|^k \right] \right)^{\frac{k-i}{k}} \left(\mathbb{E} \left[|\hat{z}_2|^k \right] \right)^{\frac{i}{k}}.$$

To see this, note that it is equivalent to show that

$$\mathbb{E} (|\hat{z}_1^{k-i} \hat{z}_2^i|) \leq \left(\mathbb{E} \left[|\hat{z}_1^{k-i}|^{\frac{k}{k-i}} \right] \right)^{\frac{k-i}{k}} \left(\mathbb{E} \left[|\hat{z}_2^i|^{\frac{k}{i}} \right] \right)^{\frac{i}{k}}.$$

But, this is immediate from Hölder's inequality, and we are done. ■

Proof of Proposition 2: For any integer $k \geq 2$, we can rewrite the downside risk measure as

$$\text{DR}_k(\tilde{z}) = (-1)^k \left(\mathbb{E} \left([\tilde{z} - \mathbb{E}(\tilde{z})]^- \right)^k \right)^{\frac{1}{k}} = \left(\mathbb{E} \left([\mathbb{E}(\tilde{z}) - \tilde{z}]^+ \right)^k \right)^{\frac{1}{k}},$$

where $[t]^+ = \max(t, 0)$ for $t \in \mathcal{R}$.

Consider any two random returns \tilde{z}_1 and \tilde{z}_2 , and let $\hat{z}_1 = [\mathbb{E}(\tilde{z}_1) - \tilde{z}_1]^+$ and $\hat{z}_2 = [\mathbb{E}(\tilde{z}_2) - \tilde{z}_2]^+$. Obviously, we have $\hat{z}_1 \geq 0$ and $\hat{z}_2 \geq 0$. What we need to show is that for any $0 < \lambda < 1$,

$$\left(\mathbb{E} \left([\mathbb{E}(\lambda \tilde{z}_1 + (1-\lambda)\tilde{z}_2) - \lambda \tilde{z}_1 - (1-\lambda)\tilde{z}_2]^+ \right)^k \right)^{\frac{1}{k}} \leq \lambda \left(\mathbb{E}(\hat{z}_1^k) \right)^{\frac{1}{k}} + (1-\lambda) \left(\mathbb{E}(\hat{z}_2^k) \right)^{\frac{1}{k}}.$$

Now,

$$\begin{aligned} [\mathbb{E}(\lambda \tilde{z}_1 + (1-\lambda)\tilde{z}_2) - \lambda \tilde{z}_1 - (1-\lambda)\tilde{z}_2]^+ &= [\lambda(\mathbb{E}(\tilde{z}_1) - \tilde{z}_1) + (1-\lambda)(\mathbb{E}(\tilde{z}_2) - \tilde{z}_2)]^+ \\ &\leq \lambda[\mathbb{E}(\tilde{z}_1) - \tilde{z}_1]^+ + (1-\lambda)[\mathbb{E}(\tilde{z}_2) - \tilde{z}_2]^+ \\ &= \lambda \hat{z}_1 + (1-\lambda) \hat{z}_2, \end{aligned}$$

where the inequality follows from Jensen's inequality using that $[\cdot]^+$ is a convex function.

Therefore, it is sufficient to show that

$$\left(\mathbb{E}(\lambda \hat{z}_1 + (1-\lambda)\hat{z}_2)^k \right)^{\frac{1}{k}} \leq \lambda \left(\mathbb{E}(\hat{z}_1^k) \right)^{\frac{1}{k}} + (1-\lambda) \left(\mathbb{E}(\hat{z}_2^k) \right)^{\frac{1}{k}}.$$

The rest of the proof follows closely the proof of Proposition 1. Indeed, since \hat{z}_1 and \hat{z}_2 are nonnegative here, the arguments in the proof of Proposition 1 apply in this case to any positive k (odd or even). ■

Proof of Proposition 3: In the definition of expected shortfall we assumed the existence of a cumulative distribution function $F(\cdot)$ applied to realizations of random variables. For the sake of this proof it will be more useful to work directly with the state space Ω and with the underlying probability measure $P(\cdot)$. We first prove that $\text{ES}_\delta(\tilde{z})$ is subadditive. That is, for any two random returns \tilde{z}_1 and \tilde{z}_2 ,

$$(A13) \quad \text{ES}_\delta(\tilde{z}_1 + \tilde{z}_2) \leq \text{ES}_\delta(\tilde{z}_1) + \text{ES}_\delta(\tilde{z}_2).$$

If either \tilde{z}_1 or \tilde{z}_2 is equal to a constant with probability 1, then the result is immediate. We shall thus only consider the case in which both of them are not equal to a constant. By (3), for any random return \tilde{z} (which is not constant), $\text{ES}_\delta(\tilde{z})$ can be expressed as

$$\text{ES}_\delta(\tilde{z}) = -\frac{1}{\delta} \int_{\{\omega: \tilde{z} \leq -\text{VaR}_\delta(\tilde{z})\}} \tilde{z} dP(\omega).$$

Let \tilde{z}_1 and \tilde{z}_2 be random returns and define $\tilde{z}_3 = \tilde{z}_1 + \tilde{z}_2$. Let $\Omega_i = \{\omega \in \Omega : \tilde{z}_i \leq -\text{VaR}_\delta(\tilde{z}_i)\}$ for $i = 1, 2, 3$. Then, (A13) is equivalent to

$$\int_{\Omega_3} \tilde{z}_3 dP(\omega) \geq \int_{\Omega_1} \tilde{z}_1 dP(\omega) + \int_{\Omega_2} \tilde{z}_2 dP(\omega),$$

which can be rewritten as

$$\int_{\Omega_3} \tilde{z}_1 dP(\omega) + \int_{\Omega_3} \tilde{z}_2 dP(\omega) \geq \int_{\Omega_1} \tilde{z}_1 dP(\omega) + \int_{\Omega_2} \tilde{z}_2 dP(\omega).$$

It is sufficient to show that

$$(A14) \quad \int_{\Omega_3} \tilde{z}_1 dP(\omega) \geq \int_{\Omega_1} \tilde{z}_1 dP(\omega),$$

and

$$(A15) \quad \int_{\Omega_3} \tilde{z}_2 dP(\omega) \geq \int_{\Omega_2} \tilde{z}_2 dP(\omega).$$

For brevity, we will only prove (A14). The proof of (A15) is parallel.

Define

$$\begin{aligned} \Omega_4 &= \{\omega \in \Omega : \tilde{z}_1 \leq -\text{VaR}_\delta(\tilde{z}_1), \tilde{z}_3 \leq -\text{VaR}_\delta(\tilde{z}_3)\}, \\ \Omega_5 &= \{\omega \in \Omega : \tilde{z}_1 \leq -\text{VaR}_\delta(\tilde{z}_1), \tilde{z}_3 > -\text{VaR}_\delta(\tilde{z}_3)\}, \text{ and} \\ \Omega_6 &= \{\omega \in \Omega : \tilde{z}_1 > -\text{VaR}_\delta(\tilde{z}_1), \tilde{z}_3 \leq -\text{VaR}_\delta(\tilde{z}_3)\}. \end{aligned}$$

Clearly, $\Omega_4 \cap \Omega_5 = \emptyset$, $\Omega_4 \cup \Omega_5 = \Omega_1$, $\Omega_4 \cap \Omega_6 = \emptyset$, and $\Omega_4 \cup \Omega_6 = \Omega_3$. Thus,

$$\int_{\Omega_1} dP(\omega) = \int_{\Omega_4} dP(\omega) + \int_{\Omega_5} dP(\omega),$$

and

$$\int_{\Omega_3} dP(\omega) = \int_{\Omega_4} dP(\omega) + \int_{\Omega_6} dP(\omega).$$

By the definition of VaR, we know

$$\int_{\Omega_1} dP(\omega) = \int_{\Omega_3} dP(\omega) = \delta.$$

Thus, we obtain

$$(A16) \quad \int_{\Omega_5} dP(\omega) = \int_{\Omega_6} dP(\omega).$$

Similarly, we have

$$\int_{\Omega_1} \tilde{z}_1 dP(\omega) = \int_{\Omega_4} \tilde{z}_1 dP(\omega) + \int_{\Omega_5} \tilde{z}_1 dP(\omega),$$

and

$$\int_{\Omega_3} \tilde{z}_1 dP(\omega) = \int_{\Omega_4} \tilde{z}_1 dP(\omega) + \int_{\Omega_6} \tilde{z}_1 dP(\omega).$$

Hence,

$$\begin{aligned} & \int_{\Omega_1} \tilde{z}_1 dP(\omega) - \int_{\Omega_3} \tilde{z}_1 dP(\omega) \\ &= \int_{\Omega_5} \tilde{z}_1 dP(\omega) - \int_{\Omega_6} \tilde{z}_1 dP(\omega) \\ &\leq \int_{\Omega_5} [-\text{VaR}_\delta(\tilde{z}_1)] dP(\omega) - \int_{\Omega_6} [-\text{VaR}_\delta(\tilde{z}_1)] dP(\omega) \\ &= -\text{VaR}_\delta(\tilde{z}_1) \left[\int_{\Omega_5} dP(\omega) - \int_{\Omega_6} dP(\omega) \right] \\ &= 0, \end{aligned}$$

where the inequality follows from $\tilde{z}_1 \leq -\text{VaR}_\delta(\tilde{z}_1)$ when $\omega \in \Omega_5$ and $\tilde{z}_1 > -\text{VaR}_\delta(\tilde{z}_1)$ when $\omega \in \Omega_6$, and where the last equality follows from (A16). Therefore, (A14) is obtained, and hence $\text{ES}_\delta(\tilde{z})$ is subadditive. Since $\text{DES}_\delta(\tilde{z}) = \text{ES}_\delta(\tilde{z}) + \text{E}(\tilde{z})$ we have that DES is also subadditive.

Convexity now follows immediately from homogeneity of degree 1 and subadditivity. ■

Derivations of Systematic Risk for Applications I–V

Here we provide derivations of the systematic risk associated with different risk measures discussed in Section III.C.

Application I: This is a special case of Application II.

Application II: Consider the risk measure $R(\tilde{z}) = m_k(\tilde{z})$ for even $k \geq 2$. The risk of the market portfolio is

$$R(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}}) = m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}}) = \mathbb{E}(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}} - \mathbb{E}(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}}))^k.$$

Differentiating with respect to the weight of asset i yields

$$\begin{aligned} \frac{\partial m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})}{\partial \alpha_i^M} &= k \mathbb{E} \left[(\tilde{z}_i - \mathbb{E}(\tilde{z}_i)) (\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}} - \mathbb{E}(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}}))^{k-1} \right] \\ &= k \text{Cov} \left(\tilde{z}_i, (\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}} - \boldsymbol{\alpha}^M \cdot \mathbb{E}(\tilde{\mathbf{z}}))^{k-1} \right). \end{aligned}$$

By Theorem 4, and since $m_k(\cdot)$ is homogeneous of degree k , the systematic risk is then given by

$$(A17) \quad \mathcal{B}_i^R = \frac{\frac{\partial m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})}{\partial \alpha_i^M}}{k m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})} = \frac{\text{Cov} \left(\tilde{z}_i, (\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}} - \boldsymbol{\alpha}^M \cdot \mathbb{E}(\tilde{\mathbf{z}}))^{k-1} \right)}{m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})},$$

as required.

Now suppose alternatively that $R(\tilde{z}) = w_k(\tilde{z})$. The market portfolio risk is

$$R(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}}) = w_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}}) = (m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}}))^{\frac{1}{k}}.$$

Differentiating with respect to the weight of asset i gives

$$\frac{\partial w_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})}{\partial \alpha_i^M} = \frac{1}{k} (m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}}))^{\frac{1}{k}-1} \frac{\partial m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})}{\partial \alpha_i^M}.$$

By Theorem 4, and since $w_k(\cdot)$ is homogeneous of degree 1, the systematic risk is

$$\mathcal{B}_i^R = \frac{\frac{1}{k} (m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}}))^{\frac{1}{k}-1} \frac{\partial m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})}{\partial \alpha_i^M}}{(m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}}))^{\frac{1}{k}}} = \frac{\frac{\partial m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})}{\partial \alpha_i^M}}{k m_k(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})},$$

which is identical to (A17).

Application III: Assume $R(\tilde{z}) = \text{DR}_k(\tilde{z})$ for $k \geq 2$. The risk of the market portfolio α^M is given by

$$\text{DR}_k(\alpha^M \cdot \tilde{\mathbf{z}}) = (-1)^k \left(\text{E} \left([\alpha^M \cdot \tilde{\mathbf{z}} - \text{E}(\alpha^M \cdot \tilde{\mathbf{z}})]^- \right)^k \right)^{\frac{1}{k}}.$$

Differentiating with respect to α_i^M gives²¹

$$\begin{aligned} & \frac{\partial \text{DR}_k(\alpha^M \cdot \tilde{\mathbf{z}})}{\partial \alpha_i^M} \\ &= (-1)^k \left(\text{E} \left([\alpha^M \cdot \tilde{\mathbf{z}} - \text{E}(\alpha^M \cdot \tilde{\mathbf{z}})]^- \right)^k \right)^{\frac{1}{k}-1} \text{E} \left[(\tilde{z}_i - \text{E}(\tilde{z}_i)) \left([\alpha^M \cdot \tilde{\mathbf{z}} - \text{E}(\alpha^M \cdot \tilde{\mathbf{z}})]^- \right)^{k-1} \right] \\ &= (-1)^k \left(\text{E} \left([\alpha^M \cdot \tilde{\mathbf{z}} - \text{E}(\alpha^M \cdot \tilde{\mathbf{z}})]^- \right)^k \right)^{\frac{1}{k}-1} \text{Cov} \left[\tilde{z}_i, \left([\alpha^M \cdot \tilde{\mathbf{z}} - \text{E}(\alpha^M \cdot \tilde{\mathbf{z}})]^- \right)^{k-1} \right]. \end{aligned}$$

By Theorem 4, and since $\text{DR}_k(\cdot)$ is homogeneous of degree 1, the systematic risk is given by

$$\begin{aligned} \mathcal{B}_i^R &= \frac{\frac{\partial \text{DR}_k(\alpha^M \cdot \tilde{\mathbf{z}})}{\partial \alpha_i^M}}{\text{DR}_k(\alpha^M \cdot \tilde{\mathbf{z}})} \\ &= \frac{\text{Cov} \left[\tilde{z}_i, \left([\alpha^M \cdot \tilde{\mathbf{z}} - \text{E}(\alpha^M \cdot \tilde{\mathbf{z}})]^- \right)^{k-1} \right]}{\text{E} \left([\alpha^M \cdot \tilde{\mathbf{z}} - \text{E}(\alpha^M \cdot \tilde{\mathbf{z}})]^- \right)^k} \\ &= (-1)^k \frac{\text{Cov} \left[\tilde{z}_i, \left([\alpha^M \cdot \tilde{\mathbf{z}} - \text{E}(\alpha^M \cdot \tilde{\mathbf{z}})]^- \right)^{k-1} \right]}{(\text{DR}_k(\alpha^M \cdot \tilde{\mathbf{z}}))^k}. \end{aligned}$$

Application IV: Assume $R(\tilde{z}) = \text{DES}_\delta(\tilde{z})$ for some confidence level $0 < \delta < 1$. Let $f(z_1, \dots, z_n)$ denote the joint density function of $\tilde{\mathbf{z}}$. Since all risky assets have

²¹Note that we are essentially relying here on Leibniz's rule for differentiation under the integral. While $\left([\alpha^M \cdot \mathbf{z} - \text{E}(\alpha^M \cdot \tilde{\mathbf{z}})]^- \right)^k$ is not everywhere differentiable, it is continuous and differentiable almost everywhere. This guarantees that Leibniz's rule applies.

positive net supply and since asset prices are positive, we have $\alpha_1^M > 0$. Hence, the risk of the market portfolio α^M can be written as follows

$$\begin{aligned} & \text{DES}_\delta(\alpha^M \cdot \tilde{\mathbf{z}}) \\ &= \text{ES}_\delta(\alpha^M \cdot \tilde{\mathbf{z}}) + \text{E}(\alpha^M \cdot \tilde{\mathbf{z}}) \\ &= -\frac{1}{\delta} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{-\text{VaR}_\delta(\alpha^M \cdot \tilde{\mathbf{z}}) - \sum_{j=2}^n \alpha_j^M z_j}{\alpha_1^M} \left(\sum_{j=1}^n \alpha_j^M z_j - \text{E}(\alpha^M \cdot \tilde{\mathbf{z}}) \right) f(z_1, \dots, z_n) dz_1 \dots dz_n. \end{aligned}$$

Differentiating $\text{DES}_\delta(\alpha^M \cdot \tilde{\mathbf{z}})$ using Leibniz's rule with respect to α_i^M yields

$$\begin{aligned} \text{(A18)} \quad & \frac{\partial \text{DES}_\delta(\alpha^M \cdot \tilde{\mathbf{z}})}{\partial \alpha_i^M} \\ &= -\frac{1}{\delta} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{-\text{VaR}_\delta(\alpha^M \cdot \tilde{\mathbf{z}}) - \sum_{j=2}^n \alpha_j^M z_j}{\alpha_1^M} (z_i - \text{E}(\tilde{z}_i)) f(z_1, \dots, z_n) dz_1 \dots dz_n \\ & \quad + \frac{\text{VaR}_\delta(\alpha^M \cdot \tilde{\mathbf{z}}) + \text{E}(\alpha^M \cdot \tilde{\mathbf{z}})}{\delta} \\ & \quad \cdot \frac{\partial}{\partial \alpha_i^M} \left(\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{-\text{VaR}_\delta(\alpha^M \cdot \tilde{\mathbf{z}}) - \sum_{j=2}^n \alpha_j^M z_j}{\alpha_1^M} f(z_1, \dots, z_n) dz_1 \dots dz_n \right). \end{aligned}$$

Notice that by the definition of $\text{VaR}_\delta(\alpha^M \cdot \tilde{\mathbf{z}})$,

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{-\text{VaR}_\delta(\alpha^M \cdot \tilde{\mathbf{z}}) - \sum_{j=2}^n \alpha_j^M z_j}{\alpha_1^M} f(z_1, \dots, z_n) dz_1 \dots dz_n = \delta,$$

which is a constant, implying that the second term in (A18) is zero. Thus,

$$\begin{aligned} \frac{\partial \text{DES}_\delta(\alpha^M \cdot \tilde{\mathbf{z}})}{\partial \alpha_i^M} &= -\frac{1}{\delta} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{-\text{VaR}_\delta(\alpha^M \cdot \tilde{\mathbf{z}}) - \sum_{j=2}^n \alpha_j^M z_j}{\alpha_1^M} (z_i - \text{E}(\tilde{z}_i)) f(z_1, \dots, z_n) dz_1 \dots dz_n \\ &= -\frac{1}{\delta} \text{E} \left[\mathbf{1}_{\alpha^M \cdot \tilde{\mathbf{z}} \leq -\text{VaR}_\delta(\alpha^M \cdot \tilde{\mathbf{z}})} (z_i - \text{E}(\tilde{z}_i)) \right] \\ &= -\text{E} \left[\tilde{z}_i - \text{E}(\tilde{z}_i) \mid \alpha^M \cdot \tilde{\mathbf{z}} \leq -\text{VaR}_\delta(\alpha^M \cdot \tilde{\mathbf{z}}) \right]. \end{aligned}$$

By Theorem 4, and since $\text{DES}_\delta(\alpha^M \cdot \tilde{\mathbf{z}})$ is homogeneous of degree 1, the systematic risk is given by

$$\mathcal{B}_i^R = \frac{\frac{\partial \text{DES}_\delta(\alpha^M \cdot \tilde{\mathbf{z}})}{\partial \alpha_i^M}}{\text{DES}_\delta(\alpha^M \cdot \tilde{\mathbf{z}})} = -\frac{\text{E} \left[\tilde{z}_i - \text{E}(\tilde{z}_i) \mid \alpha^M \cdot \tilde{\mathbf{z}} \leq -\text{VaR}_\delta(\alpha^M \cdot \tilde{\mathbf{z}}) \right]}{\text{DES}_\delta(\alpha^M \cdot \tilde{\mathbf{z}})}.$$

Application V: Consider the following family of risk measures

$$R(\tilde{z}) = \theta_1 w_2(\tilde{z}) + \theta_2 \text{DR}_3(\tilde{z}) + \theta_3 w_4(\tilde{z}) + \theta_4 \text{DES}_\delta(\tilde{z})$$

for some confidence level δ and nonnegative weights $\theta_1, \dots, \theta_4$. From Lemma 1, this family of risk measures satisfies all of the conditions in Theorem 4. Moreover, it is easy to verify that when

$$R(\tilde{z}) = \sum_{j=1}^s R^j(\tilde{z}),$$

the expression for \mathcal{B}_i^R given in (10) implies

$$\mathcal{B}_i^R = \sum_{j=1}^s \frac{R^j(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})}{R(\boldsymbol{\alpha}^M \cdot \tilde{\mathbf{z}})} \mathcal{B}_i^{R^j}.$$

That is, the systematic risk takes the form of the risk-weighted average of the systematic risk associated with each of the risk components. (See also Section IV.A for further discussion of this issue as it relates to Axiom 2 in that section.)

References

- [1] **Acharya, Viral V., Lasse H. Pedersen, Thomas Philippon, and Mathew Richardson.** 2010. “Measuring Systemic Risk.” <http://docs.lhpedersen.com/MeasuringSystemicRisk.pdf>.
- [2] **Adrian, Tobias, and Markus K. Brunnermeier.** 2014. “CoVaR.” http://scholar.princeton.edu/sites/default/files/markus/files/covar_0.pdf.
- [3] **Ang, Andrew, Joseph Chen, and Yuhang Xing.** 2006. “Downside Risk.” *Review of Financial Studies* 19 (4): 1191–1239.
- [4] **Aumann, Robert J., and Roberto Serrano.** 2008. “An Economic Index of Riskiness.” *Journal of Political Economy* 116 (5): 810–836.
- [5] **Aumann, Robert J., and Lloyd S. Shapley.** 1974. *Values of Non-Atomic Games*. Princeton: Princeton University Press.
- [6] **Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath.** 1999. “Coherent Measures of Risk.” *Mathematical Finance* 9 (3): 203–228.
- [7] **Bansal, Ravi, and Amir Yaron.** 2004. “Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles.” *Journal of Finance* 59 (4): 1481–1509.
- [8] **Barro, Robert J.** 2006. “Rare Disasters and Asset Markets in the Twentieth Century.” *Quarterly Journal of Economics* 121 (3): 823–866.
- [9] **Barro, Robert J.** 2009. “Rare Disasters, Asset Prices, and Welfare Costs.” *American Economic Review* 99 (1): 243–264.
- [10] **Billera, Louis J., and David C. Heath.** 1982. “Allocation of Shared Costs: A Set of Axioms Yielding A Unique Procedure.” *Mathematics of Operations Research* 7 (1): 32–39.
- [11] **Billera, Louis J., David C. Heath, and Joseph Raanan.** 1978. “Internal Telephone Billing Rates—A Novel Application of Non-Atomic Game Theory.” *Operations Research* 26 (6): 956–965.

- [12] **Billera, Louis J., David C. Heath, and Robert E. Verrecchia.** 1981. “A Unique Procedure for Allocating Common Costs from A Production Process.” *Journal of Accounting Research* 19 (1): 185–196.
- [13] **Campbell, John Y., and John H. Cochrane.** 1999. “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior.” *Journal of Political Economy* 107 (2): 205–251.
- [14] **Cass, David, and Joseph E. Stiglitz.** 1970. “The Structure of Investor Preferences and Asset Returns, and Separability in Portfolio Allocation: A Contribution to the Pure Theory of Mutual Funds.” *Journal of Economic Theory* 2 (2): 122–160.
- [15] **Chen, Hui, Scott Joslin, and Ngoc-Khanh Tran.** 2012. “Rare Disasters and Risk Sharing with Heterogeneous Beliefs.” *Review of Financial Studies* 25 (7): 2189–2224.
- [16] **Cochrane, John H.** 2001. *Asset Pricing*. Princeton: Princeton University Press.
- [17] **Dybvig, Philip H., and Fang Liu.** 2015. “On Investor Preferences and Mutual Fund Separation.” <http://ssrn.com/abstract=2530186>.
- [18] **Dybvig, Philip H., and Stephen A. Ross.** 1982. “Portfolio Efficient Set.” *Econometrica* 50 (6): 1525–1546.
- [19] **Epstein, Larry G., and Stanley E. Zin.** 1989. “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework.” *Econometrica* 57 (4): 937–969.
- [20] **Fama, Eugene F., and Kenneth R. French.** 1992. “The Cross-Section of Expected Stock Returns.” *Journal of Finance* 47 (2): 427–465.
- [21] **Fama, Eugene F., and James D. MacBeth.** 1973. “Risk, Return, and Equilibrium: Empirical Tests.” *Journal of Political Economy* 81 (3): 607–636.

- [22] **Foster, Dean P., and Sergiu Hart.** 2009. "An Operational Measure of Riskiness." *Journal of Political Economy* 117 (5): 785–814.
- [23] **Foster, Dean P., and Sergiu Hart.** 2013. "A Wealth-Requirement Axiomatization of Riskiness." *Theoretical Economics* 8 (2): 591–620.
- [24] **Gabaix, Xavier.** 2008. "Variable Rare Disasters: A Tractable Theory of Ten Puzzles in Macro-Finance." *American Economic Review* 98 (2): 64–67.
- [25] **Gabaix, Xavier.** 2012. "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance." *Quarterly Journal of Economics* 127 (2): 645–700.
- [26] **Gourio, François.** 2012. "Disaster Risk and Business Cycles." *American Economic Review* 102 (6): 2734–2766.
- [27] **Hadar, Josef, and William R. Russell.** 1969. "Rules for Ordering Uncertain Prospects." *American Economic Review* 59 (1): 25–34.
- [28] **Hanoch, Giora, and Haim Levy.** 1969. "The Efficiency Analysis of Choices Involving Risk." *Review of Economic Studies* 36 (107): 335–346.
- [29] **Hansen, Lars Peter, and Scott F. Richard.** 1987. "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models." *Econometrica* 55 (3): 587–613.
- [30] **Hart, Sergiu.** 2011. "Comparing Risks by Acceptance and Rejection." *Journal of Political Economy* 119 (4): 617–638.
- [31] **Harvey, Campbell R., and Akhtar Siddique.** 2000. "Conditional Skewness in Asset Pricing Tests." *Journal of Finance* 55 (3): 1263–1295.
- [32] **Huang, Xin, Hao Zhou, and Haibin Zhu.** 2009. "A Framework for Assessing the Systemic Risk of Major Financial Institutions." *Journal of Banking and Finance* 33 (11): 2036–2049.

- [33] **Jean, William H.** 1971. “The Extension of Portfolio Analysis to Three or More Parameters.” *Journal of Financial and Quantitative Analysis* 6 (1): 505–515.
- [34] **Kadan, Ohad, and Fang Liu.** 2014. “Performance Evaluation with High Moments and Disaster Risk.” *Journal of Financial Economics* 113 (1): 131–155.
- [35] **Kane, Alex.** 1982. “Skewness Preference and Portfolio Choice.” *Journal of Financial and Quantitative Analysis* 17 (1): 15–25.
- [36] **Kraus, Alan, and Robert H. Litzenberger.** 1976. “Skewness Preference and the Valuation of Risk Assets.” *Journal of Finance* 31 (4): 1085–1100.
- [37] **Lettau, Martin, Matteo Maggiori, and Michael Weber.** 2014. “Conditional Risk Premia in Currency Markets and Other Asset Classes.” *Journal of Financial Economics* 114 (2): 197–225.
- [38] **Levy, Moshe, and Richard Roll.** 2010. “The Market Portfolio May Be Mean/Variance Efficient After All.” *Review of Financial Studies* 23 (6): 2464–2491.
- [39] **Lintner, John.** 1965a. “The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets.” *Review of Economics and Statistics* 47 (1): 13–37.
- [40] **Lintner, John.** 1965b. “Security Prices, Risk, and Maximal Gains from Diversification.” *Journal of Finance* 20 (4): 587–615.
- [41] **Luenberger, David G.** 1969. *Optimization by Vector Space Methods*. Hoboken: John Wiley & Sons.
- [42] **Machina, Mark J.** 1982. “‘Expected Utility’ Analysis without the Independence Axiom.” *Econometrica* 50 (2): 277–323.
- [43] **Mossin, Jan.** 1966. “Equilibrium in a Capital Asset Market.” *Econometrica* 34 (4): 768–783.

- [44] **Müller, Sigrid M., and Mark J. Machina.** 1987. “Moment Preferences and Polynomial Utility.” *Economics Letters* 23 (4): 349–353.
- [45] **Nielsen, Lars T.** 1989. “Asset Market Equilibrium with Short-Selling.” *Review of Economic Studies* 56 (3): 467–473.
- [46] **Nielsen, Lars T.** 1990. “Existence of Equilibrium in CAPM.” *Journal of Economic Theory* 52 (1): 223–231.
- [47] **Nielsen, Lars T.** 1992. “Positive Prices in CAPM.” *Journal of Finance* 47 (2): 791–808.
- [48] **Powers, Michael R.** 2007. “Using Aumann-Shapley Values to Allocate Insurance Risk: The Case of Inhomogeneous Losses.” *North American Actuarial Journal* 11 (3): 113–127.
- [49] **Rockafellar, R. Tyrrell, Stan Uryasev, and Michael Zabarankin.** 2006a. “Generalized Deviations in Risk Analysis.” *Finance and Stochastics* 10 (1): 51–74.
- [50] **Rockafellar, R. Tyrrell, Stan Uryasev, and Michael Zabarankin.** 2006b. “Master Funds in Portfolio Analysis with General Deviation Measures.” *Journal of Banking and Finance* 30 (2): 743–778.
- [51] **Ross, Stephen A.** 1976. “The Arbitrage Theory of Capital Asset Pricing.” *Journal of Economic Theory* 13 (3): 341–360.
- [52] **Ross, Stephen A.** 1978. “Mutual Fund Separation in Financial Theory—The Separating Distributions.” *Journal of Economic Theory* 17 (2): 254–286.
- [53] **Rothschild, Michael, and Joseph E. Stiglitz.** 1970. “Increasing Risk I: A Definition.” *Journal of Economic Theory* 2 (3): 225–243.
- [54] **Rubinstein, Mark.** 1973. “The Fundamental Theorem of Parameter-Preference Security Valuation.” *Journal of Financial and Quantitative Analysis* 8 (1): 61–69.

- [55] **Samet, Dov, Yair Tauman, and Israel Zang.** 1984. “An Application of the Aumann-Shapley Prices for Cost Allocation in Transportation Problems.” *Mathematics of Operations Research* 9 (1): 25–42.
- [56] **Shapley, Lloyd S.** 1952. *A Value for n-Person Games*. Santa Monica: RAND Corporation.
- [57] **Sharpe, William F.** 1964. “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk.” *Journal of Finance* 19 (3): 425–442.
- [58] **Tarashev, Nikola A., Claudio E. V. Borio, and Kostas Tsatsaronis.** 2010. “Attributing Systemic Risk to Individual Institutions.” BIS Working Paper No. 308.
- [59] **Tobin, James.** 1958. “Liquidity Preferences as Behavior Towards Risk.” *Review of Economic Studies* 25 (2): 65–86.
- [60] **Wachter, Jessica A.** 2013. “Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?” *Journal of Finance* 68 (3): 987–1035.

Online Appendix

Mean-Risk Preferences and Expected Utility

Background

One would wonder how the mean-risk preferences considered in Section III are related to the commonly assumed von Neumann-Morgenstern utility. It is widely known that a von Neumann-Morgenstern investor with a quadratic utility function only cares about the mean and the variance of his investments in the sense that he prefers a high expected wealth and a low variance. In this sense, the mean-risk preference is consistent with the von Neumann-Morgenstern utility when variance is used as the risk measure. Alternatively, when returns are distributed according to a two-parameter elliptical distribution (normal being a special case), mean-variance preferences can also be supported by expected utility. These instances, however, are quite restrictive. First, the quadratic utility is not very intuitive since it implies increasing absolute risk aversion. Second, elliptical distributions, being determined by the first two moments only, limit our ability to describe the dependence of risk on high distribution moments and other risk characteristics. Thus, in general, mean-variance preferences are not consistent with expected utility. The approach taken in this paper is much more general, allowing for a variety of risk measures. Whether a particular risk measure is consistent with expected utility depends on the actual choice of the risk measure. For example, risk measures that are simple linear combinations of raw moments up to the k^{th} degree can be represented by a k^{th} degree polynomial (Müller and Machina (1987)), generalizing the mean-variance result.

While in general the preferences defined in (6) cannot be supported by expected utility, they are often consistent with expected utility *locally*. The idea is based on Machina's (1982) "Local Utility Function." To facilitate this approach we first restrict attention to risk measures that depend on the distribution of the random variables only. Thus, we consider risk measures that are functions from the distribution of realizations to the reals rather than functions from the random variables themselves. Practically, this does not present a binding restriction since all the examples in this paper and all standard risk measures only rely on the distribution of

realizations anyway. In this case the preferences in (6) can be written as

$$U(\zeta) = V(E(F_{\zeta \cdot \tilde{y}}), R(F_{\zeta \cdot \tilde{y}})),$$

where $F_{\zeta \cdot \tilde{y}}$ is the cumulative distribution of the random variable $\zeta \cdot \tilde{y}$. When the random variable of interest is clear, we will omit it from the notation and write the utility as $U(F) = V(E(F), R(F))$.

According to Machina (1982), if the realizations of all random variables are contained in some bounded and closed interval I and $U(F)$ is Fréchet differentiable with respect to the L^1 norm,²² then for any two distributions F_1, F_2 on I there exists $u(\cdot; F_1)$ differentiable almost everywhere on I such that

(A19)

$$U(F_2) - U(F_1) = \int_I u(y; F_1) dF_2(y) - \int_I u(y; F_1) dF_1(y) + o(\|F_2 - F_1\|),$$

where $\|\cdot\|$ denotes the L^1 norm. That is, starting from a wealth distribution F_1 , if an investor moves to another “close” distribution F_2 , then he compares the utility from these two distributions as if he is maximizing his expected utility with a local utility function $u(\cdot; F_1)$.

The key to applying Machina’s result is to find sufficient conditions on the risk measure which guarantee that $U(F)$ is Fréchet differentiable. This can be done in many ways. Next we provide one simple but effective approach which is sufficient to validate many popular risk measures as consistent with local expected utility.

Risk Measures as Functions of Moments

Let $\mu_k^F = \int y^k dF(y)$ be the k^{th} raw moment given distribution F , and $m_k^F = \int (y - \mu_1^F)^k dF(y)$ be the k^{th} central moment given distribution F . Consider risk measures which are a function of a finite number of (raw or central) moments. We denote such risk measures by $R(\mu_{j_1}^F, \dots, \mu_{j_l}^F, m_{k_1}^F, \dots, m_{k_n}^F)$. We assume that R is differentiable in all arguments. The utility function in (6) then takes the form

$$(A20) \quad U(F) = V(\mu_1^F, R(\mu_{j_1}^F, \dots, \mu_{j_l}^F, m_{k_1}^F, \dots, m_{k_n}^F)),$$

²²Fréchet differentiability is an infinite dimensional version of differentiability. The idea here is that $U(F)$ changes smoothly with F , where changes in F are topologized using the L^1 norm. See Luenberger (1969, p. 171).

where V is differentiable in both mean and risk. This class of utility functions is quite general and it allows the risk measure to depend on a large number of high distribution moments. We then have the following proposition.

Proposition 4 *If $U(F)$ takes the form (A20) then for any two distributions F_1, F_2 on I there exists $u(\cdot; F_1)$ differentiable almost everywhere on I such that (A19) holds.*

Proof: We need to show that $U(F)$ is Fréchet differentiable. By the chain rule for Fréchet differentiability (Luenberger (1969, p. 176)), we know that if both μ_k^F and m_k^F are Fréchet differentiable for any k , then so is $U(\cdot)$. The Fréchet differentiability of μ_k^F is obvious, since

$$\mu_k^{F_2} - \mu_k^{F_1} = \int_I y^k dF_2(y) - \int_I y^k dF_1(y) = -k \int_I (F_2(y) - F_1(y)) y^{k-1} dy.$$

Now we show that m_k^F is Fréchet differentiable. We have

$$\begin{aligned} m_k^F &= \int (y - \mu_1^F)^k dF(y) \\ &= \int \sum_{i=0}^k \frac{k!}{i!(k-i)!} y^i (\mu_1^F)^{k-i} dF(y) \\ &= \sum_{i=0}^k \frac{k!}{i!(k-i)!} (\mu_1^F)^{k-i} \int y^i dF(y) \\ &= \sum_{i=0}^k \frac{k!}{i!(k-i)!} (\mu_1^F)^{k-i} \mu_i^F, \end{aligned}$$

which is a differentiable function of the μ_i^F 's. By the chain rule, it follows immediately that m_k^F is also Fréchet differentiable. This completes the proof. ■

Sufficient Conditions for Positive Prices

Here we provide a sufficient condition for the positivity of equilibrium prices following the approach of Nielsen (1992). Let $\zeta \in \mathcal{R}^{n+1}$ be a bundle. Denote the gradient of investor j 's utility function at ζ by $\nabla U^j(\zeta) = (U_0^j(\zeta), \dots, U_n^j(\zeta))$,

where a subscript designates a partial derivative in the direction of the i^{th} asset. Also, let $\gamma^j(\zeta) = - (V_2^j (E(\zeta \cdot \tilde{\mathbf{y}}), R(\zeta \cdot \tilde{\mathbf{y}}))) / (V_1^j (E(\zeta \cdot \tilde{\mathbf{y}}), R(\zeta \cdot \tilde{\mathbf{y}}))) > 0$ be the marginal rate of substitution of the expected payoff of the bundle for the risk of the bundle. This is the slope of investor j 's indifference curve in the expected payoff-risk space. For brevity we often omit the arguments of this expression and use $\gamma^j(\zeta) = -V_2^j/V_1^j$.

Proposition 5 *Assume that for each asset i there is some investor j such that $E(\tilde{y}_i) > \gamma^j(\zeta) R_i(\zeta \cdot \tilde{\mathbf{y}})$ for all ζ . Then, prices of all assets are positive in all equilibria.*

Proof: At an equilibrium, all investors' gradients point in the direction of the price vector. So the price of asset i must be positive in any equilibrium if there is some investor j such that $U_i^j(\zeta) > 0$ for all ζ . Recall that

$$U^j(\zeta) = V^j(E(\zeta \cdot \tilde{\mathbf{y}}), R(\zeta \cdot \tilde{\mathbf{y}})).$$

Thus,

$$\begin{aligned} U_i^j(\zeta) &= V_1^j E(\tilde{y}_i) + V_2^j R_i(\zeta \cdot \tilde{\mathbf{y}}) \\ &= V_1^j [E(\tilde{y}_i) - \gamma^j(\zeta) R_i(\zeta \cdot \tilde{\mathbf{y}})], \end{aligned}$$

where $R_i(\zeta \cdot \tilde{\mathbf{y}})$ denotes the partial derivative of $R(\zeta \cdot \tilde{\mathbf{y}})$ with respect to ζ_i .

Since $V_1^j > 0$, $U_i^j(\zeta) > 0$ corresponds to

$$E(\tilde{y}_i) - \gamma^j(\zeta) R_i(\zeta \cdot \tilde{\mathbf{y}}) > 0,$$

as required. ■

Note that $\gamma^j(\cdot)$ can serve as a measure of risk aversion for investor j . We can thus interpret this proposition as follows. If each asset's expected return is sufficiently high relative to some investor's risk aversion and the marginal contribution of the asset to total risk, then this asset will always be desirable by some investor, and so, its price will be positive in any equilibrium.