Skewness-Aware Asset Allocation: New Theoretical Observations and Empirical Evidence

Chee-Kiat Low* Dessislava Pachamanova† Melvyn Sim‡

8 September 2009

Abstract

This paper presents a new measure of skewness, skewness-aware deviation, that can be linked to prospective satisficing risk measures and tail risk measures such as Value-at-Risk. We show that this measure of skewness arises naturally also when one thinks of maximizing the certainty equivalent for an investor with a negative exponential utility function, thus bringing together the mean-risk, expected utility, and prospective satisficing measures frameworks for an important class of investor preferences. We generalize the idea of variance and covariance in the new skewness-aware asset pricing and allocation framework. We show via computational experiments that the proposed approach results in improved and intuitively appealing asset allocation when returns follow real-world or simulated skewed distributions. We also suggest a skewness-aware equivalent of the classical CAPM beta, and study its consistency with the observed behavior of the stocks traded at the NYSE between 1963 and 2006.

Key words: Skewness, optimal portfolio allocation, beta

1 Introduction

Markowitz (1952) suggested a framework for asset allocation that set the fundamentals of modern portfolio theory. He conjectured that investors try to maximize their expected return while minimizing risk (defined as their portfolio return's variance). In addition to laying out a theory for investors' decision-making under uncertainty, Markowitz's work led to the development of the theory of the valuation of risky assets by Sharpe (1964) and Lintner (1965), whose Capital Asset

*NUS Business School, National University of Singapore, 1 Business Link, Singapore 117592. Email: cheekiat@nus.edu.sg.

†Division of Mathematics and Sciences, Babson College, Babson Park, MA 02457, USA. E-mail: dpachamanova@babson.edu. The research of the author was supported by a grant from the Babson Faculty Research Fund.

‡NUS Business School, National University of Singapore and Singapore-MIT-Alliance. Email: dscsimm@nus.edu.sg. The research of the author was supported by NUS academic research grants R-314-000-066-122 and R-314-000-068-122.

Pricing Model (CAPM) and its further enhancements are ubiquitous in financial applications even today.

Markowitz's mean-variance framework and the CAPM have come under much scrutiny over the last 50 years. It is now well-known that defining risk as the variance of future returns is only justified when future returns are assumed to be multivariate normal (more precisely, elliptically distributed), or when investors have quadratic utility functions that are only defined by their means and variances. Both of these assumptions do not hold in practice. Asset returns frequently exhibit non-normal behavior. This is particularly true of certain types of assets, such as hedge funds, emerging markets, or assets with credit risk, and is even more evident during financial crises. The quadratic utility function assumption also suffers from a number of problems, the most important of which is that it implies that at certain wealth levels, investors prefer less wealth to more wealth. It also assumes that investors are as sensitive to upside deviations as they are to downside deviations of asset returns, which is intuitively flawed.

More general assumptions on investors' utility functions can, of course, be made. For example, Kraus and Litzenberger (1976) derive a three-moment CAPM when investors are assumed to be concerned about the mean, standard deviation, and skewness (defined as the cube root of the third moment) of asset returns. Harvey and Siddique (2000) suggest an asset pricing model that incorporates conditional skewness. Hwang and Satchell (1999) develop a higher-order moment CAPM, and study its performance on emerging markets data. Harvey et al. (2004) propose a method for optimal portfolio selection with higher moments using a Bayesian decision framework. More recently, Brandt et al. (2007) suggest incorporating higher order moments for returns in the portfolio allocation process implicitly by parametrizing the portfolio weights as a function of firm characteristics, and estimating the coefficients by optimizing the average utility an investor would have obtained over a historical time period.

A second school of thought in the quantitative finance literature has emphasized incorporating considerations for skewness directly into mean-risk portfolio allocation using asymmetric and quantile-based risk measures, such as semi-variance (Markowitz 1959), Value-at-Risk (Jorion 2000) and Conditional Value-at-Risk (Rockafellar and Uryasev 2000). The interest in this approach has been spurred by developments in the theory of stochastic dominance, which have allowed investor preferences to be characterized in terms of the quantile functions of their investments (Levy 1992).

Finally, many have questioned the assumptions behind the expected utility framework altogether. In particular, the experiments of Allais (1953) and Ellsberg (1961) have shown that expected utility maximization is inconsistent with some observable behavioral preferences. To address this deficiency, a weighting function may be applied on the cumulative distribution function of uncertain outcomes. Frameworks using such a weighting function include the rank-dependent utility model (RDU) of Quiggin (1993) and the cumulative prospect theory (CPT) of Tversky and Kahneman (1992). However, including a weighting function often leads to computationally intractable optimization problems. In many applications of RDU or CPT, it is ignored; hence the models no longer serve one of their core purposes.

This paper introduces a new mean-risk allocation framework that can be linked to classical theories of utility maximization, distributional skewness measures, asset allocation and pricing. At the same time, the new framework can be linked to the recently developed theory of Prospective Satisficing Measures (Brown et al. (2009) and Brown and Sim (2009)) and is able to resolve the classical paradoxes of Allais (1953) and Ellsberg (1961). This approach evaluates the quality of portfolios based on their ability to achieve the desired financial goals or targets, which are often natural for investors to specify, as opposed to the risk-tolerance type parameters. Moreover, unlike RDU or CPT approaches, it is computational tractable in the form of conic quadratic optimization problem.

We introduce a generalized measure of skewness that in a sense measures the distance to the tails of the distribution. It represents distributional asymmetry in a different way than standard deviation, lower partial moments or quantile-based risk measures, although, as we show, it can be linked to them. We propose a definition of covariance of asset returns in the skewness-aware context, and present an alternative mean-risk approach that can be related to the certainty equivalent for CARA utility functions, thus illustrating the relationship between our mean-risk optimization framework and utility theory for an important class of investor preferences. We also derive a counterpart to the CAPM beta - a skewness-aware beta - and study whether it is consistent with returns observed in the market.

As our computational experiments illustrate, the suggested approach results in an improved and intuitively appealing asset allocations when returns follow real-world or simulated skewed distributions. Moreover, the skewness-aware beta appears more consistent with the observed risk-return tradeoff in studies of the stocks traded at the NYSE. While these computational results are preliminary, they indicate this generalized measure for incoporating skewness in portfolio allocation and risk attribution deserves a second look.

The paper is structured as follows. In Section 2, we introduce the skewness-aware deviation of a random variable, discuss important properties of the latter skewness measure, and show its relation to the problem of optimizing the certainty equivalent for a negative exponential utility function. Section 3 presents the concepts of skewness-aware variance, covariance, and standard deviation functions. Section 4 discusses a skewness-aware mean-risk model, and relates it to the optimization of a specific Prospective Satisficing Measure. Section 5 contains computational experiments with simulated and real market data that study portfolio performance under the new optimization framework, as well as a computational study of the behavior of the classical and the skewness-aware beta for historical market returns. Section 6 concludes with a summary of observations.

2 Skewness-Aware Deviation

Let $\tilde{z}$ be a random variable with zero mean. Consider the following measure of the variability of $\tilde{z}$, which we name *skewness-aware deviation*:

$$\xi^2(\tilde{z}) \triangleq \sup_{\theta>0} \left\{ \frac{2}{\theta^2} \ln(\mathbb{E}(\exp(-\theta\tilde{z})) \right\}. \tag{1}$$

Note that ξ^2 is related to the logarithm of the moment-generating function of the random variable $\tilde{z}$. Thus, at the intuitive level, it incorporates information about the characteristics of the entire probability distribution of $\tilde{z}$. However, the skewness-aware deviation is not directly related to a specific higher order moment of the distribution. We prefer to avoid working with higher moments, as their inclusion in a portfolio allocation framework requires a lot of data for estimation, and does not necessarily result in tractable convex optimization formulations. As we will show later, the skewness-aware deviation is equivalent to the variance when the random variables are normal, and provides a bound to the Value-at-Risk measure.

The generalized deviation can be estimated analytically or numerically from a set of data. Given a sample of realizations for the random variable $\tilde{z}$, one can obtain an estimate of its value by performing, for example, a binary search on θ.

2.1 Relation to Variance

The following two propositions concern the relationship between the skewness-aware deviation and the variance of a random variable. The proofs of the two propositions are in the appendix.

Proposition 1 *If the zero-mean random variable $\tilde{z}$ is normally distributed, then*

$$\xi^2(\tilde{z}) = \sigma^2(\tilde{z}),$$

where $\sigma^2(\tilde{z})$ denotes the variance of $\tilde{z}$.

Proposition 2 *For any random variable $\tilde{z}$ with zero mean,*

$$\xi^2(\tilde{z}) \geq \sigma^2(\tilde{z}) \geq 0.$$

In addition to its relation to variance, the skewness-aware deviation measure describes probabilistically the tails of the distribution, and can be linked to quantile-based risk measures such as VaR and CVaR. We will discuss these links in more detail in the context of asset allocation in Section 3.

2.2 Relation to the Expected Utility Maximization Framework

The expected utility maximization framework addresses a much wider class of problems than Markowitz's mean-variance framework, and includes the latter framework as a special case. Well-known utility functions include the quadratic utility function $U(w) = w - \frac{b}{2}w^2$, the power utility function $U(w) = w^\alpha, 0 \leq \alpha \leq 1$, and the negative exponential utility function $U(w) = -\exp(-\kappa w), \kappa > 0$.

An equivalent way of formulating the investor's problem of maximizing expected utility over the set of forecasted values for future portfolio returns is to optimize the investor's certainty equivalent. The certainty equivalent, defined as the certain return that makes an investor indifferent between taking that return and making a risky investment $\tilde{v}$, is formally given by the equation

$$C(\tilde{v}) \stackrel{\Delta}{=} U^{-1}(\mathbb{E}(U(\tilde{v}))),$$

i.e., $U(C(\tilde{v})) = \mathbb{E}(U(\tilde{v}))$. By maximizing the certainty equivalent, one incorporates both assumptions about the underlying distribution of future returns, and about preferences in the portfolio selection problem.

Consider an investor with a negative exponential utility function. The certainty equivalent of a random variable $\tilde{v}$ if the investor has a negative exponential utility function is given by

$$C_\theta(\tilde{v}) \stackrel{\Delta}{=} -\frac{1}{\theta}\ln(\mathbb{E}(\exp(-\theta\tilde{v})).$$

Therefore, $\xi^2(\tilde{z})$ can be interpreted as the maximum marginal decrease in the certainty equivalent of $\tilde{z}$ for a negative exponential utility function:

$$\xi^2(\tilde{z}) = \sup_{\theta > 0}\left\{-\frac{2}{\theta}C_\theta(\tilde{z})\right\}.[1] \tag{2}$$

[1] Given a random variable $\tilde{z}$ with zero mean, the certainty equivalent of the negative exponential utility function, $C_\theta(\tilde{z})$, is always less than or equal to 0 since $C_\theta(\tilde{z}) \leq \mathbb{E}(\tilde{z}) = 0$ if $\theta > 0$ ($\theta > 0$ implies that investors are risk-averse).

In other words, $\xi^2(\tilde{z})$ measures the greatest marginal decrease in the certainty equivalent with respect to the Arrow-Pratt absolute risk aversion parameter. It provides a bound for the possible decrease in certainty equivalent return for all investors with negative exponential utility functions.

Note that the fact that the skewness-aware deviation is related to the certainty equivalent of the exponential utility function does not mean that maximizing the skewness-aware deviation of a portfolio will produce the same allocation as maximizing the expected utility of an investor with a negative exponential utility. (This will become explicit in Section 4.) As we mentioned in the introduction, expected utility maximization theory has certain drawbacks. We work outside the theory, but draw this parallel in order to provide intuition about the skewness-aware deviation.

3 Skewness-Aware Variability Measures

This section sets up the framework for skewness-aware traditional measures of variability that we will use to generalize Markowitz's framework in Section 4. We introduce the concepts of skewness-aware variance, covariance, and standard deviation functions based on the definition of the skewness-aware deviation presented in the previous section.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a measure space. We define a family of functions that is affinely dependent on a set of M independent random variables on Ω. Let

$$\mathcal{Z} = \left\{ \tilde{y} \ : \ \exists (y_0, \boldsymbol{y}) \in \mathcal{R}^{M+1} \ : \ \tilde{y} = y_0 + \underbrace{y_1 \tilde{z}_1 + \ldots + y_M \tilde{z}_M}_{= \boldsymbol{y}' \tilde{\boldsymbol{z}}} \right\},$$

where $\tilde{z}_1, \ldots, \tilde{z}_M$ are zero means, independent random variables defined on Ω. Moreover, assume that each random variable $\tilde{z}_j$ has finite and positive $\xi^2(\tilde{z}_j)$, $\xi^2(-\tilde{z}_j)$. One can think of $\tilde{y}$ as asset returns, and of $\tilde{z}_1, \ldots, \tilde{z}_M$ - as independent factors affecting asset returns.

We define the skewness-aware variance function $\mathrm{Var}_{\mathrm{sk}} : \mathcal{Z} \to \mathcal{R}$ as follows:

$$\mathrm{Var}_{\mathrm{sk}}(y_0 + \tilde{\boldsymbol{z}}' \boldsymbol{y}) = \boldsymbol{y}' \boldsymbol{\phi}(\boldsymbol{y})$$

where

$$\boldsymbol{\phi}(\boldsymbol{y}) = [\phi_1(\boldsymbol{y}) \ldots \phi_M(\boldsymbol{y})]' \qquad \phi_j(\boldsymbol{y}) = \begin{cases} \xi^2(\tilde{z}_j) y_j & \text{if } y_j \geq 0 \\ \xi^2(-\tilde{z}_j) y_j & \text{otherwise.} \end{cases}$$

The skewness-aware variance function has the following properties (for proofs, see the appendix):

Proposition 3 **(a)**

$$\mathrm{Var}_{\mathrm{sk}}(y_0 + \tilde{\boldsymbol{z}}' \boldsymbol{y}) = 0$$

if and only if $\boldsymbol{y} = \boldsymbol{0}$,

(b)

$$\mathrm{Var}_{\mathrm{sk}}(k \tilde{y}) = k^2 \mathrm{Var}_{\mathrm{sk}}(\tilde{y}) \qquad \forall k > 0, \tilde{r} \in \mathcal{Z}$$

(c)

$$\mathrm{Var}_{\mathrm{sk}}(y_0 + \tilde{\boldsymbol{z}}' \boldsymbol{y}) = \min \left\{ \boldsymbol{u}' \boldsymbol{u} \ : \ \exists \boldsymbol{u} \in \mathcal{R}^M \ : \ u_j \geq \sqrt{\xi^2(\tilde{z}_j)} y_j, \ u_j \geq -\sqrt{\xi^2(-\tilde{z}_j)} y_j, \ j = 1, \ldots, M \right\}.$$

Note that unlike variance, the skewness-aware variance is sensitive to the direction of return movement. Hence, $\text{Var}_{\text{sk}}(\tilde{y}) = \text{Var}_{\text{sk}}(-\tilde{y})$ does not necessarily hold unless $\xi^2(\tilde{z}_j) = \xi^2(-\tilde{z}_j)$ for all $j = 1, \ldots, M$.

Associated with the definition of variance, we define a skewness-aware covariance function, $\text{Cov}_{\text{sk}} : \mathcal{Z} \times \mathcal{Z} \rightarrow \mathcal{R}$, as follows:

$$\text{Cov}_{\text{sk}}(s_0 + \tilde{\boldsymbol{z}}'\boldsymbol{s}, t_0 + \tilde{\boldsymbol{z}}'\boldsymbol{t}) = \boldsymbol{s}'\boldsymbol{\phi}(\boldsymbol{t}).$$

Correspondingly, we define a skewness-aware standard deviation function analogous to standard deviation, $\sigma_{sk} : \mathcal{Z} \rightarrow \mathcal{R}$, as

$$\sigma_{sk}(\tilde{y}) = \sqrt{\text{Var}_{\text{sk}}(\tilde{y})}.$$

The skewness-aware standard deviation has the following properties:

Proposition 4 **(a)**

$$\sigma_{sk}(y_0 + \tilde{\boldsymbol{z}}'\boldsymbol{y}) = 0$$

if and only if $\boldsymbol{y} = \boldsymbol{0}$*,*

(b) Positive Homogeneity*:*

$$\sigma_{sk}(k\tilde{y}) = k\sigma_{sk}(\tilde{y}) \qquad \forall k > 0, \tilde{r} \in \mathcal{Z}$$

(c) Subadditivity*:*

$$\sigma_{sk}(\tilde{s} + \tilde{t}) \leq \sigma_{sk}(\tilde{s}) + \sigma_{sk}(\tilde{t})$$

for all $\tilde{s}, \tilde{t} \in \mathcal{Z}$*.*

For proofs, please refer to the appendix.

In order to provide some intution about the relationship between the skewness-aware standard deviation and the classical standard deviation, we illustrate the behavior of the skewness-aware standard deviation with a simple numerical example in which $\tilde{y}$ equals a single factor $\tilde{z}$ whose probability distribution depends on a parameter ω.

Let $\tilde{z}(\omega)$ have the following probability distribution:

$$\tilde{z}(\omega) = \begin{cases} \frac{\sqrt{\omega(1-\omega)}}{\omega} & \text{with probability } \omega, \\ -\frac{\sqrt{\omega(1-\omega)}}{(1-\omega)} & \text{with probability } (1-\omega). \end{cases} \tag{3}$$

Note that the mean and the standard deviation of $\tilde{z}(\omega)$ are the same for all $\omega \in (0, 1)$ - they are 0 and 1, respectively. However, the degree of symmetry of $\tilde{z}(\omega)$ can be different. Higher values for ω (e.g., $\omega = 0.9$) result in large negative values and small upside gains.

Define the *asymmetry ratio* of the random variable $\tilde{z}$ with zero mean as

$$\rho(\tilde{z}) \stackrel{\Delta}{=} \sigma_{sk}(\tilde{z})/\sigma(\tilde{z}). \tag{4}$$

Since $\sigma_{sk}(\tilde{z}) \geq \sigma(\tilde{z})$ and $\sigma_{sk}(\tilde{z}) = \sigma(\tilde{z})$ for normal distribution, one can interpret the ratio $\rho(\tilde{z})$ as a deviation benchmark against the normal distribution. Namely, $\rho(\tilde{z}) > 1$ indicates that the left tail of the distribution of the random variable $\tilde{z}$ is "heavier" than the left tail of the normal distribution. Similarly, $\rho(-\tilde{z}) > 1$ indicates that the right tail of the distribution of the random variable $\tilde{z}$ is heavier than the right tail of the normal distribution. In particular, given an arbitrary distribution, $\sigma_{sk}(\tilde{z})$ and $\sigma_{sk}(-\tilde{z})$ characterize the partial tail deviations with respect to the influence on the certainty equivalent of $\tilde{z}$ for the negative exponential utility function. Although there are

ω	$\rho(\tilde{z}(\omega))$	$\rho(-\tilde{z}(\omega))$	Skewness	Kurtosis
0.01	1.000	3.282	9.849	98.010
0.10	1.000	1.422	2.667	8.111
0.30	1.000	1.060	0.873	1.762
0.50	1.000	1.000	0.000	1.000
0.70	1.060	1.000	-0.873	1.762
0.90	1.422	1.000	-2.667	8.111
0.99	3.282	1.000	-9.849	98.010

Table 1: Values of $\rho(\tilde{z})$ and $\rho(-\tilde{z})$ for a two-point distribution.

other statistical approaches for measuring deviation skewness, the ratios $\rho(\tilde{z})$ and $\rho(-\tilde{z})$ have a very specific link to mean-risk portfolio optimization and utility theory.

The corresponding ratios $\rho(\tilde{z}(\omega))$ and $\rho(-\tilde{z}(\omega))$ for different values of ω are presented in Table 1. Observe that as ω increases above 0.5, the distribution of $\tilde{z}(\omega)$ becomes skewed to the left, thereby increasing the marginal decrease in certainty equivalent $\sigma_{sk}(\tilde{z}(\omega))$ relative to the standard deviation.

The skewness-aware standard deviation provides bounds on the tails of return distributions as illustrated in the following proposition. This fact illustrates a link between the new deviation measure and quantile-based risk measures.

Proposition 5 *Given* $\tilde{y} \in \mathcal{Z}$,

$$\mathbb{P}\left(\tilde{y} - \mathbb{E}(\tilde{y}) < -a\sigma_{sk}(\tilde{y})\right) \leq \exp(-a^2/2).$$

Similarly,

$$\mathbb{P}\left(\tilde{y} - \mathbb{E}(\tilde{y}) > a\sigma_{sk}(-\tilde{y})\right) \leq \exp(-a^2/2).$$

The proof of Proposition 5 is provided in the appendix. Note that by using the skewness-aware standard deviation in portfolio allocation formulations, one can impose Value-at-Risk-type constraints, in which the probability of the portfolio return being in the tail is impacted by the specification of the parameter a.

Finally, let us define the function $\psi : \mathcal{Z} \to \mathcal{R}$:

$$\psi(y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}) = \frac{\mathbb{E}(y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})}{\sigma_{sk}(y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})} = \frac{y_0}{\sqrt{\boldsymbol{y}'\boldsymbol{\phi}(\boldsymbol{y})}}.$$

One can think of $\psi(.)$ as a skewness-aware risk-reward function, similar to the Sharpe ratio.

In view of Proposition 5, it is easy to see that the probability that an asset's return is nonnegative can be bounded as follows:

$$\mathbb{P}(\tilde{y} \geq 0) \geq 1 - \exp(-\psi(\tilde{y})^2/2).$$

4 The Skewness-Aware Asset Allocation Framework

Consider a factor model

$$\tilde{\boldsymbol{r}} = \hat{\boldsymbol{r}} + \boldsymbol{A}\tilde{\boldsymbol{z}}, \tag{5}$$

in which $\hat{\boldsymbol{r}}$ is a vector of expected (nominal) returns, and the uncertain returns $\tilde{r}_j = \boldsymbol{e}_i'\hat{\boldsymbol{r}} + \boldsymbol{e}_i'\boldsymbol{A}\tilde{\boldsymbol{z}} \in \mathcal{Z}$. We note that $\tilde{\boldsymbol{r}}'\boldsymbol{x} = \hat{\boldsymbol{r}}'\boldsymbol{x} + (\boldsymbol{A}'\boldsymbol{x})'\tilde{\boldsymbol{z}} \in \mathcal{Z}$ as well.

The skewness-aware Markowitz model can be formulated as follows:

$$\begin{array}{rll} s_{sk}^2(r_p) = & \min & \mathrm{Var}_{\mathrm{sk}}(\tilde{\boldsymbol{r}}'\boldsymbol{x}) \\ & \text{s.t.} & \boldsymbol{x}'\hat{\boldsymbol{r}} = r_p \\ & & \boldsymbol{x}'\boldsymbol{e} = 1 \end{array} \tag{6}$$

In formulation (6), r_p denotes the required portfolio return.

Note that, based on Proposition 3(c), the skewness-aware Markowitz model can be formulated as a quadratic optimization model:

$$\begin{array}{rlll} s_{sk}^2(r_p) = & \min & \boldsymbol{u}'\boldsymbol{u} & \\ & \text{s.t.} & u_j \geq \sigma_{sk}(\tilde{z}_j)y_j & \forall j = 1,\ldots,M \\ & & u_j \geq -\sigma_{sk}(-\tilde{z}_j)y_j & \forall j = 1,\ldots,M \\ & & \boldsymbol{y} = \boldsymbol{A}'\boldsymbol{x} & \\ & & \boldsymbol{x}'\hat{\boldsymbol{r}} = r_p & \\ & & \boldsymbol{x}'\boldsymbol{e} = 1. & \end{array} \tag{7}$$

4.1 Skewness-Aware Efficient Frontier

We next show that the frontier is a convex curve in the $r_p - s_{sk}(r_p)$ plane. In other words, by performing portfolio optimization in the skewness-aware framework, we trace an efficient frontier with similar properties to the Markowitz efficient frontier.

Proposition 6 *The frontier curve $s_{sk}(r_p)$ is convex in r_p.*

Let $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$ be the optimal portfolio at levels r_p^1 and r_p^2. Let

$$(r_p^\lambda, \boldsymbol{x}^\lambda) = \lambda(r_p^1, \boldsymbol{x}^1) + (1-\lambda)(r_p^2, \boldsymbol{x}^2)$$

for some $\lambda \in [0,1]$. Observe that $\hat{\boldsymbol{r}}'\boldsymbol{x}^\lambda = r_p^\lambda$, and $\boldsymbol{e}\boldsymbol{x}^\lambda = 1$. Using the properties of the skewness-aware deviation measure from Proposition (4), we have

$$\begin{array}{rcl} \lambda s_{sk}(r_p^1) + (1-\lambda)s_{sk}(r_p^1) & = & \lambda\sigma_{sk}(\tilde{\boldsymbol{r}}'\boldsymbol{x}^1) + (1-\lambda)\sigma_{sk}(\tilde{\boldsymbol{r}}'\boldsymbol{x}^2) \\ & = & \sigma_{sk}(\lambda\tilde{\boldsymbol{r}}'\boldsymbol{x}^1) + \sigma_{sk}((1-\lambda)\tilde{\boldsymbol{r}}'\boldsymbol{x}^2) \\ & \geq & \sigma_{sk}(\lambda\tilde{\boldsymbol{r}}'\boldsymbol{x}^1 + (1-\lambda)\tilde{\boldsymbol{r}}'\boldsymbol{x}^2) \\ & \geq & s_{sk}(r_p^\lambda). \end{array}$$

■

4.2 Portfolio Allocation in the Presence of a Riskless Asset

Suppose now that there exists a riskless asset with rate of return r_f $(r_p > r_f)$. The portfolio return is $\tilde{\boldsymbol{r}}'\boldsymbol{x} + (1 - \boldsymbol{e}'\boldsymbol{x})r_f$. The expected return is therefore $\hat{\boldsymbol{r}}'\boldsymbol{x} + (1 - \boldsymbol{e}'\boldsymbol{x})r_f$ and the skewness-aware variance is $\mathrm{Var}_{\mathrm{sk}}(\tilde{\boldsymbol{r}}'\boldsymbol{x})$. The investment problem becomes:

$$\begin{array}{rll} s_{sk}^2(r_p) = & \min & \mathrm{Var}_{\mathrm{sk}}(\tilde{\boldsymbol{r}}'\boldsymbol{x}) \\ & \text{s.t.} & (\hat{\boldsymbol{r}} - \boldsymbol{e}r_f)'\boldsymbol{x} = r_p - r_f. \end{array} \tag{8}$$

Equivalently, it can be written as

$$\begin{array}{rlll} s_{sk}^2(r_p) = & \min & \boldsymbol{u}'\boldsymbol{u} & \\ & \text{s.t.} & u_j \geq \sigma_{sk}(\tilde{z}_j)y_j & \forall j = 1, \ldots, M \\ & & u_j \geq -\sigma_{sk}(-\tilde{z}_j)y_j & \forall j = 1, \ldots, M \\ & & \boldsymbol{y} = \boldsymbol{A}'\boldsymbol{x} & \\ & & (\hat{\boldsymbol{r}} - \boldsymbol{e}r_f)'\boldsymbol{x} = r_p - r_f. & \end{array} \tag{9}$$

If $\tilde{\boldsymbol{r}}$ follows a multivariate normal distributions, we recover the Markowitz model. When $r_p \geq r_f$, the frontier consists of a positively sloped ray in the r_p - $s_{sk}(r_p)$ plane:

$$s_{sk}(r_p) = \frac{r_p - r_f}{(\hat{\boldsymbol{r}} - \boldsymbol{e}r_f)'\boldsymbol{\Sigma}(\hat{\boldsymbol{r}} - \boldsymbol{e}r_f)},$$

where $\boldsymbol{\Sigma}$ denotes the covariance matrix of the uncertain returns $\tilde{\boldsymbol{r}}$.

We show that a similar relationship holds when $\tilde{\boldsymbol{r}}$ follows more general distributions.

Proposition 7 *Assume the following:*

1. *There exist portfolio weights $\boldsymbol{x}$, $\boldsymbol{e}'\boldsymbol{x} = 1$, such that $\mathbb{E}(\tilde{\boldsymbol{r}}'\boldsymbol{x}) > r_f$.*
2. *For all $r_p > r_f$, the optimum solution $\boldsymbol{x}^*$ of Problem (8) satisfies $\boldsymbol{e}'\boldsymbol{x}^* > 0$.*

Then, the frontier in the r_p - $s_{sk}(r_p)$ plane, with $r_p \geq r_f$, is a ray starting from the point $(r_f, 0)$ and passing through a particular point, $(r_p^*, s_{sk}(r_p^*))$, with $r_p^* = \hat{\boldsymbol{r}}'\boldsymbol{x}^\dagger$, where $\boldsymbol{x}^\dagger$ is a optimal solution to the following skewness-aware Sharpe ratio problem

$$\begin{array}{rll} Z_{SkSharpe} = & \max & \psi(\tilde{\boldsymbol{r}}'\boldsymbol{x} - r_f) \\ & \text{s.t.} & \boldsymbol{e}'\boldsymbol{x} = 1. \end{array} \tag{10}$$

Proof : The skewness-aware Sharpe ratio of Problem (10) can be expressed as

$$\begin{array}{ll} \max & \dfrac{\hat{\boldsymbol{r}}'\boldsymbol{x} - r_f}{\sigma_{sk}(\tilde{\boldsymbol{r}}'\boldsymbol{x})} \\ \text{s.t.} & \boldsymbol{e}'\boldsymbol{x} = 1. \end{array}$$

or

$$\begin{array}{ll} \max & \dfrac{(\hat{\boldsymbol{r}} - \boldsymbol{e}r_f)'\boldsymbol{x}}{\sqrt{\mathrm{Var}_{\mathrm{sk}}(\tilde{\boldsymbol{r}}'\boldsymbol{x})}} \\ \text{s.t.} & \boldsymbol{e}'\boldsymbol{x} = 1. \end{array} \tag{11}$$

since $r_f = \boldsymbol{e}'\boldsymbol{x}r_f$. Under the assumption that there exist portfolio weights $\boldsymbol{x}$, $\boldsymbol{e}'\boldsymbol{x} = 1$, such that $\mathbb{E}(\tilde{\boldsymbol{r}}'\boldsymbol{x}) > r_f$, we have $Z_{SkSharpe} > 0$. Noting that the objective function of Problem (11) is invariant to positive scaling, we relax the equality constraint and keep the numerator in the objective function constant, which is equivalent to solving the following problem:

$$\begin{array}{ll} \min & \mathrm{Var}_{\mathrm{sk}}(\tilde{\boldsymbol{r}}'\boldsymbol{y}) \\ \text{s.t.} & (\hat{\boldsymbol{r}} - \boldsymbol{e}r_f)'\boldsymbol{y} = 1. \end{array} \tag{12}$$

Indeed, since the optimum solution, $\boldsymbol{y}^*$ of Problem (12) satisfies $\boldsymbol{e}'\boldsymbol{y}^* > 0$, the solution

$$\boldsymbol{x}^\dagger = \frac{1}{\boldsymbol{e}'\boldsymbol{y}^*}\boldsymbol{y}^*$$

is optimal in the Sharpe ratio problem (10). Observe that Problem (12) is essentially the same as Problem (9) in which $r_p - r_f = 1$. Hence, we consider the following KKT optimal conditions for the quadratic optimization problem (9):

$$Y(\zeta) = \left\{ (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\delta}, \gamma) \;:\; \begin{array}{ll} 2u_j - \alpha_j - \delta_j = 0 & \forall j = 1, \ldots, M \\ -\boldsymbol{A\lambda} + \gamma(\hat{\boldsymbol{r}} - \boldsymbol{e}r_f) = \boldsymbol{0} & \\ \sigma_{sk}(\tilde{z}_j)\alpha_j - \sigma_{sk}(-\tilde{z}_j)\delta_j + \lambda_j = 0 & \forall j = 1, \ldots, M \\ u_j \geq \sigma_{sk}(\tilde{z}_j)y_j & \forall j = 1, \ldots, M \\ u_j \geq -\sigma_{sk}(-\tilde{z}_j)y_j & \forall j = 1, \ldots, M \\ \boldsymbol{y} = \boldsymbol{A}'\boldsymbol{x} & \\ \boldsymbol{\alpha} \geq 0, \boldsymbol{\delta} \geq 0 & \\ (u_j - \sigma_{sk}(\tilde{z}_j)y_j)\alpha_j = 0 & \forall j = 1, \ldots, M \\ (u_j + \sigma_{sk}(-\tilde{z}_j)y_j)\delta_j = 0 & \forall j = 1, \ldots, M \\ (\hat{\boldsymbol{r}} - \boldsymbol{e}r_f)'\boldsymbol{x} = \zeta. & \end{array} \right\},$$

where $\zeta = r_p - r_f$. Let $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{u}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*, \boldsymbol{\delta}^*, \gamma^*) \in Y(r_p^* - r_f) = Y(1)$. By inspection,

$$(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{u}^*, \boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*, \boldsymbol{\delta}^*, \gamma^*)(r_p - r_f) \in Y(r_p - r_f)$$

for all $r_p \geq r_f$ and hence, $s_{sk}(r_p) = \|\boldsymbol{u}^*\|_2(r_p - r_f)$. ■

4.3 Skewness-Aware Beta

In this section, we derive an expression for an alternative to the classical CAPM beta that allows us to measure a particular asset's risk with respect to the portfolio while incorporating considerations for skewness.

Proposition 8 *The optimum solution of Problem (9) satisfies*

$$\hat{r}_i - r_f = \beta_i(\boldsymbol{x})(r_p - r_f) \qquad i = 1, \ldots, N$$

where

$$\beta_i(\boldsymbol{x}) = \frac{[\boldsymbol{A\phi}(\boldsymbol{A}'\boldsymbol{x})]_i}{\boldsymbol{x}'\boldsymbol{A\phi}(\boldsymbol{A}'\boldsymbol{x})} = \frac{\mathrm{Cov}_{\mathrm{sk}}(\tilde{r}_i, \tilde{\boldsymbol{r}}'\boldsymbol{x})}{\mathrm{Var}_{\mathrm{sk}}(\tilde{\boldsymbol{r}}'\boldsymbol{x})}, \tag{13}$$

Proof : Observe from the constraints in the KKT conditions,

$$\begin{array}{ll} \lambda_j = -\sigma_{sk}(\tilde{z}_j)\alpha_j + \sigma_{sk}(-\tilde{z}_j)\delta_j & \\ (u_j - \sigma_{sk}(\tilde{z}_j)y_j)\alpha_j = 0 & \forall j = 1, \ldots, M \\ (u_j + \sigma_{sk}(-\tilde{z}_j)y_j)\delta_j = 0 & \forall j = 1, \ldots, M \\ 2u_j = \alpha_j + \delta_j & \forall j = 1, \ldots, M \\ \boldsymbol{\alpha} \geq 0, \boldsymbol{\delta} \geq 0 & \end{array}$$

implies that $u_j = \max\{\sigma_{sk}(\tilde{z}_j)y_j, -\sigma_{sk}(-\tilde{z}_j)y_j\}$. Indeed, since $\sigma_{sk}(\tilde{z}_j)$ and $\sigma_{sk}(-\tilde{z}_j) > 0$, if

$$u_j > \max\{\sigma_{sk}(\tilde{z}_j)y_j, -\sigma_{sk}(-\tilde{z}_j)y_j\} \geq 0,$$

then $\alpha_j = \delta_j = 0$, which contradicts the fact that $2u_j = \alpha_j + \delta_j$. By inspection, if $y_j > 0$, we have $u_j = \sigma_{sk}(\tilde{z}_j)y_j > 0 > -\sigma_{sk}(-\tilde{z}_j)y_j$, $\delta_j = 0$ and $\alpha_j = 2u_j = 2\sigma_{sk}(\tilde{z}_j)y_j$. Hence,

$$\lambda_j = -\sigma_{sk}(\tilde{z}_j)\alpha_j + \sigma_{sk}(-\tilde{z}_j)\delta_j = -2\sigma_{sk}^2(\tilde{z}_j)y_j.$$

Similarly, if $y_j < 0$, we have $u_j = -\sigma_{sk}(-\tilde{z}_j)y_j > 0 > \sigma_{sk}(\tilde{z}_j)y_j$, $\alpha_j = 0$ and $\delta_j = 2u_j = -2\sigma_{sk}(-\tilde{z}_j)y_j$. Hence, $\lambda_j = -2\sigma_{sk}^2(-\tilde{z}_j)y_j$. Finally, when $y_j = 0$, we have $u_j = 0$, implying $\alpha_j = \delta_j = 0$ and $\lambda_j = 0$. Hence,

$$\lambda_j = -2\phi_j(\boldsymbol{y}).$$

Now, since the KKT conditions require

$$\begin{aligned} \gamma(\hat{\boldsymbol{r}} - \boldsymbol{e}r_f) &= \boldsymbol{A}\boldsymbol{\lambda} \\ (\hat{\boldsymbol{r}} - \boldsymbol{e}r_f)'\boldsymbol{x} &= r_p - r_f, \end{aligned}$$

we have

$$\gamma \underbrace{(\hat{\boldsymbol{r}} - \boldsymbol{e}r_f)'\boldsymbol{x}}_{=r_p - r_f} = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{\lambda} = -2\boldsymbol{x}'\boldsymbol{A}\boldsymbol{\phi}(\boldsymbol{A}'\boldsymbol{x}).$$

Hence, substituting γ, we have

$$\hat{\boldsymbol{r}} - \boldsymbol{e}r_f = \frac{\boldsymbol{A}\boldsymbol{\phi}(\boldsymbol{A}'\boldsymbol{x})}{\boldsymbol{x}'\boldsymbol{A}\boldsymbol{\phi}(\boldsymbol{A}'\boldsymbol{x})}.$$

Finally, since $\tilde{r}_i = \boldsymbol{e}_i'\hat{\boldsymbol{r}} + (\boldsymbol{A}'\boldsymbol{e}_i)'\tilde{\boldsymbol{z}}$ and $\tilde{\boldsymbol{r}}'\boldsymbol{x} = \boldsymbol{x}'\hat{\boldsymbol{r}} + (\boldsymbol{A}'\boldsymbol{x})'\tilde{\boldsymbol{z}}$, we have

$$\mathrm{Cov}_{\mathrm{sk}}(\tilde{r}_i, \tilde{\boldsymbol{r}}'\boldsymbol{x}) = \boldsymbol{e}_i'\boldsymbol{A}\boldsymbol{\phi}(\boldsymbol{A}'\boldsymbol{x})$$

and

$$\mathrm{Var}_{\mathrm{sk}}(\tilde{\boldsymbol{r}}'\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{\phi}(\boldsymbol{A}'\boldsymbol{x}).$$

■

The quantity β_i can be interpreted as the relative change in the skewness-aware portfolio variance when varying the weight of asset i. Note that, similarly to the classical case, $\sum_{i=1}^N x_i\beta_j(\boldsymbol{x}) = 1$, which produces a decomposition of the skewness-aware portfolio variability into the contributions of the individual assets. When all independent factors are normally distributed, the skewness-aware beta is equal to the classical beta: $\beta_j(\boldsymbol{x}) = \mathrm{Cov}(\tilde{r}_j, \tilde{\boldsymbol{r}}'\boldsymbol{x})/\mathrm{Var}(\tilde{\boldsymbol{r}}'\boldsymbol{x})$.

4.4 Relationship between Skewness-Aware Portfolio Allocation, the Classical Markowitz Model, and Expected Utility Maximization

Note that (7) is in fact a generalization of the Markowitz mean-variance model. If all factors follow normal distributions, we can choose $\boldsymbol{A} = \boldsymbol{\Sigma}^{1/2}$. In that case, $\sigma_{sk}(\tilde{z}_j) = \sigma_{sk}(-\tilde{z}_j) = 1$, and the skewness-aware model becomes the classical Markowitz model. More generally, consider the skewness-aware Markowitz model in a utility framework:

$$\begin{aligned} Z_{skM}^* = \max \quad & \hat{\boldsymbol{r}}'\boldsymbol{x} - \frac{\kappa}{2}\boldsymbol{u}'\boldsymbol{u} \\ \text{s.t.} \quad & u_j \geq \sigma_{sk}(\tilde{z}_j)y_j \quad \forall j \\ & u_j \geq -\sigma_{sk}(-\tilde{z}_j)y_j \quad \forall j \\ & \boldsymbol{y} = \boldsymbol{A}'\boldsymbol{x} \\ & \boldsymbol{x} \in X. \end{aligned} \tag{14}$$

Next, we show that the optimal solution to model (14) provides a lower bound to the optimal solution from the classical Markowitz model.

Theorem 1 *The optimal objective value of the skewness-aware model (14) satisfies*

$$Z^*_{skM} \leq Z^*_M,$$

where Z^*_M is the optimal objective function value of the classical Markowitz model for the same value of κ.

Proof : Note that under the classical Markowitz model, the variance of $\tilde{\boldsymbol{r}}'\boldsymbol{x}$ is $\boldsymbol{u}'\boldsymbol{u}$, where

$$u_j = |\sigma(\tilde{z}_j)y_j|,$$

and $\boldsymbol{y} = \boldsymbol{A}'\boldsymbol{x}$. Hence, the classical Markowitz model is special case of model (14), in which

$$\sigma_{sk}(\tilde{z}_j) = \sigma_{sk}(-\tilde{z}_j) = \sigma(\tilde{z}_j).$$

Since $\sigma_{sk}(\tilde{z}_j), \sigma_{sk}(\tilde{z}_j) \geq \sigma(\tilde{z}_j)$, we have

$$\max\{\sigma_{sk}(\tilde{z}_j)y_j, -\sigma_{sk}(-\tilde{z}_j)y_j\} \geq |\sigma(\tilde{z}_j)y_j|.$$

It follows that

$$Z^*_{skM} \leq Z^*_M.$$

■

Model (14) provides also a bound to the optimal objective function value of an investor with a negative exponential utility function. Note that under the factor model (5), the certainty equivalent of portfolio returns for an investor with a negative exponential utility function is

$$C_\kappa(\tilde{\boldsymbol{r}}'\boldsymbol{x}) = \hat{\boldsymbol{r}}'\boldsymbol{x} + \sum_j C_\kappa(y_j\tilde{z}_j),$$

This is because

$$\begin{aligned} C_\kappa(\tilde{\boldsymbol{r}}'\boldsymbol{x}) &= \hat{\boldsymbol{r}}'\boldsymbol{x} - \tfrac{1}{\kappa}\ln(\mathbb{E}(\exp(-\kappa(\underbrace{\boldsymbol{A}'\boldsymbol{x}}_{=\boldsymbol{y}})'\tilde{\boldsymbol{z}})) \\ &= \hat{\boldsymbol{r}}'\boldsymbol{x} - \tfrac{1}{\kappa}\ln(\mathbb{E}(\exp(-\kappa\boldsymbol{y}'\tilde{\boldsymbol{z}})) \\ &= \hat{\boldsymbol{r}}'\boldsymbol{x} - \tfrac{1}{\kappa}\ln(\textstyle\prod_j \mathbb{E}(\exp(-\kappa y_j\tilde{z}_j))) \quad (\tilde{z}_j \text{ are independent}) \\ &= \hat{\boldsymbol{r}}'\boldsymbol{x} - \tfrac{1}{\kappa}\textstyle\sum_j \ln(\mathbb{E}(\exp(-\kappa y_j\tilde{z}_j))). \end{aligned} \tag{15}$$

It follows that

Theorem 2 *The optimal objective value of the skewness-aware model (7) satisfies*

$$Z^*_{skM} \leq Z^*,$$

where Z^* is the optimal objective function value of the general certainty equivalent optimization model

$$\begin{aligned} Z^* = \max \quad & C_\kappa(\tilde{\boldsymbol{r}}'\boldsymbol{x}) \\ \text{s.t.} \quad & \boldsymbol{x} \in X. \end{aligned} \tag{16}$$

Proof : Let $\boldsymbol{u}$, $\boldsymbol{y}$ and $\boldsymbol{x}$ be the optimum solution of the model (14). Clearly,

$$\begin{array}{rcl} Z^*_{skM} & = & \hat{\boldsymbol{r}}'\boldsymbol{x} - \frac{\kappa}{2}\boldsymbol{u}'\boldsymbol{u} \\ & = & \hat{\boldsymbol{r}}'\boldsymbol{x} - \frac{\kappa}{2}\sum_j u_j^2 \\ & \leq & \hat{\boldsymbol{r}}'\boldsymbol{x} - \frac{\kappa}{2}\sum_j (\max\{\sigma_{sk}(\tilde{z}_j)y_j, -\sigma_{sk}(-\tilde{z}_j)y_j\})^2 \\ & \leq & \hat{\boldsymbol{r}}'\boldsymbol{x} + \sum_j C_\kappa(y_j\tilde{z}_j) \\ & = & C_\kappa(\tilde{\boldsymbol{r}}'\boldsymbol{x}) \\ & \leq & Z^*. \end{array}$$

■

We note that the fact that the optimal solution to the skewness-aware optimization problem is a lower bound to the optimal solution from the classical Markowitz model does not mean that the Markowitz model results in better allocations. On the contrary, the skewness-aware model is appropriately conservative when skewness is present in the underlying factors. As our computatonal experiments in Section 5 will illustrate, the skewness-aware model provides a better approximation of the certainty equivalent of the optimal portfolio return for investors with negative exponential utility functions, and better performance for skewed asset distributions.

4.5 Relation to Prospective Satisficing Measure of Brown et al. (2009)

As we mentioned in the introduction, despite its ubiquity, the von Neumann-Morgenstern expected utility theory is in fact inconsistent with some observable behavioral preferences. Notably, the experiments of Allais (1953) and Ellsberg (1961) have shown that one of the important axioms of expected utility theory, the independence axiom, is violated. Another problem with expected utility is the need to specify the risk tolerance parameter, which is an abstract entity, and is often difficult, if not impossible, to elicit from the decision maker.

Recently, Brown et al. (2009) (see also Brown and Sim 2009) proposed a target-based model of choice under uncertainty, which they termed Prospective Satisficing Measure (PSM). The approach can be viewed as a hybrid model, capturing in spirit two celebrated ideas: first, the satisficing concept of Simon (1955), and second, the switch between risk aversion and risk seeking popularized by the prospect theory of Tversky and Kahneman (1979). Specifically, PSM is defined on target premium, $\tilde{v} \in \mathcal{Z}$, which is the difference between an uncertain return, $\tilde{r} \in \mathcal{Z}$ and a target $\tilde{t} \in \mathcal{Z}$, which can be fixed or uncertain. In addition to possessing desirable computational qualities for portfolio optimization and being consistent with second order stochastic dominance, PSM is also able to resolve the classical paradoxes of Allais (1953) and Ellsberg (1961). Brown et al. (2009) fully characterize PSM as a dual of risk measures.

Consider the Entropic PSM,[2] defined as follows

$$\gamma(\tilde{v}) \triangleq \sup\left\{\kappa \in [-\infty,\infty)\backslash\{0\} \;:\; \underbrace{-\frac{1}{\kappa}\ln\mathbb{E}(\exp(-\kappa\tilde{v}))}_{=C_\kappa(\tilde{v})} \geq 0\right\},$$

where $\tilde{v} \in \mathcal{Z}$ is the target premium. Essentially, the Entropic PSM can be viewed as the risk tolerance parameter at which the corresponding certainty equivalence of the target premium breaks even.

[2] The result we will present with respect to the Entropic PSM can be also extended to the Homogenized Entropic PSM described in Brown et al. (2009).

We will now use the skewness-aware portfolio optimization model to write out a specific portfolio allocation formulation for an investor who optimizes the PSM.

Consider a target given by $\tilde{t} = t_0 + \boldsymbol{t}'\tilde{\boldsymbol{z}}$ in which the target premium of a portfolio with given weights $\boldsymbol{x} \in X$ is $\tilde{\boldsymbol{r}}'\boldsymbol{x} - \tilde{t} = (\hat{\boldsymbol{r}}'\boldsymbol{x} - t_0) + (\boldsymbol{A}'\boldsymbol{x} + \boldsymbol{t})'\tilde{\boldsymbol{z}}$. The decision maker optimizes the PSM by solving the following problem

$$\begin{aligned} \Gamma^* = \max \quad & \gamma(\tilde{\boldsymbol{r}}'\boldsymbol{x} - \tilde{t}) \\ \text{s.t.} \quad & \boldsymbol{x} \in X. \end{aligned} \tag{17}$$

We assume that the decision maker selects a target that is reasonably achievable by the underlying returns. Specifically, there exists $\boldsymbol{y} \in X$ such that $\mathbb{E}(\tilde{\boldsymbol{r}}'\boldsymbol{y}) > \mathbb{E}(\tilde{t})$. Since, $\mathbb{E}(\tilde{\boldsymbol{r}}'\boldsymbol{y} - \tilde{t}) > 0$, there exists $\kappa > 0$ such that $C_\kappa(\tilde{\boldsymbol{r}}'\boldsymbol{y} - \tilde{t}) > 0$. Hence, $\Gamma^* > 0$, and we can replace κ by its reciprocal and Model (17)is equivalent to

$$\begin{aligned} (\Gamma^*)^{-1} = \min \quad & a \\ \text{s.t.} \quad & C_{1/a}(\tilde{\boldsymbol{r}}'\boldsymbol{x} - \tilde{t}) \geq 0 \\ & \boldsymbol{x} \in X, a > 0. \end{aligned} \tag{18}$$

Hence, Model (18) can also be viewed as selecting a portfolio in which the target premium has the smallest riskiness index defined by Aumann and Serrano (2008). Observe that the model involves a constraint on certainty equivalent. Under normal distribution,

$$C_{1/a}(\tilde{\boldsymbol{r}}'\boldsymbol{x} - \tilde{t}) = \hat{\boldsymbol{r}}'\boldsymbol{x} - t_0 - \frac{1}{2a}\text{Var}((\boldsymbol{A}'\boldsymbol{x} - \boldsymbol{t})'\tilde{\boldsymbol{z}}). \tag{19}$$

If the underlying factors are not normally distributed, similarly to what we have done to the Markowitz model, we can replace the variance in (19) by the skewness-aware variance to obtain a lower bound of the certainty equivalence. As a result, the optimization problem is a tractable conic quadratic optimization problem as follows:

$$\begin{aligned} (\Gamma^*_{sk})^{-1} = \min \quad & a \\ \text{s.t.} \quad & a(\hat{\boldsymbol{r}}'\boldsymbol{x} - t_0) \geq \tfrac{1}{2}\boldsymbol{u}'\boldsymbol{u} \\ & u_j \geq \sigma_{sk}(\tilde{z}_j)y_j \qquad \forall j \\ & u_j \geq -\sigma_{sk}(-\tilde{z}_j)y_j \qquad \forall j \\ & \boldsymbol{y} = \boldsymbol{A}'\boldsymbol{x} - \boldsymbol{t} \\ & \boldsymbol{x} \in X, a > 0. \end{aligned} \tag{20}$$

Moreover, $\Gamma^*_{sk} \leq \Gamma^*$ and the bound is achieved if the factors are normally distributed.

5 Computational Experiments

We study the behavior of skewness-aware portfolio allocation model (14) in computational experiments with simulated and real market data (Section 5.1), and explore the consistency of the skewness-aware beta with returns observed in the NYSE (Section 5.2).

5.1 Skewness-Aware Portfolio Allocation

This section observes some characteristics of portfolio allocations under the skewness-aware risk measure.

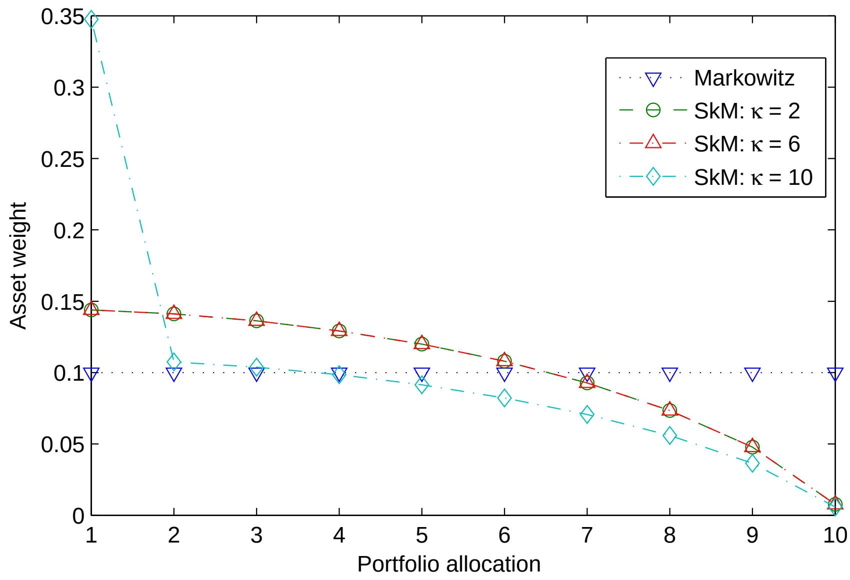

Figure 1: Optimal portfolio weights (as proportions) for assets numbered 1 through 10 resulting from the classical Markowitz and the skewness-aware Markowitz (SkM) models.

5.1.1 Controlled Experiments with Simulated Data

Consider a portfolio of $N = 10$ assets with uncertain returns $\tilde{r}_i$, $i = 1, \ldots, N$. Each return $\tilde{r}_i$ is determined by a simple single factor model $\tilde{r}_i = \hat{r}_i + \tilde{z}(\omega_i)$, where $\hat{r}_i = 1$. The factors $\tilde{z}(\omega_i)$ are independent and distributed as in (3). Similarly to the example in Section 3, all random stock returns $\tilde{r}_i$ have the same mean and standard deviation; however, depending on the parameters ω_i, $i = 1, \ldots, N$, the degree of asymmetry of each individual return distribution can be different. Higher values for ω_i result in larger negative skew. We generate values for ω_i as follows:

$$\omega_i = \frac{1}{2}\left(1 + \frac{i}{N + 0.01}\right), \quad i = 1, \ldots, N.$$

Therefore, the return distributions for stocks with high index numbers in the portfolio are more negatively skewed than those for stocks with low index numbers.

The optimal asset allocations obtained with the classical Markowitz model and with the skewness-aware Markowitz model (14) are plotted in Figure 1. One can observe that the mean-generalized deviation model results in smaller weights for the assets with more negatively skewed returns, while the Markowitz model allocates equally in all assets.

Figure 2 contains the optimal objective function values and the realized values of the certainty equivalent for a negative exponential utility function over 1000 simulated observations for asset returns from the two-point distribution (3). The mean-generalized deviation model significantly dominates the classical mean-variance model in terms of realized certainty equivalent return.

5.1.2 Experiments with Real Market Data

We apply the skewness-aware model (14) to realized returns from a comprehensive set of global stock markets. We obtain daily US$ total returns of the MSCI indices for 48 countries over the period

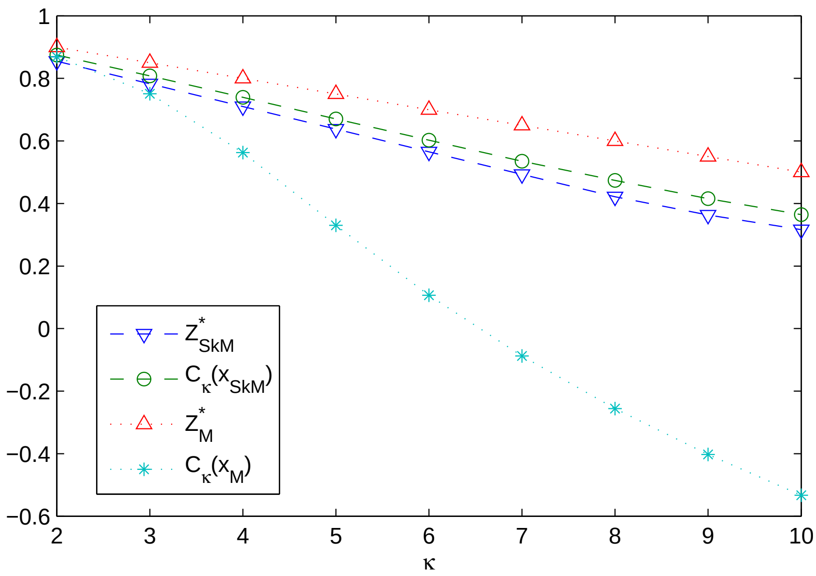

Figure 2: Plot of objectives and realized certainty equivalents.

January 2001 to August 2006 from Datastream. Data from the first 4 years are used to calibrate parameters (*training* or *in-sample* period), while data from the last 20 months are used to test the performance of model (14) against the traditional Markowitz model (*testing* or *out-of-sample* period).

Table 2 summarizes the characteristics of country returns distributions in the training and testing periods, while Table 3 checks for rank ordering stability of moments across the two periods. Countries ranked high on mean returns and standard deviation in the training period are again ranked high on the same moments in the testing period. The correlation coefficients across time for skewness and kurtosis are also positive but not significant at the 5% level.

Table 4 reports portfolio allocation across the 48 countries when there is no constraint on short-sales. In general, the total proportion of short-sales is inversely related to κ - optimal portfolios for risk-neutral investors comprise more short-sales than those for risk-averse investors. At small values of κ (towards risk neutrality), the skewness-aware allocation model generates less short-sales than the traditional Markowitz model. However, at $\kappa > 1$ (more risk aversion), total short-sales from the skewness-aware allocation model become greater than those from the traditional Markowitz model.

To illustrate the differences in allocation across the two models, we take the example of $\kappa = 1$. There are 6 common countries among the top 10 short positions from both models. The 4 different countries for the skewness-aware model are Ireland, Spain, Canada and Norway (SkM:ISCN). The 4 different countries for the traditional Markowitz model are Portugal, Japan, Denmark and Poland (M:PJDP). As a group, SkM:ISCN (average skewness rank 14) is more negatively skewed than M:PJDP (average skewness rank 23). The kurtosis of the SkM:ISCN group (average kurtosis rank 24) is higher than that of the M:PJDP group (average kurtosis rank 10). The skewness-aware model picks out countries with more negative-skewed and sharp-peak/heavy-tail returns during the training period as candidates for short selling.

In the case of the top 10 long positions, there are 8 countries common in the two models.

	January 2001 - December 2004								January 2005 - August 2006							
	Mean (%)		Std dev		Skewness		Kurtosis		Mean(%)		Std dev		Skewness		Kurtosis	
AUSTRALIA	0.07	31	0.011	5	-0.36	11	5.53	22	0.08	22	0.01	17	-0.36	13	4	13
AUSTRIA	0.11	37	0.011	5	-0.3	15	4.2	6	0.09	26	0.011	21	-0.47	7	6.53	39
BELGIUM	0.05	24	0.014	21	0.18	41	7.53	38	0.07	19	0.009	10	-0.29	19	5.16	26
CANADA	0.03	16	0.011	6	-0.48	7	6.57	31	0.09	29	0.01	15	-0.15	31	3.42	2
DENMARK	0.04	19	0.012	9	-0.26	16	5.74	25	0.09	30	0.01	16	-0.71	2	6.52	37
FINLAND	-0.02	1	0.026	45	-0.25	17	6.7	33	0.09	28	0.012	28	0.09	43	5.54	29
FRANCE	0.02	7	0.015	28	-0.05	31	4.88	13	0.07	20	0.01	12	-0.17	29	4.52	22
GERMANY	0.02	8	0.017	36	-0.06	28	4.75	11	0.07	16	0.01	18	-0.12	32	4.2	16
HONG KONG	0.02	12	0.013	15	-0.19	19	6.38	28	0.05	7	0.007	3	-0.27	21	4.29	17
IRELAND	0.05	23	0.013	16	-0.54	6	6.51	29	0.05	8	0.011	22	-0.93	1	10.69	48
ITALY	0.03	14	0.013	17	-0.24	18	5.66	23	0.05	9	0.009	7	-0.08	35	3.62	4
JAPAN	0.01	5	0.015	24	-0.1	26	4.48	9	0.07	17	0.012	25	-0.1	34	4	12
NETHERLANDS	0.00	3	0.016	31	-0.09	27	5.71	24	0.08	24	0.009	11	-0.07	36	4.84	23
NEW ZEALAND	0.11	38	0.012	10	-0.62	4	6.62	32	0.00	1	0.01	14	-0.18	27	3.27	1
NORWAY	0.07	28	0.013	14	-0.61	5	5.85	26	0.12	36	0.015	34	-0.35	17	6.53	38
PORTUGAL	0.03	13	0.011	8	-0.32	13	4.17	5	0.06	10	0.008	4	0.58	48	6.17	35
SINGAPORE	0.02	10	0.012	12	-0.05	30	5.4	18	0.07	18	0.008	6	-0.58	4	6.62	41
SPAIN	0.05	25	0.015	26	0.09	36	4.39	8	0.07	15	0.009	9	-0.05	39	4.16	15
SWEDEN	0.03	15	0.019	39	0.03	35	5.51	20	0.07	14	0.012	26	0.02	42	7.26	43
SWITZERLAND	0.02	9	0.013	18	0	33	6.55	30	0.08	23	0.009	8	-0.23	24	4.51	21
UK	0.02	11	0.012	13	-0.16	23	5.34	16	0.06	11	0.008	5	-0.05	38	4.3	18
USA	0.01	4	0.012	11	0.25	45	5.2	15	0.03	2	0.007	2	0.12	47	3.46	3
ARGENTINA	0.04	21	0.028	47	-1.07	2	20.84	45	0.19	46	0.019	42	-0.01	40	3.71	6
BRAZIL	0.07	30	0.021	41	0.14	37	6.71	34	0.16	45	0.02	44	-0.4	11	3.82	9
CHILE	0.07	27	0.011	7	-0.31	14	4.12	4	0.06	12	0.01	13	-0.48	6	4.08	14
CHINA	0.04	17	0.018	37	-0.18	20	5.37	17	0.11	35	0.012	24	-0.36	12	5.09	24
CZECH REPUBLIC	0.15	48	0.015	27	-0.11	25	4.52	10	0.13	40	0.016	35	-0.05	37	6.57	40
EGYPT	0.10	36	0.016	33	0.21	42	7.28	37	0.25	48	0.02	45	0.1	44	6.02	33
HUNGARY	0.12	40	0.015	25	-0.17	22	4.87	12	0.06	13	0.019	41	-0.16	30	3.98	11
INDIA	0.07	29	0.014	23	-0.69	3	9.4	40	0.13	39	0.015	33	-0.35	14	6.66	42
INDONESIA	0.14	47	0.021	42	-0.42	9	10.28	42	0.13	37	0.018	38	-0.42	10	8.62	45
ISRAEL	0.00	2	0.016	30	0.02	34	6.05	27	0.04	5	0.011	19	-0.34	18	6.1	34
JORDAN	0.13	43	0.01	3	-0.43	8	13.29	44	0.09	27	0.017	37	-0.35	15	6.36	36
KOREA	0.12	39	0.021	40	-0.06	29	5.12	14	0.13	38	0.013	30	-0.28	20	3.64	5
MALAYSIA	0.05	22	0.01	2	-0.38	10	9.49	41	0.04	4	0.006	1	-0.24	22	5.32	27
MEXICO	0.08	32	0.014	22	-0.04	32	5.49	19	0.13	41	0.015	32	0.11	46	5.69	32
MOROCCO	0.04	18	0.009	1	0.34	46	6.9	35	0.14	42	0.011	20	-0.2	26	5.55	30
PAKISTAN	0.13	43	0.017	34	0.24	43	7.25	36	0.14	43	0.021	46	-0.22	25	3.81	7
PERU	0.12	41	0.013	19	-0.36	12	7.81	39	0.16	44	0.017	36	-0.23	23	4.39	19
PHILIPPINES	0.02	6	0.015	29	1.82	47	23.25	46	0.10	31	0.014	31	0.02	41	5.13	25
POLAND	0.06	26	0.016	32	0.16	38	3.79	1	0.11	34	0.018	39	-0.11	33	3.82	8
RUSSIA	0.14	46	0.022	44	-0.16	24	5.51	21	0.23	47	0.019	43	-0.56	5	8.37	44
SOUTH AFRICA	0.09	33	0.014	20	-0.17	21	3.95	2	0.07	21	0.018	40	-0.46	8	5.66	31
SRI LANKA	0.13	45	0.022	43	2.63	48	51.87	48	0.10	32	0.012	27	-0.45	9	10.33	47
TAIWAN	0.04	20	0.018	38	0.17	39	4.02	3	0.03	3	0.012	23	-0.35	16	4.41	20
THAILAND	0.13	43	0.017	35	0.24	44	4.36	7	0.05	6	0.012	29	0.1	45	3.92	10
TURKEY	0.10	35	0.037	48	0.18	40	10.31	43	0.10	33	0.024	47	-0.7	3	5.53	28
VENEZUELA	0.09	34	0.028	46	-1.22	1	43.39	47	0.09	25	0.028	48	-0.18	28	8.63	46

Table 2: Summary statistics of daily returns distributions in 48 stock markets. Country ranks for each moment are provided next to the moment value.

Moments	Spearman's ρ	Normal Z-statistic	2-sided p-value
Mean	0.42	2.8793	0.0040
Std dev	0.59	4.0765	0.0000
Skewness	0.22	1.5275	0.1266
Kurtosis	0.24	1.6272	0.1037

Table 3: Correlation of moments across periods.

	$\kappa = 0.05$		$\kappa = 0.1$		$\kappa = 0.5$		$\kappa = 1$		$\kappa = 1.5$		$\kappa = 2$	
	SkM	M	SkM	M	SkM	M	SkM	M	SkM	M	SkM	M
AUSTRALIA	-4.8	-15.2	-1.7	-7.4	1.5	-1.3	1.9	-0.6	2	-0.3	2.1	-0.2
AUSTRIA	79.3	127.1	42.1	65.5	12.5	16.1	8.7	10	7.4	7.9	6.8	6.9
BELGIUM	79.1	118.2	38.9	58.4	7.1	10.5	3.3	4.5	2	2.5	1.3	1.5
CANADA	0.9	0.8	0.9	2	-1.8	3	-2.1	3.1	-2.3	3.1	-2.3	3.2
DENMARK	-58.4	-81.6	-25.1	-39.3	-1.2	-5.5	0.6	-1.3	1.2	0.1	1.5	0.8
FINLAND	-9.4	-22.4	-4.4	-11	-0.2	-1.8	0.3	-0.7	0.5	-0.3	0.6	-0.1
FRANCE	-13.4	-54.8	-7.8	-27.9	-3.3	-6.5	-2.7	-3.8	-2.3	-2.9	-2.1	-2.5
GERMANY	18.6	40.6	5.4	17.1	-4.9	-1.5	-6	-3.9	-6.5	-4.6	-6.7	-5
HONG KONG	-63	-84.3	-27.6	-40.3	-0.4	-5.1	1.3	-0.7	1.8	0.8	2.1	1.5
IRELAND	-14.9	-26.9	-6.6	-13.2	-3.1	-2.2	-3	-0.8	-3	-0.4	-3	-0.1
ITALY	1.1	-23.6	3.8	-9.3	7.1	2.3	7.5	3.7	7.6	4.2	7.6	4.4
JAPAN	-32.4	-64.5	-16.3	-32.3	-3.2	-6.5	0.1	-3.3	1.3	-2.2	1.8	-1.7
NETHERLANDS	-113.8	-124.2	-59.5	-64.2	-16.3	-16.1	-11	-10.1	-9.2	-8.1	-8.4	-7.1
NEW ZEALAND	44.3	74	24	38.8	8.1	10.7	6.2	7.2	5.5	6	5.2	5.4
NORWAY	13.2	16.9	5.5	7.7	-0.6	0.3	-1.2	-0.6	-1.4	-0.9	-1.5	-1.1
PORTUGAL	-111.8	-164.9	-51.1	-80.5	-2.5	-12.9	3.2	-4.5	5	-1.7	6	-0.3
SINGAPORE	-38.2	-70.2	-19.9	-35.6	-4.5	-8	-2	-4.5	-1.1	-3.4	-0.7	-2.8
SPAIN	33.4	96	14.4	47	-1.2	7.7	-3	2.8	-3.6	1.1	-4	0.3
SWEDEN	3.7	19.1	-2.9	6.3	-8.4	-3.9	-9	-5.1	-9.2	-5.6	-9.3	-5.8
SWITZERLAND	0.7	-41.1	-0.7	-19.2	2.1	-1.7	3.5	0.5	4	1.2	4.2	1.6
UK	17.8	8.2	17.3	9.5	15.9	10.5	15.7	10.6	15.6	10.7	15.5	10.7
USA	-29	-67.7	-7.4	-28.4	6.6	3	7.7	6.9	8.1	8.2	8.3	8.9
ARGENTINA	-3.9	-8.2	-1.9	-3.6	-0.6	0	-0.5	0.5	-0.4	0.6	-0.4	0.7
BRAZIL	4.1	7.9	0.9	3	-1.7	-1	-2.1	-1.4	-2.2	-1.6	-2.2	-1.7
CHILE	19.4	4.9	16.7	5.9	15.6	6.7	15.4	6.8	15.3	6.9	15.3	6.9
CHINA	15	22.3	6.2	10.4	-0.6	0.9	-1.1	-0.3	-1.3	-0.7	-1.4	-0.9
CZECH REPUBLIC	31.9	90.3	15	44.9	1.5	8.5	-0.2	4	-0.7	2.5	-1	1.7
EGYPT	8.6	9.1	6.7	6.6	5.5	4.7	5.3	4.4	5.2	4.3	5.2	4.3
HUNGARY	42.8	42.8	23	22.1	7.3	5.6	5.3	3.5	4.7	2.8	4.4	2.5
INDIA	8	16.3	4.3	9.7	1.3	4.4	0.9	3.8	0.8	3.6	0.7	3.5
INDONESIA	1.2	36.5	-0.1	18.2	-1	3.5	-1.2	1.7	-1.2	1	-1.2	0.7
ISRAEL	-7.8	-26.3	-3.9	-12.7	0.2	-1.7	0.9	-0.3	1.1	0.1	1.2	0.3
JORDAN	34	118.8	20.3	66.2	9.9	24.1	8.6	18.8	8.2	17.1	7.9	16.2
KOREA	46.1	74	22.1	36.3	2.9	6.2	0.3	2.4	-0.6	1.2	-1	0.6
MALAYSIA	-19.6	-61.2	-4.1	-24.7	4.1	4.6	5	8.2	5.3	9.4	5.4	10
MEXICO	29.3	44.1	15.8	22.7	6.2	5.6	5.1	3.5	4.7	2.8	4.5	2.4
MOROCCO	-9.4	-62.6	2.6	-22.3	14.7	10	16.1	14.1	16.6	15.4	16.8	16.1
PAKISTAN	36.6	52	21.1	28.2	9	9.2	7.4	6.8	6.9	6	6.7	5.6
PERU	22.6	71.2	11.8	36.8	3.4	9.4	2.3	5.9	1.9	4.8	1.7	4.2
PHILIPPINES	-12.5	-70	-6.6	-33.6	0	-4.5	1.7	-0.8	2.2	0.4	2.5	1
POLAND	-20.2	-35.4	-9.1	-17.3	0.9	-2.8	2.2	-1	2.7	-0.4	2.9	-0.1
RUSSIA	22.5	39.7	11.8	20.5	3.4	5.2	2.4	3.2	2	2.6	1.9	2.3
SOUTH AFRICA	5.4	8.9	2.2	3.8	-0.5	-0.3	-0.9	-0.8	-1.1	-1	-1.2	-1.1
SRI LANKA	14.8	43.1	7.9	22.9	2.6	6.8	1.9	4.8	1.7	4.2	1.6	3.8
TAIWAN	-26.1	-40.1	-11.8	-19.5	-0.2	-3	1.2	-0.9	1.6	-0.3	1.9	0.1
THAILAND	54.1	53.6	28.4	27	8.2	5.7	5.5	3	4.7	2.2	4.2	1.7
TURKEY	0.3	3.9	-0.4	1.6	-1	-0.3	-1.1	-0.5	-1.1	-0.6	-1.1	-0.7
VENEZUELA	-0.5	5.1	-0.4	3.1	-0.3	1.5	-0.3	1.3	-0.2	1.3	-0.2	1.2

Table 4: Weights of each country (%) in optimal portfolios across different values of κ.

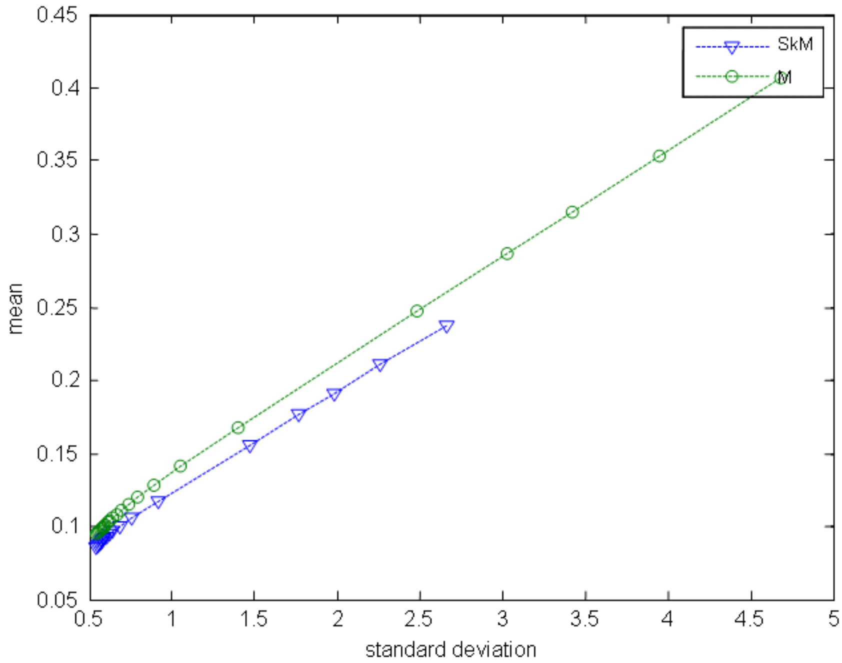

Figure 3: Realized efficient frontiers for the Markowitz and the SkM models.

The two different countries for the skewness-aware model are Thailand and Italy (SkM:TI). The two different countries for the traditional Markowitz model are Peru and Malaysia (M:PM). The SkM:TI group is more strongly and positively skewed (average rank skewness 31) than the M:PM group (average skewness rank 11). On the other hand, the SkM:TI group is much lower in kurtosis than the M:PM group (average kurtosis rank 15 vs. 40). The SkM model places long positions on countries with more positive-skewed and round-peak/thin-tail returns.

The out-of-sample performance of the models is depicted in Figures 3 and 4. The mean-standard deviation efficient frontier is higher for the traditional Markowitz model compared to that for the skewness-aware model. However, when we look at realized values of certainty equivalent, the skewness-aware model is superior for all values of κ.

The SkM model's superiority lies in the model's ability to separate markets that are positively skewed from those that are negatively skewed. Further, it is also able to distinguish the heavy positive tail from the heavy negative tail. However, we would caution that these abilities of the SkM model can be translated into superior performance over the traditional Markowitz model only if the third and fourth higher moments of returns distribution are relatively stable over time. In our market data experiment, although the autocorrelations of skewness and kurtosis are not strongly positive, they are apparently sufficient for the superior performance of the SkM model.

5.2 Risk and Return Characteristics of Portfolios Sorted by Classical and Skewness-Aware Beta

In this section, we test whether a higher sensitivity to market return as measured by the skewness-aware beta translates into higher realized returns. We compare the effect of the skewness-aware

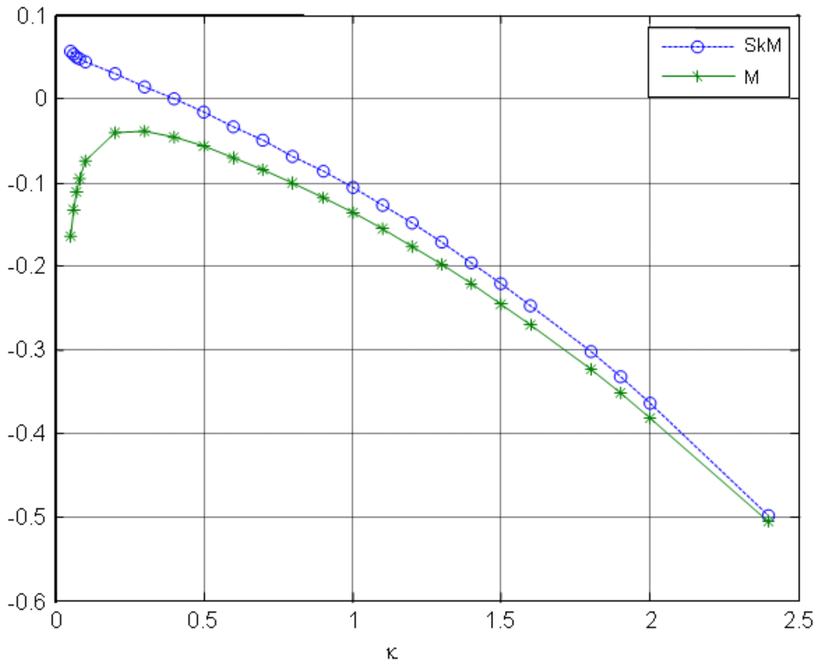

Figure 4: Realized certainty equivalents for the Markowitz and the SkM models.

beta to the classical beta. Technically, the skewness-aware beta is applicable only when the market portfolio is the optimal portfolio for investors according to (9). However, given the ubiquity of use of the classical beta (whose validity is also based on assumptions about investor behavior or the shape of return distributions) in reporting the sensitivity of a stock's return to the market return, it is important to study whether there are analytically similar, but more accurate, estimates of the risk-return characteristics of assets with skewed return distributions.

We consider historical data on daily stock returns traded at the NYSE from 1963 till 2006. Starting at the beginning of 1970, stocks are used for estimation only if they have full daily return history for the previous 7 years (1963-1969). (We selected 7 years as the window length, because the number of observations is large enough for stable estimation, but does not extend too far into the past.) The classical beta and the skewness-aware beta are computed for each stock using the 7-year window of training returns, for a total of 37 time periods. The classical beta is the slope of the OLS regression of stock daily returns against a market proxy of equally-weighted daily returns of all the stocks in our sample. The skewness-aware beta is computed from (13).

First, we sort the stocks at every time period by their betas or skewness-aware betas into three portfolios using the 30th and 70th percentiles as breakpoints. These equally-weighted portfolios are held for the next 12 months and are rebalanced at the beginning of the following year by re-sorting on classical betas or skewness-aware betas, respectively, computed from rolling seven-year windows. This procedure is repeated up to year 2006 with 1999-2005 as the last training window, and the average betas and realized returns for each group of portfolios over the 37 time periods are presented in Table 5 and Figure 5.

As Table 5 and Figure 5 illustrate, portfolios sorted by both classical beta and skewness-aware

		Low-30%	Mid-40%	High-30%
Portfolio sorted by classical beta	Average daily returns	0.0622%	0.0694%	0.0706%
	Average beta	0.460	0.969	1.577
Portfolios sorted by skewness-aware beta	Average daily returns	0.0626%	0.0695%	0.0701%
	Average skewness-aware beta	0.454	0.961	1.595

Table 5: Average realized returns and betas.

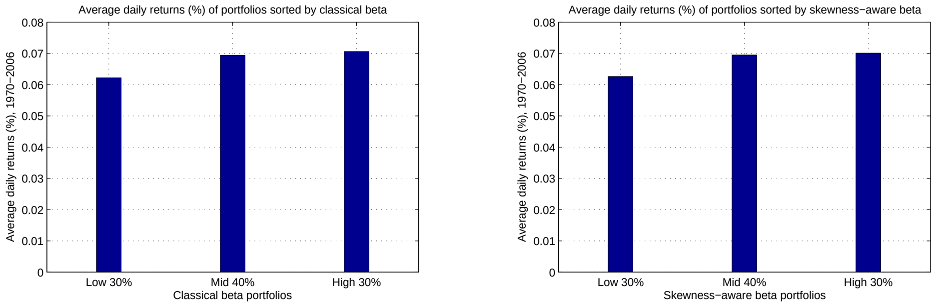

Figure 5: Bar charts of realized average returns for sorted portfolios according to classical beta and skewness-aware beta.

beta exhibit logical risk-return tradeoff characteristics - portfolios in the upper percentiles in terms of skewness-aware betas also deliver higher realized returns on average.

Next, we repeat the empirical study after partitioning the NYSE stock returns into three samples according to the individual stock's Fisher's g_1 (skewness) value. Sample I comprises stocks with negatively skewed returns ($g_1 < -0.1$) in the 7-year window. These are arguably less attractive stocks because of their negatively biased return distributions. As illustrated in Table 6 and Figure 6, portfolios constructed by sorting on skewness-aware beta continue to reveal a monotonic high-risk-high-return relation. However, this relation is not satisfied for the classical beta-sorted portfolios in this sample. In Samples II (Fisher's $-0.1 < g_1 < 0.1$) and III (Fisher's $g_1 > 0.1$), the monotonic risk-return relation is present in both the classical and the skewness-aware beta-sorted portfolio sets.

These experiments provide evidence that skewness is priced in market returns, and suggest that the skewness-aware beta may be a good measure for estimating its effect. Although the

		Low-30%	Mid-40%	High-30%
Sample I (Negative skewness)	Classical beta	0.0524%	0.0623%	0.0613%
	Skewness-aware beta	0.0505%	0.0615%	0.0646%
Sample II (Zero skewness)	Classical beta	0.0521%	0.0598%	0.0601%
	Skewness-aware beta	0.0520%	0.0578%	0.0628%
Sample III (Positive skewness)	Classical beta	0.0695%	0.0714%	0.0776%
	Skewness-aware beta	0.0701%	0.0724%	0.0757%

Table 6: Average realized daily returns (1970-2006) for portfolios sorted by classical beta or skewness-aware beta. Classical beta and skewness-aware beta are estimated with equally-weighted in-sample average as market proxy. Stocks are dividided into three samples according to skewness.

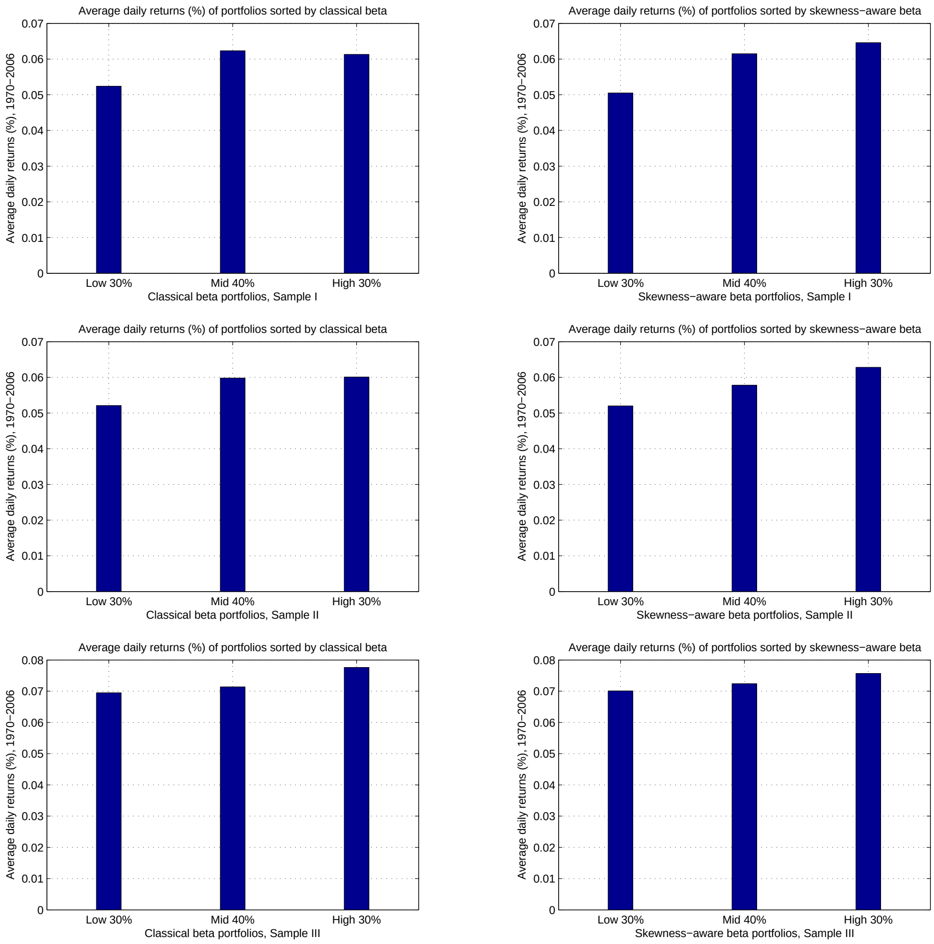

Figure 6: Bar charts of realized average returns for each of the three samples of returns. (Each row compares the results for portfolios sorted by the classical beta and the skewness-aware beta for Samples I, II and III, respectively.)

skewness-aware beta is empirically similar to the classical beta for this data set, the consideration of skewness enables the new skewness-aware beta to provide a more robust risk-based description of the cross-section of security returns than the classical beta.

We reran the experiments using value-weighted, rather than equally weighted portfolios. The results are consistent with the results presented in this section, and are available from the authors upon request. In fact, when we use value weighted portfolios (both for classical beta regression and ex-post portfolio returns accumulation), the monotonic risk-return relation is maintained for the skewness-aware beta, but weakened for the classical beta.

6 Concluding Remarks

This paper introduced a new measure of distributional skewness which allowed for generalizing the concepts of variance, covariance, and standard deviation. Minimizing skewness-aware measures of variability can be linked to expected utility maximization, Prospect Satisficing Measures, and optimizing tail-deviation measures such as Value-at-Risk, and can be used to generalize the classical Markowitz portfolio allocation framework. Experiments with simulated and real market data indicate that the proposed approach accounts for skewness in asset returns when deciding optimal asset allocations. It does it in an intuitively appealing and computationally efficient way, by implicitly incorporating higher moment distributional information. Computational studies of the historical return behavior of stocks traded at the NYSE also indicate that a skewness-aware beta may be a better way for accounting for the sensitivity of stocks to market returns than the classical CAPM beta.

References

Allais, M. (1953): Le Comportement de l'Homme Rationel devant le Risque: Critique des Portulates et Axiomes de l'Ecole Americaine. *Econometrica*, 21(4), 503-546

Aumann, R., and R. Serrano (2008): An Economic Index of Riskiness, Journal of Political Economy, 116(5), 810-836

Brandt, M., P. Santa-Clara, and R. Valkanov (2007): Parametric Portfolio Policies: Exploiting Characteristics in the Cross Section of Equity Returns. Forthcoming in *Review of Financial Studies*

Brown, D.B., E.G. De Giorgi and M. Sim (2009): A Satisficing Alternative to Prospect Theory. *Working Paper*. Available at SSRN: `http://ssrn.com/abstract=1406399`

Brown, D. B. and M. Sim (2009): Satisficing Measures for Analysis of Risky Positions. *Management Science*, 55(1), 71-84

Chamberlain, G. (1983): A Characterization of the Distributions That Imply Mean-Variance Utility Functions. *Journal of Economic Theory*, 29: 185-201

Ellsberg, D. (1961): Risk, Ambiguity, and the Savage Axioms.*Quarterly Journal of Economics*, 75(4), 643-669

Fama, E. (1976): *Foundations of Finance*, Basic Books, New York, NY

Harvey, C., M. Liechty, J. Liechty and P. Muller (2004): Portfolio Selection with Higher Moments. *Working paper*, Duke University

Hwang, S. and S. Satchell (1999): Modeling Emerging Market Risk Premia Using Higher Moments. *International Journal of Finance and Economics*, 4: 271-296

Jorion, P. (2000): *Value at Risk: The New Benchmark for Managing Financial Risk*, McGraw-Hill

Kahneman, D., and A. Tversky (1979): Prospect Theory: An Analysis of Decisions Under Risk. *Econometrica*, 47(2), 263-291

Kraus, A. and R. Litzenberger (1976): Skewness Preference and the Valuation of Risky Assets. *Journal of Finance*, 31(4): 1085-1100

Levy, H. (1992): Stochastic Dominance and Expected Utility: Survey and Analysis. *Management Science*, 38(4): 555-593

Lintner, J. (1965): The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *Review of Economics and Statistics*, 47(1): 13-37

Markowitz, H.M. (1952): Portfolio Selection. *Journal of Finance*, 7:77-91

Markowitz, H. M. (1959): *Portfolio Selection: Efficient Diversification of Investments.* John Wiley & Sons, New Jersey

Markowitz, H.M., Todd, P., Xu, G., and Yamane, Y. (1993): Computation of Mean-Semivariance Efficient Sets by the Critical Line Algorithm. *Annals of Operations Research*, 45: 307-317

Quiggin, J. (1982): A Theory of Anticipated Utility. *Journal of Economic Behavior and Organization*, 3(4), 225-243

Rockafellar, R.T. and S. Uryasev (2000): Optimization of Conditional Value-At-Risk. *The Journal of Risk*, 2(3): 21-41

Sharpe, W. (1964): Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance*, 19: 425-442

Simon, H. A. (1955): A Behavioral Model of Rational Choice. *Quarterly Journal of Economics*, 69(1), 99-118

Tversky, A., and D. Kahneman (1992): Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*. 5, 297-323

von Neumann, J., and Morgenstern, O. (1944): *Theory of Games and Economic Behavior.* Princeton University Press, Princeton, NJ

Appendix: Proofs

Proof of Proposition 1:

The moment generating function of a random variable $\tilde{z} \sim \mathcal{N}(0, \sigma)$ is given by

$$\mathbb{E}(\exp(\theta\tilde{z})) = \exp(\theta^2\sigma^2/2).$$

Hence,

$$\begin{aligned}\xi^2(\tilde{z}) &= \sup_{\theta>0}\left\{\frac{2}{\theta^2}\ln(\mathbb{E}(\exp(-\theta\tilde{z}))\right\} \\ &= \sup_{\theta>0}\left\{\frac{2}{\theta^2}\ln(\exp(\theta^2\sigma^2/2))\right\} \\ &= \sigma^2(\tilde{z}).\end{aligned}$$

■

Proof of Proposition 2:

For notational convenience, let $\xi^2 = \xi^2(\tilde{z})$ and $\sigma^2 = \sigma^2(\tilde{z})$. Notice that for all $\theta > 0$, we have

$$\mathbb{E}\left(\exp(-\theta\tilde{z})\right) = 1 + \frac{1}{2}\theta^2\sigma^2 + \sum_{k=3}^{\infty}\frac{(-\theta)^k\mathbb{E}(\tilde{z}^k)}{k!}$$

and

$$\exp\left(\frac{\theta^2\xi^2}{2}\right) = 1 + \frac{1}{2}\theta^2\sigma_{sk}^2 + \sum_{k=2}^{\infty}\frac{(\theta\xi)^{2k}}{2^k k!}.$$

From the definition of ξ^2 in (1), we have

$$\xi^2 \geq \frac{2}{\theta^2}\ln(\mathbb{E}(\exp(-\theta\tilde{z})),$$

or, equivalently,

$$\exp\left(\frac{\theta^2\xi^2}{2}\right) \geq \mathbb{E}\left(\exp(-\theta\tilde{z})\right) \qquad \forall\theta > 0.$$

If we choose θ to be close to zero, then

$$\frac{1}{2}\theta^2\sigma^2 \leq \frac{1}{2}\theta^2\xi^2.$$

Thus, $\xi^2 \geq \sigma^2$. ■

Proof of Proposition 3:

The first two conditions are straightforward. Note that

$$y_j\phi_j(\boldsymbol{y}) = \begin{cases} \xi^2(\tilde{z}_j)y_j^2 & \text{if } y_j \geq 0 \\ \xi^2(-\tilde{z}_j)y_j^2 & \text{otherwise} \end{cases}$$

Hence, noting that $\xi(\tilde{z}_j),\ \xi(-\tilde{z}_j) > 0$, we have

$$\boldsymbol{y}'\boldsymbol{\phi}(\boldsymbol{y}) = \sum_{j=1}^{M}(\max\{\xi(\tilde{z}_j)y_j, -\xi(-\tilde{z}_j)y_j\})^2$$

Moreover,

$$\begin{aligned} \mathrm{Var}_{\mathrm{sk}}(y_0 + \tilde{\boldsymbol{z}}'\boldsymbol{y}) &= \min\left\{\boldsymbol{u}'\boldsymbol{u} \ :\ \exists\boldsymbol{u} \in \mathcal{R}^M \ :\ u_j = \max\{\xi(\tilde{z}_j)y_j, -\xi(-\tilde{z}_j)y_j\},\ j = 1,\ldots,M\right\} \\ &= \min\left\{\boldsymbol{u}'\boldsymbol{u} \ :\ \exists\boldsymbol{u} \in \mathcal{R}^M \ :\ u_j \geq \xi(\tilde{z}_j)y_j,\ u_j \geq -\xi(-\tilde{z}_j)y_j,\ j = 1,\ldots,M\right\}. \end{aligned}$$

■

Proof of Proposition 4:

The first two conditions are straightforward. From Proposition 3(c), we write

$$\sigma_{sk}(y_0 + \tilde{\boldsymbol{z}}'\boldsymbol{y}) = \min\left\{\|\boldsymbol{u}\|_2 \ :\ \exists\boldsymbol{u} \in \mathcal{R}^M \ :\ u_j \geq \sigma_{sk}(\tilde{z}_j)y_j, u_j \geq -\sigma_{sk}(-\tilde{z}_j)y_j,\ j = 1,\ldots,M\right\},$$

which is a convex optimization problem. Let

$$\sigma_{sk}(y_0^1 + \tilde{\boldsymbol{z}}'\boldsymbol{y}^1) = \|\boldsymbol{u}^1\|_2$$

for some

$$u_j^1 \geq \xi(\tilde{z}_j)y_j^1, u_j^1 \geq -\xi(-\tilde{z}_j)y_j^1,\ j = 1,\ldots,M$$

and

$$\sigma_{sk}(y_0^2 + \tilde{\boldsymbol{z}}'\boldsymbol{y}^2) = \|\boldsymbol{u}^2\|_2$$

for some

$$u_j^2 \geq \xi(\tilde{z}_j)y_j^2, u_j^2 \geq -\xi(-\tilde{z}_j)y_j^2,\ j = 1,\ldots,M.$$

Let

$$(y_0^3, \boldsymbol{y}^3, \boldsymbol{u}^3) = (y_0^1, \boldsymbol{y}^1, \boldsymbol{u}^1) + (y_0^2, \boldsymbol{y}^2, \boldsymbol{u}^2).$$

Clearly, we have

$$u_j^3 \geq \sigma_{sk}(\tilde{z}_j) y_j^3, u_j^3 \geq -\sigma_{sk}(-\tilde{z}_j) y_j^3, \ j = 1, \ldots, M.$$

Therefore

$$\begin{array}{rcl} \sigma_{sk}(y_0^1 + \tilde{\boldsymbol{z}}'\boldsymbol{y}^1) + \sigma_{sk}(y_0^2 + \tilde{\boldsymbol{z}}'\boldsymbol{y}^2) & = & \|\boldsymbol{u}^1\|_2 + \|\boldsymbol{u}^2\|_2 \\ & \geq & \|\boldsymbol{u}^1 + \boldsymbol{u}^2\|_2 \\ & = & \|\boldsymbol{u}^3\|_2 \\ & \geq & \sigma_{sk}(y_0^1 + y_0^2 + \tilde{\boldsymbol{z}}'(\boldsymbol{y}^1 + \boldsymbol{y}^2)). \end{array}$$

■

Proof of Proposition 5:

Since $\tilde{y} \in \mathcal{Z}$, let $\tilde{y} = y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}$ and it suffices to show that

$$\mathbb{P}\left(\boldsymbol{y}'\tilde{\boldsymbol{z}} < -a\sqrt{\boldsymbol{y}'\boldsymbol{\phi}(\boldsymbol{y})}\right) \leq \exp(-a^2/2).$$

The proof of the upper-tail bound follows from switching the sign of $\tilde{y}$. Indeed,

$$\begin{array}{rll} & \mathbb{P}\left(\boldsymbol{y}'\tilde{\boldsymbol{z}} < -a\underbrace{\sqrt{\boldsymbol{y}'\boldsymbol{\phi}(\boldsymbol{y})}}_{=s}\right) & \\ = & \mathbb{P}\left(\exp(-\theta\boldsymbol{y}'\tilde{\boldsymbol{z}}) > \exp(\theta a s)\right) & \forall\theta > 0 \\ \leq & \mathbb{E}(\exp(-\theta\boldsymbol{y}'\tilde{\boldsymbol{z}}))/\exp(\theta a s) & \text{Markov inequality.} \end{array}$$

Since $\tilde{z}_j$ are independently distributed, we have

$$\mathbb{E}(\exp(-\theta\boldsymbol{y}'\tilde{\boldsymbol{z}})) = \exp\left(\sum_{j=1}^{M} \ln \mathbb{E}(\exp(-\theta y_j \tilde{z}_j))\right).$$

However, since

$$\sigma_{sk}^2(\tilde{z}_j) \stackrel{\Delta}{=} \sup_{\theta>0}\left\{\frac{2}{\theta^2}\ln(\mathbb{E}(\exp(-\theta\tilde{z}))\right\},$$

we have

$$\ln(\mathbb{E}(\exp(-\theta\tilde{z}_j)) \leq \frac{\sigma_{sk}^2(\tilde{z}_j)\theta^2}{2} \qquad \forall\theta > 0.$$

Suppose $y_j > 0$, we have

$$\ln(\mathbb{E}(\exp(-\theta y_j\tilde{z}_j)) \leq \frac{\sigma_{sk}^2(\tilde{z}_j)(y_j\theta)^2}{2} \qquad \forall\theta > 0.$$

Likewise, when $y_j < 0$, we have

$$\ln(\mathbb{E}(\exp(-\theta y_j\tilde{z}_j)) = \ln(\mathbb{E}(\exp(-\underbrace{(-\theta y_j)}_{>0}(-\tilde{z}_j))) \leq \frac{\sigma_{sk}^2(-\tilde{z}_j)(-y_j\theta)^2}{2} = \frac{\sigma_{sk}^2(-\tilde{z}_j)(y_j\theta)^2}{2} \qquad \forall\theta > 0.$$

Therefore by inspection, for all $\theta > 0$, we have

$$\begin{aligned} \ln(\mathbb{E}(\exp(-\theta y_j \tilde{z}_j)) &\leq \begin{cases} \dfrac{\sigma_{sk}^2(\tilde{z}_j) y_j^2 \theta^2}{2} & \text{if } y_j \geq 0 \\ \dfrac{\sigma_{sk}^2(-\tilde{z}_j) y_j^2 \theta^2}{2} & \text{otherwise} \end{cases} \\ &= \frac{y_j \phi_j(\boldsymbol{y}) \theta^2}{2} \end{aligned}$$

Therefore

$$\mathbb{E}(\exp(-\theta \boldsymbol{y}' \tilde{\boldsymbol{z}})) = \exp\left(\sum_{j=1}^{M} \ln \mathbb{E}(\exp(-\theta y_j \tilde{z}_j))\right) \leq \exp\left(\boldsymbol{y}' \boldsymbol{\phi}(\boldsymbol{y}) \frac{\theta^2}{2}\right) = \exp\left(\frac{s^2 \theta^2}{2}\right) \qquad \forall \theta > 0$$

Finally, we have

$$\mathbb{P}\left(\boldsymbol{y}' \tilde{\boldsymbol{z}} < -as\right) \leq \exp\left(\frac{s^2 \theta^2}{2} - \theta a s\right) \qquad \forall \theta > 0,$$

and the bound follows by choosing an optimal $\theta^* = a/s$. ■