Aggregate Investor Preferences and Beliefs in Stock Market: A Stochastic Dominance Analysis

Yi Fang

Center for Quantitative Economics, Jilin University, Changchun, China; and Business school, Jilin University, Changchun, China

Correspondence to: Business school, Jilin University, 2699 Qianjin Street, Changchun City, China.

Postcode: 130012.

Email: Danielfang@163.com.

Telephone: 0086-13354318008.

Fax: 0086-0431-512708994.

Aggregate Investor Preferences and Beliefs in Stock Market: A Stochastic Dominance Analysis

Abstract

This paper analyzes whether market portfolio is efficient related to benchmark portfolios formed on size, value, momentum and reversal with various utility theories by using stochastic dominance criteria. Our results support prospect theory including assumption of loss aversion at monthly and yearly horizons, which indicate the market utility is S-shaped, and steeper for losses than for gains. However the findings don't provide convincing evidence for positive skewness preference. Therefore, it should probe into asset pricing model and financial puzzles by prospect theory preferences. And, in the market, it may be difficult to benefit from the asset through its features on skewness, or other higher order central moment. In addition, for testing stochastic dominance efficiency, we also develop several bootstrap procedures that have favorable property in statistical size and power.

Keywords: Stock market efficiency; Bootstrap; Stochastic dominance; Prospect theory; Loss aversion; Skewness preference

JEL classification: C4; D4; D8; G1

Aggregate Investor Preferences and Beliefs in Stock Market: A Stochastic Dominance Analysis

1. Introduction

The aggregate of investor preferences and beliefs in stock market is the starting point of economics study and finance research, and is a much-debated topic in financial economics. Several asset pricing anomalies suggest that the market portfolio is significantly mean-variance (MV) inefficient relative to the stock portfolios formed on variables such as market capitalization (size), book-to-mark equity ratio (value), price momentum, and price reversal.[1] So it should extend or change traditional quadratic form utility to understand the market. Moreover, various risk preferences should be investigated with the pricing model by introducing alternative classes of utilities.

This paper uses the implied risk preferences to test three popular and competing utility theories. The first is the traditional expected utility theory with the assumption of global risk averse, that is, the utility function is everywhere concave. The second is the prospect theory (PT) of Kahneman and Tversky (1979), which assumes an S-shaped utility function that is risk seeking for losses and risk averse for gains,. The third theory, named Markowitz utility theory stemming from Markowitz (1952) and Thaler and Johnson (1990), indicates that contrary to prospect theory with a reverse S-shaped utility, that is, investors may risk averse for losses and risk seeking for gains.

The paper adopts stochastic dominance (SD) method, introduced by Hadar and Russell (1969), Hanoch and Levy (1969), to identify aggregate risk preferences. Considering anomalies, it analyzes whether the market portfolio is SD efficient relative to benchmark portfolios formed on size, value,

[1] See for example, Fama and French, 1992, 1993, 1996; Jegadeesh and Titman, 1993; Conrad and Kaul, 1998. However, Levy and Roll (2010) thought the market portfolio is significantly MV inefficient with ex-post parameters, but it may be not true with ex-ante parameters. However, Levy and Roll (2010) also considered their research doesn't constitute a proof of the empirical validity of the CAPM, but it shows that the model can not be rejected. Moreover, their research also didn't examine other type preferences. Hence, it also can't reject any other type preference.

momentum and reversal with various preferences.

And, we find the bootstrap method of Post and Levy (2005) may easily commit Type I error (rejecting the null when it is true). Therefore, we develop two bootstrap testing procedures for SD efficiency. One procedure adjusts the bootstrap statistic of Post and Levy (2005) corresponding to various significance levels. Another procedure shifts the entire distance between the original estimator of statistic and zero, which is an extended implementation of the method of Linton et al. (2005) for a critical estimation with full-sample bootstrap. However, there is a boundary effect, which may result in inconsistent in the bootstrap statistics. Hence, following the Simar and Wilson (1998), we also introduce the smoothed bootstrap statistics. The simulation shows the statistics of the new bootstrap procedures have favorable statistical properties for both size and power with large sample size. Even with small sample size, the statistics also have satisfactory statistical size.

Moreover, we further impose restriction of three order derivative on utility function to examine skewness preference. Many empirical evidences imply that the perception of risk is more complex than variance. Especially, the phenomena of positive skewness[2] and kurtosis preference[3] have attracted much attention among scholars. Accounting for the kind of preference, we adopt the SD criterion of Wong and Chan (2008) and extend the empirical examination for the assumption of positive skewness preference.

Furthermore, we test SD conditions that catch an important aspect of PT, namely, loss aversion as suggested by Benarzi and Thaler (1995, 1999). Baucells and Heukamp (2006) put forward loss aversion play a central role in behavioral decision research in PT. It captures the psychological intuition that losses loom larger than gains, and is very important explanation for many economic and

[2] See for example, Kraus and Litzenberger, 1976; Friend and Westerfield, 1980; Harvey and Siddique, 2000.
[3] See for example, Dittmar(2002).

financial puzzles.[4] Here, we incorporate the feature with preference condition, which introduced by Wakker and Tversky (1993), into SD criterion of S-shaped utility to analyze investor behavior.

In addtion, we investigate the market efficient not only on monthly data but also on yearly data. Hansson and Persson (2000) put forward the recommendation that investors with long investment horizons tilt their portfolios toward stocks is commonplace and an investor can gain from time diversification. Recently, Levy and Duchin (2004) also considered the investors are diverse at their planned investment horizons and the optimal investment decision of an investor may change at different horizons. The study on yearly data will discover the affection of longer horizon on asset equilibrium price and aggregate preferences. However, our evidences of yearly data are the same as those of monthly data, which are consistent with the prospect theory of Kahneman and Tversky (1979) and indicate the aggregate preferences are satisfied with an S-shaped utility function and assumption of loss aversion. The findings provide profound understanding of the capital market and asset pricing for investment horizon.

The remainder of this paper is organized as follows. Section 2 reviews methods of empirical study on aggregate preferences and the method of SD. Section 3 introduces the SD efficiency criteria. Section 4 investigates the test statistics on bootstrap methodology. The section also puts forward new bootstrap test procedures. Section 5 presents empirical findings of the aggregate investor preferences of US stock market. Finally, section 6 gives the conclusions.

2. The aggregate preferences and stochastic dominance method

2.1. The research of aggregate preferences

Many researchers investigate the individual risk preference mainly by psychological experiments.

[4] For example, the endowment effect (Thaler, 1980), and equity premium puzzle (Benartzi and Thaler, 1995, 1999).

However, the experimental and empirical results show that people's preferences are diverse. Moreover, there is much controversy about the experimental results.[5]

Furthermore, stock market is a complex system, which is a network of heterogeneous components that interact nonlinearly, to give rise to emergent behavior. Mauboussin (2002) considered that it is not possible to understand the stock market by paying attention to individual analysis. So it can’t inference the aggregate preferences straightforward by individuals risk preferences.

Other studies presuppose risk preference of investor and examine the specific preference. However, Post (2003) considered economic theory gives minimum guidance for the specification of utility function.[6] Thus, even the examination can not reject their presuppose preference, you can not confirm the assumption is the only correct option because you haven’t test any other risk preference. So the results are not sufficient.

To better understand the aggregate preferences of the market, it can not depend too much on the researches for individual, especially on the experiments for individual. Moreover, it should not just test whether the aggregate preferences is consistent with a presuppose risk preference alone. It should examine various risk preferences under a general framework and find out the aggregate preferences of the market.

Recently, to circumvent these problems, scholars use SD method to study the aggregate preferences[7] by various criteria for different risk preferences. Firstly, it can directly analyze the aggregate preferences rather than individual preference by using the benchmark portfolios and market portfolio.

[5] For example, Levy and Levy (2002) considered they find support for the Markowitz utility function in an experiment, which is the opposite of the PT. However, Wakker (2003) pointed that all the data of Levy and Levy are perfectly consistent with the predictions of the PT, if they don’t neglect the probability weighting function. Baltussen, Post and Van Vliet (2006) further put forward that they find severe violations of the PT and the Markowitz utility in a classroom choice experiment with mixed gambles and moderate probabilities.

[6] In various assumptions, only non-satiation can be accepted widely.

[7] For example, Post and Levy (2005) analyzed the aggregate preferences of American stock market; Fong et al (2008) investigated the aggregate preferences of internet stocks.

Secondly, the data is generated from the market, which is the real reaction of investors. Thirdly, the SD criteria can rely only on general preference and belief assumptions because they do not require a parametric specification.

2.2. Tests for stochastic dominance

The existing tests for SD include asymptotic method and bootstrap approach. However, the asymptotic testing procedures lack statistical power. Bootstrap can account for sampling error and deal with true distribution, and it can yield more powerful results. Therefore, the paper focuses on the methodology of bootstrap.

Simar and Wilson (1998) elaborated the sensitivity on efficiency scores by a general bootstrap methodology. They proposed a bias-adjusted idea which measures the efficiency related to a non-parametric condition based on observed data resulting from an underlying data generating process (DGP), which can be well used in situations where the sampling distribution is difficult or impossible to obtain analytically. The method can be viewed as a Mean Bias (MB) bootstrap, and an important procedure on the analysis of non-parametric efficiency scores.

Post and Levy (2005) introduced a bootstrap statistic for test of SD efficiency based on MB. They also adopted various SD criteria with restrictions on first order and second order derivative of utility function, and tested the market portfolio efficiency to investigate the aggregate investor preferences in the face of infinitely many choice alternatives using the bootstrap statistic and an asymptotic statistic developed by Post (2003).

But, we find that the statistic of MB may underestimate the bias-corrected magnitude when the statistic is not obeying the symmetrical distribution. In this case, the test will have high frequency on rejecting the null when it is true.

In the paper, we introduce new bootstrap statistics and take them compare with the bootstrap statistic of Post and Levy (2005), and re-examine the aggregate preferences of investors like Post and Levy (2005) for risk averse and risk seeking. Furthermore, we extend the examination for positive skewness preference and loss aversion.

3. Stochastic Dominance Efficiency Criteria

3.1. General Model

As Post and Levy (2005), it considers a single-period, portfolio-based model of a competitive capital market. Supposing that the investment universe consists of N assets, one of which is a riskless asset, we use the index set $\mathbf{I} \equiv \{1,...,N\}$ to denote the different assets with N for the riskless asset. The excess returns $\boldsymbol{x} \in \Re^N$ are treated as $N-1$ random variables and 0 with a continuous joint cumulative distribution function (CDF) $G(\cdot): \Re^N \to [0,1]$.[8] Investor may diversify among the assets, and we use $\boldsymbol{\lambda} \in \Re^N$ for a vector of portfolio weights belong to the portfolio possibilities set $\boldsymbol{\Lambda} \equiv \{\boldsymbol{\lambda} \in \Re_+^N : \boldsymbol{e}^{\mathrm{T}}\boldsymbol{\lambda} = 1\}$.[9] The assumption excludes short selling and risk-free borrowing because it doesn't affect the latter test of the market portfolio.[10]

The utility $u: \Re \to \Re$ is defined on excess return, which is differentiable and increasing. Investors select portfolio $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau}$ to maximize the expected utility. It represents all admissible utility functions by $U \equiv \{u : u'(x) \geq 0, \forall x \in \Re\}$, with $u'(x)$ for the marginal utility.[11] Note that Post (2003) and Post and Levy (2005) assume $U \equiv \{u : u'(x) \geq 1, \forall x \in \Re\}$. However the definition doesn't allow utility to be weakly increasing. To circumvent the problem, we let zero as lower bound of $u'(x)$.

---

[8] Throughout the paper, it uses $\Re^N$ for an *N*-dimensional Euclidean space, and $\Re_+^N$ denotes the positive orthant. To distinguish between vectors and scalars, we use a bold font for vectors and a regular font for scalars.

[9] Moreover, in this paper, all vectors are column vectors and it uses $\boldsymbol{x}^{\mathrm{T}}$ for the transpose of $\boldsymbol{x}$. And $\boldsymbol{e}$ denotes a unity vector with dimensions conforming to the rules of matrix algebra.

[10] The detailed discussion sees Post and Van Vliet (2006).

[11] For the assumption of non-satiation, the marginal utility is non-negative.

Actually, the CDF is generally not known, and information is typically limited to a discrete set of time series observations, say $\boldsymbol{X} \equiv (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_K)$. Here, $\boldsymbol{x}_k \equiv (x_{1k,\ldots,}x_{Nk})^{\mathrm{T}}$.[12]

Supposing that the observations generated from the CDF are independent random variables and since the timing of the draws is inconsequential, the observations are free to be labeled by their ranking with respect to the evaluated portfolio, i.e., $\boldsymbol{X} \equiv (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_K)$ with $\boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau} < \boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau} < \cdots < \boldsymbol{x}_K^{\mathrm{T}}\boldsymbol{\tau}$, and indexed by $\Theta \equiv \{k\}_{k=1}^{K}$. Using the observations, we can construct the empirical distribution function (EDF) as:

$$F(\boldsymbol{x}) \equiv \frac{1}{K}\sum_{k=1}^{K} 1(\boldsymbol{x}_k \le \boldsymbol{x}) \quad (1)$$

By using $F(\boldsymbol{x})$ in place of CDF $G(\boldsymbol{x})$, it may characterize different empirical distribution of SD efficiency criteria by different classes of utility functions, associated with different sets of restrictions on marginal functions by:

$$U(\Psi) = \{u \in U : u'(x) \ge u'(y), \forall (x, y) \in \Psi\} \quad (2)$$

Here, $\Psi \subseteq \Re^2$ represents the restrictions that are placed on marginal utility.

Definition 1. Portfolio $\boldsymbol{\tau} \in \boldsymbol{\Lambda}$ is empirically $U(\Psi) - SD$ efficient if and only if:

$$\min_{u \in U(\Psi)} \left\{ \max_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} \left\{ \int u(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\lambda}) dF(\boldsymbol{x}) - \int u(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau}) dF(\boldsymbol{x}) \right\} \right\} = 0 \Leftrightarrow \quad (3)$$

$$\min_{u \in U(\Psi)} \left\{ \max_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} \left\{ \sum_{k \in \Theta} \left(u(\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{\lambda}) - u(\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{\tau}) \right) / K \right\} \right\} = 0 \quad (4)$$

Alternatively, portfolio $\boldsymbol{\tau} \in \boldsymbol{\Lambda}$ is empirically $U(\Psi) - SD$ inefficient if and only if:

$$\min_{u \in U(\Psi)} \left\{ \max_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} \left\{ \sum_{k \in \Theta} \left(u(\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{\lambda}) - u(\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{\tau}) \right) / K \right\} \right\} > 0 \quad (5)$$

According to the efficiency criteria, Post (2003) and Post and Levy (2005) put forward a test statistic for SD criteria. The test essentially checks if the necessary first-order condition for portfolio optimality holds for some utility function $u \in U(\Psi)$. Specifically, if the portfolio $\boldsymbol{\tau} \in \boldsymbol{\Lambda}$ is optimal

[12] Note Post and Levy (2005) use *t* and *T* to stand for time. They also use T to represent the transpose operator. In order to make a distinction, we use *k* and *K* for time.

for some utility function $u \in U(\Psi)$, i.e., $\boldsymbol{\tau} = \underset{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}}{\arg\max} \int u(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\lambda})dG(\boldsymbol{x})$ then all assets must lie on or below the tangent hyperplane, that is:

$$\int u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau} - x_i)dG(\boldsymbol{x}) \ge 0, \quad \forall i \in \mathbf{I} \tag{6}$$

The first-order condition for optimality relative to the EDF, i.e., $\boldsymbol{\tau} = \underset{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}}{\arg\max} \sum_{k\in\Theta} u(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\lambda}) / K$, is the following sample equivalent of (6):

$$\sum_{k\in\Theta} u'(\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{\tau})(\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{\tau} - x_{ik})/K \ge 0, \quad \forall i \in \mathbf{I} \tag{7}$$

It introduces *K*-stage linear utility function with the intercept $\boldsymbol{\alpha} \equiv (\alpha_1, \cdots, \alpha_K) \in \Re^K$ and the slope $\boldsymbol{\beta} \equiv (\beta_1, \cdots, \beta_K) \in \Re^K$. So, $\boldsymbol{\beta}$ represents the gradient vector $\left(u'(\boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau}), ..., u'(\boldsymbol{x}_K^{\mathrm{T}}\boldsymbol{\tau})\right)^{\mathrm{T}}$ for some $u \in U(\Psi)$. Introducing a variable θ, the test statistic is then yielded:

$$\xi(\boldsymbol{\tau}, F(\boldsymbol{x}), U(\Psi)) = \min_{\beta\in B(\Psi),\theta} \left\{\theta : \sum_{k\in\Theta} \beta_k (\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{\tau} - x_{ik}) / K + \theta \ge 0, \forall i \in \mathbf{I}\right\} \tag{8}$$

with

$$B(\Psi) \equiv \left\{\boldsymbol{\beta} \in \left[0, \infty\right\rangle^K : \beta_k \ge \beta_s, \forall k, s \in \Theta : (\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{\tau}, \boldsymbol{x}_s^{\mathrm{T}}\boldsymbol{\tau}) \in \Psi\right\} \tag{9}$$

The admissible set $B(\Psi)$ represents the restrictions on the gradient vector that follow characteristics of the specific utility function. The test statistic $\xi(\boldsymbol{\tau}, F(\boldsymbol{x}), U(\Psi))$ basically measures the smallest possible maximum pricing error relative to the well-behaved utility function. Consequently, the statistic can be viewed as an efficiency measure.

Theorem 1. Portfolio $\boldsymbol{\tau} \in \boldsymbol{\Lambda}$ is empirically efficient if and only if $\xi(\boldsymbol{\tau}, F(\boldsymbol{x}), U(\Psi)) = 0$. Alternatively, portfolio $\boldsymbol{\tau} \in \boldsymbol{\Lambda}$ is empirically inefficient if and only if $\xi(\boldsymbol{\tau}, F(\boldsymbol{x}), U(\Psi)) > 0$.

Here, the objective is to test the null hypothesis that portfolio $\boldsymbol{\tau} \in \boldsymbol{\Lambda}$ is empirically efficient by the examination of $\xi(\boldsymbol{\tau}, F(\boldsymbol{x}), U(\Psi))$, i.e. $H_0 : \xi(\boldsymbol{\tau}, F(\boldsymbol{x}), U(\Psi)) = 0$. Moreover, the test statistic involves a linear objective function and can impose a finite set of linear constraints for various preferences. Consequently, the test statistic can be easily solved by using straightforward linear programming.

3.2. Linear constraints for risk averse and risk seeking

Considering both on risk averse and risk seeking behavior, there are three SD criteria are most concerned by scholars, which includes Second-order Stochastic Dominance (SSD), Prospect Stochastic Dominance (PSD) and Markowitz Stochastic Dominance (MSD). We adopt the different criteria accounting for first and second derivative of utility function, which based on three different sets of restrictions on marginal utility, i.e., Ψ_{SSD} for SSD, Ψ_{PSD} for PSD, and Ψ_{MSD} for MSD. Using $z \in \Theta$ for the first observation in the domain of gains, $\boldsymbol{x}_{z-1}^{\mathrm{T}}\boldsymbol{\tau} < 0 \leq \boldsymbol{x}_{z}^{\mathrm{T}}\boldsymbol{\tau}$. Then, it can be obtained that[13]:

$$B(\Psi_{SSD}) \equiv \left\{ \boldsymbol{\beta} \in \left[0, \infty\right\rangle^{K} : \beta_1 \geq \beta_2 \geq \cdots \geq \beta_K \right\} \tag{10}$$

$$B(\Psi_{PSD}) \equiv \left\{ \boldsymbol{\beta} \in \left[0, \infty\right\rangle^{K} : \beta_1 \leq \beta_2 \leq \cdots \leq \beta_{z-1}; \beta_z \geq \beta_{z+1} \geq \cdots \geq \beta_K \right\} \tag{11}$$

$$B(\Psi_{MSD}) \equiv \left\{ \boldsymbol{\beta} \in \left[0, \infty\right\rangle^{K} : \beta_1 \geq \beta_2 \geq \cdots \geq \beta_{z-1}; \beta_z \leq \beta_{z+1} \leq \cdots \leq \beta_K \right\} \tag{12}$$

SSD indicates decision-making and equilibrium under uncertainty traditionally use expected utility functions with the assumption of non-satiation and global risk averse, that is, the utility is monotone increasing and concave everywhere. However, there is compelling evidence that many decision makers are risk seeking.

In view of the PT, Levy and Wiener (1998) presented PSD corresponding to the S-shaped utility that is convex for losses and concave for gains (risk seeking for losses and risk averse for gain).[14] The criterion makes the important points that people might be both risk averse as well as risk seeking. Moreover, the criterion is supported by psychological experiments and can explaining many financial

---

[13] Note that the definition of β is different from Post (2003) and Post and Levy(2005) because we let zero as lower bound of $u'(x)$.

[14] Gains and losses are typically measured relative to a subjective reference point. As Post and Levy (2005) and Fong et al (2008), we set reference point at zero. The use of excess returns implies that the nominal reference point effectively equals the riskless rate.

market anomalies.[15]

However many researchers consider that the behavior of investors may be contrary to the PT. In particular, investors are more risk seeking following gains and more risk averse following loses, and a reverse S-shaped utility function may be more descriptive of actual behavior.[16] Levy and Levy (2002) developed MSD for the reverse S-shaped utility function, which is also an important criterion to understand the market preferences.[17]

3.3. Linear constraints for positive skewness preference

For the positive skewness preference, it pays attention to third derivative of utility function. Considering the preference, Whitmore (1970) put forward three-order SD (TSD) criterion for risk averse. Recently, Wong and Chan (2008) extended the criteron for both risk averse and risk seeking. According to them, we also impose three different sets of restrictions on marginal utility following equations (10), (11) and (12), i.e., Ψ_{T-SSD} for SSD, Ψ_{T-PSD} for PSD, and Ψ_{T-MSD} for MSD, with skewness preference. Essentially, the restriction is $u'''(.) \geq 0$ whether for risk averse or risk seeking. And, it can be obtained that:

$$B(\Psi_{T-SSD}) \equiv \left\{ \boldsymbol{\beta} \in \left[0,\infty\right\rangle^K : \beta_1 \geq \beta_2 \geq \cdots \geq \beta_K ; \frac{\beta_1 - \beta_2}{\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau}} \geq \frac{\beta_2 - \beta_3}{\boldsymbol{x}_3^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau}} \geq \cdots \geq \frac{\beta_{K-1} - \beta_K}{\boldsymbol{x}_K^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_{K-1}^{\mathrm{T}}\boldsymbol{\tau}} \right\} \quad (13)$$

$$\begin{aligned} B(\Psi_{T-PSD}) \equiv \Big\{ & \boldsymbol{\beta} \in \left[0,\infty\right\rangle^K : \beta_1 \leq \beta_2 \leq \cdots \leq \beta_{z-1} ; \beta_z \geq \beta_{z+1} \geq \cdots \geq \beta_K ; \\ & \frac{\beta_1 - \beta_2}{\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau}} \geq \frac{\beta_2 - \beta_3}{\boldsymbol{x}_3^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau}} \geq \cdots \geq \frac{\beta_{z-1} - \beta_z}{\boldsymbol{x}_z^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_{z-1}^{\mathrm{T}}\boldsymbol{\tau}} ; \frac{\beta_z - \beta_{z+1}}{\boldsymbol{x}_{z+1}^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_z^{\mathrm{T}}\boldsymbol{\tau}} \geq \frac{\beta_{z+1} - \beta_{z+2}}{\boldsymbol{x}_{z+2}^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_{z+1}^{\mathrm{T}}\boldsymbol{\tau}} \geq \cdots \geq \frac{\beta_{K-1} - \beta_K}{\boldsymbol{x}_K^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_{K-1}^{\mathrm{T}}\boldsymbol{\tau}} \Big\} \end{aligned} \quad (14)$$

$$\begin{aligned} B(\Psi_{T-MSD}) \equiv \Big\{ & \boldsymbol{\beta} \in \left[0,\infty\right\rangle^K : \beta_1 \geq \beta_2 \geq \cdots \geq \beta_{z-1} ; \beta_z \leq \beta_{z+1} \leq \cdots \leq \beta_K ; \\ & \frac{\beta_1 - \beta_2}{\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau}} \geq \frac{\beta_2 - \beta_3}{\boldsymbol{x}_3^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau}} \geq \cdots \geq \frac{\beta_{z-1} - \beta_z}{\boldsymbol{x}_z^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_{z-1}^{\mathrm{T}}\boldsymbol{\tau}} ; \frac{\beta_z - \beta_{z+1}}{\boldsymbol{x}_{z+1}^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_z^{\mathrm{T}}\boldsymbol{\tau}} \geq \frac{\beta_{z+1} - \beta_{z+2}}{\boldsymbol{x}_{z+2}^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_{z+1}^{\mathrm{T}}\boldsymbol{\tau}} \geq \cdots \geq \frac{\beta_{K-1} - \beta_K}{\boldsymbol{x}_K^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_{K-1}^{\mathrm{T}}\boldsymbol{\tau}} \Big\} \end{aligned} \quad (15)$$

3.4. Linear constraints for loss aversion

---

[15] See for example, Kahneman and Tversky (1979), Tversky and Kahneman (1992), Benartzi and Thaler (1995), Barberis and Huang (2001) and Barberis et al (2001).

[16] See for example, Markowitz (1952) and Thaler and Johnson (1990).

[17] Levy and Wiener (1998) confined MSD rule only to non-extreme outcomes, i.e., the reversed S-shaped range of the Markowitz utility.

Loss aversion means investors are distinctly more sensitive to losses than to gains, which is an important feature of PT. It is expressed mathematically as a steeper utility function for losses than for gains. Wakker and Tversky (1993) introduced a corresponding condition possessing the preference, which defines the class of utility function as follows[18]:

$$U(\Psi) = \{u \in U : u'(-x) \geq u'(x), \forall x > 0\} \tag{16}$$

For the EDF, the condition entails comparisons between the negative and the positive domain. According to them, we also impose restrictions on marginal utility following equation (11), i.e., Ψ_{LA-PSD} for PSD with loss aversion. It also uses $z \in \Theta$ for the first observation in the domain of gains, $\boldsymbol{x}_{z-1}^{\mathrm{T}}\boldsymbol{\tau} < 0 \leq \boldsymbol{x}_z^{\mathrm{T}}\boldsymbol{\tau}$. Then, it can be obtained that:

$$\begin{aligned} B(\Psi_{LV-PSD}) \equiv \{\boldsymbol{\beta} \in \left[0,\infty\right\rangle^K : \beta_1 \leq \beta_2 \leq \cdots \leq \beta_{z-1}; \beta_z \geq \beta_{z+1} \geq \cdots \geq \beta_K; \\ \beta_i \geq \beta_j, \forall i < z \mid \boldsymbol{x}_j^{\mathrm{T}}\boldsymbol{\tau} < \boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\tau} \leq \boldsymbol{x}_{j+1}^{\mathrm{T}}\boldsymbol{\tau}, j \geq z; \\ \beta_i \geq \beta_K, \forall i < z \mid \boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\tau} \geq \boldsymbol{x}_K^{\mathrm{T}}\boldsymbol{\tau}\} \end{aligned} \tag{17}$$

4. Statistics and Inference

The EDF is very sensitive to sampling variation and the test results are likely to be affected by sampling error in a non-trivial way. So, statistical method must be employed to make inferences about the true efficiency classification. In this section, we discuss the test procedures developed by Post and Levy (2005) and propose new bootstrap testing procedures.

4.1. The mean bias statistic of bootstrap

Since the estimator of ξ is obtained from finite samples, the corresponding measures of efficiency are sensitive to sampling variations of the obtained frontier. The values of ξ are often overestimated. Simar and Wilson (1998) proposed a bootstrap strategy through reasonable assumptions regarding the DGP to analyze the sensitivity of estimator of the efficiency measure.

---

[18] Baucells and Heukamp (2006) considered the condition is stronger than the condition originally put forward in Kahneman and Tversky (1979).

The strategy is based on the idea that the known bootstrap distributions will reproduce the original unknown sampling distributions of the estimators of interest. Supposing that a DGP, p, can be obtained by a reasonable estimator $\hat{p}$ from the observed data. For the SD efficiency statistic ξ, let $\hat{\xi}$ be the estimator with the observed sample, and $\hat{\xi}_b$ be the estimator with bootstrap pseudo-sample. Then, we have:

$$(\hat{\xi}_b - \hat{\xi}) \mid \hat{p} \sim (\hat{\xi} - \xi) \mid p \tag{18}$$

Simar and Wilson (1998) thought the bias of $\hat{\xi}$ for ξ, i.e., $bias_p = E_p(\hat{\xi}) - \xi$ can be obtained from its bootstrap estimation, i.e., $bias_{\hat{p}} = E_{\hat{p}}(\hat{\xi}_b) - \hat{\xi} = \frac{1}{B}\sum_{b=1}^{B}\hat{\xi}_b - \hat{\xi}$, and B is the number of bootstrap pseudo-samples. Here, both $bias_p$ and $bias_{\hat{p}}$ are mean bias. Post and Levy (2005) considered the bias-corrected statistic of $\hat{\xi}_b$ should be shifted by $2bias_{\hat{p}}$, i.e.:

$$\hat{\xi}_{b,MB} = \hat{\xi}_b - 2bias_{\hat{p}} = \hat{\xi}_b + 2\hat{\xi} - 2\frac{1}{B}\sum_{b=1}^{B}\hat{\xi}_b \tag{19}$$

Here, $\hat{\xi}_{b,MB}$ is called the Mean Bias (MB) statistic which shifts the distance of bootstrap statistic by the mean bias. Post and Levy (2005) also bootstrapped pseudo-samples with B times. Then, it arrives at the probability:

$$P_{MB} = \#\{\hat{\xi}_{b,MB} < 0\} \big/ B \tag{20}$$

At the significance level of a, it rejects the efficiency when $P_{MB} < a$ and accept the efficiency when $P_{MB} \geq a$.

For the efficient portfolio $\xi = 0$ and supposing $(\hat{\xi}_b - \hat{\xi}) \mid \hat{p} \sim \hat{\xi} \mid p \sim F$, we get:

$$\hat{\xi}_b - 2bias_{\hat{p}} = \hat{\xi}_b - 2E_p(\hat{\xi}) \tag{21}$$

In this way, the bias-corrected estimators $\hat{\xi}_{b,MB}$ are centered at $\hat{\xi} - bias_p = \hat{\xi} - E_p(\hat{\xi})$. They adjust the distance of the bias between the original estimator and the mean of the bootstrap estimators. When $\hat{\xi}$ is relative large at extreme right-hand quantile of distribution $\hat{\xi} \mid p$, i.e., the P_{MB} is smaller than a given significance level, it can reject the efficiency assumption. Otherwise it should accept the assumption. For the efficient portfolio, $\xi = 0$, we assume a sample estimator of ξ is $\hat{\xi}^{1-a}$, which is at quantile $1-a$ of distribution p. Considering an estimator of a bootstrap pseudo-sample of the sample is $\hat{\xi}_b^{1-a,a_0}$, with $(\hat{\xi}_b^{1-a,a_0} - \hat{\xi}^{1-a})$ at quantile a_0 of $\hat{p}$. Then $\hat{\xi}_b^{1-a,a_0} - \hat{\xi}^{1-a} = \hat{\xi}^{a_0}$ because $(\hat{\xi}_b^{1-a} - \hat{\xi}^{1-a}) \mid \hat{p} \sim \hat{\xi} \mid p$. As $\xi = 0$, $\hat{\xi}_b^{1-a,a_0}$ can be asymptotically viewed as the sum of $\hat{\xi}^{1-a}$ and $\hat{\xi}^{a_0}$. Then, we can obtain:

$$\hat{\xi}_b^{1-a,a_0} - 2E_p(\hat{\xi}) = (\hat{\xi}^{1-a} - E_p(\hat{\xi})) + (\hat{\xi}^{a_0} - E_p(\hat{\xi})) \tag{22}$$

At the significance level of a, we should accept the efficiency assumption when $\hat{\xi}$ on the left-hand side of $(1-a)$ quantile of distribution $\hat{\xi} \mid p$. Then, with $a_0 \le a$, it should be $\hat{\xi}_b^{1-a,a_0} - 2E_p(\hat{\xi}) \le 0$. For $a_0 = a$, we can have $\hat{\xi}_b^{1-a,a_0} - 2E_p(\hat{\xi}) = (\hat{\xi}^{1-a} - E_p(\hat{\xi})) + (\hat{\xi}^{a} - E_p(\hat{\xi}))$. So, if $\hat{\xi} \mid p$ is a symmetrical distribution, we can obtain that $(\hat{\xi}^{1-a} + \hat{\xi}^{a})/2 = E_p(\hat{\xi})$, then $\hat{\xi}_b^{1-a,a_0} - 2E_p(\hat{\xi}) = 0$; thus, it will reject the efficiency when $\hat{\xi}$ is larger than the value at $(1-a)$ quantile of p and the probability to make Type I error equals to *a*. However when the distribution of $\hat{\xi} \mid p$ is asymmetric, such as the distribution with a long right tail or positive skewness, it may underestimate the bias-corrected magnitude to $\hat{\xi}_b$ at the right tail. Then, it may result in high frequency on $\hat{\xi}_b - 2E_p(\hat{\xi}) > 0$, and the probability to make Type I error is greater than a. Therefore, the statistical size of MB is sensitive to the shape of distribution.

4.2. The level adjust mean bias statistic of bootstrap

For general distribution, in order to make the probability of Type I error equal to a at the significance level of a, we can define the bias of the MB statistic by estimating:

$$bias_{MB}^{a} = (\hat{\xi}^{1-a} - E_p(\hat{\xi})) + (\hat{\xi}^{a} - E_p(\hat{\xi})) = \hat{\xi}_b^{a} + \hat{\xi}_b^{1-a} - 2E_{\hat{p}}(\hat{\xi}) \tag{23}$$

For a, it can calculate the bias-corrected statistic as:

$$\hat{\hat{\xi}}_{b,LAMB} = \hat{\hat{\xi}}_{b,MB} - bias_{MB}^{a} = \hat{\hat{\xi}}_{b} - \hat{\hat{\xi}}_{b}^{a} - \hat{\hat{\xi}}_{b}^{1-a} + 2\hat{\xi} \quad (24)$$

Here, $\hat{\hat{\xi}}_{b,LAMB}$ is called the Level Adjust Mean Bias (LAMB) statistic which adjusts the bias of MB statistic at specific significance level. Bootstrapping pseudo-samples with B times, it arrives at the probability:

$$P_{LAMB} = \#\{\hat{\hat{\xi}}_{b,LAMB} < 0\} / B \quad (25)$$

At the significance level of a, the efficiency is rejected when $P_{LAMB} < a$ and accepted when $P_{LAMB} \geq a$. If the distribution $\hat{\xi} \mid p$ is symmetrical, it can asymptotically yield $bias_{MB}^{a} = 0$. Then, the procedures of LAMB and MB yield consistent results. But, if the distribution is characterized by a long right tail or positive skewness, the procedure of LAMB will involve more favorable statistical size. Furthermore, whatever the shape of distribution $\hat{\xi} \mid p$ is, the P_{LAMB} is always asymptotically equal to the predefined level of the significance. So, the statistical size of LAMB is not sensitive to the shape of distribution.

4.3. The entire-distance bias statistic of bootstrap

Linton et al. (2005) proposed a full-sample bootstrap procedure for estimating the critical values of a suitably extended Kolmogorov-Simrnov test for SD amongst the K-competing states. The procedure computes the bootstrap distribution of the statistic conditional on the original sample and takes the critical value from the distribution. Here, we extend the method for our SD efficiency test.

According to equation (18), $(\hat{\hat{\xi}}_{b} - \hat{\xi})$ can be viewed as a reasonable estimator for $(\hat{\xi} - \xi)$. So, the interval of $(\hat{\xi} - \xi)$ can be obtained from the distribution of $(\hat{\hat{\xi}}_{b} - \hat{\xi})$. For the efficiency portfolio, $\xi = 0$, it can compare the quantile of $\hat{\xi}$ in the distribution of $(\hat{\hat{\xi}}_{b} - \hat{\xi})$ with the specific significance level to test the SD efficiency. Then, it defines

$$\hat{\hat{\xi}}_{b,EDB} = \hat{\hat{\xi}}_{b} - \hat{\xi} \quad (26)$$

Here, $\hat{\xi}_{b,EDB}$ can be called the Entire Distance Bias (EDB) statistic which shifts the distance of bootstrap statistic by the entire distance between the original estimator and zero. And, the probability is calculated with B times bootstrap pseudo-samples:

$$P_{EDB} = \#\{\hat{\xi}_{b,EDB} > \hat{\xi}\} / B \tag{27}$$

At the significance level of a, it rejects the efficiency when $P_{EDB} < a$ and accepts the efficiency when $P_{EDB} \geq a$. Considering the efficient portfolio, since $\hat{\xi}^{1-a}$ is at quantile $(1-a)$ of p, it is known that $\hat{\xi}_{b,EDB} > \hat{\xi}^{1-a}$ only if $(\hat{\xi}_{b,EDB} - \hat{\xi}^{1-a})$ on the right-hand side of $(1-a)$ quantile of p. Otherwise, $\hat{\xi}_{b,EDB} \leq \hat{\xi}^{1-a}$. Consequently, the relative probability, of which the efficient portfolio is wrongly classified as being inefficient, is a. The test of EDB directly compares $\hat{\xi}$ with the distribution of $\hat{\xi}_{b,EDB}$. Hence, the statistical size of EDB is also not sensitive to the shape of distribution. It can be found the tests of efficiency by EDB statistic and LAMB statistic will get consistent results.

4.4. The smoothed bootstrap

Simar and Wilson (1998) addressed the key to valid implementation of the bootstrap is to find a reasonable estimate $\hat{p}$ of the DGP, p. As they discussed, it should chose a consistent estimator of the distribution F of ξ to establish the validity. Unfortunately, the distribution $\hat{F}$ by standard bootstrapping, which bootstrapped pseudo-samples from the observed sample, will provide poor estimator of F. Especially, the estimators of $\hat{\xi}_b$ may produce a large number of ostensibly efficient units, and raises boundary problem in the extreme left tail. As suggested by Simar and Wilson (1998), one way to improve the estimation of F is smoothed the standard bootstrap distribution $\hat{F}$ by using the reflection method.[19] Formally, we take into account the boundary condition that $\xi > 0$, and

[19] See, Silverman (1986), for details.

estimate the kernel density from the set of the $2n$ values $\{-\hat{\xi}_1, -\hat{\xi}_2, ..., -\hat{\xi}_n, \hat{\xi}_1, \hat{\xi}_2, ..., \hat{\xi}_n\}$, which are symmetrically distributed around 0. Then it can have

$$\hat{G}_h(\xi) = \frac{1}{2nh}\sum_{i=1}^{n}\Phi\left(\frac{\xi - \hat{\xi}_i}{h}\right) + \Phi\left(\frac{\xi + \hat{\xi}_i}{h}\right) \tag{28}$$

Here, $\Phi(.)$ is Gaussian kernel density function; h is the bandwidth controlling the scale of the kernel function, and $h = 1.06\min(s_{2n}, r_{2n}/1.34)(2n)^{-1/5}$. Where s_{2n} is the empirical standard deviation of the $2n$ reflected data; r_{2n} is interquartile range of the $2n$ reflected data. The density estimate of F is then obtained through:

$$\hat{F}_{S,h}(\xi) = \begin{cases} \hat{G}_h(\xi), & if\ \xi \geq 0 \\ 0, & otherwise \end{cases} \tag{29}$$

Let $\hat{\xi}_b^*, b = 1, 2, \ldots, B$ are the estimator by standard bootstrapping, and ε_b^* is a random deviate drawn from the standard normal, $b = 1, 2, \ldots, B$. Now consider the following random generator:

$$\tilde{\xi}_b^* = \begin{cases} \hat{\xi}_b^* + h\varepsilon_b^*, & if\ \hat{\xi}_b^* + h\varepsilon_b^* \geq 0 \\ -(\hat{\xi}_b^* + h\varepsilon_b^*), & otherwise \end{cases} \tag{30}$$

Then, it can obtain the sequence of the smoothed bootstrap:[20]

$$\hat{\xi}_b^S = \mu + \frac{\tilde{\xi}_b^* - \mu}{\sqrt{1 + h^2/\sigma^2}} \tag{31}$$

Where μ and σ are the empirical mean and variance of $\hat{\xi}_b^*, b = 1, 2, \ldots, B$. Thus, the smoothed bootstrap procedure is: (1) estimate $\hat{\xi}_b^*, b = 1, 2, \ldots, B$, by standard bootstrapping; (2) calculate $\tilde{\xi}_b^*, b = 1, 2, \ldots, B$, with equation (30); (3) compute $\hat{\xi}_b = \hat{\xi}_b^S, b = 1, 2, \ldots, B$, with equation (31); (4) compute P_{MB}, P_{LAMB} and P_{EDB} with equations (20), (25) and (27).

It can be found that the smoothed procedure will remove many efficient units, i.e., $\hat{\xi}_b = 0$ estimated by standard bootstrapping, and reflect the units with $\tilde{\xi}_b^S \geq 0$. Therefore, comparing the distribution of the statistics by standard bootstrapping with by smoothed bootstrapping, it can be

[20] See, Efron and Tibshirani(1993), Silverman (1987), and Simar and Wilson (1998), for details..

drawn two conclusions. Firstly, the left tail of the former will fatter than that of the latter. Secondly, the quantiles at the left tail of the former should be smaller than that of the latter.

Since P_{EDB} is determined at the right tail of the distribution, the boundary problem should have small impact on the standard bootstrap. On the contrary, both P_{MB} and P_{LAMB} are determined at the left tail of the distribution. Therefore, the boundary problem should influence on the two standard bootstrap procedures. However, the influence may be not very clear. For example, for the two procedures, the fatter left tail will increase the frequencies to reject the null, while the smaller quantile of left tail will decrease the frequencies to decline the null. Since the influence is complex and the smoothed method has improved the statistics for the boundary problem, we only compare the standard procedure and the smoothed procedure by simulation.

From what has been stated above, we can't come to a simple conclusion whether the smoothed procedure can involve more favorable statistical size and power or not, though the smoothed bootstrap can improve the consistency of the distribution. Then, we will illustrate the comparison by simulation.

4.5. Simulation experiment

In order to test the effectiveness of various bootstrap statistics and compare the different approaches, we do a simple simulation by 5 risky assets with the 5 Fama and French stock portfolios formed on BE/ME and a single riskless asset with the one-month US Treasury bill because a detailed simulation will beget a too large computation burden for the statistics of bootstrap. The joint population moments are equal to the sample moments of the monthly excess returns of the 5 BE/ME stock portfolios during sample period from Jan 1933 to Dec 2007. We draw random samples from the multivariate normal population distribution through Monte-Carlo simulation. We analyze two different test portfolios on

SD criterion for MV efficiency.[21] The first test portfolio is the tangency portfolio (TP), which is MV efficient. So, it may analyze the statistical size by the relative frequency of the random samples in which this portfolio is wrongly classified as being inefficient. The second test portfolio is the equal weighted portfolio (EP), which is known to be MV inefficient. Hence, it may analyze the statistical power by the ability to correctly classify the portfolio as being inefficient.[22]

[Insert Figure 1 Here]

Figure 1 shows the mean-variance diagram of excess returns of the 5 portfolios (clear dot), TP, EP, as well as the mean-variance efficient frontier (OD). The TP consists of 85.48% and 14.52% invested in the third and the fifth BE/ME portfolios (portfolio B3 and B5 in the Figure 1) respectively. And, the third and the fifth BE/ME portfolios also construct efficient frontier OD with riskless asset.

This simulation experiment is performed for sample sizes of 50, 100, 400 and 800, and at significance levels of 10%, 5% and 1%. The procedures of the SD tests on the MV efficiency to TP and EP are as follows:

Step 1. According to the joint sample moment, generate a simulated sample named sample 1 with N observations. Calculate the value of $\hat{\xi}_{1,PT}$ for TP and the value of $\hat{\xi}_{1,PE}$ for EP. Then, generate 1000 pseudo-samples from the sample 1, and calculate the bootstrap estimators $\hat{\xi}_{1,j,PT}$, $\hat{\xi}_{1,j,PE}(j=1,2,...,1000)$ for TP and EP respectively.[23] Compute P_{MB}, P_{LAMB} and P_{EDB}, with equations (20), (25) and (27), by the standard bootstrap methods and the smoothed bootstrap

[21] It adopts SD test on MV efficiency as Post and Van Vliet (2006). The admissible set of the restrictions on the gradient vector of MV utility function is $B(\Psi_{MVSD}) \equiv \left\{ \boldsymbol{\beta} \in [0,\infty\rangle^K : \beta_k = c + d\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{\tau}, k=1,...,K \right\}$.

[22] The power depends on the degree of inefficiency of the evaluated portfolio. That we use the Sharp Ratio of EP to divide the Sharp Ratio of TP measures the degree of the inefficiency of different EPs constructed from different Fama and French stock portfolios, which include the 5 portfolios formed on BE/ME, the 5 portfolios formed on size, the 6 portfolios formed on size and value, the 6 portfolios formed on size and momentum, and the 6 portfolios formed on size and short-term reversal during sample period from Jan 1933 to Dec 2007. The result shows that the degree of the EP constructed from the 5 BE/ME portfolios is 'medium', which is higher than that of the 6 size and value portfolios and the 6 size and momentum portfolios, but lower than that of the 5 size portfolios and the 6 size and short-term reversal portfolios. Hence the paper selects the set of the 5 BE/ME portfolios to simulation.

[23] According to Hall (1986), Simar and Wilson (1998) put forward that it can sample 1000 times to ensure adequate coverage of the confidence intervals.

procedures of Section *4.4* respectively.

Step 2. Repeat Step 1 for 999 times.

Step 3. Compute the statistical size as the rejection rate for TP and the statistical power as the rejection rate for EP.

[Insert Figure 2 Here]

Figure 2 shows the kernel density estimation of the distributions to $\hat{\xi}$ with various sample sizes. The estimation adopts the Epanechnikov kernel function. And, the bandwidths are selected by maximizing the likelihood cross-validation function discussed by Silverman (1986). Clearly, the distribution of the portfolio TP is asymmetric even with the large sample size, which may result in high frequency of the statistic of MB on rejecting the efficient portfolio.

[Insert Table 1 Here]

Descriptive statistics for $\hat{\xi}$ are also included in Table 1.

For the portfolio TP, with increasing sample size, the mean and the median approach 0,[24] and the standard deviation and the range decrease gradually. The skewness is always positive and the kurtosis is larger than three. Moreover the test statistic of JB indicates that $\hat{\xi}$ is not consistent with the normal distribution. Especially, the statistic of JB also indicates the positive skewness is significant as the curve of TP in figure 2, which may results in lower statistical size with the MB statistics.

For the portfolio EP, the mean and median trend to decline with increasing sample size. The standard deviation and the range also decrease gradually. However, the skewness and the kurtosis decline much faster than those of the TP. The test statistic of JB also declines much more significantly.

[Insert Table 2 Here]

[24] Since TP is MV efficient, the value of ξ is 0. Consequently, it is likely that $\hat{\xi}$ is gradually close to 0 when the sample size increases.

Table 2 includes statistical properties of the statistics. It includes the statistical size and power of the different bootstrap procedures for various sample sizes and at different significance levels. The results of TP give the size of the statistics and the results of EP show the power of the statistics. Each cell includes both the results of the standard and the smoothed bootstrap, and the result of the former is in the bracket.

For both the standard and the smoothed bootstrap statistics of MB

As the sample size increases, both the statistical powers increase obviously. In fact, at the significance level of 10%, both powers of the two statistics beyond 0.8 with sample size of 800. However, both statistical sizes of the two statistics always exceed the nominal significance level and don't show a downward trend with increasing sample sizes. Consequently, for the two statistics, both probabilities of Type I error are greater than the corresponding significance level with any sample size. As previously analyzed, since the distribution of $\hat{\xi}$ is characterized with positive skewness and fat right tail, the two statistics of MB should underestimate the bias, which results in high possibility of abandoning the true.

For both the standard and the smoothed bootstrap statistics of LAMB and EDB

It can be found that the four procedures have roughly consistent performances. The rejection rates for the efficient portfolio TP are always close to the significance level with any sample size, even with small sample size of 50. However, the rejection rates for TP as the efficient portfolio are always less than those of the corresponding MB statistics. All of the four statistics lack power with small sample sizes. Fortunately, the test powers of the two statistics also increase obviously with increasing sample size. For example, with a sample size of 800, all of the p-values of the four procedures are beyond 0.5 at the significance level of 5%.

Comparing the smoothed bootstrap statistics and the standard bootstrap statistics

For the statistics of EDB, as we expect that the boundary problem have little influence on the standard bootstrap, both the standard and smoothed procedures have consistent performance on tests not only of TP but also of EP. The boundary problem should affect the results on the standard procedure of MB, for the test of TP, but both the statistical sizes of the standard and smoothed procedures of MB are similar, and it can't find significant change of the smoothed bootstrap. Moreover, we also can draw the same conclusion on the standard and smoothed procedures of LAMB for the test of TP. Interestingly, for the test of EP by using the method LAMB, all of the statistical powers of the standard procedure are lower than the smoothed procedure. However, the opposite evidences can be found on the test of EP by using the method MB. Therefore, there is no reliable evidence to suggest how effect on the statistical power result form the boundary problem.

In summary, our simulation findings are in line with the theoretical discussion in Section *4.4*. Firstly, it can be found that the statistics of LAMB and EDB by using smoothed bootstrapping and the EDB by using standard bootstrapping own the best statistical properties. All of them appear sufficiently powerful to be of practical use in application. Though there is left boundary problem, it should have minimal effect on the standard bootstrap of EDB because the statistic is determined at the right tail of the distribution of $\hat{\xi}$. Secondly, the smoothed MB statistic prefers to classify the efficient portfolio incorrectly as being inefficient even in large sample for the shape of the distribution of $\hat{\xi}$. Especially, it should act with caution when the statistic reject the null with significant asymmetric distribution of $\hat{\xi}$. But, the property of the power of the statistic is better than both statistics of the smoothed LAMB and the smoothed EDB.[25] Thirdly, for the standard procedures of LAMB and MB, the asymptotic

[25] This paper does not adjust the bias of the MB statistics for skewness as Efron (1987), which center the median of the distribution on the statistics. Since the distribution is positive skewness in most cases, the correction of the skewness will result in the increasing occurrence probability of Type I error.

properties of the statistics are very vague with the boundary problem.

5. Empirical results

5.1. Data

Post and Levy (2005) analyzed aggregate investor preferences and beliefs by testing whether the SSD, PSD or MSD criteria could rationalize the market portfolio on monthly data. They used two sets of benchmark portfolios, i.e., 25 Fama and French benchmark portfolios from Jul 1963 to Oct 2001, and 27 benchmark portfolios formed on value, size and momentum from Jul 1963 to Dec 1994, which is described in Carhart et al. (1996) and is used in Carhart (1997). Using the monthly data, we also investigate the aggregate preferences and beliefs by examining whether the CRSP all-share index, a popular proxy for the stock market portfolio, is efficient with the three SD criteria. To improve the test power, we focus on a longer 75-year sample period from Jan 1933 to Dec 2007 with 900 monthly observations. In addition, we also examine on yearly data. The sample period is from 1933 to 2007 with 75 observations. This paper analyzes three sets of benchmark portfolios including the 25 Fama and French benchmark portfolios formed on size and value, the 25 portfolios formed on size and momentum, and the 25 portfolios formed on size and short-term reversal.[26] These three sets of benchmark portfolios can capture the effects of size, value, momentum and reversal. To calculate the excess returns, we subtract the risk-free rate, which is defined as the US one month T-bill rate maintained by Ibbotson. All data are obtained from the data library on the homepage of Kenneth French.

[Insert Figure 3 Here]

[26] The Fama and French market portfolio can not be constructed exactly as a convex combination of the benchmark portfolios. Therefore, the proxies for the individual assets include the market portfolio as the 27th asset. For the inefficient portfolio, the approach does not affect the test results. But, for the efficient portfolio, it forces the test statistic to take a zero value.

Figure3 shows the six sets of benchmark portfolios in mean and standard deviation space on both monthly and yearly excess returns. The figure includes the individual assets (the clear dots), the market portfolio (MP), the mean-variance tangency portfolio (TP), and the mean-standard deviation frontier (O-TP). Obviously, on both the monthly and yearly data, the MP is inefficient in terms of mean-variance analysis related to all benchmark sets. These empirical mean and standard deviation diagrams do not reveal whether the mean-variance classification is statistically significant or not.

[Insert Table 3 Here]

Table 3 provides some descriptive statistics for excess returns of the portfolios. It addresses the skewness that plays an important role in analyzing of aggregate preferences and portfolio efficient. For example, both declining absolute risk aversion and risk seeking imply positive skewness preference, and the MV efficient can be safely employed only in the case of elliptical distribution of returns, which is a class of symmetric distribution. We give the significance of skewness at the 5% level with t-test, i.e., $t = skewness \Big/ \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}}$, where n is sample size. On monthly data, it can be found that the majority of market portfolio and benchmark portfolios has significant positive skewness, and only three benchmark portfolios have significant negative skewness, and the remaining three benchmark portfolios don’t have significant skewness. On yearly data, it also can be seen that there are many portfolios with significant positive skewness. But the number only is 44, which is much less than the number of monthly data. Moreover, all of the other portfolios don’t have significant skewness. Obviously, the results of monthly data are very different from Levy and Duchin (2004). They found that logistic distribution, which is a special case of the elliptical, fits individual asset returns and portfolio returns best. However, we illustrate a remarkable positive skewness of the returns with different sample period and different portfolios. Moreover, it also indicates the elliptical distribution can be rejected in most cases. Though we can't further evaluate the difference between their results and our results, our results still provide strong negative evidences to the MV efficient.

Now, we can't obtain more information about the risk profile of assets unless the investor utility is quadratic. In addition, the figure and the table constructed by ex post data have little statistical sense, and are silent on loss aversion. Those provide the motivation for testing whether the market portfolio is efficient in terms of different SD criteria with various utilities.

5.2. Full-sample results for risk averse and risk seeking

We adopt the bootstrap test procedures are as follows: firstly, generate 2000 pseudo-samples $\hat{X}_b, b = 1, \cdots, 2000$ and calculate the $\hat{\xi}_b(\Psi)$ for each pseudo-sample; secondly, calculate the P_{MB}, P_{LAMB} and P_{EDB} with both the standard and the smoothed methods. To obtain more convincing results, it also calculates the asymptotic statistic P_{PV} referring to Post and Van Vliet (2006). Here, reject the efficiency if the p-value is smaller than or equal to the significance level of 5%.

[Insert Table 4 Here]

Table 4 summarizes the results of the SSD, PSD and MSD criteria for full-samples.

The results of the asymptotical statistics provide little meaningful information about the aggregate investor preferences to us. The statistic can not reject the null for all of the three sets of benchmark portfolios with all of the three SD criteria except the 25 size and short-term reversal portfolios with MSD criterion on monthly data. It may result from low statistical power of the statistic.[27]

Fortunately, it can learn about the aggregate investor preferences from the bootstrap tests. Furthermore, the results of the standard and the corresponding smoothed bootstrap procedures are consistent. Based on the results, the SSD criterion performs the worst and PSD criterion performs the

[27] Post and Van Vliet (2006) considered the test of Post (2003) may lead to erroneous conclusion about the true efficiency classification. So they relaxed the restrictive null of Post (2003) to give protection against the error, which class the efficient as inefficient. Though their test involves a more favorable statistical size, the test has a less favorable statistical power in detecting inefficient portfolios in small samples.

best not only on monthly data but also on yearly data.

The market portfolio is significantly SSD inefficient for all of the six bootstrap statistics with all of the three sets of benchmark portfolios on both monthly and yearly horizon. Consequently, we reject the SSD criterion. These results are in line with those of Post (2003) and Post and Levy (2005). It is suggested that no concave utility function can rationalize the market portfolio. And it also indicates that the MV efficient with quadratic utility should be surely denied, which confirms the results of the skewness in table 3. Hence, the aggregate of the investor preferences should be not globally risk averse, and we should pay more attention to investigate the local risk seeking behavior.

The MSD criterion is rejected by all of the six bootstrap statistics with all of the three sets of benchmark portfolios on yearly data and with the 25 size and short-term reversal portfolios on monthly data. At the significance level of 5%, all of the p-values of bootstrap procedures are close to zero. In addition, the MSD criterion is also rejected by both the standard and smoothed statistics of MB with the 25 Fama and French benchmark portfolios on monthly data. These results apparently indicate that the reversed S-shaped utility function cannot rationalize the market portfolio.

The PSD criterion only has been rejected by the standard and smoothed statistics of MB with the 25 size and short-term reversal portfolios on monthly data, while all of the p-values of the four LAMB and EDB statistics are large than 0.125. And it can be found the $\hat{\xi}_b$'s skewness(=1.422) is positive. The Section *4.1* suggests the statistics of MB may commit Type I error when the skewness of the distribution of $\hat{\xi}$ is positive. In addition, the simulation also shows that the statistical sizes of MB statistics are often greater than the desired level of significance. Therefore, the results of LAMB and EDB should be more reliable. Then, the prospect theory established by Kahneman and Tversky (1979) may be best to capture investor preferences, which indicates that the aggregate utility is S-shaped.

We think the results of PSD and MSD are different from many experimental evidences supporting reverse S-shaped preferences, but the results don't deny those results of experiments because the individual preferences is not equal to the aggregate preferences as discussed in Section *2.1*.

The results of PSD and MSD are opposite to Post and Levy (2005). We consider our results may be more reasonable. There are four main reasons. Firstly, we test the efficiency by using a larger sample size, including pre-1963 period and post-2001 period. And with increasing sample size, the power of the bootstrap statistics will be greatly improved. Secondly, the statistics of MB has smaller statistical size than both of the statistics of LAMB and EDB in identifying the efficient portfolio. Thirdly, we standardize betas by $\sum_{t=1}^{T}\beta_t/T=1$, which consider weakly increasing and decreasing to utility letting beta equal to zero, which Post and Levy (2005) didn't account for. Fourthly, we also study on yearly data and the results strengthen our findings of monthly data, which give strong backing to the S-shaped preferences rather than the reverse S-shaped preferences.

5.3. Results of rolling window analysis for PSD

To obtain more convincing results, we further employ a rolling window analysis on monthly horizon. With 60-month steps, we consider all 240-month samples from Jan 1933 to Dec 2007. We compute the p-values of the tests of PSD including all of the 12 subsamples.

[Insert Table 5 Here]

Table 5 reports the results. For the tests by the asymptotic statistic, it can be observed that we accept the PSD efficiency in most cases, and only reject the PSD efficiency in subsample 9 with the 25 Fama and French benchmark portfolios and the 25 size and momentum benchmark portfolios. At the same time, it can be found that we accept the PSD efficiency in all of subsamples with all of the bootstrap statistics. Hence, our findings are robust. It also indicates that the risk preferences of the market don't change through time, and the tests of market portfolio efficiency don't affected by time variation of sample.

To sum up, we can believe that the rolling window results further reinforce our conclusion that the market is PSD efficient and the aggregate utility of investors is S-shaped. The investors are risk-aversion for gains and risk-seeking for loses. Then, they are willing to pay a premium for stocks that give downside protection in the bull market and upside potential in the bear market.

5.4. Results for positive skewness preference

Considering positive skewness preference, we test PSD efficiency with linear constraints as equation (14) and the results are shown in Table 6. The asymptotical statistics also can not reject the null for all of the three sets of benchmark portfolios on both monthly returns and yearly returns. However, all of the bootstrap statistics reject the assumption on skewness preference for the sets of the 25 size and short-term reversal portfolios on monthly data. Moreover, all of the bootstrap statistics also deny the assumption on skewness preference for the sets of the 25 size and momentum portfolios on yearly data.

[Insert Table 6 Here]

In fact, the table also illustrates the results of various bootstrap procedures are not exactly the same. For example, for the 25 size and short-term reversal portfolios on yearly returns, the statistic of smoothed MB is 0.0009, which is less than 5%, while all of the remaining bootstrap statistics don't reject the null. Since the $\hat{\xi}_b$'s skewness(=0.982) is positive, the smoothed MB procedures may be erroneously declined efficient portfolio. However, the statistics of LAMB and EDB are very close to 5% with the maxima is 0.0675.

On all accounts, our empirical evidences don't entirely support positive skewness preference. It should not be surprised with our results of skewness preferences, which have been discussed in a number of previous researches. Barberis and Huang (2008) considered it is hard to forecast a security's future skewness: past skewness, the most obvious potential predictor, has little actual predictive power. The time variation in skewness also can be found from our empirical results. Table 3 in our paper, for the 25 Fama and French portfolio, shows the skewness of twenty three portfolios is significant positive. On the contrary, for the 25 Fama and French portfolio, from the table 1 in Post and Levy (2005), only four portfolios are with positive skewness but not anyone is significant, and all of the remaining twenty one portfolios are with negative skewness and fourteen are significant, due to different sample period. In addition, the time variation of skewness also may be the vital to reconcile our study with Levy and Duchin (2004). Here the results don't fully deny the skewness preference in other theory analyze and experiments. But, in the real market world, the skewness of asset is time variation. We think it is a reasonable ground for the aggregate preferences don't have skewness feature. And it also indicates that one should be caution with the idea of profiting by skewness.

5.5. Results for loss aversion

The tests on PSD efficiency inspires us to further examine the assumption of loss aversion with the restrictions referring to equation (17), and the results are shown in Table 7. The evidences strong claim the loss aversion is a prominent feature of the aggravate preferences. On both the monthly and yearly returns, for all of the asymptotical statistics and the bootstrap statistics, the assumption of loss aversion can't be declined. The discovery can help us understand the equity premium when the market representative investor has the prospect theory preferences.[28]

[Insert Table 7 Here]

[28] See, Benartzi and Thaler (1995).

5.6. Investment horizon

The practitioners' advices indicate that the optimal investment decision rule is horizon dependent. However, it is a controversial view.[29] Many related researches only narrow to a specific utility function, especially with the MV preferences. Our study imposing less restrictive assumptions on utility function examines various preferences, so the conclusions are more general. Our results only fully confirm the PT preferences, with S-shaped utility and loss aversion, rather than expected utility preferences and Markowitz utility preferences. Hence, the results indicate who study the investment horizon should be under the framework of PT instead of other utility theory. For example, Levy and Duchin (2004) found the asset return is with logistic distribution on short horizon, but with positively skewness distribution on long horizon. It implies strong support for MV preferences at short horizons and the optimal investment decision of an investor should change at long horizons. But, it can't know what aggregate preferences has the market at long horizons. Our results of descriptive statistics in table 1 and the analysis of time variation of the skewness on monthly data have challenged the finding at short horizon of Levy and Duchin (2004). And, the results of SD test further declined the conclusion of Levy and Duchin (2004) at short horizon. In addition, the examination of SD also suggests the market equilibrium should be reached with PT preferences at monthly horizon and yearly horizon. So it also provides a stimulus for further research of market puzzles based on PT at different horizons.

Our results suggest it doesn't change the type of the aggregate preferences from monthly horizon to yearly horizon. Firstly, the individual investor may be various in his planned investment horizon. However, the preference of the representative investor is always PT preference. Secondly, the results don't mean the degree of risk aversion, risk seeking and loss aversion of the aggregate preferences is entirely independent

[29] For example, it can see the supportive evidences from Hansson and Persson (2000), Levy and Duchin (2004), and the oppositive findings from Samuelson (1994), Bodie, Merton and Samuelson (1992).

on investment horizon. Unfortunately, it can't compare the degree at different horizons. Because Post and Levy (2005) had stressed that the example utility functions, estimating by the SD testing, are not unique.

6. Conclusions

It can be argued that the bootstrap procedures of smoothed LAMB, smoothed EDB and standard EDB are useful for analyzing the SD efficiency. And，the three statistics own the best statistical properties for both size and power with large sample size. Even with small sample size, they also have satisfactory statistical size. Moreover, it can be concluded that the bootstrap tests of MB may easily commit the Type I error.

Many asset pricing anomalies are hard to understand in the context of the expected utility paradigm and globally risk-aversion preference. Under a very general framework of risk preference, the paper investigates the aggregate investor preferences and beliefs of the US stock market by examining enduring puzzles in finance: market size, value, price momentum, and price reversal effects in stock returns. Fortunately, it can be seen that inferences about the aggregate preferences in our study are not heavily affected by the exact test procedure. Moreover, at monthly and yearly horizon, the findings are consistent. Our results reject SSD efficiency of the market portfolio. Thus it should pay particular attention to non-expected utility theories and risk seeking preference for further research. Moreover, our results reject the criterion of MSD, on the contrary, accept the criterion of PSD. Encouragingly, our examination also accepts loss aversion-a vital assumption in the framework of the PT. Therefore, the market is efficient and the prospect theory may be a more prominent non-expected utility theory. It also suggests the aggregate utility function of the representative investor is in line with the PT, who will adopt different risk attitudes when they face various prospects. The aggregate of investors' preferences is not globally risk-aversion, but risk-aversion for gains and risk-seeking for losses, and more sensitive to losses, i.e., the utility is S-shaped, and steeper for losses than for gains. Therefore, it should probe into asset pricing model and financial anomalies by S-shaped and loss aversion preferences. In fact, the study of Barberis and Hung (2008) show PT can exactly produce the CAPM formula. Thus, the S-shape risk preference may be suffices to explain many asset pricing anomalies.

However, our examination isn't convincing of positive skewness preference. Furthermore, the results also can reject kurtosis preference and higher order preferences. Hence, in the market, it may be difficult to benefit from the asset through its features on skewness, kurtosis, or other higher order central moments.

It should be noted that our results only can illustrate the aggregate investor preferences of the market, and can't reveal the preferences for individuals. Hence, it can't reject skewness preference, kurtosis preference or higher order preferences for individual. For the researchers, who deeply believe higher order preferences, we consider the examination may provide a stimulus for further research for differences between the aggregate behavior characteristics and individual behavior features. And, our study also can't decline the expected utility preference, the Markowitz utility preference, and the MV preference for individual. However, it is worth mentioning that our findings may reveal the general laws of the market. In the further, it should attach importance to studies on heterogeneous components and emergent behavior, which may result in disappearance of individual preferences in aggregate view.

Of course, it is doubtful whether the results are trustworthy because the SD test assumes that return observations are serially identical and independent distributed (IID). Fortunately, Sharpe (2007) showed that an investor who holds the market portfolio will satisfy an equilibrium equation[30] in the static-state by the analysis of marginal utility and state prices. The equation is consistent with the

[30] The equation is $\sum_{k\in\Theta} u'(\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{\lambda})\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{\lambda}/K = \sum_{k\in\Theta} u'(\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{\lambda})x_{ik}/K = \sum_{k\in\Theta} u'(\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{\lambda})r_f/K, \forall i \in \mathbf{I}$, with r_f is risk free rate.

first-order condition for optimality of our test. Therefore, our results may not be biased by the IID assumption.

In addition, there are reasons to doubt the reliability of the findings because we only test with ex-post parameters. Recently, Levy and Roll (2010) found slight variations in parameters, well within estimation error bounds, suffice to make the market proxy MV efficient. Their findings also suggested ex-ante MV efficiency is consistent with the observed parameters. How can our research be reconciled with their study? Firstly, their results mainly inference the efficient based on the parameters are “closest” their observed sample rather than the population parameters. Levy and Roll (2010) also considered their research doesn’t constitute a proof of the empirical validity of the CAPM, but it shows that the model can not be rejected. Secondly, our SD method is testing for the efficiency of a given portfolio with respect to the utility with all possible parameters for one type preference instead of a utility with specific parameters only. Thirdly, the bootstrap procedures inference the efficient statistics in view of the bias with sampling variation, and it is encouraging to see that our results, including descriptive statistics, SD statistics of full sample, and SD statistics of rolling window on different horizon, support each other. Hence, we consider the influence of parameters on our systematic results should be little.

Acknowledgments

I thank Kenneth French for providing me with the data. I appreciate the valuable comments of the anonymous referees and detailed suggestions of the editor. In addition, financial support by the National Natural Science Foundation of China (Grant No. 71001044) is gratefully acknowledged. Any remaining errors are my own.

References

Baltussen, G., Post, G.T., Van Vliet, P., 2006. Violations of cumulative prospect theory in mixed gambles with moderate probabilities. Management Science 52, 1288-1290.

Barberis, N., Huang, M., 2001. Mental accounting, loss aversion, and individual stock returns. Journal of Finance 56, 1247-1292.

Barberis, N., Huang, M., 2008. Stock as lotteries: the implication of probability weighting for security prices. American Economic Review 98, 2066-2100.

Barberis, N., Huang, M., Santos, T., 2001. Prospect theory and asset prices. Quarterly Journal of Economics 116, 1-53.

Baucells, M., Heukamp, F.H., 2006. Stochastic dominance and cumulative prospect theory. Management Science, 52, 1409-1423.

Benartzi, S., Thaler, R.H., 1995. Myopic loss aversion and the equity premium puzzle. Quarterly Journal of Economics 110, 73-92.

Benartzi, S., Thaler, R.H., 1999. Risk aversion or myopia? Choices in repeated gambles and retirement investment. Management Science 45, 364-381.

Bodie, Z., Merton, R.C., Samuelson, W.F., 1992. Labor supply flexibility and portfolio choice in a life cycle model. Journal of Economics Dynamics and Control, 16, 427-49.

Carhart, M.M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57-82.

Carhart, M.M., Krail, R.J., Stevens, R.I., Welch, K.E., 1996. Testing the conditional CAPM. Working Paper, University of Chicago.

Conrad, J.S., Kaul, G., 1998. An anatomy of trading strategies. Review of Financial Studies 11, 489-519.

Dittmar, R.F., 2002. Nonlinear Pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. Journal of Finance 57, 368-403.

Efron, B., 1987. Better bootstrap confidence intervals. Journal of the American Statistical Association 82, 171-201.

Efron, B., Tibshirani, R. J., 1993. An introduction to the bootstrap. Chapman and Hall, New York.

Fama, E.F., French, K.R., 1992. The Cross-section of expected stock returns. Journal of Finance 47, 427-465.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3-56.

Fama, E.F., French, K.R., 1996. Multifactor explanations of asset pricing anomalies. Journal of Finance 51, 55-84.

Fong, W., Lean, H., Wong, W., 2008. Stochastic dominance and behavior towards risk: The market for internet stocks. Journal of Economic Behavior and Organization 68, 194–208.

Friend, I., Westerfield, R., 1980. Coskewness and capital asset pricing. Journal of Finance 35, 897-913.

Hadar, J., Russell, W.R., 1969. Rules for ordering uncertain prospects. American Economic Review

59, 25-34.
Hall, P., 1986. On the number of bootstrap simulations required to construct a confidence interval. Annals of Statistics 14, 1453-1462.
Hanoch, G., Levy, H., 1969. The efficiency analysis of choices involving risk. Review of Economic Studies 36, 335-346.
Hansson, H., Persson, M., 2000. Time diversification and estimation risk. Financial Analysts Journal, 56(5), 2000, 55-62.
Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing tests. Journal of Finance 55, 1263-1295.
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65-91.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47, 263-292.
Kraus, A., Litzenberger, R., 1976. Skewness preferences and the valuation of risk assets. Journal of Finance 31, 1085-1100.
Levy, H., Duchin, R., 2004. Asset return distributions and the investment horizon. Journal of Portfolio Management 30, 47-62.
Levy, H., Wiener, Z., 1998. Stochastic dominance and prospect dominance with subjective weighting functions. Journal of Risk and Uncertainty 16, 147-163.
Levy, M., Levy, H., 2002. Prospect theory: much ado about nothing? Management Science 48, 1334-1349.
Levy, M., Roll, R., 2010. The market portfolio may be mean/variance efficient after all. Review of Financial Studies 23, 2464-2491.
Linton, O., Maasoumi, E., Whang, Y., 2005. Consistent testing for stochastic dominance under general sampling schemes. Review of Economics studies 72, 735-765.
Markowitz, H., 1952. The utility of wealth. Journal of Political Economy 60, 151-158.
Mauboussin, M.J., 2002. Revisiting market efficiency: The stock market as a complex adaptive system. Journal of Applied Corporate Finance 14, 47-55.
Post, G.T., 2003. Empirical tests for stochastic dominance efficiency. Journal of Finance 58, 1905-1931.
Post, G.T., Levy, H., 2005. Does risk seeking drive stock prices? A stochastic dominance analysis of aggregate investor preferences and beliefs. Review of Financial Studies 18, 925-953.
Post, T., Van Vliet., P., 2006. Downside risk and asset pricing. Journal of Banking and Finance 30, 823-849.
Sharpe, W.F., 2007. Expected utility asset allocation. Financial Analysts Journal 63, 18-30.
Silverman, B.W., 1986. Density estimation for statistics and data analysis. Chapman and Hall, London.
Silverman, B.W., 1987. The bootstrap: to smooth or not to smooth. Biometrika 74: 469-479.

Simar, L., and Wilson, P.W., 1998. Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models. Management Science 44, 49-61.

Samuelson, P.A., 1994. The long-term case for equities and how it can be oversold. Journal of Portfolio Management 16, 15-24.

Thaler, R.H., 1980. Toward a positive theory of consumer choice. Journal of Economic Behavior and Organization, 1, 39-60.

Thaler, R.H., Johnson, E.J., 1990. Gambling with the house money and trying to break even: the effects of prior outcomes on risky choice. Management Science 36, 643–660.

Tversky, A, Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty 5, 297-323.

Wakker, P.P., 2003. The data of Levy and Levy (2002) "Prospect theory: much ado about nothing?" actually support prospect theory. Management Science 49, 978-981.

Wakker, P.P., Tversky, A., 1993. An axiomatization of cumulative prospect theory. Journal of Risk and Uncertainty 7, 147-176.

Whitmore, G.A., 1970. Third-degree stochastic dominance. American Economic Review 60, 457-459.

Wong, W., Chan, R., 2008. Prospect and Markowitz stochastic dominance. Annals of Finance 4, 105-129.

Table 1
Descriptive statistic of the stochastic dominance efficient statistics. It contains descriptive statistic of the statistics $\hat{\xi}$ for stochastic dominance criterion with mean-variance (MV) efficient. It analyzes two different test portfolios for various sample sizes (T= 50, 100, 400, 800). The first test portfolio is the tangency portfolio (TP), which is MV efficient. The second test portfolio is the equal weighted portfolio (EP), which is known to be MV inefficient. JB is test statistic of Jarque-Bera, and * is significance at the 1% level at least. The results are based on 1000 random samples from a multivariate normal population distribution with joint moments equal to the sample moments of monthly excess returns of the 5 Fama and French BE/ME stock portfolios and the one-month US Treasury bill during the period from Jan 1933 to Dec 2007.

Portfolio	TP				EP			
Sample size	50	100	400	800	50	100	400	800
Mean	0.2390	0.1484	0.0610	0.0409	0.2836	0.1958	0.1162	0.0971
Median	0.1932	0.1186	0.0426	0.0277	0.2481	0.1778	0.1073	0.0940
Maximum	1.6994	1.0179	0.3973	0.2749	1.6909	0.9561	0.4117	0.2866
Minimum	0.0000	0.0000	0.0000	0.0000	0.0081	0.0129	0.0185	0.0132
Standard deviation.	0.2144	0.1418	0.0676	0.0452	0.1743	0.1119	0.0551	0.0372
Skewness	1.4282	1.4330	1.3808	1.3229	1.9236	1.6840	1.0583	0.7284
Kurtosis	6.5945	6.2737	5.0401	4.7199	10.6096	9.0407	5.0848	4.4432
JB	878.317*	788.828*	491.186*	414.941*	3029.486*	1993.093*	367.765*	175.208*

Table 2
Statistical properties of different bootstrap procedures. It includes the statistical size and power of the different bootstrap procedures for mean-variance efficient by stochastic dominance criterion with various sample sizes (T= 50, 100, 400, 800) and significance levels (10%, 5%, 1%). It shows the probabilities that a given portfolio is classified as efficient with bootstrap statistics by the mean bias procedure (P_{MB}), the level adjust mean bias (P_{LAMB}), and the entire distance bias procedure (P_{EDB}). The results of tangency portfolio (TP) give the size of the statistics and the results of the equal weighted portfolio (EP) show the power of the statistics. Each cell includes both the results of the standard and the smoothed bootstrap, and the result of the former is in the bracket. The results are based on 1000 random samples from the multivariate normal population distribution with joint moments equal to the sample moments of monthly excess returns of the 5 Fama and French BE/ME stock portfolios and the one-month US Treasury bill during the period from Jan 1933 to Dec 2007. And all of the bootstrap statistics are computed with 1000 preudo-samples.

Portfolio	TP			EP		
Significance level	10%	5%	1%	10%	5%	1%
T=50						
P_{MB}	0.142 (0.142)	0.084 (0.078)	0.025 (0.025)	0.211 (0.214)	0.130 (0.139)	0.057 (0.072)
P_{LAMB}	0.121 (0.117)	0.050 (0.049)	0.007 (0.008)	0.155 (0.153)	0.065 (0.064)	0.011 (0.011)
P_{EDB}	0.121 (0.118)	0.050 (0.048)	0.008 (0.009)	0.155 (0.154)	0.065 (0.064)	0.011 (0.011)
T=100						
P_{MB}	0.138 (0.135)	0.083 (0.084)	0.030 (0.030)	0.278 (0.288)	0.207 (0.214)	0.090 (0.107)
P_{LAMB}	0.124 (0.121)	0.053 (0.053)	0.008 (0.008)	0.228 (0.226)	0.114 (0.113)	0.016 (0.015)
P_{EDB}	0.124 (0.123)	0.053 (0.053)	0.008 (0.008)	0.228 (0.229)	0.012 (0.113)	0.015 (0.016)
T=400						
P_{MB}	0.145 (0.138)	0.094 (0.083)	0.035 (0.024)	0.602 (0.604)	0.492 (0.497)	0.264 (0.289)
P_{LAMB}	0.132 (0.131)	0.063 (0.058)	0.011 (0.011)	0.537 (0.537)	0.337 (0.333)	0.106 (0.09)
P_{EDB}	0.132 (0.130)	0.059 (0.058)	0.010 (0.011)	0.536 (0.578)	0.336 (0.333)	0.101 (0.095)
T=800						
P_{MB}	0.144 (0.137)	0.093 (0.076)	0.047 (0.032)	0.811 (0.814)	0.717 (0.727)	0.482 (0.512)
P_{LAMB}	0.132 (0.130)	0.055 (0.055)	0.012 (0.014)	0.767 (0.761)	0.579 (0.574)	0.233 (0.221)
P_{EDB}	0.132 (0.130)	0.054 (0.055)	0.012 (0.013)	0.767 (0.763)	0.575 (0.574)	0.231 (0.227)

Table 3
Descriptive statistics of portfolios. It contains descriptive statistic for the excess returns of portfolios on both monthly and yearly data. The excess returns are computed from the raw return observations by subtracting the return on the one-month US Treasury bill. The portfolios include the CRSP all-share index, the 25 Fama and French benchmark portfolios formed on size and value, the 25 portfolios formed on size and momentum, and the 25 portfolios formed on size and short-term reversal. The monthly sample period is from Jan 1933 to Dec 2007 and the yearly sample period is from 1933 to 2007. Here, S.D. is standard deviation, and Ske is skewness, and Kur is Kurtosis. * represents the skewness is significance at the 5% level at least. All data are obtained from the data library on the homepage of Kenneth French.

		Monthly Data				Yearly Data			
		Mean	S.D.	Ske	Kur	Mean	S.D.	Ske	Kur
		Market Portfolio							
		0.731	4.833	0.137*	6.498	9.415	18.955	-0.083	-0.008
BE/ME	Size	25 Fama and French Portfolio							
Growth	Small	0.600	10.792	1.818*	14.879	6.661	36.495	0.728*	1.107
2	Small	0.996	9.502	3.736*	52.700	12.548	33.411	0.495*	0.446
3	Small	1.201	8.294	1.884*	16.482	15.868	31.563	0.320	-0.088
4	Small	1.375	7.876	3.075*	38.042	18.960	35.032	1.276*	3.856
Value	Small	1.551	8.742	3.296*	35.739	21.339	37.967	1.316*	3.616
Growth	2	0.777	7.833	0.445*	5.623	10.133	30.464	0.369	-0.072
2	2	1.106	7.282	1.939*	23.942	14.495	29.359	1.099*	3.956
3	2	1.188	6.810	2.210*	26.177	16.070	28.882	0.918*	2.652
4	2	1.246	6.895	1.839*	21.144	17.060	30.676	1.237*	4.106
Value	2	1.369	7.687	1.297*	12.997	18.272	31.220	0.619*	1.087
Growth	3	0.821	7.152	0.953*	9.866	10.702	28.393	1.063*	4.973
2	3	1.018	6.108	0.379*	7.720	13.499	25.199	0.742*	3.262
3	3	1.078	6.126	1.262*	16.920	14.253	24.736	0.519*	1.103
4	3	1.147	6.049	1.166*	13.233	15.354	25.472	0.424	0.303
Value	3	1.268	7.387	1.147*	12.377	16.881	30.060	0.655*	1.013
Growth	4	0.753	5.830	-0.055	3.127	9.561	21.696	0.128	0.466
2	4	0.845	5.735	0.833*	13.708	11.026	23.016	1.079*	5.125
3	4	1.035	5.578	0.382*	7.754	13.691	23.681	0.857*	2.738
4	4	1.041	5.918	0.686*	8.246	13.969	25.169	0.476*	0.744
Value	4	1.203	7.655	1.505*	17.412	15.840	32.527	1.630*	5.735
Growth	Big	0.661	4.957	0.137*	4.716	8.521	18.986	-0.171	-0.470
2	Big	0.682	4.723	-0.042	4.074	8.705	17.104	-0.041	-0.276
3	Big	0.817	4.811	0.768*	11.205	10.548	18.934	0.690*	1.491
4	Big	0.874	5.715	1.627*	20.355	11.214	22.332	0.855*	2.628
Value	Big	0.986	6.960	0.913*	11.061	12.580	26.026	0.390	0.672
Momentum	Size	25 Size and Momentum Portfolios							
Low	Small	0.688	9.951	2.858*	22.532	10.227	43.130	1.156*	1.843
2	Small	1.221	8.523	3.907*	46.696	17.290	41.511	2.470*	11.931
3	Small	1.413	7.868	3.219*	39.932	20.204	40.174	2.236*	9.748
4	Small	1.646	8.616	4.219*	53.658	23.708	47.153	3.504*	20.608
High	Small	1.794	8.418	1.729*	19.185	25.388	42.537	2.201*	10.108
Low	2	0.455	8.629	1.669*	14.688	6.116	33.654	0.733*	0.774
2	2	1.024	7.453	2.631*	27.978	13.937	33.041	2.090*	10.173
3	2	1.079	6.404	0.982*	11.460	14.380	27.933	1.426*	5.375
4	2	1.371	6.891	2.271*	30.129	18.986	32.775	1.944*	9.202

High	2	1.611	7.473	0.764*	10.554	21.939	31.650	0.530*	0.904
Low	3	0.454	8.145	1.492*	12.252	5.379	27.906	0.213	-0.200
2	3	0.857	6.681	1.428*	14.303	10.969	25.388	0.744*	2.560
3	3	0.945	6.159	1.179*	13.459	12.191	23.772	0.495*	1.589
4	3	1.113	5.905	0.523*	10.539	15.246	27.166	1.228*	4.550
High	3	1.484	6.759	-0.003	4.785	19.746	26.976	0.193	-0.050
Low	4	0.452	7.581	0.912*	10.413	5.259	25.630	0.104	0.224
2	4	0.733	6.276	1.255*	16.099	8.967	21.607	0.591*	1.944
3	4	0.839	5.824	1.503*	18.635	10.780	22.380	1.095*	5.207
4	4	1.069	5.654	1.021*	14.765	14.074	23.726	1.828*	8.854
High	4	1.407	6.241	-0.194*	3.183	18.975	26.904	0.547*	1.187
Low	Big	0.322	7.644	-2.040*	38.759	4.292	28.187	0.021	3.807
2	Big	0.607	5.439	1.562*	19.415	7.620	18.838	0.068	0.514
3	Big	0.616	4.948	0.475*	9.757	7.652	17.467	-0.074	0.159
4	Big	0.817	4.855	0.195*	4.734	10.397	18.129	0.157	-0.067
High	Big	1.046	5.564	-0.269*	2.767	13.887	23.088	0.030	-0.261
Reversal	Size	25 Size and Short-term Reversal Portfolios							
Low	Small	2.200	10.092	2.892*	25.527	36.207	83.570	5.122*	35.343
2	Small	1.459	8.848	2.881*	31.430	20.731	43.498	1.834*	5.691
3	Small	1.297	8.406	3.343*	35.074	18.154	40.262	2.240*	9.714
4	Small	0.895	8.270	4.255*	59.368	12.461	33.251	0.376	-0.292
High	Small	0.084	8.494	1.868*	18.157	1.438	30.854	0.506*	0.499
Low	2	1.737	8.757	2.025*	20.215	25.687	51.917	3.391*	19.099
2	2	1.423	7.227	1.909*	21.817	19.567	32.179	1.109	4.008
3	2	1.144	6.915	1.905*	22.275	15.303	27.904	0.416	1.000
4	2	0.898	7.005	2.305*	27.146	11.743	27.297	0.745	2.274
High	2	0.348	7.175	0.691*	10.037	4.504	26.394	0.298	0.081
Low	3	1.561	7.831	0.768*	6.909	21.864	38.608	1.934*	8.232
2	3	1.215	6.516	0.911*	9.851	16.064	25.490	0.294	1.142
3	3	1.114	6.361	1.407*	15.389	14.876	26.225	0.669*	2.107
4	3	0.816	6.137	1.281*	18.800	10.643	23.881	0.389	0.854
High	3	0.380	6.579	0.685*	11.288	4.547	22.869	0.000	-0.196
Low	4	1.373	7.380	1.044*	12.258	18.135	29.387	0.415	1.001
2	4	1.157	5.980	0.495*	8.189	15.051	22.497	0.159	0.653
3	4	0.976	5.527	0.506*	6.576	12.645	21.480	0.694*	2.294
4	4	0.760	5.703	1.465*	23.341	9.838	21.451	0.377	1.464
High	4	0.488	6.220	0.583*	10.851	6.445	26.151	1.732*	8.506
Low	Big	0.931	6.108	0.706*	6.400	11.701	22.205	0.309	0.093
2	Big	0.841	5.071	0.297*	5.494	11.041	20.099	-0.216	0.009
3	Big	0.735	4.819	0.421*	6.629	9.523	18.136	-0.287	-0.086
4	Big	0.686	4.858	1.111*	17.092	8.883	19.258	0.284	0.629
High	Big	0.431	5.406	1.176*	18.797	5.418	19.080	0.161	2.050

Table 4
Results for risk averse and risk seeking. It includes the results of aggregate preferences for risk averse and risk seeking, which test whether the market portfolio (i.e., the CRSP all-share index) is efficient on both monthly horizon (the sample period is from Jan 1933 to Dec 2007) and yearly horizon (the sample period is from 1933 to 2007) by various stochastic dominance criteria including Second-order Stochastic Dominance (SSD), Prospect Stochastic Dominance (PSD) and Markowitz Stochastic Dominance (MSD). It tests the efficient relative to all portfolios formed from the one-month US Treasury bill and a set of risky benchmark portfolios (the 25 Fama and French benchmark portfolios formed on size and value, the 25 portfolios formed on size and momentum, or the 25 portfolios formed on size and short-term reversal). It shows the observed value for the test statistic $\hat{\xi}$, the skewness of the bootstrap statistics $\hat{\xi}_b$, and the p-value that the market portfolio is classified as efficient with asymptotic statistic P_{PV} referring to Post and Van Vliet (2006), and bootstrap statistics by the mean bias procedure (P_{MB}), the level adjust mean bias (P_{LAMB}), and the entire distance bias procedure (P_{EDB}). Moreover, for the bootstrap statistics, they also include both the results of the standard and the smoothed bootstrap, and the result of the former is in the bracket. In addition, all of the bootstrap statistics are computed with 2000 preudo-samples. Here P_{LAMB} adjusts for significance level of 5%. * represents significance at 5% level at least.

Criteria	$\hat{\xi}$	Skewness of $\hat{\xi}_b$	P_{PV}	P_{MB}	P_{LAMB}	P_{EDB}
Monthly Data						
25 Fama and French Portfolios						
SSD	0.232	0.578	0.7890	0.0010* (0.0005*)	0.0025* (0.0010)*	0.0140* (0.0140*)
PSD	0.013	1.546	1.0000	0.7925 (0.7815)	0.9360 (0.9275)	0.6650 (0.6355)
MSD	0.145	1.422	0.9845	0.0350* (0.0290*)	0.3675 (0.3880)	0.1100 (0.1070)
25 Size and Momentum Portfolios						
SSD	0.402	0.237	0.5059	0.0000* (0.0005*)	0.0005* (0.0005*)	0.0015* (0.0010*)
PSD	0.083	0.859	1.0000	0.5510 (0.5350)	0.7350 (0.7160)	0.4150 (0.4140)
MSD	0.117	0.756	0.9989	0.3365 (0.3230)	0.4380 (0.3745)	0.2895 (0.2820)
25 Size and Short-term Reversal Portfolios						
SSD	0.546	-0.107	0.0767	0.0000* (0.0000*)	0.0000* (0.0000*)	0.0000* (0.0000*)
PSD	0.128	1.175	0.9929	0.0000* (0.0000*)	0.5140 (0.4835)	0.1280 (0.1260)
MSD	0.655	0.128	0.0346*	0.0000* (0.0000*)	0.0000* (0.0000*)	0.0000* (0.0000*)
Yearly data						
25 Fama and French Portfolios						
SSD	3.698	0.732	0.8169	0.0000* (0.0095*)	0.0165* (0.0130*)	0.0315* (0.0300*)
PSD	0.570	1.076	0.9996	0.5145 (0.4850)	0.7600 (0.7380)	0.3310 (0.3310)
MSD	5.509	0.476	0.9068	0.0000* (0.0000*)	0.0000* (0.0000*)	0.0105* (0.0090*)
25 Size and Momentum Portfolios						
SSD	5.208	0.443	0.6566	0.0025* (0.0025*)	0.0025* (0.0025*)	0.0090* (0.0080*)
PSD	1.833	0.520	0.9535	0.2215 (0.2175)	0.3165 (0.3240)	0.1870 (0.1825)
MSD	7.914	0.175	0.3642	0.0000* (0.0000*)	0.0000* (0.0000*)	0.0005* (0.0000*)
25 Size and Short-term Reversal Portfolios						
SSD	5.004	0.380	0.7435	0.0000* (0.0000*)	0.0015* (0.0120*)	0.0080* (0.0070*)
PSD	0.588	0.601	0.9998	0.6320 (0.6175)	0.7225 (0.6880)	0.5860 (0.5735)
MSD	7.836	0.094	0.4808	0.0000* (0.0000*)	0.0000* (0.0000*)	0.0005* (0.0005*)

Table 5
Results of rolling window analysis for Prospect Stochastic Dominance (PSD). It includes the results of rolling window analysis, which test whether the market portfolio (i.e., the CRSP all-share index) is efficient for PSD. With 60-month steps, it considers all 240-month samples from Jan 1933 to Dec 2007（including 12 subsamples）. It tests the efficient relative to all portfolios formed from the one-month US Treasury bill and a set of risky benchmark portfolios (the 25 Fama and French benchmark portfolios formed on size and value, the 25 portfolios formed on size and momentum, or the 25 portfolios formed on size and short-term reversal). It shows the observed value for the test statistic $\hat{\xi}$, the skewness of the bootstrap statistics $\hat{\xi}_b$, and the p-value that the market portfolio is classified as efficient with asymptotic statistic P_{PV} referring to Post and Van Vliet (2006), and bootstrap statistics by the mean bias procedure (P_{MB}), the level adjust mean bias (P_{LAMB}), and the entire distance bias procedure (P_{EDB}). Moreover, for the bootstrap statistics, they also include both the results of the standard and the smoothed bootstrap, and the result of the former is in the bracket. In addition, all of the bootstrap statistics are computed with 2000 preudo-samples. Here P_{LAMB} adjusts for significance level of 5%. * represents significance at the 5% level at least.

Subsample	$\hat{\xi}$	Skewness of $\hat{\xi}_b$	P_{PV}	P_{MB}	P_{LAMB}	P_{EDB}
			25 Fama and French Portfolio			
1	0.003	0.956	1.0000	0.8870(0.8845)	0.9520(0.9415)	0.9680(0.9245)
2	0.006	1.266	1.0000	0.8985(0.8925)	0.9445(0.9430)	0.9270(0.9255)
3	0.017	1.401	0.9998	0.8825(0.8830)	0.9335(0.9345)	0.9075(0.9130)
4	0.010	1.422	1.0000	0.8690(0.8675)	0.9385(0.9355)	0.8670(0.8535)
5	0.010	1.737	1.0000	0.8660(0.8640)	0.9400(0.9330)	0.8650(0.8670)
6	0.023	1.281	1.0000	0.8550(0.8555)	0.9370(0.9410)	0.8825(0.8850)
7	0.043	0.901	0.9999	0.8220(0.8225)	0.9140(0.9015)	0.8005(0.8050)
8	0.070	0.635	0.8571	0.8750(0.8745)	0.9150(0.9145)	0.8785(0.8780)
9	0.162	0.322	0.0493*	0.9105(0.9025)	0.9205(0.9230)	0.9240(0.9190)
10	0.106	0.443	0.2189	0.9015(0.9055)	0.9415(0.9435)	0.9315(0.9410)
11	0.057	0.874	0.9681	0.8645(0.8670)	0.9240(0.9285)	0.8755(0.8840)
12	0.022	0.923	0.9998	0.8680(0.8500)	0.9370(0.9330)	0.8985(0.8955)
			25 Size and Momentum Portfolios			
1	0.072	0.545	0.9681	0.8555(0.8500)	0.8825(0.8910)	0.8475(0.8540)
2	0.057	0.443	0.9981	0.7675(0.7560)	0.8405(0.7955)	0.7385(0.7385)
3	0.023	0.543	0.9999	0.8935(0.8840)	0.9175(0.9060)	0.8785(0.8775)
4	0.039	1.582	0.9978	0.8535(0.8590)	0.9425(0.9425)	0.9025(0.9165)
5	0.064	0.902	0.9347	0.8720(0.8650)	0.9285(0.9280)	0.8825(0.8900)
6	0.067	1.037	0.9723	0.8580(0.8585)	0.9280(0.9315)	0.8775(0.8835)
7	0.110	0.438	0.9750	0.8485(0.8440)	0.8865(0.8985)	0.8600(0.8560)
8	0.065	0.357	0.9696	0.8350(0.8315)	0.8760(0.8710)	0.8225(0.8195)
9	0.196	0.062	0.0342*	0.7315(0.7290)	0.7185(0.7325)	0.7280(0.7185)
10	0.143	0.679	0.5402	0.6870(0.6930)	0.7985(0.8080)	0.6130(0.6185)
11	0.069	0.731	0.9994	0.7990(0.7935)	0.8685(0.8540)	0.7545(0.7495)
12	0.037	0.931	0.9999	0.8560(0.8470)	0.9280(0.9190)	0.8485(0.8420)
			25 Size and Short-term Reversal Portfolios			
1	0.138	0.744	0.9900	0.7755(0.7765)	0.8700(0.8785)	0.7525(0.7510)
2	0.054	1.612	0.9498	0.8275(0.8260)	0.9315(0.9335)	0.8430(0.8470)
3	0.014	1.871	0.9997	0.8435(0.8380)	0.9405(0.9355)	0.8235(0.8065)
4	0.011	2.376	1.0000	0.8300(0.8275)	0.9410(0.9345)	0.7540(0.7445)
5	0.005	1.503	1.0000	0.8450(0.8335)	0.9455(0.9395)	0.8735(0.7970)
6	0.039	1.683	0.9936	0.7775(0.7800)	0.9085(0.8995)	0.6495(0.6565)
7	0.086	1.124	0.9519	0.7640(0.7655)	0.8785(0.8770)	0.6975(0.6935)
8	0.032	1.234	0.9905	0.8760(0.8765)	0.9415(0.9425)	0.9145(0.9215)
9	0.130	0.798	0.3356	0.7665(0.7635)	0.8740(0.8720)	0.7175(0.7005)
10	0.072	1.187	0.6821	0.8455(0.8520)	0.9330(0.9325)	0.8660(0.8630)
11	0.053	0.587	0.9453	0.9000(0.8965)	0.9335(0.9270)	0.9090(0.9125)
12	0.047	0.693	0.9434	0.8625(0.8655)	0.9110(0.9125)	0.8700(0.8715)

Table 6

Results for positive skewness preference. It includes the results of aggregate preferences for positive skewness preference under the framework of prospect theory, which test whether the market portfolio (i.e., the CRSP all-share index) is efficient on both monthly horizon (the sample period is from Jan 1933 to Dec 2007) and yearly horizon (the sample period is from 1933 to 2007) with Prospect Stochastic Dominance. It tests the efficient relative to all portfolios formed from the one-month US Treasury bill and a set of risky benchmark portfolios (the 25 Fama and French benchmark portfolios formed on size and value, the 25 portfolios formed on size and momentum, or the 25 portfolios formed on size and short-term reversal). It shows the observed value for the test statistic $\hat{\xi}$, the skewness of the bootstrap statistics $\hat{\xi}_b$, and the p-value that the market portfolio is classified as efficient with asymptotic statistic P_{PV} referring to Post and Van Vliet (2006), and bootstrap statistics by the mean bias procedure (P_{MB}), the level adjust mean bias (P_{LAMB}), and the entire distance bias procedure (P_{EDB}). Moreover, for the bootstrap statistics, they also include both the results of the standard and the smoothed bootstrap, and the result of the former is in the bracket. In addition, all of the bootstrap statistics are computed with 2000 preudo-samples. Here P_{LAMB} adjusts for significance level of 5%. * represents significance at 5% level at least.

$\hat{\xi}$	Skewness of $\hat{\xi}_b$	P_{PV}	P_{MB}	P_{LAMB}	P_{EDB}
			Monthly Data		
			25 Fama and French Portfolios		
0.038	1.430	1.0000	0.6025 (0.5975)	0.8590 (0.8685)	0.3770 (0.3570)
			25 Size and Momentum Portfolios		
0.101	0.763	1.0000	0.4520 (0.4585)	0.6570 (0.6160)	0.3430 (0.3470)
			25 Size and Short-term Reversal Portfolios		
0.176	1.228	0.9650	0.0000* (0.0000*)	0.0355* (0.0000*)	0.0490* (0.0485*)
			Yearly data		
			25 Fama and French Portfolios		
1.326	0.982	0.9893	0.0000* (0.0000*)	0.2215 (0.3090)	0.0630 (0.0605)
			25 Size and Momentum Portfolios		
2.848	0.407	0.8540	0.0000* (0.0000*)	0.0035* (0.0035*)	0.0210* (0.0240*)
			25 Size and Short-term Reversal Portfolios		
1.743	0.389	0.9812	0.0090* (0.0535)	0.0675 (0.0615)	0.05350 (0.05350)

Table 7

Results for loss aversion. It includes the results of aggregate preferences for loss aversion under the framework of prospect theory, which test whether the market portfolio (i.e., the CRSP all-share index) is efficient on both monthly horizon (the sample period is from Jan 1933 to Dec 2007) and yearly horizon (the sample period is from 1933 to 2007) with Prospect Stochastic Dominance. It tests the efficient relative to all portfolios formed from the one-month US Treasury bill and a set of risky benchmark portfolios (the 25 Fama and French benchmark portfolios formed on size and value, the 25 portfolios formed on size and momentum, or the 25 portfolios formed on size and short-term reversal). It shows the observed value for the test statistic $\hat{\xi}$, the skewness of the bootstrap statistics $\hat{\xi}_b$, and the p-value that the market portfolio is classified as efficient with asymptotic statistic P_{PV} referring to Post and Van Vliet (2006), and bootstrap statistics by the mean bias procedure (P_{MB}), the level adjust mean bias (P_{LAMB}), and the entire distance bias procedure (P_{EDB}). Moreover, for the bootstrap statistics, they also include both the results of the standard and the smoothed bootstrap, and the result of the former is in the bracket. In addition, all of the bootstrap statistics are computed with 2000 preudo-samples. Here P_{LAMB} adjusts for significance level of 5%. * represents significance at 5% level at least.

$\hat{\xi}$	Skewness of $\hat{\xi}_b$	P_{PV}	P_{MB}	P_{LAMB}	P_{EDB}
		Monthly Data			
		25 Fama and French Portfolios			
0.013	1.408	1.0000	0.8335 (0.8290)	0.9350 (0.9305)	0.8255 (0.8200)
		25 Size and Momentum Portfolios			
0.083	1.006	1.0000	0.6575 (0.6495)	0.8395 (0.8445)	0.4950 (0.4885)
		25 Size and Short-term Reversal Portfolios			
0.128	1.253	0.9926	0.1460 (0.1320)	0.6325 (0.6110)	0.1610 (0.1540)
		Yearly data			
		25 Fama and French Portfolios			
0.718	0.690	0.9995	0.4930 (0.4620)	0.7090 (0.6960)	0.3685 (0.3785)
		25 Size and Momentum Portfolios			
1.961	0.457	0.9458	0.1715 (0.1720)	0.2720 (0.2640)	0.1645 (0.1660)
		25 Size and Short-term Reversal Portfolios			
0.588	0.596	0.9999	0.7015 (0.7055)	0.7925 (0.7795)	0.6570 (0.6530)