DOWNSIDE RISK AND ASSET PRICING

Thierry Post and Pim van Vliet

ERIM REPORT SERIES *RESEARCH IN MANAGEMENT*	
ERIM Report Series reference number	ERS-2004-018-F&A
Publication	2004
Number of pages	28
Email address corresponding author	gtpost@few.eur.nl
Address	Erasmus Research Institute of Management (ERIM)
	Rotterdam School of Management / Rotterdam School of Economics
	Erasmus Universiteit Rotterdam
	P.O. Box 1738
	3000 DR Rotterdam, The Netherlands
	Phone: + 31 10 408 1182
	Fax: + 31 10 408 9640
	Email: info@erim.eur.nl
	Internet: www.erim.eur.nl

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:
www.erim.eur.nl

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

BIBLIOGRAPHIC DATA AND CLASSIFICATIONS		
Abstract	We analyze if the value-weighted stock market portfolio is second-order stochastic dominance (SSD) efficient relative to benchmark portfolios formed on size, value, and momentum. In the process, we also develop several methodological improvements to the existing tests for SSD efficiency. Interestingly, the market portfolio is SSD efficient relative to all benchmark sets. By contrast, the market portfolio is inefficient if we replace the SSD criterion with the traditional mean-variance criterion. Combined these results suggests that the mean-variance inefficiency of the market portfolio is caused by the omission of return moments other than variance. Especially downside risk seems to be important for rationalizing asset pricing puzzles in the 1970s and the early 1980s.	
Library of Congress Classification (LCC)	5001-6182	Business
	4001-4280.7	Finance Management, Business Finance, Corporation Finance
	HG 4529.5	Portfolio management
Journal of Economic Literature (JEL)	M	Business Administration and Business Economics
	G 3	Corporate Finance and Governance
	G 12	Asset pricing
European Business Schools Library Group (EBSLG)	85 A	Business General
	220 A	Financial Management
	220 T	Quantitative methods for financial management
Gemeenschappelijke Onderwerpsontsluiting (GOO)		
Classification GOO	85.00	Bedrijfskunde, Organisatiekunde: algemeen
	85.30	Financieel management, financiering
	85.33	Beleggingsleer
Keywords GOO	Bedrijfskunde / Bedrijfseconomie	
	Financieel management, bedrijfsfinanciering, besliskunde	
	Effectenhandel, Portfolio-analyse, Stochastische modellen, risico's, 197X-198X	
Free keywords	Stock market efficiency, asset pricing, SSD, lower partial moments, downside risk.	

Downside Risk and Asset Pricing

Thierry Post* and Pim van Vliet

Abstract

We analyze if the value-weighted stock market portfolio is second-order stochastic dominance (SSD) efficient relative to benchmark portfolios formed on size, value, and momentum. In the process, we also develop several methodological improvements to the existing tests for SSD efficiency. Interestingly, the market portfolio is SSD efficient relative to all benchmark sets. By contrast, the market portfolio is inefficient if we replace the SSD criterion with the traditional mean-variance criterion. Combined these results suggests that the mean-variance inefficiency of the market portfolio is caused by the omission of return moments other than variance. Especially downside risk seems to be important for rationalizing asset pricing puzzles in the 1970s and the early 1980s.

Keywords:

Stock market efficiency, asset pricing, SSD, lower partial moments, downside risk.

Downside Risk and Asset Pricing

THIERRY POST* and PIM VAN VLIET

Abstract

We analyze if the value-weighted stock market portfolio is second-order stochastic dominance (SSD) efficient relative to benchmark portfolios formed on size, value, and momentum. In the process, we also develop several methodological improvements to the existing tests for SSD efficiency. Interestingly, the market portfolio is SSD efficient relative to all benchmark sets. By contrast, the market portfolio is inefficient if we replace the SSD criterion with the traditional mean-variance criterion. Combined these results suggests that the mean-variance inefficiency of the market portfolio is caused by the omission of return moments other than variance. Especially downside risk seems to be important for rationalizing asset pricing puzzles in the 1970s and the early 1980s.

EFFICIENCY OF THE STOCK MARKET PORTFOLIO is a much-debated topic in financial economics. Asset pricing models that employ a representative investor, including the mean-variance based CAPM, predict that the market portfolio is efficient. At first glance, market portfolio efficiency is also consistent with the popularity of passive mutual funds and exchange traded funds that track broad value-weighted equity indexes. Nevertheless, several empirical studies suggest that the market portfolio is highly and significantly inefficient. Most notably, the market portfolio seems mean-variance (MV) inefficient relative to stock portfolios formed on variables such as market capitalization (size), book-to-market equity ratio (value) and price momentum (see for instance Basu (1977), Banz (1981), Fama and French (1992) (1993) and Jegadeesh and Titman (1993)).

The empirical results may reflect a fundamental flaw in representative investor models in general. Alternatively, the results may reflect the flaws of the MV criterion. It is well known that asset returns cannot be described by mean and variance alone. For example, the monthly returns of many stocks exhibit positive skewness and excess kurtosis. Also, a wealth of psychological research on decision-making under risk suggests that the perception of risk is more complex than variance. Especially the phenomena of skewness preference and loss aversion have attracted much attention among financial economists. This provides a rationale for replacing the MV criterion with a more general efficiency criterion that accounts for higher-order central moments (such as skewness and kurtosis) and lower partial moments (such as expected loss and semi-variance).

One popular approach to extend the MV criterion is by changing the maintained assumptions on investor preferences. If we do not restrict the shape of the return

distribution, then the MV criterion is consistent with expected utility theory only if utility takes a quadratic form.[1] Extensions can be obtained by using alternative classes of utility. For example, Kraus and Litzenberger (1976) and Harvey and Siddique (2000) assume that utility can be approximated using a third-order polynomial (to account for skewness), and Dittmar (2002) uses a fourth-order polynomial (to account for skewness and kurtosis).

Unfortunately, the researcher faces two possible obstacles in implementing this approach. First, economic theory does not forward strong predictions about the shape of investor preferences and the return distribution. The theory specifies general regularity conditions such as nonsatiation (an increasing utility function) and risk aversion (a concave utility function), but not a functional form for preferences and distributions. This introduces a serious risk of specification error. For example, a fourth-order polynomial is not sufficiently flexible to account for lower partial moments, which require a kinked utility function. Second, it is often difficult to impose the regularity conditions in practice. The regularity conditions are needed to ensure that the results are economically meaningful. Also, these regularity conditions can improve the power of statistical tests. Further, from a mathematical perspective, a concave increasing utility function is required in order to justify the common approach of checking the first-order condition to test for market portfolio efficiency.[2] Unfortunately, it can be difficult to impose the regularity conditions on a parametric utility function. For example, we cannot restrict a quadratic utility function to be globally increasing and we cannot restrict a cubic utility function to be globally concave (see, for example, Levy (1969)). Moreover, not imposing the regularity conditions frequently leads to severe violations of these conditions. For example, Dittmar (2002, Section III-d) shows that the apparent explanatory power of a quartic utility function (which accounts for skewness and kurtosis) disappears if we impose necessary conditions for risk aversion.

These empirical and theoretical considerations provide strong arguments for using the criterion of second-order stochastic dominance (SSD; Hadar and Russell (1969), Hanoch and Levy (1969) and Levy and Hanoch (1970)). This criterion avoids parameterization of investor preferences and the return distribution, and at the same time ensures that the regularity conditions of nonsatiation and risk aversion are satisfied. In brief, the market portfolio must be SSD efficient for all asset pricing models that use a nonsatiable and risk-averse representative investor, regardless of the specific functional form of the utility function and the return distribution.

In this paper, we analyze if the stock market portfolio is SSD efficient. To implement the SSD criterion, we first extend Post's (2003) empirical test for SSD efficiency in several respects. Most notably, we derive the asymptotic sampling distribution of the SSD test statistic under the true null of efficiency rather than the restrictive null of equal means that was used earlier. This extension is intended to avoid rejection of efficiency in cases where the market portfolio is efficient but the assets have substantially different means. Also, we derive a linear programming test for MV efficiency that can be compared directly with the SSD test. This allows us to attribute

differences between the two tests to omitted moments exclusively. With the resulting tests, we show that the value weighted CRSP all-share index is SSD efficient relative to common benchmark portfolios formed on size, value, and momentum. By contrast, we find that the market portfolio is significantly MV inefficient relative to value and momentum, consistent with the existing evidence on these puzzles. The SSD criterion is especially successful in rationalizing MV inefficiencies that occur in the 1970s and the early 1980s. This suggests that the asset pricing puzzles that occur in the MV framework can be explained by omitted return moments during this period. The difference with the results of Post (2003), who rejects SSD efficiency, can be attributed to Type I error caused by the use of the restrictive null of equal means and the use of an extended sample period.

We stress that our motivation for testing market portfolio efficiency follows from the *revealed preferences* of (some) investors rather than from first principles. In asset pricing theories, market portfolio efficiency generally follows from underlying assumptions about investor preferences.[3] By contrast, our motivation for testing efficiency lies in the popularity of passive mutual funds and exchange traded funds that track broad value-weighted equity indexes that strongly resemble the market portfolio (or at least popular proxies for the market portfolio). In other words, (some) investors reveal a preference for market indexes, and our objective is to rationalize their choice and to analyze their preferences. [4]

The remainder of this study is structured as follows. Section I introduces the notation, assumptions, definitions and tests that will be used throughout the text. Section II discusses the data used in our analysis. Section III empirically analyzes the SSD efficiency of the market portfolio. Finally, Section IV summarizes our conclusions and presents directions for further research.

I. Methodology

A. Assumptions

We consider a single-period, portfolio-based model of investment that satisfies the following assumptions:

Assumption 1 Investors are nonsatiable and risk averse and they choose investment portfolios to maximize the expected utility associated with the return of their portfolios. Throughout the text, we will denote utility functions by $u : \Re \to P$, $u \in U_{SSD}$, with U_{SSD} for the set of increasing and concave, once continuously differentiable, von Neumann-Morgenstern utility functions, and P for a nonempty, closed, and convex subset of $\Re$.[5,6,7] For the sake of comparison, we will also use the subset of quadratic utility functions that underlies mean-variance efficiency, that is, $U_{MV} \equiv \{u \in U_{SSD} : u(x) = ax + 0.5bx^2\}$. We follow the definition of MV efficiency by Hanoch and Levy (1970): a portfolio is efficient if and only if there exists an increasing and concave, quadratic utility function that

rationalizes the portfolio. Further, we will use u' to denote the first-order derivative or marginal utility. If the utility function belongs to U_{SSD}, then marginal utility is a positive and decreasing function. In the special case of quadratic utility, marginal utility is a positive and decreasing, linear function.[8]

Assumption 2 The investment universe consists of N-1 risky assets and a riskless asset. Throughout the text, we will use the index set $I \equiv \{1, \cdots, N\}$ to denote the different assets, with N for the riskless asset. The returns $\boldsymbol{x} \in \Re^N$ are serially independent and identically distributed (IID) random variables with a continuous joint cumulative distribution function (CDF) $G : \Re^N \rightarrow [0,1]$.

Assumption 3 Investors may diversify between the assets, and we will use $\boldsymbol{\lambda} \in \Re^N$ for a vector of portfolio weights. We focus on the case where short sales are not allowed, and the portfolio weights belong to the portfolio possibilities set $\Lambda \equiv \{\boldsymbol{\lambda} \in \Re_+^N : \boldsymbol{e}^{\mathrm{T}} \boldsymbol{\lambda} = 1\}$, with $\boldsymbol{e}$ for a unity vector with dimensions conforming to the rules of matrix algebra. The simplex Λ excludes short sales. Short selling typically is difficult to implement in practice due to margin requirements and explicit or implicit restrictions on short selling for institutional investors (see, for instance, Sharpe (1991) and Wang (1998)).[9]

Under these assumptions, the investors' optimization problem can be summarized as $\max_{\boldsymbol{\lambda} \in \Lambda} \int u(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{\lambda}) dG(\boldsymbol{x})$. A given portfolio, say $\boldsymbol{\tau} \in \Lambda$, is optimal for a given utility function $u \in U_{SSD}$ if and only if the first-order condition is satisfied:

$$\int u'(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau})(x_i - \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau}) dG(\boldsymbol{x}) \leq 0 \quad \forall i \in I, \tag{1}$$

The inequality should hold with strict equality for all assets that are included in the evaluated portfolio, that is, $i \in I : \tau_i > 0$. If all assets are included in the evaluated portfolio $(\boldsymbol{\tau} > \boldsymbol{0})$, as is true for the value-weighted market portfolio, then inequality (1) automatically reduces to $\int u'(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau})(x_i - \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau}) dG(\boldsymbol{x}) = 0 \quad \forall i \in I$.[10] Interestingly, this equality represents the least favorable case for our test of efficiency, that is, the probability of a Type I error (wrongly classifying an efficient portfolio as inefficient) achieves its maximum if all assets are included. The proof to Theorem 1 in Section IC uses this result.

Following the asset pricing terminology, we refer to violations of the first-order condition as pricing errors. We may measure the maximum pricing error as:[11]

$$\zeta(\boldsymbol{\tau},G,u) \equiv \max_{i \in I}\left\{\int u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})(x_i - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})dG(\boldsymbol{x})\right\}. \tag{2}$$

Portfolio $\boldsymbol{\tau} \in \Lambda$ is optimal relative to $u \in U_{SSD}$ if and only if $\zeta(\boldsymbol{\tau},G,u)=0$. The SSD efficiency criterion basically checks if this condition is satisfied for some $u \in U_{SSD}$. Similarly, the MV efficiency criterion checks if the first-order condition is satisfied for some $u \in U_{MV}$. To test for SSD efficiency or MV efficiency, we introduce the following measure:

$$\xi(\boldsymbol{\tau},G,U) \equiv \min_{u \in U} \zeta(\boldsymbol{\tau},G,u). \tag{3}$$

with $U \in \{U_{SSD}, U_{MV}\}$.

The efficiency criteria can equivalently be formulated in terms of the set of utility functions that rationalize the evaluated portfolio:

$$\Psi(\boldsymbol{\tau},G,U) \equiv \left\{u \in U : \zeta(\boldsymbol{\tau},G,u)=0\right\}. \tag{4}$$

Note that the evaluated portfolio may be optimal for multiple utility functions, and hence $\Psi(\boldsymbol{\tau},G,U)$ may contain multiple elements.

Definition 1 *Portfolio $\boldsymbol{\tau} \in \Lambda$ is efficient if and only if it is optimal for at least some $u \in U$, $U \in \{U_{SSD}, U_{MV}\}$, that is, $\xi(\boldsymbol{\tau},G,U)=0$, or, equivalently, $\Psi(\boldsymbol{\tau},G,U)$ is non-empty. Portfolio $\boldsymbol{\tau} \in \Lambda$ is inefficient if and only if it is not optimal for all $u \in U$, that is, $\xi(\boldsymbol{\tau},G,U)>0$, or, equivalently, $\Psi(\boldsymbol{\tau},G,U)$ is empty.*

To test the null of efficiency, that is, $H_0 : \xi(\boldsymbol{\tau},G,U)=0$, we need full information on the CDF $G(\boldsymbol{x})$. In practical applications, $G(\boldsymbol{x})$ generally is not known and information is limited to a discrete set of T time series observations.

Assumption 4 The observations are serially independently and identically distributed (IID) random draws from the CDF. Throughout the text, we will represent the observations by the matrix $\mathbf{X} \equiv (\boldsymbol{x}_1 \cdots \boldsymbol{x}_T)$, with $\boldsymbol{x}_t \equiv (x_{1t} \cdots x_{Nt})^{\mathrm{T}}$. Since the timing of the draws is inconsequential, we are free to label the observations by their ranking with respect to the evaluated portfolio, that is, $\boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau} < \boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau} < \cdots < \boldsymbol{x}_T^{\mathrm{T}}\boldsymbol{\tau}$.

Using the observations, we can construct the following empirical distribution function (EDF):

$$F_{\mathbf{X}}(x) \equiv \frac{1}{T} \sum_{t=1}^{T} 1(x_t \leq x), \tag{5}$$

Since the observations are assumed to be serially IID, $F_{\mathbf{X}}(x)$ is a consistent estimator for $G(x)$, and we may use $\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U)$ as a consistent estimator for $\xi(\boldsymbol{\tau}, G, U)$.

B. Linear programming test statistics

Following Post (2003), we may derive the following linear programming formulation for $\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U_{SSD})$:[12]

$$\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U_{SSD}) = \min_{\boldsymbol{\beta} \in \mathrm{B}_{SSD}, \theta} \left\{ \theta : \sum_{t=1}^{T} \beta_t (x_t^{\mathrm{T}} \boldsymbol{\tau} - x_{it}) / T + \theta \geq 0 \quad \forall i \in \mathrm{I} \right\}, \tag{6}$$

with $\mathrm{B}_{SSD} \equiv \left\{ \boldsymbol{\beta} \in \Re_+^T : \beta_1 \geq \beta_2 \geq \cdots \geq \beta_T; \sum_{t=1}^{T} \beta_t / T = 1 \right\}$. In this formulation, $\boldsymbol{\beta}$ represents the gradient vector $(u'(x_1^{\mathrm{T}} \boldsymbol{\tau}) \cdots u'(x_T^{\mathrm{T}} \boldsymbol{\tau}))^{\mathrm{T}}$ for some well-behaved utility function $u \in U_{SSD}$. B_{SSD} represents the restrictions on the gradient vector that follow from the assumptions of nonsatiation and risk aversion and the standardization $\sum_{t=1}^{T} u'(x_t^{\mathrm{T}} \boldsymbol{\tau}) / T = 1$.

Note that the original test of Post (2003, Thm 2) uses the standardization $\beta_T = 1$ rather than $\sum_{t=1}^{T} \beta_t / T = 1$. The original standardization has an important drawback. Specifically, the higher the degree of risk aversion of the utility function, the higher the values of all betas. Hence, increasing the level of risk aversion tends to inflate the value of test statistic relative to the case with risk neutrality ($\boldsymbol{\beta} = \boldsymbol{e}$). This lowers Post's (2003, Thm 3) p-value for testing efficiency, possibly leading to erroneous rejections of efficiency, as the p-value is based on the risk neutral case and it does not account for the level of the betas. To circumvent this problem, we use the standardization $\sum_{t=1}^{T} \beta_t / T = 1$ in this study.[13] This standardization allows for risk aversion without inflating the test statistic, because the average level of the betas is fixed. Also, the novel standardization allows utility to be weakly increasing, as some betas may equal zero.

Using linear interpolation, we may recover a full utility function from the optimal solution ($\boldsymbol{\beta}^*$) as $p(x|\boldsymbol{\beta}^*) \equiv \int_{z=-\infty}^{x} p'(z|\boldsymbol{\beta}^*) dz$, with

$$p'(z|\boldsymbol{\beta}^*) = \begin{cases} \beta_1^* & z \leq \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau} \\ \beta_1^* + (\beta_2^* - \beta_1^*)(z - \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau})/(\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau}) & \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau} \leq z \leq \boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau} \\ \vdots \\ \beta_{T-1}^* + (\beta_T^* - \beta_{T-1}^*)(z - \boldsymbol{x}_{T-1}^{\mathrm{T}}\boldsymbol{\tau})/(\boldsymbol{x}_T^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_{T-1}^{\mathrm{T}}\boldsymbol{\tau}) & \boldsymbol{x}_{T-1}^{\mathrm{T}}\boldsymbol{\tau} \leq z \leq \boldsymbol{x}_T^{\mathrm{T}}\boldsymbol{\tau} \\ \beta_T^* & z \geq \boldsymbol{x}_T^{\mathrm{T}}\boldsymbol{\tau} \end{cases} . \tag{7}$$

We can derive a linear programming test for mean-variance efficiency in the spirit of (6). As discussed in Section IA, mean-variance analysis is the special case of SSD where utility takes a quadratic form and marginal utility takes a linear form. Put differently, the gradient vector $\boldsymbol{\beta}$ must belong to $\mathrm{B}_{MV} \equiv \mathrm{B}_{SSD} \cap \{\boldsymbol{\beta} \in \Re^T : \beta_t = a + b\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau} \quad \forall t \in \Theta\}$. Hence, we obtain a linear programming test for MV efficiency by simply adding the restrictions $\beta_t = a + b\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau} \quad \forall t \in \Theta$:[14]

$$\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U_{MV}) = \min_{\boldsymbol{\beta} \in \mathrm{B}_{<V}, \theta} \left\{ \theta : \sum_{t=1}^{T} \beta_t (\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau} - x_{it})/T + \theta \geq 0 \quad \forall i \in \mathrm{I} \right\}. \tag{8}$$

This test statistic differs in several respects from the traditional MV efficiency tests, such as the Gibbons, Ross and Shanken (1989) or GRS test. First, the test is consistent with SSD by adhering to the Hanoch and Levy (1969) definition of MV efficiency. For example, the GRS test may classify the market as inefficient if all assets have the same mean. However, in this case, the market is SSD efficient, because the market is optimal for the risk neutral investor. Second, our MV test excludes short selling; see Section IA. Third, our test focuses on the maximum pricing error rather than a weighed average of all squared pricing errors, so as to allow for the case where not all assets are included in the evaluated portfolio.

C. Asymptotic sampling distribution

Our objective is to test the null hypothesis that $\boldsymbol{\tau} \in \Lambda$ is SSD or MV efficient, that is, $H_0 : \xi(\boldsymbol{\tau}, G, U) = 0$, $U \in \{U_{SSD}, U_{MV}\}$. Again, since the observations are serially IID, $F_{\mathbf{X}}$ is a consistent estimator for $G(\boldsymbol{x})$, and $\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U)$ is a consistent estimator for $\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U)$. However, $F_{\mathbf{X}}(\boldsymbol{x})$ generally is very sensitive to sampling variation and the test results are likely to be affected by sampling error in a nontrivial way. The applied researcher must therefore have knowledge of the sampling distribution in order to make inferences about the true efficiency classification. Post (2003) derived the asymptotic sampling distribution of his SSD test statistic under the null hypothesis that all assets have the same mean, that is, $H_1 : E[\boldsymbol{x}] = \mu\boldsymbol{e}$, $\mu \in \Re$. In general, H_1 gives a sufficient condition for the true null of efficiency, that is, H_0. In fact, under the null, all portfolios $\boldsymbol{\lambda} \in \Lambda$ are efficient, because they are optimal for investors with utility function $v(x) \equiv x$,

that is, the risk neutral investors. However, H_1 does not give a necessary condition for H_0, and rejection of H_1 generally does not imply rejection of H_0 and there is no guarantee that H_1 is sufficiently close to H_0. Hence, the sampling distribution H_1 may lead to erroneous conclusions about H_0. The purpose of this section is to analyze the asymptotic sampling distribution of $\xi(\tau, G, U) = 0$ under the true null of efficiency rather than the null of equal means.

Using $\Phi(\cdot | \mathbf{0}, \Sigma(\tau, G, u))$ for a N-dimensional multivariate normal distribution function with mean $\mathbf{0}$ and (singular) variance-covariance matrix $\Sigma(\tau, G, u) \equiv (\mathbf{I} - e\tau^\top)\Omega(G, u)(\mathbf{I} - e\tau^\top)^\top$, with $\Omega(G, u) \equiv \int (u'(x^\top \tau))^2 xx^\top) dG(x)$, the following theorem characterizes the asymptotic sampling distribution under the true null:

THEOREM 1 *Asymptotically, the p-value* $\Pr[\xi(\tau, F_{\mathbf{X}}, U) > y | H_0]$, $U \in \{U_{SSD}, U_{MV}\}$, $y \geq 0$, *is bounded from above by* $\Gamma(y, \Sigma(\tau, G, u)/T) \equiv (1 - \int_{z \leq ye} d\Phi(z | \mathbf{0}, \Sigma(\tau, G, u)/T))$ *for all* $u \in \Psi(\tau, G, U)$.

The theorem provides an upper bound to the p-value $\Pr[\xi(\tau, F_{\mathbf{X}}, U) > y | H_0]$. It is difficult to derive the exact p-value, because $\Psi(\tau, G, U)$ generally contains multiple elements under H_0. Also, the theorem considers the least favorable case where all assets are included in the evaluated portfolio, and $\Gamma(y, \Sigma(\tau, G, u)/T)$ represents an upper bound for the p-value in cases where some assets are excluded. While it is possible to identify an element of $\Psi(\tau, G, U)$ (see below), we generally do not know the element that minimizes the p-value $\Gamma(y, \Sigma(\tau, G, u)/T)$. Nevertheless, the upper bound can be used in the same way as the true p-value. Specifically, we may compare $\Gamma(\xi(\tau, F_{\mathbf{X}}, U), \Sigma(\tau, G, u))$, $u \in \Psi(\tau, G, U)$, with a predefined level of significance $a \in [0,1]$, and reject efficiency if $\Gamma(\xi(\tau, F_{\mathbf{X}}, U), \Sigma(\tau, G, u)) \leq a$. Equivalently, we may reject efficiency if the observed value of $\xi(\tau, F_{\mathbf{X}})$ is greater than or equal to the critical value $\Gamma^{-1}(a, \Sigma(\tau, G, u))$ $\equiv \inf_{y \geq 0}\{y : \Gamma(y, \Sigma(\tau, G, u)) \leq a\}$, $u \in \Psi(\tau, G, U)$. The statistical size or the probability of a Type I error (wrongly classifying an efficient portfolio as inefficient) of this approach is almost always smaller than the nominal significance level a.

Two results are useful for implementing the above approach in practice. First, computing p-values and critical values requires the variance-covariance matrix $\Sigma(\tau, G, u)$ for some $u \in \Psi(\tau, G, U)$. Unfortunately, G is not known. Nevertheless, we know that $F_{\mathbf{X}}$ converges to G and that $p(x | \beta^*)$ asymptotically belongs to $\Psi(\tau, G, U)$

under H_0. Hence, we may estimate $\Sigma(\tau, G, u)$ in a distribution-free and consistent manner using the sample equivalent $\Sigma(\tau, F_{\mathbf{X}}, p(x|\boldsymbol{\beta}^*))$ with elements

$$\omega_{ij}(\tau, F_{\mathbf{X}}, p(x|\boldsymbol{\beta}^*)) = \sum_{t=1}^{T} \beta_t^{*2}(x_{it} - \mathbf{x}_t^{\mathrm{T}}\tau)(x_{jt} - \mathbf{x}_t^{\mathrm{T}}\tau)/T, \ i, j \in \mathrm{I}.$$

Second, we may approximate $\Gamma(\xi(\tau, F_{\mathbf{X}}, U), \Sigma(\tau, F_{\mathbf{X}}, p(x|\boldsymbol{\beta}^*)))$, using Monte-Carlo simulation. In this paper, we will use the following approach. We first generate $S=10,000$ independent standard normal random vectors $\mathbf{w}_s \in \mathfrak{R}^{N-1}$, $s \in \{1, \cdots, S\}$, using the RNDN function in Aptech Systems' GAUSS software. Next, each random vector $\mathbf{w}_s$ is transformed into a multivariate normal vector $\mathbf{z}_s \in \mathfrak{R}^N$ with variance-covariance matrix $\Sigma(\tau, G, u)$ by using $\mathbf{z}_s = (\mathbf{I} - e\tau^{\mathrm{T}})\mathbf{D}(G, u)\mathbf{w}_s$, where $\mathbf{D}(G, u) \in \mathfrak{R}^{N \times (N-1)}$ is a matrix with the first $(N-1)$ rows equal to the Cholesky factor of the nonsingular $(N-1) \times (N-1)$ variance-covariance matrix of risky assets, and with zeros for the Nth row. Finally, $\Gamma(\xi(\tau, F_{\mathbf{X}}, U), \Sigma(\tau, G, u))$ is approximated by the relative frequency of the transformed vectors $\mathbf{z}_s$, $s \in \{1, \cdots, S\}$, that fall outside the integration region $\{\mathbf{z} \in \mathfrak{R}^N : \mathbf{z} \le y e\}$.

Theorem 1 subtly differs from Post's characterization of the sampling distribution under H_1. That characterization used the variance-covariance matrix $\Sigma(\tau, G, v)$, with $v(x) = x$, in place of $\Sigma(\tau, G, u)$. This replacement is valid only if $v \in \Psi(\tau, G, U)$, that is, if the evaluated portfolio is optimal for the risk neutral investor. This reflects the replacement of the null of efficiency (H_0) by the null of equal means (H_1); under H_1, all portfolios are optimal for the risk neutral investor. Obviously, it is relatively simple to reject H_1 and hence the p-values and critical values under this null are likely to underestimate the true values under H_0. Consequently, a test procedure that uses the sampling distribution under H_0 will involve a more favorable statistical size (=relative frequency of Type I error) and a less favorable statistical power (=one minus the relative frequency of Type II error) than a test procedure that uses the sampling distribution under H_1.

D. Simulation experiment

Using a simulation experiment, Post (2003, Section IIIC) demonstrates that his SSD test procedure based on H_1 involves low power in small samples generated from the return distribution of the well-known 25 Fama and French stock portfolios formed on size and value. Since H_0 is more general than H_1, the power of our test procedure based on H_0 will be even worse. Nevertheless, the procedure may be sufficiently powerful to be of practical use for data sets with a smaller cross-section. The lack of power in the Post experiment probably reflects the difficulty of estimating a 25-dimensional multivariate

return distribution in a nonparametric fashion. It is likely that the power increases (at an increasing rate) as the length of the cross-section is reduced to for example ten benchmark portfolios, which is common in asset-pricing tests.

To shed some light on the statistical properties of our test procedure based on H_0 in smaller cross-sections, we extend the original simulation experiment. The simulations involve ten assets with a multivariate normal return distribution. The joint population moments are equal to the sample moments of the monthly excess returns of ten Fama and French stock portfolios formed on B/M during Post's sample period from July 1963 to October 2001.[15] We will analyze the statistical properties of our SSD test procedure (reject efficiency if and only if $\Gamma(\xi(\boldsymbol{\tau},F_{\mathbf{X}},U_{SSD}),\boldsymbol{\Sigma}(\boldsymbol{\tau},F_{\mathbf{X}},p(x|\boldsymbol{\beta}^*)) \leq a)$ and our MV test procedure (reject efficiency if and only if $\Gamma(\xi(\boldsymbol{\tau},F_{\mathbf{X}},U_{MV}),\boldsymbol{\Sigma}(\boldsymbol{\tau},F_{\mathbf{X}},p(x|\boldsymbol{\beta}^*)) \leq a)$ by the rejection rates of these procedures for certain test portfolios in random samples drawn from this multivariate normal distribution.

For the true distribution, the equal weighted portfolio (EP) is known to be MV and SSD inefficient.[16] Hence, we may analyze the statistical power of the SSD and MV test procedures by the ability to correctly classify the EP inefficient. By contrast, the ex ante tangency portfolio (TP) is MV and SSD efficient and we may analyze the statistical size by the relative frequency of random samples in which this portfolio is wrongly classified as inefficient.[17]

We draw 10,000 random samples from the multivariate normal population distribution through Monte-Carlo simulation. To each sample, we also add 'observations' for a riskless asset with a return of zero in every month (recall that we use excess returns). For every random sample, we apply the MV and SSD test procedures to the efficient TP and the inefficient EP. For both procedures, we compute the size as the rejection rate for TP and the power as the rejection rate for EP. This experiment is performed for a sample size (T) of 10 to 4,000 observations and for a significance level (a) of 2.5, five, and ten percent.

Figure 1 shows the results. The size is generally substantially smaller than the nominal level of significance a, and it converges to zero. In fact, the size is smaller than one percent for samples as small as 100 observations and with a level of significance as high as ten percent. Presumably, this reflects our focus on the least favorable distribution, which minimizes Type I error.

As discussed in Section IC, $\xi(\boldsymbol{\tau},F_{\mathbf{X}},U), U \in \{U_{SSD},U_{MV}\}$, converges to $\xi(\boldsymbol{\tau},G,U)$, and we expect minimal Type II error in large samples. Indeed, for both procedures, the power goes to unity as we increase the sample size. However, in small samples, the SSD procedure is substantially less powerful than the MV procedure. For example, using a ten percent significance level, the MV procedure achieves a rejection rate of about 50 percent already for samples of about 300 observations. By contrast, The SSD procedure achieves this rejection rate only for samples of about 700 observations. At that sample size, the rejection rate of the SSD procedure increases rapidly. Interestingly, this sample

size is not uncommon for this type of application. For example, our empirical tests will use samples of 840 monthly observations. Hence, the SSD procedure appears sufficiently powerful to be of practical use in this type of application.

[Insert Figure 1 about here]

E. Further tests

Our main results rest on the MV and SSD tests derived in Section IB. Nevertheless, we will also make use of three other tests in order to interpret our test results. First, as discussed in Section IB, our MV tests uses the Hanoch-Levy definition of MV efficiency that requires nonsatiation. To determine if the MV results are due to the difference between this definition and the traditional MV efficiency definition (TMV, which does not require nonsatiation), we consider two alternative tests that do adhere to the traditional definition. First, we apply TMV, a relaxed version of our MV test that drops nonsatiation.[18] In addition, we also apply the well-known GRS test of market efficiency. Note that this test not only drops nonsatiation, but it also allows for short sales. In addition, we also apply Post's (2003) original SSD efficiency test. Recall that this test uses the sampling distribution under the null of equal means rather than the null of efficiency and hence it may erroneously reject efficiency if the assets have different means. Furthermore, inclusion of this statistic is interesting because the benchmark sets and sample periods differ.

The below Venn-diagram illustrates the various tests used in this study. The TMV efficient set is not a proper subset of the SSD efficient set. By contrast, the MV efficient set is the intersection of the TMV efficient set and the SSD efficient set.
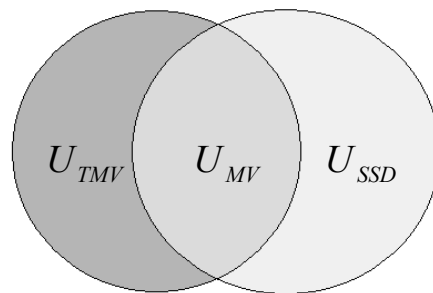
II. Data

We analyze if the CRSP all-share index is efficient. This value-weighted index consists of all common stocks listed on NYSE, AMEX, and Nasdaq. To proxy the investment universe of individual assets, we use three different sets of ten portfolios that cover the three well-known puzzles of size, value, and momentum. To calculate monthly excess returns, we subtract the risk-free rate defined as the US 30 day T-bill rate maintained by Ibbotson.

First, we use the widely used decile portfolios formed on market capitalization. The size portfolios are formed using NYSE decile breakpoint data for the total sample period also when (smaller) Nasdaq and AMEX shares are added to the CRSP database. Second, we use ten benchmark portfolios formed on book-to-market-equity ratio (value). For detailed data description and selection procedures we refer to Fama and French (1992) (1993)). In our analysis of size and value, we focus on a long 70-year sample period starting in January 1933 to December 2002 (840 months). Third, we use ten portfolios sorted on price momentum as described in Jegadeesh and Titman (2001). This benchmark set ranges from January 1965 to December 1998 (408 months).[19] These three sets of decile portfolios capture the separate effects of size, value, and momentum.

It is perhaps even more interesting to consider the combined effect of the different puzzles. For this purpose, we also analyze the 25 Fama and French (FF25) benchmark portfolios formed on size and value and the 27 Carhart (C27) portfolios formed on value, size, and momentum.[20] For the former benchmark set, data are available also from January 1933 to December 2002 (840 months). These particular 25 portfolios have become the central piece in the canon of empirical literature. For the latter benchmark set, data are limited to the period ranging from July 1963 to December 1994 (378 months). These two benchmark data sets allow us to determine if the same type of utility function can explain the various puzzles combined. Also, they allow us to analyze if the SSD criterion can capture the interaction between the size, value, and momentum effects that may occur if investors can adopt a mixed strategy of investing in for instance small value stocks or small value winners.

One should bear in mind that empirical research carried out before 1999 could give different results due to the 'delisting bias' first noticed by Shumway (1997)) Especially Nasdaq listed stocks with low market capitalizations were severely affected by this bias (Shumway and Warther (1999)). In reaction to these publications CRSP has carried out a series of projects to improve the quality of the delisting returns database. By the end of 1999, two major historical data research projects were completed: issues that delisted due to liquidation and issues that delisted because they were dropped by an exchange (CRSP (2001)). These projects have reduced the selection bias present in the database and have made the CRSP database much more reliable.[21] Because momentum (single and triple sorted) portfolios rely on older versions of the CRSP database, we treat these with more caution than the size and value sorted portfolios.

Table I gives descriptive statistics for the five sets of benchmark portfolios: ten size portfolios, ten value portfolios, ten momentum portfolios, 25 size-value portfolios (FF25) and 27 value-size-momentum portfolios (C27). Clearly, the returns of these portfolios do not obey a normal distribution. This provides an important rationale for adopting the SSD test, which account for the full return distribution rather than mean and variance alone.

[Insert Table I about here]

Figure 2 shows the five benchmark sets in mean-standard deviation space, including the individual assets (the clear dots), the market portfolio (the black square), the MV tangency portfolios (black dots), and the mean-standard deviation frontier (with and without the risk free asset). Clearly, the market portfolio is (ex post) inefficient in terms of mean-variance analysis relative to all benchmark sets. The distance from the mean-standard deviation frontier is smallest for the size portfolios and it is largest for the momentum portfolios. For example, in the momentum data set, it is possible to achieve an average monthly return of 0.33% per month (or 4.0% per annum) in excess of the market average and given the market standard deviation. If size, value, and momentum are combined the distance to the MV-frontier increases. Of course, these ex post mean-standard deviation diagrams do not reveal if the MV classification is statistically significant. Also, the diagrams are silent on return moments other than mean-variance (such as higher-order central moments and lower partial moments). The next section gives more details on the mean-variance efficiency and SSD-efficiency of the market portfolio.

[Insert Figure 2 about here]

III Results

A. Full sample results

Table II summarizes the results for the full sample from January 1933 to December 2002. The table includes a separate cell for every combination of the five data sets (size, value, momentum, FF25 and C27) and the five efficiency tests (SSD, MV, Post, TMV and GRS). Each cell includes the value of the test statistic and the associated p-value. Also, each cell includes the identity of the optimal portfolio (OP) or the benchmark portfolio with the largest pricing error.

[Insert Table II about here]

Panel A compares the results for the MV and SSD criteria. Using a significance level of ten percent, the market portfolio is highly and significantly MV inefficient relative to all benchmark sets. For example, we find a test statistic of 0.317 in the value data set. This figure can be interpreted as a maximal pricing error of 0.317% per month (or 3.8% per annum). The associated p-value is 0.032, suggesting that the market is significantly MV inefficient relative to the value portfolios. Similar results are found for the other four data sets. In other words, the three asset pricing puzzles (size, value, and momentum) are clearly present in our data sets. The strongest evidence against MV efficiency is found in the triple-sorted C27 data set that combines the three puzzles, with a maximum pricing error of 0.795 (or 9.5% per annum) and a p-value of 0.000.

In contrast to the MV results, we find that the market portfolio is SSD-efficient relative to all five benchmark sets; in all cases, the test statistic becomes insignificant. For example, for the value portfolios, the test statistic falls from 0.317 to 0.170 (or 2.0% per annum) and the p-value increases from 0.032 to 0.431. Similar results are found for the other benchmark sets. For the triple-sorted C27 data set, the maximum pricing error falls from 0.795 to 0.444 (or 5.3% per annum) and the p-value increases from 0.000 to 0.331. Combined, the MV and SSD results suggest that omitted moments (such as higher-order central moments and lower partial moments) may explain MV inefficiency of the market portfolio.

Panel B of Table II includes the three alternative tests for market portfolio efficiency (see section IE). First, the panel shows the results of the Post test for SSD efficiency, which differs from our SSD test, because it uses the restrictive null of equal means rather than the true null of SSD efficiency. Clearly, using the wrong null lowers the p-values in all data sets. In fact, for the triple-sorted C27 data set, market portfolio efficiency is rejected with more than 90 percent confidence. However, SSD efficiency can not be rejected for the other four data sets, even if we use the Post test. These results help to understand the difference between our findings and the finding in Post (2003, Section IV) that the market portfolio is SSD inefficient relative to the FF25 portfolios. First, including the pre-1963 data increases the p-value to 0.175 (was 0.031). Second, replacing the null of equal means with the null of SSD efficiency further increases the p-value further to 0.355.

Second, Panel B includes the results for the MV test after dropping the regularity condition of nonsatiation, which yields a test that is in line with the traditional definition of MV efficiency (TMV). The condition of non-satiation is binding (the MV and TMV test results are different) for the size, value and FF25 data sets, but not for the momentum and C27 data sets. Presumably, this can be explained by the broad range of the market return for the 1933-2003 period. During this period, the market return ranged from -/-23.67 to 38.17 percent. A quadratic utility function can exhibit only little curvature if it is to be monotone increasing over such a broad interval. During the more recent periods for the momentum and C27 portfolios, the market return interval is narrower (-/-23.09% to 16.05%) and the two tests coincide. The effect of dropping nonsatiation is most substantial for the size data set. In fact, the size effect is no longer significant if we drop nonsatiation; the maximum pricing error falls from 0.341 (or 4.1% per annum) to 0.222 (or 2.7% per annum) and the p-value increases from 0.049 to 0.144. By contrast, the value and momentum effects remain strong and significant.

Third, we give the results of the GRS test. Although this test differs in various ways from our TMV test, the results are fairly similar. We find slightly higher pricing errors and lower levels of significance. However, the market portfolio is efficient relative to size portfolios and inefficient relative to all other benchmark sets. It is encouraging to see that inferences about mean-variance efficiency in our study are not heavily affected by the exact test procedure.

B. Downside risk

How can we explain the differences between the MV and the SSD results? In principle, any omitted higher-order central moment (other than mean and variance) or lower partial moment could explain why the market seems MV inefficient but SSD efficient. To gain further insight in what explains SSD efficiency, it is useful to look at Figure 3. This figure shows the optimal utility function $p(x|\boldsymbol{\beta}^*)$ for the MV, TMV and SSD tests. As discussed in Section IB, this utility function is constructed from the optimal solution for the utility gradient vector ($\boldsymbol{\beta}^*$). Two results are noteworthy. First, the figure shows that under the more restrictive MV criterion, the optimal quadratic utility function for the size, value and FF25 data sets exhibits less curvature than under the less restrictive TMV criterion, so as to ensure nonsatiation for the entire sample range of market returns. Second and most important, the SSD utility functions exhibit "crash-o-phobia" or a strong aversion for downside risk. For all five data sets, the optimal utility function assigns a high weight (marginal utility) to large losses. In case of the size, momentum, FF25 and C27 data set, the utility function is very sensitive to losses that exceed about -/- 12 percent. In the value data set, the critical return level is even lower, at about -/- 20 percent. The quadratic utility function is not sufficiently flexible to allow for these patterns of downside risk aversion. Hence, it seems that downside risk may explain why the market portfolio is MV inefficient.

[Insert Figure 3 about here]

We do not intend to reduce investment risk to a single measure, because we believe that risk is a multi-dimensional concept; in fact, this is one of our reason for using the SSD criterion in the first place. Still, it is useful to quantify downside risk so as to illustrate its explanatory power for SSD efficiency. It is especially useful to analyze the contribution of the mean-variance optimal portfolio (MVOP) to the downside risk of the market portfolio. After all, the market portfolio is classified as MV inefficient because MVOP's contribution to the variance risk of the market portfolio cannot explain the high average return of MVOP. Following Bawa and Lindenberg (1977), Price, Price and Nantell (1982) and Harlow and Rao (1989), we measure the downside risk of the market portfolio by the second lower partial moment (LPM):

$$LPM_2(m,\boldsymbol{\tau}|G) \equiv \int_{\boldsymbol{x}^\mathrm{T}\boldsymbol{\tau} \leq m}(m - \boldsymbol{x}^\mathrm{T}\boldsymbol{\tau})^2 dG(\boldsymbol{x}) \tag{9}$$

This statistic measures the average squared deviation below the target rate of return m. We may measure the contribution of MVOP to the LPM of the market portfolio by the lower partial moment (LPM) market beta:

$$\beta_i(m,\boldsymbol{\tau}|G) \equiv \frac{\int\limits_{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau}\leq m}(m - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})(x_i)dG(\boldsymbol{x})}{\int\limits_{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau}\leq m}(m - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})dG(\boldsymbol{x})} \tag{10}$$

Since the true return distribution G is not known, we use the following sample equivalent:

$$\beta_i(m,\boldsymbol{\tau}|F_{\mathbf{x}}) = \frac{\frac{1}{T}\sum\limits_{t=1}^{T}\max(m - \boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau},0)(x_{i,t})}{\frac{1}{T}\sum\limits_{t=1}^{T}\max(m - \boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau},0)(\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau})} \tag{11}$$

Figure 4 shows MVOP's LPM beta ⊟ threshold levels (m) ranging from 0 to -/-25 percent for the five benchmark portfolio sets. Clearly, for all data sets, the systematic downside risk of MVOP increases as the target return is lowered. In other words, the portfolio that seems superior to the market in mean-variance terms becomes riskier during market downturns. The pattern of increasing betas during heavy stock market losses is most noticeable for the triple sorted C27 dataset.

The pattern of the LPM betas in Figure 4 is similar to the pattern of the optimal utility functions in Figure 3. The optimal utility functions tend to kink at return levels where the LPM beta of the optimal portfolio is high. For example, in the value data set, the largest increase in downside risk occurs for target returns below -/-20 percent.

The increase in LPM beta as we lower the target return is in the range of 0.08 to 0.32. At first sight, these increases may seem relatively small. However, combined with a beta premium of about 0.387 to 0.714 percent per month (or 4.6% to 8.6% per annum, see Table I), these increases yield a substantial increase in the predicted risk premium for OP.

[Insert Figure 4 about here]

C. Rolling window analysis

How robust are the findings for the sample period under consideration? It is possible that equity return distributions are conditionally normal, but unconditionally non-normal. After all, the risk profile of stocks and the risk preferences of investors change through time.[22] Tests for market portfolio efficiency could be affected by this time variation in risk and risk premia.

Therefore the market portfolio is possibly MV inefficient in the total sample, but MV efficient in the subsamples. Perhaps the SSD efficiency classification picks up this conditional pattern in the risk return relation. In addition, the degree of efficiency may

change over time. For example, Gibbons, Ross and Shanken (1989) find no full sample size effect, but reject MV efficiency during the early 1960s, late 1970s and early 1980s.

To control for structural variation in risk and risk premia, we employ a rolling window analysis. With 1-month steps, we consider all 120-month samples from January 1933 to December 2002 (721 samples in total). For every subsample, we compute the p-values of the MV and SSD tests. The TMV results coincide with the MV results in the large majority of the subsamples and are not reported separately.

[Insert Figure 5 about here]

Interestingly, we observe in Figure 5 that the market portfolio is MV efficient from the early 1930s to the late 1950s. The first empirical tests of the CAPM depended heavily on this sample period. For example, the Black, Jensen and Scholes (1972) sample period ranged from 1931 to 1967. However, starting in 1960s to the early 1990s we find serious violations of MV efficiency for all sorts based on size, value, and momentum. The celebrated Fama and French (1993) (1996) papers focus on this particularly anomalous sample period.

However, during the 1970s and the early 1980s, the MV and SSD results diverge strongly. Specifically, MV efficiency is consistently rejected, with p-values reaching levels far below ten percent. The MV-results are very similar across the various benchmarks sets with significant inefficiencies concentrated in one particular anomalous period, the period during the 1970s and the early 1980s. By contrast, SSD efficiency is not rejected relative to size, value, or momentum for mostly all subsamples. Only in a few subsamples during the 1960s for value, the SSD p-value falls slightly below the ten percent level. In all other subsamples and for all benchmark sets we cannot reject SSD efficiency. As is true for the full sample, the implicit SSD utility functions during the 1970s and the early 1980s assign a high weight to large losses. Again, this suggests that downside risk can rationalize the asset pricing puzzles.

Figure 6 further illustrates the explanatory power of downside risk. The figure hows a rolling window analysis of the standard market beta (the contribution to the variance of the market portfolio) of MVOP (the portfolio with the largest MV pricing error) and the LPM market beta. In every subsample, the threshold return (m) is set at the 2.5th percentile of the market return distribution. Hence, the LPM market beta measures the contribution to the left tail of the market return distribution. During most subperiods, the downside risk of MVOP is smaller than or equal to the standard market beta. Interestingly, during these periods, the market portfolio generally is efficient in terms of both the MV criterion and the SSD criterion. By contrast, during the 1970s and early 1980s, when the MV and SSD efficiency classifications diverge, MVOP's downside risk also sharply increases; MVOP involves substantially more downside risk than measured by the standard market beta. Again, the distinctive historical pattern of downside risk is found for all data sets. The optimal portfolio is riskier than it seems in the mean-variance framework, because its contribution to the left tail of the market

return distribution is much larger than its contribution to market variance. The SSD criterion picks up the downside risk, overlooked in the MV framework, and uses that additional source of risk to rationalize the MV inefficiency of the market portfolio. The stock market is SSD efficient.

[Insert Figure 6 about here]

IV Conclusions

1. The value-weighted stock market portfolio is MV inefficient but SSD efficient relative to benchmark portfolios formed on size, value, and momentum. The SSD criterion is especially successful in rationalizing the persistent MV inefficiencies that occur in the 1970s and the early 1980s. During this period, the mean-variance tangency portfolio has relatively high downside risk, and no other portfolio yields a significantly better trade-off between mean and downside risk than the market portfolio.

2. We may ask if SSD efficiency of the market portfolio simply reflects a lack of power due to the use of minimal assumptions. Indeed, the SSD test uses few assumptions and hence it gives only a necessary test for market portfolio efficiency. Nevertheless, we see the use of minimal assumptions as the strength rather than the weakness of our approach; SSD reduces the specification error that follows from ad hoc parameterizations, while increasing power by imposing economically meaningful regularity conditions. However, it is true that the tests will lack power in small samples. Still, we use a relatively long time-series (840 months for the size, value and FF25 portfolios) and a narrow cross-section (ten benchmark portfolios for the size, value and momentum portfolios). Our simulation study shows that the SSD test is powerful for this type of data set.

3. Presumably, the difference with the results of Post (2003), who rejects SSD efficiency, can be attributed to Type I error caused by the use in this study of the restrictive null of equal means rather than the true null of SSD efficiency and the focus on the post-1963 period. By contrast, our analysis uses the sampling distribution under the true null and also includes the pre-1963 period.

4. Our analysis assumes that return observations are serially identical and independent distributed (IID). However, there exists a wealth of evidence that the risk/return characteristics of securities show structural and cyclical variation. Notwithstanding the arguments in favour of time-variation, our main finding is that an unconditional model can explain the size, value and momentum puzzles. Implicitly we control for structural variation and find SSD-efficiency in virtually all subsamples. Still, we may ask how our findings are affected by cyclical variation in risk and risk premia. Future research could focus on answering this question.

However, there are several reasons for scepticism about the ability of conditional MV models to explain unconditional MV inefficiency. First, there is a large risk of specification error, as a conditional model has to specify how each aspect of investor preferences and the return distribution depends on the state-of-the-world. Unfortunately, economic theory gives minimal guidance about the evolution of investor preferences and the return distribution. Second, the problem of imposing the regularity conditions is very severe for conditional models; with a conditional model, we have to make sure that the utility function is well-behaved for all possible states-of-the-world. In fact, the results of many conditional asset pricing studies can be shown to reflect severe violations of the basic conditions. For example, Lettau and Ludvigson (2001) suggest that the cyclical variation of investor preferences explains the size and value puzzles. However, their results imply strong risk-seeking behaviour during economic expansion periods and imposing risk aversion can be shown to nullify their results.[23] Third, if time-variation is the cause for unconditional MV inefficiency, then it is difficult to understand why the largest violations of MV efficiency occur in the historical subperiods (see Section IIIB). Since the risk profile of firms and the risk preferences of investors are less likely to change substantially within historical subperiods than in the full sample period, we would expect less severe violations of efficiency in the subsamples if the market were conditionally MV efficient.

5. We use a simple single-period, portfolio-oriented, rational model of a competitive and frictionless capital market that is very similar to the traditional CAPM. Our explanation rests solely on a generalization of the way risk is measured in the CAPM. Of course, alternative explanations may exist for the MV inefficiencies, for example based on a multi-period, consumption-oriented model, a model with imperfect competition or market frictions, or a model where investor behave according to non-expected utility theory. It may be impossible to empirically distinguish between some of these explanations and our explanation based on downside risk. For example, liquidity effects and downside risk may be indistinguishable, because liquidity typically dries up when the largest losses occur and, in turn, liquidity dry-up may cause or amplify the losses. Similarly, the subjective overweighing of the probability of large losses will result in similar predictions as a high marginal utility for large losses. Our point is simply that a simple risk-based generalization of the CAPM suffices to rationalize the size, value, and momentum puzzles.

Appendix

Proof to Theorem 1: By construction, $\xi(\boldsymbol{\tau}, F_{\mathbf{X}})$ is bounded from above by

$$\zeta(\boldsymbol{\tau}, F_{\mathbf{X}}, u) = \min_{i \in \mathrm{I}} \left\{ \sum_{t=1}^{T} u'(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\tau})(x_{it} - \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\tau}) / T \right\}, \text{ and hence}$$

$$\Pr[\xi(\boldsymbol{\tau}, F_{\mathbf{X}}) > y \mid H_0] \leq \Pr[\zeta(\boldsymbol{\tau}, F_{\mathbf{X}}, u) > y \mid H_0], \tag{P1}$$

for all $u \in U$. Interestingly, we can derive the asymptotic sampling distribution of $\zeta(\boldsymbol{\tau}, F_{\mathbf{X}}, u)$, $u \in \Psi(\boldsymbol{\tau}, G, U)$, under H_0 from known results. Since the observations $\boldsymbol{x}_t$, $t = 1, \cdots, T$, are serially IID, the vectors $u'(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\tau})(\boldsymbol{x}_t - \boldsymbol{e} \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\tau})$, $t = 1, \cdots, T$, are also serially IID. In general, these vectors have mean $\boldsymbol{\mu}(\boldsymbol{\tau}, G, u) \equiv \int u'(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau})(\boldsymbol{x} - \boldsymbol{e} \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau}) dG(\boldsymbol{x})$ and variance-covariance matrix $\int (u'(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau})(\boldsymbol{x} - \boldsymbol{e} \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau}) - \boldsymbol{\mu}(\boldsymbol{\tau}, G, u))(u'(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau})(\boldsymbol{x} - \boldsymbol{e} \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau}) - \boldsymbol{\mu}(\boldsymbol{\tau}, G, u))^{\mathrm{T}} dG(\boldsymbol{x})$. We adhere to the statistical convention of using the least favorable distribution that maximizes the *p*-value under the null. Under H_0, $\boldsymbol{\mu}(\boldsymbol{\tau}, G, u) \leq 0$, and $\Pr[\zeta(\boldsymbol{\tau}, F_{\mathbf{X}}, u) > y \mid H_0]$ is maximal if $\boldsymbol{\mu}(\boldsymbol{\tau}, G, u) = 0$. (Note that this represents the case where all assets are included in the evaluated portfolio, that is, $\boldsymbol{\tau} > \boldsymbol{0}$; in this case, the first-order condition (1) must hold with strict equality.) Hence, for the least favorable distribution, the vectors $u'(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\tau})(\boldsymbol{x}_t - \boldsymbol{e} \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\tau})$, $t = 1, \cdots, T$, are serially IID draws with mean $\boldsymbol{0}$ and variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\tau}, G, u) \equiv \int u'(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau})^2 (\boldsymbol{x} - \boldsymbol{e} \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau})(\boldsymbol{x} - \boldsymbol{e} \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau})^{\mathrm{T}} dG(\boldsymbol{x})$ $= (\mathbf{I} - \boldsymbol{e} \boldsymbol{\tau}^{\mathrm{T}}) \boldsymbol{\Omega}(G, u)(\mathbf{I} - \boldsymbol{e} \boldsymbol{\tau}^{\mathrm{T}})^{\mathrm{T}}$. Therefore, the Lindeberg-Levy central limit theorem implies that the vector $\sum_{t=1}^{T} u'(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\tau})(\boldsymbol{x}_t - \boldsymbol{e} \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\tau}) / T$ obeys an asymptotically normal distribution with mean $\boldsymbol{0}$ and variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\tau}, G, u) / T$. Hence, asymptotically, $\zeta(\boldsymbol{\tau}, F_{\mathbf{X}}, u)$ is the largest order statistic of N random variables with a joint normal distribution, and we find

$$\Pr[\zeta(\boldsymbol{\tau}, F_{\mathbf{X}}, u) > y \mid H_0] = (1 - \int_{z \leq ye} d\Phi(z \mid \boldsymbol{0}, \boldsymbol{\Sigma}(\boldsymbol{\tau}, G, u) / T)). \tag{P2}$$

Combining (P1) and (P2), we asymptotically find

$$\Pr[\xi(\boldsymbol{\tau}, F_{\mathbf{X}}) > y \mid H_0] \leq (1 - \int_{z \leq ye} d\Phi(z \mid \boldsymbol{0}, \boldsymbol{\Sigma}(\boldsymbol{\tau}, G, u) / T)), \tag{P3}$$

for all $u \in \Psi(\boldsymbol{\tau}, G, U)$.

REFERENCES

Banz, Rolf W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.

Basu, Sanjoy, 1977, Investment performance of common stocks in relationship to their price-earnings ratios: A test of the efficient market hypothesis, *Journal of Finance* 32, 663-682.

Bawa, Vijay, and Eric B. Lindenberg, 1977, Capital market equilibrium in a mean-lower partial moment framework, *Journal of Financial Economics* 5, 189-200.

Black, Fischer, Michael C. Jensen, and Myron Scholes, 1972, The capital asset pricing model: some empirical tests, *Studies in the theory of capital markets* (Praeger).

CRSP, Center for Research in Security Prices, 2001, CRPS Delisting returns, (University of Chicago, Chicago).

Dittmar, Robert F., 2002, Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns, *Journal of Finance* 57, 369-403.

Dybvig, Phillip, and Stephen Ross, 1982, Portfolio efficient sets, *Econometrica* 50, 1525-1546.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

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

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.

Fama, Eugene F., and Kenneth R. French, 2002, The equity premium, *Journal of Finance* 57, 637-659.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121-1152.

Hadar, J., and W.R. Russell, 1969, Rules for ordering uncertain prospects, *American Economic Review* 59, 25-34.

Hanoch, Giora, and Haim Levy, 1969, The efficiency analysis of choices involving risk, *Review of Economic Studies* 36, 335-346.

Hanoch, Giora, and Haim Levy, 1970, Efficient portfolio selection with quadratic and cubic utility, *Journal of Business* 43, 181-189.

Harlow, W.V., and Ramesh K.S. Rao, 1989, Asset pricing in a generalized mean-lower partial moment framework: theory and evidence, *Journal of Financial and Quantitative Analysis* 24, 285-311.

Harvey, Campbell R., and Akhtar Siddique, 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263-1295.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jegadeesh, Narasimhan, and Sheridan Titman, 2001, Profitability of momentum strategies: An evaluation of alternative explanations, *Journal of Finance* 56, 699-720.

Kraus, Alan, and Robert H. Litzenberger, 1976, Skewness preference and the valuation of risk assets, *Journal of Finance* 31, 1085-1100.

Levy, Haim, 1969, Comment: A utility function depending on the first three moments, *Journal of Finance* 24, 715-719.

Levy, Haim, and G. Hanoch, 1970, Relative effectiveness of efficiency criteria for portfolio selection, *Journal of Financial and Quantitative Analysis* 5, 63-76.

Petkova, Ralitsa, and Lu Zhang, 2003, Is value riskier than growth?, (University of Rochester, NY).

Post, Thierry, 2003, Empirical tests for stochastic dominance efficiency, *Journal of Finance* 58, 1905-1931.

Post, Thierry, and Haim Levy, 2004, Does risk seeking drive asset prices?, *Review of Financial Studies* Forthcoming.

Price, Kelly, Barbara Price, and Timothy J. Nantell, 1982, Variance and lower partial moment measures of systematic risk: some analytical and empirical results, *Journal of Finance* 37, 843-855.

Sharpe, William, 1991, Capital asset prices with and without negative holdings, *Journal of Finance* 64, 489-509.

Shumway, Tyler, 1997, The delisting bias in CRPS data, *Journal of Finance* 54, 2361-2379.

Shumway, Tyler, and Vincent A. Warther, 1999, The delisting in CRSP's Nasdaq data and its implications for the size effect, *Journal of Finance* 54, 2361-2379.

Wang, Zhenyu, 1998, Efficiency loss and constraints on portfolio holdings, *Journal of Financial Economics* 48, 359-375.

Table I
Descriptive Statistics Data Sets

The table shows descriptive statistics for the monthly excess returns of the value-weighted CRSP all-share market portfolio and the size, value, momentum, FF25 and C27 data sets. The sample period is from January 1933 to December 2002 (T=840) for the size, value and FF25 data sets, from January 1965 to December 1998 (T=408) for the momentum data set and from July 1963 to December 1994 (T=378) for the C27 data set. Excess returns are computed from the raw return observations by subtracting the return on the one-month US Treasury bill. The size, value and FF25 data are taken from the homepage of Kenneth French, the momentum data are from Jegadeesh (2001) and the C27 data are from Carhart (1997).

		Mean	Stdev.	Skewness	Kurtosis	Min	Max
	Panel A: Single sort: Size, BM (T=840 months)						
	Market	0.714	4.937	0.156	9.18	-23.67	38.17
10 Size Portfolios	Small	1.328	9.425	2.988	29.49	-34.59	95.97
	2	1.173	8.368	2.399	28.24	-32.93	33.30
	3	1.107	7.432	1.599	21.00	-29.65	95.19
	4	1.047	7.041	1.311	16.88	-30.07	78.59
	5	1.019	6.706	1.086	15.77	-28.89	64.25
	6	0.962	6.298	0.799	13.05	-26.89	57.52
	7	0.928	6.105	0.774	13.77	-29.07	53.07
	8	0.854	5.638	0.379	10.11	-24.90	54.42
	9	0.811	5.327	0.656	12.95	-23.80	45.86
	Big	0.657	4.665	0.093	7.95	-22.99	49.10
10 Value Portfolios	Growth	0.630	5.366	0.193	7.32	-23.30	38.45
	2	0.695	5.104	-0.192	6.68	-25.19	71.70
	3	0.676	4.967	-0.196	6.87	-26.47	28.74
	4	0.723	5.365	1.210	19.30	-24.26	27.24
	5	0.843	4.982	0.752	13.89	-24.58	56.29
	6	0.898	5.276	0.621	13.40	-26.20	46.15
	7	0.912	5.778	1.424	18.55	-25.62	48.79
	8	1.077	5.910	1.074	14.79	-29.08	59.17
	9	1.136	6.943	1.470	18.83	-30.87	52.43
	Value	1.164	8.215	1.496	18.97	-45.76	62.24
	Panel B: Single sort: Momentum (T=408 months)						
	Market	0.513	4.472	-0.505	5.42	-23.09	16.05
10 Mom. Portfolios	Loser	-0.105	6.878	0.001	5.20	-28.59	28.98
	2	0.374	5.674	-0.085	5.50	-22.68	20.05
	3	0.520	5.156	-0.150	5.80	-23.01	26.82
	4	0.586	4.873	-0.334	6.47	-25.18	25.13
	5	0.606	4.709	-0.491	6.87	-25.90	23.85
	6	0.640	4.701	-0.694	6.93	-26.61	22.51
	7	0.666	4.770	-0.923	7.26	-27.80	20.72
	8	0.750	4.995	-1.041	7.30	-29.18	18.98
	9	0.860	5.391	-1.038	6.73	-30.27	17.31
	Winner	1.123	6.561	-0.882	5.52	-32.74	15.84
	Panel C: Double sort, Size/Value (T=840 months)						
	Market	0.714	4.937	0.156	9.18	-23.67	38.17

23

S	G	0.575	11.006	1.872	17.81	-42.95	99.95
S	2	0.992	9.193	2.119	22.22	-34.49	87.77
S	3	1.229	8.513	2.039	20.36	-36.52	77.98
S	4	1.426	8.017	3.125	41.40	-34.00	105.62
S	V	1.523	8.937	3.346	38.38	-34.15	105.30
2	G	0.734	7.960	0.462	8.27	-33.32	54.01
2	2	1.049	7.286	1.760	24.40	-32.37	81.58
2	3	1.170	6.974	2.300	30.12	-28.27	81.55
2	4	1.242	6.930	1.860	23.70	-28.28	71.97
2	V	1.329	7.816	1.204	14.71	-34.41	57.98
3	G	0.802	7.244	0.930	12.36	-30.03	59.99
3	2	0.989	6.240	0.424	10.79	-29.49	44.72
3	3	1.030	6.182	1.059	16.17	-28.34	56.03
3	4	1.136	6.087	1.073	15.04	-24.99	54.46
3	V	1.250	7.410	1.029	13.98	-36.04	62.32
4	G	0.731	5.961	-0.073	5.84	-26.02	32.80
4	2	0.812	5.870	0.966	17.24	-29.45	58.25
4	3	1.010	5.696	0.500	12.02	-27.13	48.35
4	4	1.003	5.997	0.694	11.10	-29.48	47.42
4	V	1.183	7.841	1.595	20.44	-34.75	75.34
B	G	0.662	5.078	0.128	7.43	-22.61	35.27
B	2	0.648	4.799	-0.033	6.94	-23.53	27.95
B	3	0.819	4.915	0.771	13.76	-22.46	46.40
B	4	0.891	5.823	1.470	19.37	-27.14	60.77
B	V	0.943	7.083	0.799	13.92	-34.62	56.24

(Left margin label: 25 Size /Value Portfolios)

Panel D: Triple sort, Value/Size/Momentum (T=378 months)

			Mean	Stdev.	Skewness	Kurtosis	Min	Max
	Market		0.387	4.399	-0.394	5.60	-23.09	16.05
G	S	L	-0.279	6.644	0.003	5.78	-31.41	30.45
G	S	M	0.349	6.070	-0.461	6.19	-31.33	27.54
G	S	W	0.933	6.755	-0.704	5.49	-33.52	19.45
G	M	L	-0.089	5.952	0.088	5.76	-26.49	27.54
G	M	M	0.143	5.341	-0.461	5.21	-26.81	18.96
G	M	W	0.895	6.033	-0.550	5.26	-29.90	20.73
G	B	L	0.105	5.218	0.178	5.21	-20.40	25.40
G	B	M	0.250	4.507	-0.161	4.88	-20.73	17.86
G	B	W	0.611	5.348	-0.298	5.03	-23.68	21.21
N	S	L	0.306	5.975	0.474	8.36	-24.81	37.51
N	S	M	0.631	4.957	-0.201	8.04	-27.05	27.02
N	S	W	1.157	6.088	-0.730	6.57	-32.60	23.17
N	M	L	0.400	5.396	0.501	6.47	-19.08	31.20
N	M	M	0.513	4.498	-0.344	7.72	-25.68	23.05
N	M	W	0.772	5.325	-0.896	6.68	-30.54	17.52
N	B	L	0.438	4.801	0.477	5.29	-19.66	21.74
N	B	M	0.367	4.211	0.158	5.31	-15.57	21.03
N	B	W	0.595	4.856	-0.347	5.56	-24.00	19.95
V	S	L	0.552	6.440	0.974	10.77	-28.26	45.34
V	S	M	1.062	5.583	0.199	9.22	-29.09	33.99
V	S	W	1.341	6.246	-0.289	7.56	-32.00	30.20
V	M	L	0.582	6.078	0.469	7.68	-25.79	38.07
V	M	M	0.909	5.303	-0.016	9.04	-28.30	31.03
V	M	W	1.287	5.881	-0.806	8.23	-33.86	25.76
V	B	L	0.600	5.569	0.718	7.41	-16.99	35.34
V	B	M	0.589	4.752	-0.053	5.72	-23.47	20.41
V	B	W	0.923	5.352	-0.254	5.85	-24.78	22.84

(Left margin label: 27 Value/Size/Momentum Portfolios)

Table II
Efficiency of the stock market portfolio

We test if the CRSP all-share index is SSD and MV efficient in the size, value, momentum, FF25 and C27 data sets. Results are shown for the full samples. Each cell contains the test statistic (Theta), the associated p-value, and the identity of the optimal portfolio (OP) or the portfolio with the largest pricing error. Cells are coloured grey if the p-value falls below 10% and efficiency is rejected with at least 90% confidence. Panel A includes our main results and compares the results of our SSD and MV tests. Panel B shows the outcomes of three alternative tests: the original Post (2003) test, the traditional mean variance (TMV) test (without the restriction of non-satiation), and finally the Gibbons, Ross, Shanken (1989) test.

	N	T		Panel A: SSD versus MV		Panel B: Alternative tests		
				SSD	MV	Post	TMV	GRS
Size	10	840	Theta	0.056	0.341	0.184	0.222	0.314
			p-value	0.647	0.049	0.231	0.144	0.391
			OP	Small	Small	Small	Small	Small
Value	10	840	Theta	0.170	0.317	0.211	0.291	0.312
			p-value	0.431	0.032	0.135	0.041	0.022
			OP	8	8	8	8	8
Mom.	10	408	Theta	0.206	0.349	0.215	0.349	0.455
			p-value	0.355	0.051	0.175	0.051	0.000
			OP	Winner	Winner	Winner	Winner	Winner
FF25	25	840	Theta	0.230	0.579	0.363	0.411	0.556
			p-value	0.446	0.014	0.140	0.075	0.000
			OP	14	SV	SV	SV	SV
C27	27	378	Theta	0.444	0.795	0.469	0.795	0.902
			p-value	0.311	0.000	0.048	0.000	0.000
			OP	VSW	VSW	VSW	VSW	VSW

Figure 1

Statistical properties of the SSD test procedure

The figure displays the size and power for various numbers of time-series observations (T) and for a significance level (a) of 2.5, five and ten percent. The results are based on 10,000 random samples from a multivariate normal distribution with joint moments equal to the sample moments of the monthly excess returns of the ten B/M portfolios and the U.S. Treasury bill for the period from July 1963 to October 2001. Size is measured as the relative frequency of random samples in which the efficient tangency portfolio (TP) is wrongly classified as inefficient. Power is measured as the relative frequency of random samples in which the inefficient equally weighted portfolio (EP) is correctly classified as inefficient.

26
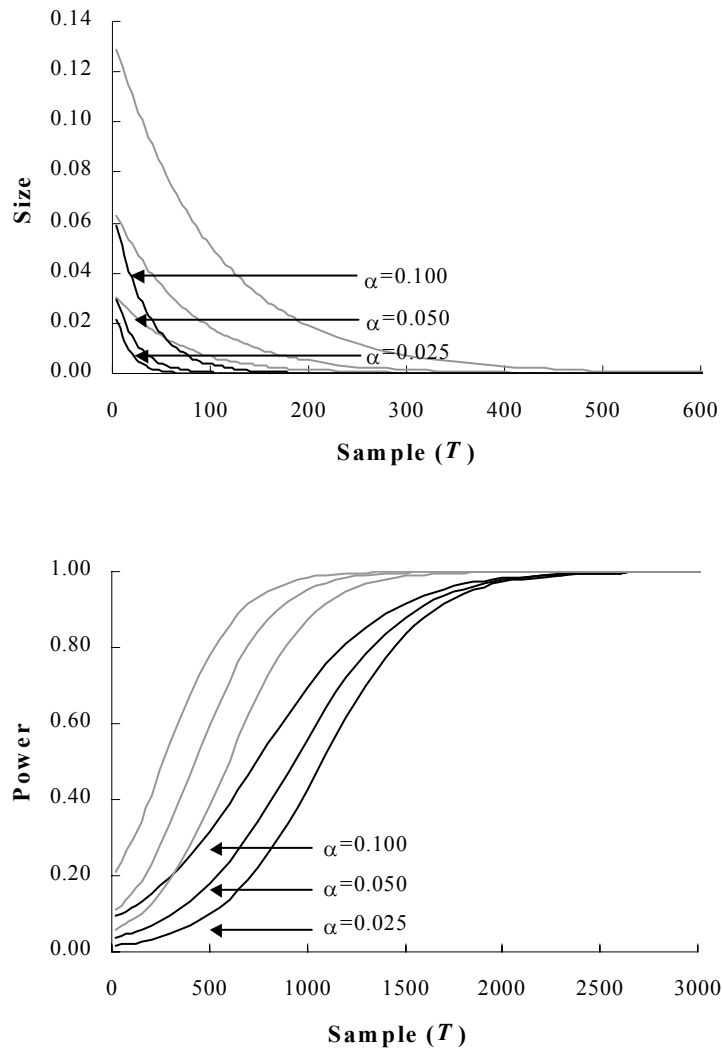
Figure 2
Mean-standard deviation diagrams

This figure shows the mean-standard deviation diagram for the size, value, momentum, FF25 and C27 data sets. We show the mean excess return and the standard deviation of the individual benchmark portfolios (clear dots) and the efficient frontier with and without the risk-free asset. The market portfolio is labelled "M" and the mean-variance tangency portfolio "TP".
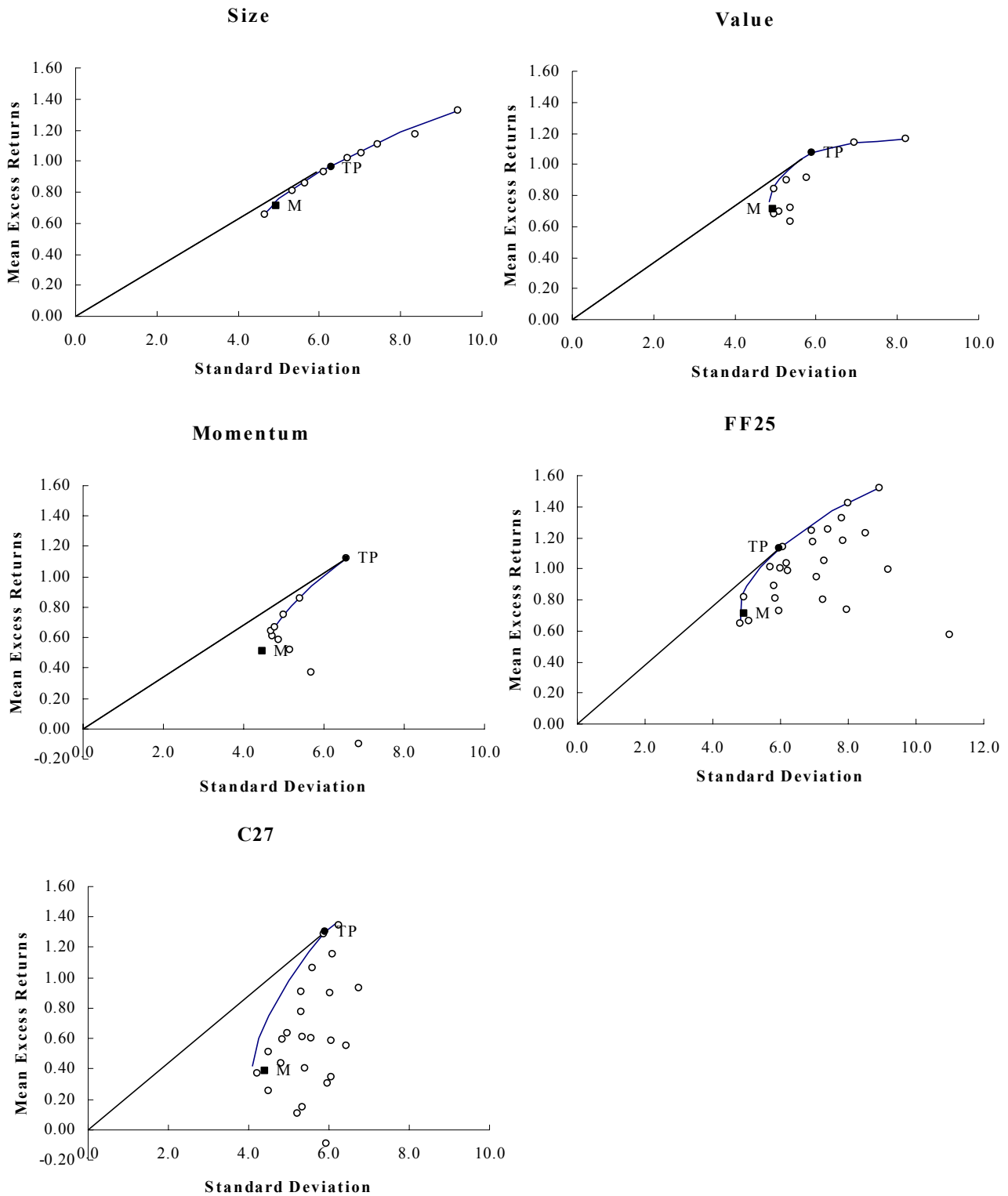
Figure 3
Optimal utility functions

This figure shows the optimal utility functions $p(x|\boldsymbol{\beta}^*)$ for the MV, TMV and SSD tests and the size, value, momentum, FF25 and C27 data sets. The black line shows the utility function under the SSD criterion, the dark grey line shows the quadratic utility function under the TMV criterion, while the grey line shows the quadratic utility function under the MV criterion (with nonsatiation over the sample range of market return). The utility functions are standardized such that $p(0|\boldsymbol{\beta}^*) = 0$ and $\sum_{t=1}^{T} p'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau}|\boldsymbol{\beta}^*)/T = 1$.

Size

Value

Momentum

FF25

C27

28
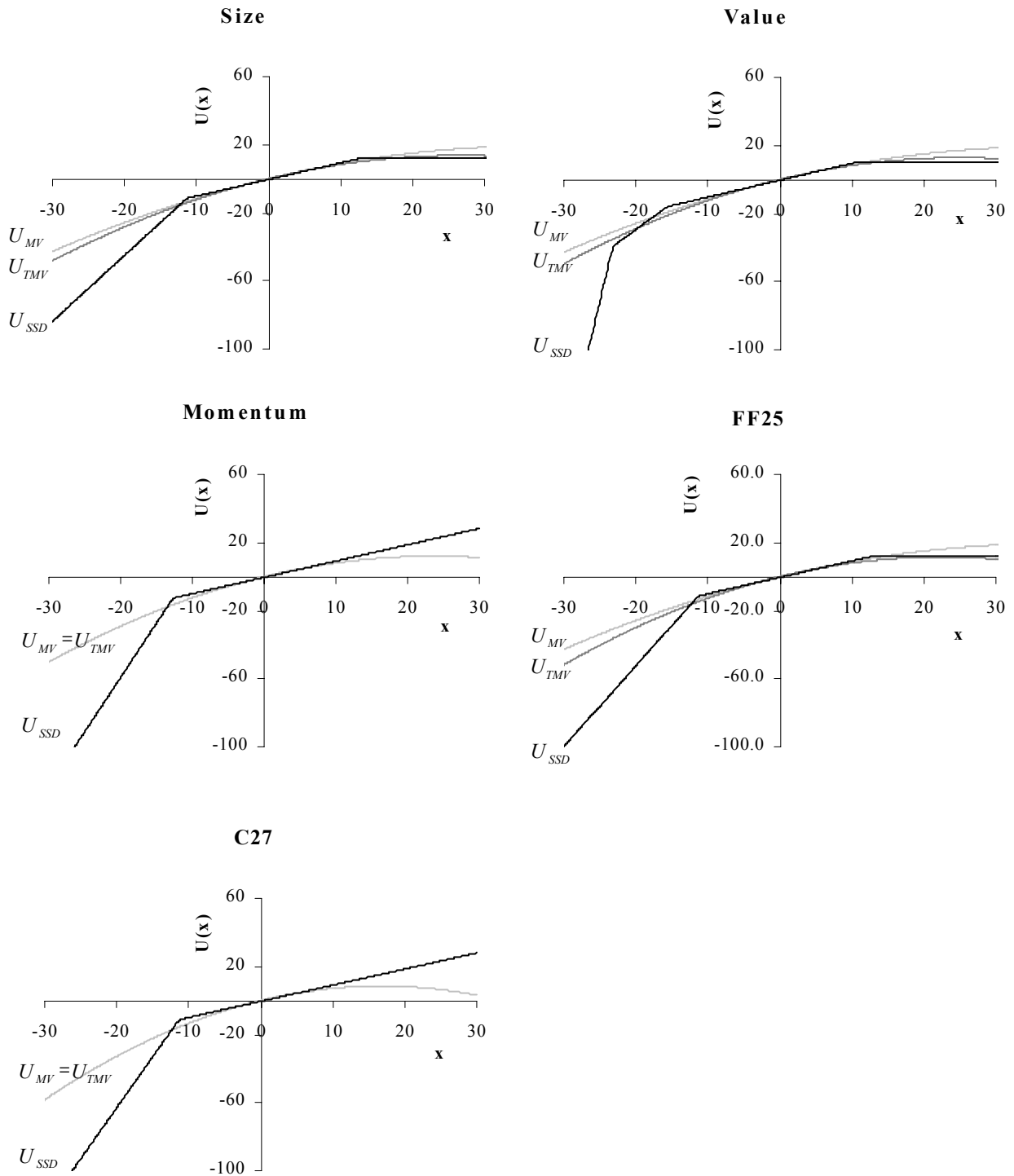
Figure 4
Downside risk of MV-optimal portfolio for different threshold values

This figure shows the lower partial moment beta $\beta_i\left(m, \tau | F_{\mathbf{x}}\right)$ of MVOP (the portfolio with the maximal MV pricing error) for different threshold levels (m) and for each of the five data sets. The downside risk of MVOP increases as the threshold return is lowered. Hence, the systematic risk of the seemingly attractive portfolios (Small/Value/Winner) increases during market crashes.
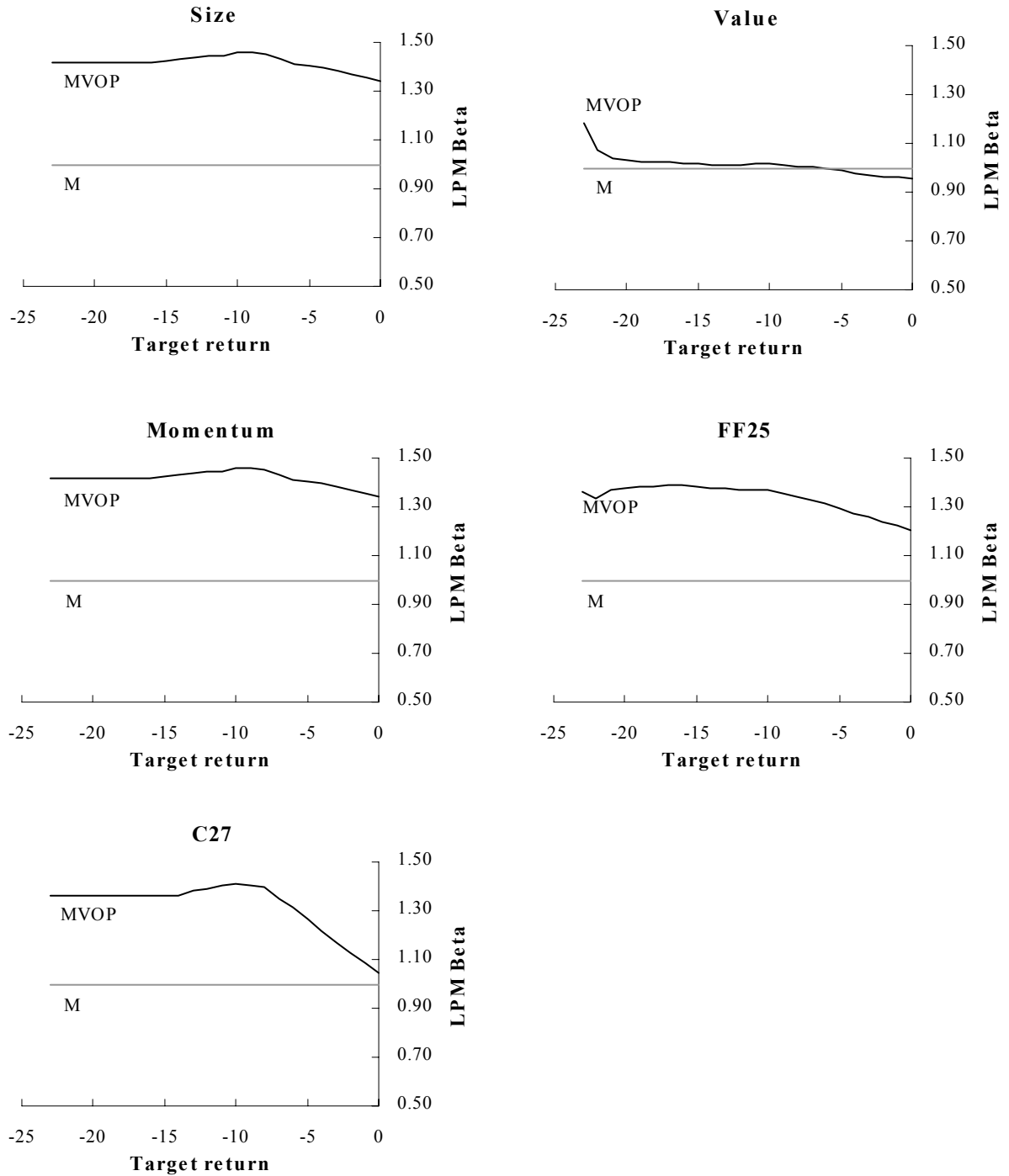

29

Figure 5

Rolling window analysis of efficiency tests

This figure shows the p-values for the MV and SSD efficiency tests in each of the five data sets using a rolling 120-months period (1-month steps). The grey line represents the p-value of the MV test and the dark line shows the SSD p-value. The figure reads as follows: consider the 1980 observation in the value data set, which represents the 120-month period starting in January 1975 to December 1984. For this period, the stock market is MV inefficient (p=0.004) but SSD efficient (p=0.377). During the 1970s and early 1980s, MV efficiency can be rejected relative to all data sets (p<0.05), while SSD efficiency cannot be rejected (p>0.2).
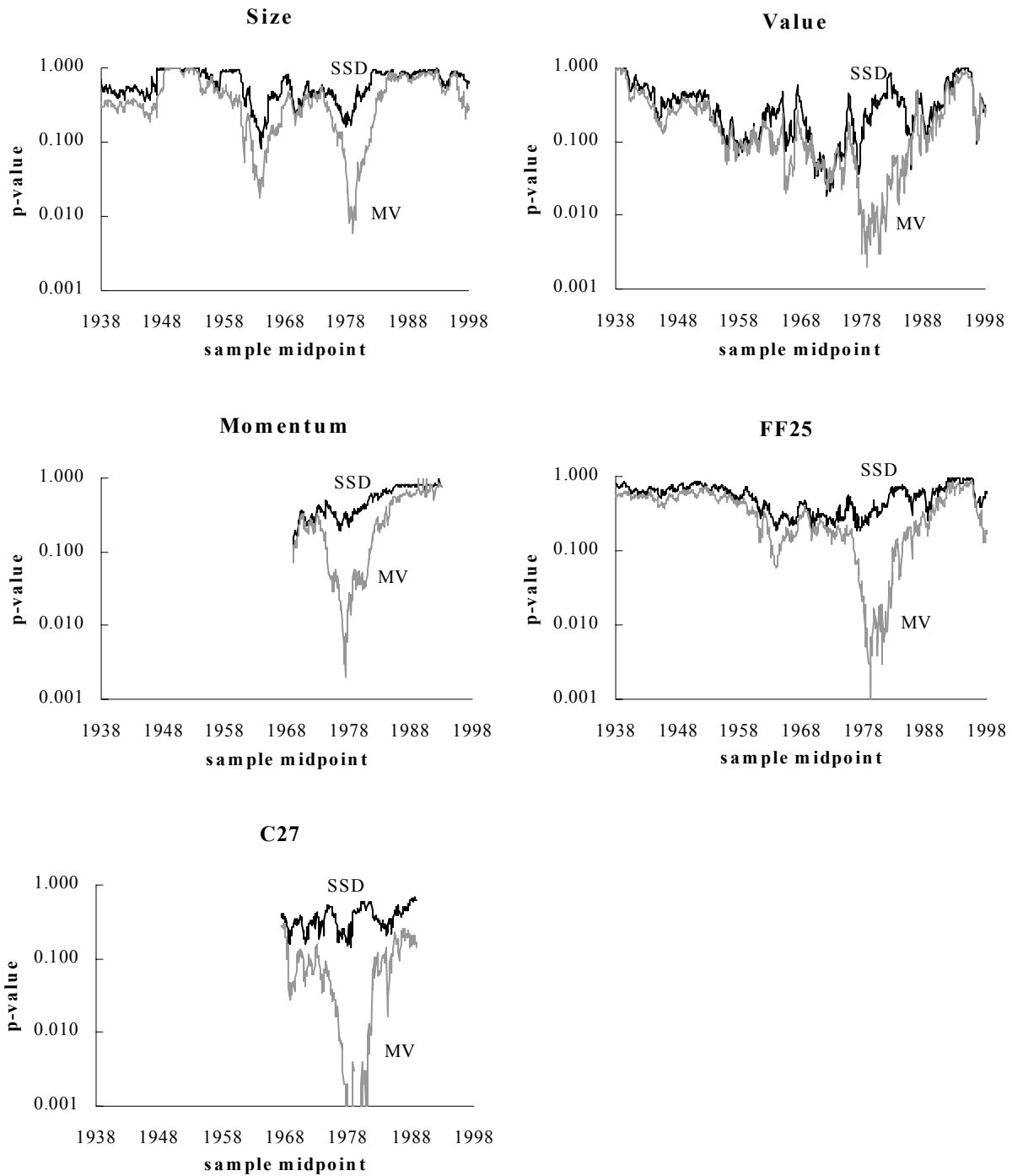
Figure 6

LPM beta of tangency portfolio

This figure shows the lower partial moment (LPM) beta $\beta_i(m, \tau | F_{\mathbf{x}})$ of MVOP (the portfolio with the maximal MV pricing error) compared to the traditional MV beta, using a rolling 120-months period (1-month steps). The rolling LPM betas of the optimal portfolio coincide with the standard MV betas during most subsamples. During the most sample periods, MV beta is higher than the LPM beta. Interestingly, during these periods, the market portfolio is MV-efficient. However, during the anomalous 1970s and early 1980s downside risk of the optimal portfolio is higher than co-variance risk.
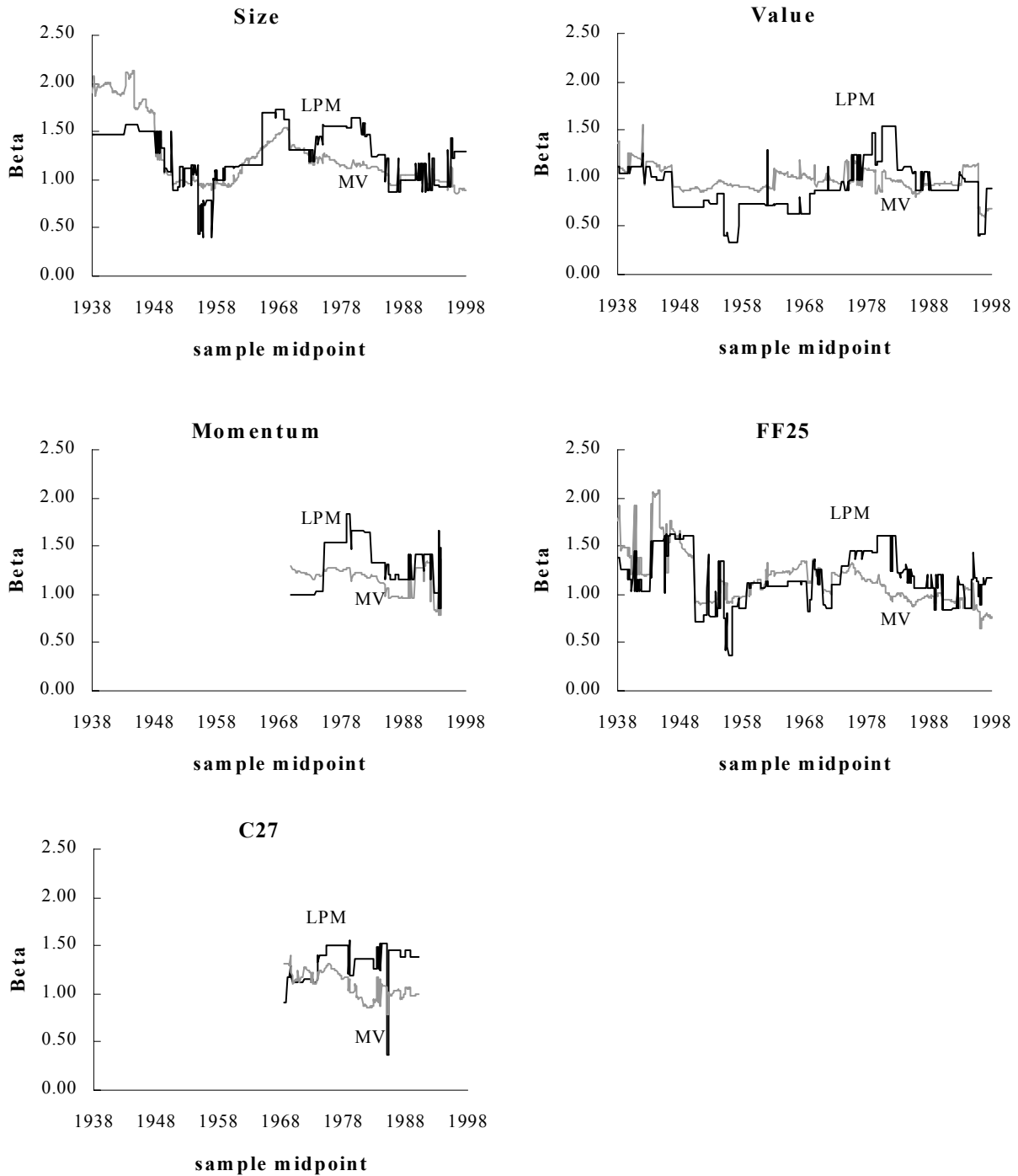
[1] Less restrictive assumptions are obtained if we do restrict the shape of the return distribution; see, for example, Berk (1997).

[2] For a non-concave utility function the first-order condition is a necessary but nut sufficient condition for establishing a global maximum for expected utility; the first-order condition also applies for possible local optima and a possible global minimum. This important result is sometimes ignored when using non-concave utility functions.

[3] Market portfolio efficiency does not follow directly from our maintained assumptions about investor preferences and beliefs. In fact, Dybvig and Ross (1982) demonstrate that the SSD efficient set generally is not convex and hence the market portfolio may be inefficient.

[4] Of course, we could directly analyze the efficiency of actual funds. Nevertheless, for the sake of data availability and comparability, we focus on the CRSP all-share index, which is used in many comparable studies.

[5] Throughout the text, we will use $\Re^N$ for an N-dimensional Euclidean space, and $\Re_+^N$ denotes the positive orthant. Further, to distinguish between vectors and scalars, we use a bold font for vectors and a regular font for scalars. Finally, all vectors are column vectors and we use $\boldsymbol{x}^\mathrm{T}$ for the transpose of $\boldsymbol{x}$.

[6] Post (2003) does not assume that the utility function is continuously differentiable, so as to allow for, for instance, piecewise linear utility functions. However, in practice, we typically cannot distinguish between a kinked utility function and a smooth utility function with rapidly changing marginal utility. Nevertheless, using subdifferential calculus, we may obtain exactly the same characterization of the sampling distribution if utility is not continuously differentiable.

[7] Post (2003) requires utility to be strictly increasing. To remain consistent with the original definition of SSD, we require a weakly increasing utility function. This is one of our reasons for adopting a novel standardization for the gradient vector; see Section IB.

[8] Of course, if marginal utility is decreasing, then it can be positive only over a bounded interval. In our analysis, we require marginal utility to be positive and decreasing over the sample interval of returns on the market portfolio.

[9] Nevertheless, we may generalize our analysis to include (bounded) short selling. The SSD test is based on the first-order optimality conditions for optimizing a concave objective function over a convex set. The analysis can be extended to a general polyhedral portfolio possibilities set. We basically have to check whether there exists an increasing hyperplane that supports the extreme points of the portfolio possibilities set. One approach is to enumerate all extreme points and to include all extreme points as virtual assets.

[10] This equality is a variation to the well-known Euler equation $\int u'(\boldsymbol{x}^\mathrm{T}\boldsymbol{\tau})x_i dG(\boldsymbol{x}) = c \quad \forall i \in \mathrm{I}, c \geq 0$.

[11] Our focus on the maximum error reflects the short sales restriction. A negative 'pricing error' for a given asset does not constitute a violation of the first-order condition if the asset is not included in the evaluated portfolio ($\tau_i = 0$), as the investor can then improve the evaluated portfolio only by short selling the asset (which is not allowed) and not by reducing the weight of the asset in the portfolio. By contrast, a positive error is always problematic, because an investor can then improve the evaluated portfolio by increasing the weight of the asset and decreasing the weight of the other assets included in the portfolio. Of course, we generally do not know the number or the identity of the positive pricing errors in advance. However, the maximum pricing error is always positive.

[12] This LP problem can be solved with minimal computational burden, even with spreadsheet software run on a desktop computer. Nevertheless, for applications where the number of time-series observations (T) is very large (for example, thousands of observations), it is useful to use a simplified formulation. A simplified version can be obtained by using $\boldsymbol{b} \equiv (b_1 \cdots b_T)^\mathrm{T}$, with $b_s \equiv \beta_s - \beta_{s+1}$, $s = 1, \cdots, T-1$ and $b_T \equiv \beta_T$. Specifically, substituting $\beta_t = \sum_{s=1}^{T} b_s$ in Equation (6) and rearranging terms yields

$$\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U_{SSD}) = \min_{\boldsymbol{b} \in \Re_+^T, \theta} \left\{ \theta : \sum_{s=1}^{T} b_s \sum_{t=1}^{s} (\boldsymbol{x}_t^\mathrm{T}\boldsymbol{\tau} - x_{it})/T + \theta \geq 0 \quad \forall i \in \mathrm{I} \right\}. \tag{6'}$$

While Equation (6) involves T variables and $N+T$-1 constraints, Equation (6') involves T variables and only N constraints, which yields a large reduction in computational burden. We have effectively removed the T-1 restrictions $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_T$, which are now satisfied by construction, as $\beta_t = \sum_{s=1}^{T} b_s$ and $b_t \geq 0 \; \forall t \in \Theta$ imply $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_T$. This simplification is similar to the one used by Post (2003, Proof to Thm 2) to arrive at his simplified dual test statistic.

[13] Also, we will derive the p-value under the null of efficiency rather than the null of equal means.

[14] In fact, the additional restrictions simplify the problem, because there are now only two unknown variables. Substituting $\beta_t = a + b\boldsymbol{x}_t^\mathrm{T}\boldsymbol{\tau}$ in equation (6) and rearranging terms, we find $\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U_{MV}) =$

$$\min_{a,b,\theta}\left\{\theta : \sum_{t=1}^{T}(a+b\boldsymbol{x}_t^\mathrm{T}\boldsymbol{\tau})(\boldsymbol{x}_t^\mathrm{T}\boldsymbol{\tau}-x_{it})/T+\theta \geq 0 \quad \forall i \in \mathrm{I}; a+b\boldsymbol{x}_t^\mathrm{T}\boldsymbol{\tau} \geq 0 \quad \forall t \in \Theta; \sum_{t=1}^{T}(a+b\boldsymbol{x}_t^\mathrm{T}\boldsymbol{\tau})/T = 1; b \leq 0\right\}.$$ The full

utility function is simply $p(x|\boldsymbol{\beta}^*) = a^*x + \frac{1}{2}b^*x^2$, with (a^*, b^*) for the optimal solution.

[15] The power depends on the degree of inefficiency of the evaluated portfolio. We selected the set of ten B/M portfolios, because the degree of inefficiency of the equal weighted portfolio is 'medium'; it is higher than for the ten size portfolios, but lower than for the momentum portfolios.

[16] It is possible to achieve a substantially higher mean given the standard deviation of EP and hence EP is MV inefficient. Since we assume a normal distribution in the simulations, the SSD criterion coincides with the MV criterion and EP is also SSD inefficient.

[17] The tangency portfolio consists of 18.22%, 2.04% and 79.74% invested in the fifth, sixth and eight B/M portfolios respectively.

[18] This means that we no longer require the gradient vector to be non-negative. It is straightforward to verify that this yields a necessary and sufficient test for mean-variance efficiency in the traditional definition but with short sales excluded.

[19] We thank Jegadeesh Titman for sharing his data with us. We also have price momentum data for a longer period, but this benchmark set is not publicly available. To make our results comparable and replicable we confine ourselves to easily accessible and widely known databases.

[20] We thank Mark Carhart for generously providing us with these data.

[21] For example, consider the Fama and French (1996) sample period ranging from July 1963 to December 1993 (366 months). In FF Table I the average monthly excess return on the BV (SG) portfolio was 0.71 (0.31), but with the new 2002 database (employed in this study) this number is downwards revised to 0.59 (0.25).

[22] For example, Fama and French (2002) argue that structural economic changes have lowered the expected equity premium during the last four decades. Also, we know that the risk of stocks has changed along, for example the betas of value stocks have structurally declined in the long run (for example see: Petkova and Zhang (2003)).

[23] It can also be shown that the market is significantly MV inefficient and that residual size and value effects exist in the Lettau-Ludvigson study, even if we do not require risk aversion. Result available upon request from the authors.

Publications in the Report Series Research* in Management

ERIM Research Program: "Finance and Accounting"

2004

Corporate Finance In Europe Confronting Theory With Practice
Dirk Brounen, Abe de Jong and Kees Koedijk
ERS-2004-002-F&A
http://hdl.handle.net/1765/1111

Downside Risk And Asset Pricing
Thierry Post and Pim van Vliet
ERS-2004-018-F&A

An Alternative Decomposition Of The Fisher Index
Winfried G. Hallerbach
ERS-2004-022-F&A
http://hdl.handle.net/1765/1220

The effects of systemic crises when investors can be crisis ignorant
Erik Kole, Kees Koedijk & Marno Verbeek
ERS-2004-027-F&A
http://hdl.handle.net/1765/1270

The Eco-Efficiency Premium Puzzle
Jeroen Derwall, Nadja Günster, Rob Bauer and Kees Koedijk
ERS-2004-043-F&A
http://hdl.handle.net/1765/1296

Shareholders' Voting at General Meetings: Evidence from the Netherlands
Abe De Jong, Gerard Mertens And Peter Roosenboom
ERS-2004-039-F&A
http://hdl.handle.net/1765/1332

Do banks influence the capital structure choices of firms?
Petra Daniševská, Abe de Jong & Marno Verbeek
ERS-2004-040-F&A
http://hdl.handle.net/1765/1333