



Nonparametric portfolio efficiency measurement with higher moments

Jens J. Krüger¹

Received: 24 October 2019 / Accepted: 18 July 2020 / Published online: 1 August 2020
© The Author(s) 2020

Abstract

The paper considers a nonparametric approach to determine portfolio efficiency using specific directions toward the portfolio frontier function. This approach allows for a straightforward incorporation of higher moments of the returns distribution beyond mean and variance. The nonparametric approach is extended by the computation of optimal directions endogenously by maximizing the distance toward the portfolio frontier as a novel methodological feature. An empirical application to Fama–French portfolios demonstrates the applicability of the nonparametric approach. The results show that the optimal directions to the frontier depend on the portfolio considered as well as on the period for which the moments are estimated. Skewness in particular plays a role in determining the optimal direction, whereas kurtosis seems to be less crucial.

Keywords Finance · Portfolio choice · Directional distance functions · Skewness and kurtosis

JEL Classification G11 · C14

1 Introduction

Portfolio selection in the legacy of Markowitz (1952, 1959) is traditionally concerned with finding a compromise between a high return and low risk. This setting can be criticized, however, and in particular, it neglects the influence of higher moments of the returns distribution beyond mean and variance on the utility of an investor.¹ The consideration of those higher moments also has a long tradition in finance with Samuelson

¹ See, e.g., Hens and Rieger (2010, p. 48ff.) for a discussion of the limitations of mean–variance utility.

✉ Jens J. Krüger
krueger@vwl.tu-darmstadt.de

¹ Department of Law and Economics, Technical University of Darmstadt, Hochschulstr. 1, 64289 Darmstadt, Germany

(1970) being one of the first contributions. There is a steady stream of articles following with a special focus on the third moment, skewness, but also considering the fourth moment, kurtosis (see, e.g., Kraus and Litzenberger 1976; Scott and Horvath 1980; Harvey and Siddique 2000; Jondeau and Rockinger 2006; Harvey et al. 2010 and the numerous references cited therein).

Assuming a specific utility function is restrictive and an approach which is workable without such an assumption would be valuable. One such approach applies methods originating from nonparametric productivity analysis, mostly based on variants of data envelopment analysis (DEA). DEA is usually applied to determine productive efficiency in general input–output relationships. In finance applications, the mean return is usually taken as the output variable, whereas measures of risk or transaction cost components are treated as inputs. This follows from the simple logic that the output should be maximized, whereas the inputs are to be minimized. There are numerous published articles reporting empirical analyses along these lines. See Murthi et al. (1997), Morey and Morey (1999) and Edirisinghe and Zhang (2007) as representative examples and Glawischnig and Sommersguter-Reichmann (2010) for a succinct summary.

Lamb and Tee (2012) point to various problems with applications of DEA for portfolio performance evaluation. One of these problems is the inability of most DEA analyses to take account of the effects of diversification. Better adapted to the task of portfolio performance evaluation is the nonparametric methodological approach put forward by Briec and Kerstens with their various co-authors based on the concept of the shortage function (Luenberger 1992, 1995) which is also known as the directional distance function (DDF) of Chambers et al. (1996). In contrast to DEA analyses which rely on linear programming solvers for their computation, the Briec–Kerstens approach leads to nonlinear constraints and requires the application of numerical methods for nonlinear optimization.

As mentioned above, higher moments of returns such as skewness and kurtosis also play a role for portfolio choice beyond mean and variance as additional sources of risk. Conventional parametric and Bayesian approaches to the analysis of portfolio selection under higher moments are explained in Jondeau and Rockinger (2006) and Harvey et al. (2010), respectively. The Briec–Kerstens approach also allows for a straightforward extension to incorporate higher moments (skewness, kurtosis) in addition to mean and variance into the nonparametric approach. This is developed by Briec et al. (2004) for the mean–variance analysis and extended by Briec et al. (2004) to a mean–variance–skewness analysis. Briec and Kerstens (2010) provide a generalization of the previous papers to moments of arbitrary order including partial moments. By this development, the mean–variance–skewness–kurtosis case is also covered. This case is also covered in the related analysis of Boudt et al. (2020a) who also apply shrinkage methods for the estimation of higher moments as we do subsequently.

In this paper, we follow the Briec–Kerstens approach to analyze portfolios constructed by Fama and French for various selected periods. The primary advantage of this approach is that there is no need to assume specific functional forms, neither for the frontier function nor for the distribution of the returns. We are able to determine appropriate benchmarks for evaluating the portfolios without having to rely on a specific theoretical model. In addition, the approach permits a straightforward incor-

poration of higher moments (of course at the expense of increasing computational complexity). We introduce the novel methodological feature of an endogenous determination of the direction toward the frontier function in the moment space for this type of methodological approach.

The results obtained with this approach show that the more efficient portfolios tend to be composed of small (rather than large) stocks, of high (rather than low) book-to-market value stocks, of stocks with high (rather than low) operating profitability and of stocks with low (rather than high) investment. There are major differences of the distance to the frontier and the optimal direction along which this distance is measured across the subperiods considered. In addition to the roles of mean and variance, skewness tends to play a larger role compared to kurtosis.

The paper proceeds with a discussion of the influence of higher moments on the utility derived from the final wealth of an investment in Sect. 2. This is followed by an outline of the nonparametric approach to compute portfolio efficiency with its extensions in Sect. 3. The results obtained with this approach for a selection of univariate sorts of portfolios and various subperiods are presented and discussed in Sect. 5. Beforehand, in Sect. 4 the data are explained in detail and statistical description is provided. The results are summarized, and conclusions are drawn in the final Sect. 6.

2 Higher moments and utility

This section relates the higher moments of the returns distribution to the utility of the investor holding a particular portfolio of assets (which may be stocks, bonds, funds or portfolios). The exposition is mainly linked to Jondeau and Rockinger (2006), but here, it is more strongly focused on the central utility motives of non-satiation, risk aversion, prudence and temperance corresponding to the first four moments of the returns distribution. For expositional convenience, we consider an investor holding a portfolio of n risky assets with returns collected in the vector $\mathbf{R} = (R_1, \dots, R_n)'$. The returns are characterized by a joint distribution $F(\mathbf{R})$. At the end of a particular period, wealth is generated by $W = 1 + \mathbf{x}'\mathbf{R}$ where $\mathbf{x} = (x_1, \dots, x_n)'$ denotes a vector of nonnegative portfolio weights summing to unity.

The utility derived from this wealth is given by the utility function $u(W)$. By a Taylor approximation up to fourth order around the expected wealth $E(W)$, we get

$$\begin{aligned}
 u(W) \approx & u(E(W)) + (W - E(W)) \cdot u'(E(W)) + \frac{(W - E(W))^2}{2!} \cdot u''(E(W)) \\
 & + \frac{(W - E(W))^3}{3!} \cdot u'''(E(W)) + \frac{(W - E(W))^4}{4!} \cdot u''''(E(W)) \quad (1)
 \end{aligned}$$

for the utility function. The appropriate order of the Taylor approximation is not clear, and adding more terms may even be harmful for the approximation quality under certain circumstances (see Hlawitschka 1994). Dittmar (2002) provides evidence in favor of a cubic pricing kernel which implies preferences over the first four moments of the returns distribution.

According to Arrow (1971), investors generally have utility functions with a positive marginal utility of wealth ($u'(\cdot) > 0$, implying non-satiation) decreasing marginal utility of wealth ($u''(\cdot) < 0$, representing risk aversion) and non-increasing absolute risk aversion. The latter property is associated with $u'''(\cdot) > 0$ and implies prudence as a necessary condition. A prudent investor values precautionary savings as a protection against loss. A further characteristic of the utility function is so-called temperance associated with its fourth derivative $u''''(\cdot) < 0$ (Gollier and Pratt 1996).

Taking expectations, we get

$$E(u(W)) \approx u(E(W)) + \frac{1}{2}\text{Var}(W) \cdot u''(E(W)) + \frac{1}{6}\text{Skew}(W) \cdot u'''(E(W)) + \frac{1}{24}\text{Kurt}(W) \cdot u''''(E(W)) \quad (2)$$

for expected utility. We see that expected utility is driven by the variance $\text{Var}(W) = E((W - E(W))^2)$ representing risk and reducing expected utility for a risk-averse investor. Skewness $\text{Skew}(W) = E((W - E(W))^3)$ is larger than zero if the probability of large positive returns is larger than the probability of negative returns of the same size (right-skewed distribution) and smaller than zero if the distribution of W is left-skewed with a higher probability of below-mean returns. Stock returns are frequently negatively skewed since very large losses are more likely than very large gains. The importance of skewness has already been discussed by Kraus and Litzenberger (1976). They point out that expected utility increases with skewness for an investor with non-increasing absolute risk aversion. Investors thus tend to have preference for positive skewness. A larger value for kurtosis $\text{Kurt}(W) = E((W - E(W))^4)$ indicates a higher probability for extreme market events associated with large gains or losses, therefore representing tail heaviness.² In the subsequent analysis, we will consider the three relevant cases with a second, third and fourth order separately.

3 Computing portfolio efficiency

Starting point of this nonparametric approach to portfolio analysis was Morey and Morey (1999) who were the first to integrate the optimization of the portfolio weights into a nonparametric efficiency analysis framework. Their quadratic optimization programs are specified either to maximize mean return for the actual variance of an asset or to minimize the variance for a given return. This mean–variance (MV) analysis has been extended to a simultaneous enhancement of mean return and reduction in variance by Brier et al. (2004). They use the device of the Luenberger (1992, 1995) shortage function, also used as directional distance function (DDF) by Chambers et al. (1996) for the formulation of the respective optimization problems. This framework

² Usually, this is negatively valued by investors and is associated with the temperance motive discussed above. As a general rule derived by Scott and Horvath (1980), investors with consistent preferences have a preference for large odd moments (mean, skewness) and small even moments (variance, kurtosis). Concerning the weight of the moments, Scott and Horvath argue that kurtosis is of less importance since reducing variance should also be instrumental for reducing kurtosis.

has been extended by Briec et al. (2004) to a mean–variance–skewness (MVS) analysis and to a general moment space (including partial moments) by Briec and Kerstens (2010). This also covers the mean–variance–skewness–kurtosis (MVSK) analysis.³

3.1 Portfolio moments

The treatment of portfolio choice is based on a vector of returns for a cross section of n assets $R_i, i = 1, \dots, n$, stacked to a vector $\mathbf{R} = (R_1, \dots, R_n)'$ with expectation $\boldsymbol{\mu} = E(\mathbf{R})$ and a covariance matrix $\mathbf{V} = \text{Var}(\mathbf{R}) = E(\mathbf{r}\mathbf{r}')$ with $\mathbf{r} = \mathbf{R} - \boldsymbol{\mu}$.⁴

In addition to these first two moments, skewness and kurtosis may also be relevant for portfolio selection, as explained above. For a scalar random variable Y , skewness is defined as $s = E((Y - \mu_Y)^3)$ with $\mu_Y = E(Y)$. Positive (negative) skewness $s > 0$ ($s < 0$) indicates a right-skewed (left-skewed) distribution with a higher probability of large positive (negative) returns. In the case of a symmetric (about the mean) distribution, we have $s = 0$. Scalar skewness can be estimated by $\hat{s} = T^{-1} \sum_{t=1}^T (y_t - \bar{y})^3$ for a time-series sample of observations $y_t, t = 1, \dots, T$, and $\bar{y} = T^{-1} \sum_{t=1}^T y_t$. Scalar kurtosis is likewise defined as $k = E((Y - \mu_Y)^4)$ and can be estimated by $\hat{k} = T^{-1} \sum_{t=1}^T (y_t - \bar{y})^4$. It measures the extent of fat tails of a distribution with $k > 3\sigma^4$ for fat-tailed distributions relative to the benchmark case of a normal distribution. Fat tails are associated with a larger probability of extreme returns.

To deal with portfolio skewness, we need the $n \times n^2$ co-skewness matrix $\mathbf{S} = E(\mathbf{r}\mathbf{r}' \otimes \mathbf{r}')$ with elements of the type $E(r_i r_j r_h)$.⁵ This comprises not only the skewness of asset i , i.e., $E(r_i^3)$, but also co-skewness elements of the sort $E(r_1 r_3^2)$ or $E(r_1 r_2 r_3)$, for example. Likewise, we define the $n \times n^3$ co-kurtosis matrix $\mathbf{K} = E(\mathbf{r}\mathbf{r}' \otimes \mathbf{r}' \otimes \mathbf{r}')$ with typical element $E(r_i r_j r_h r_l)$. An example for the specific arrangement of the co-skewness and co-kurtosis terms in the matrices \mathbf{S} and \mathbf{K} can be found in Jondeau and Rockinger (2006, p. 36) and is defined by the form of indexing used in (3).

These moments can be estimated by using a data sample for a vector of returns $\mathbf{R}_t = (R_{1t}, \dots, R_{nt})', t = 1, \dots, T$. Defining $\mathbf{r}_t = (r_{1t}, \dots, r_{nt})'$ correspondingly with $r_{it} = R_{it} - \hat{m}_i, t = 1, \dots, T$, we estimate the mean vector and the moment matrices (listed jointly with their typical elements) by

$$\begin{aligned} \hat{\mathbf{m}} &= T^{-1} \sum_{t=1}^T \mathbf{R}_t \quad \text{with} \quad \hat{m}_i = T^{-1} \sum_{t=1}^T R_{it} \\ \hat{\mathbf{V}} &= T^{-1} \sum_{t=1}^T \mathbf{r}_t \mathbf{r}_t' \quad \text{with} \quad \hat{v}_{i,j} = T^{-1} \sum_{t=1}^T r_{it} r_{jt} \\ \hat{\mathbf{S}} &= T^{-1} \sum_{t=1}^T \mathbf{r}_t \mathbf{r}_t' \otimes \mathbf{r}_t' \quad \text{with} \quad \hat{s}_{i,n(j-1)+h} = T^{-1} \sum_{t=1}^T r_{it} r_{jt} r_{ht} \end{aligned}$$

³ See Gouriéroux and Monfort (2005) and Sentana (2009) for alternative econometric approaches and Lai (1991) for a mathematical programming approach also taking account of skewness.

⁴ Alternatively, R_i may be interpreted as excess returns as in Joro and Na (2006).

⁵ As usual, ‘ \otimes ’ denotes the Kronecker product. See Harville (2008, p. 337ff.) for the definition and rules.

$$\hat{K} = T^{-1} \sum_{t=1}^T \mathbf{r}_t \mathbf{r}'_t \otimes \mathbf{r}'_t \otimes \mathbf{r}'_t \quad \text{with} \quad \hat{k}_{i,n^2(j-1)+n(h-1)+l} = T^{-1} \sum_{t=1}^T r_{it} r_{jt} r_{ht} r_{lt}, \tag{3}$$

respectively. We find the scalar moments of asset i at $\hat{m}_i = \hat{\boldsymbol{\mu}}_i$, $\hat{v}_i = \hat{V}_{ii}$, $\hat{s}_i = \hat{S}_{i,n(i-1)+i}$ and $\hat{k}_i = \hat{K}_{i,n(n+1)(i-1)+i}$.

There are many parameters to be estimated in higher moment matrices leading to noisy estimates, and their number is growing with dimensionality. Shrinkage estimation intends to improve estimation precision in terms of mean square error at the expense of some bias. We perform a shrinkage estimation of the covariance matrix and the higher moment matrices with shrinkage toward independence. Given independence the co-skewness and co-kurtosis matrices in particular would have many zeros and the estimator is shrunk towards this target. Shrinkage methods for covariance matrix estimation are outlined by Ledoit and Wolf (2003, 2004) as well as by Boudt et al. (2020a) and Martellini and Ziemann (2010) for higher moment matrices (co-skewness and co-kurtosis). Computations are performed using the R-package ‘PerformanceAnalytics’ (Peterson and Carl 2020).

For an investor holding a portfolio of assets with shares $0 \leq x_i \leq 1$ (collected in a vector $\mathbf{x} = (x_1, \dots, x_n)'$ with $\sum_{i=1}^n x_i = 1$), we have a portfolio return of $R(\mathbf{x}) = \sum_{i=1}^n x_i R_i = \mathbf{x}' \mathbf{R}$ with expectation $m_p(\mathbf{x}) = E(R(\mathbf{x})) = \mathbf{x}' \boldsymbol{\mu}$ and variance $v_p(\mathbf{x}) = \text{Var}(R(\mathbf{x})) = \mathbf{x}' \mathbf{V} \mathbf{x}$. Portfolio skewness is computed by $s_p(\mathbf{x}) = \text{Skew}(R(\mathbf{x})) = \mathbf{x}' \mathbf{S}(\mathbf{x} \otimes \mathbf{x})$ and portfolio kurtosis by $k_p(\mathbf{x}) = \text{Kurt}(R(\mathbf{x})) = \mathbf{x}' \mathbf{K}(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x})$.

For the portfolios defined by \mathbf{x} , we estimate the portfolio moments by

$$\begin{aligned} \hat{m}_p(\mathbf{x}) &= \mathbf{x}' \hat{\boldsymbol{\mu}} = \sum_{i=1}^n x_i \hat{m}_i \\ \hat{v}_p(\mathbf{x}) &= \mathbf{x}' \hat{\mathbf{V}} \mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n x_i x_j \hat{v}_{i,j} \\ \hat{s}_p(\mathbf{x}) &= \mathbf{x}' \hat{\mathbf{S}}(\mathbf{x} \otimes \mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n \sum_{h=1}^n x_i x_j x_h \hat{s}_{i,n(j-1)+h} \\ \hat{k}_p(\mathbf{x}) &= \mathbf{x}' \hat{\mathbf{K}}(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n \sum_{h=1}^n \sum_{l=1}^n x_i x_j x_h x_l \hat{k}_{i,n^2(j-1)+n(h-1)+l} \end{aligned} \tag{4}$$

with the notation clarifying the dependence on \mathbf{x} .

3.2 Directional distance functions

The cornerstone of the nonparametric approach to portfolio efficiency measurement is the so-called shortage function introduced by Luenberger (1992, 1995) in a production context. Chambers et al. (1996) propose the same concept as a directional distance function (DDF) for measuring the distance of an input–output point to the production

frontier function along a particular direction. They have pointed out the interpretation of the distance measure as an efficiency measure.

The methodological extension pursued in this work is to treat the direction vector involved in computing the directional distances as a part of the optimization problem for calculating the optimal distances to the portfolio frontier function which is interpreted as portfolio efficiency. The following exposition directly turns to the MVSK case, while the MVS and MV cases can be easily obtained by deleting the respective constraints for skewness and kurtosis.

The formal definition of the DDF is based on the admissible set of all nonnegative vectors \mathbf{x} summing up to unity,

$$\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}, \mathbf{1}'\mathbf{x} = 1\}, \tag{5}$$

with $\mathbf{0}$ and $\mathbf{1}$ as conformable vectors of zeros and ones, respectively. Considering a vector of portfolio moments $\mathbf{z}_p(\mathbf{x})$, e.g., $\mathbf{z}_p(\mathbf{x}) = (m_p(\mathbf{x}), -v_p(\mathbf{x}))'$ for the mean–variance case (where the negative sign of the variance indicates that this moment is to be minimized), the so-called disposal representation set is the admissible moment space

$$\mathcal{D} = \{\mathbf{z} \in \mathbb{R}^{|\mathbf{z}|} : \mathbf{x} \in \mathcal{P}, \mathbf{z} \leq \mathbf{z}_p(\mathbf{x})\}, \tag{6}$$

where $|\mathbf{z}|$ denotes the length of the vector \mathbf{z} (i.e., 2 in the MV case, 3 in the MVS case and 4 in the MVSK case).

The shortage function or DDF as our measure of portfolio efficiency is defined as

$$S_g(\mathbf{x}) = \sup\{\delta \geq 0 : \mathbf{z}_p(\mathbf{x}) + \delta \mathbf{g} \in \mathcal{D}\} \tag{7}$$

with the vector \mathbf{g} controlling the direction in which moments are to be expanded or contracted. In the mean–variance case, we would specify $g_1 > 0$ and $g_2 < 0$ to get an expansion of the mean jointly with a contraction of the variance. The portfolio efficiency measure δ (which actually is a measure of inefficiency) is to be maximized simultaneously with \mathbf{x} to find a point on the portfolio frontier function along the direction \mathbf{g} . In other words, the value of δ is chosen jointly with \mathbf{x} to maximize the moments along the direction \mathbf{g} so that the vector $\mathbf{z}_p(\mathbf{x}) + \delta \mathbf{g}$ remains in \mathcal{D} and therefore hits the frontier. For points on the frontier, we obtain $\delta = 0$.

For computing the value of the shortage function, we use the basic DDF portfolio optimization problem which can be stated as

$$\begin{aligned} & \max_{\delta, \mathbf{x}} && \delta \\ & \text{s.t.} && \hat{m}_i + \delta g_m \leq \hat{m}_p(\mathbf{x}) \\ & && \hat{v}_i - \delta g_v \geq \hat{v}_p(\mathbf{x}) \\ & && \hat{s}_i + \delta g_s \leq \hat{s}_p(\mathbf{x}) \\ & && \hat{k}_i - \delta g_k \geq \hat{k}_p(\mathbf{x}) \\ & && \mathbf{1}'\mathbf{x} = 1 \\ & && \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{8}$$

with the restriction $\mathbf{x} \geq \mathbf{0}$ excluding the possibility of short sales.⁶ In addition, the vector of portfolio weights is required to sum to unity $\mathbf{1}'\mathbf{x} = 1$. If specific institutional restrictions on trading in a market exist, these could be introduced by additional restrictions of the general form $\mathbf{A}\mathbf{x} \leq \mathbf{b}$. Program (8) has to be computed for each asset $i \in \{1, \dots, n\}$ under consideration.⁷

The directions g_m, g_v, g_s, g_k (now assuming all ≥ 0) collected in the vector $\mathbf{g} = (g_m, g_v, g_s, g_k)'$ are to be supplied by the analyst. A frequent choice in the literature is to base the directions on the moments of the asset i under evaluation, i.e., $\mathbf{g} = (|\hat{m}_i|, \hat{v}_i, |\hat{s}_i|, \hat{k}_i)'$. In the subsequent empirical application, we compute (8) for the single fixed direction with maximizing only the mean ($\mathbf{g} = (|\hat{m}_i|, 0, 0, 0)'$) minimizing only the variance ($\mathbf{g} = (0, \hat{v}_i, 0, 0)'$) maximizing only the skewness ($\mathbf{g} = (0, 0, |\hat{s}_i|, 0)'$) or minimizing only the kurtosis ($\mathbf{g} = (0, 0, 0, \hat{k}_i)'$) as derived from the effect of the moments on expected utility (2). All these choices share the particular advantage of making the efficiency measure δ invariant to units of measurement. This measure can then be interpreted as a percentage improvement of a single moment or of a mixture of all moments.⁸

We also treat a case named as fixed where the direction vector is specified as $\mathbf{g} = (|\hat{m}_i|, \hat{v}_i, |\hat{s}_i|, \hat{k}_i)'$ and allows computing the distance to the portfolio frontier function in the direction of simultaneously increasing mean and skewness while reducing variance and kurtosis. The label 'fix' is here to be understood in the sense of being fixed for the specific values of the moments of asset i under evaluation but not in the sense of being fixed across the different assets in the sample. In particular, this direction vector is not the result of an optimal choice of the directions. This optimal choice is considered next.

3.3 Endogenous direction choice

The choice of the direction vector exerts an influence on the measured portfolio efficiency. In this respect, Kerstens et al. (2012) discuss several ways to choose the direction vector for mean–variance problems. They conclude that the choice of direction vector affects the relative ranking of the portfolios. However, they are not determining an optimal choice of the direction vector. In the following, we pursue this methodological extension. It is inspired by the work on direction choice for directional distance functions outside of portfolio efficiency analysis by Färe et al. (2013) and Hampf and Krüger (2015) in particular.

⁶ In general, restrictions on \mathbf{x} could be: (a) $\mathbf{x} \geq \mathbf{0}$ if lending and borrowing at a risk-free rate are possible, (b) $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{1}'\mathbf{x} \leq 1$ if only lending is possible, (c) $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{1}'\mathbf{x} = 1$ if neither lending nor borrowing is possible. Cases (a) and (b) require the presence of a riskless asset which could be achieved by basing the analysis on excess returns, see Joro and Na (2006). The exclusion of short sales is sensible since short sales are often legally restricted or prohibited at all. In addition, allowing short sales leads to a much greater dispersion of the portfolio weights and to an exaggeration of measurement error as demonstrated by Jagannathan and Ma (2003).

⁷ This program can be viewed as an extension of the quadratic optimization programs to compute pure mean–variance portfolio frontiers à la (Markowitz 1959).

⁸ As an example, we would obtain $\delta_i = (\hat{m}_p(\mathbf{x}_i) - \hat{m}_i)/|\hat{m}_i|$ for the pure mean maximizing case with \mathbf{x}_i denoting the solution value for \mathbf{x} in the case of asset i .

To endogenize the direction choice, we introduce weights $\alpha = (\alpha_m, \alpha_v, \alpha_s, \alpha_k)'$ for the directions into the optimization problem (8) and treat these weights as additional parameters to be optimized. This leads to the modified program

$$\begin{aligned}
 & \max_{\delta, \alpha, \mathbf{x}} \quad \delta \\
 & \text{s.t.} \quad \hat{m}_i + \delta\alpha_m |\hat{m}_i| \leq \hat{m}_p(\mathbf{x}) \\
 & \quad \hat{v}_i - \delta\alpha_v \hat{v}_i \geq \hat{v}_p(\mathbf{x}) \\
 & \quad \hat{s}_i + \delta\alpha_s |\hat{s}_i| \leq \hat{s}_p(\mathbf{x}) \\
 & \quad \hat{k}_i - \delta\alpha_k \hat{k}_i \geq \hat{k}_p(\mathbf{x}) \\
 & \quad \mathbf{1}'\mathbf{x} = 1 \\
 & \quad \mathbf{x} \geq \mathbf{0} \\
 & \quad \mathbf{1}'\alpha = 1 \\
 & \quad \alpha \geq \mathbf{0}
 \end{aligned} \tag{9}$$

where the direction weights α are restricted to be nonnegative and to sum to unity. This restriction is necessary to identify δ .

Program (9) is a nonlinear program anyway because of the nonlinear dependence of the portfolio moments on \mathbf{x} , see (4). In addition, the products of δ with the α weights can lead to instability of the numerical optimization procedure. Therefore, we actually solve an equivalent and presumably numerically more stable nonlinear program

$$\begin{aligned}
 & \max_{\gamma, \mathbf{x}} \quad \mathbf{1}'\gamma \\
 & \text{s.t.} \quad \hat{m}_i + \gamma_m |\hat{m}_i| \leq \hat{m}_p(\mathbf{x}) \\
 & \quad \hat{v}_i - \gamma_v \hat{v}_i \geq \hat{v}_p(\mathbf{x}) \\
 & \quad \hat{s}_i + \gamma_s |\hat{s}_i| \leq \hat{s}_p(\mathbf{x}) \\
 & \quad \hat{k}_i - \gamma_k \hat{k}_i \geq \hat{k}_p(\mathbf{x}) \\
 & \quad \mathbf{1}'\mathbf{x} = 1 \\
 & \quad \mathbf{x} \geq \mathbf{0} \\
 & \quad \gamma \geq \mathbf{0}
 \end{aligned} \tag{10}$$

with the endogenous direction weights now contained in the parameter vector $\gamma = (\gamma_m, \gamma_v, \gamma_s, \gamma_k)'$. Comparing (9) and (10), we immediately see the relation $\mathbf{1}'\gamma = \delta \cdot \mathbf{1}'\alpha = \delta$ (since $\mathbf{1}'\alpha = 1$ in (9)) which allows computing δ and retrieving $\alpha_j = \gamma_j/\delta$, $j \in \{m, v, s, k\}$.⁹ Hence, the target function value represents the same efficiency measure.

The computations are done using a sequential quadratic programming (SQP) algorithm for nonlinearly constrained gradient-based optimization by Kraft (1994). This algorithm solves a sequence of quadratic programming approximations and is able to deal with both equality and inequality constraints. For the procedure, we need the derivatives $\partial m_p/\partial \mathbf{x} = \boldsymbol{\mu}$, $\partial s_p/\partial \mathbf{x} = 2\mathbf{V}\mathbf{x}$, $\partial s_p/\partial \mathbf{x} = 3\mathbf{S}(\mathbf{x} \otimes \mathbf{x})$ and $\partial k_p/\partial \mathbf{x} = 4\mathbf{K}(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x})$. Details of the derivation can be found in the Appendix. We use the implementation in the R-package 'nloptr'.¹⁰ The programs prove to be

⁹ A similar idea related to the so-called directional Färe–Lovell efficiency measure is discussed in a production context by Briec (2000) and in a portfolio setting by Kerstens et al. (2011a).

¹⁰ See <http://ab-initio.mit.edu/wiki> for further details on the package and the algorithms.

computationally very well behaved. Bric and Kerstens (2010) discuss conditions for the existence of a global optimum and the possibility of local optima. Starting several times from randomly drawn starting configurations leads only to minor changes of the solutions.

4 Fama–French portfolios

For the empirical application, we use the data which provide the basis for the influential work of Fama and French (e.g., 1993, 1995, 1996, 2015) on market efficiency. These data are generously provided by Kenneth French on his homepage, where the full details of the definitions can be found.¹¹ To keep the presentation of the results manageable, we stick to the univariate sorts of stocks into portfolios by their market equity, book-to-market ratio, operating profitability and investment. These portfolios are used to form the four factors in addition to market excess returns in the five-factor model of Fama and French (2015). We use the value-weighted returns for these sorts formed on quintiles and add the returns of the market portfolio as a further asset.

In particular, we use the following portfolios as assets (with their self-explanatory abbreviations):

- MKT: Returns of a broad value-weighted index as the market portfolio, formed on the universe of all CRSP firms incorporated in the USA and listed on the NYSE, AMEX or NASDAQ.
- ME1, ME2, ME3, ME4, ME5: Portfolios formed on size as measured by market equity quintiles at the end of June each year, including all NYSE, AMEX and NASDAQ stocks for which we have market equity data for June.¹²
- BM1, BM2, BM3, BM4, BM5: Portfolios formed on quintiles of the book-to-market equity ratio at the end of June each year.
- OP1, OP2, OP3, OP4, OP5: Portfolios formed on quintiles of operating profitability (annual revenues minus cost of goods sold, interest expense, and selling, general, and administrative expenses divided by book equity for the last fiscal year) at the end of June each year.
- INV1, INV2, INV3, INV4, INV5: Portfolios formed on quintiles of investment (defined as the change in total assets from the fiscal year ending in year $t - 2$ to the fiscal year ending in $t - 1$, divided by total assets in $t - 2$) at the end of June each year.

Used are the monthly returns data over the period July 1963 until February 2020 for these 21 portfolios. Data from this archive is also used in other work on nonparametric portfolio efficiency analysis, e.g., Brandouy et al. (2010) and Branda (2015). Measurement errors and outliers may have detrimental effects on the estimation of the moments of the returns distribution. Chopra and Ziemba (1993) demonstrate the effects of measurement errors on the selection of optimal portfolios in a mean–variance framework. For dealing with this potential problem, we rely on the method of Boudt

¹¹ To be found at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

¹² The quintiles are indicated by 1 to 5 corresponding to the labels Lo20, Qnt2, Qnt3, Qnt4, Hi20 in the original data files.

et al. (2008) who propose an algorithm using methods from robust statistics to mitigate the influence of outliers. This method is preferred to using so-called L -moments as in Kerstens et al. (2011b). In addition and as outlined above, shrinkage estimation is used as a further device to reduce the effects of estimation error on the moment estimates.

Table 1 reports descriptive statistics for the monthly returns series over the entire sample period which are used for the subsequent computations. Note that the Boudt et al. (2008) method for cleansing the data is used here for the entire time series, whereas it is applied to the observations of each of the subperiods in the subsequent analysis. Reported are the minimum, median, maximum, mean, standard deviation, skewness coefficient, kurtosis coefficient, Jarque–Bera normality test (Jarque and Bera 1980) first-order autocorrelation coefficient, and the ERS unit-root test outcome (Elliott et al. 1996) for each of the 21 portfolio returns.

We find positive mean and median returns for all portfolios, while the dispersion is quite large. Skewness is negative, and the kurtosis coefficient is larger than three throughout. Analogous descriptive statistics without cleansing (results not shown) reveal that means and standard deviations are about the same, but the returns are more negatively skewed, and kurtosis is larger for all portfolios (i.e., median kurtosis is about 5). Both skewness and kurtosis lead to strong rejections of normality for all portfolios by the Jarque–Bera test. Regarding temporal dependence, we recognize a rather low degree of first-order autocorrelation, except maybe for the portfolio ME1 where autocorrelation is nearly 0.2. This finding corresponds to the strong rejections of the unit-root hypothesis by the ERS test.

5 Results

We consider the frontier functions for four subperiods. These subperiods are defined by relying on the stock market history of the USA as discussed in chapter 1 of Shiller (2015) and by Engle (2009, p. 94ff.). To support this discussion, Fig. 1 shows the time series of the Standard and Poors 500 performance index. The data at monthly frequency are retrieved from Yahoo Finance.¹³ The time-series length is restricted to the period since July 1963 when the portfolio data are also available.

The figure shows the index time series together with the NBER recession bars¹⁴ indicated by gray shades and major stock market events marked by vertical solid lines. These are (from left to right) the first oil price shock (January 1973), the Black Monday stock market crash (October 1987), the burst of the dot-com bubble (March 2000), the bankruptcy of Lehman Brothers (September 2008), the stock market selloff just before the Greece stock market crash (September 2015) and the global stock market downturn after September 2018.

The dashed vertical lines define the four periods for which we perform the non-parametric portfolio analysis and discuss the results subsequently. These periods are characterized as follows:

¹³ Used is the series ‘GSPC Historical Prices | S&P 500 Stock–Yahoo! Finance’ at monthly frequency.

¹⁴ Based on information from the NBER to be found at www.nber.org/cycles.

Table 1 Descriptive statistics

	Min	Median	Max	Mean	Sdev	Skew	Kurt	JB	AR1	ERS
MKT	-13.29	1.25	14.76	0.94	4.26	-0.31	3.72	25.51***	0.05	-13.44***
ME1	-19.17	1.37	20.83	1.10	6.03	-0.19	4.05	34.94***	0.19	-13.71***
ME2	-17.62	1.59	19.39	1.15	5.66	-0.32	3.79	29.30***	0.12	-12.36***
ME3	-16.01	1.43	17.72	1.12	5.16	-0.33	3.66	24.46***	0.11	-10.95***
ME4	-15.02	1.39	16.74	1.09	4.84	-0.27	3.70	21.83***	0.08	-11.13***
ME5	-12.73	1.15	14.02	0.90	4.08	-0.26	3.71	21.76***	0.01	-13.93***
BM1	-14.48	0.99	15.78	0.91	4.63	-0.26	3.60	17.76***	0.06	-13.73***
BM2	-13.43	1.18	14.95	0.98	4.31	-0.21	3.75	21.20***	0.05	-14.48***
BM3	-12.80	1.31	14.35	1.00	4.12	-0.20	3.92	28.30***	0.03	-14.19***
BM4	-13.20	1.27	14.66	1.04	4.25	-0.32	4.07	44.12***	0.06	-11.14***
BM5	-15.79	1.49	17.79	1.20	5.07	-0.39	4.15	54.78***	0.08	-11.11***
OP1	-16.92	1.14	17.62	0.82	5.38	-0.33	3.74	27.53***	0.10	-13.06***
OP2	-13.55	1.15	14.86	0.86	4.34	-0.28	3.87	29.87***	0.04	-14.67***
OP3	-13.08	1.23	14.56	0.96	4.22	-0.29	3.92	33.35***	0.03	-13.80***
OP4	-13.28	1.20	14.73	0.94	4.28	-0.29	3.72	23.90***	0.02	-11.20***
OP5	-13.23	1.20	14.91	1.07	4.27	-0.24	3.75	22.50***	0.06	-14.43***
INV1	-14.23	1.31	16.19	1.16	4.62	-0.18	3.88	25.42***	0.08	-10.84***
INV2	-12.04	1.12	13.60	1.01	3.92	-0.24	3.80	24.43***	-0.01	-12.70***
INV3	-12.37	1.19	13.90	0.96	3.99	-0.29	3.86	30.77***	0.02	-14.80***
INV4	-13.26	1.13	14.71	0.97	4.26	-0.24	3.73	21.42***	0.05	-13.94***
INV5	-16.97	1.26	18.61	0.91	5.43	-0.34	3.81	31.21***	0.08	-14.44***

Significance is indicated by *** on 1%, ** on 5% and * 10% level

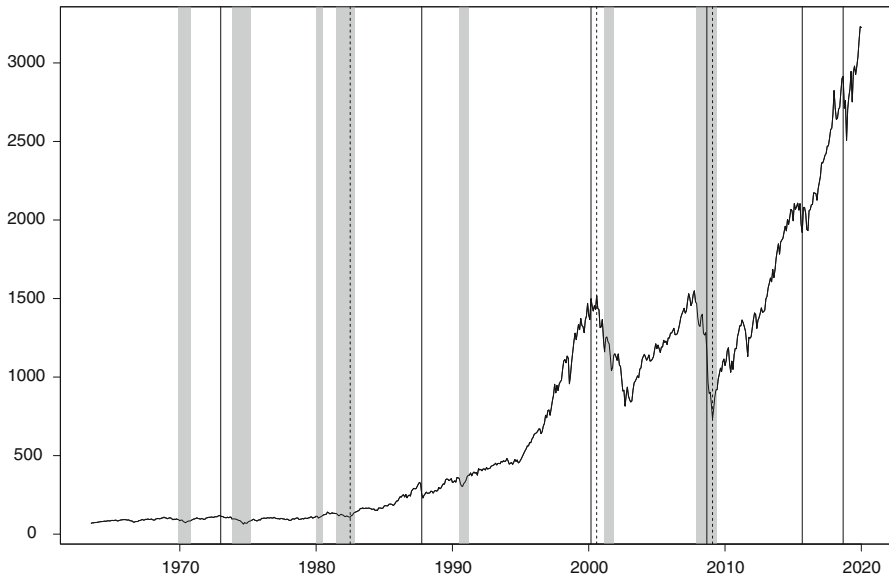


Fig. 1 S&P 500 time series

- Period 1 (1963:7–1982:7): This is the period before the start of the Millennium Boom phase, including the 1974/5 recession.
- Period 2 (1982:8–2000:8): This period is associated with a rapidly rising index disturbed by several crises such as the so-called Millennium Boom (or Bubble) including the October 1987 crash, the Tequila crisis of 1994, the Asian currency crisis of 1996/97, the October 1997 anniversary crash and the LTCM/Russian default crisis of 1998.
- Period 3 (2000:9–2009:2): This period is a period of stagnation and decline. It includes 9/11, the ownership society boom 2003-7 and the world financial crisis 2008/9 leading into the Great Recession.
- Period 4 (2009:3–2020:2): This is the boom period since March 2009, fueled by expansive monetary policy and low interest rates, only interrupted by two temporary downturns.

We see that the four periods are rather different regarding the specific events and the factors driving the stock market development. In particular, periods 2 and 4 are marked boom phases of stock market development.

Table 2 reports the means, standard deviations, skewness and kurtosis coefficients for the returns across all portfolios (jointly with their standard deviations) for the four periods. The first period is characterized by a small mean, a comparably large standard deviation, skewness near zero and excess kurtosis. By contrast, the second period with the Millennium Boom is characterized by a much larger mean return, a slightly smaller standard deviation, more negative skewness and about the same degree of excess kurtosis. The third period is more similar to the first but has a negative mean return, the largest standard deviation, the smallest (in the sense of most negative)

Table 2 Moments in the subperiods

Period:	Mean		Standard deviation		Skewness		Kurtosis	
	Mean	Sdev	Mean	Sdev	Mean	Sdev	Mean	Sdev
1963:7–1982:7	0.77	0.21	4.74	0.71	− 0.06	0.09	3.78	0.15
1982:8–2000:8	1.58	0.11	4.37	0.53	− 0.22	0.08	3.81	0.31
2000:9–2009:2	− 0.20	0.34	4.99	0.84	− 0.69	0.16	3.68	0.38
2009:3–2020:2	1.38	0.08	4.27	0.68	− 0.21	0.12	3.42	0.19

skewness and also some excess kurtosis. Finally, the fourth period after the decay of the Great Recession is a pronounced stock market boom phase with a large mean return, a standard deviation comparable to the second period, small skewness and mild excess kurtosis.

The presentation of the results of the portfolio efficiency analysis is structured in the following three subsections. First, we show the MV frontiers for the four periods. Then, we analyze the results for the nonparametric portfolio efficiency measures for the MV, MVS and MVS_K cases with fixing single directions towards the frontier at either mean or variance or skewness or kurtosis. Finally, we turn to the results with the efficiency determined in *all* relevant directions, either fixed or endogenously chosen by optimization.

5.1 Mean–variance frontiers

Figure 2 shows the mean–variance frontier functions for the four periods defined above jointly in a single plot. The efficient portfolio frontiers are computed by quadratic programming using the Rmetrics suite [Würtz et al. (2009) in the R-package ‘fPortfolio’] and relying on the method of Goldfarb and Idnani (1983). The efficient part of the frontier is depicted as a solid line, whereas the inefficient part is plotted in dashes. The positions of the individual portfolios are indicated by dots and their abbreviations defined above. We find the four frontier functions to be separated with the best performance during periods 2 (red) and 4 (green) with the highest means and the lowest dispersion of variances. The worst performance is found in period 3 (blue) with very small and even negative returns and the largest dispersion of variances. Period 1 (black) is located between these extremes.

We find quite large heterogeneity of the portfolio positions relative to the efficient frontier. Consistently on the frontier or close to the frontier during periods 1 and 3 are mainly the portfolios formed by stocks with small equity and high book-to-market ratios. During the boom periods 2 and 4, these portfolios are farther away from the frontier. Instead, the portfolios consisting of firms with lower investment and higher operating profitability are either on the frontier or close to the frontier.

Concerning the individual portfolios, we always find the market portfolio in between the others, although, admittedly, this is not always easy to spot visually. This is quite natural since the market portfolio is constructed from a value-weighted average of all returns, whereas the other portfolios are formed from systematically selected subsets.

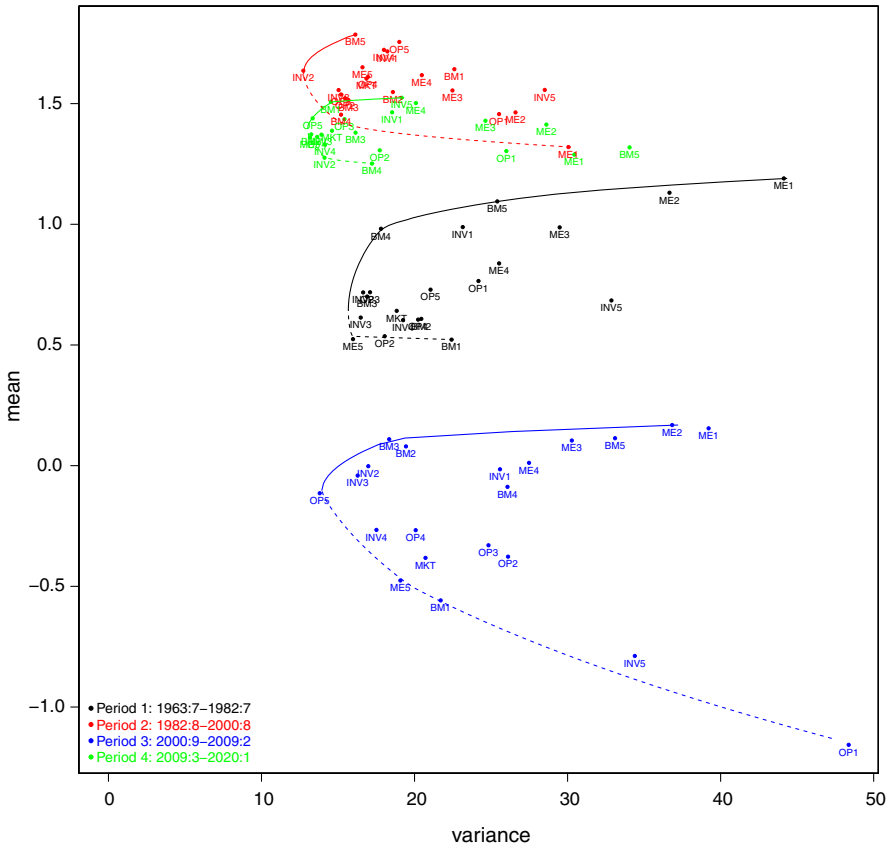


Fig. 2 Mean–variance frontier functions

The portfolios formed on size (i.e., market equity) ME1 to ME5 are quite dispersed with the lowest variances for the largest stocks (ME5) and the largest variances for the smallest stocks (ME1). We observe a weakly positive association of mean and variance for these portfolios in periods 1 and 3, while the association appears to be weakly negative in periods 2 and 4.

For the portfolios formed on book-to-market equity ratio, BM1 to BM5, we find BM1 among the portfolios with the lowest mean returns in periods 1 and 3, and in-between in the other periods. BM5, with the highest book-to-market ratio, is among the portfolios with the largest mean returns and a medium variance (with the most recent period being somewhat exceptional).

In the case of the portfolios formed on operating profitability, OP1 to OP5, we observe lower variances, with the separation of the low profitability stocks in OP1 in period 3. The association of mean and variance across these portfolios is clearly negative in periods 3 and 4. In contrast, the association is rather unclear in periods 1 and 2. The lowest profitability stocks in OP1 are consistently far away from the efficient frontier.

Finally, concerning the portfolios formed on investment, INV1 to INV5, the stocks in the fifth quintile with the highest investment (INV5) perform considerably worse in terms of their mean–variance relation compared to the stocks with smaller investment for all periods. This is different for the portfolios with lower investment, such as INV2 and INV3, which are quite close to the efficient frontier in all periods. An explanation may be that investments are frequently associated with adjustment costs caused by the reorganization within the firms and need time for becoming productive and leading to profits.

5.2 Single fixed directions

Figure 3 shows profile lines for the δ -values resulting from the efficiency measurement in a single fixed direction. The directions are tuned to the pure cases of mean maximization only ($\delta_{m \text{ only}}$), variance minimization only ($\delta_{v \text{ only}}$), skewness maximization only ($\delta_{s \text{ only}}$) and kurtosis minimization only ($\delta_{k \text{ only}}$). The direction vectors for these four cases used in the computation of program (8) are specified as $\mathbf{g} = (|\hat{m}_i|, 0, 0, 0)'$, $\mathbf{g} = (0, \hat{v}_i, 0, 0)'$, $\mathbf{g} = (0, 0, |\hat{s}_i|, 0)'$, $\mathbf{g} = (0, 0, 0, \hat{k}_i)'$ for portfolio i . The rows of the figure pertain to the four periods, and the columns pertain to the configuration of moments considered in the portfolio efficiency measurement (MV, MVS, MVSK). The abscissa lists the portfolios proceeding from the market portfolio to the portfolios sorted on market equity, book-to-market value, operating profitability and finally investment. On the ordinate, the respective δ -values are connected by the profile lines.

For the profile lines resulting from the MV analysis in the first column, we observe a larger portfolio inefficiency in the mean direction (solid line) during periods 1 and 3. These are the more stagnant phases of stock market development characterized by several crises leading to the Great Recession as seen from Fig. 1. The variance direction (dashed line) is more relevant during periods 2 and 4 which are the boom phases. Here, the portfolios tend to be farther away from the frontier in the variance direction, while they are closer to the frontier in the mean direction.

This corresponds to the configurations analyzed in Fig. 2, where we also observe relatively larger distances to the frontier in the vertical direction for periods 1 and 3 and relatively larger horizontal distances for periods 2 and 4. We also find zero or very low δ -values more frequently occurring for particular portfolios, revealing that these portfolios are on or close to the frontier. These tend to be portfolios consisting of small stocks (e.g., ME1 or ME2) of stocks with high book-to-market ratios (e.g., BM4 or BM5) of stocks with high operating profitability (e.g., OP5) and low investment (e.g., INV1 or INV2).

When the skewness dimension is added to compute the distances toward the MVS frontier which are shown in the second column, some major changes are visible. The profile lines of the δ -values in the skewness-only direction, shown by the dash-dotted lines, appear to be dominating in all except the third period where the distances are larger in the mean-only direction. During that particularly turbulent period, the large firm portfolios are much farther away from the frontier in the mean direction even when skewness is taken into account. The portfolios on or close to the frontier are similar to those of the MV case, i.e., the small rather than large stocks, high rather

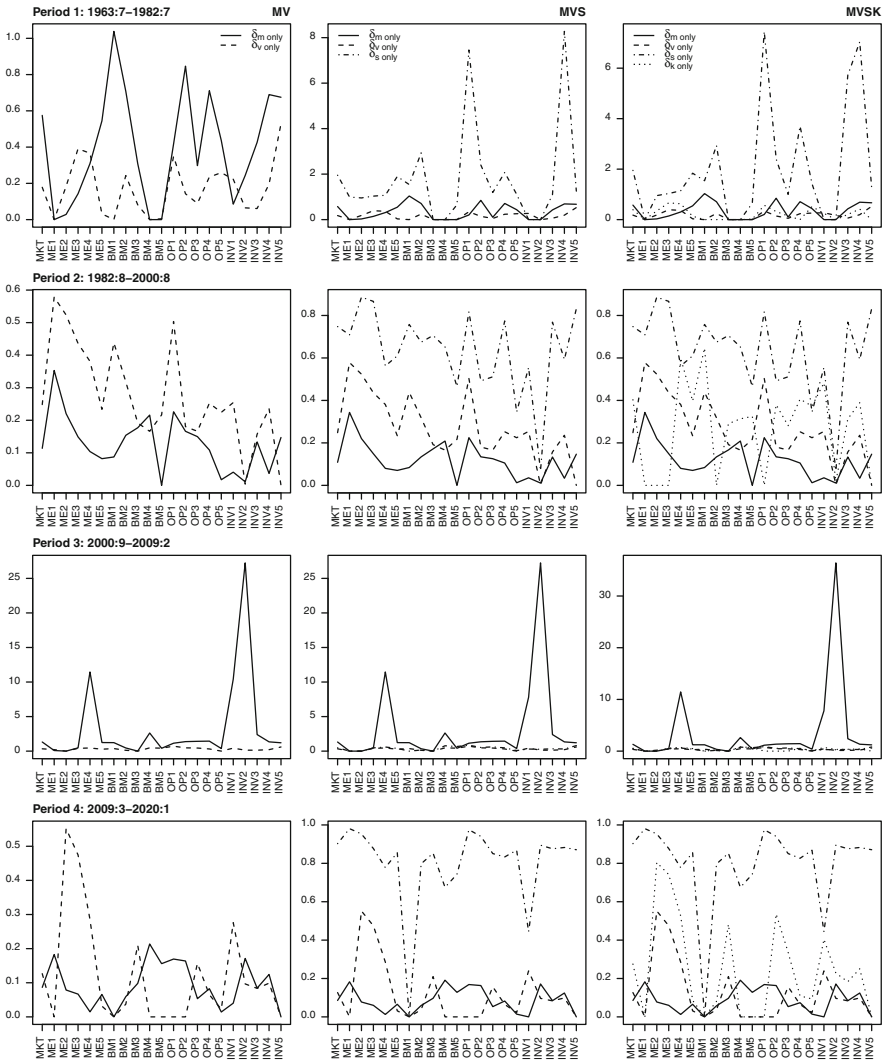


Fig. 3 δ -Values for single fixed directions

than low book to market, high rather than low profitability and low rather than high investment portfolios.

Introducing kurtosis to reach a full MVSK analysis for which the results reported in the third column does not lead to major changes of that pattern. The dotted profile line for the kurtosis-only variant either rests on a low level or does not exert a large impact on the other profile lines when compared to the MVS analysis. In the boom periods 2 and 4, the dotted lines are somewhere between the dashed and the dash-dotted lines for the variance-only and skewness-only cases, respectively. Hence, there is a certain distance to the frontier recognizable in the kurtosis direction, but this has

only a minor effect on the distances in the other directions. This resembles the finding that adding skewness to a mean–variance analysis has a much larger impact on the analysis than adding kurtosis (see Bricc and Kerstens 2010). From the findings here, one may conjecture that the addition of kurtosis does open a distinct dimension in which some portfolios are closer to the frontier than others but does not alter the space spanned by the other three moments substantially.

5.3 Multiple and optimized directions

We now turn from the consideration of directions pertaining to a single moment to the analysis with multiple directions combining increasing mean and skewness with decreasing variance and kurtosis toward the frontier function. The direction vector is either fixed at $\mathbf{g} = (|\hat{m}_i|, \hat{v}_i, |\hat{s}_i|, \hat{k}_i)'$ for portfolio i in program (8) in the MVSK case or endogenously determined by the optimization in program (10). The direction vectors for the MV and MVS cases are $\mathbf{g} = (|\hat{m}_i|, \hat{v}_i)'$ and $\mathbf{g} = (|\hat{m}_i|, \hat{v}_i, |\hat{s}_i|)'$, respectively. These variants with fixed multiple and optimized directions lead to the efficiency measures indicated by δ_{fix} and δ_{opt} , respectively.

In Fig. 4, these measures are plotted in the same way as previously with the four periods in the rows and the three moment configurations in the columns. We immediately see that the profile lines for the fixed directions lead to the smaller δ -values as compared to the profile lines with the direction determined to maximize δ (with the exception of the efficient portfolios where the optimal direction vector is indeterminate).

The profile lines are now overall closer to the zero line, and correspondingly, the portfolios are overall closer to being efficient in periods 2 and 4 compared to periods 1 and 3. The more efficient portfolios tend to be again those identified previously, i.e., the small, high book to market, high profitability and low investment portfolios. We again find larger changes of the profile lines when moving from MV to MVS compared to moving from MVS to MVSK. Thus, the impact of adding skewness is considerable, while the impact of adding kurtosis is minor. Overall, the level of the profile lines increases as more moments are considered.

Correspondingly, Fig. 5 depicts the optimal direction weights γ from solving program (10). When compared to the solutions for the δ -values from the single directions in Fig. 3, we find a close resemblance of the dominating profile lines.

In this respect, we find a larger distance in the direction of an increasing mean for period 1 in the MV case as well as in period 3 for all three configurations of moments. The distance is larger in the direction of reducing variance for most portfolios in the boom periods 2 and 4 in the MV case. Adding skewness shows that the distance to the frontier is considerable in the direction of increasing skewness in all periods except the third. Again, it appears that adding kurtosis seems to add a really distinct dimension. We observe considerable magnitudes of the direction weights in the kurtosis direction, while the weights for the other directions remain largely unaffected. This is particularly evident for the boom phases in the second and the fourth period.

Portfolio management can retrieve substantial information from the results. The δ -value indicates the extent of the overall possible improvement of portfolio efficiency.

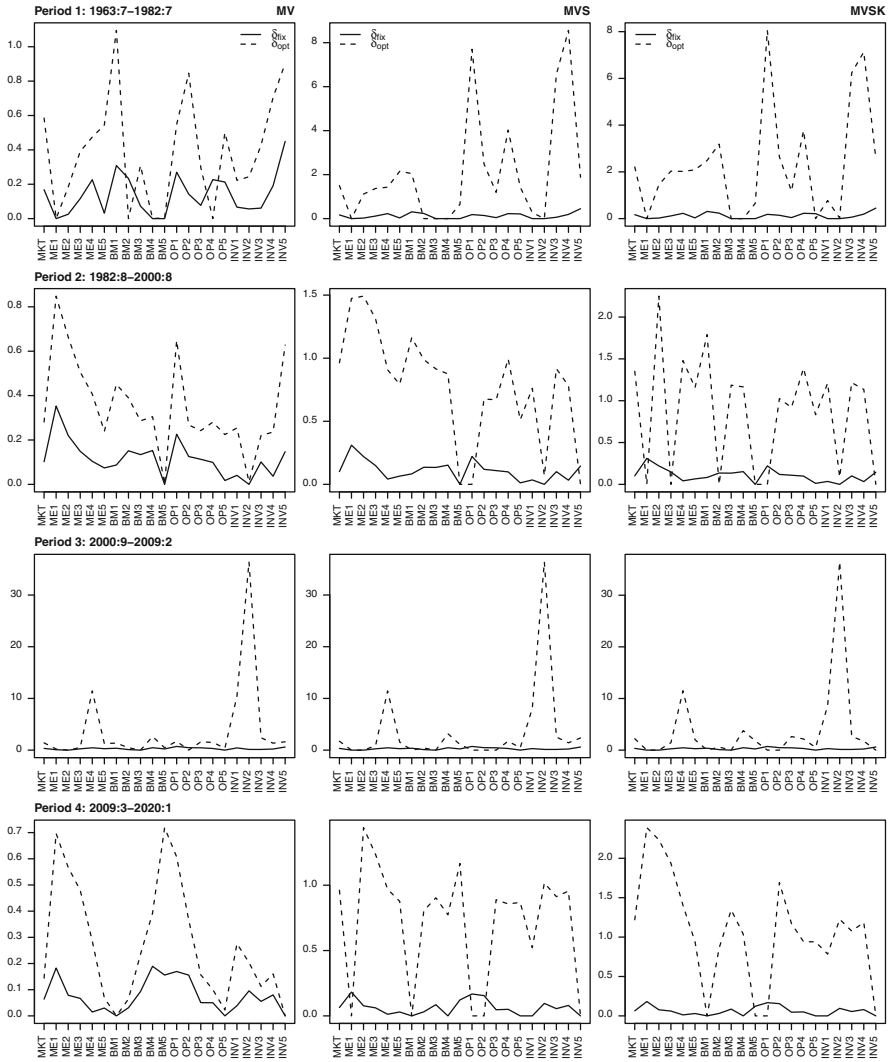


Fig. 4 δ -values for multiple and optimized directions

It represents the distance of the MVS K position of the portfolio under consideration to an efficient point on the frontier function. The γ -values give a dissection of the overall portfolio efficiency in the directions of the moments under consideration. This can be analyzed for single directions (i.e., the distance to the frontier if just one moment is maximized or minimized) or in all directions with an optimally chosen mix. In addition, the x -weights show to which extent other portfolios or assets are used to construct the efficient point on the portfolio frontier function.

To illustrate this further, we look at the results for portfolio ME1 in the final period in greater detail. The actual moment estimates ($\hat{m}_i, \hat{v}_i, \hat{s}_i, \hat{k}_i$) for $i = ME1$

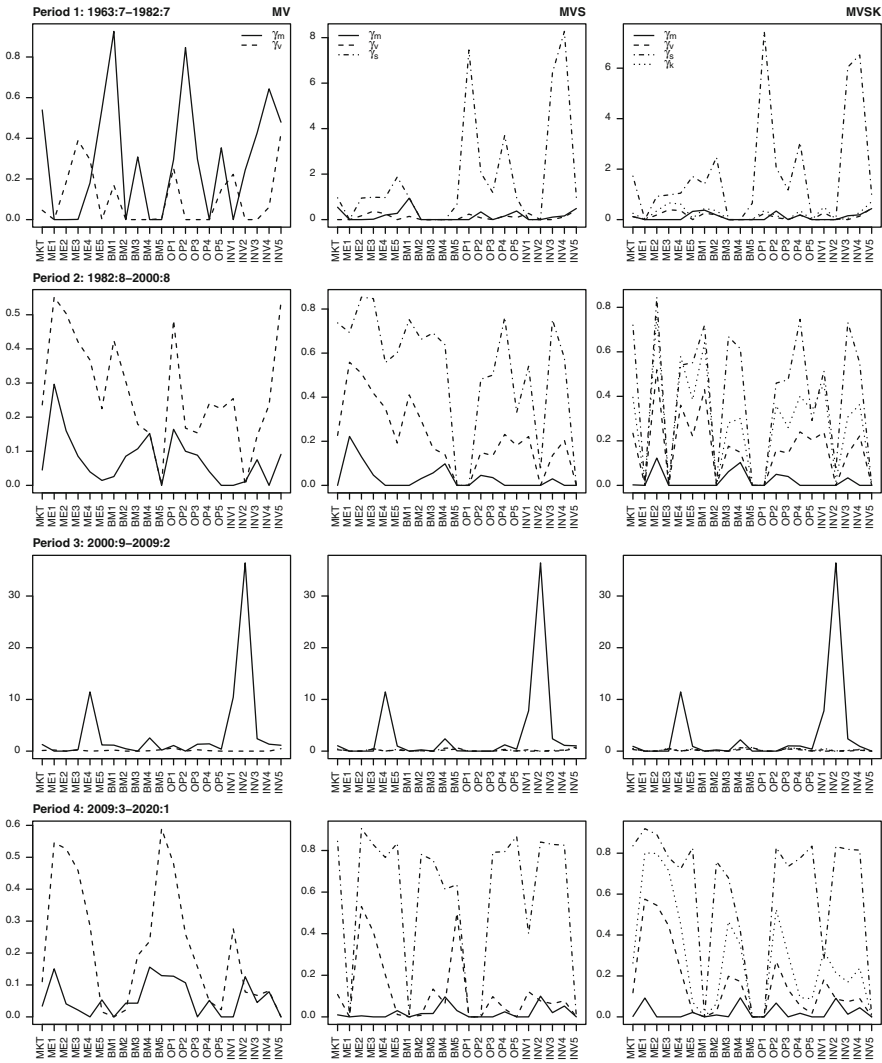


Fig. 5 γ -values for optimized directions

are (1.29, 30.44, -33.64, 2733.19). From the solution of the nonlinear program (10) with endogenous direction δ choice, we obtain an inefficiency measure of $\delta_i = 2.389$ which is the sum of the γ -values (0.092, 0.575, 0.920, 0.802). These values show that ME1 is quite close to the frontier regarding the mean return, farther away regarding risk (variance) and even more distant regarding the higher moments. An improvement of the utility from an investment in these directions would be possible by increasing return slightly, reducing variance substantially as well as increasing skewness and reducing kurtosis by larger margins. The point on the frontier function reached by moving along the direction given by the γ -values has the coordinates (1.41, 12.93, -2.69, 541.84)

stated in the order of the moments. This point can be replicated by a convex combination of 6 portfolios ME5 (0.10) BM1 (0.20) BM2 (0.27) OP4 (0.18) OP5 (0.17) and INV3 (0.08), where the x -weights are stated in parentheses and all other weights are equal to zero.

6 Conclusion

In this work, we pursue a nonparametric approach adapted from the literature on the measurement of productive efficiency using directional distance functions to the problem of determining optimal distances to portfolio frontier functions in multiple dimensions. This approach permits the straightforward incorporation of higher moments beyond mean and variance. We find that this nonparametric approach leads to several interesting findings.

First, we find that the efficient (or close to efficient) portfolios are those composed of small stocks, those with a high book-to-market ratio, those with a high operating profitability as well as those with low investment. This pattern is quite consistent across all moment configurations. Second, in the mean–variance setting we find larger distances to the frontier in a dominating mean direction for the more stagnating periods. In contrast, the distances to the frontier are much larger in the variance direction for the booming periods. Third, adding the skewness dimension reveals a large role for skewness in the sense of many portfolios being far away from the frontier in this direction (except for one subperiod where the distances in the mean and variance directions remain considerable). Fourth, adding the kurtosis dimension appears not to be as important as adding skewness. The results not change much compared to the mean–variance–skewness setting.

One deficiency of the above analysis is that the portfolio optimization programs are purely deterministic and prone to be sensitive to outliers. Such outliers are surely present in the form of stock market crashes. In the analysis of this paper, the possible effect of those outliers is mitigated by cleansing the returns series before computing the moments as well as by using shrinkage estimators for the second, third and fourth moments. Future research could deal with this issue in another way by attempting to adapt stochastic methods for estimating directional distance functions as proposed by Simar and Vanhems (2012) and Simar et al. (2012) to the portfolio setting. Another promising line of research would be the exploration of the predictive content in the distances to the portfolio frontier, and the specific directions in which these distances are established. Those investigations could proceed along the lines of the research surveyed by Rapach and Zhou (2013).

Acknowledgements Open Access funding provided by Projekt DEAL. I am grateful to an anonymous referee for suggestions to improve the paper. The usual disclaimer applies.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted

by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix

In this appendix, we provide the details of the calculation of the derivatives of portfolio mean, variance and in particular skewness and kurtosis with respect to the portfolio weights. In the literature, this is usually claimed to be straightforward [see, e.g., Jondeau and Rockinger (2006, p. 37)] and nowhere spelled out in detail. The present author only found this to be straightforward after recognizing the specific problem structure here. To beware others from spending much time on replicating these results, we spell out the calculations in full detail here.

For the portfolio mean $m_p = \mathbf{x}'\boldsymbol{\mu}$, we immediately obtain $\partial m_p / \partial \mathbf{x} = \boldsymbol{\mu}$. Likewise for the portfolio variance $v_p = \mathbf{x}'\mathbf{V}\mathbf{x}$, we easily get $\partial v_p / \partial \mathbf{x} = 2\mathbf{V}\mathbf{x}$ since \mathbf{V} is a symmetric square matrix (see Harville (2008) or Magnus and Neudecker (1999)).

The derivative of the portfolio skewness $s_p = \mathbf{x}'\mathbf{S}(\mathbf{x} \otimes \mathbf{x})$ is less straightforward and requires paying attention to the specific structure of $\mathbf{S} = E(\mathbf{r}\mathbf{r}' \otimes \mathbf{r}')$. Since

$$\mathbf{x}'\mathbf{S}(\mathbf{x} \otimes \mathbf{x}) = E(\mathbf{x}'(\mathbf{r}\mathbf{r}' \otimes \mathbf{r}')(\mathbf{x} \otimes \mathbf{x})) = E(\mathbf{x}'\mathbf{r}\mathbf{r}'\mathbf{x} \otimes \mathbf{r}'\mathbf{x}) = E(y \cdot y \cdot y)$$

with $y = \mathbf{r}'\mathbf{x}$ being scalar, we can apply the product rule of differentiation and the Kronecker product rules $(\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C})(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}) = \mathbf{A}\mathbf{x} \otimes (\mathbf{B} \otimes \mathbf{C})(\mathbf{x} \otimes \mathbf{x}) = \mathbf{A}\mathbf{x} \otimes \mathbf{B}\mathbf{x} \otimes \mathbf{C}\mathbf{x}$ forwards and backwards [assuming conformable matrices \mathbf{A} , \mathbf{B} and \mathbf{C} , see Harville (2008, p. 341)]. Acknowledging the equivalence of ordinary multiplication and the Kronecker product for scalar factors we obtain

$$\begin{aligned} \frac{\partial s_p}{\partial \mathbf{x}} &= E\left(\frac{\partial}{\partial \mathbf{x}} y \cdot y \cdot y\right) \\ &= E\left(\frac{\partial y}{\partial \mathbf{x}} \cdot y \cdot y + y \cdot \frac{\partial y}{\partial \mathbf{x}} \cdot y + y \cdot y \cdot \frac{\partial y}{\partial \mathbf{x}}\right) \\ &= 3E\left(\frac{\partial y}{\partial \mathbf{x}} \cdot y \cdot y\right) \\ &= 3E(\mathbf{r} \cdot \mathbf{r}'\mathbf{x} \cdot \mathbf{r}'\mathbf{x}) \text{ using } \frac{\partial y}{\partial \mathbf{x}} = \frac{\partial \mathbf{r}'\mathbf{x}}{\partial \mathbf{x}} = \mathbf{r} \\ &= 3E(\mathbf{r}\mathbf{r}'\mathbf{x} \otimes \mathbf{r}'\mathbf{x}) \\ &= 3E((\mathbf{r}\mathbf{r}' \otimes \mathbf{r}')(\mathbf{x} \otimes \mathbf{x})) \\ &= 3E(\mathbf{r}\mathbf{r}' \otimes \mathbf{r}')(\mathbf{x} \otimes \mathbf{x}) \\ &= 3\mathbf{S}(\mathbf{x} \otimes \mathbf{x}). \end{aligned}$$

The same reasoning can be applied to calculate the derivative for portfolio kurtosis $k_p = \mathbf{x}'\mathbf{K}(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x})$. Here, we have $\mathbf{K} = E(\mathbf{r}\mathbf{r}' \otimes \mathbf{r}' \otimes \mathbf{r}')$, and thus,

$$\mathbf{x}'\mathbf{K}(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}) = E(\mathbf{x}'(\mathbf{r}\mathbf{r}' \otimes \mathbf{r}' \otimes \mathbf{r}')(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}))$$

$$= E (\mathbf{x}' \mathbf{r} \mathbf{r}' \mathbf{x} \otimes \mathbf{r}' \mathbf{x} \otimes \mathbf{r}' \mathbf{x}) = E (y \cdot y \cdot y \cdot y)$$

with $y = \mathbf{r}' \mathbf{x}$ being scalar as above. Using the same approach as above for skewness, we get:

$$\begin{aligned} \frac{\partial k_p}{\partial \mathbf{x}} &= E \left(\frac{\partial}{\partial \mathbf{x}} y \cdot y \cdot y \cdot y \right) \\ &= E \left(\frac{\partial y}{\partial \mathbf{x}} \cdot y \cdot y \cdot y + y \cdot \frac{\partial y}{\partial \mathbf{x}} \cdot y \cdot y + y \cdot y \cdot \frac{\partial y}{\partial \mathbf{x}} \cdot y + y \cdot y \cdot y \cdot \frac{\partial y}{\partial \mathbf{x}} \right) \\ &= 4E \left(\frac{\partial y}{\partial \mathbf{x}} \cdot y \cdot y \cdot y \right) \\ &= 4E (\mathbf{r} \cdot \mathbf{r}' \mathbf{x} \cdot \mathbf{r}' \mathbf{x} \cdot \mathbf{r}' \mathbf{x}) \text{ again using } \frac{\partial y}{\partial \mathbf{x}} = \frac{\partial \mathbf{r}' \mathbf{x}}{\partial \mathbf{x}} = \mathbf{r} \\ &= 4E (\mathbf{r} \mathbf{r}' \mathbf{x} \otimes \mathbf{r}' \mathbf{x} \otimes \mathbf{r}' \mathbf{x}) \\ &= 4E ((\mathbf{r} \mathbf{r}' \otimes \mathbf{r}' \otimes \mathbf{r}') (\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x})) \\ &= 4E (\mathbf{r} \mathbf{r}' \otimes \mathbf{r}' \otimes \mathbf{r}') (\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}) \\ &= 4\mathbf{K} (\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}). \end{aligned}$$

References

Arrow KJ. Essays in the theory of risk-rearing. Chicago: Markham Publishing; 1971.

Boudt K, Peterson B, Croux C. Estimation and decomposition of downside risk for portfolios with non-normal returns. *J Risk*. 2008;11:79–103.

Boudt K, Cornilly D, Van Holle F, Willems F. Algorithmic portfolio tilting to harvest higher moment gains. *Heliyon*. 2020a;6:e03516.

Boudt K, Cornilly D, Verdonck T. A coskewness shrinkage approach for estimating the skewness of linear combinations of random variables. *J Finance Econ*. 2020b;18:1–23.

Branda M. Diversification-consistent data envelopment analysis based on directional-distance measures. *Omega*. 2015;52:65–76.

Brandouy O, Bricc W, Kerstens K, Van de Woestyne I. Portfolio performance gauging in discrete time using a luenberger productivity indicator. *J Bank Finance*. 2010;34:1899–910.

Bricc W. An extended Färe-Lovell technical efficiency measure. *Int J Prod Econ*. 2000;65:191–9.

Bricc W, Kerstens K. Portfolio selection in multidimensional general and partial moment space. *J Econ Dyn Control*. 2010;34:636–56.

Bricc W, Kerstens K, Lesourd JB. Single-period Markowitz portfolio selection, performance gauging, and duality: a variation on the Luenberger shortage function. *J Opt Theory Appl*. 2004;120:1–27.

Bricc W, Kerstens K, Jokung O. Mean-variance-skewness portfolio performance gauging: a general shortage function and dual approach. *Manag Sci*. 2007;53:135–49.

Chambers RG, Chung Y, Färe R. Benefit and distance functions. *J Econ Theory*. 1996;70:407–19.

Chopra VK, Ziemba WT. The effect of errors in means, variances, and covariances on optimal portfolio choice. *J Portfolio Manag*. 1993;19:6–11.

Dittmar RF. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross-section of equity returns. *J Finance*. 2002;57:369–403.

Edirisinghe N, Zhang X. Generalized DEA model of fundamental analysis and its application to portfolio optimization. *J Bank Finance*. 2007;31:3311–35.

Elliott G, Rothenberg TJ, Stock JH. Efficient tests for an autoregressive unit root. *Econometrica*. 1996;64:813–36.

Engle RF. Anticipating correlations: a new paradigm for risk management. Princeton: Princeton University Press; 2009.

- Fama EF, French KR. Common risk factors in the returns on stocks and bonds. *J Finance Econ*. 1993;33:3–56.
- Fama EF, French KR. Size and book-to-market factors in earnings and returns. *J Finance*. 1995;50:131–56.
- Fama EF, French KR. Multifactor explanations of asset pricing anomalies. *J Finance*. 1996;51:55–84.
- Fama EF, French KR. A five-factor asset pricing model. *J Finance Econ*. 2015;116:1–22.
- Färe R, Grosskopf S, Whittaker G. Directional output distance functions: endogenous constraints based on exogenous normalization constraints. *J Prod Anal*. 2013;40:267–9.
- Glawischning M, Sommersguter-Reichmann M. Assessing the performance of alternative investments using non-parametric efficiency measurement approaches: is it convincing? *J Bank Finance*. 2010;34:295–303.
- Goldfarb D, Idnani A. A numerically stable dual method for solving strictly convex quadratic programs. *Math Prog*. 1983;27:1–33.
- Gollier C, Pratt JW. Risk vulnerability and the tempering effect of background risk. *Econometrica*. 1996;64:1109–23.
- Gourieroux C, Monfort A. The econometrics of efficient portfolios. *J Empl Finance*. 2005;12:1–41.
- Hampf B, Krüger JJ. Optimal directions for directional distance functions: an exploration of potential reductions of greenhouse gases. *Am J Agric Econ*. 2015;97:920–38.
- Harvey CR, Siddique A. Conditional skewness in asset pricing tests. *J Finance*. 2000;55:1263–95.
- Harvey C, Liechty J, Liechty M, Müller P. Portfolio selection with higher moments. *Quant Finance*. 2010;10:469–85.
- Harville DA. Matrix algebra from a statistician's perspective. New York: Springer; 2008.
- Hens T, Rieger MO. Financial economics: a concise introduction to classical and behavioral finance. New York: Springer; 2010.
- Hlawitschka W. The empirical nature of Taylor-series approximations to expected utility. *Am Econ Rev*. 1994;84:713–9.
- Jagannathan R, Ma T. Risk reduction in large portfolios: why imposing the wrong constraints helps. *J Finance*. 2003;58:1651–84.
- Jarque CM, Bera AK. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Econ Lett*. 1980;6:255–9.
- Jondeau E, Rockinger M. Optimal portfolio allocation under higher moments. *Eur Finance Manag*. 2006;12:29–55.
- Joro T, Na P. Portfolio performance evaluation in a mean-variance-skewness framework. *Eur J Oper Res*. 2006;175:446–61.
- Kerstens K, Mounir A, Van de Woestyne I. Geometric representation of mean-variance-skewness portfolio frontier based upon the shortage function. *Eur J Oper Res*. 2011a;210:81–94.
- Kerstens K, Mounir A, Van de Woestyne I. Non-parametric frontier estimates of mutual fund performance using C- and L-moments: some specification tests. *J Bank Finance*. 2011b;33:1190–201.
- Kerstens K, Mounir A, Van de Woestyne I. Benchmarking mean-variance portfolios using a shortage function: the choice of direction vector. *J Oper Res Soc*. 2012;63:1199–212.
- Kimball MS. Precautionary savings in the small and in the large. *Econometrica*. 1990;58:53–73.
- Kraft D. Algorithm 733: TOMP—Fortran modules for optimal control calculations. *ACM Trans Math Softw*. 1994;20:262–81.
- Kraus A, Litzenberger RH. Skewness preference and the valuation of risk assets. *J Finance*. 1976;31:1085–100.
- Lai TY. Portfolio selection with skewness: a multiple objective approach. *Rev Quant Finance Account*. 1991;1:293–305.
- Lamb JD, Tee KH. Data envelopment analysis models of investment funds. *Eur J Oper Res*. 2012;216:687–96.
- Ledito O, Wolf M. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J Empl Finance*. 2003;10:603–21.
- Ledito O, Wolf M. A well-conditioned estimator for large-dimensional covariance matrices. *J Multivar Anal*. 2004;88:365–411.
- Luenberger DG. New optimality principles for economic efficiency and equilibrium. *JOTA*. 1992;75:221–64.
- Luenberger DG. Microeconomic theory. New York: McGraw-Hill; 1995.
- Magnus JR, Neudecker H. Matrix differential calculus with applications in statistics and econometrics. revised ed. New York: Wiley; 1999.

- Markowitz H. Portfolio election. *J Finance*. 1952;7:77–91.
- Markowitz H. Portfolio selection: efficient diversification of investments. New York: Wiley; 1959.
- Martellini L, Ziemann V. Improved estimates of higher-order comoments and implications for portfolio selection. *Rev Finance Stud*. 2010;23:1467–502.
- Morey MR, Morey RC. Mutual fund performance appraisals: a multi-horizon perspective with endogenous benchmarking. *Omega*. 1999;27:241–58.
- Murthi BPS, Choi YK, Desai P. Efficiency of mutual funds and portfolio performance measurement: a non-parametric approach. *Eur J Oper Res*. 1997;98:408–18.
- Peterson BG, Carl P. PerformanceAnalytics: econometric tools for performance and risk analysis, R package version 2.0.4, 2020. <https://CRAN.R-project.org/package=PerformanceAnalytics>. Accessed 27 Apr 2020.
- Rapach D, Zhou G. Forecasting stock returns. In Elliott G, Timmermann A (eds.) *Handbook of economic forecasting*, 2013;2A:329–383.
- Samuelson PA. The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *Rev Econ Stud*. 1970;37:537–42.
- Scott RC, Horvath PA. On the direction of preference for moments of higher order than the variance. *J Finance*. 1980;35:915–9.
- Sentana E. The econometrics of mean-variance efficiency tests: a survey. *Econom J*. 2009;12:C65–101.
- Shiller RJ. *Irrational exuberance*. 3rd ed. Princeton: Princeton University Press; 2015.
- Simar L, Vanhems A. Probabilistic characterization of directional distances and their robust versions. *J Econom*. 2012;166:342–54.
- Simar L, Vanhems A, Wilson PW. Statistical inference for DEA estimators of directional distances. *Eur J Oper Res*. 2012;220:853–64.
- Würtz D, Chalabi Y, Chen W, Ellis A. *Portfolio optimization with R/Rmetrics*, Rmetrics eBook, Rmetrics Association and Finance Online, Zürich, 2009.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.