Journal of Economic Dynamics & Control 34 (2010) 636–656

Contents lists available at ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

Portfolio selection in multidimensional general and partial moment space ☆

Walter Briec [a], Kristiaan Kerstens [b,*]

[a] University of Perpignan, GEREM, 52 avenue de Villeneuve, F-66000 Perpignan, France
[b] CNRS-LEM (UMR 8179), IESEG School of Management, 3 rue de la Digue, F-59000 Lille, France

ARTICLE INFO

Article history:
Received 2 January 2008
Accepted 13 October 2009
Available online 10 November 2009

JEL classification:
G11

Keywords:
Shortage function
Efficient frontier
K-moment portfolios

ABSTRACT

This paper develops a general approach for the single period portfolio optimization problem in a multidimensional general and partial moment space. A shortage function is defined that looks for possible increases in odd moments and decreases in even moments. A main result is that this shortage function ensures sufficient conditions for global optimality. It also forms a natural basis for developing tests on the influence of additional moments. Furthermore, a link is made with an approximation of an arbitrary order of a general indirect utility function. This non-parametric efficiency measurement framework permits to differentiate mainly between portfolio efficiency and allocative efficiency. Finally, information can, in principle, be inferred about the revealed risk aversion, prudence, temperance and other higher-order risk characteristics of investors.

1. Introduction

Maintaining strong assumptions on probability distributions and von Neumann–Morgenstern utility functions, Markowitz (1952) initiated modern portfolio theory by trading off risk and expected return of a portfolio. In a static context, he defines an efficient frontier of portfolios whose expected return cannot improve unless one is willing to assume more risk. This parametric approach where utility depends on the first two moments of a random variable's distribution is only consistent with the von Neumann–Morgenstern axioms of choice underlying expected utility (EU) theory when: (i) asset processes follow normal distributions, or (ii) investors have quadratic utility functions (e.g., Samuelson, 1967). However, as further developed in detail below, (i) many empirical studies cast doubt on the normality hypothesis of portfolio returns and (ii) point out that investors may well care about higher moments. In particular, they seem to prefer positive skewness and small kurtosis. Finally, Samuelson (1970) showed convincingly that the mean–variance (MV) approach is only appropriate if: (i) returns follow compact distributions and (ii) portfolio decisions are recurrent, such that the risk parameter becomes sufficiently small. Otherwise, higher moments are needed, since the quadratic approximation is not locally of high contact.

Meanwhile a large empirical literature has convincingly shown that normality of asset returns can be rejected. In particular, the distributional characteristics of a variety of financial and other economic variables (assets, options, hedge funds, etc.) indicate skewness and extreme kurtosis (see, e.g., Jondeau and Rockinger, 2003, or Kim and White, 2004). This stylized fact pertains to developed as well as emerging financial markets. Furthermore, it is also clear that traditional

☆ We thank the editor, Prof. C. Chiarella, two referees, as well as O. Brandouy and I. Van de Woestyne for their most constructive comments. The usual disclaimer applies.
* Corresponding author. Tel.: +33 320545892.
E-mail address: k.kerstens@ieseg.fr (K. Kerstens).

0165-1889/$ - see front matter
doi:10.1016/j.jedc.2009.11.001

Markowitz portfolio theory does not manage to diversify away this systematic skewness and kurtosis by increasing portfolio size (e.g., Gibbons et al., 1989). Hence, non-diversifiable skewness and kurtosis have become important research topics in asset valuation research. For instance, within the framework of the Capital Asset Pricing Model (CAPM) Rubinstein (1973) was the seminal contribution on multi-moment asset pricing. More recently, one does find articles on three-moment (e.g., Kraus and Litzenberger, 1976) and four-moment (e.g., Dittmar, 2002) asset pricing models. Higher expected returns compensate investors bearing systematic variance and kurtosis risks, while investors forego return to benefit from increasing systematic skewness. Recent books summarising the debate on this topic and the ensuing need for multi-moment portfolio theories are Berényi (2003) and Jurczenko and Maillet (2006).

A positive preference for skewness and negative preference for kurtosis has been postulated in explaining financial behavior since at least Scott and Horvath (1980). Skewness preference is one potential explanation for investors holding imperfectly diversified portfolios, just like it can contribute to explaining observed behavior in other areas (e.g., betting (see Golec and Tamarkin, 1998)). However, general choice theoretic foundations for the signs of third, fourth and higher derivatives of the von Neumann–Morgenstern utility function have only rather recently been developed. While a positive third derivative of EU (known as prudence) is meanwhile widely accepted (see Kimball, 1990), the sign of fourth and higher derivatives of EU is still typically met with some skepticism. Recent work linking the sign of EU derivatives to behavior towards risk firmly establishes a general preference for odd moments and an aversion to even moments. As to kurtosis aversion, Menezes and Wang (2005) define outer risk in terms of a transfer of actuarially neutral noise from the center of a distribution to its tail. They show that outer risk aversion (i.e., investors disliking greater outer risk) presupposes a negative fourth derivative of the von Neumann–Morgenstern utility function (known as temperance). Assuming that individuals dislike both: (i) a certain reduction in wealth and (ii) adding a zero-mean independent noise random variable to the wealth distribution, Eeckhoudt and Schlesinger (2006) manage to define a set of preferences over simple lotteries (having equal likelihoods for all outcomes) that offer a behavioral characterization of the general mathematical assumption that derivatives of the EU function alternate in sign. This confirms the generality of the large class of mixed risk aversion utility functions to characterize behavior towards risk, initially proposed in the seminal articles of Brockett and Golden (1987) and Caballé and Pomansky (1996).[1] Furthermore, since the signs of derivatives of utility coincide with preferences for n-th degree stochastic dominance, these lottery preference interpretations are also compatible with stochastic-dominance preferences.

Taking these mixed risk aversion preference structures for granted, the question is now how one can empirically conduct risk analysis in portfolio choice without imposing strong assumptions on either the return distribution or the investor preferences. Directly translating the alternating signs of the derivatives of the EU function, stochastic dominance (SD) criteria compare the expected utilities of the probability functions related to two risky prospects. While the financial literature mainly focuses on the first, second, and third degree SD criteria (Levy, 2006), generalizations to n-th degree SD exists since at least Thistle (1993). While the SD criteria respect the necessary and sufficient conditions for EU maximisation, they suffer from serious practical drawbacks in that pairwise comparisons between the alternative choices must be carried out (which requires information on the entire return distribution) and this renders the evaluation of diversification strategies extremely difficult.[2] Empirical work up to fourth-order and fifth-order SD seems to be available in the literature (e.g., Vinod, 2004; Tehranian, 1980).

By contrast, the empirical appeal of the traditional MV approach is entirely due to its ability to easily test and build diversification strategies that are efficient. The development of more general procedures to include higher moments when constructing portfolios has been severely hampered by computational problems (e.g., Markowitz, 1991). This contribution introduces a general procedure allowing for general higher moments in portfolio choice following the mixed risk aversion preference structures, even though it is well known that these moment orderings meet the necessary, but not the sufficient, conditions for EU maximization under strong additional assumptions on probability distributions and investor's utility functions.[3] This added generality should be weighted against the cost of having a theoretical imperfect solution compared to the SD approach.

Our approach reflects the basic conviction that a general procedure to describe the boundary of a higher dimensional, possibly non-convex multi-moment portfolio set and to select a boundary point in function of certain risk preferences necessitates employing a generalized distance function. Briec et al. (2004) integrate the *shortage function* (interpreted as an efficiency measure) into the Markowitz model and develop a dual framework to assess the degree of satisfaction of investors' preferences (the latter idea mounts back to Farrar, 1962).[4] They decompose portfolio performance into portfolio

[1] In fact, Brockett and Golden (1987) refer to "completely monotone" utility functions. Jondeau and Rockinger (2006, p. 34) speak about strict consistency of moment preferences, whereby the direction of preference is independent of wealth level. Notice that mixed risk aversion utility functions also help shedding light on various issues in insurance (self-protection, willingness-to-pay, and background risk): see Dachraoui et al. (2004). Furthermore, for multivariate decisions under risk, a similar class of simple lotteries allows to sign the cross derivatives of such utility functions (see Eeckhoudt et al., 2007), revealing the generality of this new approach.

[2] Recently, some progress has been made in terms of assessing efficient portfolio diversification according to SD criteria (see, e.g., Kuosmanen, 2004).

[3] Brockett and Kahane (1992) provide examples invalidating this leap from derivatives of utility functions to preferences for general moments of arbitrary distributions for the MV case as well as in general.

[4] In production theory, a generalized distance (shortage) function that simultaneously looks for reductions in inputs and expansions in outputs and that is dual to the profit function has been introduced by Luenberger (1995). The distance function is used in consumer theory to position consumption bundles relative to a reference utility level and it is dual to the expenditure function (see, e.g., Luenberger, 1992).

and allocative efficiency. Moreover, via the shadow prices associated with the efficiency measure, duality yields information about investors' risk aversion. The distance function is estimated using a non-parametric approach to approximate the true, unknown portfolio frontier (see Varian, 1983).

This work has been extended in Briec et al. (2007) to the non-convex mean–variance–skewness (MVS) space.[5] Here, we generalize this shortage function to the multi-moment portfolio problem to account for a preference for odd moments in addition to an aversion to even moments. The extension of the shortage function to the multi-moment space is straightforward, because a distance (gauge) function offers a perfect representation of multidimensional choice sets and can position any point relative to the boundary (frontier) of the set. However, it seems to have been unnoticed in the literature that this shortage (distance) function respects sufficient conditions for a global optimum on non-convex, multidimensional choice sets. The decomposition of portfolio performance into portfolio and allocative efficiency dissociates a description of the boundary of the portfolio choice set from the choice of an ideal point on this boundary according to well-defined investor preferences. This allows one to break away from the dominant approach in finance to postulate approximations of EU that necessitate relevant risk parameters, while investors have no opportunity within these approaches to obtain a clear idea on the multitude of efficient portfolios on the boundary of the choice set open to them, let alone that they know which of these boundary points they would happen to prefer. This new approach clearly separates both steps in portfolio analysis. Again, via the shadow prices associated with the efficiency measure in multi-moment portfolio space, duality yields under certain conditions information about investors' higher order risk preferences.

In a portfolio context, the shortage function projects any (in)efficient portfolio exactly on the possibly non-convex multidimensional moment portfolio frontier. In general, this shortage function accomplishes several goals of both theoretical and practical importance. First, portfolio performance is rated by measuring the distance between a portfolio and its optimal benchmark projection onto the multidimensional moment efficient frontier. Apart from a rating tool, this distance also reveals something about the goodness-of-fit of the maintained model (see Färe and Grosskopf, 1995). Second, the shortage function is a non-parametric estimate of the inner bound of the true, unknown portfolio frontier. Third, the shortage function evaluates odd moment expansions and even moment contractions simultaneously. Finally, the shortage function has a dual interpretation as a portfolio efficiency distance and could, in principle, reveal (shadow) risk parameters minimizing portfolio inefficiency.

While we develop this approach based on nonlinear programming for the multidimensional moment model with short sales excluded, it is good to stress that this offers a valid general framework for any other traditional portfolio extension (e.g., short selling, risk-free asset, buy-in thresholds for assets, cardinality constraints restricting the number of assets, transaction round lot restrictions, etc.). This contribution therefore paves the way to any portfolio selection approach consistent with a higher order Taylor expansion of the EU function in terms of moment preferences, as ideally dictated by the number of moments that turn out to count in statistically explaining portfolio choice behavior.

We claim that no such general procedure has so far been described to handle multi-moment portfolios. In the recent literature one can find various general approaches to estimating efficient portfolios including higher moments (see Adler and Kritzman, 2006; Gourieroux and Monfort, 2005; Harvey et al., 2004; Jondeau and Rockinger, 2006; Sharpe, 2007, amongst others). But, until now not a single generally valid framework seems to have emerged to handle third-order (accounting for skewness: see, for instance, Harvey and Siddique, 2000), fourth-order (accounting for skew and kurtosis: see, for instance, Dittmar, 2002), or higher degree polynomial forms for the EU function. Furthermore, all of the above approaches focus on selecting an ideal boundary point in function of certain risk preferences using an approximation of the EU function.

Apart from the few recent shortage function applications in the literature (e.g., Lozano and Guttiérez, 2008 or Bacmann and Benedetti, 2009), we are unaware of any non-utility based general procedure that moves beyond a three dimensional moment space in portfolio selection. Lai (1991) determines MVS optimal portfolios via a multi-objective programming approach. Jana et al. (2007) in a similar vein propose fuzzy programming to solve similar multi-objective non-linear programming problems. For instance, de Athayde and Flôres (2003, 2004) come up with analytical solutions for the mean–skewness–kurtosis portfolio frontier under restrictive assumptions, but at the cost of ignoring the variance dimension. Thus, general non-utility based procedures for multi-moment portfolios do not seem to be currently available.

Another strand in the literature focuses on lower partial moments rather than general moments (see Bawa, 1975 and Fishburn, 1977) to model the concern for deviations below a target return. It is well-known that mean lower partial moments (LPM) models always satisfy the necessary (though not the sufficient) conditions for EU theory under some strong assumptions on investor preferences (but, in the absence of any assumptions on probability distributions). Basically, investor utility should only depend on the mean and the partial moment appearing in the bi-criteria problem. For instance, while Fishburn (1977) proves this result only for second degree partial moments (i.e., lower semi-variance), Gotoh and Konno (2000) prove the same result for third degree partial moments (i.e., lower semi-skewness), and Ogryczak and Ruszczyński (2001) demonstrate the same for higher order partial moments.[6] Notice that parallel arguments for

[5] There exist alternative primal contributions that have tried to solve the MVS portfolio problem. For instance, de Athayde and Flôres (2004) develop an analytical solution to characterize the MVS portfolio frontier assuming a risk-free asset and shorting by minimizing the variance for given mean and skewness.

[6] As a matter of fact, Ogryczak and Ruszczyński (2001) demonstrate that sufficient conditions for EU theory can be respected by limiting the weights on the relevant risk measure in these bi-criteria problems.

developing an upper partial moment approach can also be found in the literature (see Holthausen, 1981). Given the discussion on general moments, it must be clear that the shortage function offers a way of comparing the goodness-of-fit of these various bi-criteria problems.

In addition, Konno et al. (1993) have been adding a lower semi-skewness to a given mean lower semi-variance model. To the extent that this is useful and for the ease of the exposition in parallel with the general moments, it is clear that the shortage function can also serve to assess the eventual extensions of these bi-criteria problems by including further lower partial moments of higher order compatible with more general investor preferences that are function of multiple lower partial moments (instead of just one in the case of the bi-criteria approach).[7]

Thus, this work mainly responds to a practical need for portfolio selection and management tools and develops a general theory for portfolio selection under multidimensional general and lower partial moments, while acknowledging that the relation with EU maximization is at best imperfect unless additional strong conditions are imposed. In addition, the case of upper partial moments is added for reasons of symmetry and it may as well provide a complementary basis for later developments seeking for a combination of lower and upper partial moments. Indeed, in line with Holthausen (1981), Balzer (2001) pleads for combining lower and upper partial moments up to the fourth order.[8] In view of the arguments for mixed risk aversion utility functions developed above, this plea of Balzer (2001) could eventually be extended up to the n-th moment. This would resemble the work by Kahneman and Tversky (1979) on prospect theory and Gul (1991) on disappointment aversion, both in the non-EU tradition, in that one treats losses and gains asymmetrically and approximates the utility of losses and gains by their respective successive partial moments. However, this development we leave for future work. Currently, it is good to underscore that models based on general, lower partial, or upper partial moments assume very different investor attitudes and that modelers should be aware of these differences when selecting any of these models.

Section 2 introduces the basic building blocks for the analysis, introduces the shortage function, studies its axiomatic properties, and formulates a general principle for testing the impact of moments on the approximation. The next section studies the link between the shortage function and the direct and indirect higher order moment utility functions. An empirical illustration using a small sample of 30 assets from the London Stock Exchange is provided in Section 4. Conclusions and issues for future work are summarized in the final section.

2. Efficient portfolios in *K*-moment space

2.1. A general class of moments

To introduce some basic notation and definitions, consider the problem of selecting a portfolio from n financial assets. Let $R_1, \ldots, R_n$ be random returns of assets $1, \ldots, n$. Assets are characterized by a set of moments of an arbitrary order. A portfolio $x=(x_1, \ldots, x_n)$ is simply a vector of weights specified over these n financial assets that sums to unity ($\sum_{i=1,\ldots,n} x_i = 1$). If shorting is impossible, then these weights must satisfy non-negativity conditions ($x_i \geq 0$). Therefore, the set of admissible portfolios can be written in general as[9]

$$\mathfrak{I} = \left\{ x \in \mathbb{R}^n : \sum_{i=1,\ldots,n} x_i = 1, x \geq 0 \right\}. \tag{2.1}$$

To be able to focus on higher moments and for notational convenience, we adopt the following general formulation. The return of portfolio x is defined as $R(x) = \sum R_i x_i$. Let $K \subset \mathbb{N} \backslash \{0\}$ be the index set of moments considered. We suppose that K is finite i.e., $|K| < +\infty$, where $|K|$ stands for the cardinality of K. This section intends to construct a general class of moments including as a special case, usual moments, lower partial moments and upper partial moments. To do this we consider for all $k \in K$ the functions $\psi_0 : \mathbb{R} \longrightarrow \mathbb{R}_+$ and $\psi_1 : \mathbb{R} \longrightarrow \mathbb{R}_+$ defined, respectively, as

$$\psi_0(w) = \min\{0, w\} \quad \text{and} \quad \psi_1(w) = \max\{0, w\}. \tag{2.2}$$

To construct this generalized class of moments we also introduce the function ψ_λ that is defined for all $\lambda \in [0,1]$ by

$$\psi_\lambda(w) = \frac{1}{\max\{(1-\lambda), \lambda\}} [(1-\lambda)\psi_0(w) + \lambda \psi_1(w)]. \tag{2.3}$$

Obviously, we have $\max\{(1-\lambda), \lambda\} \geq 1/2$ for all $\lambda \in [0,1]$. These notations are clearly coherent, since we have

$$\psi_{1/2}(w) = \frac{1}{\frac{1}{2}} \left[\frac{1}{2} \psi_0(w) + \frac{1}{2} \psi_1(w) \right] = \psi_0(w) + \psi_1(w) = w. \tag{2.4}$$

---

7 Harlow and Rao (1989) mention an approximation argument similar to Samuelson (1970) to justify the use of the traditional bi-criteria LPM models.

8 Scherer and Martin (2005) also plea to combine lower and upper semi-variances into a single model. Another similar proposal is found in Cumova et al. (2006).

9 This set of admissible portfolios can be easily adapted for additional constraints (e.g., transaction costs) that can be written as linear functions of asset weights: see Briec et al. (2004).

This means that ψ_λ comes down to the identity for $\lambda=\frac{1}{2}$. Notice that (i) if $\lambda=0$, then $\psi_\lambda=\psi_0$ is concave; and (ii) if $\lambda=1$, then $\psi_\lambda=\psi_1$ is convex.

Using this notation, a moment of order k and level λ is defined by

$$m_{k,\lambda}(x)=\begin{cases} E[\psi_\lambda(R(x))] & \text{if } k=1,\\ E[(\psi_\lambda(R(x)-E[R(x)]))^k] & \text{if } k\neq 1,\end{cases} \tag{2.5}$$

for all $\lambda\in[0,1]$. Notice that the return can also equal a target return (i.e., $R(x)=R_\tau$) as in, e.g., Fishburn (1977).[10] We then obtain three cases:

(a) If $\lambda=0$, then the level is 0 and $m_{k,0}(x)$ represents a lower partial moment of order k.
(b) If $\lambda=\frac{1}{2}$, then the level is $\frac{1}{2}$ and $m_{k,1/2}(x)$ represents the standard moment of order k.
(c) If $\lambda=1$, then the level is 1 and $m_{k,1}(x)$ represents the upper partial moment of order k.

Obviously, $m_{2,1/2}(x)$ denotes the portfolio variance, $m_{3,1/2}(x)$ is the portfolio skewness, and so on.

In the remainder of the paper we denote $\mathbb{N}^*=\mathbb{N}\backslash\{0\}$. Moreover, for all $K\subset N^*$, let $B_K=\bigcup_{k\in K}\{e_k\}$ be the canonical basis of $\mathbb{R}^K$.[11] We consider the K-moment representation function $m_{K,\lambda}:\mathfrak{I}\longrightarrow\mathbb{R}^K$ of level λ defined as

$$m_{K,\lambda}(x)=\sum_{k\in K} m_{k,\lambda}(x)e_k. \tag{2.6}$$

This function summarizes all moments k of level λ in the index set K characterizing a portfolio. For instance, the MV lower partial moment model is represented by $m_{\{1,2\},0}(x)=(E[\min\{0,R(x)\}],Var[\min\{0,R(x)\}])$, where the notation for the portfolio semi-variance is obvious. As another example, in the MVS ordinary moment case, we have $m_{\{1,2,3\},1/2}(x)=(E[R(x)], Var[R(x)],Sk[R(x)])$ where $Sk[R(x)]$ stands for the skewness.

It is useful to define the moment representation set of level λ for the set $\mathfrak{I}$ of portfolios as the range of $m_{K,\lambda}$ on $\mathfrak{I}$:

$$\mathcal{M}_{K,\lambda}=m_{K,\lambda}(\mathfrak{I})=\{m_{K,\lambda}(x):x\in\mathfrak{I}\}, \tag{2.7}$$

where $\lambda\in[0,1]$. This set can be extended by defining a moment (portfolio) *disposal representation set of level* λ:

$$\mathcal{DR}_{K,\lambda}=\mathcal{M}_{K,\lambda}-\mathcal{C}_{K,+}, \tag{2.8}$$

where

$$\mathcal{C}_{K,+}=\prod_{k\in K}(-1)^{k-1}\mathbb{R}_+ \tag{2.9}$$

is called the free disposal cone. For example, in the MV case, the free disposal cone is $\mathcal{C}_{\{1,2\},+}=(-\mathbb{R}_+)\times\mathbb{R}_+$. One can rewrite this disposal representation set $\mathcal{DR}_{K,\lambda}$ as follows:

$$\mathcal{DR}_{K,\lambda}=\{m'\in\mathbb{R}^K:\exists x\in\mathfrak{I}, m'\in m_{K,\lambda}(x)-\mathcal{C}_{K,+}\}. \tag{2.10}$$

Adding a cone defines a sort of "free disposal hull" of the moment representation of feasible portfolios. We denote the interior of this free disposal cone as $\mathcal{C}_{K,++}=\prod_{k\in K}(-1)^{k-1}\cdot\mathbb{R}_{++}$.

When evaluating portfolio efficiency, one must be able to identify a subset of weakly efficient portfolios.

Definition 2.1. The set of weakly efficient portfolios is defined, in the simplex, as

$$\Theta_{K,\lambda}(\mathfrak{I})=\{x\in\mathfrak{I}:(-1)^{k-1}m_{k,\lambda}(x)<(-1)^{k-1}m_{k,\lambda}(z)\forall k\in K\Longrightarrow z\neq\mathfrak{I}\}.$$

The set of weakly efficient portfolios is the set of all portfolios that are not strictly dominated in $|K|$- dimensional space. The power notation ensures that the even moments are as big as possible and the odd moments as small as possible. The weakly efficient subset can also be expressed with respect to the interior of the free disposal cone. Namely, we have $\Theta_{K,\lambda}(\mathfrak{I})=\{x\in\mathfrak{I}:m_{K,\lambda}(z)\in m_{K,\lambda}(x)+\mathcal{C}_{K,++}\Longrightarrow z\notin\mathfrak{I}\}=\{x\in\mathfrak{I}:(m_{K,\lambda}(x)+\mathcal{C}_{K,++})\cap\mathcal{M}_{K,\lambda}=\emptyset\}$. This definition suffices to define the set of weakly efficient portfolios.

One can also define a set of *strongly* efficient portfolios, but the weak formulation simplifies to some extent the results. To simplify notation we define a standard partial order based upon the dominance criterion defined above. We say that portfolio z is not dominated by portfolio x if $(-1)^{k-1}m_{k,\lambda}(z)\geq(-1)^{k-1}m_{k,\lambda}(x)$ for all $k\in K$ and we denote $z\succcurlyeq_{K,\lambda}x$. Equivalently, this means that $m_{K,\lambda}(x)\in m_{K,\lambda}(z)-\mathcal{C}_{K,+}$.

Along this line, $x\sim_{K,\lambda}z$ means that $m_{K,\lambda}(x)=m_{K,\lambda}(z)$. If such a condition does not hold, then we denote $x\nsim_{K,\lambda}z$. In addition, we denote $z\succ_{K,\lambda}x$ if $z\succcurlyeq_{K,\lambda}x$ and $z\nsim_{K,\lambda}x$. Using these notations, the subset of strongly efficient portfolios is defined as follows:

[10] In the empirical literature, this target return is often set at the risk-free return or a market return.
[11] $\mathbb{R}^K$ is the $|K|$- dimensional vector space indexed on K.

Definition 2.2. The set of strongly efficient portfolios is defined, in the simplex, as

$$\Xi_{K,\lambda}(\mathfrak{I}) = \{x \in \mathfrak{I} : z \succ_{K,\lambda} x \Longrightarrow z \notin \mathfrak{I}\}.$$

Obviously, $\Xi_{K,\lambda} \subset \Theta_{K,\lambda}$. All results in this contribution, except one, focus on weakly efficient portfolios.[12]

Similar to its role in the MV approach (Briec et al., 2004), the next subsection introduces the shortage function (Luenberger, 1995) as a performance indicator for the K-moment portfolio optimization problem.

2.2. Characterization of efficient portfolios using the shortage function

This subsection introduces the shortage function and studies its properties in the context of multidimensional moment portfolio theory, including lower partial and upper partial moments. Basic properties of the subset $\mathcal{DR}_{K,\lambda}$ on which the shortage function is defined have been discussed in Briec et al. (2004) for the MV model. It is possible to extend their definition to obtain an efficiency measure suitable for the general K-moment portfolio selection problem of level λ.

Definition 2.3. For all $\lambda \in [0,1]$, the function $S_{K,\lambda} : \mathfrak{I} \times (\mathcal{C}_{K,+} \backslash \{0\}) \longrightarrow \mathbb{R}_+$ defined as

$$S_{K,\lambda}(x;g) = \sup\{\delta : m_{K,\lambda}(x) + \delta g \in \mathcal{DR}_{K,\lambda}\}$$

is the shortage function for portfolio x of level λ in the direction of vector g.

This shortage function is a perfectly suitable portfolio management efficiency indicator because of its elementary properties. Since these properties carry over from the MV into the K-moment space, we state them without extensive comments.

Proposition 2.4. *For all $\lambda \in [0,1]$, $S_{K,\lambda}$ satisfies the following properties*:

(a) *If $g \in \mathcal{C}_{K,++}$, then we have*: $S_{K,\lambda}(x;g) = 0 \Longleftrightarrow x \in \Theta_{K,\lambda}(\mathfrak{I})$ (*weak efficiency*).
(b) *$S_{K,\lambda}$ is weakly monotonic on $\mathfrak{I}$, i.e., $z \succcurlyeq_{K,\lambda} x$ implies that*: $0 \leq S_{K,\lambda}(z;g) \leq S_{K,\lambda}(x;g)$.
(c) *If $g \in \mathcal{C}_{K,++}$, then $S_{K,\lambda}$ is continuous with respect to x.*

All the proofs in this contribution are relegated in appendix A. If the value of the shortage function is zero, then the portfolio is situated on the weakly efficient K-moment frontier of level λ.[13] A positive value indicates its degree of portfolio inefficiency. This inefficiency interpretation of the shortage function also leads to its use as a goodness-of-fit indicator that assesses the extent to which a maintained model fits observed portfolio choice behavior (Färe and Grosskopf, 1995). Secondly, a weakly dominated portfolio in terms of general or partial moment characteristics is classified as weakly less efficient. Finally, this shortage function is continuous as long as the direction vector does not contain any zero component.

The representation set $\mathcal{DR}_{K,\lambda}$, defined by expression (2.10), can be directly used to compute the shortage function by nonlinear optimization methods. Assume a sample of m portfolios $x^1, \cdots, x^m$. The shortage function for a specific portfolio x^j whose performance needs to be gauged ($S_{K,\lambda}(x^j;g)$) is computed by solving the following nonlinear program in K-moment space of level λ[14]:

$$\begin{aligned} \sup_{\delta,z} \quad & \delta \\ s.t. \quad & (-1)^{k-1} m_{k,\lambda}(x^j) + \delta g_k \leq (-1)^{k-1} m_{k,\lambda}(z), \quad k \in K, \qquad (P_{K,\lambda}) \\ & \sum_{i=1,\ldots,n} z_i = 1, \quad z_i \geq 0, \quad i = 1 \cdots n. \end{aligned}$$

Thus, gauging the performance of a sample of m portfolios requires computing one mathematical program for each portfolio in turn to determine its position with respect to the boundary of the choice set.[15] Combinations of moments of the portfolios in the sample constituting the portfolio frontier are situated on the RHS of ($P_{K,\lambda}$). The evaluated portfolio is represented on the LHS of ($P_{K,\lambda}$). Maximizing δ attempts to augment its odd moments and reduce its even moments in the direction of vector g. If $\delta = 0$, then the evaluated portfolio is efficient and on the boundary of the relevant portfolio frontier.

---

[12] Weak efficiency as a basic criterion in portfolio theory is introduced by analogy to the theoretical use of shortage (distance) functions in developing basic duality relations in consumption and production theory (see Cornes, 1992). Furthermore, it is a priori impossible to know to which extent portfolio applications using the shortage function introduced below could benefit from using a strong rather than a weak notion of efficiency.

[13] To guarantee strongly efficient solutions, it is possible to employ a different type of distance function (see, e.g., Briec, 2000 for such a solution in a production context: this could be easily transposed to a portfolio context).

[14] To save space, from this point on the indication 'of level λ' is suppressed whenever possible, since it applies in general.

[15] For example, in the MVS ordinary moment case (i.e., $m_{\{1,2,3\},1/2}(x) = (E[R(x)], Var[R(x)], Sk[R(x)])$ this nonlinear program reads as follows:

$$\sup_{\delta,z} \ \delta \text{ s.t. } \ E[R(x^j)] + \delta g_E \leq E[R(z)], \quad Var[R(x^j)] + \delta g_V \geq Var[R(z)], \quad Sk[R(x^j)] + \delta g_S \geq Sk[R(z)], \quad \sum_{i=1,\ldots,n} z_i = 1, \quad z_i \geq 0, \ i = 1 \cdots n.$$

If $\delta > 0$, then there are combinations of portfolios that yield higher odd moments and lower even moments. Hence, the evaluated portfolio is situated below the boundary and inefficient.

Remark 2.5. The shortage function is always well-defined and an infeasibility of its corresponding optimization problem $(P_{K,\lambda})$ cannot occur as long as it is defined with respect to each portfolio (x^j). Since by definition $m_{K,\lambda}(x) \in \mathcal{M}_{K,\lambda}$ for all $x \in \Im$, it follows that there is some $\delta \geq 0$ such that $m_{K,\lambda}(x)+\delta g \in \mathcal{DR}_{K,\lambda}$. Hence, the affine line spanned from $m_{K,\lambda}(x)$ in the direction of g meets the disposal representation set $\mathcal{DR}_{K,\lambda}$ which contains $\mathcal{M}_{K,\lambda}$. In the case where some vector $m \in \mathbb{R}^K$ does not lie in $\mathcal{DR}_{K,\lambda}$, then the direction of g may be infeasible at m if g contains a zero component.[16]

Clearly, if $\lambda = 0$ (i.e., the case of lower partial moments), then $m_{k,0}$ is a concave function. This mathematical program can then be converted to a convex optimization program (see Luenberger, 1984). Moreover, if $\lambda \in \{\frac{1}{2}, 1\}$ (i.e., the case of standard or upper partial moments) and $K \subset \{1\} \cup 2\mathbb{N}^*$ (i.e., portfolio models combining mean and even moments only), then this program also involves convex constraints. To simplify the exposition in the remainder, we identify two *convexity conditions*:

- $C1$: $\lambda \in \{\frac{1}{2}, 1\}$ and $K \subset \{1\} \cup 2\mathbb{N}^*$;
- $C2$: $\lambda = 0$.

These properties are now summarized in the following proposition:

Proposition 2.6. *For all $K \subset \mathbb{N}^\star$, if either C 1 or C 2 holds, then $\mathcal{DR}_{K,\lambda}$ is convex.*

The next proposition shows convincingly how the shortage function guarantees a global optimal solution for a large class of problems computed using the possibly non-convex, nonlinear program $(P_{K,\lambda})$ for all $\lambda \in [0,1]$ and for all $K \subset \mathbb{N}^*$.

Proposition 2.7. *Let (δ^*, z^*) be a local optimum of $(P_{K,\lambda})$. If either C 1 or C 2 holds, then (δ^*, z^*) is a global maximum of $(P_{K,\lambda})$. In such cases, if the first order and second order Kuhn–Tucker conditions hold at point (δ^*, z^*), then (δ^*, z^*) is a global maximum of $(P_{K,\lambda})$.*

One can slightly refine this result extending our analysis to the general case including odd moments. Paralleling Briec et al. (2007), it is easy to establish a specific condition involving local optimality of any portfolio model containing odd moments. Due to non-convexities, it is well-known that in these cases it is not guaranteed that a local optimum is a global optimum. To the best of our knowledge, the shortage function is the only tool providing a global optimal solution for the K-moment portfolio approach, including partial moments**,** in the cases enumerated in Proposition 2.7.

Remark 2.8. The shortage function can equally be applied in case of other sets of admissible portfolios. Also the above Proposition 2.7 holds true for a wide range of variations on the set of admissible portfolios. For reasons of space, we just focus on two cases:

(i) Availability of a risk-free asset: The results do not impose any particular structure on the moment matrices and tensors.[17] Hence, all statements also apply when there is a risk-free asset.
(ii) Possibility of short selling: The shortage function can be transposed to the case with short sales by dropping the non-negativity constraint $(x \geq 0)$. However, in such a case the set of feasible portfolios is not bounded. Hence, it is necessary to assume that the return of each asset is positive to ensure that the optimal portfolio necessarily belongs to the simplex.

As stated in the introduction, this contribution offers a perfectly general framework for analyzing any other traditional extension of empirical portfolio models (e.g., buy-in thresholds, cardinality constraints on the number of assets, transaction round lot restrictions, etc.) responding to specific investors' needs.

[16] Theoretical treatments of the distance function in developing duality relations in consumption and production theory often ignore the possibility of infeasibilities (see Cornes, 1992). From Remark 2.5 infeasibilities never occur in the context of our contribution. However, in situations where one focusses on, for instance, the multidimensional reconstruction of the efficient frontier rather than the measurement of portfolio efficiency, this question may be important. In such a context an extended shortage function can be defined on $\mathbb{R}^K$ to compute the frontier points, but in some circumstances this shortage function may well be ill-defined. In production theory Briec and Kerstens (2009) have exhaustively explored the circumstances (e.g., the axioms of production) under which infeasibilities may occur for the shortage function in both parametric and non-parametric production models alike. We are unaware of other work focusing on eventual similar problems of the use of distance functions in consumption, production or portfolio applications. Thus, given the novelty of this approach in portfolio theory, it is impossible to currently say much more about this issue in general.

[17] For instance, positive-definiteness of the variance-covariance matrix is not required.

2.3. Selection function and strongly efficient portfolios

In this subsection we introduce a multivalued map whose values necessarily contain a strongly efficient portfolio. Doing so, it is immediate to determine wether or not a given portfolio is strongly efficient. To understand the basic idea it is useful to note that one can equivalently rewrite the shortage function as follows:

$$S_{K,\lambda}(x;g) = \sup\{\delta : m_{K,\lambda}(x) + \delta g \leq m_{K,\lambda}(z), z \in \Im\}. \tag{2.11}$$

Definition 2.9. The multivalued map $\zeta : \Im \longrightarrow 2^{\Im}$ defined by

$$\zeta_{K,\lambda}(x) = \{z \in \Im : m_{K,\lambda}(x) + S_{K,\lambda}(x;g)g \leq m_{K,\lambda}(z)\}$$

is called a portfolio selection function.

This function provides a useful tool to characterize strongly efficient portfolios. An immediate implication is that for all $z^* \in \zeta_{K,\lambda}(x^j)$, the vector $(S_{K,\lambda}(x^j;g), z^*)$ is a solution of program $(P_{K,\lambda})$. We have the following statements.

Proposition 2.10. *For all $\lambda \in [0,1]$, the following conditions hold true*:

(a) *$\zeta_{K,\lambda}$ has non-empty values on $\Im$ (i.e., for all $x \in \Im$ we have $\zeta_{K,\lambda}(x) \neq \emptyset$).*
(b) *For all $x \in \Im$, $\zeta_{K,\lambda}(x)$ contains some strongly efficient portfolio (i.e., $\zeta_{K,\lambda}(x) \cap \Xi_{K,\lambda}(\Im) \neq \emptyset$).*
(c) *If $z \in \zeta_{K,\lambda}(x)$, then it is weakly efficient (i.e., $\zeta_{K,\lambda}(x) \subset \Theta_{K,\lambda}(\Im)$).*

The above result has a corollary that establishes that program $(P_{K,\lambda})$ yields a set of weakly efficient portfolio solutions that contains at least a strongly efficient solution. Hence, if the solution is unique, the solution is strongly efficient.

Corollary 2.11. *For all $\lambda \in [0,1]$, if (δ^*, z^*) is a solution of $(P_{K,\lambda})$, then z^* is weakly efficient (i.e., $z^* \in \Theta_{K,\lambda}(\Im)$). Moreover, if there is a unique solution z^*, then z^* is strongly efficient (i.e., $z^* \in \Xi_{K,\lambda}(\Im)$).*

Thus, the optimal portfolio weights resulting from program $(P_{K,\lambda})$ guarantee weak efficiency. However, there is always also a strongly efficient solution.

2.4. Testing the impact of general and partial moments

This subsection provides procedures to determine the influence of changing the set of general and partial moments considered in measuring portfolio performance. Using the goodness-of-fit interpretation of the shortage function (see Färe and Grosskopf, 1995), the goal is to outline a basis for the development of statistical tests about the relevance of including additional moments when approximating the EU function based on finite data sets. It is well-known that the quality of moment approximations of EU is an empirical issue (e.g., Hlawitschka, 1994). In the same vein, one can expect that the approximation quality of a partial series of a Taylor expansion of the shortage function needs to be empirically assessed.

The following definition measures in a straightforward manner the influence of the choice between two different subsets of moments K and K' in measuring portfolio performance. To simplify the statement, for all $g \in \mathcal{C}_{K \cup K',+}$, let $g_K = \sum_{k \in K} g_k e_k$ and $g_{K'} = \sum_{k \in K'} g_k e_k$ denote the orthogonal projection of g onto $\mathbb{R}^K$ and $\mathbb{R}^{K'}$, respectively.

Definition 2.12. For all $\lambda \in [0,1]$, the measure $\Delta_\lambda : (2^{\mathbb{N}^*} \backslash \emptyset) \times (2^{\mathbb{N}^*} \backslash \emptyset) \times \Im \times \mathcal{C}_{K,+} \longrightarrow \mathbb{R}$ defined as $\Delta_\lambda(K, K', x; g) = S_{K,\lambda}(x; g_K) - S_{K',\lambda}(x; g_{K'})$ is called a measure of moment impact.

For instance, suppose we compare two models $m_{\{1,2,3\},\frac{1}{2}}(x)$ and $m_{\{1,2,4\},\frac{1}{2}}(x)$. Both contain 3 moments, but differ in that the former adds the skewness to the MV model while the latter adds kurtosis. Then, $\Delta_\lambda(\{1,2,3\},\{1,2,4\},x;g)$ measures the relative goodness-of-fit of both models. In the case of lower partial moments, one could use the measure of moment impact to test, for instance, whether a mean semi-variance model fits the data better or worse than a mean semi-skewness model.

It is trivial to establish that the shortage function is decreasing when the set of moments increases. This has the following immediate consequence for the measure of moment impact when we consider two proper subsets of moments K and K':

Proposition 2.13. *For all $\lambda \in [0,1]$, if $K \subset K'$, then $\Delta_\lambda(K, K', x; g) \geq 0$.*

This proposition describes the effect of changing the set of moments (hence, constraints) on the difference between maximal value functions when considering two proper subsets of moments. E.g., the two models $m_{\{1,2,3\},\frac{1}{2}}(x)$ and $m_{\{1,2,3,4\},\frac{1}{2}}(x)$ differ only in that the latter adds kurtosis to a basic MVS model. Then, $\Delta_\lambda(\{1,2,3\},\{1,2,3,4\},x;g)$ must be semi-positive. It is only zero when both shortage functions obtain identical values, which would imply that adding the kurtosis constraint would not have had an impact on the objective function.

This definition and proposition offer a starting point for developing tests for the relevance of including specific additional moments in the approximation of the EU function. Indeed, a test measuring the role played by a specific moment k' in evaluating portfolio performance is now straightforwardly defined as:

Definition 2.14. For all $\lambda \in [0,1]$, the measure $I_\lambda : \mathbb{N}^* \times (2^{\mathbb{N}^*} \backslash \emptyset) \times \Im \times \mathcal{C}_{K,+}$ defined as $I_\lambda(k', K, x; g) = \Delta_\lambda(K, K \cup \{k'\}, x; g)$ is called a measure of moment k' impact on K-moment space.

For instance, assuming we start from a traditional MV model it is possible to test for the impact of adding the skewness (i.e., $I_\lambda(\{3\}, \{1,2\}, x; g)$) and thereafter to check whether adding the kurtosis adds any value (i.e., $I_\lambda(\{4\}, \{1,2,3\}, x; g)$).

An open challenge is to transform these exact goodness-of-fit tests, capturing the economic significance of deviations from rational behavior in portfolio decisions, into a statistical test (Varian, 1990). Given the inherent downward bias of any boundary estimator due to the dependency of the boundary on the portfolios in the sample, the small sample error and bias of these non-parametric frontier estimators can be probably be improved upon using simulated (bootstrapped) empirical distributions (see Simar and Wilson, 2000 for a successful implementation of this strategy for monotone boundaries in a production context). However, a crucial difference between the production and portfolio context is that perturbed observations are sufficient to compute bootstrap efficiencies in the former context while the return observations in the latter first need to be transformed into moment statistics (mean, variances and covariances, etc.). The transposition of the successful bootstrapping framework of these authors in a portfolio frontier framework is therefore not straightforward.

In the empirical application below we ignore this bias issue and we simply employ a test statistic developed by Li (1996) and refined by Fan and Ullah (1999) for dependent and independent observations alike to measure the difference between two densities of shortage functions.[18] Under the null hypothesis that both distributions are identical and the alternative hypothesis that they are different, this test statistic asymptotically follows a standard normal-distribution (for small samples, a bootstrap approximation can be employed).

Clearly, much more research is needed to arrive at proper test statistics capable to tackle the bias issue in the portfolio boundary estimation context. We end by noting that recently several authors started contributing to the statistical analysis of the shortage function framework in a portfolio context. For instance, Bacmann and Benedetti (2009) use Bayesian inference methods to address estimation risk using multivariate skewed distributions. Jurczenko et al. (2008) replace the classical moments by the far more robust L-moments, while Jurczenko and Yanou (2010) employ the even more robust trimmed L-moments. The next section studies the shortage function from a duality standpoint.

3. Indirect utility in *K*-moment space and duality result

3.1. Preferences and approximations

Portfolio selection has always been conceived as a two-step procedure. In the MV world, tracing the efficient set of portfolios is a first step to select an optimal portfolio on the boundary of the set for a given preference structure. To provide a dual interpretation of the shortage function, a corresponding general indirect utility function must be defined. Suppose that $K = \{1, \ldots, \overline{k}\}$. The link between the $\overline{k}$- th order derivatives of the utility function and the $\overline{k}$- th order moments is illustrated by taking a Taylor expansion of the EU of the final wealth w_f of an investor around his expected wealth $\overline{w}$ as follows:

$$u(w_f) = u(\overline{w}) + \sum_{k=1}^{\overline{k}} \frac{u^{(k)}(\overline{w})}{k!}(w_f - \overline{w})^k + \cdots. \tag{3.1}$$

This can be rewritten as

$$E[u(w_f)] = E[u(\overline{w})] + \sum_{k=1}^{\overline{k}} \frac{u^{(k)}(\overline{w})}{k!} E[(w_f - \overline{w})^k] + \cdots. \tag{3.2}$$

Finally, this leads to the expression:

$$E[u(w_f)] = u(\overline{w}) + \sum_{k=2}^{\overline{k}} \frac{u^{(k)}(\overline{w})}{k!} m_k^{\frac{1}{2}}(w_f) + \cdots. \tag{3.3}$$

Clearly, one supposes negative (positive) even (odd) derivatives of the EU function for behavior representing mixed risk aversion (Caballé and Pomansky, 1996). Taylor series expansions of the EU function have certain well-known limitations. Loistl (1976) already indicated that if a utility function is polynomial of degree m, then its value can be expressed via a

[18] This test statistic was probably first used by Kumar and Russell (2002) in a production frontier context and it has gained some popularity since then: see their Appendix (p. 546) for technical details.

finite Taylor series expansion; while if a utility function is not a polynomial function, then its value can be expressed via an infinite Taylor series expansion. Since no polynomial utility function is part of the class of mixed risk aversion utility functions (Brockett and Golden, 1987, p. 956), we know the second case prevails. Hlawitschka (1994) expanded on two important points in this respect: (i) when a Taylor series diverges, then the truncation at MV may provide good approximation, and (ii) when a Taylor series converges, then adding more terms may actually worsen the approximation (since the usefulness of Taylor series approximations is purely an empirical matter and one can say very little about the behavior of partial series).[19]

One can also extend this approach to the case of lower partial moments ($\lambda = 0$). This we do considering the Taylor expansion:

$$u(w_f - \overline{w}) = u(0) + \sum_{k=1}^{\overline{k}} \frac{u^{(k)}(0)}{k!}(w_f - \overline{w})^k + \cdots. \tag{3.4}$$

We can then deduce that

$$u(\psi(w_f - \overline{w})) = u(0) + \sum_{k=1}^{\overline{k}} \frac{u^{(k)}(0)}{k!}\psi^k(w_f - \overline{w}) + \cdots. \tag{3.5}$$

Consequently,

$$E[u(\psi(w_f - \overline{w}))] = u(0) + \sum_{k=1}^{\overline{k}} \frac{u^{(k)}(0)}{k!}E[\psi_0^k(w_f - \overline{w})^k] + \cdots. \tag{3.6}$$

Finally, this leads to the expression

$$E[u(\psi(w_f - \overline{w}))] = u(0) + \sum_{k=1}^{\overline{k}} \frac{u^{(k)}(0)}{k!}m_k^0(w_f) + \cdots. \tag{3.7}$$

Now, let us define the function u_0 as

$$u_0(w_f) = \begin{cases} u(\psi(w_f - \overline{w})) - u(0) & \text{if } w_f - \overline{w} \le 0, \\ 0 & \text{if } w_f - \overline{w} \ge 0. \end{cases} \tag{3.8}$$

Notice that we have $u_0(\overline{w}) = 0$. One can then deduce that:

$$E[u_0(w_f)] = \sum_{k=1}^{\overline{k}} \frac{u^{(k)}(0)}{k!}m_k^0(w_f) + \cdots. \tag{3.9}$$

Using a similar procedure one can establish a parallel result in the context of upper partial moments ($\lambda = 1$). We have

$$E[u_1(w_f)] = \sum_{k=1}^{\overline{k}} \frac{u^{(k)}(0)}{k!}m_k^1(w_f) + \cdots, \tag{3.10}$$

where

$$u_1(w_f) = \begin{cases} u(\psi(w_f - \overline{w})) - u(0) & \text{if } w_f - \overline{w} \ge 0, \\ 0 & \text{if } w_f - \overline{w} \le 0. \end{cases} \tag{3.11}$$

3.2. Duality result

Along these lines, we define a general K-moment utility function and a corresponding indirect utility function. To simplify the notations, we first introduce a K-inner product $\langle \cdot, \cdot \rangle_K : \mathbb{R}^K \times \mathbb{R}^K \longrightarrow \mathbb{R}$ defined as

$$\langle \mu, m \rangle_K = \sum_{k \in K} (-1)^{k-1} \mu_k m_k \tag{3.12}$$

for all $(\mu, m) \in \mathbb{R}^K \times \mathbb{R}^K$. As an example, in the MVS case, we have $\langle \mu, m \rangle_{\{1,2,3\}} = \mu_1 m_1 - \mu_2 m_2 + \mu_3 m_3$.

Definition 3.1. For all $\lambda \in [0,1]$, letting $\mu \in \mathbb{R}_+^K$, the function $U_{K,\lambda,\mu} : \mathfrak{I} \longrightarrow \mathbb{R}$ defined as

$$U_{K,\lambda,\mu}(x) = \langle \mu, m_{K,\lambda}(x) \rangle_K$$

[19] Jondeau and Rockinger (2006, p. 34) formulate a condition for a smooth convergence of the Taylor series expansion such that any additional moment improves the quality of the approximation: it boils down to imposing that the preference weighted odd central moments should not be dominated by consecutive preference weighted even central moments. Nothing is known about the plausibility of this condition in empirical distributions.

is called the general K-moment utility function of level λ. The function $V_{K,\lambda}:\mathbb{R}^K_+\longrightarrow\mathbb{R}$ defined as

$$V_{K,\lambda}(\mu)=\sup\left\{U_{K,\lambda,\mu}(x):\sum_{i=1,\dots,n}x_i=1,x\geq 0\right\}$$

is called the indirect general K-moment utility function of level λ.

For instance, the usual MV utility function can then be written $U_{\{1,2\},1/2,\mu}(x)=\langle\mu,m_{\{1,2\},1/2}(x)\rangle_{\{1,2\}}=\mu_1 m_{1,1/2}(x)-\mu_2 m_{2,1/2}(x)=\mu_1 E[R(x)]-\mu_2 Var[R(x)]$.

Caballé and Pomansky (1996) defined the n-th order index of absolute risk aversion, a generalization of the Arrow-Pratt absolute risk aversion index to higher orders, as follows: $A_k(\overline{w})=u^{(k+1)}(\overline{w})/u^{(k)}(\overline{w})$ for $k=1,2,\dots$. In the context of the indirect general K-moment utility function, the ratios $A_1(\overline{w})=\mu_2/\mu_1\geq 0$, $A_2(\overline{w})=\mu_3/\mu_2\geq 0$, and $A_3(\overline{w})=\mu_4/\mu_3\geq 0$ represent the degree of absolute risk-aversion, prudence, respectively, temperance.[20] Therefore, the maximum value function for the decision maker is simply determined for a given vector of risk parameters $\mu>0$. Knowledge of these parameters allows normally selecting a unique efficient portfolio among those on the weakly efficient frontier maximizing the decision maker's direct general K-moment utility function.

The next result is useful to highlight the role of convexity in duality:

Lemma 3.2. *For all $K\subset\mathbb{N}^*$ and all $\mu\geq 0$, if either C 1 or C 2 holds, then $U_{K,\lambda,\mu}$ is concave and $V_{K,\lambda}$ is convex.*

This is generally not the case whenever there is some uneven moment included ($k\in 2\mathbb{N}+1$).

Before establishing duality relations in our framework, it is first useful to make a distinction between overall, allocative, and portfolio efficiency when evaluating portfolio performance. Similar to analogous distinctions in micro-economics (see Cornes, 1992), the next definition clearly separates these concepts from one another.

Definition 3.3. For all $\lambda\in[0,1]$, the Overall Efficiency ($OE_{K,\lambda}$) index is the quantity:

$$OE_{K,\lambda}(x,\mu;g)=\sup\{\delta:\langle\mu,m_{K,\lambda}(x)+\delta g\rangle_K\leq V_{K,\lambda}(\mu)\}.$$

The allocative efficiency ($AE_{K,\lambda}$) index is the quantity:

$$AE_{K,\lambda}(x,\mu;g)=OE_{K,\lambda}(x,\mu;g)-S_{K,\lambda}(x;g).$$

The Portfolio Efficiency ($PE_{K,\lambda}$) index is the quantity: $PE_{K,\lambda}(x;g)=S_{K,\lambda}(x;g)$.

Portfolio efficiency ensures portfolios are situated on the possibly non-convex boundary of the portfolio frontier. Such points need not maximize the indirect general K-moment utility function. *Allocative efficiency* indicates the necessary adjustment along the boundary of efficient portfolios to achieve the point maximizing the indirect general K-moment utility function. *Overall efficiency* requires the simultaneous achievement of both these objectives. More precisely, $OE_{K,\lambda}$ is the ratio of (i) the difference between indirect general K-moment utility function (Definition 3.1) and the value of the direct general K-moment utility function for the evaluated portfolio, and (ii) the normalized value of the direction vector g for given parameter vector μ:

$$OE_{K,\lambda}(x,\mu;g)=\frac{V_{K,\lambda}(\mu)-U_{K,\lambda,\mu}(x)}{\langle\mu,g\rangle_K}. \tag{3.13}$$

Obviously, these definitions lead to the following additive decomposition identity:

$$OE_{K,\lambda}(x,\mu;g)=AE_{K,\lambda}(x,\mu;g)+PE_{K,\lambda}(x;g). \tag{3.14}$$

Luenberger (1995) has proven a duality result between the expenditure function and the shortage function. In an analogous way, the following result establishes that the shortage function can be computed over the dual of the K-moment space. Since the representation set $\mathcal{DR}_{K,\lambda}$ is generally non-convex when $\lambda\in\{\frac{1}{2},1\}$ and thereby incompatible with a dual representation, we first consider the special case where the set K only contains mean and even moments to ensure convexity of the disposal representation set. When $K\subset\{1\}\cup 2\mathbb{N}^*$ (i.e., portfolios combining mean and even moments only), then clearly the support function of the representation set $\mathcal{DR}_{K,\lambda}$ is the indirect general K-moment utility function ($V_{K,\lambda}$). If $\lambda=0$, then the disposal representation set is constructed from the lower partial moments and is convex. However, if $\lambda=1$, then the disposal representation set is constructed from upper partial moments and it is not convex. These properties have already been summarized in Proposition 2.6 above. One can then establish the following property:

Proposition 3.4. *If either C 1 or C 2 holds, then*:

$$\mathcal{DR}_{K,\lambda}=\bigcap_{\mu\in\mathbb{R}^K_+}\{m\in\mathbb{R}^K:\langle\mu,m\rangle_K\leq V_{K,\lambda}(\mu)\}.$$

[20] Notice that for $k\in K$ and $k\geq 4$, the ratio $A_k(\overline{w})=\mu_k/\mu_{k-1}\geq 0$ sometimes continues to be labeled as the degree of absolute temperance. Thus, ratios based on fifth and higher order moment parameters of this indirect utility function are in general no longer further differentiated. Though, one should add that Eeckhoudt and Schlesinger (2006) mention the notion of edginess related to the fifth derivative of the EU function.

Now we can formulate a duality result between the shortage function and the indirect general K-moment utility function. This shows that the shortage function is economically meaningful, because it can represent behavior compatible with a general K-moment utility function.

Proposition 3.5. *If either C 1 or C 2 holds and* $g \neq 0$, *then the shortage function* $S_{K,\lambda}$ *has the following properties*:

(a) *For all* $x \in \Im$: $S_{K,\lambda}(x;g) = \inf_{\mu \geq 0}\{V_{K,\lambda}(\mu) - U_{K,\lambda,\mu}(x) : \langle \mu, g \rangle_K = 1\}$.
(b) *For all* $\mu \in \mathbb{R}^K_+$: $V_{K,\lambda}(\mu) = \sup_{x \in \Im}\{U_{K,\lambda,\mu}(x) - S_{K,\lambda}(x;g)\langle \mu, g \rangle_K\}$.

The shortage function guarantees that any portfolio is projected onto the weakly efficient portfolio frontier. A slightly weaker result is available in utility space in the convex case. For a given weakly efficient portfolio $x \in \Theta_{K,\lambda}(\Im)$ (i.e., $S_{K,\lambda}(x;g) = 0$), the optimal value of the indirect utility function can be achieved by a utility function with a proper choice of risk parameters (μ). This corollary is a direct consequence of the preceding duality result.

Corollary 3.6. *Assume that a portfolio x is weakly efficient, i.e.,* $x \in \Theta_{K,\lambda}(\Im)$, *if either C1 or C2 holds, then there exists a general K-moment utility function of level* λ ($U_{K,\lambda,\mu}$) *such that* $U_{K,\lambda,\mu}(x) = V_{K,\lambda}(\mu)$.

Next, to handle the general non-convex case, we introduce what we term the hyper-shortage function.[21] This is a kind of concave regularized version of the shortage function in K-moment space.

Definition 3.7. The function $\overline{S}_{K,\lambda} : \Im \times (\mathcal{C}_{K,+} \backslash \{0\}) \longrightarrow \mathbb{R}_+$ defined as

$$\overline{S}_{K,\lambda}(x;g) = \inf_{\mu \geq 0}\{V_{K,\lambda}(\mu) - U_{K,\lambda,\mu}(x) : \langle \mu, g \rangle_K = 1\}$$

is the hyper-shortage function for portfolio x of level λ in the direction of vector g.

Thus, by definition, this hyper-shortage function is dual to the indirect general K-moment utility function. This complements the above duality result that is limited to the convex case only.

The preceding duality result (Proposition 3.5) combined with the above definition of the hyper-shortage function have now the following immediate consequence:

Corollary 3.8. *If either C*1 or *C*2 *holds, then*: $S_{K,\lambda} = \overline{S}_{K,\lambda}$.

Thus, when limiting attention to even moments only or to lower partial moments only, then the shortage function equals the hyper-shortage function. However, in general the shortage function does not equal the hyper-shortage function. For instance, if K is not a subset of $\{1\} \cup 2\mathbb{N}^*$, then $S_{K,\lambda} \neq \overline{S}_{K,\lambda}$, because in such a case $\mathcal{DR}_{K,\lambda}$ is not convex.

To measure the impact of convexity related to the inclusion of odd moments on portfolio performance, we introduce the *convexity efficiency* ($CE_{K,\lambda}$) index which is defined as the quantity:

$$CE_{K,\lambda}(x;g) = \overline{S}_{K,\lambda}(x;g) - S_{K,\lambda}(x;g).$$

Convexity efficiency measures the difference between the shortage functions computed on both $\mathcal{DR}_{K,\lambda}$ and its convex hull. In the decomposition specified above for the convex case solely, it appears as a part of $AE_{K,\lambda}$. Clearly, for cases with even moments only or lower partial moments only this convexity efficiency component is zero: if either $C1$ or C_2 holds, then $CE_{K,\lambda}(x;g) = 0$. But, in general, the above decomposition (3.14) can be extended by adding a non-zero convexity efficiency component (that disappears in the convex case).

Fig. 1 illustrates the basic OE decomposition for the MV model. The shortage function seeks to improve a given portfolio in the direction of both an increased mean return and a reduced risk. For example, let us focus on an inefficient portfolio denoted by point A. This portfolio A is projected onto the weak efficient frontier at point B. However, this point is not optimal in view of investor preferences, while point C does maximize the MV utility function. Assuming for simplicity that $\|g\| = \|(g_{Var}, g_E)\| = 1$ (where $\|.\|$ is the usual Euclidean metric), one can straightforwardly see that $OE_{K,\lambda}(.) = \|CA\|$, $PE_{K,\lambda}(.) = \|BA\|$ and $AE_{K,\lambda}(.) = \|CB\|$. Notice that $CE_{K,\lambda}(.) = 0$ in this convex MV case.

Fig. 2 illustrates another part of the OE decomposition for an MVS model. We use a small sample of 35 assets that are part of the French CAC40 index observed between February 1997 and October 1999 (see Briec et al., 2007 for details). For a skewness level of 2.49, we have generated a section of the frontier in the MV subspace. The resulting dot plot is clearly somewhat non-convex.[22] A solid line has been added to convexify this empirically derived frontier section. Let us now focus on a fictitious inefficient portfolio denoted by the point A that is superimposed on this empirical frontier section derived from these 35 assets. It is easy to observe that the shortage function and the hyper-shortage functions project point A onto two different boundaries (compare point B on the non-convex section and point C on the convexified section). The

[21] This parallels in a portfolio context the hyper-benefit function of Luenberger (1992) in consumer theory.

[22] Notice that in principle, the shortage function approach can be used to reconstruct MVS portfolio frontiers (and beyond) using either 2D or 3D grids based upon the empirical domain of mean, variance and skewness of the basic set of observed portfolios (x^j). Apart from Remark 2.5, technical issues surrounding such reconstruction are topics of currently ongoing research.

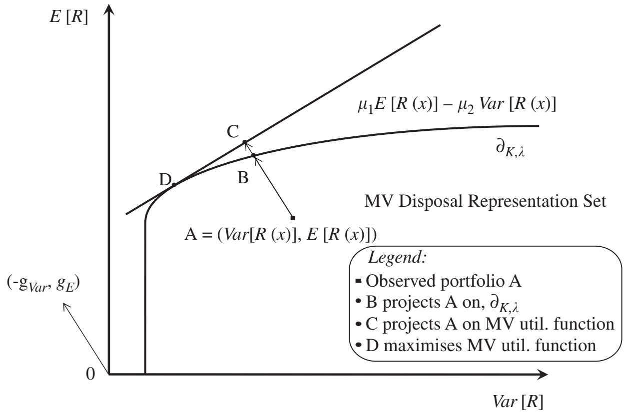

Fig. 1. *Shortage function and OE decomposition in MV model.* In this figure we illustrate the case where $K=\{1,2\}$ and $\lambda=\frac{1}{2}$. An inefficient portfolio (point A) is projected onto the weak efficient frontier (denoted $\partial_{K,\lambda}$) at point B. Point B is not optimal for given investor preferences. Point D maximizes the MV utility function (point C yields same level). Notice that the disposal representation set $\mathcal{DR}_{K,\lambda}$ is convex.

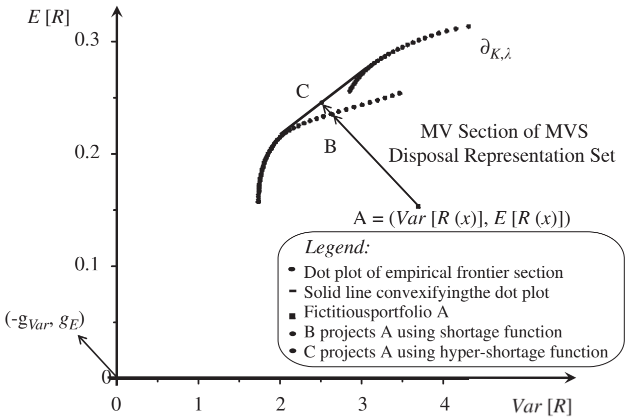

Fig. 2. *MV section of the MVS disposal representation set: shortage function and hyper-shortage function*. Sample of 35 assets (part of French CAC40 index): starting from a MVS model and given a skewness level of 2.49, a section of the frontier in the MV subspace is generated. In this case we have $K=\{1,2,3\}$ and $\lambda=\frac{1}{2}$. One can see that the disposal representation set $\mathcal{DR}_{K,\lambda}$ is not convex. The solid line convexifies this MV section of the MVS disposal representation set. A fictitious inefficient portfolio (point A) is superimposed on this frontier section. The shortage and hyper-shortage functions project point A onto 2 different boundaries: point B (C) on the non-convex (convexified) section. The difference between both projections reflects a positive convexity efficiency for this point A.

difference between both projections reflects the existence of a positive Convexity Efficiency component for this particular fictitious point A.

We conclude this subsection with two remarks: one on the nature of the dual solutions; and another one on the development of tests regarding the influence of additional moments.

First, computed solutions for the general K-moment utility function only guarantee local optimality (in contrast to the shortage function approach). Under certain circumstances, it is possible to infer the nature of the solution obtained for the general K-moment utility function (see Briec et al., 2007 for details). Furthermore, we are unaware of a way to compute the hyper-shortage function, which precludes netting out Convexity Efficiency from $AE_{K,\lambda}$ in general. Under certain circumstances, it is possible to check whether the Convexity Efficiency equals zero or is larger than zero, though its exact value remains unknown (see Briec et al., 2007).

Second, in contrast to the primal approach based upon the shortage function, it is difficult to assess the usefulness of additional moments in the approximation of the EU function using the above indirect utility approach. The basic problem is that the successive maximum values of the value function of the general K-moment utility function (hence, the values of the $OE_{K,\lambda}$ components) cannot be a priori ordered when moments are added, since these value functions depend on the specific risk parameters postulated for the added moments. Therefore, the shortage function approach may well be a more promising way forward to develop proper statistical test procedures.

Recently, utility maximization for portfolio choice has again become the focus of research via the development of the method of Full-Scale Optimization (FSO) (see, e.g., Adler and Kritzman, 2006; Gourieroux and Monfort, 2005; Maringer and Parpas, 2009; Sharpe, 2007). But,this development also exemplifies the above mentioned problems. Empirical return distributions (or theoretical distributions, or combinations of both) are used in their entirety, and the choice of utility function is completely flexible (uncompromised by mathematical convenience, but in practice limited up to at most four moments), portfolio allocations are then optimized to maximize a variety of utility functions and the best fitting utility maximizing portfolio is retained. However, this absence of simplifying assumptions—which yields its theoretical appeal—implies an enormous computational cost: since the optimization problem is non-convex, one needs a gradient-based technique (assuming a single maximum exists) or a grid search. For larger problems with many assets (common in portfolio optimisation practice), heuristic search algorithms (e.g., differential evolution (DE)) can offer a way out. However, because the objective function is non-concave, it is impossible to guarantee global optimality in such a dual approach and one can at best only verify whether conditions of local optimality are satisfied.

3.3. *Results assuming differentiability*

This subsection studies properties of the shortage function that presume differentiability at the point where the function is evaluated. For this purpose, the *adjusted K-moment risk characteristics correspondence of level* λ, $\mu_{K,\lambda}:\mathfrak{I}\times(\mathcal{C}_{K,+}\backslash\{0\})\longrightarrow 2^{[\mathbb{R}^K]}$ is introduced for all $\lambda\in[0,1]$:

$$\mu_{K,\lambda}(x;g)=\underset{\mu}{\operatorname{argmin}}\{V_{K,\lambda}(\mu)-U_{K,\lambda,\mu}(x):\langle\mu,g\rangle_K=1,\mu\geq 0\}. \tag{3.15}$$

Notice that $2^{[\mathbb{R}^K]}$ is the collection of all the subsets (power set) of $\mathbb{R}^K$. In the remainder, $\mu_{k,\lambda}(x;g)$ denotes the k-th component of $\mu_{K,\lambda}(x;g)$. This function implicitly characterizes the agent's risk aversion, prudence, temperance, etc.[23] The fact that, at least in principle, absolute risk aversion, prudence, temperance, and other risk characteristics can be revealed using this adjusted K-moment risk characteristics function expands the possibilities to directly optimize the K-moment approximation of EU based on more realistic parameters.

The next result shows that the hyper-shortage function increases in the odd moments and decreases in the even moments.

Proposition 3.9. *Let $\overline{S}_{K,\lambda}$ be the hyper-shortage function defined on $\mathfrak{I}$. At the points where $\overline{S}_{K,\lambda}$ is differentiable in x, for all $k\in K$:* $\partial\overline{S}_{K,\lambda}(x;g)/m_{k,\lambda}(x;g)|_{m_{k',\lambda}(x;g)=Ct,k'\neq k}=(-1)^{k-1}\mu_k(x;g)$.

Unfortunately, we are unaware of a way to compute the hyper-shortage function and thus to obtain these shadow prices. However, in the convex case (i.e., if either $C1$ or $C2$ holds), then following Corollary 3.8 we have $\overline{S}_{K,\lambda}=S_{K,\lambda}$, and consequently: $\partial S_{K,\lambda}(x;g)/m_{k,\lambda}(x,g)|_{m_{k',\lambda}(x;g)=Ct,k'\neq k}=(-1)^{k-1}\mu_{k,\lambda}(x;g)$. When limiting attention to even moments or to lower partial moments, this result shows that changes of the shortage function with respect to x are identical to the variation of the indirect utility function, computed with respect to the adjusted K-moment risk characteristics function. Furthermore, this same variation can be linked to the moment matrices of each asset.

In fact, under some regularity conditions and assuming that either $C1$ or $C2$ holds, the adjusted K-moment risk characteristics function can be obtained from the Kuhn–Tucker multipliers of the mathematical program $(P_{K,\lambda})$ computing the shortage function. This is demonstrated in the next proposition.

Proposition 3.10. *Assume that either $C1$ or $C2$ holds. Let $j\in\{1\cdots m\}$ such that that program $(P_{K,\lambda})$ has a regular optimal solution. For $k\in K$, let $\eta_k\geq 0$ be the Kuhn–Tucker multipliers of the first $|K|$ constraints in $(P_{K,\lambda})$. If the shortage function is differentiable at point x^j, then*:

(a) *We have*:

$$\left.\frac{\partial S_{K,\lambda}(x;g)}{\partial m_{k,\lambda}(x)}\right|_{\substack{x=x^j\\ m_{k',\lambda}(x;g)=m_{k',\lambda}(x^j;g)\\ k\neq k'}}=(-1)^{k-1}\eta_k,$$

(b) *The adjusted K-moment risk characteristics correspondence of level λ is single valued and identical to the Kuhn–Tucker multipliers*: $\mu_{k,\lambda}(x;g)=\eta_k$.

In line with Corollary 3.6, in the convex case and assuming minimal regularity properties, it is possible for any portfolio x^j to achieve the indirect utility function using the optimal solution (δ^*,x^*) of the program $(P_{K,\lambda})$ and the optimal risk characteristics $(\mu(x^j;g))$ determined by the adjusted K-moment risk characteristics function.

[23] This name is inspired by the adjusted price function in consumer theory (see Luenberger, 1995).

Proposition 3.11. *For a portfolio x^j, $j \in \{1, \cdots, m\}$ and under either C 1 or C 2, if (δ^*, x^*) is a solution to program $(P_{K,\lambda})$ and assuming the adjusted K-moment risk characteristics function of level λ is single valued at x^j, then*: $U_{K,\lambda,\mu^*}(x^*) = V_{K,\lambda}(\mu^*)$, *where* $\mu^* = \mu_{K,\lambda}(x^j; g)$.

It is also possible to link the adjusted K-moment risk characteristics function and some kind of Marshalian demand for each asset. First, let us introduce the matrix of derivatives:

$$\mathcal{B}_{k,i}(\lambda) = \left[\frac{\partial m_{k,\lambda}(x)}{\partial x_i}\right]_{k,i}. \tag{3.16}$$

Moreover, assuming the maximum is unique and given a vector of risk characteristics we define a "Marshalian" demand for assets by

$$x_{K,\lambda}(\mu) = \text{argmax}\{U_{K,\lambda,\mu}(x) : x \in \Im\}. \tag{3.17}$$

At points where this Marshalian demand is single valued and differentiable in μ, one can then define some kind of Slutsky substitution matrix:

$$\mathcal{S}_{i,k}(\lambda) = \left[\frac{\partial x_{K,\lambda,i}(\mu_k)}{\partial \mu_k}\right]_{i,k}. \tag{3.18}$$

As shown in the next proposition, this Slutsky matrix can be linked to the matrix B.

Proposition 3.12. *Let $\overline{S}_{K,\lambda}$ be the hyper-shortage function. Let $\mathcal{D}(K)$ be the $|K| \times |K|$ diagonal matrix defined by $\mathcal{D}_{k,k}(K) = (-1)^{k-1}$. At the points where $\overline{S}_{K,\lambda}$ is differentiable in x, we have*:

(a) $$\mathcal{B}(\lambda)\mathcal{S}(\lambda) = \frac{1}{\langle \mu, g \rangle_K} I - \frac{1}{(\langle \mu, g \rangle_K)^2} \mu \times g\mathcal{D}(K);$$

(b) $$\mathcal{B}(\lambda)^T \mathcal{S}(\lambda)^T = \frac{1}{\langle \mu, g \rangle_K} I - \frac{1}{(\langle \mu, g \rangle_K)^2} \mathcal{D}(K) g \times \mu;$$

(c) $$\mathcal{B}(\lambda)\mathcal{B}(\lambda)^+ = I - \frac{1}{(\langle \mu, g \rangle_K)^2} \mathcal{D}(K) g \times g\mathcal{D}(K).$$

Obviously, if either $C1$ or $C2$ holds, then the above result holds at points where $S_{K,\lambda}$ is differentiable in x.

4. Empirical illustration

To illustrate the feasibility of this new approach, we compute the decomposition of $OE_{K,\frac{1}{2}}$ for a small sample of 30 "blue chips" stocks quoted on the London Stock Exchange between January 1990 and May 2001. The sample contains 2874 observations on continuously compounded total rates of returns for all assets.

For reasons of space, the empirical analysis is limited to an empirical analysis focusing on the first four centered moments. To show the flexibility of this new approach, we contrast two basic models: a mean–variance (MV) model, and a mean–kurtosis (MK) model. To illustrate approximating the indirect EU function to the 4-th order, we add to both these models the skewness and kurtosis, respectively, the variance and skewness to test for the impact of including additional moments in different sequences.

Summarizing the computational procedure, we start with solving the program $(P_{K,1/2})$ to obtain $PE_{K,1/2}$.[24] Then, solving the mathematical program corresponding to maximizing the direct 4-th order moment EU function over the set $\mathcal{DR}_{K,\lambda}$ with parameters $\mu_1 = 1$, $\mu_2 = 1.5$, $\mu_3 = 2$, and $\mu_4 = 2$ yields the indirect general K-moment EU function in Definition 3.1. These parameters fix absolute risk aversion $(A_1(\overline{w}) = 1.5)$, prudence $(A_2(\overline{w}) = 1.33)$ and temperance $(A_3(\overline{w}) = 1)$ around conventional values (though one should realize that little is known on especially the latter parameters). Finally, applying the decomposition in Definition 3.3 and using (3.13) leads to the decomposition results in Tables 1 and 2 for the MV, respectively, the MK models.

The first part of Table 1 summarises the basic MV results, while the second part describes the impact of adding the third and fourth moments in various sequences. The first part of Table 2 starts from a MK model and the second part of this same table checks the contribution of second and third moments in various orders. Comparing the sample averages for $PE_{K,1/2}$ in the lower part of Tables 1 and 2 and recalling the goodness-of-fit interpretation of the shortage function, one observes that the MV model fits the sample data better than the MK model. This could indicate that a MK model as such would not be a good substitute for the traditional MV model. Starting from these two models, it is clear that adding skewness, respectively, skewness and variance makes a difference, while adding kurtosis to the first model has an ignorable impact. The first line in Table 3 contains the results for the Li (1996) test statistic confirming these conclusions for Tables 1 and 2.

[24] A technical issue when computing $(P_{K,1/2})$ is the choice of direction vector. To obtain a proportional interpretation, the direction vector equals the moments of the evaluated asset (i.e., $g = \|m_{k,\lambda}(x^j)\|$).

Table 1
Mean–variance model: impact of adding up to four moments.

		Mean–variance			Impact measures			
		OE	*AE*	*PE*	$I(\{3\},\{1,2\})$[a]	$I(\{4\},\{1,2\})$	$I(\{4\},\{1,2,3\})$	$I(\{3\},\{1,2,4\})$
1	Assd.Brit.Foods	0.782	0.032	0.750	0.750	0.000	0.000	0.750
2	Allied Domecq	0.780	0.000	0.780	0.780	0.000	0.000	0.780
3	Abbey National	0.804	0.793	0.011	0.011	0.002	0.000	0.009
4	Bae Systems	0.899	0.066	0.834	0.000	0.000	0.000	0.000
5	Baa	0.712	0.269	0.443	0.000	0.000	0.000	0.000
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
25	Imp.Chm.Inds.	0.819	0.005	0.813	0.060	0.000	0.000	0.060
26	Invensys	0.879	0.097	0.782	0.000	0.000	0.000	0.000
27	Kingfisher	0.815	0.235	0.580	0.013	0.000	0.000	0.013
28	Land Securities	0.571	0.002	0.569	0.109	0.000	0.002	0.110
29	Legal General	0.825	0.385	0.440	0.056	0.000	0.000	0.056
30	Marks Spencer Group	0.817	0.001	0.816	0.794	0.000	0.000	0.794
	Mean	0.804	0.174	0.631	0.328	0.000	0.000	0.328
	St. Dev.	0.067	0.251	0.252	0.367	0.000	0.000	0.367
	Max	0.899	0.885	0.885	0.883	0.002	0.002	0.883

[a] Notation has been simplified to save space.

Table 2
Mean–kurtosis model: impact of adding up to four moments.

		Mean–Kurtosis			Impact measures			
		OE	*AE*	*PE*	$I(\{3\},\{1,4\})$[a]	$I(\{2\},\{1,4\})$	$I(\{2\},\{1,3,4\})$	$I(\{3\},\{1,2,4\})$
1	Assd.Brit.Foods	0.981	0.015	0.966	0.966	0.216	0.000	0.750
2	Allied Domecq	0.981	0.001	0.980	0.980	0.200	0.000	0.780
3	Abbey National	0.976	0.967	0.009	0.009	0.000	0.000	0.009
4	Bae Systems	1.000	0.002	0.998	0.000	0.164	0.164	0.000
5	Baa	0.981	0.304	0.678	0.000	0.235	0.235	0.000
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
25	Imp.Chm.Inds.	0.982	0.000	0.982	0.047	0.169	0.182	0.060
26	Invensys	0.999	0.007	0.992	0.000	0.209	0.209	0.000
27	Kingfisher	0.983	0.256	0.727	0.001	0.147	0.159	0.013
28	Land Securities	0.845	0.015	0.830	0.297	0.260	0.074	0.110
29	Legal General	0.983	0.489	0.493	0.051	0.053	0.057	0.056
30	Marks Spencer Group	0.987	0.000	0.987	0.965	0.171	0.000	0.794
	Mean	0.978	0.194	0.784	0.406	0.154	0.075	0.328
	St. Dev.	0.027	0.310	0.310	0.458	0.077	0.091	0.367
	Max	1.000	0.996	0.998	0.993	0.266	0.266	0.883

[a] Notation has been simplified to save space.

The lower part adds some additional transitions between models that have been tested for. For instance, adding an even moment to a portfolio model containing first, third and another even moment does not seem to add much value.

Now focusing attention to $OE_{K,1/2}$, it is clear that in both basic models $PE_{K,1/2}$ is the dominant source of inefficiency, while $AE_{K,1/2}$ is secondary in importance for the postulated parameters. Actually, quiet a few individual assets lead to portfolio projections rather close to the optimal risk characteristics postulated, resulting in near-zero values for their $AE_{K,1/2}$ score.

The distribution of optimal portfolio weights in the shortage function approaches is reported in Table 4 in a condensed format. For each model variation, one finds the number of average non-zero weights, and the mean and standard deviation of these portfolio weights. Comparing MV and MK results first, one observes that the former implies a higher diversification with on average lower weights and less dispersion among weights. Adding the skewness dimensions always leads to fewer non-zero weights, resulting in higher average and more dispersed weights. Extending the MK model with a variance dimension increases the number of non-zero weights, lowers average weights, but increases its dispersion.

While this analysis has so far been limited to average results at the sample level, we now discuss some results at the individual level. In addition to the average results for the MV and MK models and their sequentially added moments,

Table 3
Li (1996) test statistic for various portfolio models: results.

From/To model	MVS	MVK	MVSK	From/To model	MKS	MKV	MVSK
MV	6.695[a]	−7.45E − 06	6.694[a]	MK	7.638[a]	10.272[a]	13.104[a]
MVS		6.688[a]	0.000	MKS		9.757[a]	0.6735
MVK			6.687[a]	MKV			6.687[a]

[a] Test statistic significant at 1% level.

Table 4
Optimal portfolio weights.

K	# Non-0 weights[a]	Avg. weight[a]	St. dev. weight
$\{1,2\}$	10.541	0.095	0.209
$\{1,4\}$	7.550	0.132	0.217
$\{1,2,3\}$	3.118	0.321	0.428
$\{1,2,4\}$	10.541	0.095	0.404
$\{1,3,4\}$	2.628	0.380	0.404
$\{1,2,3,4\}$	3.118	0.321	0.428

[a] Geometric mean.

Tables 1 and 2 equally list detailed individual results for part of the sample (to save some space).[25] To develop some intuition, in the first example we briefly comment how the portfolio performance changes depend on the added moments when the starting point of the portfolio optimisation is a single asset. This asset can be arbitrary chosen or it may reflect certain moment characteristics valued particularly by the investor. We also briefly comment on the $OE_{K,1/2}$ decomposition.

Example 4.1. In the MV model, the single asset "Abbey National" obtains a $PE_{K,1/2}$ of 0.011 implying that mean return can be improved by 1.1% and its risk can be reduced by the same amount. In the MK model, its $PE_{K,1/2}$ equals 0.009 meaning that mean return could only be improved by 0.9% and kurtosis diminished by an equal percentage.

Extending the MV model with the skewness leads to an impact measure of $I_{1/2}(\{3\},\{1,2\},x;g)$ = 0.011, meaning that the difference between $PE_{K,1/2}$ under MV and MVS models equals 0.011. Thus, in the MVS model this asset must be efficient ($I_{1/2}(\{3\},\{1,2\},x;g) = 0.011 = 0.011 - 0$). Extending the same MV model with the kurtosis yields an impact measure of $I_{1/2}(\{4\},\{1,2\},x;g) = 0.002$. Applying the same reasoning one can infer that its $PE_{K,1/2}$ in the MV–kurtosis (MVK) model must be around 0.009, which is only marginally below the MV efficiency score. Thus, adding skewness to the MV model has a larger impact than adding kurtosis.

Adding now a fourth moment to the previous three dimensional models results in the following impact measures: $I_{1/2}(\{4\}\{1,2,3\},x;g) = 0.000$ for the MVS model, and $I_{1/2}(\{3\},\{1,2,4\},x;g) = 0.009$ for the MVK model. This means that adding kurtosis to a MVS model adds nothing (i.e., the asset remains efficient), while adding skewness to a MVK model makes the asset become efficient. A similar reasoning can be applied starting from the MK model when interpreting the last four columns in Table 2.

Looking at the efficiency decomposition for the MV model for the same asset, it is clear that only about 1.1% of its poor performance is due to $PE_{K,1/2}$ (i.e., operating below the MV portfolio frontier), while the remaining 79.3% of the performance gap is due to $AE_{K,1/2}$ (i.e., choosing a wrong mix of return and risk given the postulated risk aversion parameters). This adds up to an $OE_{K,1/2}$ performance gap of 80.4%. In the same vein, it is possible to interpret the decomposition results for the MK model when interpreting the last four columns in Table 2. Of course, it is useful to reiterate that $OE_{K,1/2}$ depends on the specification of risk characteristics about which little is known.

For each of these portfolio models, one can of course obtain information on the optimal portfolio return, risk, skewness and kurtosis at the frontier. To clarify this issue, the second example mentions the frontier projections and their resulting moment characteristics for a selection of different portfolio models.

Example 4.2. The asset "Land Securities" has a return of 0.0177, a risk of 1.4653, a skewness of 0.3046, and a kurtosis of 10.6128. Determining an optimal portfolio using this asset as a starting point yields the results listed in Table 5. The MV model leads to an increase in return and reduction in risk that is rather substantial. However, when we add the skewness,

[25] The complete Tables 1 and 2 are in Appendix B.

Table 5
Optimal portfolio characteristics in different models for "Land Securities".

	Return	Risk	Skewness	Kurtosis	# Non-0 weights	Avg. weight
Initial situation	0.0177	1.4653	0.3046	10.6128		
MV	0.0278	0.6309	0.0380	1.7860	22.000	0.045
MVS	0.0272	0.7904	0.4449	6.1744	16.000	0.063
MVK	0.0278	0.6309	0.0380	1.7860	22.000	0.045
MVSK	0.0275	0.7926	0.4444	5.7410	16.000	0.062

then it is clear that the optimal results for the MVS model are less spectacular in terms of returns and risk. But, this model manages to increase the skewness relative to the starting point. Thus, it becomes clear that the good performance of the MV model in both return and risk dimensions is due to its neglect of the skewness which had substantially fallen compared to the initial situation. When adding the kurtosis to the MV model, close to nothing happens: the results are almost indistinguishable. Finally, when adding the kurtosis to the previous three dimensional models, one ends up with an even less spectacular improvement in terms of return and risk compared to the MV model, but now both the skewness and kurtosis improve compared to the initial situation. By contrast, the MV model coincidentally ended up with a better kurtosis compared to the starting point, but at the cost of a substantial loss in skewness. Notice furthermore that the number of non-zero weights as well as the mean of these portfolio weights follow patterns close to the ones described before. Obviously, these frontier projection are computed based on the optimal portfolio weights. Details on the optimal portfolio weights are suppressed for reasons of space.

Of course, this is but one sample of financial data, replication studies on different asset classes and markets are needed to verify whether any of these tentative conclusions about the relative fit of the various portfolio models can be corroborated.

5. Conclusions

A general method for benchmarking portfolios in the non-convex K-moment space has been proposed utilizing the shortage function (see Luenberger, 1995). In this higher-order moment portfolio problem, portfolio efficiency is evaluated by simultaneously looking for reductions in even moments and expansions in odd moments. In the finance literature, these moment preferences are traditionally related to the general class of mixed risk aversion utility functions proposed by Caballé and Pomansky (1996). In the convex case, this shortage function is linked via a duality result to a multidimensional moment approximation of a general indirect utility function. This duality result forms the basis to distinguish between portfolio efficiency and allocative efficiency. Under non-convexity, a convexity efficiency component is defined that is related to the difference between the shortage function and the hyper-shortage function, the latter being defined relative to a convexified representation set. An empirical illustration illustrated the computational tractability of this new, general approach for both general and partial moments.

This shortage function framework projects portfolios relative to a non-parametric estimate approximating the unknown true frontier. This shortage function always achieves global optimality, has the interpretation of an efficiency measure gauging the performance of portfolios, and forms a natural basis for testing the impact of additional higher moments in the approximation. An additional virtue is that sound economic interpretations are available thanks to duality with a general, higher order Taylor expansion of the EU function. In contrast, no global optimal solution can be guaranteed for this more traditional indirect utility function approach, that furthermore depends on risk parameters about which little practical knowledge is available.

These results indicate that future research should probably focus on developing portfolio optimization methods using this shortage function framework. Of course, a wide range of potential further improvements can be listed. A major limitation of the current analysis is that lots of statistical issues are ignored. We sketch a few examples of open issues. For one, while it is well-known that the estimation errors in means are more important than errors in variance-covariance matrices, whereby errors in variances weight heavier than errors in covariances (see, e.g., Chopra and Ziemba, 1993), little seems to be known about errors in the estimation of higher order moment matrices. Second, Kim and White (2004) recently raise the issue of the robustness of current ways of computing the higher moments, i.e., skewness and kurtosis, and indicate that results on the skewness and kurtosis of stock market returns are heavily influenced by outliers. Another desirable extension is to develop multiperiod deterministic or stochastic formulations of this shortage function approach (see, e.g., Chellathurai and Draviam, 2007 or Zenios et al., 1998).

Appendix A. Proofs of lemmas and propositions

Proof of Proposition 2.4. (a) If $x \notin \Theta_{K,\lambda}(\Im)$, then there is some $m' \in \mathcal{DR}_{K,\lambda}$ such that $(-1)^{k-1}m' > (-1)^{k-1}m_{k,\lambda}(x)$ for all $k \in K$. Therefore, it is immediate to see that $S_{K,\lambda}(x;g) > 0$. Conversely, if $S_{K,\lambda}(x;g) > 0$, then $m_{K,\lambda}(x) + S_{K,\lambda}(x;g).g \in \mathcal{DR}_{K,\lambda}$. Therefore,

$x \notin \Theta_{K,\lambda}(\mathfrak{I})$. (b) is an immediate consequence of the definition of $\mathcal{DR}_{K,\lambda}$. (c) Since the function $x \mapsto m_{K,\lambda}(x)$ is continuous on $\mathfrak{I}$, using the argument developed in Luenberger (1992), the proof is immediate. □

Proof of Proposition 2.6. Let us prove the first part of the result. Assume that $m, m' \in \mathcal{DR}_{K,\lambda}$ for $\lambda \in \{0,1\}$. There exists $x, x' \in \mathfrak{I}$ such that $m_1 \leq m_{1,\lambda}(x)$, $m'_1 \leq m_{1,\lambda}(x')$ and $m_k \geq m_{k,\lambda}(x)$, $m'_k \geq m_{k,\lambda}(x')$ for all $k \in K$. Let $\theta, \theta' \in [0,1]$ such that $\theta+\theta'=1$. Since $K \subset \{1\} \cup 2\mathbb{N}^*$, the functions $m_k^{\lambda}(\cdot)$ are convex. Therefore, $\theta m_1 + \theta' m'_1 \leq \theta m_{1,\lambda}(x) + \theta' m_{1,\lambda}(x') = m_1(\theta x + \theta' x')$ and $\theta m_k + \theta' m'_k \geq \theta m_{k,\lambda}(x) + \theta' m_{k,\lambda}(x') \geq m_{k,\lambda}(\theta x + \theta' x')$. Hence, $\theta m + \theta' m' \in \mathcal{DR}_{K,\lambda}$. This proves convexity of $\mathcal{DR}_{K,\lambda}$.

Let us prove the last part of the statement ($\lambda = -1$). Suppose that $m_K, m'_K \in \mathcal{DR}_{K,\lambda}$. We need to prove that for all $\theta, \theta' \in [0,1]$ with $\theta+\theta'=1$, we have $\theta m_K + \theta' m'_K \in \mathcal{DR}_{K,\lambda}$. If $m_K, m'_K \in \mathcal{DR}_{K,\lambda}$, then there exists $x, x' \in \mathfrak{I}$ such that $(-1)^{k-1} m_{k,\lambda}(x) \geq (-1)^{k-1} m_k$ and $(-1)^{k-1} m'_{k,\lambda}(x') \geq (-1)^{k-1} m'_k$ for all $k \in K$. Moreover, for all $k \in 2\mathbb{N}^*$, $m_{k,\lambda}(\cdot)$ is convex. Consequently,

$$(-1)^{k-1} m_{k,\lambda}(\theta x + \theta' x') \geq (-1)^{k-1}(\theta m_k + \theta' m'_k).$$

Since for all $k \in 2\mathbb{N}^*+1$, $m_{k,\lambda}(\cdot)$ is concave we deduce that

$$(-1)^{k-1} m_{k,\lambda}(\theta x + \theta' x') \geq (-1)^{k-1}(\theta m_k + \theta' m'_k).$$

Consequently, since $\theta x + \theta' x' \in \mathfrak{I}$ we deduce that $\theta m_K + \theta' m'_K \in \mathcal{DR}_{K,\lambda}$ which ends the proof. □

Proof of Proposition 2.7. Clearly, $m_{k,0}$ is a convex function for all $k \in 2\mathbb{N}^*$. Moreover, it is -by construction- concave for all $k \in \{1\} \cup 2\mathbb{N}^*$. Since for all $k \in \{1\} \cup 2\mathbb{N}^*$, $m_{k,\lambda}$ is convex for $\lambda \in \{1/2, 1\}$, we deduce that, under the assumptions in $C1$ and $C2$, the domain:

$$D_K = \{(\delta, z) \in \mathbb{R}_+^{n+1} : (-1)^{k-1} m_{k,\lambda}(x^j) + \delta g_k \leq (-1)^{k-1} m_{k,\lambda}(z), k \in K, z \in \mathfrak{I}\}$$

is convex. We have $S_{K,\lambda}(x^j; g) = \max\{\delta : (\delta, z) \in D\}$. Hence, if (δ^*, z^*) is a local maximum, then it is global maximum which ends the proof. □

Proof of Proposition 2.10. (a) By construction, $m_{K,\lambda}(x) + S_{K,\lambda} g \in \mathcal{D}_{K,\lambda}$. By definition $\mathcal{D}_{K,\lambda} = \mathcal{M}_{K,\lambda} - \prod_{k \in K} (-1)^{k-1} \mathbb{R}_+$. Consequently, there is some $z \in \mathfrak{I}$ such that $m_{K,\lambda}(z) \geq m_{K,\lambda}(x) + S_{K,\lambda} g$. Hence, $\zeta_{K,\lambda}(x) \neq \emptyset$. (b) Suppose that $z^* \in \zeta_{K,\lambda}(x)$, this implies that $m_{K,\lambda}(z^*) \in \mathcal{D}_{K,\lambda}$. Let us denote

$$A^* = \{z \in \mathfrak{I} : z \succcurlyeq_{K,\lambda} z^*, z \nsim_{K,\lambda} z^*\}.$$

Since for all $k \in K$ $m_{k,\lambda}(z)$ is continuous in z, it follows that A^* is a closed subset of $\mathfrak{I}$. Moreover, since $\mathfrak{I}$ is bounded, A^* is a compact subset of $\mathbb{R}^n$. Let us consider the map $z \mapsto \sum_{k \in K} m_{k,\lambda}(z)$ defined on $\mathfrak{I}$. This map is continuous and, from the compactness of A^*, its achieves its maximum at some $\bar{z}$. By construction, its maximum is strongly efficient and, consequently, $\bar{z}$ is strongly efficient. However, since $\bar{z} \succcurlyeq_{K,\lambda} z^*$, it follows that $m_{K,\lambda}(\bar{z}) \geq m_{K,\lambda}(z^*) = m_{K,\lambda}(x) + S_{K,\lambda}(x; g) g$. Consequently, $\bar{z} \in \zeta_{K,\lambda}(x)$, which ends the proof of (b). (c) Suppose that x^* is not weakly efficient and let us show a contradiction. If x^* is not weakly efficient then, by hypothesis, there exists some $m' \in \mathcal{DR}_{K,\lambda}$ such that $(-1)^{k-1} m'_k > (-1)^{k-1} m_{k,\lambda}(x^*)$ for all $k \in K$. Hence, there exists some $\delta' > 0$ such that $m_{K,\lambda}(x^*) + \delta' g \in \mathcal{DR}_{K,\lambda} \leq m'$. However, for $j = 1 \cdots m$ and for all $k \in K$, we have $(-1)^{k-1} m_{k,\lambda}(x^j) + \delta^* g_k \leq (-1)^{k-1} m_{k,\lambda}(x^*)$. Hence, we deduce that $(-1)^{k-1} m_{k,\lambda}(x^j) + (\delta^* + \delta') g_k \leq (-1)^{k-1} m_{k,\lambda}(x^*) + \delta' g_k$. Since $m_{k,\lambda}(x^*) + \delta' g \in \mathcal{DR}_{K,\lambda}$ and $\delta^* + \delta' > \delta^*$, this contradicts the fact that δ^* is an optimal solution of $(P_{K,\lambda})$. □

Proof of Proposition 2.13. Let D_K be defined as in the proof of Proposition 2.7. We have obviously $D_{K'} \subset D_K$. Since $S_{K,\lambda}(x; g_K) = \max\{\delta : (\delta, y) \in D_{K,\lambda}\}$, we deduce the result. □

Proof of Lemma 3.2. The proof is similar to that of Proposition 2.6. Therefore, it is omitted. □

Proof of Proposition 3.4. We have established that $\mathcal{DR}_{K,\lambda}$ is convex. Since:

$$\mathcal{DR}_{K,\lambda} = \mathcal{DR}_{K,\lambda} - \prod_{k \in K} (-1)^{k-1} \mathbb{R}_+$$

we deduce that if $\mu \notin \mathbb{R}_+^K$, then:

$$\sup\left\{\sum_{k \in K} (-1)^{k-1} \mu_k . m_k : m \in \mathcal{DR}_{K,\lambda}\right\} = +\infty.$$

Now clearly the function $m \to \sum_{k \in K} (-1)^{k-1} \mu_k . m_k$ is linear on $\mathbb{R}^K$. Moreover, $V_{K,\lambda}(\mu) = \sup\{\sum_{k \in K} (-1)^{k-1} \mu_k . m_k : m \in \mathcal{DR}_{K,\lambda}\}$. Therefore, from the convex separation theorem, we deduce the result. □

Proof of Proposition 3.5. The proof is straightforward from Proposition 2.6 and Luenberger (1992). □

Proof of Proposition 3.9. The proof is obtained by the standard envelope theorem. □

Proof of Proposition 3.10. The proof is similar to the one established in Briec et al. (2004). Therefore, it is omitted. □

Proof of Proposition 3.11. Since (δ^*, x^*) is solution of $(P_{K,\lambda})$, we have

$$(-1)^{k-1} m_{k,\lambda}(x^*) \geq (-1)^{k-1}(m_{k,\lambda}(x^j) + \delta^* g_k)$$

for all $k \in K$. Consequently,

$$\sum_{k \in K} \mu_{k,\lambda}(x^j)(-1)^{k-1} m_{k,\lambda}(x^*) \geq \sum_{k \in K} \mu_{k,\lambda}(x^j)(-1)^{k-1}(m_{k,\lambda}(x^j) + \delta^* g_k).$$

Since $\sum_{k \in K} (-1)^{k-1} \mu_{k,\lambda}(x^j) g_k = 1$, we obtain

$$U_{K,\lambda,\mu^*}(x^*) \geq U_{K,\lambda,\mu^*}(x^j) + \delta^*,$$

where $\mu^* = \mu_{K,\lambda}(x^j)$. However, since either $C1$ or $C2$ holds, we have $\delta^* = V_{K,\lambda}(\mu^*) - U_{K,\lambda,\mu^*}(x^j)$. Hence, $U_{K,\lambda,\mu^*}(x^*) \geq V_{K,\lambda}(\mu^*)$. By definition, we have $U_{K,\lambda,\mu^*}(x^*) \leq V_{K,\lambda}(\mu^*)$ which ends the proof. □

Proof of Proposition 3.12. (a) Let $\overline{\mu} = \mu / (\sum_{k \in K} (-1)^{k-1} \mu_k g_k)$. We have

$$\frac{\partial \overline{\mu}_k}{\partial \mu_k} = \sum_{i=1 \cdots n} \frac{\partial \overline{\mu}_k}{\partial x_i} \frac{\partial x_i}{\partial \mu_k} = \frac{1}{\sum_{k \in K} (-1)^{k-1} \mu_k g_k} - \frac{\mu_k (-1)^{k-1} g_k}{\left(\sum_{k \in K} (-1)^{k-1} \mu_k g_k\right)^2}.$$

Moreover, if $k \neq k'$:

$$\frac{\partial \overline{\mu}_k}{\partial \mu_{k'}} = \sum_{i=1 \cdots n} \frac{\partial \overline{\mu}_k}{\partial x_i} \frac{\partial x_i}{\partial \mu_{k'}} = -\frac{\mu_k (-1)^{k'-1} g_{k'}}{\left(\sum_{k \in K} (-1)^{k-1} \mu_k g_k\right)^2}.$$

But:

$$[\mathcal{B}(\lambda)\mathcal{S}(\lambda)]_{k,k'} = \left[\sum_{i=1 \cdots n} \frac{\partial \overline{\mu}_k}{\partial x_i} \frac{\partial x_i}{\partial \mu_{k'}}\right]_{k,k'}$$

and the result holds. (b) is obtained by taking the transpose in (a). (c) This follows by combining (a) and (b). □

Appendix B. Supplementary data

The complete Tables 1 and 2 associated with this article can be found in the online version at doi:10.1016/j.jedc.2009.11.001.

References

Adler, T., Kritzman, M., 2006. Mean–variance versus full-scale optimization in and out of sample. Journal of Asset Management 7, 302–311.
Bacmann, J.F., Benedetti, S.M., 2009. Modeling Returns of Hedge Fund Strategies in a Bayesian Asset Allocation Framework, Mimeo.
Balzer, L., 2001. Investment risk: a unified approach to upside and downside returns. In: Sortino, F., Satchell, S. (Eds.), Managing Downside Risk in Financial Markets. Buttersworth–Heinemann, Oxford, pp. 103–155.
Bawa, V.S., 1975. Optimal rules for ordering uncertain prospects. Journal of Financial Economics 2, 95–121.
Berényi, Z., 2003. Risk and performance evaluation with skewness and kurtosis for conventional and alternative investments. Peter Lang, Frankfurt.
Briec, W., 2000. An extended Färe-Lovell technical efficiency measure. International Journal of Production Economics 65, 191–199.
Briec, W., Kerstens, K., 2009. Infeasibilities and directional distance functions: with application to the determinateness of the Luenberger productivity indicator. Journal of Optimization Theory and Applications 141, 55–73.
Briec, W., Kerstens, K., Lesourd, J.-B., 2004. Single period Markowitz portfolio selection, performance gauging and duality: a variation on the Luenberger shortage function. Journal of Optimization Theory and Applications 120, 1–27.
Briec, W., Kerstens, K., Jokung, O., 2007. Mean–variance–skewness portfolio performance gauging: a general shortage function and dual approach. Management Science 53, 135–149.
Brockett, P.L., Golden, L.L., 1987. A class of utility functions containing all the common utility functions. Management Science 33, 955–964.
Brockett, P.L., Kahane, Y., 1992. Risk, return, skewness and preference. Management Science 38, 851–866.
Caballé, J., Pomansky, A., 1996. Mixed risk aversion. Journal of Economic Theory 71, 485–513.
Chellathurai, T., Draviam, T., 2007. Dynamic portfolio selection with fixed and/or proportional transaction costs using non-singular stochastic optimal control theory. Journal of Economic Dynamics and Control 31, 2168–2195.
Chopra, V.K., Ziemba, W.T., 1993. The effect of errors in means, variances, and covariances on optimal portfolio choice. Journal of Portfolio Management 19, 6–11.
Cornes, R., 1992. Duality and Modern Economics. Cambridge University Press, Cambridge.
Cumova, D., Moreno, D., Nawrocki, D., 2006. The critical line algorithm for UPM–LPM parametric general asset allocation problem with allocation boundaries and linear constraints. In: Gregoriou, G.N. (Ed.), Asset Allocation and International Investments. Palgrave Macmillan, London, pp. 80–95.
Dachraoui, K., Dionne, G., Eeckhoudt, L., Godfroid, P., 2004. Comparative mixed risk aversion: definition and application to self-protection and willingness to pay. Journal of Risk and Uncertainty 29, 261–276.
de Athayde, G., Flôres, R., 2003. Incorporating skewness and kurtosis in portfolio optimization: a multidimensional efficient set. In: Satchell, S., Scowcroft, A. (Eds.), Advances in Portfolio Construction and Implementation. Butterworth–Heinemann, Oxford, pp. 243–257.
de Athayde, G.M., Flôres Jr., R.G., 2004. Finding a maximum skewness portfolio—a general solution to three-moments portfolio choice. Journal of Economic Dynamics and Control 28, 1335–1352.

Dittmar, R., 2002. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross-section of equity returns. Journal of Finance 57, 369–403.
Eeckhoudt, L., Schlesinger, H., 2006. Putting risk in its proper place. American Economic Review 96, 280–289.
Eeckhoudt, L., Rey, B., Schlesinger, H., 2007. A good sign for multivariate risk taking. Management Science 53, 117–124.
Fan, Y., Ullah, A., 1999. On goodness-of-fit tests for weakly dependent processes using Kernel method. Journal of Nonparametric Statistics 11, 337–360.
Färe, R., Grosskopf, S., 1995. Nonparametric tests of regularity, Farrell efficiency and goodness of fit. Journal of Econometrics 69, 415–425.
Farrar, D.E., 1962. The Investment Decision Under Uncertainty. Prentice-Hall, Englewood Cliffs, NJ.
Fishburn, P.C., 1977. Mean-risk analysis with risk associated with below-target returns. American Economic Review 67, 116–126.
Gibbons, M.R., Ross, S.A., Shanken, J., 1989. A test of the efficiency of a given portfolio. Econometrica 57, 1121–1152.
Golec, J., Tamarkin, M., 1998. Bettors love Skewness, not risk, at the horse track. Journal of Political Economy 106, 205–225.
Gotoh, J-Y., Konno, H., 2000. Third degree stochastic dominance and mean-risk analysis. Management Science 46, 289–301.
Gourieroux, C., Monfort, A., 2005. The econometrics of efficient portfolios. Journal of Empirical Finance 12, 1–41.
Gul, F., 1991. A theory of disappointment aversion. Econometrica 59, 667–686.
Harlow, W.V., Rao, R.K.S., 1989. Asset pricing in a generalized mean-lower partial moment framework: theory and evidence. Journal of Financial and Quantitative Analysis 24, 285–311.
Harvey, C.R., Liechty, J.C., Liechty, M.W., Müller, P., 2004. Portfolio selection with higher moments. Duke University, Fuqua School of Business, DP FRPS04-123.
Harvey, C.R., Siddique, A., 2000. Conditional Skewness in Asset Pricing Tests. Journal of Finance 55, 1263–1295.
Hlawitschka, W., 1994. The empirical nature of Taylor-series approximations to expected utility. American Economic Review 84, 713–719.
Holthausen, D.M., 1981. A risk-return model with risk and return measured as deviations from a target return. American Economic Review 67, 116–126.
Jana, P., Roy, T.K., Mazumder, S.K., 2007. Multi-objective mean–variance–skewness model for portfolio optimization. Advanced Modeling and Optimization 9, 181–193.
Jondeau, E., Rockinger, M., 2003. Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. Journal of Economic Dynamics and Control 27, 1699–1737.
Jondeau, E., Rockinger, M., 2006. Optimal portfolio allocation under higher moments. European Financial Management 12, 29–55.
Jurczenko, E., Maillet, B., 2006. Multi-moment Asset Allocation and Pricing Models. Wiley, New York.
Jurczenko, E., Maillet, B., Merlin, P., 2008. Efficient frontier for robust higher-order moment portfolio selection. CES Working Paper.
Jurczenko, E., Yanou, G., 2010. Fund of hedge funds portfolio selection: a robust non-parametric multi-moment approach. In: Watanabe, Y. (Ed.), Emerging Trends in Hedge Fund Strategies. Nova Science, New York, forthcoming.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47, 263–291.
Kim, T-H., White, H., 2004. On more robust estimation of skewness and kurtosis. Finance Research Letters 1, 56–73.
Kimball, M.S., 1990. Precautionary saving in the small and in the large. Econometrica 58, 53–73.
Konno, H., Shirakawa, H., Yamazaki, H., 1993. A mean–absolute deviation–skewness portfolio optimization model. Annals of Operations Research 45, 205–220.
Kraus, A., Litzenberger, R.H., 1976. Skewness preference and the valuation of risky assets. Journal of Finance 31, 1085–1100.
Kumar, S., Russell, R.R., 2002. Technological change, technological catch-up, and capital deepening: relative contributions to growth and convergence. American Economic Review 92, 527–548.
Kuosmanen, T., 2004. Efficient diversification according to stochastic dominance criteria. Management Science 50, 1390–1406.
Lai, T.Y., 1991. Portfolio selection with skewness: a multiple-objective approach. Review of Quantitative Finance and Accounting 1, 293–305.
Levy, H., 2006. Stochastic Dominance: Investment Decision Making under Uncertainty, second ed Springer, Berlin.
Li, Q., 1996. Nonparametric testing of closeness between two unknown distribution functions. Econometric Reviews 15, 261–274.
Loistl, O., 1976. The erroneous approximation of expected utility by a Taylor's series expansion: analytic and computational results. American Economic Review 66, 904–910.
Lozano, S., Guttiérez, E., 2008. TSD—consistent performance assessment of mutual funds. Journal of the Operational Research Society 59, 1352–1362.
Luenberger, D.G., 1984. Linear and Nonlinear Programming, second ed Addison-Wesley, Reading, MA.
Luenberger, D.G., 1992. Benefit function and duality. Journal of Mathematical Economics 21, 461–481.
Luenberger, D.G., 1995. Microeconomic Theory. McGraw-Hill, New York.
Maringer, D., Parpas, P., 2009. Global optimization of higher order moments in portfolio selection. Journal of Global Optimization 41, 219–230.
Markowitz, H., 1952. Portfolio selection. Journal of Finance 7, 77–91.
Markowitz, H., 1991. Foundations of portfolio theory. Journal of Finance 46, 469–478.
Menezes, C.F., Wang, X.H., 2005. Increasing outer risk. Journal of Mathematical Economics 41, 875–886.
Ogryczak, W., Ruszczyński, A., 2001. On consistency of stochastic dominance and mean–semideviation models. Mathematical Programming 89, 217–232.
Rubinstein, M.E., 1973. The fundamental theorem of parameter-preference security valuation. Journal of Financial and Quantitative Analysis 8, 61–69.
Samuelson, P.A., 1967. General proof that diversification pays. Journal of Financial and Quantitative Analysis 2, 1–13.
Samuelson, P.A., 1970. The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. Review of Economic Studies 37, 537–542.
Scott, R.C., Horvath, P.A., 1980. On the direction of preference for moments of higher order than the variance. Journal of Finance 35, 915–919.
Scherer, B., Martin, D., 2005. Introduction to Modern Portfolio Optimization with NUOPT and S-Plus. Springer, New York.
Sharpe, W.F., 2007. Expected utility asset allocation. Financial Analysts Journal 63, 18–30.
Simar, L., Wilson, P.W., 2000. A general methodology for bootstrapping in non-parametric frontier models. Journal of Applied Statistics 27, 779–802.
Tehranian, H., 1980. Empirical studies in portfolio performance using higher degrees of stochastic dominance. Journal of Finance 35, 159–220.
Thistle, P.D., 1993. Negative moments, risk aversion and stochastic dominance. Journal of Financial and Quantitative Analysis 28, 301–311.
Varian, H.R., 1983. Nonparametric tests of models of investor behavior. Journal of Financial and Quantitative Analysis 18, 269–278.
Varian, H.R., 1990. Goodness-of-fit in optimizing models. Journal of Econometrics 46, 125–140.
Vinod, H.D., 2004. Ranking mutual funds using unconventional utility theory and stochastic dominance. Journal of Empirical Finance 11, 353–377.
Zenios, S.A., Holmer, M.R., McKendall, R., Vassiadou-Zeniou, C., 1998. Dynamic models for fixed-income portfolio management under uncertainty. Journal of Economic Dynamics and Control 22, 1517–1541.