Accepted Manuscript

Portfolio analysis with DEA: prior to choosing a model

Tarnaud Albane Christine , Leleu Hervé

PII: S0305-0483(17)30191-3
DOI: 10.1016/j.omega.2017.02.003
Reference: OME 1757

To appear in: *Omega*

Received date: 27 October 2015
Revised date: 20 February 2017
Accepted date: 23 February 2017

Please cite this article as: Tarnaud Albane Christine , Leleu Hervé , Portfolio analysis with DEA: prior to choosing a model, *Omega* (2017), doi: 10.1016/j.omega.2017.02.003

Portfolio analysis with DEA: prior to choosing a model

TARNAUD Albane Christine
IÉSEG School of Management & LEM-CNRS (UMR 9221)
3, rue de la Digue, 59 000 Lille, France, +33 (0)3 20 54 58 92
a.tarnaud@ieseg.fr (corresponding author)

LELEU Hervé
LEM-CNRS (UMR 9221) & IÉSEG School of Management
3, rue de la Digue, 59 000 Lille, France, +33 (0)3 20 54 58 92
h.leleu@ieseg.fr

Abstract

This paper aims at providing answers to the questions raised in Cook, Tone & Zhu (2014) in the context of portfolio analysis with Data Envelopment Analysis (DEA). This reflection leads to define the financial production process as the generation of a distribution of returns by an initial investment. The main contribution of the paper is therefore to consider risks of various orders – mean return, variance of returns, and moments of higher order – as output variables and propose a set of axioms accordingly to supplement the definition of 'financial' technology set. In particular, this revisited set of axioms offers the advantage of allowing a generalization to multi-moment frameworks, and the resulting portfolio possibility set allows taking into account preferences for increases in risk that have remained ignored in applied studies with DEA although discussed in economic theory. We provide illustrations to show the effects of this contribution on the measures of technical efficiency and ranking of portfolios on a sample set of US common stocks; it shows how the proposed adjustments result in providing rankings that are more consistent with standard risk-return ratios in finance.

Keywords : Data Envelopment Analysis; Portfolio Frontier; Model orientation; Mean-Variance; Risk preferences

Portfolio analysis with DEA: prior to choosing a model

1. Introduction

Since the seminal work of Markowitz (1952) on portfolio selection, several tools, models and approaches for decision-making have been developed in the financial and economics literature to evaluate the performance of portfolios of financial assets. The mean-variance approach introduced by Markowitz relies on the construction of a frontier relative to which portfolio performance is measured. Parallel to this literature, a methodology for performance measurement of decision-making units (DMUs) was being developed in the economics and operational research literature with Data Envelopment Analysis (DEA), a non-parametric tool. The junction between portfolio selection through quadratic optimization and the methodology with DEA inherited from operational research occurred with Sengupta (1989), but it took until Murthi, Choi & Desai (1997) to identify DEA as an "*extremely useful technique for measuring efficiency*" of mutual funds. While they used a CCR [1] model on mutual funds, the following contribution of McMullen & Strong (1998) used a BCC [2] model. Premachandra, Powell & Shi (1998) then introduced stochastic DEA and studied stock indexes, and Morey & Morey (1999) used DEA for multi-horizon portfolio analysis. Since then, numerous studies have transposed the whole methodology used in production theory and operational research to the study of portfolios of financial assets with DEA without necessarily questioning the accuracy of such transposal. Though these works contributed to the elaboration of a general approach for measuring single-period portfolio efficiency in multi-moments frameworks (see Briec & Kerstens, 2010), some adjustments can still be proposed in order to make the approach suited to the analysis of financial assets, by so much as the definition of the underlying technology or the choice of a model orientation. At the crossroads of risk and lottery theory, a part of the literature in economics emphasizes on two major changes that have not been given much attention until now in the literature on multi-criteria decision-making with DEA: multi-moment frameworks ought to replace the simple mean-variance framework and the desirability of increases in risk measures ought to be considered. These changes drive the adjustments proposed in the paper.

In a recent article, Cook, Tone & Zhu (2014) list several modeling issues raised by an ill-adapted transposal of DEA models to various fields of research. In order to bring adequate solutions to these issues and ensure proper modeling, they also list a series of questions that should be answered prior to any analysis with DEA. In this article we intend to question the definition of a technology in the context of portfolio analysis through the identification of DMUs, the proper selection of input and output variables and the definition of a set of axioms. We propose answers to the questions raised in Cook, Tone & Zhu (2014) by identifying what can be the purpose of performance measurement and analysis and how it can impact the identification of DMUs or model orientation. We propose to apply a similar treatment to risk as the one used for byproducts in weakly disposable DEA models and show the consequences on the definition of the technology and model orientations. We also propose to modify the set of axioms inherited from production theory to take into account the correlations between assets' returns, the possibility of riskless investments and the implications of risk reduction on the level of expected return. We finally

[1] as introduced in Charnes, Cooper & Rhodes (1978)

[2] as introduced in Banker, Charnes & Cooper (1984)

provide an illustration of how these changes in the definition of the technology set and the model orientations have an impact on the efficiency scores and ranking of a set of portfolios of US common stocks.

2. Definition of a financial technology prior to the analysis

2.1. "What is the purpose of the performance measurement and analysis?"

The first question raised in Cook, Tone & Zhu (2014) relates to the purpose of performance measurement and analysis. Identifying the purpose of performance measurement prior to the analysis is a key concern in a field in which most studies have applied an identical methodology and transposed models from production theory to financial assets of various natures, from individual stocks to portfolios like mutual funds, hedge funds or CTAs. In both portfolio theory and production theory, performance evaluation can be considered under the two complementary angles of technical and allocative efficiency. Technical efficiency of financial assets can provide information on the return on investment relative to the various costs incurred, independently from any system of prices (e.g. with no regards to the decision-makers' preferences in portfolio analysis). Allocative efficiency is estimated relative to a profit-maximizing strategy (a utility-maximizing strategy in portfolio analysis, provided that the parameters of the utility function are known). While most studies in the literature focus on technical efficiency, Briec, Kerstens & Lesourd (2004) and Briec, Kerstens & Jokund (2007) also show how economic efficiency can be reached.

Regarding the study of financial assets with DEA, we propose to distinguish between three purposes. One purpose is to provide a ranking of investments for portfolio or asset selection when one investment is to be selected rather than the others. The analysis is led from the perspective of investors and implies measuring technical or economic efficiency of portfolio selection. A second purpose is to assess efficiency of portfolio construction or fund management. In that case the analysis would require measuring technical efficiency of portfolios relative to the frontier of their holdings, which has not yet been done in the literature to the best of our knowledge. Allocative efficiency of portfolio management or portfolio selection can also be measured in order to assess to which extend fund managers or investors succeed in reaching their individual objectives regarding either the fund's orientation or their respective preferences towards risks. [3] A third purpose is to assess the efficiency of financial markets. In this case the level of aggregation of the technology should replicate the one of the studied market, from a set of large-cap stocks only to the set of all assets on the US market for instance. Technical efficiency would then be measured to determine how far is the set of all assets on the markets from the market frontier. Further analysis can then be made on the determinants of portfolio inefficiency in order to study the drivers of the funds' performance in two-steps DEA (as in Galagedera & Silvapulle, 2002), which should theoretically converge to the results of a fundamental analysis.

[3] Allocative efficiency of portfolio management for already constituted funds would then have to be considered from the perspective of the funds' managers, as the investors in a fund are not the decision-makers and their utility functions can differ substantially, as discussed in Ballestero & Pla-Santamaria (2004).

Depending on the purpose of the study, the definition of technology set through the identification of DMUs or assumption of convexity will vary. While efficiency of portfolio selection can be measured relative to a set of already constituted portfolios, efficiency of portfolio construction ought to be measured relative to the set of assets that can be included in portfolios (either individual assets – a fund's holdings for instance – or already constituted portfolios like mutual funds to be invested in a fund of funds). Efficiency of the financial markets ought to be measured relative to the set of all individual assets that this market is composed of. [4] In the case of portfolio selection (when the objective is to invest fully in one DMU instead of composing a portfolio of DMUs), building a convex frontier does not serve the purpose of the study; rather, measuring performance relative to the frontier of a Free Disposal Hull (FDH) is consistent to rank the funds and correlations between the portfolios or funds can be ignored. In case of portfolio construction however, any element from the set frontier is built from convex combinations of individual securities. The linear correlation between the assets' distributions of prices however result in building a non-convex technology set in a moment-based framework.

Though performance evaluation can accurately be based on past records, its corollary risk measurement may also require resorting to fundamental analysis if it has a predictive intent, so that expectations about future prices can be formed from accounting information and used in the decision-making process. If this is especially true for individual stocks, performance assessment and risk measurement of investment funds can still accurately rely on historical records on the grounds that higher performance demonstrates superior management skills that can back up more favorable expectations about the funds' future performance. Predicting price and return remains a difficult challenge; as a consequence, whenever the objective of the study relates to portfolio selection in order to achieve future performance, decision-makers deal with the records of past returns to form their expectations. Most studies on portfolio performance with DEA until now have consequently adopted a retrospective approach, though Briec & Kerstens (2009) introduce a few thoughts for a prospective approach.

Cook, Tone & Zhu (2014) remind the importance of spending more time, prior to the analysis, determining what matters to the study-maker ("*the precise measures deemed important by management*"), which leads to the corollary questions of the choice of a theoretical framework to study technical efficiency and the identification of the study-makers' or decision-makers' preferences to study allocative efficiency. The economics literature on both aspects reveals that two major changes ought to be taken into consideration: first, the inclusion of higher moments of the distribution of returns in the analysis, and second, a possible preference for assets with a higher variance.

Regarding the theoretical framework, though we observe that the mean-variance framework has been very popular in the literature on performance measurement of financial portfolios with DEA until now (see the table in Appendix), several additional criteria have been proposed and reveal the need to consider more general multi-moment frameworks instead. The choice of a theoretical framework impacts the definition of the technology by determining the set of inputs, outputs and the definition of a set of regularity conditions. In this paper we propose a new treatment of risk as an output in the mean-variance framework; this output-based approach allows for a generalization to multi-moment frameworks and to measures of risk of various orders.

[4] The literature on portfolio performance with DEA has focused until now on already constituted portfolios such as mutual funds, hedge funds or CTAs (see the table in Appendix).

Regarding the parameters of the utility functions, preference for higher mean returns and aversion to a higher variance of returns are systematically assumed by study-makers. However, a whole part of the literature in economics and empirical finance has developed around the question of riskier choices and preferences for increases in even moments (from Blum & Friend (1975) to Golee & Tamarkin (1998), Astebro (2003), or Bali, Cakici & Whitelaw, 2010). From this literature we keep here only one major finding: the choice of a random prospect with a higher variance does not necessarily imply a preference for an increase in risk [5] in a multi-moments framework. In other words, risk aversion and the choice of a portfolio of financial assets with a higher variance are not mutually exclusive. At constant mean, such choice can be attributed to the impact of higher-order moments; in any other case, it can simply result from a utility function that attributes higher utility to riskier prospects, provided that they offer a high enough return. For this reason, measuring performance relative to a set of efficient but systematically less risky DMUs is too restrictive and ought to be reconsidered.

2.2. "What are the decision-making units and the outputs and inputs to be used to characterize the performance of those DMUs?"

In portfolio analysis, DMUs can either be individual securities or portfolios of securities such as investment funds or indexes, depending on the object of the study. One specificity of such DMUs is the linear (Pearson) correlation between their distributions of prices or returns. This correlation can be opposed to the implicit assumption of independence between DMUs in production theory, as long as the definition of the 'financial technology' is based on characteristics of the distribution of returns of the financial assets. Indeed, the existence of a linear correlation between the DMUs' distributions results in making the convex combinations of the moments of the distributions a non-linear function of these moments, which in turn results in building non-convex technology sets whenever these moments are identified as input or output variables.

[Insert Figure 1 here] As illustrated in Figure 1, some degree of linear dependence between financial assets impacts the level of risk of any convex combination of financial assets. It can be reduced through diversification; still, even linear independence (a case of zero linear correlation) between DMUs would not anyway result in linear combinations of the initial DMUs' risk levels for some risk measures. The minimum level of risk when measured by the variance of a distribution of returns is a convex quadratic function of the mean return; the resulting non-convexity of the set frontier is consequently an issue only whenever we measure performance using a direction vector that follows an expansion path. These dependence relationships between financial assets result in an additional specificity of portfolio analysis with DEA: in spite of being made of efficient portfolios, the frontier may very well be composed of portfolios made of inefficient assets only. This is illustrated in Figure 1 where the two inefficient DMUs "Oppenheimer Target M." and "Fund Manager Aggressive Growth" both enter in the composition of the efficient frontier.

Cook, Tone & Zhu (2014) remind that any process assimilated to a production process has to be clearly understood prior to the selection of input and output variables and remind the importance of ensuring that they "*properly reflect, to the greatest extent possible, the "process" under study*". However, the criteria

[5] In the sense of Rotschild & Stiglitz (1970), meaning a preference for the application of a Mean-Preserving Spread.

that have been either explicitly proposed or implicitly used for the selection of inputs and outputs in the literature that uses DEA on financial assets do not necessarily relate to the notion of a "production process". One criterion relates to the nature of the interaction between input and output variables, and more precisely whether it is characterized by a causal relationship of the production-kind or not. The identification of the causal relationship, if any, considerably eases the choice of input and output variables among the set of relevant variables. Another criterion relates to the behavior of decision-makers towards input or output variables inferred from what's assumed to be their preference or aversion to these variables. This second criterion is often considered as a mean to address the need to take into consideration investors' preferences, and more especially the measures that are relevant in their opinion. However, once the first criterion has been successfully applied, the second criterion becomes pointless. Still, the choice of a theoretical framework prior to the analysis should ensure consistency in the definition of any 'financial production process', determine its level of aggregation and drive the selection of input and output variables.

In spite of a rather consensual choice in the literature on the metrics (mainly return and total risk), on their source [6] and on the choice of output variables, the multiplicity of measures that have been proposed to account for input variables shows that there is no consensus either on the theoretical framework to be used for the study, on the measures to be used to account for some risk metrics or on the input or output status attributed to various measures.

On the one hand, desirable outcomes have always been included in the set of output variables and the choice of measures of reward as outputs obtains a consensus. Over various measures of average return proposed in the literature, one can find either mean or compounded return on past performance, [7] as well as expected return on future performance. Returns can either be expressed as gross or net returns [8] and as sometimes as excess return above the market's performance, either before or after tax. Minimum returns can also be found in some studies (see Wilkens & Zhu (2001), Glawischnig & Sommersguter-Reichmann, 2010) as well as the number of days/months with positive returns on a daily/monthly distribution of returns, (see Gregoriou & Zhu, 2007), upper (or higher) partial moments (see Gregoriou (2003), Gregoriou & al., 2005) or consecutive gains (see Gregoriou & Zhu, 2007). McMullen & Strong (1998) also take into consideration the returns over various time-horizons. Traditional performance indicators as the Sharpe and Treynor ratios, the Jensen's alpha or the reward-to-half-variance index (see Basso & Funari, 2005) have been considered as output measures, as well as other desirable outcomes such as a positive skewness of the distribution of returns (see Wilkens & Zhu, 2001), indicators of stochastic dominance or of the ethical orientation of a fund. In these latter cases a qualitative indicator can be added to the set of output variables (the ethical factor of Basso & Funari (2003) or the stochastic dominance indicator of Basso & Funari (2001, 2005) and Kuosmanen, 2005).

[6] Lamb & Tee (2012) remind that we generally "*estimate the measures from the series of returns and classify some as return measures and some as risk measures*"

[7] Arithmetic means of the distribution of returns assume withdrawal of gains while geometric means of the distribution of (gross) returns assume reinvestment of past gains.

[8] What we refer to as the 'gross return' is similar to the 'capitalization factor' of Basso & Funari (2007).

On the other hand, various costs associated to investment as well as undesirable outcomes have always been included in the set of input variables. [9] Murthi, Choi & Desai (1997) consider standard deviation of the returns as an input variable, as well as the transaction costs that managers incur "*in order to generate the return*" such as an expense ratio (management fees, marketing and operational expenses), additional loads for some funds (sales charges, redemption fees) or management turnover. Similarly, McMullen & Strong (1998) consider standard deviation of returns, sales charges, expense ratio and minimum investment as inputs. Eling (2006) includes the minimum investment or the lock-up period in the set of input variables, as well as an indicator of trading activity or excess kurtosis. Basso & Funari (2001) propose various risk measures (standard deviation of returns, root of the half-variance or beta coefficient) and additional costs as input variables. Choi & Murthi (2001) also consider managerial skills, market and institutional factors in their set of input variables. Wilkens & Zhu (2001), Galagedera & Silvapulle (2002), Basso & Funari (2003) then used many of the above-listed measures on multiple time horizons. Glawischnig & Sommersguter-Reichmann (2010) introduced lower partial moments as new measures of risk and input variables. Eling (2006), Gregoriou & Zhu (2007) or Branda (2015) also use various drawdowns (maximum or average drawdown, standard deviation of drawdown, Value-at-Risk, conditional Value-at-Risk or Modified Value-at-Risk), the beta factor and residual volatility or tracking error, and Gregoriou & Zhu (2005) use the proportion of negative returns in a distribution of returns as an input.

For all contributions listed above risk measures are treated as input variables and return measures as output variables, which can be explained by two main reasons. On the one hand, decision-making in production is based on input reduction and output augmentation and decision-making in finance is generally based on risk reduction and return augmentation. On the other hand the frontier of efficient portfolios is similar in shape to a production frontier; the analogy has then been made for long between efficiency analysis in production and performance analysis in finance. This analogy and the desirability for return and commonly accepted undesirability of risk have led numerous authors to consider the risk-return relationship of financial assets as equivalent to an input-output relationship. We propose instead an approach that both reflects the 'financial production process' by treating risks of various orders as outputs in any multi-moment framework and includes the possibility of increases in risk measures.

2.3. Resulting financial technology in a mean-variance framework

Under a mean-variance framework, defining the relationship between the level of second-order risk (measured by the variance or standard deviation of returns) and the realized return on investment as a 'production' relationship would lead to an erroneous and incomplete representation of the technology. On the one hand, no functional form can express the expectation of a higher return as a result of a riskier investment. On the other hand, the positivity of the risk-return relationship has been proven wrong on some categories of assets (on alternative investments for instance) or in case of the so-called 'leverage effect' (when the past returns are negatively correlated to the future volatility of some stocks, on short-term horizons). The risk-return relationship, often considered positive, is consequently no appropriate support for the representation of the technology.

[9] The only exception can be found in Devaney & Weber (2005) and Devaney, Morillon & Weber (2016) who include risk measures in the set of output variables.

As the measures associated to risk and return both have the same source (the distribution of returns), it then seems consistent to treat them all as outputs. [10] Devaney, Morillon & Weber (2016) remind for instance that "*in a statistical sense it is not the case that the second moment of the return distribution (risk) determines the first moment of the distribution (return)*". Similarly, higher moments of the distribution to be included in multi-moment framework (as skewness or kurtosis) and any measure characterizing the distribution of returns could be regarded as outputs. A similar understanding of the 'financial technology' can be found in Anderson & al. (2004) who consider that any benefit arising from an investment is an output and the investment itself is the input. However, while they consider that the level of risk is "taken" by the investor and is therefore part of the initial investment made in the portfolio, we consider that it cannot be quantified a priori and is therefore not 'taken' but rather generated simultaneously to the distribution of returns. A timing assumption also underlies any production process: output generation must be preceded by the supply of some input, as it results from the latter and the production process necessarily takes some time. This sequence is realized here: all outputs are generated simultaneously after the initial investment has been made.

Following our proposal to treat risks of various orders as outputs, any analysis based on a multi-moment framework implies working on the output correspondence of the production possibility set, which ensures that the set is closed. One convenient consequence is that performance measurement relative to the frontier can be made regardless of the system of preferences: models are no more constrained to be oriented towards risk reduction only along with return augmentation, which was the case when risk was considered as an input.

From the above specifications a financial technology can be defined as in equation (1). We assume an output vector $\mathbf{y} = (f_1, f_2, \dots, f_n)$ with f_i the random variable "risk of order n". [11] As reminded in Lamb & Tee (2012), any measure from the set of possible portfolios is a function $g\colon \mathfrak{I} \to \mathbb{R}$, with the value of this measure being a population statistic (a sample statistic in practice) and a real number (rather than a random variable). It is always assumed that the portfolios' returns are realizations of the random variable. Staying in the simple mean-variance framework, we can assume f_1 and f_2 to be the first two non-standardized moments that characterize the portfolios' distributions of returns. For an initial investment x [12] (or a vector of inputs $\mathbf{x} \in \mathbb{R}_+^n$ with $n \in \mathbb{N}$ the number of input variables), the output vector can then be defined as $\mathbf{y} = (f_1, f_2) = (\mu, \sigma^2)$ and the financial technology set T can be defined on its output correspondence $P(x)$ as such as in (1).

$T(x, \mu, \sigma^2) = \{(x, \mu, \sigma^2) : x$ can produce a distribution of returns with the first two moments $(\mu, \sigma^2)\}$

and $P(x) = \{(\mu, \sigma^2) : (x, \mu, \sigma^2) \in T\}$ for all $x \in \mathbb{R}^+$ **(1)**

[10] Liu, Zhou, Liu & Xiao (2015) provide a similar argument.

[11] We use the terminology of the lottery theory that designates as risk of order 1 the risk of loss, measured by the 1st moment of the distribution of returns, risk of order 2 the risk of variability measured by the 2nd moment, risk of order 3 the downside risk measured by the 3rd moment, risk of order 4 the outer risk measured by the 4th moment, etc.

[12] The above redefinition of a 'financial production process' mainly questions the causal link between risk and return to conclude that they are both generated by an initial investment. Yet, several other input variables may be proposed in addition to this initial investment to complete this definition, such as the mandatory presence of a market, some necessary degree of liquidity or the presence of intermediaries. Indeed, an initial amount available for investment could generate no distribution of return at all if there was no market for instance. One could however argue that the notion of initial investment implicitly assumes that these requirements are met.

It is also important to notice that the distribution of returns is generally expressed not in monetary units but as rates of return on investment. In this case, the output sets would be the same for any level of input x, which would translate into the following equality: $P(x) = P = \{(\mu, \sigma^2) : (\mu, \sigma^2) \in T\}$ for all $x \in \mathbb{R}^+$.

Let R_j be the return on DMU j with $j \in J$ at time t. We consider R_j as a random variable defined on the probability space (Ω^j, F^j, P^j), with Ω^j the sample space (or set of all possible outcomes) of the variable R_j, $F^j = (F_1^j, F_2^j, \dots, F_E^j) = \{F_e^j : e \in E\}$ the set of events that can influence the outcomes of the variable R_j, with $E \in \mathbb{N}^*$ the number of possible events and $P^j = (p_1^j, p_2^j, \dots, p_E^j) = \{p_e^j : e \in E\}$ for all j the assignment of probabilities to every event contained in F^j. Let's also assume that for any DMU i, μ_i is the mean return of a distribution of returns, q_i is the share of DMU i in a portfolio, σ_i is the standard deviation of the distribution of periodic returns and ρ_{ij} is the coefficient of linear correlation with the distribution of a DMU j. The set of admissible activity vectors [13] that represents all possible combinations of shares q_j of initial investment in portfolio j can be defined as in equation (2).

$$\mathfrak{I} = \left\{ \mathbf{q} \in \mathbb{R}^J : \sum_{j=1}^{J} q_j \leq 1 \ , \ q_j \geq 0 \text{ for all } j \right\}^{14} \quad \text{and } \mathfrak{I} \neq \emptyset \qquad \textbf{(2)}$$

Depending on the theoretical framework selected, the representation of the set of possible portfolios is then expressed as the set of all the related measures such that $\mathbf{q} \in \mathfrak{I}$. The portfolio possibility set defined in (1) on the output correspondence can then be redefined from the sample set of observed DMUs and a set of admissible activity vectors $\mathfrak{I}$ as subsets of output vectors P as in equation (3) below, if free disposability was assumed on outputs.

$$P = \{(\mu, \sigma^2) \ : \ \mathbf{q^T \mu} \geqq \mu \ , \ \mathbf{q^T \Omega q} \leqq \sigma^2 \ , \ \mathbf{q} \in \mathfrak{I}\} \qquad \textbf{(3)}$$

with:

- $\boldsymbol{\mu}$ the $(n \times 1)$ vector of mean returns of the n observed DMUs
- $\boldsymbol{\Omega}$ the $(n \times n)$ matrix of covariances of the n observed DMUs and $\boldsymbol{\sigma^2}$ the $(n \times 1)$ vector of variance of returns of the n DMUs
- $\mathbf{q^T}$ the transpose of $\mathbf{q}$

The representation of the set of possible portfolios in the mean-variance framework can also be defined as the set of all mean-variance combinations of portfolios such that $\mathbf{q} \in \mathfrak{I}$ as in equation (4), with GR_{MV} a non-convex set.

$$GR_{MV} = \{ (E_P, V_P) : \mathbf{q} \in \mathfrak{I} \} \qquad \textbf{(4)}$$

with $E_P = \sum_{j=1}^{J} q_j \mu_j$ the mean return and $V_P = \sum_{j=1}^{J} \sum_{k=1}^{J} q_j q_k \rho_{jk} \sigma_j \sigma_k$ the variance of returns

[13] Referred to as the "set of admissible portfolios" or "portfolio possibility set" in Briec, Kerstens & Lesourd (2004), but as this terminology can be confusing we rather keep "portfolio possibility set" to refer to the financial technology set T.

[14] Following the treatment of short sales commonly accepted in the literature, short sales are treated like negative purchases, such that any negative q_j indicates that portfolio j is sold short. If short sales are allowed, the share invested in each portfolio is no more constrained to be non-negative. A lower bound to q_j consequently determines whether or not short sales are allowed.

2.4. “What is the appropriate model orientation?”

Based on the literature on decision-makers’ preferences in the expected utility framework, the following model orientations appear theoretically grounded in an output-oriented mean-variance model that assumes a utility function of mean and variance only.

(1) ***Return augmentation at constant variance***, as proposed in Morey & Morey (1999) is recommended whenever the parameters of the utility functions are unknown or if it is assumed that the tolerance of decision-makers for risk is equivalent to the evaluated DMU’s variance. If the level of risk carried by the evaluated funds reflects the maximum level of risk that the funds’ managers are willing to face, this orientation provides information on how much more return could have been generated for that level of risk. In no case is this orientation related to risk-neutrality. [15]

(2) ***Risk reduction at constant mean***, as proposed in Morey & Morey (1999), allows to measure to which extend fund managers succeed in reducing the risk of their portfolio at a given mean return – or investors in the evaluated portfolios succeed in selecting the less risky portfolio at a given level of required return. This orientation is accurate whenever the parameters of the utility function (consequently the coefficient of risk aversion) are unknown but risk aversion is assumed to be shared by all decision-makers: under such conditions any portfolio with a lower variance will be preferred, at a constant mean.

(3) The ***simultaneous risk reduction and return augmentation*** (the Efficiency Improvement Possibility (EIP) function introduced in Briec, Kerstens & Lesourd (2004) and used in Briec, Kerstens & Jokung (2007) and Briec & Kerstens, 2009) can also be considered when the parameters of the utility function are known and allocative or economic efficiency is measured. However, if the parameters remains unknown, nothing can theoretically justify this orientation even though all DMUs with a lower variance and a higher return dominate the set of observed DMUs. Indeed, only the evaluation relative to a portfolio with a lower variance at a constant mean guarantees that the choice results from risk-aversion (risk-lovers could prefer a portfolio with a higher return and a lower variance for the higher return it provides). Unless the parameters of the utility function are known, nothing can theoretically justify a measurement of performance relative to a DMU that has neither the same level of risk nor the same level of return. [16] When the parameters remain unknown but risk aversion is a key assumption, we recommend to use the second model orientation, as evaluation relative to a DMU with lower variance implies risk aversion at constant mean only.

(4) A ***simultaneous augmentation of risk and return*** has not been considered so far in the literature with DEA on financial portfolios. Yet, this orientation is especially legitimate whenever it is assumed that the higher the expectation on return, the higher the level of risk investors are ready to take, even under the assumption of risk aversion. If risk and return are considered as outputs, it is simply an output-based Debreu-Farrell measure of technical efficiency. This radial measure, by keeping the ratio risk/return constant, guides the evaluated DMU towards the frontier along the ‘expansion path’, its output mix. As proved by Russell (1985), radial measures present several desirable properties especially when market

[15] Assuming risk neutrality would merely imply removing any risk measure from the model and evaluate the DMUs relative to the one that provides the highest return.

[16] It could for instance lead to measure the performance of a fund relative to another fund that has a different objective regarding risk management.

prices – or preferences in our context are unknown. In a production framework, one desirable property of the radial measure is that it provides information of the variation in revenue, independently of the value of prices, known or not (Russell, 1985). For portfolio analysis under the expected utility framework and for a utility function that can be expressed as a linear function of the first two moments as in equation (5), such direction can provide information on the variation in utility that results from reaching the efficient frontier, regardless of the parameters of the utility function. A radial expansion of the observed DMU (E_o, V_o) by a factor n results in (E_o^*, V_o^*) as in equation (6), and the utility U^* associated to this projection to the frontier can be expressed as in equation (7). This direction is also consistent with the assumption of jointness introduced in section 3.4. If expected future returns are used instead of the mean return on past records, they should naturally be positive – according to theory – as they would exceed the risk-free rate of return. The expansion path would in that case be particularly relevant.

$$U(E, V) = \mu E - \rho V \qquad \text{with } \mu \text{ and } \rho \text{ the parameters of the utility function,} \tag{5}$$

with $\mu \geq 0$ and $\rho > 0$ in case of risk-aversion and $\rho < 0$ in case of a preference for risk

$$(E^*, V^*) = \big((1+n)E, (1+n)V\big) \tag{6}$$

$$U_o^* = (1+n)U_o \text{ with } U_o^* = U(E_o^*, V_o^*) = \mu(1+n)E_o - \rho(1+n)V_o \tag{7}$$

The above direction implies positive mean returns, of which a real sample is often not made as it deals with mean returns on past records. As no relevant model orientation would consider a decrease in mean return, an alternative consists in using the direction of the regression line in the mean-variance space, and equation (6) would then become equation (8). It would also be justified in cases where the R-squared of the regression is positive and significant, so that it can support the idea of a positive risk-return relationship on the market of the studied sample of portfolios. [17]

$$(E^*, V^*) = \big((1+\delta\, g_E)E, (1+\delta\, g_V)V\big) \tag{8}$$

with g_E/g_V the slope of the regression line and δ a scalar.

2.5. "What is an appropriate number of DMUs, given the number of inputs and outputs chosen?"

One additional issue to the identification of input and output variables often raised in the literature and mentioned in Cook, Tone & Zhu (2014) relates to the appropriate number of DMUs to constitute a sample. A reciprocal question relates to the maximum amount of variables to be allowed in the set of input and output variables, knowing that additional variables most often result in an increase in the number of efficient DMUs. This phenomenon is sometimes referred to as the *curse of dimensionality* and is especially crucial for non-parametric estimators (see Simar & Wilson, 2000).

[17] By way of example, for the data used in our illustration (sections 4 and 5), the R^2 of the regression on the returns of the random portfolios over 3-year time windows are equal to 0.1559, 0.1597, 0.4104, 0.2576, 0.4979, 0.3008, 0.5106 and 0.2498 for the periods ranging from 2005 to 2007 to 2012 to 2015, respectively.

This question can be addressed in two way. On a purely theoretical and statistical basis, one can study the influence of the number of DMUs and the number of input and output variables on the speed of convergence of the estimator (see Kneip, Park & Simar, 1998). On an empirical basis, numerous authors recommend restricting as much as possible the number of input and output and some rules of thumb relating the number of input and outputs to the number of DMUS have been proposed without being theoretically grounded. Empirical testing is also a way to deal with this issue. By varying the sample size and the number of inputs and outputs, we could test if efficiency scores are robust across different specifications. A recent contribution of Liu, Zhou, Liu & Xiao (2015) deals with how well various DEA models used in performance measurement of financial portfolios approximate the portfolio frontier.

3. Modifications to the traditional set of axioms

The set of axioms underlying the definition of the technology in the literature that studies performance of financial assets with DEA has been transposed from the set of axioms proposed in production theory. This set of axioms that we would define as 'traditional' usually consists in assuming no free lunch, the possibility of inaction, free or weak disposability on input and output variables, and that the technology set is convex, bounded and closed. Various assumptions on returns to scale are also proposed.

As financial assets differ in their nature and dependence relationships from production units studied in production theory or operations research, the definition of the related technology cannot rely on a strictly similar set of axioms. Still, to the best of our knowledge, every study on financial assets with DEA has assumed free disposability on both input and output variables for instance. Based on the above redefinition of DMUs, input and output variables, we propose here a revisited set of axioms to be applied in a multi-moment framework.

3.1. Non-convexity of portfolio possibility sets

As a consequence of the linear correlation between the assets' distributions, convexity cannot be imposed as regularity condition for a financial technology. We saw that independence between production units could in no way be translated into independence between the distributions of returns of financial assets; it would instead be equivalent to an assumption of perfect linear correlation between them in a mean-standard deviation framework only. But linear independence or perfect linear correlation between financial assets are only particular cases that may never be observed in reality. Moreover, though linear independence between the distributions imply a null correlation, the reciprocal is not true: independence is no way implied by an observed zero linear correlation.

In the simplest multi-moment framework that only considers the first two moments of the distribution of returns, the mean return of any convex combination of DMUs (of any portfolio of assets) can be expressed as a linear function of the individual mean returns of the assets (see equation (4)). On the contrary, the variance of any convex combination of DMUs cannot be expressed as a linear function of the individual variances of returns of the assets (see equation (4)). In case of a perfect linear correlation between assets i and j such that $\rho_{ij} = 1$, diversification does not impact the portfolio's total risk. The portfolio variance σ_P^2 is then equal to $\sum_{i=1}^{n}\sum_{j=1}^{n} q_i q_j \sigma_i \sigma_j$. In case all assets in the portfolio were independent, the portfolio variance σ_P^2 would be equal to $\sum_{i=1}^{n} {q_i}^2 {\sigma_i}^2$ and diversification would impact (reduce) the total portfolio's risk. When variance is the measure of second-order risk, convex combinations of assets can consequently never result in linear combinations of both the assets' risk and return measures, which would ensure the convexity of the technology set in a mean-variance framework. [18]

One way of making the technology set convex even when correlations are taken into consideration is to assume free disposability on the risk dimensions under the approach that defines risk as an input. However, as long as risk is defined as an output, even free disposability on the risk dimensions would not result in extending the output set into a convex set. On Figure 1 for instance, free disposability on the risk measure would extend the set 'to the left' and result in a non-convex output set due to the set frontier 'on the right'.

Another way of making the technology set convex would be to ignore the linear correlation between DMUs and work with standard deviation instead of variance, but this solution leads to wrong estimates of efficiency scores. As noticed by Lozano & Gutierrez (2008), all linear programming approaches used to measure efficiency of mutual funds – except for Daraio & Simar (2006) – have overestimated risk by considering convex combinations of the DMUs' respective levels of risk to account for portfolio risk. For Brandouy, Kerstens & Van de Woestyne (2013) "*it could at best be considered a type of linear approximation of a possibly non-linear portfolio model*".

A third way of obtaining a convex technology set for financial assets would be to choose risk measures that display no linear dependence, but that would lead to reconsider the choice of a theoretical framework. For instance, Lamb & Tee (2012) and Branda (2015) recently proposed to work with "diversification-consistent models" that solve this issue by using other risk measures than variance or standard deviation.

Obtaining a convex technology set of financial assets in a multi-moment framework would then be at the cost of either making the unrealistic assumption of free disposability on the risk measures, ignoring the linear correlations between the distributions of assets' prices, or even modifying the theoretical framework itself. In order to stay in line with a framework much favored by the literature on decision-making, we think that convexity should not be imposed as a regularity condition.

3.2. Axioms of "no free lunch" and the possibility of inaction

The axiom of "no free lunch" has often been considered in financial analysis as equivalent to an assumption of fair pricing on the markets. However, Barberis & Thaler (2003) remind that this equivalence holds in efficient markets only, and while correct pricing implies no free lunch, the opposite is not true. Market inefficiency does not necessarily imply free lunches, and market inefficiency is certainly not to be deduced from the sole inability of investors to generate excess return over the market's

[18] If standard deviation of returns – which can be considered as an equivalent measure to variance in terms of information – was the measure of second-order risk, linearity of the risk and return measures of the combinations and the assets could only be observed in the unrealistic case where all assets portfolio were perfectly linearly correlated ($\rho_{ij} = 1$), such that $\sigma_P = \sum_{i=1}^{n} q_i \sigma_i$ (the squared root of expression $\sigma_P^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} q_i q_j \sigma_i \sigma_j$ above) – or whenever a risk-free asset is included in the portfolio, such that $\sigma_P = \sum_{j=1}^{n-1} q_i \sigma_i$, which makes the optimization problem trivial.

return. Arbitrage strategies are led by rational investors (the 'arbitrageurs') and are possible as long as some assets are mispriced on the market. Though the strict definition of arbitrage refers to some riskless profit opportunities, these arbitrages are not necessarily riskless, as rational investors still lack some information on inefficient markets.

If $(\mathbf{x}, \mathbf{y}) \in T$ and $\mathbf{x} = 0$, then $\mathbf{y} = 0$: no 'free lunch' when the risk measure is the input $\mathbf{x}$ **(9)**

$P(\mathbf{0}) = \mathbf{0}$: no 'free lunch' when the risk measure serves as an output **(10)**

The axiom of 'no free lunch' can be expressed as in equation (9), for any non-negative vector $\mathbf{x}$ of input and any vector $\mathbf{y}$ of output and production possibility set T such that $T = \{(\mathbf{x}, \mathbf{y}) \ : \ \mathbf{x}\, can\ produce\ \mathbf{y}\}$. In an approach that assimilates risk to an input, the axiom of 'no free lunch' implies that if no second-order risk characterizes the asset, no return can be generated, which contradicts the existence of assets that are considered free of risk and at the same time generate positive returns (such as T-bills). The axiom of 'no free lunch' therefore precludes the introduction of risk-free assets in the portfolio or the proper analysis of any portfolio with a guaranteed minimum return under an approach that would consider risk as an input. Such drawback could however be overcome by using specific measures of return: in case the excess return over the risk-free rate is used instead of the mean return, this axiom allows taking a risk-free asset into consideration and implies that if no second-order risk characterizes the asset, no excess return can be generated over the risk-free rate. If however second-order risk is defined as an output, the axiom can then be defined as in equation (10), and simply implies that for a distribution of returns to be generated, there must be some strictly positive initial investment.

If short selling or any kind of leverage was allowed it would then be necessary to define another input that would be specific to that kind of investment and account for the initial operation implied (borrowing the shares, finding a counterpart, arbitraging). Then, though no single cost can theoretically be incurred, the action undertaken by the investor would be taken into account and the principle of 'no free lunch' would still hold. A zero input vector would remain specific to the very specific case of 'doing nothing' that excludes short selling.

A second axiom of the "possibility of inaction" – sometimes referred to as "doing nothing is feasible" – could be expressed in two ways depending on the assumption of disposability on inputs. On the one hand, the 'raw' axiom of inaction only assumes the possibility of producing no output from a zero vector of input. On the other hand, the 'extended' axiom of inaction (or axiom of 'near' inaction) adds a disposability component to the input variables and assumes the possibility of producing nothing from any non-negative level of input. In an approach that assimilates risk to an input, 'raw' inaction implies that riskless holdings that generate no return belong to the technology set, such that $(\mathbf{0}, \mathbf{0}) \in T$. The origin of the production possibility set can in this case represent any holding generating neither risk nor return (cash holdings or cash equivalent, or any theoretical DMU obtained from the free disposal of a riskless investment). The inclusion of such assets in the technology set can however be hampered by some returns to scale assumptions when no such holding is observed and included in the set directly from the sample set of DMUs. For a classical DEA under VRS or NDRS, the origin does not belong to the set T, except if the axiom of convexity always accepted under the DEA-production approach was rejected.

When risk is assimilated to an output, the representation of cash holdings is the origin of each output set of the output correspondence. The raw axiom of inaction ($\mathbf{0} \in P(\mathbf{x})$ for $\mathbf{x} = \mathbf{0}$) then implies that making no initial investment in any portfolio is possible and will result in no generation of a distribution of returns (and consequently a zero-risk and a zero-return). Still, it does not ensure that the origin of the set belongs to any output set regardless of the level of input (initial investment). It therefore fails to ensure that holding cash or cash equivalents is allowed for any initial amount to be invested. While this axiom relates to the possibility of holding cash under an approach that assimilates risk to an input, it only relates to the possibility of doing nothing under the output-oriented approach, which matches its initial meaning.

3.3. Disposability assumptions on outputs

To the best of our knowledge, every portfolio analysis with DEA until now has assumed free disposability on input and output variables (which means on both risks and returns). Free disposability on inputs seems consistent when the initial investment is considered as an input. Free disposability of return when considered as an output is consistent as well, as any return on an investment can be disposed of: once perceived, returns can be kept, reinvested or even wasted. By contrast, the intangible nature of risk seems inconsistent with free disposal. Moreover, when risk is identified as an input variable, assuming free disposability implies the possibility of increasing the level of risk of an investment at a constant level of return. In such case, the addition of any risky asset with a zero mean return to the portfolio would correspond to such increase. But still, such possibility would depend on the selected risk measure, which prevents us from considering the assumption of free disposability on the risk variable as a generally accepted assumption. Assuming free disposability of risk implies the possibility of full reduction of the risk measure at a constant level of return, and no more the feasibility of unlimited increases in risk. Such reduction will always come at the cost of hedging; costly disposability [19] may then be more accurate. Moreover, assuming free disposability on the risk dimension would also prevent any inclusion of risk-loving behaviors in the study as it precludes any projection to ‘the right’ of the technology set. But as mentioned in Färe & Grosskopf (2003), the disposability assumptions are properties of the technology while the choice of a direction relates to the following step of performance measurement. Only arguments that relate to the definition of the technology should be used to determine which disposability assumptions are appropriate.

The rationale to treat risk as an input is then similar to what makes some authors treat any detrimental variable as an input. The idea that it incurs a cost, together with the natural assumption that decision-makers try to decrease their costs, leads to consider every variable that is to be decreased as an input. In production theory, the same rationale is used in models that assimilate byproducts to freely disposable inputs (introduced by Hailu & Veeman, 2001) with negative shadow price associated to these “bad” outputs. It however implies that no positive value can be attributed to these byproducts, which is a clear limit for portfolio analysis when we consider the progress of the literature on risk-loving behaviors. Moreover, as emphasized in Färe & Grosskopf (2003) considering byproducts as inputs would lead to inconsistencies with both the traditional set of axioms and physical laws. As these byproducts are technically produced by the inputs, they should be considered as outputs. This argument of technical feasibility can also be put forward to support our choice of treating risk measures as output variables, as we considered an initial investment generates a distribution of returns and that both the mean return, the

[19] We use “costly disposability” in this paper preferably to “weak disposability” to account for the cost of hedging mentioned above.

risk or higher moments characterizing this distribution are just statistics of the distribution of a single random variable.

A second limit of the DEA models used until now on financial assets is that they do not take into consideration any jointness between the so-called 'good' and 'bad' outputs. The inclusion of jointness in the model implies considering the 'bad' variable as an output variable and no more an input variable in order to comply with the definition of jointness and even null jointness, and that's what Färe & Grosskopf (2004) proposed. For this reason, we choose in our approach not to refer to risk or any other output variable as a 'good' or 'bad' outputs but rather identify them as "intended outputs" and "joint outputs". Though we agree on the positivity of shadow prices associated to intended outputs, we leave the characterization of joint outputs as 'good' or 'bad' to the choice of decision-makers according to their own preferences. We then impose no a priori assumption of negativity on shadow prices associated to joint outputs that could potentially be positively valued by some decision-makers. Joint outputs can then either be desired or rejected, but if no uniform preference is assumed among investors it can be interesting to consider the possibility of the expansion path that would increase both. As illustrated in sections 4 and 5, even though costly disposability is not imposed on the undesirable output but free disposability is kept – as it would have been if it has been treated as an input – differences in scores and in rankings can be observed due to the application of jointness (when the DMUs are evaluated relative to the risk-averse frontier).

3.4. Introducing costly disposability and jointness in the models

Costly disposability as it is usually modeled consists in assuming three elements at the same time. A first element is jointness (see in Färe & Grosskopf, 2003) that can be defined as in (11)[20] and introduced through a factor τ on the joint output variables. A second element is null jointness and can be defined as in (12). A third element consists in assuming costly disposability on the output that is not freely disposable, meaning relaxing the positivity constraint on the shadow price associated to the constraint.

Jointness of μ and σ^2: if $(\mu, \sigma^2) \in P$ and $0 \leq \tau \leq 1$ then $(\tau\mu, \tau^2\sigma^2) \in P$ **(11)**

Null jointness of μ and σ^2: if $(\mu, \sigma^2) \in P$, $\sigma^2 = 0$ implies $\mu = 0$ **(12)**

We saw that when free disposability was assumed on both risk and return, the portfolio possibility set could be defined on the output correspondence as P in equation (3). After jointness is introduced to the definition of the technology it can be redefined as P_2 in equation (13). A related model – oriented towards risk reduction only here – is proposed in equation (14) with an efficiency scores δ^2 that can be interpreted as the highest reduction in variance at a constant level of return to reach the set frontier.

$$P_2 = \left\{(\mu, \sigma^2) \ : \ \tau^2\left(\mathbf{q}^\mathbf{T}\boldsymbol{\mu}\right) \geqq \mu \,, \tau\mathbf{q}^\mathbf{T}\boldsymbol{\Omega}\mathbf{q} \leqq \sigma^2 \,, 0 \leq \tau \leq 1 \,, \mathbf{q} \in \Im\right\} \quad \textbf{(13)}$$

$\min\{\delta^2\}$ s.t. **(14)**

[20] If standard deviation was selected as a measure of risk, then equation (11) could be expressed as follows: jointness of μ and σ: if $(\mu, \sigma) \in P$ and $0 \leq \tau \leq 1$ then $(\tau\mu, \tau\sigma) \in P$

- $\tau \sum_{j=1}^{n} q_j \mu_j \geq \mu_{j_0}$
- $\tau^2 \sum_{j=1}^{n} \sum_{k=1}^{n} q_j q_k \sigma_{jk} \leq \delta^2 \sigma_{j_0}^2$
- $\sum_{j=1}^{J} q_j = 1$
- $q_j \geq 0$ for all j
- $0 \leq \tau \leq 1$

The third element of costly disposability can be modeled by an equality in the constraint related to the output that is not freely disposable. Introducing this new element, the portfolio possibility set can then be defined as P_3 in equation (15). A related model – oriented towards risk reduction only here – is proposed in equation (16) with an efficiency scores δ^3 that can be interpreted as the highest reduction in variance to reach the set frontier.

$$P_3 = \{(\mu, \sigma^2) \ : \tau(\mathbf{q}^\mathbf{T}\boldsymbol{\mu}) \geqq \mu \,,\ \tau^2\mathbf{q}^\mathbf{T}\boldsymbol{\Omega}\mathbf{q} = \sigma^2 \ , 0 \leq \tau \leq 1 \,, \mathbf{q} \in \mathfrak{I}\} \tag{15}$$

$\min\{\delta^3\}$ s.t. **(16)**

- $\tau \sum_{j=1}^{n} q_j \mu_j \geq \mu_{j_0}$
- $\tau^2 \left(\sum_{j=1}^{n} \sum_{k=1}^{n} q_j q_k \sigma_{jk}\right) = \delta^3 \sigma_{j_0}^2$
- $\sum_{i=1}^{J} q_j = 1$
- $q_j \geq 0$ for all j
- $0 \leq \tau \leq 1$

[Insert Figure 2 here] As reminded by Lamb & Tee (2012), though the definition of a portfolio possibility set and its frontier "*are […] rarely discussed*", they are fundamental as "*any DEA model estimates how far a DMU is from the frontier of its production possibility set*". While models (14) and (16) should obtain the same results, the choice of any other direction can result in obtaining efficiency scores that differ from one model to the other, as P_3 relaxes the positivity constraint on the shadow prices of variances compared to P_2. Figure 2 shows how a strict transposal of the traditional set of axioms with Variable Returns to Scale resulted in building a technology set $T(x, y)$ in the majority of studies on financial assets with DEA. It also shows how considering both risk and return measures as outputs allows to build the output set P of equation (3), similar to $T(x, y)$ in the elements it contains and in preventing any model orientation that would simultaneously increase risk and return. Adding jointness to P builds the output set P_2; adding costly disposability to P_2 builds the output set P_3. The latter allows for performance measurement in the direction of a simultaneous increase in risk and return.[21] As illustrated in Figure 2, minimizing risk only will not change the efficiency scores obtained from a set P_2 to a set P_3. However, assuming return augmentation or a simultaneous increase in risk and return will result in increases in efficiency scores for some DMUs with the highest levels of risk or return. On the output set P_3 of Figure 2, DMU 16 would be deemed efficient if evaluated with a model oriented towards mean augmentation only. DMUs 1, 7, 13, 15, 20, 21, 23 and 25 on the other hand would obtain a lower inefficiency score

[21] In practice, including cash in the sample of DMUs comes down to assuming jointness; both result in building a similar technology set, regardless of the measure selected for risk. As the programs can be made more complex depending on the jointness factor adapted to the selected measures of risk (τ will be used with standard deviation and τ^2 with variance), it may be easier in practice to include cash in the sample set of DMUs.

compared to the scores obtained with a direction of variance reduction only. In case of a simultaneous increase in risk and return, the inefficiency score of DMU 13 for instance would substantially decrease.

It could also be accurate to impose a limit to the share of cash holdings or other assets in the portfolios. Porter & Gaumnitz (1972) for instance add a maximum requirement of 5% of the portfolio invested in each asset (excluding cash holdings). In this latter case, it is necessary to constraint activity vectors to be lower than this upper limit. A revisited axiom of a ‘restricted possibility of inaction’ could then be proposed to allow for the study of any kind of risky portfolios, from fully riskless to fully risky portfolios (what Porter & Gaumnitz (1972) refer to as cases where no full investment is required). The definition of this ‘restricted possibility of inaction’ would depend on a factor q_{Cash} to range from a fully restricted to a full access to cash holdings. Let’s define the share $q_{Cash} \in [0,1]$ as the maximum proportion of cash holdings allowed in the portfolio. The restricted possibility of inaction can then be defined as in equation (17). The particular case $q_{Cash} = 1$ ensures the inclusion of the origin of the output set in the possibility set. A lower limit to the factor of jointness such that $(1 - q_{Cash}) \leq \tau \leq 1$ would then have similar effects and P_3 rewritten as in equation (18).

$$\{\mathbf{0},\mathbf{0}\} \in P \text{ and } q_{Cash} \in [0,1] \qquad \text{: restricted possibility of inaction} \qquad \textbf{(17)}$$

$$P_3 = \left\{(\mu,\sigma^2): \tau\left(\mathbf{q}^{\mathbf{T}}\boldsymbol{\mu}\right) \geqq \mu\ ,\ \tau^2\mathbf{q}^{\mathbf{T}}\boldsymbol{\Omega}\mathbf{q} = \sigma^2\ ,\ (1 - q_{Cash}) \leq \tau \leq 1\ ,\ 0 \leq q_{Cash} \leq 1\ ,\ \mathbf{q} \in \Im\right\} \qquad \textbf{(18)}$$

3.5. Returns to scale assumptions

Jointness is a key assumption in the context of portfolio analysis: on the one hand, we know that on the financial markets risk can only be reduced through diversification or at the cost of hedging, but once the higher degree of diversification has been reached only the inclusion of some riskless assets like cash in portfolios can further reduce risk. On the other hand, the expected return-risk relationship is positive and can justifies the assumption of jointness as well on an expected return-risk framework. Assuming jointness or imposing the inclusion of cash in the set of DMUs is then relevant for portfolio performance measurement (both solutions will deem similar efficiency measurement in a mean-variance analysis). To this regards, Liu, Zhou, Liu & Xiao (2015) propose to include cash in the analysis by replacing the constraint $\sum_{j=1}^{m} \lambda_j = 1$ (the convexity constraint of activity vectors λ_j for a set of m DMUs j) by a constraint $\sum_{j=1}^{m} \lambda_j \leq 1$ that is actually a mix of the convexity constraint and the Non-Increasing Returns to Scale (NIRS) constraint on some scale parameter θ that should be such that $0 \leq \theta \leq 1$. The new ‘convexity constraint’ they propose should consequently be written $\sum_{j=1}^{m} z_j \leq 1$ with $z_j = \theta\lambda_j$. Their answer to the need of including a risk-free asset in the portfolio possibility set is to assume NIRS, which includes the origin of the set to the set of possible portfolios but ensures the inclusion of a risk-free asset only if and only if excess returns (above the risk-free rate) are considered instead of returns, which is not the case in most studies and should not be a condition for the set to be consistent. An alternative way of dealing with this matter is to propose jointness as we did in this article.

Moreover, as underlined by Brandouy, Kerstens & Van de Woestyne (2013), “*the very notion of returns to scale may not necessarily be directly transposed to the finance context*”. It translates into different implications than in the previous studies with DEA: once redefined the ‘financial production process’ of generation of a distribution of return by an initial investment, the question rather becomes the following: "to which extent does an increase in the initial investment (input) result in an increase in return and risks (outputs)?". In this case, the notion of returns to scale can very well be transposed to the finance context and is a consistent object of study. But no such link can be made between risk and return and making no returns to scale assumption (assume Variable Returns to Scale – VRS) on output-oriented risk-return models seems more accurate. In a recent contribution, Lamb & Tee (2012) assert that a model with NIRS is appropriate for the study of investment funds. However, this choice is based on the rejection of the other possible RTS assumptions: Constant Returns to Scale (and we suppose Non-Decreasing Returns to Scale) are rejected due to the infeasibility of funds with infinitely high levels of risk and return that are non-attainable in practice, which is actually one of the above arguments to rejected free disposability on risk. The rejection of VRS is due to the fact that it violates the axiom of the possibility of inaction in the input-output space; this problem is solved as soon as risk is defined as an output.

It is also important to notice that the measures chosen to account for input and output variables may once again require a specific treatment: if the distribution of returns is expressed not in monetary units but as rates of return on investment, then scale invariance may even be assumed. Return being traditionally expressed as rates of return, an increase in the quantity invested (free of charges) should remain constant. An assumption of scale invariance can be made on return if expressed as a rate of return and if all additional costs are excluded. In the particular case of some measures like standard deviation of rates of returns, the feasible set of outputs combinations would be the same for all any level of input $\mathbf{x}$. Scale invariance in a multi-moment framework could then be assumed and would translate into equation (19).

Scale invariance : $P(\lambda\mathbf{x}) = P(\mathbf{x}) = P$ for all $\lambda > 0$ and $\mathbf{x} \geq \mathbf{0}$ **(19)**

3.6. Handling negative data

In production theory, positivity conditions are usually assumed on input and output variables (see the conditions in Färe, Grosskopf & Lovell, 1994). While these conditions make sense in a production framework, they have to be adapted since financial assets frequently exhibit negative returns, for instance during recession periods. Basso & Funari (2007) for instance report that 79% of the mutual funds in their sample exhibit a negative average rate of return, and that 86% of these funds exhibit a negative excess return. Dealing with negative data is then a matter of importance not only regarding negative returns, but also regarding skewness and all odd mathematical moments of the distribution that potentially enter into the set of output variables. Moreover, if negative values of return measures are especially problematic when using popular ratios of financial performance (Sharpe, Treynor or reward-to-half-variance indexes), the capacity of DEA to handle negative measures provides supplementary material to advocate for its use on financial assets.

The literature in operations research provides various solutions to handle negative data. An early treatment consists in either performing a change in variables in order to make all data positive (by adding a number – sometimes arbitrary – to all values of a variable) or simply excluding any DMU with negative input or output from the sample. But since the contribution of Ali & Seiford (1990), the translation invariance property for efficiency measures has become a key requirement for any model allowing zero or negative data that should be translated prior to the analysis. Translation invariance ensures finding the same optimal solutions using the original data and the translated data. Lovell & Pastor (1995) and Pastor

(1996) showed that both the additive model of Charnes, Cooper, Golany, Seiford & Stutz (1985) and the weighted additive model of Pastor (1994) are translation invariant. On the contrary, the CCR model, the variant and invariant multiplicative models of Charnes, Cooper, Seiford & Stutz (1982, 1983) and the extended additive model of Charnes, Cooper, Rousseau & Semple (1987) are not translation invariant. They also show that the BCC model is translation invariant only when the translation affects either input variables (when the model is output-oriented) or output variables (when the model is input-oriented) but not both at the same time. Still, as shown in the table of Appendix, many studies with DEA used either BCC or CCR model on financial assets and exclude negative data from their sample.

While Wilkens & Zhu (2001) propose a change in variables for inputs (outputs) when using an input-oriented (output-oriented) model with a radial efficiency measure to satisfy translation invariance, Seiford & Zhu (2002) distinguish between three kinds of translation invariance depending on which outcome is left unchanged after the data is transformed. ‘Classification invariance’ leaves the classification of DMUs as efficient or inefficient unchanged, ‘ordering invariance’ leaves the ranking of DMUs according to their efficiency scores unchanged, and ‘solution invariance’ (often referred to as simply translation invariance) leaves the efficiency scores unchanged. They also propose an approach that ensures classification invariance by correcting a posteriori the classification of DMUs as efficient or inefficient.

Silva Portela, Thanassoulis & Simpson (2004) then propose a model based on directional distance functions that prevents DMUs with negative output to be deemed efficient while some other DMUs with positive output would be deemed inefficient. Pastor & Ruiz (2007), Thanassoulis & Silva Portela & Despic (2008) and Kerstens & Van de Woestyne (2011) provide a more complete review of the literature on that aspect. Following their conclusions we propose to use directional distance functions that ensures translation invariance of the models (see the translation property of Lemma 2.2 in Chambers, Chung & Färe, 1998). The choice of direction vectors must then be in line with the three model orientations that are consistent with theory, as discussed in section 2.4.

5. Key results

5.1. Performance of portfolio selection and performance of portfolio construction

As explained in section 2.1, efficiency of portfolio selection (when the decision-maker wants to invest in one single portfolio) could be measured relative to the FDH, so that the portfolios can be benchmarked one to the others. On the other hand, efficiency of portfolio construction can be measured either relative to the DEA frontier of portfolios (if the decision-maker wants to constitute a portfolio of already constituted portfolios available on the market) or relative to the DEA frontier of individual assets that could compose the portfolios (if the study-maker wants for instance to assess how good was a fund manager at selecting the assets on the market to compose the fund).

When measuring efficiency of portfolio selection, more portfolios are obviously deemed efficient when evaluated relative to a Free Disposal Hull (FDH) frontier than relative to a DEA frontier. Building an FDH technology that includes cash holdings in the sample set of portfolios (to account for the possibility inaction) results in an average number of 8.5 efficient portfolios over the 8 periods when the model is oriented towards risk reduction only, and 7.75 efficient portfolios when the model is oriented towards mean augmentation only. This result makes sense considering these tools serve a different purpose: when using FDH, the study-maker does not look for information on how distant the portfolios are from the potentially attainable frontier – which is the case when using DEA – but rather which portfolio performs better using the sample set as a benchmark. A DEA technology defined as in equation (20) results in an average number of 2.25 efficient portfolios when the model is oriented towards risk reduction only and 1 efficient DMU when the model is oriented towards mean augmentation only. On each time window, this efficient portfolio is the one with the highest mean return; such result would be found by any model oriented towards mean augmentation only when the correlations are taken into consideration, as for all other elements of the efficient frontier, risk is reduced further than any observed level. This specificity is illustrated in Figure 2 where DMU 26 does not belong to the efficient frontier. Finally, a DEA technology as defined in equation (20) from the sample of 920 US stocks results in deeming none of the 50 portfolios efficient, whatever the model orientation or time window. This result makes sense as none of the 50 randomly generated portfolios could replicate any of the 'perfectly' built portfolios of assets of the efficient frontier.

[Insert Table 1] As reported in Table 1, evaluating the efficiency of portfolio construction relative to the DEA frontier of the portfolios themselves or relative to the frontier of their potential holdings results in an average increase in inefficiency scores of 31.33 % over the 8 rolling time windows (among DMUs that are deemed inefficient when performance is measured relative to the DEA frontier of portfolios). The maximal increase in efficiency scores amounts to 465.73 %. [Insert Table 2] Table 2 provides the variations in inefficiency scores obtained from a model that measures efficiency relative to frontier of portfolios to a model that measures efficiency relative to the frontier of individual assets in Period 1, with both models oriented towards variance reduction only. The average observed variation is + 18.5 % on all the 50 portfolios (excluded portfolio 29 that belongs to the efficient frontier of portfolios). For that period, all DMUs with a positive mean return experience an increase in their efficiency score, from + 0.02 % for portfolio 32 – one of the portfolios with the lowest positive mean return on that period – to + 439 % for portfolio 20. As illustrated by the example of portfolio 20, the DMUs that obtained the lowest efficiency scores by being closer to the frontier of portfolios are now as further from the frontier of individual assets than the others. As a consequence, their increase in inefficiency scores is very high, and the same is observed on other time windows. Similarly, the lowest average impact is observed on periods where more funds exhibit negative mean returns.

5.2. Spearman Rank correlations with the Sharpe ratio

[Insert Tables 3 & 4] In order to assess the impact of using a model orientation that simultaneously maximizes risk and return in a technology set defined on the output correspondence, we compared both the inefficiency scores and the rankings obtained from the following models:

- a ‘traditional’ model – described as case (3) in table 4 – oriented towards risk reduction only on a technology that assimilates risk to an input and return to an output, assumes free disposability on risk and return, thus ensuring the technology set is convex by allowing for infinite increases in risk,
- the model of equation (23) – described as case (4) in table 4 – oriented towards risk and return augmentation in the direction of the slope of the regression line of the sample set, on a technology defined in equation (20) that assimilates both risk and return measures to outputs, assumes free disposability on the return measure but costly disposability on the risk measure, with jointness and null jointness between risk and return, and which results in a non-convex technology set.

As illustrated in Table 3, the inefficiency scores increase on average and the impact on the ranking of portfolios is quite substantial: the DMUs gain or lose around 22 ranks on average on each period. A closer look to these variations on the 50 portfolios in period 1 shows in Table 4 a strong decrease of inefficiency scores for DMUs carrying the highest levels of risk, and a strong increase in inefficiency scores for DMUs carrying the lowest levels of risk. The model we propose in equation (23) consequently attributes lower inefficiency scores to risky DMUs, provided that they offer a high enough return.

As the mean-variance framework is fundamental in finance – from the Modern Portfolio Theory to the Capital Asset Pricing Model and its extensions – any methodology proposed with DEA in a mean-variance framework is consistent with the standard approaches in finance. Regarding the Sharpe ratio in particular, Eling (2006) already showed that a DEA approach results in obtaining “*the same ranking and evaluation of the investments as is made by the Sharpe ratio*” in the particular case where the excess return over the risk-free rate and standard deviation are used as measures of return and risk, respectively (as the Sharpe ratio is defined as the ratio of excess return to standard deviation), and the CCR model is used (hence Constant Returns to Scale are assumed). Any other model would however result in a different ranking. In order to assess the applicability in finance of defining the technology set as in equation (20) and using the model of equation (23), we calculated the Spearman rank correlation between the inefficiency scores provided by this model and the Sharpe ratios [24] of the evaluated portfolios on each period. We did so for the two types of portfolio construction: using the definition provided in equation (20) we defined a first output set from our sample of 50 random portfolios, and a second output set from our sample of 920 US stocks that could compose the 50 portfolios. We also calculated Spearman's rank correlation between the Sharpe ratios and the scores generated by a 'traditional model' described as case (3) in Table 4. [Insert Table 5] As shown in Table 5, the rank correlation with the Sharpe ratio is quite high and the highest (except for period 5) for the model that measures performance relative to the frontier of potential portfolio's holdings using a direction that simultaneously increases risk and return measures. The rank correlation between the Sharpe ratio and the model that measures performance relative to the frontier of portfolios using the same direction is quite high as well compared to the one between the Sharpe ratio and the 'traditional' model in the direction of variance reduction only (except for periods 5 to 7, knowing that each period overlaps with the next one by two years).

5.3. Impact of the introduction of jointness to the technology sets

[Insert Tables 6 & 7] Tables 6 & 7 illustrate the impact of the introduction of jointness in the definition of a technology described in case (5) of Table 7 to obtain the technology described in case (6), using a model oriented towards risk reduction only. When models are oriented towards risk reduction and jointness is ignored, the potential risk reduction is underestimated for DMUs for which performance is measured relative to a "weakly efficient" part of the frontier, or what Färe, Grosskopf & Lovell (1994) define as the Weak Efficiency Subset (see Figure 2). The introduction of jointness introduces the possibility of reducing risk through the introduction of cash holdings (a riskless asset with a null return). This way, potential risk reduction of the DMUs with a negative mean return is equal to 100% of their respective levels of risk when jointness is introduced to the model. As long as costly disposability with jointness and null jointness are assumed, performance can be measured relative to the "strongly efficient" part of the frontier for any DMUs with a positive return whatever the model orientation; it ensures that for any decrease in return risk is reduced as well as much as possible. As illustrated in the results of Period 1 (grey rows in Table 7), jointness implies that portfolios with the lowest – yet positive – mean returns in the sample could reduce risk even further. We therefore observe a higher increase in inefficiency scores for DMUs with the lowest positive mean returns as well as losses in their respective rank. On the other hand, introducing jointness attributes a higher rank to the DMUs with higher levels of risk and returns. This evolution in the evaluation of performance the riskiest DMUs is in line with what the literature on preferences that allow for a favorable evaluation of increases in risk. As briefly mentioned at the end of section 2.1, this literature shows how reaching higher levels of risk is not in contradiction with risk aversion. This more favorable evaluation is either due to the simultaneous increase in return together with parameters of the utility function that attribute a higher gain in utility to this increase in return than the loss in utility incurred by the increase in risk, or to a simple preference for increases in risk. This evolution then allows for performance measurement of risky portfolios under the most favorable conditions, which is in line with an underlying principle of the use of DEA.

[24] As our dataset consists in distributions of returns (not excess returns over a risk-free rate), we considered cash as the riskless benchmark to calculate our Sharpe ratio here. It is then a simple ratio of return over standard deviation, which ensures consistency with the model (23) that assumes jointness and null jointness between risk and return.

6. Conclusion

In this paper we provide arguments to support the idea that performance measurement of portfolios of financial assets with Data Envelopment Analysis should not rely on a technology defined through a production process that assimilates risk to an input generating some return. Financial assets differ in their nature and the dependence relationships between their prices from traditional production units, which result in a financial production process that we propose to understand as the generation of a distribution of returns from an initial investment. The set of axioms inherited from production theory is revisited accordingly and we showed that beyond its theoretical basis, the assimilation of risk to an output results in convenient consequences on the consistency of this revisited set of axioms. The resulting definition of a financial technology in a multi-moment framework is also well-suited to the study of financial assets in that it allows taking into consideration model orientations that consider the possibility of an increase in risk measures. We provided illustrations to show how this new definition of the technology and the new model orientations could impact efficiency scores and rankings of the portfolios. These illustrations also reveal huge potential increases in efficiency scores for riskier DMUs, which leads to question the systematic choice of a direction towards risk reduction on markets that are theoretically recognized as efficient. Unless the theoretical frameworks are ill-adapted, such variations in efficiency scores lead to reconsider either the assumption of market efficiency, the definition of the technology or the model orientations. The definition of the technology we provide here also allows including a range of preferences that remain ignored by the practitioners in finance, though studied in the literature. We also provide an example of how a model orientation that takes into account the possibility for increases in variance can obtain rankings of the evaluated portfolios that are in line with traditional ratios used in finance.

References

Ali, A. I. & Seiford, L. M. (1990), Translation Invariance in Data Envelopment Analysis, *Operations Research Letters*, 9(6), pp. 403-405

Anderson, R. I., Brockman, C. M., Giannikos, C. & McLeod, R. W. (2004), A Non-Parametric Examination of Real Estate Mutual Fund Efficiency, *International Journal of Business and Economics*, 3(3), pp. 225-238

Astebro, T. (2003), The Return to Independent Invention: Evidence of Unrealistic Optimism, Risk Seeking or Skeweness Loving?, *The Economic Journal*, 113(484), pp. 226-239

Bali, T. G., Cakici, N. & Whitelaw, R. F. (2010), Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics*, 99(2), pp. 427-446

Ballestero, E. & Pla-Santamaria, D. (2004), Selecting portfolios for mutual funds, *OMEGA*, 32, pp. 385-394

Banker, R. D., Charnes, A. & Cooper, W. W. (1984), Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis, *Management Science*, 30(9), pp. 1078-1092

Barberis, N. & Thaler, R. (2003), A Survey of Behavioral Finance, in Constantinides, G., Harris, M. & Stulz, R. M., Handbook of the Economics of Finance, 1B, Financial Markets and Asset Pricing, North Holland, Chapter 18, pp. 1053-1128

Basso, A. & Funari, S. (2001), A data envelopment analysis approach to measure the mutual fund performance, *European Journal of Operational Research*, 135, pp. 477-492

Basso, A. & Funari, S. (2003), Measuring the Performance of Ethical Mutual Funds: A DEA Approach, *Journal of the Operational Research Society*, 54(5), pp. 521-531

Basso, A. & Funari, S. (2005), A generalized performance attribution technique for mutual funds, *Central European Journal of Operations Research*, 13(1), pp. 65-84

Basso, A. & Funari, S. (2007), DEA models for ethical and non ethical mutual funds, *Mathematical Methods in Economics and Finance*, 2(1), pp. 21-40

Blume, M. & Friend, I. (1973), A New Look at the Capital Asset Pricing Model, *The Journal of Finance*, 28(1), pp. 19-33

Branda, M. (2015), Diversification-consistent data envelopment analysis based on directional-distance measures, *OMEGA*, 52, pp. 65-76

Branda, M. & Kopa, M. (2014), On relations between DEA-risk models and stochastic dominance efficiency tests, *Central Europe Journal of Operational Research*, 22(1), pp. 13-35

Brandouy, O. & Kerstens, K. & Van de Woestyne, I. (2013), Backtesting Superfund Portfolio Strategies Based on Frontier-Based Mutual Fund Ratings, in: Pasourias, F. (Eds.), Efficiency and Productivity Growth: Modelling in the Financial Services Industry, New York, Wiley, pp. 135-170.

Briec, W. & Kerstens, K. (2009), Multi-horizon Markowitz portfolio performance appraisals: A general approach, *OMEGA*, 37, pp. 50-62

Briec, W. & Kerstens, K. (2010), Portfolio Selection in Multidimensional General and Partial Moment Space, *Journal of Economic Dynamics & Control*, 34, pp. 646-656

Briec, W., Kerstens, K. & Jokung, O. (2007), Mean-Variance-Skewness Portfolio Performance Gauging: A General Shortage Function and Dual Approach, *Management Science*, 53(1), pp. 135-149

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, pp. 1-27

Chambers, R. G., Chung, Y. & Färe, R. (1998), Profit, Directional Distance Functions, and Nerlovian Efficiency, Journal of Optimization Theory and Applications, 98(2), pp. 351-364

Charnes, A., Cooper, W. W., Golany, B., Seiford, L. M. & Stutz, J. (1985), Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions, *Journal of Econometrics*, 30, pp. 91-107

Charnes, A., Cooper, W. W. & Rhodes, E. (1978), Measuring the efficiency of decision making units, *European Journal of Operational Research*, 2(6), pp. 429-444

Charnes, A., Cooper, W. W., Rousseau, J. J. & Semple, J. H. (1987), Data Envelopment Analysis and Axiomatic Notions of Efficiency and Reference Sets, Research Report CCS 558, Center for Cybernetic Studies, University of Texas, Austin, TX, (USA)

Charnes, A., Cooper, W.W., Seiford, L. M. & Stutz, J. (1982), A Multiplicative Model for Efficiency Analysis, *Socio-Economic Planning Sciences*, 16(5), pp. 223-224

Charnes, A., Cooper, W.W., Seiford, L. M. & Stutz, J. (1983), Invariant multiplicative efficiency and piecewise Cobb-Douglas envelopments, *Operations Research Letters*, 2(3), pp. 101-103

Choi, Y. K. & Murthi, B. P. S. (2001), Relative Performance Evaluation of Mutual Funds: A Non-Parametric Approach, *Journal of Business Finance and Accounting*, 28(7 & 8), pp. 853-976

Cook, W. D., Tone, K. & Zhu, J. (2014), Data envelopment analysis: Prior to choosing a model, *OMEGA*, 44, pp. 1-4

Daraio, C. & Simar, L. (2006), A robust nonparametric approach to evaluate and explain the performance of mutual funds, *European Journal of Operational Research*, 175, pp. 516-542

Devaney, M., Morillon, T. & Weber, W. (2016), Mutual fund efficiency and tradeoffs in the production of risk and return, *Managerial Finance*, 42(3), pp. 225-243

Devaney, M. & Weber, W. (2005), Efficiency, Scale Economies, and the Risk/Return Performance of Real Estate Investment Trusts, *The Journal of Real Estate Finance and Economics*, 31(3), pp. 301-317

Eling, M. (2006), Performance measurement of hedge funds using data envelopment analysis, *Financial Markets Portfolio Management*, 20, pp. 442-471

Färe, R. & Grosskopf, S. (2003), Nonparametric Productivity Analysis with Undesirable Outputs: Comments, *American Journal of Agricultural Economics*, 85(4), pp. 1070-1074

Färe, R. & Grosskopf, S. (2004), Modeling undesirable factors in efficiency evaluation: Comment, *European Journal of Operational Research*, 157, pp. 242-245

Färe, R., Grosskopf, S. & Lovell, C. A. K. (1994), Production Frontiers, Cambridge University Press

Galagedera, Don U. A. & Silvapulle, P. (2002), Australian mutual fund performance appraisal using data envelopment analysis, *Managerial Finance*, 28(9), pp. 60-73

Glawischnig, M. & Sommersguter-Reichmann, M. (2010), Assessing the performance of alternative investments using non-parametric efficiency measurement approaches: Is it convincing?, *Journal of Banking and Finance*, 34, pp. 295-303

Golee, J. & Tamarkin, M. (1998), Bettors love skewness, not risk, at the horse track, *Journal of Political Economy*, 106, pp. 205-225

Gregoriou, G. N. (2003), Performance appraisal of funds of hedge funds using data envelopment analysis, *Journal of Wealth Management*, 5, pp. 88-95

Gregoriou, G. N., Rouah, F., Satchell, S. E. & Diz, F. (2005), Simple and cross efficiency of CTAs using data envelopment analysis, *European Journal of Finance*, 11, pp. 393-409

Gregoriou, G. N., Sedzro, K., Zhu, J. (2005), Hedge fund performance appraisal using data envelopment analysis, *European Journal of Operational Research*, 164, pp. 555-571

Gregoriou, G. N. & Zhu, J. (2005), Evaluating Hedge Fund and CTA Performance: Data Envelopment Analysis Approach, Vol. 279, John Wiley & Sons Inc.

Gregoriou, G. N. & Zhu, J. (2007), Data Envelopment Analysis, *Journal of Portfolio Management*, 33(2), pp. 120-132

Grinblatt, M. & Titman, S. (1989), Mutual Fund Performance: an Analysis of Quarterly Portfolio Holdings, *Journal of Business*, 62, pp. 394-415

Hailu, A. & Veeman, T. S. (2001), Non-Parametric Productivity Analysis with Undesirable Outputs: An Application to the Canadian Pulp and Paper Industry, *American Journal of Agricultural Economics*, 83(3), pp. 605-616

Kerstens, K. & Van De Woestyne, I. (2011), Negative data in DEA: a simple proportional distance function approach, *Journal of the Operational Research Society*, 62, pp. 1413-1419

Kneip, A., Park, B. U. & Simar, L. (1998), A note on the convergence of nonparametric DEA estimators for production efficiency scores, *Journal of Econometrics*, 14, pp. 783-793

Kuosmanen, T. (2005), Weak Disposability in Nonparametric Production Analysis with Undesirable Outputs, *American Journal of Agricultural Economics*, 87(4), pp. 1077-1082

Lamb, J. D. & Tee, K.-H. (2012), Data envelopment analysis models of investment funds, *European Journal of Operational Research*, 216, pp. 687-696

Liu, W., Zhou, Z., Liu, D. & Xiao, H. (2015), Estimation of portfolio efficiency via DEA, *OMEGA*, 52, pp. 107-118

Lovell, C. A. K. & Pastor, J. T. (1995), Units invariant and translation invariant DEA models, *Operations Research Letters*, 18, pp. 147-151

Lozano, S. & Gutierrez, E. (2008), TSD-Consistent Performance Assessment of Mutual Funds, *Journal of Operational Research Society*, 59(10), pp. 1352-1362

Markowitz, H. M. (1952), Portfolio selection, *Journal of Finance*, 7(1), pp. 77-91

McMullen, P. R. & Strong, R. A. (1998), Selection of Mutual Fund Using Data Envelopment Analysis, *Journal of Business and Economic Studies*, 4(1), pp. 1-14

Nguyen-Thi-Thanh, H. (2007), On the use of data envelopment analysis in hedge fund performance appraisal, *Money Macro and Finance (MMF)*, No 131, Research Group Conference 2006, Money Macro and Finance Research Group

Morey, M. R. & Morey, R. C. (1999), Mutual fund performance appraisals: a multi-horizon perspective with endogenous benchmarking, *OMEGA*, 27, pp. 241-258

Murthi, B. P. S., Choi, Y. K. & Desai, P. (1997), Efficiency of mutual funds and portfolio performance measurement: A non-parametric approach, *European Journal of Operational Research*, 98, pp. 408-418

Pastor, J. T. (1994), New Additive DEA Models for handling Zero and Negative Data, Departamento de Estadistica e Investifacion Operativa, Universidad de Alicante, 03071 Alicante, Spain

Pastor, J. T. (1996), Translation invariance in data envelopment analysis: A generalization, *Annals of Operations Research*, 66, pp. 93-102

Pastor, J. T. & Ruiz, J. L. (2007), Variables with negative values in DEA, in Zhu, J. & Cook, W. D., Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis, Springer US, pp. 63-84

Porter, R. B. & Gaumnitz, J. E. (1972), Stochastic Dominance vs. Mean-Variance Portfolio Analysis: An Empirical Evaluation, *American Economic Review*, 62(3), pp. 438-446

Premachandra, I., Powell, J. G. & Shi, J. (1998), Measuring the Relative Efficiency of Fund Management Strategies in New Zealand Using a Spreadsheet-based Stochastic Data Envelopment Analysis Model, *OMEGA*, 26(2), pp. 319-331

Rotschild, M. & Stiglitz, J. E. (1970), Increasing Risk I: A definition, *Journal of Economic Theory*, 2, pp. 225-243

Russell, R. R. (1985), Measures of technical efficiency, *Journal of Economic Theory*, 35, pp. 109-126

Seiford, L. M. & Zhu, J. (2002), Classification invariance in data envelopment analysis, in Misra, J. C. (Eds.), Uncertainty and Optimality, Probability, Statistics and Operations Research, World Scientific

Sengupta, J. K. (1989), Nonparametric Tests of Efficiency of Portfolio Investment, *Zeitschrift für Nationalökonomie (Journal of Economics)*, 50(1), pp. 1-15

Sengupta, J. K. (2003), Efficiency tests for mutual fund portfolios, *Applied Financial Economics*, 13, pp. 869-876.

Silva Portela, M. C., Thanassoulis, E. & Simpson, G. (2004), Negative Data in DEA: a Directional Distance Approach Applied to Bank Branches, *The Journal of the Operational Research Society*, 55(10), pp. 1111-1121

Simar, L. & Wilson, P. W. (2000), Statistical inference in nonparametric frontier models: The state of the art, *Journal of Productivity Analysis*, 13, pp. 49-78

Thanassoulis, E., Silva Portela, M. & Despic, O. (2008), DEA – The mathematical programming approach to efficiency analysis, in Fried, H. Lovell, C. & Schmidt, S. (2008), The Measurement of Productive Efficiency and Productivity Change, Oxford University Press: New York, pp. 251-420

Wilkens, K. & Zhu, J. (2001), Portfolio Evaluation and Benchmark Selection: A Mathematical Programming Approach, *Journal of Alternative Investments*, 4(1), pp. 9-19

Tables

Table 1

Period :	1	2	3	4	5	6	7	8	
Average variation of the inefficiency scores	+ 18.49%	+ 0.17%	+ 29.37%	+ 42.80%	+ 47.62%	+ 44.25%	+ 42.54%	+ 25.38%	**+ 31.33%**
Maximal variation of the inefficiency scores	+ 439.04%	+ 5.53%	+ 465.73%	+ 300.45%	+ 152.00%	+ 255.05%	+ 173.43%	+ 115.77%	**+ 465.73%**

Table 2

Portfolio	**Inefficiency measured relative to the frontier of portfolios** (max. decrease in variance x10,000)		**Variation in inefficiency score**	**Variation in inefficiency score**
	(1)	**(2)**		
1	9.9546	10.0001	+ 0.46%	0.46%
2	10.1046	10.1046	-	0.00%
3	8.4394	9.5424	+ 13.07%	13.07%
4	9.3128	9.8448	+ 5.71%	5.71%
5	11.5676	11.7613	+ 1.67%	1.67%
6	9.9253	9.9253	-	0.00%
7	7.4708	9.0971	+ 21.77%	21.77%
8	8.9192	9.1987	+ 3.13%	3.13%
9	6.1587	10.0372	+ 62.98%	62.98%
10	6.2278	6.9042	+ 10.86%	10.86%
11	8.2682	8.2815	+ 0.16%	0.16%
12	6.5989	8.2815	+ 25.50%	25.50%
13	9.4629	9.4629	-	0.00%
14	7.8341	8.9087	+ 13.72%	13.72%
15	8.0689	8.0689	-	0.00%
16	10.8961	10.9176	+ 0.20%	0.20%
17	8.6236	8.6472	+ 0.27%	0.27%
18	8.0120	8.0447	+ 0.41%	0.41%
19	9.8947	9.8947	-	0.00%
20	1.3323	7.1819	+ 439.04%	439.04%
21	8.3576	8.4066	+ 0.59%	0.59%
22	6.5900	7.6386	+ 15.91%	15.91%
23	9.2063	9.7758	+ 6.19%	6.19%
24	11.1728	12.6549	+ 13.27%	13.27%
25	8.5485	8.5485	-	0.00%
26	5.2502	9.8191	+ 87.02%	87.02%
27	8.0792	8.0792	-	0.00%
28	7.6970	7.6970	-	0.00%
29	0.0000	12.3186	+ ∞	+ ∞
30	7.7328	10.3006	+ 33.21%	33.21%
31	10.0856	10.4479	+ 3.59%	3.59%
32	9.1568	9.1583	+ 0.02%	0.02%
33	7.1797	7.3625	+ 2.55%	2.55%
34	6.3273	6.3419	+ 0.23%	0.23%
35	6.6642	8.5939	+ 28.96%	28.96%
36	7.3006	7.7032	+ 5.51%	5.51%
37	9.7791	9.7825	+ 0.03%	0.03%
38	9.8340	9.8340	-	0.00%
39	9.6831	11.0551	+ 14.17%	14.17%
40	6.5511	6.7022	+ 2.31%	2.31%
41	7.8975	7.8975	-	0.00%
42	7.5574	7.5574	-	0.00%
43	8.5853	8.5853	-	0.00%
44	7.8766	10.2989	+ 30.75%	30.75%
45	8.5711	11.9207	+ 39.08%	39.08%
46	8.6645	8.6645	-	0.00%
47	9.3078	11.3419	+ 21.85%	21.85%
48	7.0874	7.0874	-	0.00%
49	9.0816	9.0816	-	0.00%
50	7.7177	7.8517	+ 1.74%	1.74%

Average variation : **+ 18.49%**

Max : **+ 439.04%**

Evaluated DMUs:

- 50 random portfolios

(1)

- technology P^1 built from the sample set of 50 random portfolios
- model orientation : variance minimization

(2)

- technology P^1 built from the sample set of 920 US stocks
- model orientation : variance minimization

Table 3

Period :	1	2	3	4	5	6	7	8	
Average variation of the inefficiency scores	- 15.34%	+ 9.98%	+ 276.46%	+ 195.12%	+ 772.56%	+ 510.06%	+ 492.07%	- 6.88%	**+ 279.26%**
Maximal variation of the inefficiency scores	+ 187.60%	+ 1,103.22%	+ 5,076.68%	+ 2,075.97%	+ 3,212.69%	+ 1,791.49%	+ 1,601.74%	+ 152.94%	**+ 5,076.68%**
Average variation of the ranks	22.94	21.50	21.14	21.18	22.10	22.70	22.54	21.62	**21.97**
Maximal variation of the ranks	49	49	48	48	48	46	47	46	**49**

Table 4

Portfolio	(3) Inefficiency (max. decrease in variance x10,000)	(3) Rank	(4) Inefficiency (max. decrease in variance x10,000)	(4) Rank	Variation in inefficiency score	Variation in ranking
1	5.8103	42	2.6982	17	- 53.56%	+ 25
2	5.9135	43	2.5950	15	- 56.12%	+ 28
3	5.2311	31	2.9536	23	- 43.54%	+ 8
4	5.6415	37	2.6685	16	- 52.70%	+ 21
5	7.5755	49	0.9325	5	- 87.69%	+ 44
6	5.7342	41	2.7743	19	- 51.62%	+ 22
7	4.6592	26	3.3831	26	- 27.39%	-
8	5.0117	30	3.3222	25	- 33.71%	+ 5
9	4.9442	28	2.1920	14	- 55.67%	+ 14
10	2.6794	5	5.6047	47	+ 109.17%	- 42
11	4.0908	18	4.2474	36	+ 3.83%	- 18
12	5.2839	33	2.0909	10	- 60.43%	+ 23
13	5.2718	32	3.2367	24	- 38.60%	+ 8
14	4.6040	25	3.5881	28	- 22.07%	- 3
15	3.8778	15	4.6307	39	+ 19.42%	- 24
16	6.7271	46	1.7814	9	- 73.52%	+ 37
17	4.4568	22	3.8814	31	- 12.91%	- 9
18	3.8545	14	4.4837	37	+ 16.32%	- 23
19	5.7035	40	2.8050	20	- 50.82%	+ 20
20	1.2977	2	0.4587	3	- 64.66%	- 1
21	4.2170	19	4.1212	34	- 2.27%	- 15
22	3.3399	8	4.8590	43	+ 45.48%	- 35
23	5.5679	34	2.7363	18	- 50.86%	+ 16
24	8.2527	50	0.0000	1	- 100.00%	+ 49
25	4.3574	20	4.1511	35	- 4.73%	- 15
26	4.4832	24	1.6064	8	- 64.17%	+ 16
27	3.8880	16	4.6205	38	+ 18.84%	- 22
28	3.5058	10	5.0027	45	+ 42.70%	- 35
29	0.0000	1	0.0000	1	-	-
30	5.6165	36	2.1511	12	- 61.70%	+ 24
31	6.2574	44	2.0944	11	- 66.53%	+ 33
32	4.9672	29	3.5413	27	- 28.71%	+ 2
33	3.1766	7	5.1614	46	+ 62.48%	- 39
34	2.1512	3	6.1869	50	+ 187.60%	- 47
35	4.0795	17	3.8771	30	- 4.96%	- 13
36	3.5104	11	4.8140	42	+ 37.14%	- 31
37	5.5914	35	2.9171	22	- 47.83%	+ 13
38	5.6428	38	2.8657	21	- 49.21%	+ 17
39	6.6798	45	1.4327	7	- 78.55%	+ 38
40	2.5156	4	5.8226	49	+ 131.46%	- 45
41	3.7063	13	4.8022	41	+ 29.57%	- 28
42	3.3663	9	4.9719	44	+ 47.70%	- 35
43	4.3942	21	4.1143	33	- 6.37%	- 12
44	5.6553	39	2.1573	13	- 61.85%	+ 26
45	7.0019	48	0.5074	4	- 92.75%	+ 44
46	4.4733	23	4.0352	32	- 9.79%	- 9
47	6.8007	47	1.1260	6	- 83.44%	+ 41
48	2.8962	6	5.6123	48	+ 93.78%	- 42
49	4.8904	27	3.6181	29	- 26.02%	- 2
50	3.6646	12	4.6736	40	+ 27.53%	- 28
Averages :	**4.6604**		**3.2782**		**+ 15.34%**	**+ 22.94**
					Max	**+ 49**
					Min	**- 47**

Evaluated DMUs:

- 50 random portfolios

(3)

- technology built from the sample set of 50 random portfolios under the following assumptions :
  - input : risk measure
  - output : return measure
  - free disposability on risk and return
  - convexity of the technology set
- model orientation : variance minimization

(4)

- technology P^1 built from the sample set of 50 random portfolios under the following assumptions
  - input : initial investment
  - outputs : risk and return measures
  - free disposability on the return measure
  - costly disposability on the risk measure
  - jointness and null jointness between risk and return
  - no convexity of the technology set
- model orientation : simultaneous mean and variance maximization

Table 5

Spearman's rank correlation between the Sharpe ratio with cash used as the riskless asset and …	Period							
	1	2	3	4	5	6	7	8
… inefficiency scores of a 'traditional' model described as case (3) in Table 5	-0.101	0.013	0.007	0.190	0.429	0.254	0.192	0.066
… inefficiency scores of model (23) in the technology set built from the 50 random portfolios of US stocks	0.517	0.696	0.771	0.611	-0.105	0.027	0.089	0.634
… inefficiency scores of model (23) in the technology set built from the 920 individual US stocks that could compose the portfolios	0.999	0.986	0.952	0.853	0.388	0.698	0.565	0.998

Table 6

Period :	1	2	3	4	5	6	7	8	
Average variation of the inefficiency scores	+ 80.54%	+ 137.56%	+ 94.24%	+ 50.48%	+ 26.84%	+ 40.01%	+ 31.01%	+ 37.48%	**+ 62.27%**
Maximal variation of the inefficiency scores	+ 194.12%	+ 691.64%	+ 1,210.70%	+ 445.39%	+ 148.19%	+ 227.27%	+ 164.59%	+ 93.40%	**+ 1,210.70%**
Average variation of the ranks	+ 6.8	+ 4.36	+ 5.44	+ 6.36	+ 5.32	+ 6.88	+ 5.92	+ 8.72	**+ 6.23**
Maximal variation of the ranks	+ 24	+ 9	+ 26	+ 18	+ 14	+ 19	+ 17	+ 22	**+ 26**

Table 7

Portfolio	(5) Inefficiency (max. decrease in variance x10,000)	(5) Rank	(6) Inefficiency (max. decrease in variance x10,000)	(6) Rank	Variation in inefficiency score	Variation in ranking
1	5.8103	42	9.9546	45	+ 71.33%	- 3
2	5.9135	43	10.1046	47	+ 70.87%	- 4
3	5.2311	31	8.4394	27	+ 61.33%	+ 4
4	5.6415	37	9.3128	38	+ 65.08%	- 1
5	7.5755	49	11.5676	50	+ 52.70%	- 1
6	5.7342	41	9.9253	44	+ 73.09%	- 3
7	4.6592	26	7.4708	14	+ 60.34%	+ 12
8	5.0117	30	8.9192	33	+ 77.97%	- 3
9	4.9442	28	6.1587	4	+ 24.56%	+ 24
10	2.6794	5	6.2278	5	+132.43%	
11	4.0908	18	8.2682	25	+102.12%	- 7
12	5.2839	33	6.5989	9	+ 24.89%	+ 24
13	5.2718	32	9.4629	39	+ 79.50%	- 7
14	4.6040	25	7.8341	19	+ 70.16%	+ 6
15	3.8778	15	8.0689	23	+108.01%	- 8
16	6.7271	46	10.8961	48	+ 61.97%	- 2
17	4.4568	22	8.6236	31	+ 93.49%	- 9
18	3.8545	14	8.0120	22	+107.86%	- 8
19	5.7035	40	9.8947	43	+ 73.48%	- 3
20	1.2977	2	1.3323	2	+ 2.67%	
21	4.2170	19	8.3576	26	+ 98.19%	- 7
22	3.3399	8	6.5900	8	+ 97.31%	
23	5.5679	34	9.2063	36	+ 65.35%	- 2
24	8.2527	50	11.1728	49	+ 35.38%	+ 1
25	4.3574	20	8.5485	28	+ 96.19%	- 8
26	4.4832	24	5.2502	3	+ 17.11%	+ 21
27	3.8880	16	8.0792	24	+107.80%	- 8
28	3.5058	10	7.6970	16	+119.55%	- 6
29	0.0000	1	0.0000	1	-	
30	5.6165	36	7.7328	18	+ 37.68%	+ 18
31	6.2574	44	10.0856	46	+ 61.18%	- 2
32	4.9672	29	9.1568	35	+ 84.35%	- 6
33	3.1766	7	7.1797	12	+126.02%	- 5
34	2.1512	3	6.3273	6	+194.12%	- 3
35	4.0795	17	6.6642	10	+ 63.36%	+ 7
36	3.5104	11	7.3006	13	+107.97%	- 2
37	5.5914	35	9.7791	41	+ 74.90%	- 6
38	5.6428	38	9.8340	42	+ 74.27%	- 4
39	6.6798	45	9.6831	40	+ 44.96%	+ 5
40	2.5156	4	6.5511	7	+160.42%	- 3
41	3.7063	13	7.8975	21	+113.08%	- 8
42	3.3663	9	7.5574	15	+124.50%	- 6
43	4.3942	21	8.5853	30	+ 95.38%	- 9
44	5.6553	39	7.8766	20	+ 39.28%	+ 19
45	7.0019	48	8.5711	29	+ 22.41%	+ 19
46	4.4733	23	8.6645	32	+ 93.69%	- 9
47	6.8007	47	9.3078	37	+ 36.87%	+ 10
48	2.8962	6	7.0874	11	+144.71%	- 5
49	4.8904	27	9.0816	34	+ 85.70%	- 7
50	3.6646	12	7.7177	17	+110.60%	- 5
	4.6604		**8.0923**		**+ 80.54%**	**+ 6.8**
				Max	**+ 194.12%**	**+ 24**
				Min	**+ 2.67%**	**- 9**

Evaluated DMUs:

- 50 random portfolios

(5)

- technology built from the sample set of 50 random portfolios under the following assumptions
  - input : initial investment
  - outputs : risk and return measures
  - free disposability on the return measure
  - free disposability on the risk measure
  - no jointness on the measures of risk and return
  - convexity of the technology set
- model orientation : variance minimization

(6)

- technology built from the sample set of 50 random portfolios under the following assumptions
  - input : initial investment
  - outputs : risk and return measures
  - free disposability on the return measure
  - free disposability on the risk measure
  - jointness and null jointness on the measures of risk and return
  - convexity of the technology set
- model orientation : variance minimization

Figures

Figure 1 Feasible combinations of three funds in a mean-standard deviation space

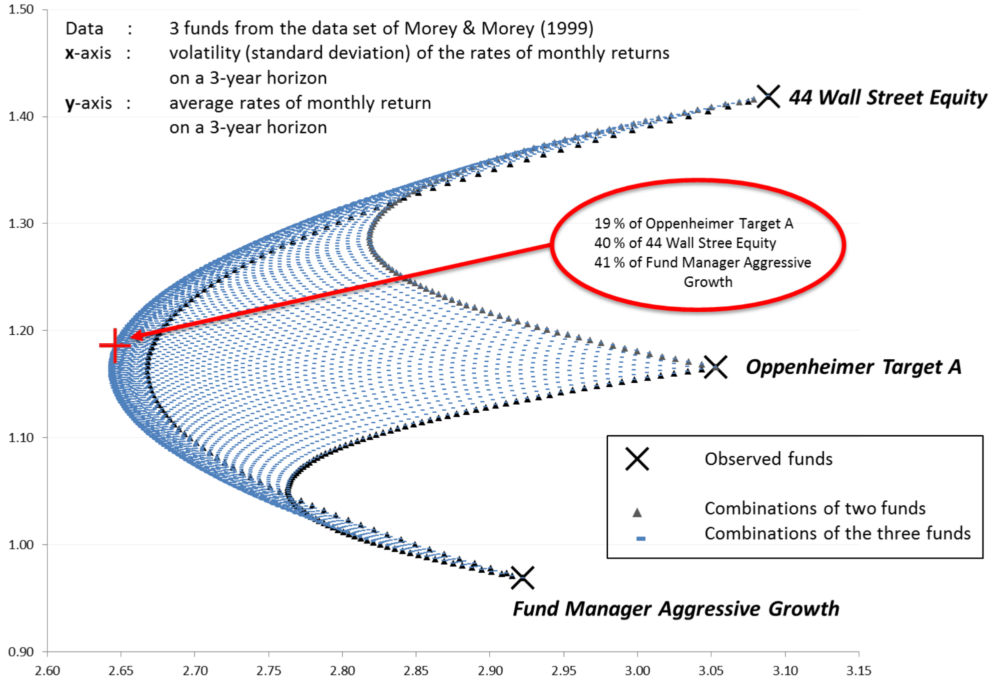
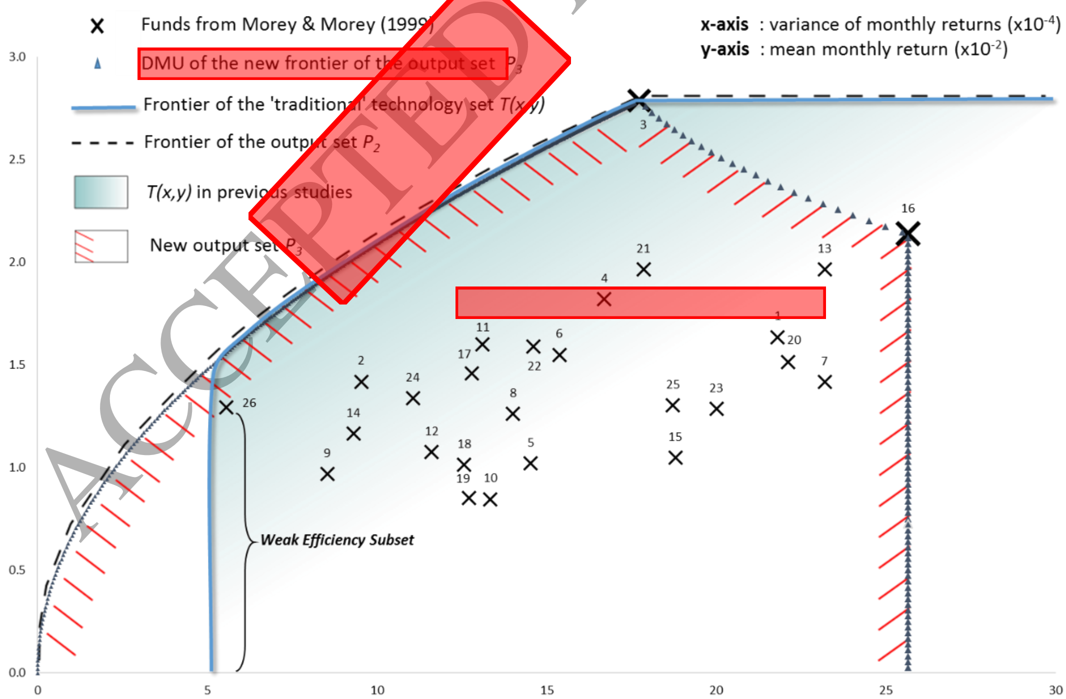


Figure 2 Portfolio possibility sets under various approaches



Appendix

	Input variables	Risk measures	Output variables	Return measures	Model	Data	Convexity of the set	Considers correlations?	Superior moments
Murthi, Choi & Desai (1997)	**Risk measure** Expense ratio Front-end Load Turnover	Standard deviation of returns	**Return measure**	Excess return	Input-oriented DEA CCR	MUTUAL FUNDS	YES	NO	NO
McMullen & Strong (1998)	**Risk measure** Expense ratio Sales charges Minimum initial investment	Standard deviation of returns	**Return measures**	Return before tax	Input and output-oriented DEA BCC	MUTUAL FUNDS	YES	NO	NO
Premachandra, Powell & Shi (1998)	Initial investment (dollar value)	-	Portfolio's total market value	Difference between market value and a benchmark return	Input oriented SDEA based on DEA CCR with slacks	Portfolios of stocks	YES	NO	N/A
Morey & Morey (1999)	**Risk measures** (systematic and non-systematic risks)	Variance of returns	**Return measures**	Mean returns	Input and output-oriented DEA BCC	MUTUAL FUNDS	YES	YES	NO
Basso & Funari (2001)	**3 risk measures** 2 investment costs	Standard deviation of returns, standard semi-deviation of returns, beta coefficient	**Return measure** Stochastic dominance indicator	Expected (excess) return	Input-oriented DEA CCR	MUTUAL FUNDS	YES	Indirectly, in the beta coefficient	Indirectly (stochastic dominance indicator of degree 3)
Choi & Murthi (2001)	**Risk measure** Expense ratio Loads Turnover ratio	Standard deviation of annualized 3-year return	**Return measure**	Gross returns (before any deduction of expenses)	Input-oriented DEA BCC with slacks	MUTUAL FUNDS	YES	NO	NO
Wilkens & Zhu (2001)	**Risk measure** % of neg. Returns	Standard deviation of returns	**Return measures** Skewness	Average return Minimum return	Input-oriented DEA BCC	CTAs	YES	NO	YES (of order 3)
Galagedera & Silvapulle (2002)	**Risk measures** Expense ratio Minimum initial investment Sales charges Entry fee	Standard deviation of returns	**Return measures**	Gross returns	Input-oriented DEA BCC	MUTUAL FUNDS	YES	NO	NO
Sengupta (2003)	**Risk measures** Load Expense ratio Turnover	beta coefficient	**Return measures** Skewness	Mean return Skewness	Input-oriented DEA BCC	MUTUAL FUNDS	YES	NO	YES (of order 3)
Basso & Funari (2003)	**Risk measures** Subscription and redemption costs	Standard deviation of returns, standard semi-deviation of returns, beta coefficient	**Return measure** Ethical indicator	Expected (excess) return	Input and Output-oriented DEA CCR with slacks	Ethical MUTUAL FUNDS	YES	NO	NO
Gregoriou (2003)	**Risk measures**	Lower partial moments of orders 1, 2 and 3	**Return measures**	Higher partial moments of orders 1, 2 and 3	Input-oriented DEA BCC, cross and super efficiency	HEDGE FUNDS	YES	NO	YES (of order 3)
Andersen, Brockman,	**Risk measure** Front Load,	Standard deviation of	**Return measure**	Actual returns	Input-oriented DEA	Real Estate MUTUAL	YES	NO	NO

Giannikos & McLeod (2004)	Deferred load, Marketing & distribution fees Other expenses	returns			CCR	FUNDS			
Briec, Kerstens & Lesourd (2004)	**Risk measure**	Variance of returns	**Return measure**	3-year rate of return	Directional distance DEA model	INVESTMENT FUNDS	NO, but free disposability makes it convex	YES	NO
Gregoriou, Rouah, Satchell & Diz (2005)	**Risk measures**	Lower partial moments of orders 1, 2 and 3	**Return measures**	Higher partial moments of orders 1, 2 and 3	Input-oriented DEA BCC and cross efficiency	CTAs	YES	NO	YES (of order 3)
Gregoriou, Sedzro & Zhu (2005)	**Risk measures**	Lower partial moments of orders 1, 2 and 3	**Return measures**	Higher partial moments of orders 1, 2 and 3	Input-oriented DEA BCC, Cross and super efficiency	HEDGE FUNDS	YES	NO	YES (of order 3)
Daraio & Simar (2006)	**Risk measure** Expense ratio Turnover Fund size	Standard deviation of returns	**Return measure**	Total return	Output-oriented FDH	MUTUAL FUNDS	NO (FDH)	NO	NO
Eling (2006)	**Risk measures**	Standard deviation of returns, Lower Partial Moments of ranks 1 to 3	**Return measures**	Excess return, Higher Partial Moments of ranks 1 to 3, average return, skewness, minimum returns	CCR and BCC models, 2 super efficiency models	HEDGE FUNDS	YES	NO	YES
Basso & Funari (2007)	**Risk measure** Initial fee Exit fee	Standard deviation of returns	**Return measure** Initial capital invested (= 1) Ethical indicator	Mean return, capitalization factor	Input-oriented CCR model with slacks	MUTUAL FUNDS	YES	NO	NO
Nguyen-Thi-Thanh (2007)	**Risk measure** Excess Kurtosis	Standard deviation of returns	**Return measure** Skewness	Average return	Input-oriented DEA CCR with slacks	HEDGE FUNDS	YES	NO	YES (of orders 3 & 4)
Glawischnig & Sommersguter-Reichmann (2010)	**Risk measures**	Standard deviation, Lower partial moments of orders 0 to 3, Maximum drawdown period	**Return measures**	Average and compounded returns, Max. consecutive gain, Upper partial moments of orders 1 to 4	Input-oriented DEA BCC with slacks	MANAGED FUTURES	YES	NO	YES (of orders 3 & 4)
Branda & Kopa (2012)	**Risk measures**	Standard deviation, Value-at-Risk, Conditional and drawdown Value-at-Risk	**Return measure**	Gross mean returns	Mean-Risk, Input-oriented DEA CCR and Stochastic Dominance	Stock Indices	YES	NO	no, but use 1^{st} & 2^{nd} order Stochastic Dominance