Ca' Foscari University of Venice

Department of Economics

Working Paper

Monica Billio, Massimiliano Caporin and Michele Costola

Backward/forward optimal combination of performance measures for equity screening

ISSN: 1827-3580
No. 13/WP/2012

Working Papers
Department of Economics
Ca' Foscari University of Venice
No. 13/WP/2012
ISSN 1827-3580

Backward/forward optimal combination of performance measures for equity screening

Monica Billio
Università Ca' Foscari Venezia

Massimiliano Caporin
Università di Padova

Michele Costola
Università di Padova

Abstract
We introduce a novel criterion for performance measure combination designed to be used as an equity screening algorithm. The proposed approach follows the general idea of linearly combining existing performance measures with positive weights and the combination weights are determined by means of an optimisation problem. The underlying criterion function takes into account the risk-return trade-off potentially associated with the equity screens, evaluated on a historical and rolling basis. By construction, performance combination weights can vary over time, allowing for changes in preferences across performance measures. An empirical example shows the benefits or our approach compared to naive screening rules based on the Sharpe ratio.

Keywords
Performance measures, combining performance measures, portfolio allocation, equity screening, differential evolution.
JEL Codes
C44, C58, C61, G11, G17.

Address for correspondence:
Monica Billio
Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S.Giobbe
30121 Venezia - Italy
Phone: (++39) 041 2349170
Fax: (++39) 041 2349176
e-mail: billio@unive.it

1 Introduction

The investment process is given by the collection of actions taken by a portfolio manager, starting from the definition of the investment objectives and the associated strategic allocation, and including the construction of tactical asset allocation, security selection choices, and general rules for portfolio monitoring (see for example Grinhold and Kahn 1999). The security selection step focuses on identifying of the most promising investment opportunities, represented by specific assets. Different approaches might be employed at this stage, inspired by technical analysis or based on a more fundamental point of view. In general, security selection methodologies can be classified as qualitative or quantitative. The latter presumes the existence and the use of some quantitative tools. The broad class of quantitative security selection instruments includes the so-called equity screening rules, methodologies whose to rank a large set of asset in order to focus attention on the best ones or to exclude the worst ones. Screening rules can be used directly as security selection tools or might represent a first step in a security selection program; in fact, they restrict the investment universe to a reasonably limited set of assets, to be analysed in greater detail by analysts. However, screening rules should not be used directly as asset allocation tools (for instance by directly investing in the assets with the highest rank), since they do not control for the correlation across assets. Relevant and relatively simple examples of screening rules are given by performance measures; these are quantities that, in most cases, represent a remuneration per unit of risk, or risk adjusted returns. In the last decade, the financial economics literature has discussed a large number of alternative performance measures; see the surveys by Aftalion and Poncet (2003), Le Sourd (2007), Bacon (2008), Cogneau and Hubner (2009a,b) and Caporin et al. (2011). The available performance indices can be classified into large families, as suggested by Caporin et al. (2011), to highlight their differences: relative performance measures (rewards per unit of risk), absolute performance measures (risk-adjusted measures referred to a benchmark or to a set of risk factors), measures derived from utility functions and measures expressed as functions of return distribution features. Note also that performance measures belonging to the same class are heterogeneous since they can be based on different quantities (such as utility functions, moments, partial moments or quantiles) or different information sets (different choices of risk factors). Furthermore, if performance measures are used to order assets (as equity screening rules), the ranks they produce for a common set of assets might be sensibly different; see Caporin and Lisi (2011). The last finding confirms that alternative measures have different views over assets, and the construction of an 'optimal' equity screening tool should take those different viewpoints into account.

A possible solution is the construction of a composite performance index to be used within an equity screening program. Few authors have considered this approach and, to our best knowledge, the only published reference is the work of Hwang and Salmon (2003), in which the authors propose a combination based on copula functions. We contribute to this strand of the quantitative finance literature by introducing a new approach for the construction of a composite performance index. Our proposal lies in between security selection and asset allocation, since our composite index is determined within a pre-specified equally weighted asset allocation scheme. Such a choice, despite being restrictive, is motivated by the recent contributions of De Miguel et al. (2009), who show evidence of the (statistical) equivalence between the performances of equally weighted portfolios and optimised (in a Markowitz sense) portfolios. However, the limitation of the calibrated portfolio weights might be removed, even if such an extension is not empirically considered in the current paper. The fixing of portfolio weights also simplifies the identification of the composite performance index, which is given by an optimal linear combination of a set of performance measures.

Within the performance evaluation framework, several authors have considered the problem of determining the optimal portfolio weights by maximizing different performance measures. Their aimed at finding the 'best' performance measure; see for example Farinelli et al. (2008, 2009), among others. The outcomes of these studies were not completely conclusive, since different performance measures provide superior results over different samples and different assets. This further motivates the need for a combination of performance measures, similar to what happens in the forecast combination literature. We thus relax the overwhelming restrictive assumption that a single performance measure provides superior results over all time periods. Furthermore, still referring to Farinelli et al. (2008, 2009), we note that these studies optimize the portfolio weights in order to choose a performance measure. Our approach is the reverse, since we fix the portfolio weights and optimise the weights of a linear convex combination of performance measures.

In this paper, we introduce a novel criterion for performance measure combination designed to be used as an equity screening algorithm. The combination criterion follows the general idea of linearly combining existing performance measures with positive weights. Moreover, these weights are determined by means of an optimisation problem. The underlying criterion function takes into account the risk-return trade-off potentially associated with the equity screens, evaluated on a historical and rolling basis. By construction, and due to the rolling evaluation approach, our method provides performance combination weights that can vary over time, thus allowing for changes in preferences across performance measures. The proposed approach is implicitly robust to the dynamic features of the returns densities, as these will affect the evaluation of performance measures that are the inputs of our screening algorithm. The final product of the linear combination of performance measure will be a composite performance index, which can then be used to create asset screens.

Apart from introducing our composite index, we discuss several implementation issues that further clarify the methodology. These include the selection of performance measures, their evaluation and the optimisation of the objective function with respect to the combination weights.

Finally, we present an empirical application that illustrates the use of our screening algorithm in a simplified portfolio allocation. We show how combined performance indices might be used for equity asset screening.

The remainder of the paper proceeds as follows. In Section 2 we define the investment objective and introduce our screening algorithm. In Section 3, we discuss several implementation aspects. Section 4 contains an empirical example, and Section 5 concludes the paper.

2 The investment problem and the objective function

Our purpose is to focus on the equity selection problem faced by an investor (a portfolio manager). The investor is willing to allocate his portfolio over a set of assets, and wants to select the assets using a combination of several performance measures. We presume the investor follows a myopic allocation rule: he chooses at time t the assets to form a portfolio with an investment horizon of one period, ending at $t+1$. To be consistent with this myopic objective, the equity screening is based on a criterion function depending on the expectations for time $t+1$ of a set of performance measures. Given the time t information set, the investor first determines the performance measures expectations for the next period, and then computes the composite performance index. In turn, this index is used as an equity screening tool, and helps the investor identify the most interesting assets -those with higher composite performance index value. Finally, the selected assets are introduced in the portfolio, with weights to be determined.

We further assume the investor includes in the portfolio M assets chosen from a larger group containing N assets. Note that $M << N$ in order to avoid excessive transaction and rebalancing costs, while M should not be too small, otherwise diversification benefits will tend to vanish. In this study, we fix $M = 25$, 50, or 100. Those values are reasonable in small and medium-sized managed portfolios, and will allow us to verify if changes in the number of assets will provide relevant variations in the portfolio turnover and, as a consequence, on rebalancing costs. Identifying the optimal number of assets might also be accomplished following given criterion. However, such a strategy is not considered in this work.

For simplicity, and to focus on the advantages of using a combination of performance measures, we impose an equally weighted portfolio composition. This implies that all assets included in the portfolio will have weight $1/M$. At first glance, this might seem a restrictive assumption. However, equally weighted portfolios have been shown to have performances comparable to, if not better than, optimised portfolios; see De Miguel et al. (2009). The motivation of this finding is that equally weighted portfolios avoid to introduce in the empirical comparison the estimation errors associated with the weights estimation (this will happen, for instance, if weights are determined within a mean-variance approach).

Given the choice of M and an investment universe of N assets, the objective of the investor is to select the M assets to be included in her portfolio following an optimality criterion based on a combination of performance measures. The main contribution we provide in this paper is the peculiar screening rule we propose, which is based on an optimized linear convex combination of performance measures.

In general terms, a simple screening rule orders assets using a single performance measure. For instance, we could order the N assets by computing the Sharpe ratios of all assets, order them, and invest using an equally weighted strategy in the M assets with the highest Sharpe ratios. However, when asset return densities deviate from normality, higher order moments, partial moments or quantiles may have additional informative content. As a result, more general and flexible performance measures could provide different asset rankings.

Therefore, our aim is to propose a more efficient screening rule the purpose of which is to combine a set of performance measures. The combination will take advantage of different views on the assets, or, similarly, of different information, including the asset returns density, the relation between asset returns and risk factors and the use of alternative utility functions.

Let us first introduce some notations. We define for each asset j at time t a composite performance index, $CI_{j,t}$, which is a function of Q performance indices $p_{i,j,t}$, where $i = 1, 2, ...Q$, and $j = 1, 2, ...N$. Note that this index is computed using the information set up to time t, I^t, and we presume that $E\left[CI_{j,t+1}|I^t\right] = CI_{j,t}$, and $E\left[p_{i,j,t+1}|I^t\right] = p_{i,j,t}$. We impose the simplifying assumption that the set of performance measures is fixed and known a-priori (thus, the value of Q is fixed over time, and the Q performance measures used in the combination do not change over time; for instance, if the Sharpe ratio is included from the beginning, it will not be substituted after some time with the Sortino ratio).

A first composite index may be easily produced by summing up, or averaging, the ranks separately obtained by the Q performance measures over a set of assets, as

$$CI_{j,t} \quad = \quad \sum_{i=1}^{Q} \text{Rank}\left(p_{i,j,t}\right), \tag{1}$$

where Rank(.) denotes the rank of asset j within the asset set with respect to measure i (therefore, the asset rank is a number between 1 and N). The best assets are those with smaller values of the composite index, since lower ranks imply higher values of the performance measures. In this case, the performance measures will all have the same weight, independent from the facts that i) some subsets may provide the same ordering of assets or ii) they contain the same information. This is not infrequent and the studies by Eling and Shuchmacher (2008), Eling (2009), Eling et al. (2009) and Caporin and Lisi (2011) show evidence of extremely high rank correlations across a number of widely used performance measures. The second issue could be controlled by the appropriate selection of performance measures, but the first limitation could only be removed by optimally combining performance measures. We thus define the following more general composite index for asset j at time t:

$$CI_{j,t}\left(w_1, w_2, ...w_Q\right) = \sum_{i=1}^{Q} w_i p_{i,j,t}, \quad j = 1, 2, ..., N, \tag{2}$$

$$w_i \geq 0, \quad i = 1, 2, ..., Q, \qquad \sum_{i=1}^{Q} w_i = 1,$$

where the weights are imposed to be positive and to sum to one. Note the weights are the same for all assets and are time-invariant. Assume for a while the weights are known; then, given the composite indices for each asset, we determine the portfolio composition by investing $1/M$ of the actual wealth in the M assets with the highest score of the composite index $CI_{j,t}$. However, weights $\mathbf{w} = \{w_1, w_2, ...w_Q\}$ have to be estimated, or optimally determined. We propose to determine the weights by maximizing the following criterion function:

$$\max_{\mathbf{w}} f\left(\mathbf{w}\right) = \frac{1}{m} \sum_{l=t-m+1}^{t} r_{p,l} - \lambda \frac{1}{m} \sum_{l=t-m+1}^{t} \left(r_{p,l} - \mu_p\right)^2, \tag{3}$$

$$r_{p,l} = \frac{1}{M} \sum_{j \in \mathcal{A}_t(\mathbf{w})} r_{j,l}, \tag{4}$$

where $\mathcal{A}_t\left(\mathbf{w}\right)$ is the set of the M assets with the highest score of the $CI_{j,t}\left(\mathbf{w}\right)$ index (note this set depends on the choice of the weights' vector); $r_{p,l}$ is the time l return of the equally weighted portfolio over the assets included in $\mathcal{A}_t(\mathbf{w})$; the first term of the criterion function is the average return of the portfolio over the last m observations; the second term of the criterion function is similar to a risk measure, weighted by a risk aversion coefficient λ; the risk measure depends on the choice of μ_p, which could be set either to the average portfolio return, thus making the second term equivalent to the portfolio variance, or it could be set equal to the average return of a benchmark over the last m observations, making the second term equivalent to a variance tracking error; the risk aversion coefficient could be set to values between 2 and 50, mimicking the standard choices in the mean-variance framework.

The overall criterion function is similar to a mean-variance utility function. However, (3) is not optimized with respect to the portfolio weights, which are fixed, but with respect to the performance measure weights. The intuition behind this criterion function is that we are determining the weights which, using up-to-date information, would have maximized the difference between the return and the risk of the allocated portfolio (where the risk is weighed by a risk aversion coefficient). The risk could be monitored in absolute terms, by using the portfolio variance, or in relative terms, by comparing the portfolio returns to those of a benchmark index, which is a reasonable choice if the investor is an investment manager. The criterion function is a backward evaluated mean-variance function which is used to forward allocate the portfolio.

Note that in the function $f(\mathbf{w})$, the weights are fixed, but when we apply it over time, the weights become time-varying since they update with respect to the changing relevance of the underlying performance measures. This fact could be related to the evidence provided in Caporin and Lisi (2011) that shows time-variation in the rank correlation across performance measures, suggesting their informative content is not stable over time. This could also be associated with a change over time of the asset return densities, or equivalently, of their moments and quantiles. In fact, if these elements vary over time, performance measures vary over time and their relevance might change over time.

We note our performance combination is generated in the cross-section of assets. As a result, the combination weights might change over time and, within the same time window, they might differ across different choices of the N assets. In fact, our proposal takes into account the dependence across performance measures within the assets' cross-section. On the contrary, our combined index does not use the potential dependence across performance measures at the single asset level. The extension of our approach in that direction is left for future research.

The proposed approach for the evaluation of the composite performance index entails a number of implicit assumptions. From a statistical point of view, the construction of performance measures at the single asset level implies our focus is on the marginal distributions of each asset included in the analysis. As a consequence, the composite index in (2) is an equity screening tool since it does not provide optimal asset allocation. However, a relevant aspect is not taken directly into account, i.e. the correlation across assets. In fact, the asset dependence only has an implicit role in the portfolio return and risk but it is not an element directly appearing in the criterion function. However, given that (3) penalises excessive risks and that the portfolio is equally weighted, the effect of correlations is partially sterilised. It might be possible that assets highly correlated with relatively good performances are included in the equally weighted portfolio. For this reason, we suggest the use of (2) within an asset allocation framework, but not directly as an asset allocation tool. A second element not directly covered by our composite index is the dependence between performance measures. We suggest including performance measures with low rank correlation in (2), following Eling and Shumacher (2008), Eling (2009), Eling et al. (2009) and Caporin and Lisi (2011). We leave to future research the generalisation of the proposed approach in order to introduce a composite index depending on the correlation across performance measures. Continuing to discuss the statistical assumptions, by employing a rolling estimation of performance measures, we can capture potential time-varying features of the asset return distributions. Performance measures may vary over time as a consequence of the change in the sample used for their evaluation, but also due to changes in the return distribution. Basically, we assume that returns observed over time might be the realisation of different underlying densities. However, such a choice does not influence our composite index since it is, actually, model-free. In fact, most of the available performance measures are functions of transformations of sample data (such as utility-based performance measures), empirical moments, empirical partial moments or empirical quantiles of asset returns. If the population values of those quantities would be time-varying, our methodology will take that into account, while, if they would be time-invariant, we will only have an effect coming from sampling errors, which could be controlled by changing the size of the rolling window. To be even more clear, we do not capture with a specific model the potential time variation in the asset return density, but rather use sample rolling estimators of the relevant elements in the performance measures. We thus work on a conditional or short-term distribution, and not the unconditional one. By construction, our composite index and criterion function identify the best assets on a backward basis and use them forward within an asset allocation program. This implies we are implicitly assuming the best forecast of the composite index at the asset level is equal to the last available observation of the composite index. Furthermore, we assume the sample estimators adopted are consistent and unbiased estimators of the corresponding population quantities. In the presence of time variation of the population moments, partial moments and quantiles, our approach is clearly suboptimal. An ideal choice would be to fit a parametric model, specifying a distributional assumption and the law of motion of the density parameters. However, such a choice would expose the analysis to specification problems and potential changes in the density family (even if that could be accommodated by means of mixture models). In this light, our approach is partially non-parametric given that we always consider empirical quantities the evaluation of which does not depend on a distributional assumption.

3 Implementation issues

In the following, we discuss a number of issues that should be considered in the implementation of the composite performance index evaluation. Those elements clarify the choice of performance measures, their computation and the approach we suggest for optimising the criterion function $f(\mathbf{w})$.

Standardisation. Given a list of Q performance measures, our final purpose is the construction of a composite index. However, we must recognise that different performance measures could have different ranges, thus making their combination dependent on the scale of the chosen performance measures. For this reason, we suggest considering the standardised performance measures as inputs of the composite index. Let $p_{i,j,t}$ be a given performance measure; we suggest computing the composite index $CI_{j,t}$ using the following quantities as inputs:

$$\bar{p}_{i,j,t} = \frac{p_{i,j,t} - \min\{p_{i,j,t}\}_{j=1}^{N}}{\max\{p_{i,j,t}\}_{j=1}^{N} - \min\{p_{i,j,t}\}_{j=1}^{N}}. \tag{5}$$

Such a standardisation makes the performance indices vary between 0 and 1, thus avoiding the scale effect, and ideally putting all performance measures on the same playing field.

Investment universe. When the allocation is performed over time for different t, our approach does not require the portfolio cardinality M and the number of assets N to be fixed. These two could be changed over time, thus allowing for changes in the universe of available assets (companies may die, or might be involved in mergers and acquisitions, or new companies can be included in the investment universe) as well as for changes in the portfolio strategies (increasing/decreasing the diversification). In fact, the performance combination weights are time t specific.

Benchmark. If a benchmark is used in the criterion function, its choice also has to be carefully considered. In fact, the benchmark should be chosen to be representative of the N assets included in the analysis to evaluate an appropriate tracking error. The benchmark and the assets should thus include the effect of dead companies and whether the use of a market index coupled with a subset of currently traded companies exposes the evaluation to survivorship bias. One alternative approach that overcomes the bias and excludes dead companies is to create a synthetic benchmark using a set of N selected assets and their market values. We follow this approach for simplicity, and leave to future research the empirical applications using dead companies.

Companies' market value. Liquidity is one of the possible market constraints that could affect our modelling strategy. In fact, the selected assets could differ in terms of market value and thus liquidity, making the allocation of the optimal portfolio problematic. In extreme cases, our optimally created composite index could suggest investing in companies with small market value, whose shares might be characterised by limited liquidity. As a result, the implementation of the portfolio could be characterised by large costs (transaction costs as well as large deviations in the price due to the limited liquidity or the impossibility of creating the portfolio because some trades could not be executed in the market due to the absence of a counterpart). In order to mitigate this aspect, and thus force the optimal portfolio to invest in small caps only if their performances are really relevant, we suggest introducing market value as a further performance measure. This would capture the liquidity effect, higher the market value and higher the liquidity. Clearly, other measures of stock liquidity could be considered.

Definition of performance measure. Our approach is flexible, and the term "performance measures" could be interpreted in a wider sense. In fact, we could optimally combine a set of indicators we associate with listed companies. These indicators could be performance measures, but could also be liquidity measures, technical analysis indicators or company-specific variables (revenues, employees, balance sheet ratios). From a different viewpoint, the composite index we propose might be separately evaluated for a set of risk measures, as well as for a set of reward measures. We leave such alternative constructions to future research. From this different point of view, our criterion is similar to a multi-criteria methodology, similar in some respects, to Ballestro et al. (2007).

Evaluation of performance measures. The performance indicators chosen to build the composite index are generally computed on a given sample. In order to follow the evolution over time of the asset return densities, we suggest evaluating the performance measures over a rolling window of m observations. The value of m depends on the time frequency of observations and on the total sample length; some examples could be 60 or more months, 25 or more weeks or 40 or more days. In general terms, we suggest using between 40 and 60 observations to avoid excessive volatility in performance measure values that might induce relevant changes in the construction of the composite index, in the assets included in $\mathcal{A}_t(\mathbf{w})$ and consequently, a large turnover in the portfolio. On the contrary, longer samples could significantly smooth performance measures sequences, leading to a very low turnover, but would not capture local (medium period) changes in performance measure relative rankings.

Estimation of performance weights. The determination of the composite index requires the solution of a non-trivial optimisation problem. For each point in time, the evaluation of $f(\mathbf{w})$ conditional to a vector of weights $\mathbf{w}$ requires the following steps:

- Evaluate the performance measures $p_{i,j,t}$;
- Compute the standardized performance measures $\bar{p}_{i,j,t}$;
- Determine for each asset the composite index $CI_{j,t}(\mathbf{w}) = \sum_{i=1}^{Q} w_i \bar{p}_{i,j,t}$;
- Identify the set $\mathcal{A}_t(\mathbf{w})$;
- Obtain the ex-post return of the allocation $r_{p,l} = \frac{1}{M} \sum_{j \in \mathcal{A}_t(\mathbf{w})} r_{j,l}$ and the objective function $f(\mathbf{w})$.

The criterion function $f(\mathbf{w})$ is, however, a non-linear and non-differentiable function of the performance measure weights $\mathbf{w}$. In fact, these enter only in the construction of the set $\mathcal{A}_t(\mathbf{w})$ that contains the assets with the highest values of the index $CI_{j,t}(\mathbf{w})$. Furthermore, different values of the weights could provide the same set of 'best' assets, thus making the optimisation of $f(\mathbf{w})$ computationally demanding. We provide a graphical example to clarify this aspect. Let us assume we have three performance measures and thus two weights to be estimated (the third one is obtained through the constraint). We report in Figure (1) the value of the criterion function for all possible weight combinations. Notably, the surface has many flat areas and local maxima. On the basis of the previous comments, we conclude that optimisation methods based on derivatives of the function $f(\mathbf{w})$ are not appropriate. We suggest the use of genetic algorithms, in particular the Differential Evolution method. For a description of the algorithm, see Storn and Price (1997), Maringer (2005) and Price et al. (2005). For applications of the Differential Evolution in finance, see Maringer (2005), Gilli et al. (2008), Hagstromer and Binner (2009), Krink et al. (2009), Krink and Paterlini (2011) and Gilli and Schumann (2012), among others.

4 Equity screening with composite indices on the US market

We consider an empirical application of the composite index previously introduced within an asset allocation framework. The composite index might be seen here as an equity screening rule, and our purpose is to verify its advantages in terms of portfolio returns. We first list the performance measures we take into account, and later describe the data we consider. Moreover, we describe two alternative naive equity screening rules that are compared to our proposal. The empirical results are reported in a fourth subsection.

4.1 Selected performance measures

The results of our approach clearly depend on the choice of performance measures combined in the index $CI_{j,t}(\mathbf{w})$. We do not provide a criterion for their choice, but refer to existing studies reporting comparisons among performance measures, e.g. Eling and Shumacher (2008), Eling (2009), Eling et al. (2009) and Caporin and Lisi (2011). The following surveys might be used to select among the large set of performance measures proposed in the financial economics literature: Aftalion and Poncet (2003), Le Sourd (2007), Bacon (2008), Cogneau and Hubner (2009a,b) and Caporin et al. (2011). In the following we list the performance measures for asset i. Furthermore, we evaluate the performance measures using returns data for the range $t-m$ to $t-1$, to be used for time t equity screening. Our selection includes traditional performance measures:

- Sharpe ratio (Shapre, 1966 and 1994)

$$Sh(i,t-1,m) = \frac{\mu\left(r_{i,t-1} - rf_{t-1}, m\right)}{\sigma\left(r_{i,t-1} - rf_{t-1}, m\right)}, \tag{6}$$

  where rf_t is the risk-free rate, $\mu\left(x_{t-1}, m\right) = \frac{1}{m}\sum_{j=1}^{m} x_{t-j}$ and $\sigma^2\left(x_{t-1}, m\right) = \frac{1}{m}\sum_{j=1}^{m}\left(x_{t-j} - \mu\left(x_{i,t-\right.\right.$

- The expected return over the Mean Absolute Deviation (MAD) introduced by Konno (1990, 1991)

$$ERMAD(i,t-1,m) = \frac{\mu\left(r_{i,t-1} - rf_{t-1}, m\right)}{MAD\left(r_{i,t-1} - rf_{t-1}, m\right)}, \tag{7}$$

  where $MAD\left(x_{t-1}, m\right) = \frac{1}{m}\sum_{j=1}^{m}\left|r_{t-j} - \mu\left(x_{t-1}, m\right)\right|$;

- The Appraisal ratio, defined as

$$AR(i,t-1,m) = \frac{\alpha_i}{\sigma[\epsilon_{i,t}]}, \tag{8}$$

  where α_i is the intercept of the CAPM regression $r_{i,t} - rf_t = \alpha_i + \beta_i\left(r_t^M - rf_t\right) + \varepsilon_{i,t}$, with r_t^M being the market return; $\sigma[\epsilon_{i,t}]$ is the volatility of the CAPM regression residuals (the volatility of the idiosyncratic shocks) and the regression parameters are estimated over the range $t-1$ to $t-m$.

- The Treynor index (Treynor, 1965), or Risk Adjusted Return,

$$RaR(i,t-1,m)=\frac{\mu\left(r_{i,t-1}-rf_{t-1},m\right)}{\beta_i},\tag{9}$$

  where the risk adjustment is made using the systemic risk exposition as computed from the CAPM regression reported in the Appraisal ratio description;

- The M2 index by Modigliani and Modigliani (1997),

$$M2(r_{i,t})=\mu\left(r_{i,t-1}-r_{B,t-1},m\right)\times\frac{\sigma\left(r_{B,t}\right)}{\sigma\left(r_{i,t}\right)}+\sigma\left(rf_t\right)-\sigma\left(r_{B,t}\right),\tag{10}$$

  where $r_{B,t}$ identifies the return of a benchmark investment.

We also include measures based on the Drawdown, defined as the maximum loss an investor may suffer in the period $t-m$ to $t-1$:

$$D_t(i)=\min(D_{t-j}+r_{i,t},0)\text{ with }D_{t-m-1}=0.$$

The Drawdowns obtained in the range $t-m$ to $t-1$ can be ordered from the smallest (generally negative) to the largest (generally a zero), resulting in the ordered sequence $\bar{D}_1(r_{i,t-1}),\bar{D}_2(r_{i,t-1}),\ldots\bar{D}_m(r_{i,t-1})$. In our analysis, we use the following indicators:

- the Calmar ratio of Young (1991),

$$CR(i,t-1,m)=\frac{\mu\left(r_{i,t-1},m\right)}{-\bar{D}_1(i)};\tag{11}$$

- the Sterling ratio of Kestner (1996),

$$SR(i,t-1,m,w)=\frac{\mu\left(r_{i,t-1},m\right)}{-\frac{1}{w}\sum_{j=1}^{w}\bar{D}_j(X_{i,t})},\tag{12}$$

  where w identifies the number of values used for the drawdown measure;

- the Burke (1994) ratio,

$$BR(i,t-1,m,w)=\frac{\mu\left(r_{i,t-1},m\right)}{\left(-\frac{1}{w}\sum_{j=1}^{w}[\bar{D}_j(X_{i,t})]^2\right)^{\frac{1}{2}}}.\tag{13}$$

We also consider other measures based on Partial Moments:

- the Sortino ratio by Sortino and Van der Meer (1991),

$$Sr(i,t-1,m) = \frac{\mu\left(r_{i,t-1},m\right)}{LPM\left(r_{i,t-1},m,2\right)}, \tag{14}$$

where $LPM\left(x_{t-1},m,p\right) = \left(\frac{1}{m}\sum_{j=1}^{m}(-\min(x_{i,t-j},0))^{p}\right)^{\frac{1}{p}}$;

- the Kappa 3 measures by Kaplan and Knowles (2004),

$$K3(i,t-1,m) = \frac{\mu\left(r_{i,t-1},m\right)}{LPM\left(r_{i,t-1},m,3\right)}. \tag{15}$$

Finally, we consider a measure based on quantiles, the expected return over absolute Value-at-Risk of Dowd (2000)

$$VR(i,t-1,m,\alpha) = \frac{\mu\left(r_{i,t-1},m\right)}{|VaR(r_{i,t-1};\alpha)|}, \tag{16}$$

where $VaR(r_{i,t-1};\alpha)$ is the α-quantile of asset i returns in the period $t-1$ to $t-m$.

In the following empirical application, we consider a rolling evaluation of the performance measures over a window of 60 days. Moreover, we compute the Sterling and Burke ratios employing the five largest Drawdowns. Finally, in the VR index, we set $\alpha = 0.05$.

As mentioned in the previous section, the MV of each company will be included as an additional performance measure to penalise smaller companies.

4.2 Dataset description and benchmark construction

Our dataset is based on the constituents of the S&P Composite 1500 (the 22nd of February, 2012). The time series were downloaded from Datastream at a monthly frequency from the 31st of January 1990, to the 31st of January 2012, for a total of 265 observations. We also recovered a proxy of the risk free asset, the JP Morgan 1 Month Cash bond index. To cope with survivorship bias, we restricted the dataset to a collection of assets constantly available in the analysed sample. Following this criterion, we restricted our attention to 695 assets.[1]

Given that we exclude a relevant part of the assets included in the S&P500, and given that the index composition changes over time,[2] the S&P500 index cannot be used as a benchmark or market index to evaluate the performances of our equity screening approach. Therefore, we build a benchmark that is coherent with the selected assets. The index we construct corresponds to the value-weighted index composed of the 695 selected assets.

[1] The list of assets is available upon request.

[2] The time series of the S&P constituents is not available to us through Datastream.

Table (1) provides some information about the MV of the assets, while Figure (2) is the plot of the benchmark total returns since the end of January 1990. We note the well-known decreases in the equity benchmark in 2000-2001 and 2008. Moreover, we point out the increase in the selected equity average market value from 1990 to the present. The upward movement is matched with a somewhat stable coefficient of variation, with the exception of the value reported in 1999, just before the technology market bubble burst.

4.3 Portfolio allocation and naive equity screening rules

We apply our equity screening approach to the selected assets, estimating performance measures on rolling windows of 60 months. Starting from the end of January 1995 (the first month where 60 monthly returns are available), we identify, across the 695 assets, the 50 assets that maximise the criterion function in (3). We then create an equally weighted portfolio where each asset has weight equal to 2%, and we compute the monthly realised returns of the portfolio. The portfolio composition is modified on a monthly basis, where the criterion function (3) is optimised each month. At the end of this procedure, we have a total of 205 portfolio returns.

We apply the screening rule discussed in Section 2, combining two different specifications of the risk component in (3); we consider the portfolio variance (VO) and the tracking error volatility (TE) cases.[3] Moreover, we make use of two different values for the risk aversion parameter, 1 and 20. The first corresponds to a mild penalisation of the risk, while the second mimics the choices of a more risk-averse agent.

To evaluate the performances of our equity screening algorithm, we compare it to a naive equity screening rule based on the Sharpe ratio. Therefore, with a rolling procedure similar to that outlined above, we selected the 50 assets that have higher Sharpe ratios, and use those assets to create a second equally weighted portfolio. Performances will also be compared to those of the benchmark, computed as described in the previous subsection.

The portfolio returns are compared by means of the following approaches: standard descriptive analyses of returns, including the computation of some risk measures; a horse-race over the range February 1995 to January 2012; the weights associated with the different performance measures; the turnover of the portfolios based on the screening algorithms.

[3] We set the term μ_p equal to the average portfolio return in the first case, while μ_p is equal to the benchmark return in the second case.

4.4 Performance results

Table (2) includes the descriptive analysis of the portfolios, while Figure (3) shows the cumulated returns from 1995 to 2012. Compared to the benchmark, all equity screening-based portfolios provide higher cumulated returns. However, the risk-aversion coefficient has a relevant impact; in fact, the TE and VO portfolios with risk aversion set to 20 are less profitable than the Sharpe-based portfolios. The result is even stronger if we analyse the turnover, which is sensibly higher for higher values of the risk aversion. Comparing the Sharpe index of the portfolio returns, the Sharpe-based portfolio seems to be the preferred choice, excluding of the VO case with risk aversion set equal to 20. Such a result is a consequence of the criterion used for portfolio construction, and might be expected. In terms of risk measures, we observe that all portfolios based on screening rules are more risky than the benchmark. However, we stress that all screening-based portfolios have not been optimised to reduce the risk, but are simply based on an equally weighted allocation scheme. As a result, risk reductions might be achieved by optimising portfolio weights. Finally, we emphasise that risk measures decrease (in absolute terms) with increasing risk aversion, as expected, and for a risk aversion coefficient equal to 20 they are better than the benchmark and also preferred to those of the Sharpe-based portfolio.

Overall, the introduction of screening rules provides higher returns than the benchmark, with a preference for our proposed algorithm compared to simpler screening based only on the Sharpe ratio. Risk measures are different across screening strategies, but this is a consequence of the portfolio construction that is not optimised. The turnover induced by screening rules is influenced by the degree of risk aversion, and becomes higher and more volatile with increasing risk aversion; see Figure (4).

In our selection procedure, the weights assigned to the different performance measures have a relevant role; moreover, they change over time, and react to the different features of the returns time series. Figure (5) shows an example, while Table (3) includes the descriptive statistics for the tracking error- and volatility-based screening. We first point out that the screening algorithm we propose generally assigns a very small weight to the Sharpe ratio, independent of form of the criterion function and of the risk aversion coefficient level. Both tracking error and portfolio volatility objective functions provide similar performance measure weights when the risk aversion coefficient is equal to 1; in particular, we observe that the Modigliani-Modigliani index receives the largest weight. Other performance measures receiving a high weight are the Appraisal Ratio and the Excess Return over Mean Absolute Deviation. The other measures are characterised by very small average weights and limited standard deviations, signalling they receive a relevant weight only occasionally. When the risk aversion coefficient is increased to 20, the difference between the two forms of the criterion function leads to different average weights being assigned to the performance measures. When we focus on the tracking error-based function, the MV, the Appraisal Ratio and the Excess Return over mean absolute deviation receive a weight larger than 10%. In contrast, in the second implementation based on portfolio volatility, the Burke and RAR measures increase over 10%, while the market value falls below 10%. Even for a large risk aversion level, the performance measures with small average weight have a large standard deviation, thus confirming their limited relevance. The results support our expectation of variability in the informative content of performance measures, and might also be seen as a confirmation of the potential interest in measures going beyond the Sharpe ratio.

A relevant element to emphasise is the limited weight assigned to the market value. As a consequence, the selected assets might be characterised by small market value and thus small liquidity, possibly creating difficulties in the implementation of portfolios based on those assets. This is confirmed by Figure (6) in which we see a sharp decrease in the market value of the selected companies in the second half of the sample. Such a behavior is common across the different implementations of the screening algorithm. To force the impact of market value in the screening algorithm, we run a second set of evaluations where we constrain the weight assigned to the market value, imposing a lower bound set to 10%.

Table (2) includes the portfolio return descriptive results, showing the impact of the market value bound in terms of cumulated returns, risk measures, and Sharpe ratios. Overall, imposing minimum relevance to the companies' market value leads to a risk reduction, as Value-at-Risk, Expected Shortfall, returns volatility and range all improve. However, the total and average returns and the Sharpe ratio decrease, except in the case with the tracking error objective function and a high level of risk aversion. In addition, the turnover shows a decrease, which is larger for the cases where the risk aversion is set to 20. Comparing the market value of selected companies, we note an increase in the second part of the sample compared to the previous cases; see Figure (6). As a consequence, the introduction of a lower bound to the market value leads to the selection of equities the average market value of which is generally higher than the average market value of the benchmark. Weights assigned to the performance measures are partially affected by the constraint imposed on MV; see Table (4). However, the sets of the most influential performance measures are unchanged.

As mentioned in the previous section, the number of assets identified by our screening algorithm might be easily modified. Tables (5) and (6) contain descriptive analyses of the realised portfolio returns for 25 and 100 assets, as does Table (2) for the 50 assets case. By comparing the results across different values of M, we note that screening algorithms always beat the benchmark in terms of cumulated returns but not with respect to risk measures. Nevertheless, we observe a general reduction in the risk measures for increasing M, and an improvement in the Sharpe ratios for $M = 100$. Such a finding depends on the possibility of identifying profitable investment opportunities (that is, single assets) the performance of which might be variable over time, leading to assets being "above average quality" but not necessarily "top performers". If a small number of assets is used, the selected equities are subject to more frequent changes, as shown by the average turnover (decreasing for increasing M). As a result, when the number of selected assets increases, the performances improve. Moreover, we observe that the Sharpe ratio of the portfolios based on our screening algorithm are better than the naive approach in a few cases only, but are associated with different objective functions: when $M = 25$ the use of the tracking error-based objective function, with a large risk aversion and bounded weight of the market value, provide the best results. In contrast, when $M = 100$, the results are slightly better for the objective function using the variance of the selected portfolio and a large risk aversion level, independent from the presence of a bound on the weight of MV. The case where $M = 50$ is in the middle, with the Sharpe ratios of the naive strategy and our screening rule being very close to each other. Overall, this empirical application shows that the proposed screening algorithm is able to identify profitable investment opportunities.

5 Conclusions

We introduced a new screening algorithm that selects within an investment universe a subset of assets based on a composite index of performance measures. Such an index linearly combines different performance measures where the combination weights are derived from an optimisation problem that takes into account past performances associated with the "optimal" weights and the subsequent asset ranking. Accordingly, past performances lead to the asset selection for future allocations, suggesting the backward/forward equity screening name. We discuss several implementation issues of our screening algorithm and then present an empirical application based on US equities. The results show the advantage of our composite performance index in a simplified asset allocation framework. In fact, comparing simulated equally weighed portfolio strategies, the composite performance index here proposed provides superior results in terms of realised profits. Several aspects of our analysis might be further analysed. For instance, our study might be improved by the introduction of "dead" companies. In addition, in this paper, we do not consider the construction of contrarian (investing in assets with the lowest composite index) or long-short (long on the best assets, short on the worst ones) strategies. In fact, those approaches would be much more susceptible to survivorship bias, since they would require investing in companies with poor performance. Finally, portfolio weights might be estimated, contrary to calibrating them to the equally weighted case. Such extensions will be left to future research.

References

[1] Aftalion, F., and Poncet, P. (2003) Les techniques de mesure de performance. Paris: Economica.

[2] Bacon, C.R. (2008) Practical Portfolio Performance Measurement and Attribution. Wiley, Chichester.

[3] Ballestro, E., M. Gunther, D. Pla-Santamaria, and C. Stummer (2007) Portfolio selection under strict uncertainty: A multi-criteria methodology and its application to the Frankfurt and Vienna Stock Exchanges, European Journal of Operational Research, 181: 1476-1487.

[4] Burke, G. (1994) A Sharper Sharpe Ratio. Futures 23: 56.

[5] Caporin, M., G. Jannin, F. Lisi, and B. Maillet (2011) A survey of the four families of performance measures, manuscript.

[6] Caporin, M., and Lisi, F. (2011) Comparing and Selecting Performance Measures for Ranking Assets, Economics: The Open-Acces, Open-Assessment E-journal 5: 31 pages.

[7] Cogneau, P., and Hübner, G. (2009a) The (more than) 100 ways to measure portfolio performance - part 1: standardized risk-adjusted measures, Journal of Portfolio Measurement, 13: 56-71.

[8] Cogneau, P., and Hübner, G. (2009b) The (more than) 100 Ways to measure portfolio performance - part 2: special measures and comparison, Journal of Portfolio Measurement, 14: 56-69.

[9] De Miguel, V., Garlappi, L., and Uppal, R. (2009) Optimal versus naïve diversification: how inefficient is the 1/N portfolio strategy?, Review of Financial Studies, 22: 1915-1953.

[10] Dowd, K. (2000) Adjusting for Risk: An Improved Sharpe Ratio, International Review of Economics and Finance 9: 209-222.

[11] Eling, M. (2008) Does the measure matters in the mutual fund industry?, Financial Analyst Journal, 64: 54-66.

[12] Eling, M. and Schuhmacher, F. (2007) Does the choice of performance measure influence the evaluation of hedge funds? Journal of Banking and Finance, 31: 2632-2647.

[13] Eling, M., Farinelli, S., Rossello, D., and Tibiletti, L., (2011) One-size or tailor-made performance ratios for ranking hedge funds, Journal of Derivatives and Hedge Funds, 16: 267-277.

[14] Farinelli, S., Ferreira, M., Rossello, D., Thoeny, M. and Tibiletti, L. (2008) Beyond Sharpe ratio: optimal asset allocation using different performance ratios, Journal of Banking and Finance, 32: 2057-2063.

[15] Farinelli, S., Ferreira, M., Rossello, D., Thoeny, M. and Tibiletti, L. (2009) Optimal asset allocation aid system: from "one-size" vs "taylor-made" performance ratio, European Journal of Operational Research, 192: 209-215.

[16] Gilli, M., and Schumann, E. (2012) Heuristic optimisation in financial modelling, Annals of operations research, 193: 129-158.

[17] Gilli, M., Maringer, D., and Winker, P. (2008) Applications of heuristics in finance, in Schlottmann, D., Weinhardt, C., and Schlottmann F., eds., Handbook on information technology in finance, Ch. 26, Springer-Verlag, Berlin, Heidelberg, 2008.

[18] Grinold, R., and Kahn, R. (1999) Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk. McGraw-Hill New York.

[19] Hagstromer, B., and Binner, J.M. (2009) Stock portfolio selection with full-scale optimization and differential evolution, Applied Financial Economics, 19: 1559-1571.

[20] Hwang, S. and Salmon, M. (2003) An analysis of performance measures using copulae, in: Knigth, J., and Satchell, S. (eds), Performance measurement in finance: firms, funds and managers, Butterworth-Heinemann Finance, Quantitative Finance Series.

[21] Kaplan, P.D. and Knowles, J.A. (2004) Kappa: A Generalized Downside Risk-adjusted Performance Measure, Journal of Performance Measurement 8: 42-54.

[22] Kestner, L.N. (1996) Getting a Handle on True Performance, Futures 25: 44-47.

[23] Konno, H. (1990) Piecewise Linear Risk Functions and Portfolio Optimization, Journal of the Operations Research Society of Japan 33: 139-156.

[24] Konno, H. and Yamazaki, H. (1991) Mean-Absolute Deviation Portfolio Optimization Model and its Application to Tokyo Stock Market, Management Science 37: 519-531.

[25] Krink, T., Mittnik, S., and Paterlini, S. (2009), Differential evolution and combinatorial search for constrained index tracking, Annals of Operations Research, 172: 153-176.

[26] Krink, T., and Paterlini, S. (2011), Multiobjective optimization using differential evolution for real-world portfolio optimization, Computational Management Science, 8: 157-179.

[27] Le Sourd, V. (2007) Performance Measurement for Traditional Investment, Financial Analysts Journal, 58: 36-52.

[28] Maringer, D. (2005) Portfolio management with heuristic optimization, Advances in Computational Management Science, Volume 8, Springer.

[29] Modigliani, F. and Modigliani, L. (1997) Risk-adjusted Performance: How to Measure it and Why, Journal of Portfolio Management 23: 45-54.

[30] Price, K., Storn, R., and Lampinen, J. (2004), Differential evolution - A practical to global optimization. Springer Berlin.

[31] Sharpe, W.F. (1966) Mutual Fund Performance, Journal of Business, 39: 119-138.

[32] Sharpe, W.F. (1994) The Sharpe Ratio, The Journal of Portfolio Management, 20: 49-59.

[33] Sortino, F.A. and van der Meer, R. (1991) Downside Risk. Capturing What's at Stake in Investments Situations, Journal of Portfolio Management 17: 27-31.

[34] Storn, R., and Price, K. (1997) Differential Evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces, Journal of Global Optimization, 11: 341-359.

[35] Treynor, J.L. (1965) How to Rate Management of Invested Funds, Harvard Business Review 43: 63-75.

[36] Young, T.W. (1991) Calmar Ratio: A Smoother Tool, Futures 20: 40.

Table 1: Market value of selected equities (in millions of USD). CV denotes the coefficient of variation.

	Mean	Stdv	Min	Max	CV
Jan-90	1976	5304.59	0.79	58827	2.685
Dec-94	2505	6911.52	2.40	87193	2.759
Dec-99	8448	34849.45	1.04	602432	4.125
Dec-04	7872	22502.83	50.52	323717	2.859
Jan-12	9229	26383.81	48.97	425608	2.859

Table 2: Descriptive analysis of realized portfolio returns and of the benchmark with 50 assets. TE denotes portfolios where the criterion function considers the tracking error volatility, while VO represents portfolio where the criterion function depends on the portfolio variance. Moreover, 1 and 20 identify the risk aversion coefficient value. The columns report the cumulated returns obtained in the range February 1995 to January 2012, the annualized average monthly return and annualized variance, the minimum and maximum monthly returns, the skewness and kurtosis indices computed on monthly returns, the 5% Value-at-Risk and Expected Shortfall, the Sharpe ratio, and the average monthly turnover.

	Cum. Ret.	Ret. Ann.	Vol. Ann.	Min	Max	Skew	Kurt	Var(5%)	ES(5%)	Sharpe	Avg-TO
Bench	3.2030	0.1011	0.1538	-0.1601	0.1044	-0.5778	3.8469	0.0650	0.0963	0.1814	-
SH	6.3647	0.1404	0.1667	-0.1876	0.1121	-0.8082	4.4216	0.0681	0.1042	0.2288	0.1624
TE1	9.5338	0.1783	0.2259	-0.2477	0.1991	-0.5968	4.3472	0.0935	0.1405	0.2111	0.1822
TE20	4.2629	0.1169	0.1601	-0.1853	0.1180	-0.6989	4.2987	0.0668	0.0984	0.2002	0.2457
VO1	7.6904	0.1611	0.2100	-0.2477	0.1463	-0.7312	4.5592	0.0872	0.1319	0.2067	0.1862
VO20	4.6833	0.1186	0.1405	-0.1646	0.1360	-0.8225	5.4076	0.0573	0.0888	0.2314	0.2208
Portfolios with a lower bound on Market Value weight in the criterion function (weight less or equal to 0.1)											
TE1	8.3774	0.1686	0.2194	-0.2383	0.1662	-0.6080	4.1221	0.0911	0.1359	0.2063	0.1723
TE20	4.8260	0.1222	0.1524	-0.1777	0.1127	-0.6593	4.2221	0.0627	0.0921	0.2194	0.1972
VO1	6.8385	0.1530	0.2054	-0.2383	0.1442	-0.7387	4.3963	0.0856	0.1297	0.2013	0.1773
VO20	4.1327	0.1113	0.1361	-0.1598	0.0943	-0.9143	5.1529	0.0558	0.0861	0.2248	0.2013

Table 3: Descriptive analysis of performance measure weights in the range February 1995 to January 2012. The minimum is not included since equal to zero for all measures. RA denotes the levels of the risk aversion. TE identifies tracking-error-based screening while VO refers to screening with portfolio variance in the objective function.

	Mean	Max	St.Dev.	Mean	Max	St.Dev.
TE		RA = 1			RA = 20	
MV	0.0041	0.0911	0.0126	0.2450	0.6544	0.1758
ERMAD	0.1198	0.9987	0.2100	0.3353	0.9340	0.3191
AR	0.2904	0.9386	0.1839	0.2623	0.9588	0.2601
BR	0.0091	0.2149	0.0222	0.0335	0.4865	0.0784
SR	0.0030	0.0727	0.0095	0.0614	0.4955	0.1231
CR	0.0024	0.1074	0.0098	0.0035	0.1238	0.0162
RaR	0.0040	0.2078	0.0193	0.0140	0.3064	0.0470
M2	0.5559	0.9342	0.2175	0.0173	0.4235	0.0647
Sr	0.0018	0.1219	0.0105	0.0024	0.2780	0.0200
K3	0.0053	0.2345	0.0258	0.0195	0.6066	0.0994
VR	0.0017	0.1226	0.0100	0.0048	0.2878	0.0299
Sh	0.0026	0.3199	0.0253	0.0009	0.1770	0.0123
VO		RA = 1			RA = 20	
MV	0.0033	0.0668	0.0092	0.0100	0.1988	0.0261
ERMAD	0.1869	0.9987	0.2904	0.1634	0.9666	0.2652
AR	0.3204	0.9817	0.2213	0.3917	0.9999	0.3631
BR	0.0140	0.3627	0.0364	0.1306	0.9108	0.1814
SR	0.0081	0.1227	0.0221	0.0563	0.4404	0.1091
CR	0.0032	0.1064	0.0121	0.0204	0.3385	0.0560
RaR	0.0088	0.2840	0.0357	0.1955	0.9066	0.3111
M2	0.4390	0.9814	0.2285	0.0114	0.1933	0.0373
Sr	0.0012	0.0389	0.0042	0.0188	0.7849	0.1005
K3	0.0073	0.2518	0.0291	0.0007	0.0882	0.0064
VR	0.0075	0.4965	0.0528	0.0010	0.0673	0.0074
Sh	0.0001	0.0064	0.0007	0.0001	0.0180	0.0013

Table 4: Descriptive analysis of performance measure weights in the range February 1995 to January 2012 with MV weight with a lower bound at 10%. The minimum is not included since equal to zero for all measures. RA denotes the levels of the risk aversion. TE identifies tracking-error-based screening while VO refers to screening with portfolio variance in the objective function.

	Mean	Max	St.Dev.	Mean	Max	St.Dev.
TE		RA = 1			RA = 20	
MV	0.1035	0.1820	0.0109	0.3236	0.6889	0.1572
ERMAD	0.1065	0.8988	0.1879	0.3025	0.8406	0.2880
AR	0.2636	0.8447	0.1652	0.2358	0.8629	0.2344
BR	0.0082	0.1585	0.0179	0.0301	0.4378	0.0706
SR	0.0025	0.0851	0.0086	0.0520	0.4459	0.1060
CR	0.0021	0.0966	0.0088	0.0029	0.1114	0.0143
RaR	0.0035	0.1870	0.0173	0.0126	0.2758	0.0423
M2	0.5027	0.8408	0.1950	0.0156	0.3812	0.0582
Sr	0.0013	0.1097	0.0081	0.0021	0.2502	0.0180
K3	0.0029	0.1734	0.0132	0.0176	0.5459	0.0895
VR	0.0010	0.0559	0.0046	0.0044	0.2590	0.0269
Sh	0.0023	0.2879	0.0228	0.0008	0.1593	0.0111
VO		RA = 1			RA = 20	
MV	0.1029	0.1601	0.0082	0.1086	0.2789	0.0229
ERMAD	0.1653	0.8988	0.2604	0.1484	0.8699	0.2396
AR	0.2890	0.8835	0.1989	0.3519	0.8999	0.3280
BR	0.0126	0.3265	0.0327	0.1187	0.8289	0.1641
SR	0.0069	0.1104	0.0190	0.0516	0.3964	0.0990
CR	0.0027	0.0958	0.0104	0.0162	0.3046	0.0468
RaR	0.0080	0.2556	0.0321	0.1765	0.8160	0.2813
M2	0.3979	0.8833	0.2074	0.0096	0.1682	0.0316
SR	0.0012	0.0350	0.0043	0.0169	0.7955	0.0925
K3	0.0066	0.2267	0.0262	0.0007	0.0794	0.0060
VR	0.0067	0.4468	0.0475	0.0008	0.0583	0.0055
Sh	0.0001	0.0058	0.0006	0.0001	0.0162	0.0011

Table 5: Descriptive analysis of realized portfolio returns and of the benchmark with 25 assets. TE denotes portfolios where the criterion function considers the tracking error volatility, while VO represents portfolio where the criterion function depends on the portfolio variance. Moreover, 1 and 20 identify the risk aversion coefficient value. The columns report the cumulated returns obtained in the range February 1995 to January 2012, the annualized average monthly return and annualized variance, the minimum and maximum monthly returns, the skewness and kurtosis indices computed on monthly returns, the 5% Value-at-Risk and Expected Shortfall, the Sharpe ratio, and the average monthly turnover.

	Cum. Ret.	Ret. Ann.	Vol. Ann.	Min	Max	Skew	Kurt	Var(5%)	ES(5%)	Sharpe	Avg-TO
Bench	3.2030	0.1011	0.1538	-0.1601	0.1044	-0.5778	3.8469	0.0650	0.0963	0.1814	-
SH	5.6209	0.1354	0.1765	-0.1958	0.1182	-0.9211	4.8543	0.0732	0.1154	0.2087	0.1845
TE1	8.4082	0.1729	0.2366	-0.1937	0.2369	-0.2312	3.7493	0.0990	0.1415	0.1959	0.2180
TE20	5.9388	0.1348	0.1586	-0.1642	0.1180	-0.5272	3.8236	0.0647	0.0939	0.2313	0.3400
VO1	9.0752	0.1756	0.2294	-0.1807	0.2369	-0.1526	3.6282	0.0954	0.1328	0.2049	0.2406
VO20	4.6774	0.1197	0.1469	-0.1842	0.1120	-1.0308	6.0695	0.0603	0.0955	0.2232	0.2843
Portfolios with a lower bound on Market Value weight in the criterion function (weight less or equal to 0.1)											
TE1	8.6419	0.1723	0.2281	-0.1831	0.2032	-0.2574	3.4093	0.0950	0.1332	0.2025	0.2159
TE20	6.4369	0.1382	0.1519	-0.1573	0.1093	-0.5565	3.8731	0.0613	0.0892	0.2474	0.2475
VO1	6.6512	0.1552	0.2237	-0.1698	0.2337	-0.0969	3.5733	0.0941	0.1267	0.1873	0.2357
VO20	4.0101	0.1110	0.1439	-0.1764	0.0999	-0.9735	5.6622	0.0595	0.0962	0.2120	0.2725

Table 6: Descriptive analysis of realized portfolio returns and of the benchmark with 100 assets. TE denotes portfolios where the criterion function considers the tracking error volatility, while VO represents portfolio where the criterion function depends on the portfolio variance. Moreover, 1 and 20 identify the risk aversion coefficient value. The columns report the cumulated returns obtained in the range February 1995 to January 2012, the annualized average monthly return and annualized variance, the minimum and maximum monthly returns, the skewness and kurtosis indices computed on monthly returns, the 5% Value-at-Risk and Expected Shortfall, the Sharpe ratio, and the average monthly turnover.

	Cum. Ret.	Ret. Ann.	Vol. Ann.	Min	Max	Skew	Kurt	Var(5%)	ES(5%)	Sharpe	Avg-TO
Bench	3.2030	0.1011	0.1538	-0.1601	0.1044	-0.5778	3.8469	0.0650	0.0963	0.1814	-
SH	7.7535	0.1508	0.1599	-0.1898	0.1242	-0.8493	4.9815	0.0642	0.1005	0.2550	0.1337
TE1	8.5915	0.1652	0.1991	-0.2307	0.1262	-0.7565	4.3627	0.0817	0.1257	0.2231	0.1423
TE20	6.1371	0.1360	0.1541	-0.1635	0.1206	-0.7079	4.4306	0.0625	0.0964	0.2400	0.1761
VO1	8.1007	0.1591	0.1878	-0.2291	0.1298	-0.8484	4.7995	0.0768	0.1187	0.2284	0.1430
VO20	5.7616	0.1292	0.1346	-0.1623	0.0960	-0.9998	5.7702	0.0537	0.0873	0.2619	0.1776
Portfolios with a lower bound on Market Value weight in the criterion function (weight less or equal to 0.1)											
TE1	7.8952	0.1592	0.1950	-0.2334	0.1259	-0.8254	4.5355	0.0802	0.1236	0.2201	0.1362
TE20	6.1409	0.1356	0.1520	-0.1621	0.1146	-0.7105	4.3957	0.0615	0.0935	0.2428	0.1551
VO1	6.8876	0.1490	0.1854	-0.2334	0.1270	-0.9107	5.0157	0.0764	0.1185	0.2175	0.1408
VO20	5.7109	0.1284	0.1323	-0.1592	0.0944	-0.9877	5.5716	0.0527	0.0864	0.2649	0.1680

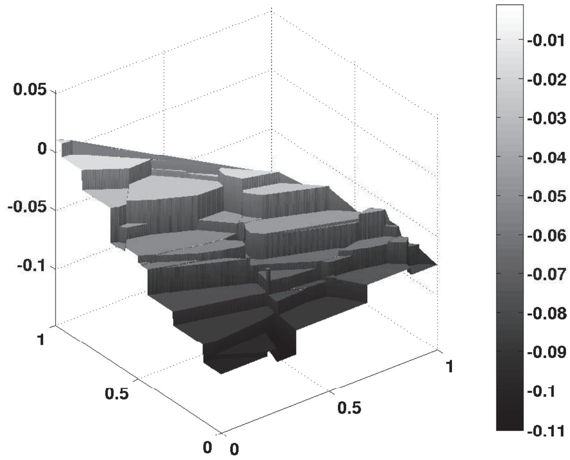


Figure 1: Surface of the objective function for the determination of the composite index weights with three performance measures. The vertical axis refers to the objective function while the other two axis report the weights of two performance measures (the third being obtained through the constraint on combination weights).

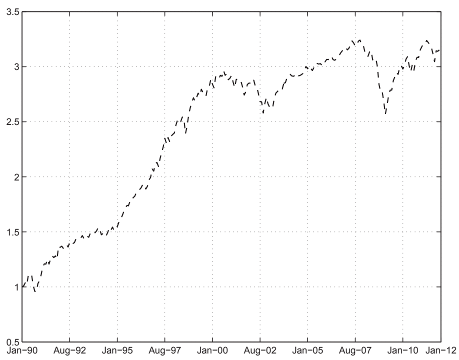
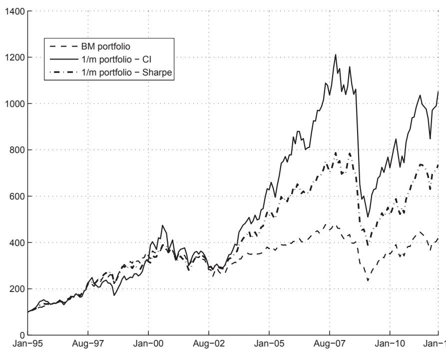
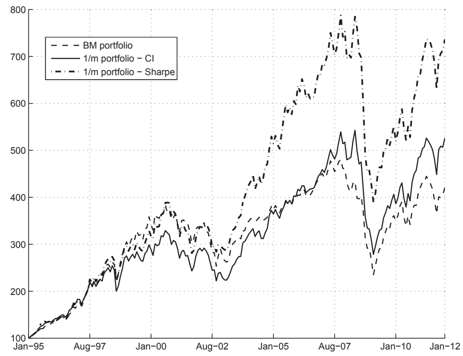


Figure 2: Cumulated returns of the benchmark in the period 1990-2012

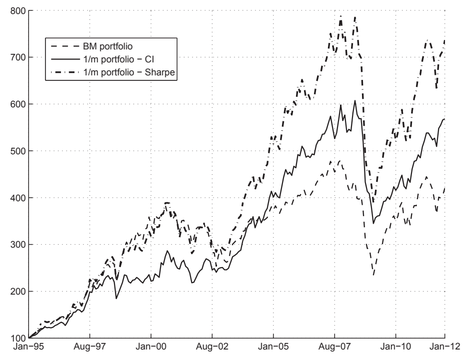


(a) TE1



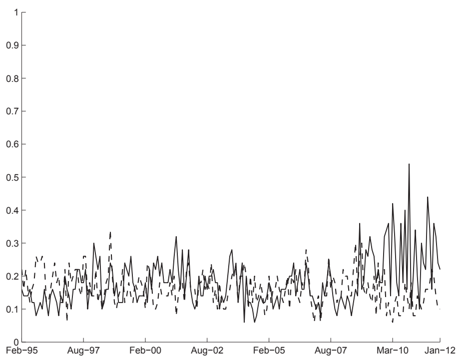
(b) TE20

(c) VO1

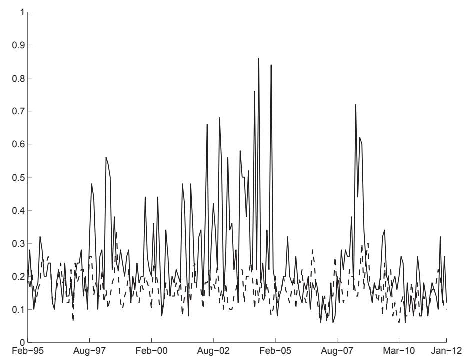


(d) VO20

Figure 3: Cumulated returns of the strategies and of the benchmark in the range 1995-2012



(a) TE1



(b) TE20

Figure 4: Turnover induced by Sharpe-based screening (dotted line) and by the TE screening with risk aversion coefficient equal to 1 (bold line - left picture) and equal to 20 (bold line - right picture)

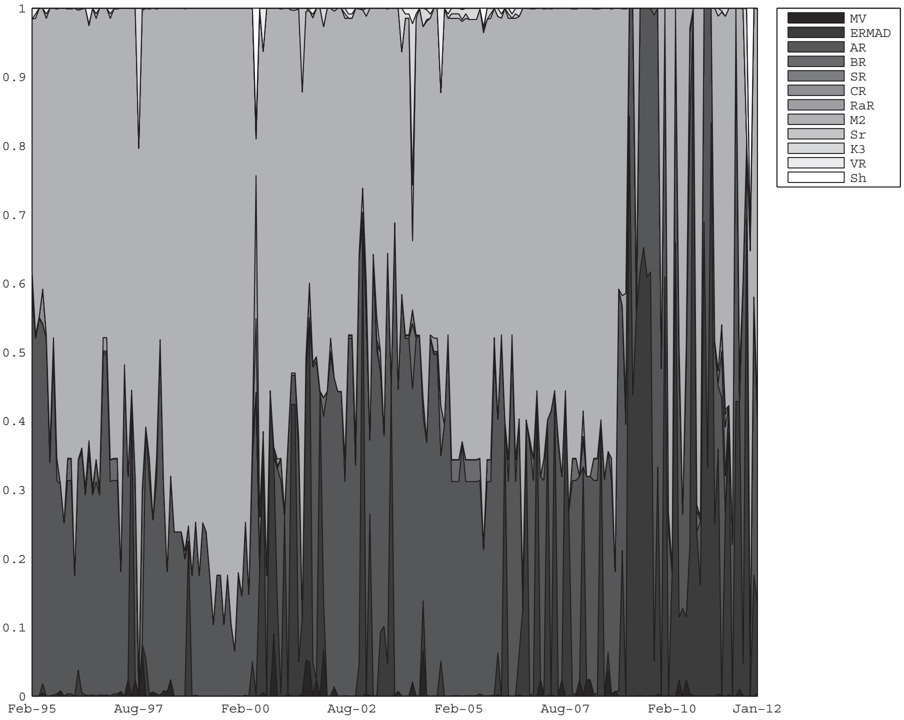
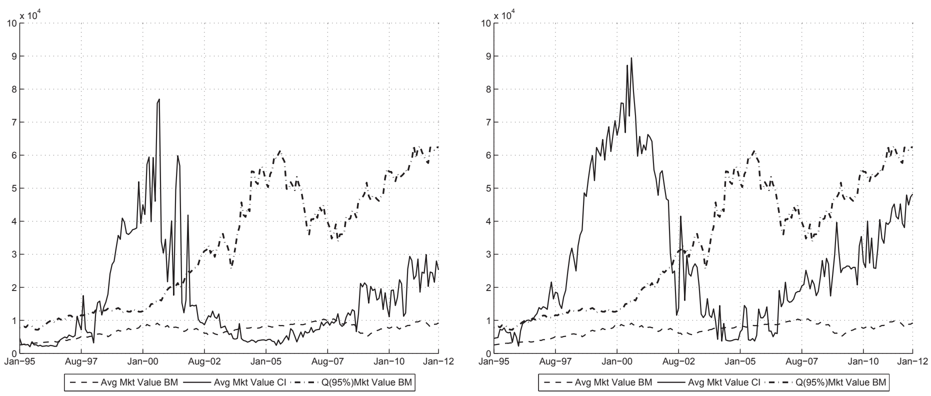


Figure 5: Performance measures weights in the TE screening with risk aversion coefficient equal to 1



(a) Free MV

(b) MV=10%

Figure 6: Average market value of the assets selected by the TE screening with risk aversion equal to 1 with unconstrained (bold line - left figure) or constrained market value weight (market value fixed at 10% - bold line - right figure). The figures also report the average market value of the companies included in the benchmark (dashed line), and the 95% quantile of the market value of the companies included in the benchmark (dotted line).