ROMA TRE
UNIVERSITÀ DEGLI STUDI

COLLANA DEL
DIPARTIMENTO DI ECONOMIA

A NEW LP MODEL FOR ENHANCED INDEXATION

Renato Bruni - Francesco Cesarone - Andrea Scozzari - Fabio Tardella

ISSN 2279-6916 Working papers
(Dipartimento di Economia Università degli studi Roma Tre) (online)

Working Paper n° 168, 2012

I Working Papers del Dipartimento di Economia svolgono la funzione di divulgare tempestivamente, in forma definitiva o provvisoria, i risultati di ricerche scientifiche originali. La loro pubblicazione è soggetta all'approvazione del Comitato Scientifico.

Per ciascuna pubblicazione vengono soddisfatti gli obblighi previsti dall'art. l del D.L.L. 31.8.1945, n. 660 e successive modifiche.

Copie della presente pubblicazione possono essere richieste alla Redazione.

esemplare fuori commercio
ai sensi della legge 14 aprile 2004 n.106

REDAZIONE:
Dipartimento di Economia
Università degli Studi Roma Tre
Via Silvio D'Amico, 77 - 00145 Roma
Tel. 0039-06-57335655 fax 0039-06-57335771
E-mail: dip_eco@uniroma3.it

DIPARTIMENTO DI ECONOMIA

A NEW LP MODEL FOR ENHANCED INDEXATION

Renato Bruni - Francesco Cesarone - Andrea Scozzari - Fabio Tardella

A New LP Model for Enhanced Indexation

Renato Bruni[1], Francesco Cesarone[2], Andrea Scozzari[3], Fabio Tardella[4]

[1] "Sapienza" Università di Roma
Dip. di Ingegneria Informatica, Automatica e Gestionale
bruni@dis.uniroma1.it
[2] Università degli Studi Roma Tre
Dipartimento di Economia
fcesarone@uniroma3.it
[3] Università degli Studi "Niccolò Cusano" - Telematica, Roma
Facoltà di Economia
andrea.scozzari@unisu.it
[4] "Sapienza" Università di Roma
Dip. Metodi e Modelli per l'Economia, il Territorio e la Finanza
fabio.tardella@uniroma1.it

Abstract

Enhanced Indexation is the problem of selecting a portfolio that should produce excess return with respect to a given benchmark index. In this work we propose a linear bi-objective optimization approach to Enhanced Indexation that maximizes average excess return and minimizes underperformance over a learning period. This can be formulated as a simple Linear Programming problem that is solved to optimality by standard LP codes. Moreover, we investigate conditions that guarantee or forbid the existence of a portfolio strictly outperforming the index. We present extensive computational analysis of the results on publicly available real-world financial datasets, including comparison with previous results, performance and diversification analysis, and empirical verification of some of the proposed theoretical results.

JEL classification: C610, D810, G110.

Keywords: Enhanced Indexation, Linear Programming, Performance Analysis, Portfolio Management, No Arbitrage condition.

1 Introduction

In the field of Asset Management, the basic Index Tracking (IT) problem consists in selecting a portfolio, possibly with a small number of assets, that best replicates (*tracks*) the performance of a given index or benchmark. This problem is usually formulated as the problem of minimizing a given distance measure, computed over a learning period, between the index and a tracking portfolio that uses at most m out of n available assets. Extensive reviews of the literature on this problem can be found in [4] and, more recently, in [6] and [9].

A more ambitious and desirable objective in Asset Management is that of outperforming a given index or benchmark. This problem has been recently addressed with various approaches under the name of Enhanced Indexation (EI) or Enhanced Index Tracking (see [6] and references therein). The portfolio obtained with this approach is sometimes called Enhanced Indexation (EI) portfolio, and its return in excess to that of the index is called excess return. In general, no guarantee of always obtaining a positive excess return can be given, so an Enhanced Indexation portfolio may sometimes underperform the index. The EI problem is often viewed as a risk-controlled active management strategy, consisting in two goals: maximizing excess return and minimizing tracking error at a controlled level.

EI models are usually built and validated using the price data of the n assets and of the benchmark index over a time period. In order to simulate practical usage, a part of this time period is considered the past (and so it is known), and the rest is considered the future (unknown at the time of portfolio selection). The past (called in-sample) is used for finding the EI portfolio, while the future (called out-of-sample) can only be used for testing the performance of the already selected portfolio.

Several approaches to the EI problem have been recently proposed. A detailed overview of the literature is given below. However, in our opinion, most existing approaches have three main limitations. First, EI bi-objective models (or their scalarizations) based on minimizing tracking error and maximizing excess return contain a contradiction in their purposes. On one hand, the first goal penalizes both positive and negative deviations from the index while, on the other hand, one seeks to maximize the mean of positive deviations. This contradiction derives from the use of a symmetric distance measure, which is not suitable for controlling the distance between the returns of the portfolio and those of the benchmark in this case, and can be avoided by using an asymmetric distance measure. Furthermore, EI is a computationally demanding task [20], and several proposed models are too complex for being practically solved to optimality for medium or large size problems. They are therefore only solved approximately by means of heuristics. Finally, several authors do not test their models on publicly available datasets, so comparison is generally impracticable.

In Section 2 we present a linear bi-objective risk-return model for the EI problem overcoming the above limitations. The proposed model consists in maximizing the excess return of the selected portfolio with respect to an index, while minimizing a downside risk measure, evaluated as the maximum underperformance with respect to the same index. This model can be formulated as a simple Linear Programming problem, and this allows for efficient solution of the basic model and reasonable computational complexity when adding further complicating constraints coming from real-world practice.

In the same section we investigate some theoretical aspects of our model. More precisely, we establish conditions for the existence of a portfolio strictly outperforming the benchmark when the number of assets is greater than the number of time periods. We then show that, when the number of time periods is greater than the number of assets, a classical No Arbitrage condition guarantees that there is no portfolio strictly outperforming the index, and that the only portfolio weakly outperfoming the index is the one realizing the index itself.

Finally, an extensive computational analysis of the behavior of the proposed model is reported in Section 3. We apply our model to eight major stock markets across the world using datasets publicly available in [5] and already used in some similar studies for the single period case. We compare the performance of our portfolios with that of the portfolios obtained in works that use the same datasets, and we further analyze such performance by using a rolling time window approach.

Furthermore, we empirically verify the theoretical results presented in Section 2. Finally, we analyze the diversification of the EI portfolios along all the efficient frontier. Results are very encouraging and show that portfolios selected by our model have a good performance and exhibit several useful properties.

1.1 Literature Overview

We sketch here an overview of the literature in the field of Enhanced Indexation. This is a relatively recent area, where quantitative approaches have been mainly developed in the last decade [6].

In a seminal study, Beasley *et al.* [4] examine the trade-off between the tracking error (for Index Tracking) and excess return. This trade-off is managed through a parameter in the objective function that weights one objective against the other. They consider five datasets drawn from major world markets, that have then been made publicly available in Beasley's OR-Library.

Alexander and Dimitriu [2] move from Index Tracking to Enhanced Indexation by means of a model based on a cointegration approach. This model constructs portfolios that track two artificial indexes: the index plus a constant (index-plus) and the index minus the same constant (index-minus). They attempt to generate excess returns by selling the plus tracking portfolio and by purchasing the minus tracking portfolio. Their methodology is tested on 30 stocks of the Dow Jones Industrial Average (DJIA) over 13 years (Jan.1990-Dec.2003), using daily closing prices and considering as a benchmark a reconstructed DJIA index. The same authors apply in [3] the same methodology also to FTSE100, CAC, and S&P100 finding evidence of excess returns (called abnormal).

Dose and Cincotti [7] present a two step methodology for selecting an enhanced indexation portfolio. First they detect a subset of assets by cluster analysis and then they set the portfolio weights by solving a bi-objective optimization problem: minimizing the root mean squared distance between the portfolio returns and the index returns, and maximizing the average difference between the portfolio returns and the index returns. They impose constraints limiting (to 50) the number of assets to be held in the portfolio, and setting lower and upper bounds on the fraction of capital invested in each asset. Their data set consists in 487 assets from S&P500 for 753 days, but considering an artificial index and in-sample and out-of-sample periods of 40 days. They show the out-of-sample average (on 16 periods) Tracking Error and average Excess Returns, but no comparison with the real S&P500 index, nor computational times, are given.

Konno and Hatagi [11] minimize the mean absolute deviation between index values plus a positive factor alpha and the enhanced indexation portfolio values and impose also minimal transaction costs. They formulate the problem as a concave minimization under linear constraints, and they solve it by means of a branch and bound algorithm. Computational results are conducted by using monthly data of 225 stocks from the Nikkei 225. They construct the portfolios on two in-sample windows (Jan.1995-Dec.1997 and Jan.1996-Dec.1998), and they test the performance of the chosen portfolio for an out-of-sample period of 12 months (Jan.1998-Dec.1998 and Jan.1999-Dec.1999). Computational times vary from few seconds to 15 min., depending on the dataset and on parameters setting.

Wu *et al.* [26] suggest a bi-objective goal programming problem, where the first goal is to constrain the tracking error to be at most a prefixed level. The second objective is to maximize the portfolio return. They consider a dataset from the Taiwan Stock Market with 426 stocks, using daily price data covering years 2002-2005. A rolling time window method is used to evaluate the performance of the selected portfolio. The training period is set to two years, whereas the following three months are the out-of-sample period (in total they consider eight testing periods). In these testing periods 6 out of 8 times the performance of the portfolio has better performance w.r.t. the market index. No computational times are presented.

Canakgoz and Beasley consider in [6] both index tracking and enhanced indexation problems, adopting a linear regression based view of the returns of the tracking (or enhanced indexation) portfolio as depending on benchmark index returns. For enhanced indexation they propose a two-stage optimization problem using a mixed-integer linear programming formulations. The first stage consists in achieving 1 as slope of the regression, while the regression intercept is maximized. In the second stage they minimize transaction costs subject to the regression coefficients (intercept and slope) obtained before. Computational results are presented for eight publicly available datasets from Beasley's OR-Library.

Koshizuka *et al.* [12] deal with Enhanced Indexation by minimizing the mean absolute deviation between index values plus a factor alpha and the enhanced indexation portfolio values and by imposing a constraint on the correlation between the weights of the selected portfolio and those of the benchmark. This model is tested on the Tokyo Stock Exchange with around 1500 assets, considering three non-overlapping time windows where the in-sample period is 3 years and the out-of-sample is 1 year. No running times are reported.

Meade and Beasley [15] propose a modified Sortino ratio as objective for selecting a portfolio outperforming the benchmarck. The authors investigate momentum strategy: they assume that performance observed in the recent past will continue into the near future. They use S&P Global 1200 and its subsets (Europe, UK, Japan, Australia, Canada, etc.) with their respective market indexes, finding evidence of significant momentum profits.

Li *et al.* [13] develop a multi-objective optimization model for the enhanced indexation problem, where the excess return is maximized and the downside standard deviation (relative to the benchmark index) is minimized. They use as decision variables the number of units of stock, so the problem becomes a non-linear-multi-objective problem with integer variables. The model is solved by an evolutionary algorithm on a subset of the data sets used in [6]. They consider as input to the model an in-sample period of 145 weeks and they test the performance of the selected portfolio on an out-of-sample of 145 weeks. No computational times are shown.

Roman *et al.* [20] apply a Second order Stochastic Dominance strategy (see also [8, 19]) to construct a portfolio whose return distribution dominates the ones of a benchmark. Empirical studies are conducted on three datasets (FTSE 100, SP 500, and Nikkei 225) using weekly returns. The authors consider the possibility of rebalancing the portfolio composition each week, for a total of ten weeks, which represents the backtesting sample. Computations require few seconds for each problem instance.

Thomaidis [25] suggests an enhanced indexation approach with prefixed investment goal (on the excess return and on the probability that the enhanced indexation portfolio underperforms the benchmark) and with risk constraints by using techniques of Fuzzy set theory. The model includes a cardinality constraint and buy-in threshold and it is formulated as a mixed-integer nonlinear programming problem. The problem is solved by using heuristics: simulated annealing, genetic algorithms, particle swarm optimization. Experimental results are presented for Dow Jones Industrial Average index with 30 securities, but no computational times are shown.

Guastaroba and Speranza [9] use a heuristic approach (called Kernel Search) for solving mixed-integer linear programming models for IT and EI including also cardinality, buy-in, and transaction costs constraints. They evaluate the efficiency and accuracy of their heuristic comparing it with a standard exact solver. Furthermore, they use the eight publicly available datasets from Beasley's OR-Library, to compare the accuracy of their IT model with that of [4], and provide examples of the out-of-sample performance of their IT and EI models on two datasets.

2 A Linear Risk-Return Enhanced Indexation Model

2.1 Formulation of the Model

We construct an enhanced indexation portfolio by observing and using the price value of n stocks and the value of the index we are tracking over $T+1$ time periods $0, 1, 2, \dots, T$. We use the following notation:

p_{it} is the price of the i-th asset at time t, with $t = 0, ..., T$;

b_t is the benchmark index value at time t, with $t = 0, ..., T$;

$R_t^I = \dfrac{b_t - b_{t-1}}{b_{t-1}}$ is the benchmark index return at time t, with $t = 1, ..., T$;

$r_{it} = \dfrac{p_{it} - p_{i(t-1)}}{p_{i(t-1)}}$ is the i-th asset return at time t, with $t = 1, ..., T$;

x is the vector whose components x_i are the fractions of a given capital invested in asset i in the enhanced indexation portfolio we are selecting.

Adopting a standard approximation we assume that

$R_t(x) = \sum_{i=1}^{n} x_i r_{it}$ is the portfolio return at time t, with $t = 1, ..., T$; so that

$\delta_t(x) = R_t(x) - R_t^I$ is the excess return, or overperformance, of the selected portfolio w.r.t. the benchmark index at time t, with $t = 1, ..., T$. Note that $-\delta_t(x)$ is the underperformance of the selected portfolio w.r.t. the benchmark index at time t.

Following a classical paradigm we would like to maximize return and, at same time, minimize risk. Thus, we propose a linear bi-objective risk-return model where the objectives are:

(a) the maximization of the (average) excess return of the selected portfolio w.r.t. a benchmark index: $\max \frac{1}{T} \sum_{t=1}^{T} (\sum_{i=1}^{n} x_i r_{it} - R_t^I) = \max \frac{1}{T} \sum_{t=1}^{T} \delta_t(x)$

(b) the minimization of the downside risk defined as the maximum underperformance w.r.t. the same index in the past periods: $\min_x \max_t -\delta_t(x)$

Note that a negative [resp. positive] value of objective (b) corresponds to a positive [resp. negative] excess return. All efficient solutions of this bi-objective problem can be found by solving a family of single objective problems depending on a parameter K (called *risk level*) that specifies the maximum allowed risk (in the sense of underperformance).

Thus the model becomes:

$$
\begin{aligned}
\phi(K) = \max_x \quad & \frac{1}{T} \sum_{t=1}^{T} \delta_t(x) \\
s.t. \quad & -\delta_t(x) \le K \quad t = 1, \dots, T \\
& \sum_{i=1}^{n} x_i = 1 \\
& x_i \ge 0 \qquad\quad i = 1, \dots, n
\end{aligned}
\tag{1}
$$

2.2 No Arbitrage and the minimum risk portfolio

In model (1), (limited) positive values of K allow to construct portfolios that might have (limited) underperformances with respect to the index in the in-sample. On the other hand, negative values of K imply that the constructed portfolios strictly overperform the index in all the in-sample. The minimum feasible value of the parameter K in model (1) can be found by solving the problem:

$$\begin{aligned} K_{min} = \min_{x,K} \quad & K \\ s.t. \quad & -\delta_t(x) \leq K \quad t = 1, \ldots, T \\ & \sum_{i=1}^{n} x_i = 1 \\ & x_i \geq 0 \qquad\quad i = 1, \ldots, n \end{aligned} \tag{2}$$

Thus, the optimal solution to this problem yields the portfolio with minimum risk. Note that the optimal value K_{min} of problem (2) is nonpositive if one can find a tracking portfolio that never underperforms the index in all past observations. This is clearly not always possible. However, when the number T of observations is smaller than the number of assets, the index returns are realized by a portfolio composed by all assets, and the assets returns are independent, the optimal portfolio selected in (2) is guaranteed to strictly overperform the index in past observations so that $K_{min} < 0$, as proved in the following theorem (see also Section 3.4).

Let $R^I = (r_1^I, \ldots, r_T^I)$ be the vector of the index returns and let $R^i = (r_{i1}, \ldots, r_{iT})$ denote the vector of returns of asset i, for $i = 1, \ldots, n$. We say that the index returns are realized by a *complete portfolio* if $R^I = \sum_i \tilde{x}_i R^i$ for some $\tilde{x}_1 > 0, \ldots, \tilde{x}_n > 0$ with $\sum_i \tilde{x}_i = 1$. We recall that points $v^1, \ldots, v^m$ in $\mathbb{R}^T$ are said to be *affinely independent* if $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$, $\sum_{i=1}^m \lambda_i = 0$, and $\sum_{i=1}^m \lambda_i v^i = 0$ imply $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0$. This is equivalent to requiring that the convex hull of these points is a polytope of dimension $m - 1$. Note that $m \leq T + 1$ randomly chosen points in $\mathbb{R}^T$ are affinely independent with probability 1. We also recall that the open mapping theorem states that the images of open sets through a surjective linear mapping are open.

Theorem 1 *Assume that $T < n$, that among the vectors $R^1, \ldots, R^n$ of the assets returns there are $T + 1$ affinely independent vectors, and that the index returns $R^I = (r_1^I, \ldots, r_T^I)$ are realized by a complete portfolio. Then there exists a portfolio that strictly overperforms the index in all past observations, i.e., $\delta_t(x) = R_t(x) - R_t^I > 0$ for $t = 1, \ldots, T$.*

Proof. Let $\Delta = \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1,\ x_i \geq 0,\ i = 1 \ldots, n\}$ be the standard simplex in $\mathbb{R}^n$ and let $F : \Delta \to \mathbb{R}^T$ be the linear mapping defined by $F(x) = \sum_{i=1}^n x_i R^i$. By assumption we have that $R^I = F(\tilde{x}) = \sum_i \tilde{x}_i R^i$ for some $\tilde{x}$ in the interior of Δ. By the open mapping theorem we then deduce that R^I belongs to the interior of $F(\Delta)$, which is the bounded polyhedron obtained as the convex hull of the points $R^1, \ldots, R^n$. Since among these points there are $T + 1$ affinely independent vectors, we have that $F(\Delta)$ is a full-dimensional polyhedron, so that the ball $B(R^I, \epsilon) = \{y \in \mathbb{R}^T : \| R^I - y \| \leq \epsilon\}$ is contained in $F(\Delta)$ for some $\epsilon > 0$. Thus, in particular, the point $R_\epsilon^I = (r_1^I + \epsilon, \ldots, r_T^I + \epsilon)$ belongs to $F(\Delta)$, so that there exists $(x'_1, \ldots, x'_n) \in \Delta$ with $F(x'_1, \ldots, x'_n) = R_\epsilon^I$. In other words the entries of R_ϵ^I are the returns of the feasible portfolio determined by the investments $(x'_1, \ldots, x'_n)$. This portfolio clearly outperforms the index in all past observations. □

Absence of arbitrage is a common assumption in financial markets. In this framework a standard No Arbitrage (NA) condition (see, e.g., [17]) requires that there exists no long-short portfolio

$y = (y_1, \ldots, y_n)$, where y_i denotes the amount of asset i purchased (if $y_i > 0$) or shorted (if $y_i < 0$), that gives a positive profit at time 0, i.e., satisfies $\sum_{i=1}^n y_i p_{i0} < 0$, and yields nonnegative returns for all periods, i.e., satisfies $\sum_{i=1}^n y_i r_{it} \geq 0$, for all $t = 1, \ldots, T$. A stronger version of the No Arbitrage condition requires in addition that every self-financing portfolio (i.e., such that $\sum_{i=1}^n y_i p_{i0} = 0$) that yields nonnegative returns for all periods must actually yield zero returns in all periods, i.e., $\sum_{i=1}^n y_i r_{it} = 0$, for all $t = 1, \ldots, T$.

We now show that, under a full column rank condition on the matrix R of returns (i.e., the matrix whose columns are the vectors R^i, $i = 1, \ldots, n$), if the index returns are realized by a portfolio and we assume that the strong No Arbitrage condition holds, then the only portfolio that weakly outperforms the index in all past observations is the one realizing the index.

Theorem 2 *Assume that the returns matrix R has full column rank, that the index returns $R^I = (r_1^I, \ldots, r_T^I)$ are realized by a portfolio, and that the strong No Arbitrage condition holds. Then the only portfolio that weakly outperforms the index in all past observations is the one realizing the index.*

Proof. Let $\tilde{x} \in \Delta$ be a portfolio realizing the index, i.e., such that $R^I = \sum_i \tilde{x}_i R^i$ and assume that there exists a portfolio $x \in \Delta$ that outperforms the index in all past observations, i.e., such that $\sum_i x_i R^i \geq R^I$ or, equivalently, $R(x - \tilde{x}) = \sum_i (x_i - \tilde{x}_i) R^i \geq 0$. Observe that the (long-short) portfolio $y = x - \tilde{x}$ is self-financing since $\sum_{i=1}^n x_i - \sum_{i=1}^n \tilde{x}_i = 1 - 1 = 0$. Then, by the strong No Arbitrage condition, we must have $R(x - \tilde{x}) = 0$ which implies $x = \tilde{x}$ by the assumption of linear independence of the columns of R. □

An immediate consequence of Theorems 1 and 2 is that arbitrage must be possible under the assumptions of Theorem 1. So, for instance, in order to assume a No Arbitrage condition one should assume that $T \geq n$. Furthermore, one can observe that, under very mild assumptions, obtaining a negative value for K_{min} becomes quite unlikely when increasing the number T of observations. Indeed, if we assume that any portfolio x has a positive probability ϵ of underperforming the index in any period t, then the probability of finding a portfolio that overperforms the index in all past observations is given by $(1 - \epsilon)^T$, which rapidly converges to zero as T increases. It is also straightforward to observe that the value of K_{min} is nondecreasing with respect to T, since increasing the in-sample window can never decrease the worst underperformance K_{min}.

Some computational experiments, described in the next Section, show that on the data sets considered there seems to be a threshold $T \simeq 1.7n$ below which $K_{min} < 0$.

2.2.1 The maximum return portfolio

The other extreme case in our bi-objective model consists in maximizing excess return regardless of the underperformance risk. This is formulated as

$$
\begin{aligned}
\delta_{max} = \quad & \max_{x} \quad \frac{1}{T} \sum_{t=1}^{T} \delta_t(x) \\
& s.t. \quad \sum_{i=1}^{n} x_i = 1 \\
& \qquad\quad x_i \geq 0 \quad i = 1, \ldots, n
\end{aligned}
\tag{3}
$$

Let I^* be the set of all indices i^* of the assets with maximum average return, i.e., such that $\sum_{t=1}^T r_{i^*t} \geq \sum_{t=1}^T r_{it}$ for all i. Then it is straightforward to show that the set of all solutions to problem (3) coincides with the set of all portfolios containing only assets with indices in I^*. Thus the maximum value K_{max} of the downside risk of the portfolios on the efficient frontier of our bi-objective model is given by

$$K_{max} = \min_{i^* \in I^*} \max_{1 \le t \le T} (R_t^I - r_{i^*t}),$$

while $\delta_{max} = \frac{1}{T} \sum_{t=1}^{T} (r_{i^*t} - R_t^I)$ is the value of the maximum average excess return with respect to the benchmark.

2.3 Properties of the efficient frontier

For every value of K between K_{min} and K_{max} the optimal solution to problem (1) provides a portfolio on the risk-return efficient frontier with average excess return $\phi(K)$, while for $K < K_{min}$ problem (1) is infeasible, and for $K > K_{max}$ the optimal solution coincides with the one for $K = K_{max}$.

The risk-return efficient frontier is thus obtained as the graph of the function $\phi(K)$ on the interval $[K_{min}, K_{max}]$. Some theoretical properties of the function $\phi(K)$ are easily derived from known results in parametric Linear Programming.

Theorem 3 *The function $\phi(K)$ is piecewise linear, concave and increasing on the interval $[K_{min}, K_{max}]$.*

For a proof in the general case see, e.g., [16].

3 Computational Analysis

In this section we test our model in two ways on some publicly available datasets: we first adopt a single period approach (SP), i.e., we consider a single in-sample window and a single subsequent out-of-sample window; then we use a rolling time window approach (RTW), i.e., we shift the in-sample window (and consequently the out-of-sample window) all over the time length of each dataset.

With the SP approach we perform a partial comparison of our model with two recent EI techniques that use the same datasets: Canakgoz and Beasley [6] and Guastaroba and Speranza [9]. However, the three models considered are quite different, so there is no direct correspondence between the parameters used. Nevertheless, such a comparison is performed by putting, as far as possible, these models in the same working conditions so as to provide a reasonable idea of their performances.

In the real world, however, the markets are in continuous evolution, and it is desirable to rebalance the portfolio from time to time in order to take new information into account. An RTW approach allows this rebalancing, thus capturing non-stationary market conditions, and is therefore better suited for practical application. For these reasons, extensive results on all datasets are reported with this approach.

Furthermore, we empirically test the theoretical properties discussed in Section 2, and we analyze the diversification of the obtained portfolios.

3.1 Data Sets

We support the view of Canakgoz and Beasley [6] that researchers should try to compare their models on a sufficiently large number of datasets, which should also be (or should be made)

publicly available. This would greatly simplify the evaluation of the quality of the proposed models. For this reason, we conduct an extensive analysis on the eight real-world datasets [5] frequently used in studies on portfolio management that are available at
`http://people.brunel.ac.uk/~mastjjb/jeb/orlib/indtrackinfo.html`.

Those datasets consist in weakly price data from March 1992 to September 1997 (i.e. 291 historical realizations) for the following capital market indexes:

- Hang Seng (Hong Kong), file indtrack1, containing 31 assets;
- DAX 100 (Germany), file indtrack2, containing 85 assets;
- FTSE 100 (UK), file indtrack3, containing 89 assets;
- S&P 100 (USA), file indtrack4, containing 98 assets;
- Nikkei 225 (Japan), file indtrack5, containing 225 assets;
- S&P 500 (USA), file indtrack6, containing 457 assets;
- Russell 2000 (USA), file indtrack7, containing 1318 assets;
- Russell 3000 (USA), file indtrack8, containing 2151 assets.

The return rates for these markets have been computed as relative variations of the quotation prices $(P_t - P_{t-1})/P_{t-1}$, thus obtaining 290 outcomes.

3.2 Single Period Performance Evaluation

For the above data sets, we compute the portfolio that gives the best excess return (w.r.t. the benchmark) for a given risk level in the in-sample window, and we then analyze (in terms of excess return) the performance of the chosen portfolio in the out-of-sample period. Sample intervals are set in order to perform the comparison respectively with [6] and [9].

Table 1 contains the first comparison. 'Best C&B' reports the best yearly (out-of-sample) percentage Average Excess Return (AER) obtained from Table 5 of [6], and the number of assets in the corresponding portfolio. 'Fixed cardinality' reports the same information for portfolios obtained by solving model (1) on in-sample [1, 145] and with out-of-sample [146, 290], as in [6], and choosing the risk level K that produces a portfolio composed by the same number of assets of [6] (making portfolios as comparable as possible). 'Bounded cardinality' reports again the best AER obtained by using model (1) on the same sample intervals, but this time choosing portfolios composed by a number of assets ranging between 5 and 10.

Observe that our portfolios provide higher AER than 'Best C&B' in half of the cases when they include the same number of assets, and in 6 out of 8 cases where we choose fewer assets. In order to evaluate the magnitude of the AER differences between the considered approaches, we also compute the average over all datasets of the AER (even if it has no direct financial interpretation). We observe that our EI portfolios, in particular those with few assets, have a better behavior.

Guastaroba and Speranza [9] report two graphs showing out-of-sample portfolio returns on the FTSE 100 and the S&P 100 for the portfolios obtained with their EI approach. More precisely, the graphs describe the cumulative returns of the portfolios, which correspond to the values of wealth after τ periods [18], given by

$$W_\tau = W_{\tau-1}(1 + R_\tau(x)) \qquad \tau = 1, \ldots, 52$$

	Best C&B [6]		EI portfolio			
			Fixed cardinality		Bounded cardinality	
	selected assets	AER	selected assets	AER	selected assets	AER
Hang Seng	10	-2.43	10	-2.51	10	-2.51
DAX 100	10	11.88	10	15.88	5	16.40
FTSE 100	10	5.01	10	10.39	5	12.59
S&P 100	10	2.30	10	11.30	5	19.39
Nikkei	10	7.81	10	4.00	7	5.62
S&P 500	40	14.78	40	11.30	8	21.84
Russell 2000	70	12.19	70	16.92	5	52.63
Russell 3000	90	22.62	90	18.59	5	69.98
Average		9.27		10.73		24.49

Table 1: Average Excess Return (AER) comparison with portfolios from Canakgoz and Beasley [6]

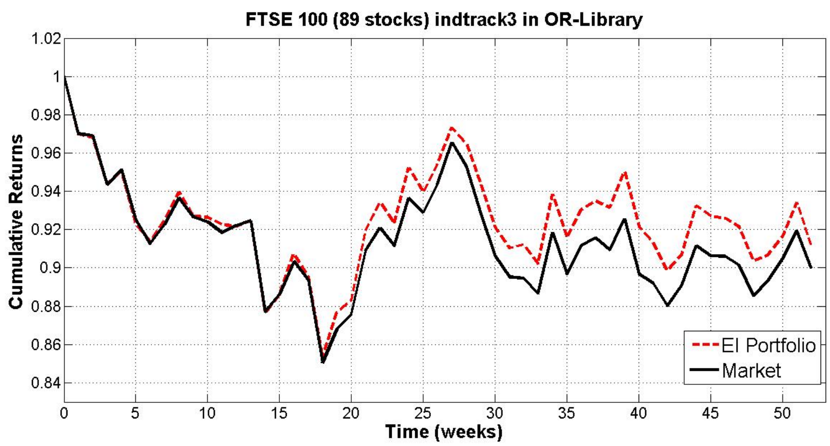


Figure 1: Out-of-sample performance for FTSE100 (SP approach)

with initial wealth $W_0 = 1$. We now present the outcome of our model for those two datasets in the same format.

Figures 1 and 2 are obtained by solving model (1) on the in-sample period [1, 104] and with out-of-sample [105, 156], as in [9]. No useful indication is given there for the choice of our risk level K. In the absence of this, we assume portfolios in [9] to be quite low risk portfolios because they closely track the index and are obtained by imposing narrow constraints on the amount of each asset ($0.01 \leq x_i \leq 0.1$ if the i-th asset is selected in the portfolio). We therefore show results corresponding to our minimum risk EI portfolios compared to the performance of the market index in the same period. The graphs show a slight overperformance of the EI portfolios with respect to the market index similar to what provided in [9]. Furthermore, in Figure 3 we report a similar experiment (same sample windows and risk level) on the largest dataset (Russel 3000) and we observe that a good excess return is achieved on all the out-of-sample periods. We remark, however, that the best performances obtained by our model generally correspond to higher risk levels, as reported in the following subsection.

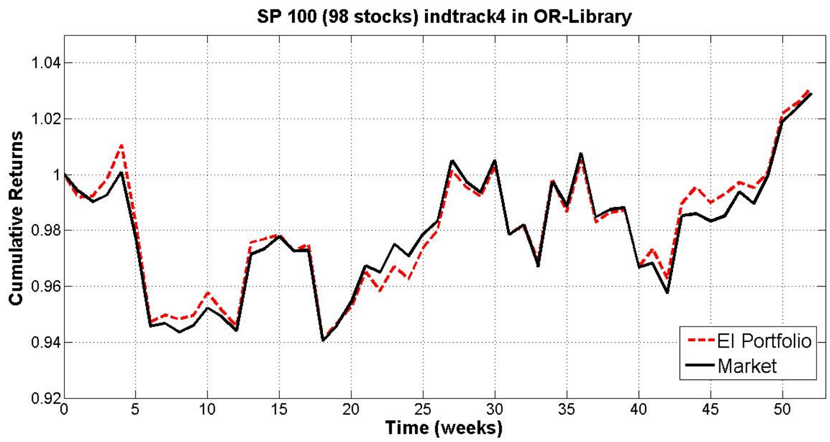


Figure 2: Out-of-sample performance for S&P100 (SP approach)

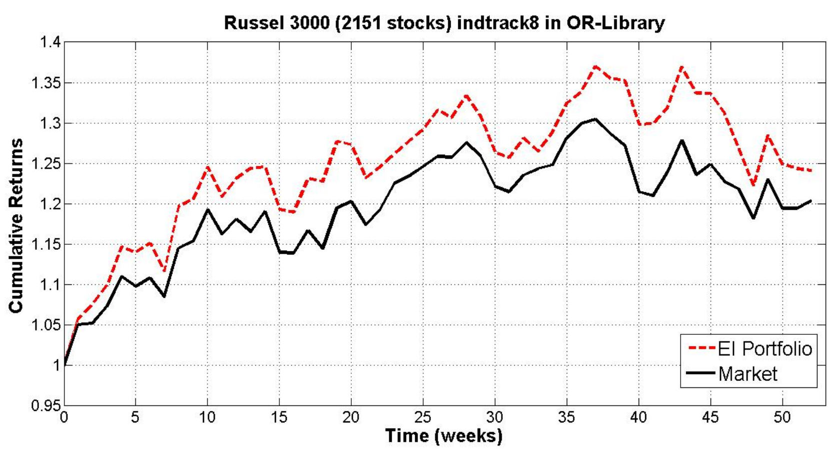


Figure 3: Out-of-sample performance for Russel 3000 (SP approach)

	assets	K_1 ($\times 10^{-2}$)	K_2 ($\times 10^{-2}$)	Market ($\times 10^{-2}$)
Hang Seng	31	0.469	**0.613**	0.456
DAX 100	85	0.567	**0.852**	0.631
FTSE 100	89	0.368	**0.486**	0.357
S&P 100	98	0.501	**0.700**	0.510
Nikkei	225	-0.049	-0.130	**-0.042**
S&P 500	457	**-0.210**	-0.893	-0.316
Russell 2000	1318	0.175	**0.567**	-0.004
Russell 3000	2151	-0.069	**0.602**	-0.297

Table 2: Out-of-sample average returns of our portfolios and of the market index

We conclude this comparison by highlighting that, with our linear programming model, the time required for computing each portfolio for FTSE 100 and S&P 100 is about 0.01 sec, while for the largest dataset, Russel 3000, the average time required for computing each of our portfolios is 0.17 sec.

3.3 Rolling Time Window Evaluation

We now allow for the possibility of changing the portfolio composition (rebalancing) during the holding period. Clearly, frequent rebalances would be useful from the optimization point of view, but would also be practically infeasible because of the implied transaction costs. Thus, although we do not explicitly take them into account in our model, we have chosen a holding period that represents a compromise between the two requirements above.

We obtain EI portfolios by solving model (1) for different values of the risk level K on in-sample intervals of 200 periods repeatedly shifted all over the dataset. More precisely, for each of those in-sample intervals, we evaluate the portfolio performance in the following 4 weeks (out-of-sample), during which no rebalances are allowed. After this, we shift the mentioned in-sample window by 4 weeks in order to cover the out-of-sample period, we recompute the optimal portfolio w.r.t. the new in-sample window and repeat. We thus obtain 22 different EI portfolios. For instance, the first in-sample is [1,200] and the corresponding out-of-sample is [201,204], the second in-sample is [5,204] and the corresponding out-of-sample is [205,208]. We consider two risk levels corresponding to the following values

$$K_1 = K_{min} \qquad K_2 = K_{min} + 1/4(K_{max} - K_{min}).$$

Table 2 reports the average out-of-sample returns of the EI portfolios compared to the corresponding average returns of the market index. Best results for each dataset are marked in bold. Observe that the EI portfolios outperform the market index in 7 out of 8 cases, and each of the two strategies K_1 and K_2 provide portfolios that outperform the market index in 5 and 6 out of 8 cases, respectively.

In addition, we report the outcomes of two standard performance measures: the Sharpe Ratio [22, 23] and the Rachev Ratio [18]. The Sharpe Ratio is the ratio between the expected return and its standard deviation, namely $P_s = E[R(x)]/\sigma(R(x))$. However, when the expected return is negative this index has no meaning, so we report "-". The Rachev Ratio is defined as the ratio between the average of the best $\beta\%$ returns of a portfolio and thet of the worst $\alpha\%$ returns. Parametes α and β have been set equal to 0.1. Sharpe and Rachev ratios were selected

	assets	K_1	K_2	Market
Hang Seng	31	0.178	**0.186**	0.170
DAX 100	85	**0.314**	0.273	0.302
FTSE 100	89	**0.236**	0.221	0.222
S&P 100	98	**0.250**	0.226	0.247
Nikkei	225	-	-	-
S&P 500	457	-	-	-
Russell 2000	1318	0.049	**0.106**	-
Russell 3000	2151	-	**0.106**	-

Table 3: Out-of-sample Sharpe Ratio values of our portfolios and of the market index

	assets	K_1	K_2	Market
Hang Seng	31	1.082	**1.280**	1.041
DAX 100	85	**1.408**	1.268	1.171
FTSE 100	89	1.233	1.065	**1.264**
S&P 100	98	1.492	**1.539**	1.510
Nikkei	225	0.932	0.847	**0.938**
S&P 500	457	**1.023**	0.910	0.920
Russell 2000	1318	0.919	**1.096**	0.902
Russell 3000	2151	**1.055**	0.933	0.889

Table 4: Out-of-sample Rachev Ratio values of our portfolios and of the market index

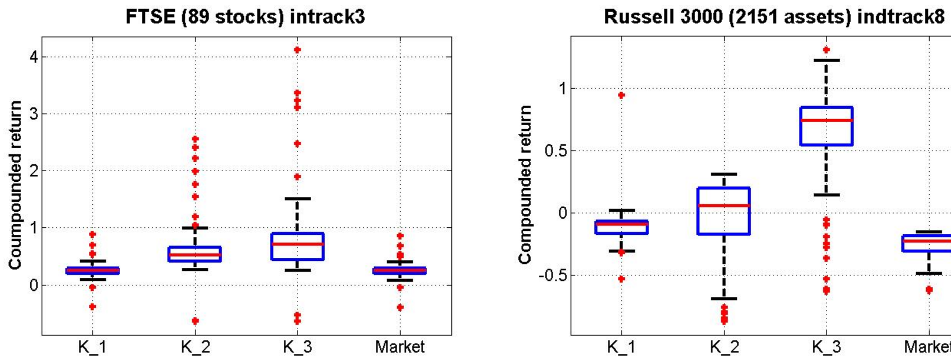
because they are somehow complementary: while the first one is more focused on the central part of the return distribution, the latter stresses its tails. Other commonly used performance indices (Sortino Ratio [24], etc.) have been also computed and their results (not shown here but available upon request) turned out to be similar to the reported ones.

Table 3 reports the Sharpe Ratio values for the EI portfolios and for the market index. Best results for each datasets are marked in bold. In this case, results obtained by the EI portfolios are always better than the benchmark. Table 4 reports the Rachev Ratio values for the EI portfolios and for the market index. Best results for each datasets are marked in bold. In 6 out of 8 cases the EI portfolios outperform the market index.

In order to better understand the behavior of our model, we compute the yearly compounded out-of-sample return CR_τ (after τ periods) of the 22 EI portfolios in the following way:

$$CR_\tau = \Big[\prod_{t=1}^{\tau}(1 + R_t(x))\Big]^{\frac{52}{\tau}} - 1 \qquad \tau = 1, \ldots, 88$$

where $R_t(x)$ is the *t-th* value of the 88 weekly out-of-sample returns (4 values for each of the 22 out-of-sample windows) of the EI portfolios. The following analysis is then performed considering 3 different risk levels: K_1, K_2, defined as above, and $K_3 = K_{min} + 1/2(K_{max} - K_{min})$. As an example, we provide the box plots of results for the FTSE 100 (Figure 4a) and for the Russell 3000 (Figure 4b) datasets. In the figures, each box represents the yearly compounded return distribution; the central mark is the median and the edges are the 25^{th} and the 75^{th} percentiles, the whiskers correspond to approximately ± 2.7 times the standard deviation, and the outliers



(a) FTSE 100

(b) Russell 3000

Figure 4: Box plot of yearly compounded return

are represented individually. The yearly compounded return distribution of the EI portfolios with minimum risk level tends to be similar to that of the market index, while for higher risk levels the EI portfolios seem preferable. Note that this happens also for the FTSE 100, where the market index was preferable according to the Rachev Ratio.

3.4 Analysis of Minimum Risk Portfolios

In order to analyze the theoretical results presented in Section 2, and the corresponding assumptions, we compute the minimum value of the maximum allowed underperformance K_{min} for the datasets described above. We then examine the sign of K_{min} with respect to the value of the ratios between the number T of the in-sample observations and the number n of assets. Specifically, $n = \{31, 85, 89, 98, 255, 457, 1318, 2151\}$, while T ranges from 10 to 290, which is the maximum number of available observations. Table 5 reports the values of K_{min} considering the market index as a benchmark. Observe that K_{min} is negative when T is not much larger than n (approximately $T < 1.7n$), while K_{min} is positive for larger values of T. This agrees with the results of Section 2 where it is shown that, under the assumptions of Theorem 1, for $T < n$ a minimum-risk portfolio strictly outperforming the in-sample market index always exists (or, equivalently, the maximum allowed underperformance K_{min} is negative).

On the other hand, for T sufficiently greater than n, the assumptions of Theorem 2 hold, so that a minimum-risk EI portfolio coincides with the benchmark if the latter is a feasible portfolio, and this implies $K_{min} = 0$. Here, the behavior exhibited for $T \leq 1.7n$ is consistent with these results. However, for $T > 1.7n$, the results appear inconsistent with Theorem 2 because for T sufficiently large K_{min} is different from 0. This apparent inconsistence could be explained by the fact that, in those cases, either the market index is not a feasible portfolio or the No Arbitrage condition does not hold, or the matrix R has no full column rank.

We now check that using a feasible portfolio as a benchmark, gives results consistent with those of Theorem 2, at least for the datasets for which this can be tested (which are the first four because they are the only ones that satisfy $T > 1.7n$).

We thus consider as a benchmark the so-called *naïve* or *uniform* portfolio, namely $R_t^I = \sum_{i=1}^{n} r_{it}/n$, which is a feasible portfolio for problem (1). Table 6 reports the corresponding values of K_{min} w.r.t. the uniform portfolio. In all instances, if $T < 1.7n$ then $K_{min} < 0$. On the other hand, when T is sufficiently large with respect to n (i.e., approximately $T > 1.7n$), K_{min}

T	Hang Seng $n = 31$ ($\times 10^{-2}$)	DAX 100 $n = 85$ ($\times 10^{-2}$)	FTSE 100 $n = 89$ ($\times 10^{-2}$)	S&P 100 $n = 98$ ($\times 10^{-2}$)	Nikkei $n = 225$ ($\times 10^{-2}$)	S&P 500 $n = 457$ ($\times 10^{-2}$)	Russell 2000 $n = 1318$ ($\times 10^{-2}$)	Russell 3000 $n = 2151$ ($\times 10^{-2}$)
10	-0,933	-1,089	-1,822	-1,153	-1,951	-3,644	-7,546	-9,091
30	-0,238	-0,440	-0,549	-0,452	-0,741	-0,846	-2,245	-2,105
50	-0,037	-0,135	-0,266	-0,258	-0,412	-0,725	-1,781	-1,715
70	0,090	-0,059	-0,186	-0,180	-0,245	-0,473	-1,347	-1,155
90	0,098	-0,029	-0,124	-0,115	-0,192	-0,413	-1,183	-0,919
110	0,219	-0,017	-0,086	-0,054	-0,153	-0,365	-1,001	-0,702
130	0,240	-0,003	-0,044	-0,035	-0,114	-0,307	-0,884	-0,613
150	0,278	0,013	-0,017	-0,009	-0,095	-0,261	-0,825	-0,550
170	0,280	0,022	0,013	0,006	-0,074	-0,224	-0,771	-0,511
190	0,284	0,029	0,025	0,024	-0,064	-0,171	-0,699	-0,449
210	0,311	0,034	0,038	0,033	-0,050	-0,150	-0,640	-0,424
230	0,311	0,041	0,045	0,042	-0,041	-0,128	-0,526	-0,396
250	0,313	1,886	0,061	0,085	-0,037	-0,100	-0,468	-0,371
270	0,321	1,905	0,080	0,093	-0,018	-0,087	-0,435	-0,350
290	0,322	2,015	0,119	0,104	-0,003	-0,067	-0,397	-0,327

Table 5: Minimum risk K_{min} with market index as benchmark

approaches zero, so the outcome is fully consistent with the results of Section 2.

3.5 Diversification Analysis and Cardinality Constraints

The diversification of a portfolio is a fundamental property in portfolio management. Diversification should be guaranteed by a good risk-return model, especially for low risk strategies. For example, in the original Markowitz model [14] diversification is obtained through variance minimization. On the other hand, another important requirement, specifically in the case of index tracking and of enhanced indexation problems, is that of finding a portfolio that uses only a limited number of assets. This is usually obtained by imposing *cardinality constraints* that typically require the use of integer variables thus greatly increasing the computational complexity of the model.

Figures 5a and 5c show the number of assets having $x_i > 0$ in the EI portfolios obtained solving model (1) with different percentages of the maximum risk level, computed as

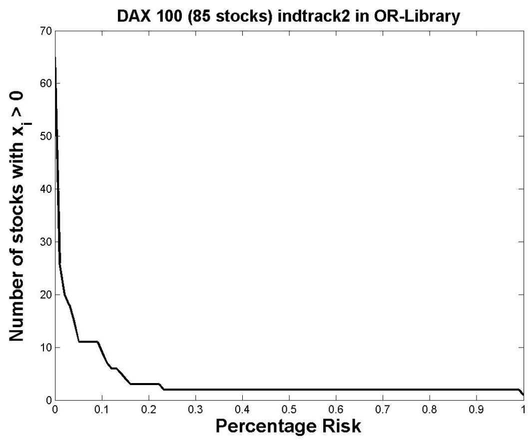
$$\tilde{K} = (K - K_{min})/(K_{max} - K_{min}).$$

Note that the minimum risk level K_{min} and the maximum risk level K_{max} correspond to $\tilde{K}$ values of 0% and 100%, respectively. The sample datasets considered here are DAX 100 and S&P 500. However, the results are very similar for all other instances. Figures 5b and 5d also report, for the same datasets, a diversification analysis obtained by computing the Herfindahl Index

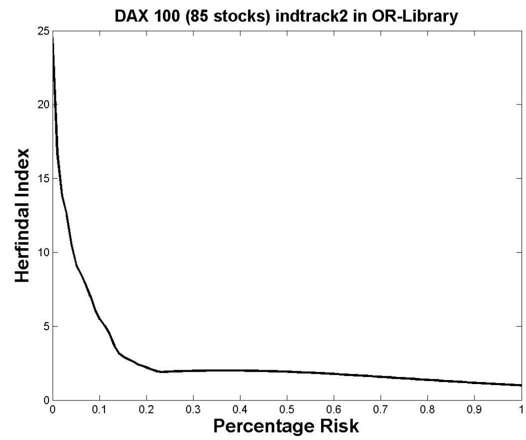
$$H(x) = (\sum_{i=1}^{n} x_i^2)^{-1},$$

which is considered a common measure of diversification (see, e.g., [1]). We point out that these two diversification analyses are practically equivalent.

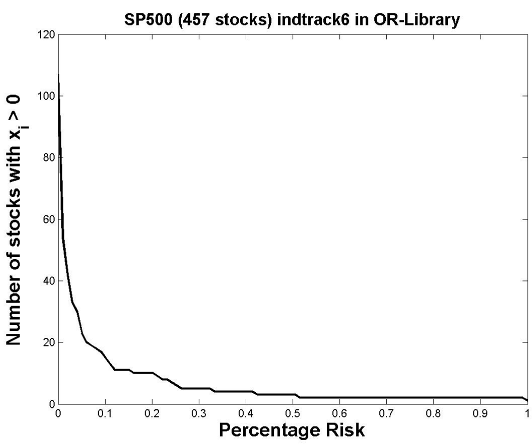
We observe that the EI portfolios obtained with our model are very diversified for small risk values, while the number of assets included in the optimal portfolios becomes rapidly small for slightly larger risk values. This avoids the use of complicating cardinality constraints for constructing the efficient frontiers. However, our model easily allows the introduction of additional



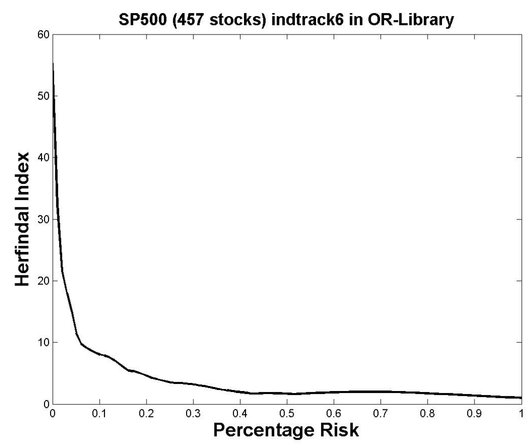
(a) DAX 100: Number of assets with positive weights vs Percentage risk



(b) DAX 100: Herfindal Index vs Percentage risk



(c) S&P 500: Number of assets with positive weights vs Percentage risk



(d) S&P 500: Herfindal Index vs Percentage risk

Figure 5: Diversification analysis

T	Hang Seng $n = 31$ $(\times 10^{-2})$	DAX 100 $n = 85$ $(\times 10^{-2})$	FTSE 100 $n = 89$ $(\times 10^{-2})$	S&P 100 $n = 98$ $(\times 10^{-2})$	Nikkei $n = 225$ $(\times 10^{-2})$	S&P 500 $n = 457$ $(\times 10^{-2})$	Russell 2000 $n = 1318$ $(\times 10^{-2})$	Russell 3000 $n = 2151$ $(\times 10^{-2})$
10	-0,813	-1,190	-1,593	-1,318	-1,788	-3,678	-7,482	-8,522
30	-0,176	-0,617	-0,510	-0,482	-0,673	-0,988	-2,127	-2,273
50	-0,039	-0,251	-0,202	-0,272	-0,353	-0,753	-1,651	-1,833
70	0	-0,134	-0,109	-0,209	-0,195	-0,540	-1,219	-1,335
90	0	-0,054	-0,055	-0,110	-0,144	-0,469	-1,054	-1,128
110	0	-0,026	-0,030	-0,049	-0,110	-0,406	-0,872	-0,937
130	0	-0,011	-0,013	-0,023	-0,077	-0,314	-0,737	-0,802
150	0	0	0	-0,010	-0,051	-0,249	-0,668	-0,717
170	0	0	0	0	-0,037	-0,232	-0,644	-0,681
190	0	0	0	0	-0,027	-0,198	-0,588	-0,633
210	0	0	0	0	-0,019	-0,173	-0,526	-0,570
230	0	0	0	0	-0,013	-0,133	-0,447	-0,487
250	0	0	0	0	-0,011	-0,112	-0,405	-0,440
270	0	0	0	0	-0,009	-0,091	-0,371	-0,409
290	0	0	0	0	-0,006	-0,081	-0,335	-0,373

Table 6: Minimum risk K_{min} with uniform portfolio as benchmark

real-world constrains. Indeed, the constrained version of model (1) with cardinality constraints and buy-in thresholds can be reformulated as a Mixed Integer Linear Program by adding n binary variables y_i:

$$\begin{array}{lll}
\max\limits_{x,y} & \dfrac{1}{T}\sum\limits_{t=1}^{T}\delta_t(x) & \\
 & & \\
s.t. & -\delta_t(x) \leq K & t = 1,\dots,T \\
 & \sum\limits_{i=1}^{n} x_i = 1 & \\
 & \ell_i y_i \leq x_i \leq u_i y_i & i = 1,\dots,n \\
 & \sum\limits_{i=1}^{n} y_i \leq m & \\
 & x_i \geq 0 & i = 1,\dots,n \\
 & y_i \in \{0,1\} & i = 1,\dots,n
\end{array} \qquad (4)$$

For moderate sizes of n this problem can be solved to optimality by general purpose mixed integer linear programming (MILP) solvers like CPLEX. Nevertheless, for larger problems, specialized and possibly approximate methods are required. As an example, in Figure 6 we report the efficient frontiers for the Hang Seng (31 assets) and DAX 100 (85 assets) datasets for 100 equally spaced values of K in $[K_{min}, K_{max}]$. This is done both in the unconstrained case and in the cardinality constrained case with at most ten assets ($m = 10$). As expected from previous diversification analysis, the two efficient frontiers coincide for all but the smallest risk levels.

Models (1) and (4) have been coded in MATLAB 7.11.0 and executed on a PC with Intel Core i3 CPU M330 2.13 GHz with 4Gb RAM under MS Windows 7, by using the exact solver CPLEX 11.0, which is called from MATLAB with the TOMLAB/CPLEX toolbox [10]. When solving model (4) for computing all the efficient frontier, computational times are 45 secs. for Hang Seng and 958 secs. for DAX 100. When solving model (1), on the other hand, the corresponding computations require less than 1 sec.

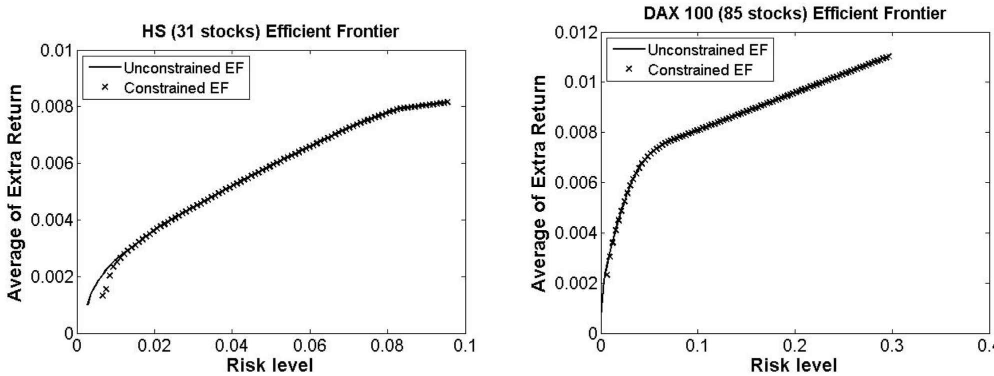
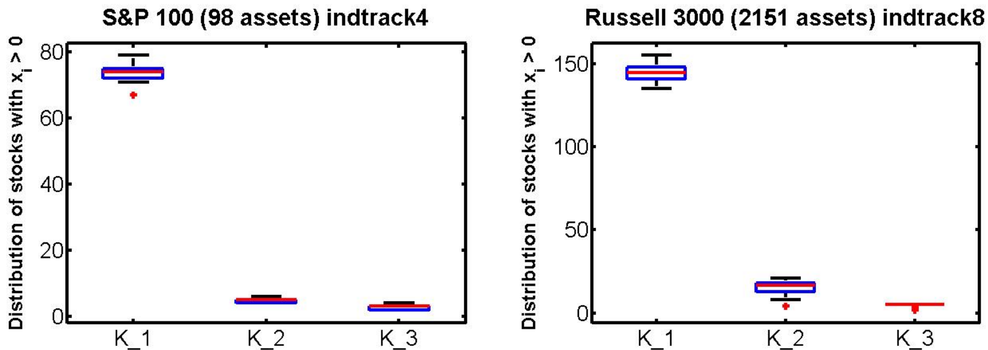


Figure 6: Examples of Cardinality Constrained Efficient Frontiers



(a) S&P 100

(b) Russell 3000

Figure 7: Distributions of the number of selected assets (RTW approach)

Finally, we analyze the number of assets selected in the portfolios obtained with the different in-sample windows in the RTW approach, again for the above described risk levels (K_1, K_2, K_3). We observe that the number of selected assets is fairly stable in all in-sample windows, as reported in Figure 7a and Figure 7b that show the box plots of the distribution of those number of assets for the datasets S&P 100 and Russel 3000. This analysis confirms that imposing cardinality constraints in our model is necessary only for the smallest risk levels.

4 Conclusions

We proposed a new simple risk-return approach to the Enhanced Indexation problem. In spite of its simplicity, our model is able to find portfolios that exhibit out-of-sample performances that seem comparable or even superior, to those reported in previous works on the same problem. We chose to avoid cluttering the presentation of our model with complicating real-world constraints also in order to highlight some theoretical connections between a No Arbitrage condition and the existence of a portfolio outperforming the index. However, the linearity of our model easily allows for the addition of further constraints coming from real-world practice such as the cardinality constraints and buy-in thresholds mentioned in Section 3, or the turn-over or UCITS constraints described in [21]. A detailed analysis of the effect of such constraints on our model is left for future research.

References

[1] A. Adam, M. Houkari, and J. P. Laurent, *Spectral risk measure and portfolio selection*, Journal of Banking and Finance **32** (2008), 1870–1882.

[2] C. Alexander and A. Dimitriu, *Indexing and statistical arbitrage*, The Journal of Portfolio Management **31** (2005), no. 2, 50–63.

[3] ______, *Indexing, cointegration and equity market regimes*, International Journal of Finance and Economics **10** (2005), 213–231.

[4] J. E. Beasley, N. Meade, and T. J. Chang, *An evolutionary heuristic for the index tracking problem*, European Journal of Operational Research **148** (2003), 621–643.

[5] J.E. Beasley, *Or-library: Distributing test problems by electronic mail*, Journal of the Operational Research Society **41** (1990), 1069–1072.

[6] N.A. Canakgoz and J.E. Beasley, *Mixed-integer programming approaches for index tracking and enhanced indexation*, European Journal of Operational Research **196** (2008), 384–399.

[7] C. Dose and S. Cincotti, *Clustering of financial time series with application to index and enhanced index tracking portfolio*, Physica A: Statistical Mechanics and its Applications **355** (2005), no. 1, 145–151.

[8] C.I. Fábián, G. Mitra, D. Roman, and V. Zverovich, *An enhanced model for portfolio choice with ssd criteria: a constructive approach*, Quantitative Finance **11** (2011), no. 10, 1525–1534.

[9] G. Guastaroba and M.G. Speranza, *Robust investment strategies with discrete asset choice constraints using DC programming*, European Journal of Operational Research **217** (2012), 54–68.

[10] K. Holmstrom, A.O. Goran, and M.M. Edvall, *Users guide for TOMLAB*, November 2007.

[11] H. Konno and T. Hatagi, *Index-plus-alpha tracking under concave transaction cost*, Journal of Industrial and Management Optimization **1** (2005), no. 1, 87–98.

[12] T. Koshizuka, H. Konno, and R. Yamamoto, *Index-plus-alpha tracking subject to correlation constraint*, International Journal of Optimization: Theory, Methods and Applications **1** (2009), 215–224.

[13] Q. Li, L. Sun, and L. Bao, *Enhanced index tracking based on multi-objective immune algorithm*, Expert Systems with Applications **38** (2011), no. 5, 6101–6106.

[14] H. M. Markowitz, *Portfolio selection: Efficient diversification of investments*, Cowles Foundation for Research in Economics at Yale University, Monograph 16, John Wiley & Sons Inc., New York, 1959.

[15] N. Meade and JE Beasley, *Detection of momentum effects using an index out-performance strategy*, Quantitative Finance **11** (2011), no. 2, 313–326.

[16] K. G. Murty, *Linear programming*, John Wiley & Sons Inc., New York, 1983.

[17] E. Z. Prisman, *Valuation of Risky Assets in Arbitrage Free Economies with Frictions*, The Journal of Finance **16** (1986), no. 3, 545–557 (English).

[18] S. Rachev, A. Biglova, S. Ortobelli, and S. Stoyanov, *Different Approaches to Risk Estimation in Portfolio Theory*, Journal of Portfolio Management **31** (2004), no. 1, 103–112.

[19] D. Roman, K. Darby-Dowman, and G. Mitra, *Portfolio construction based on stochastic dominance and target return distributions*, Mathematical Programming **108** (2006), no. 2, 541–569.

[20] D. Roman, G. Mitra, and V. Zviarovich, *Enhanced indexation based on second-order stochastic dominance*, Available at SSRN: http://ssrn.com/abstract=1776966 (2011).

[21] A. Scozzari, F. Tardella, S. Paterlini, and T. Krink, *Exact and heuristic approaches for the Index Tracking problem with UCITS constraints*, Proceedings APMOD 2012: Applied Mathematical Optimization and Modelling. Available at SSRN: http://ssrn.com/abstract=2066672 (2012).

[22] W.F. Sharpe, *Mutual fund performance*, The Journal of Business **39** (1966), no. 1, 119–138.

[23] ______, *The sharpe ratio*, The Journal of Portfolio Management **21** (1994), no. 1, 49–58.

[24] F. A. Sortino and S. Satchell, *Managing downside risk in financial markets*, Batterworth-Heinemann Finance, Oxford, 2001.

[25] N. Thomaidis, *A soft computing approach to enhanced indexation*, Natural Computing in Computational Finance (2012), 61–77.

[26] L.C. Wu, S.C. Chou, C.C. Yang, and C.S. Ong, *Enhanced index investing based on goal programming*, The Journal of Portfolio Management **33** (2007), no. 3, 49–56.