Economics Bulletin

Volume 33, Issue 2

No arbitrage and a linear portfolio selection model

Renato Bruni
Università di Roma "Sapienza", Dip. di Ingegneria Informatica, Automatica e Gestionale

Francesco Cesarone
Università degli Studi Roma Tre, Dip. di Studi Aziendali

Andrea Scozzari
Università degli Studi "Niccolò Cusano" - Telematica, Roma, Facoltà di Economia

Fabio Tardella
Università di Roma "Sapienza", Dip. Metodi e Modelli per l''Economia, il Territorio e la Finanza

Abstract

We propose a linear bi-objective optimization approach to the problem of finding a portfolio that maximizes average excess return with respect to a benchmark index while minimizing underperformance over a learning period. We establish some theoretical results linking classical No Arbitrage conditions to the existence of a feasible portfolio for our model that strictly outperforms the index. Empirical analyses on publicly available real-world financial datasets show the effectiveness of the model and confirm the described theoretical results.

Citation: Renato Bruni and Francesco Cesarone and Andrea Scozzari and Fabio Tardella, (2013) "No arbitrage and a linear portfolio selection model", *Economics Bulletin*, Vol. 33 No. 2 pp. 1247-1258.
Contact: Renato Bruni - bruni@dis.uniroma1.it, Francesco Cesarone - fcesarone@uniroma3.it, Andrea Scozzari - andrea.scozzari@unicusano.it, Fabio Tardella - fabio.tardella@uniroma1.it.
Submitted: March 20, 2013.   **Published:** May 21, 2013.

1. Introduction

Several approaches have been proposed to find portfolios that best correspond to the desires of typical investors. When the target is selecting a portfolio, possibly with a small number of assets, that best tracks the performance of a given index or benchmark, the problem is called *Index Tracking* (IT). This problem is usually formulated as the minimization of a chosen distance between the index and a tracking portfolio that uses at most m out of the n available assets. Extensive reviews of the literature on this problem can be found in Beasley *et al.* (2003) and, more recently, in Canakgoz and Beasley (2008).

A more ambitious and desirable objective is that of outperforming the given index or benchmark. This problem has been recently addressed with various approaches under the name of *Enhanced Indexation* (EI), or *Enhanced Index Tracking*. The portfolio selected in this case is sometimes called *Enhanced Indexation Portfolio* (EI portfolio), and its return in excess to that of the index is called *excess return*. After a seminal study in Beasley *et al.* (2003), quantitative approaches to EI have been: Alexander and Dimitriu (2005a), Alexander and Dimitriu (2005b), Dose and Cincotti (2005), Konno and Hatagi (2005), Roman *et al.* (2006), Wu *et al.* (2007), Canakgoz and Beasley (2008), Koshizuka *et al.* (2009), Li *et al.* (2011), Meade and Beasley (2011), Roman *et al.* (2011), Fábián *et al.* (2011), Bruni *et al.* (2012), Thomaidis (2012). 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, and can be avoided by using an asymmetric distance measure. Furthermore, EI is a computationally demanding task (see, e.g., Roman *et al.*, 2011) 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.

We present here 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 only 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, as reported in Section 2, and this allows for its efficient solution even in large markets. The simplicity of our model also allows for a theoretical analysis of the connections with classical *No Arbitrage* conditions, as explained in Section 3. 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, the No Arbitrage condition implies that there is no portfolio strictly outperforming the index, and that the only portfolio weakly outperforming the index is the one realizing the index itself. These conditions can be related to the well-known Farkas' lemma, that has several applications in different fields (see, e.g., Murty, 1983, Bruni and Bianchi, 2012). Finally, in Section 4, we provide empirical results on the performance behavior of the proposed model on eight major stock markets using datasets publicly available. We also verify empirically the theoretical results linking our model to the No Arbitrage condition.

2. A Linear Risk-Return Model

EI models are usually built and validated by using the price values of the n assets and of the benchmark index over a time interval. In order to simulate practical usage, a part of this interval is considered the past, and so it is known, and the rest is considered the future, supposed unknown at the time of portfolio selection. The past (called in-sample) is used for selecting the EI portfolio, while the future (called out-of-sample) can only be used for testing the performance of the selected portfolio. Let the in-sample be constituted by $T + 1$ time periods $0, 1, 2, \ldots, T$. We use the following notation:

p_{it} is the price of the asset i at time t, with $i = 0, \ldots, n$ and $t = 0, \ldots, T$;

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

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

$r^I_t = \dfrac{b_t - b_{t-1}}{b_{t-1}}$ is the benchmark index return at time t, with $t = 1, \ldots, T$;

x_i are the fractions of a given capital invested in asset i in the EI portfolio we are selecting;

$R_t(x) = \sum_{i=1}^{n} x_i r_{it}$ is the standard approximation of portfolio return at time t; so that

$\delta_t(x) = R_t(x) - r^I_t$ is the excess return, or overperformance, of the selected portfolio w.r.t. the benchmark index at time t, with $t = 1, \ldots, T$. 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: $\max_x \dfrac{1}{T}\sum_{t=1}^{T} \delta_t(x)$,

(b) the minimization of the downside risk, defined here as the maximum underperformance: $\min_x \max_t -\delta_t(x)$.

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 that we call *risk level*, specifying the maximum allowed risk (in the sense of underperformance), as follows:

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

3. No Arbitrage, Minimum Risk and Maximum Return

Solving model (1) for positive values of K produces portfolios that might have underperformances (in the in-sample). On the other hand, for negative values of K we obtain portfolios strictly overperforming the index (in the in-sample). However, too small negative values for K may produce infeasiblity of the model. The minimum feasible value of K can be found by solving the problem:

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

The optimal solution to this problem yields the portfolio with *minimum risk* K_{min}. Note that K_{min} is nonpositive if the optimal portfolio never underperforms the index. Clearly, this is not always possible. However, there are conditions under which the optimal portfolio is guaranteed to strictly overperform the index ($K_{min} < 0$), as proved in Theorem 1 below.

Let $R^I = (r_1^I, \ldots, r_T^I)$ be the vector of the index returns and let $R^i = (r_{i1}, \ldots, r_{iT})$ be the vector of returns of asset i, for $i = 1, \ldots, n$. We say that the index returns are realizable by a *complete portfolio* if $R^I = \sum_i \tilde{x}_i R^i$ for some $\tilde{x}$ having all strictly positive components (a complete portfolio is a portfolio containing all assets of the market). We recall that points $v^1, \ldots, v^m$ in $\mathbb{R}^T$ are *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 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 $R^I = (r_1^I, \ldots, r_T^I)$ can be realized by a complete portfolio. Then there exists a portfolio that strictly overperforms the index in all the in-sample periods, 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 the in-sample. □

Absence of arbitrage is a common assumption in financial markets. In this framework a standard No Arbitrage (NA) condition (see, e.g., Prisman, 1986) 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., a portfolio having a negative cost $\sum_{i=1}^n y_i p_{i0} < 0$, 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 some technical assumptions typically verified in practice by the matrix R of returns (i.e., the matrix whose columns are the vectors R^i, $i = 1, \ldots, n$), the strong No Arbitrage condition implies then the only portfolio that weakly outperforms the index in all the in-sample periods is the one realizing the index.

Theorem 2 *Assume that the returns matrix R has full column rank, that the index returns $R^I = (r_1^I, \ldots, r_T^I)$ can be realized by a portfolio, and that the strong No Arbitrage condition holds. Then the only portfolio that weakly outperforms the index in all the in-sample periods 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 the in-sample, 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. 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 the in-sample periods 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}.

The other extreme case in our model consists in maximizing excess return regardless of the underperformance risk. Finding the portfolio with the *maximum return* is modelled 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.

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 function $\phi(K)$ on the interval $[K_{min}, K_{max}]$. From known results in parametric Linear Programming (Murty, 1983), one can prove that the function $\phi(K)$ is piecewise linear, concave and increasing on the interval $[K_{min}, K_{max}]$.

4. Empirical Analysis

In order to allow comparison, we tested our model on publicly available real-world datasets (from `http://people.brunel.ac.uk/~mastjjb/jeb/orlib/indtrackinfo.html`; Beasley, 1990) frequently used in studies on portfolio management. Those datasets consist of weakly price data from March 1992 to September 1997 (i.e. 291 historical realizations) for the following capital market indexes: Hang Seng (Hong Kong), with 31 assets; DAX 100 (Germany), with 85 assets; FTSE 100 (UK), with 89 assets; S&P 100 (USA), with 98 assets; Nikkei 225 (Japan), with 225 assets; S&P 500 (USA), with 457 assets; Russell 2000 (USA), with 1318 assets; Russell 3000 (USA), with 2151 assets. Return rates have been computed as relative variations of the quotation prices $(p_t - p_{t-1})/p_{t-1}$, thus obtaining 290 outcomes. For the above data sets, we compute the portfolios that give the best excess return for a given risk level K in the in-sample window, and we then analyze the performances of the obtained portfolios in the out-of-sample period, by using a rolling time window scheme (RTW). Furthermore, we empirically test the theoretical properties discussed in Section 3.

4.1 Rolling Time Window Evaluation

In order to simulate practical usage, we allow for the possibility of changing the portfolio composition (rebalancing) during the holding period. More precisely, we compute EI portfolios by solving model (1) on in-sample intervals repeatedly shifted all over the dataset, and, 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 each evaluation, we shift the mentioned in-sample window by 4 weeks in order to cover the former out-of-sample period, we recompute the optimal portfolio w.r.t. the new in-sample window and repeat. We set the in-sample length at 200 periods, thus allowing 21 portfolio rebalancings over each dataset. 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 minimum and moderate resk requirements:

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

Table 1 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

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

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

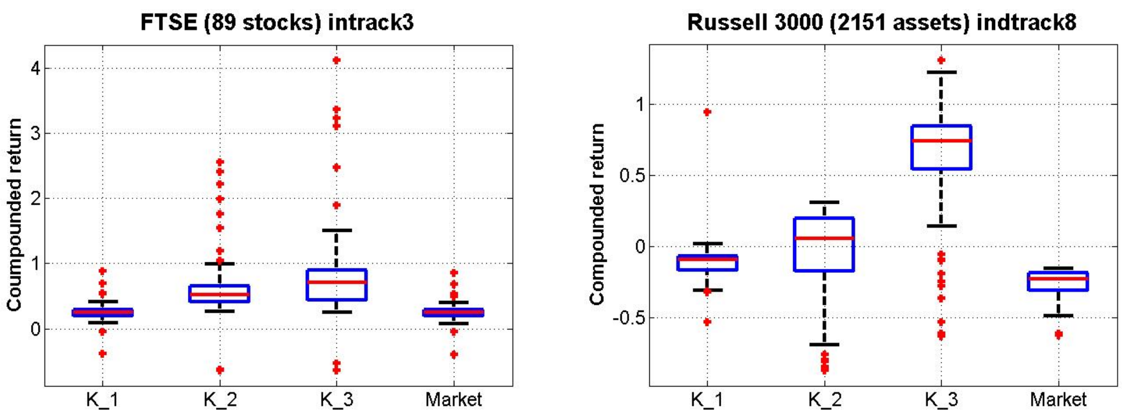
each of the two strategies K_1 and K_2 provide portfolios that outperform the market index in 5 and 6 out of 8 cases, respectively.

Moreover, we report the outcomes of two standard performance measures: the Sharpe Ratio (Sharpe, 1994 and 1996) and the Rachev Ratio (Rachev *et al.*, 2004), respectively in Tables 2 and 3, both for the computed EI portfolios and for the market index. The Sharpe Ratio is the ratio between the expected return and its standard deviation, namely $P_s = E[R(x)]/\sigma(R(x))$. However, this index has no meaning when the expected return is negative, so we report "-". The Rachev Ratio is defined as the ratio between the average of the best β% returns of a portfolio and that of the worst α% returns. Parameters α and β have been set equal to 0.1. Sharpe and Rachev ratios were selected 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. Best results for each datasets are marked in bold. Results obtained by the computed EI portfolios are always better than the benchmark by using the Sharpe Ratio analysis, while they are better than that 6 times out of 8 by using the Rachev Ratio analysis.

In order to better understand the behavior of our model, we also compute the yearly

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



(a) FTSE 100

(b) Russell 3000

Figure 1: Box plot of yearly compounded return

compounded out-of-sample return CR_τ (after τ periods) of the 22 EI portfolios, as follows:

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

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 1a) and for the Russell 3000 (Figure 1b) 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 are represented individually. The yearly compounded return distribution of the EI portfolios with minimum risk level are similar to that of the market index, or even slightly more performant than that; those of the EI portfolios with the higher risk levels are distinctly more performant than that of the market index. Note that this happens also for the FTSE 100, where the market index was preferable according to the Rachev Ratio 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	-0,933	-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 4: Minimum risk K_{min} with market index as benchmark

4.2 Analysis of Minimum Risk Portfolios

In order to analyze the theoretical results presented in Section 3, 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 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 4 reports the values of K_{min} considering the market index as a benchmark. We observe that:

(*i*) K_{min} is negative when T is not much larger than n (approximately $T < 1.7n$);

(*ii*) K_{min} is positive for larger values of T w.r.t. n (the bottom left corner).

The first observation agrees with the results of Section 3, because it shows that, under the assumptions of Theorem 1, for $T < n$ a minimum-risk portfolio strictly outperforming the market index always exists (i.e., arbitrage is possible). On the other hand, for T sufficiently greater than n the returns matrix R is expected to have full column rank. In this case, under the strong No Arbitrage condition, no portfolio that strictly outperforms the index can exist, and if the market index is a realizable portfolio, it is the optimal portfolio. If, on the contrary, the market index is not realizable, the optimal portfolio necessarily underperfoms the index. Observation (*ii*) above shows that the latter case holds here.

We now repeat the experiment using as benchmark the so-called *naïve* or *uniform* portfolio, namely $R_t^I = \sum_{i=1}^{n} r_{it}/n$, which is a feasible solution of model (1), hence a realizable portfolio. Table 5 reports the values of K_{min} obtained in this case. 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} becomes zero, so the outcome is fully consistent with the results of Section 3, at least for the datasets for which this can be tested (which are the first four because they are the only ones that satisfy $T > 1.7n$).

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

5. 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 good out-of-sample performances. We chose to avoid cluttering the presentation of our model with complicating real-world constraints also in order to highlight theoretical connections between the 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 (Cesarone *et al.*, 2012) and buy-in thresholds, or the turn-over or UCITS constraints (Scozzari *et al.*, 2012).

References

Alexander, C. and A. Dimitriu (2005) “Indexing and statistical arbitrage” *The Journal of Portfolio Managment* **31**, 50–63.

Alexander, C. and A. Dimitriu (2005) “Indexing, cointegration and equity market regimes” *International Journal of Finance and Economics* **10**, 213–231.

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

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

Bruni, R. and G. Bianchi (2012) “A Formal Procedure for Finding Contradictions into a Set of Rules” *Applied Mathematical Sciences* **6**, 6253–6271.

Bruni, R., F. Cesarone, A. Scozzari and F. Tardella (2012) "A new stochastic dominance approach to enhanced index tracking problems" *Economics Bulletin*, **32**, 3460–3470.

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

Cesarone, F., A. Scozzari, F. Tardella (2013) "A new method for mean-variance portfolio optimization with cardinality constraints" *Annals of Operations Research* **205**, 213–234.

Dose, C. and S. Cincotti (2005) "Clustering of financial time series with application to index and enhanced index tracking portfolio" *Physica A: Statistical Mechanics and its Applications* **355**, 145–151.

Fábián, C.I., G. Mitra, D. Roman, and V. Zverovich (2011) "An enhanced model for portfolio choice with ssd criteria: a constructive approach" *Quantitative Finance* **11**, 1525–1534.

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

Konno, H. and T. Hatagi (2005) "Index-plus-alpha tracking under concaev transaction cost" *Journal of Industrial and Managment Optimization* **1**, 87–98.

Li, Q., L. Sun, and L. Bao (2011) "Enhanced index tracking based on multi-objective immune algorithm" *Expert Systems with Applications* **38**, 6101–6106.

Meade, N. and J.E. Beasley (2011) "Detection of momentum effects using an index out-performance strategy" *Quantitative Finance* **11**, 313–326.

Murty, K.G. (1983) "Linear programming" John Wiley & Sons Inc.: New York.

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

Rachev, S., A. Biglova, S. Ortobelli, and S. Stoyanov (2004) "Different Approaches to Risk Estimation in Portfolio Theory" *Journal of Portfolio Management* **31**, 103–112.

Roman, D., K. Darby-Dowman, and G. Mitra (2006) "Portfolio construction based on stochastic dominance and target return distributions" *Mathematical Programming* **108**, 541–569.

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

Scozzari, A., F. Tardella, S. Paterlini, and T. Krink (2013) "Exact and heuristic approaches for the Index Tracking problem with UCITS constraints" *Annals of Operations Research* **205**, 235–250.

Sharpe, W.F. (1996) "Mutual fund performance" *The Journal of Business* **39**, 119–138.

Sharpe, W.F. (1994) "The sharpe ratio" *The Journal of Portfolio Management* **21**, 49–58.

Thomaidis, N.S. (2012) "A soft computing approach to enhanced indexation" In *Natural Computing in Computational Finance*, A. Brabazon *et al.*, Eds., Studies in Computational Intelligence **380**, Springer: Berlin, 61–77.

Wu, L.C., S.C. Chou, C.C. Yang, and C.S. Ong, (2007) "Enhanced index investing based on goal programming" *The Journal of Portfolio Management* **33**, 49–56.