

Document downloaded from:

<http://hdl.handle.net/10251/62638>

This paper must be cited as:

García García, F.; Guijarro Martínez, F.; Moya Clemente, I. (2013). A multiobjective model for passive portfolio management: an application on the S&P 100 index. *Journal of Business Economics and Management*. 14(4):758-775. doi:10.3846/16111699.2012.668859.



The final publication is available at

<http://dx.doi.org/10.3846/16111699.2012.668859>

Copyright Taylor & Francis: SSH Journals

Additional Information

This is an author's accepted manuscript of an article published in: "Journal of Business Economics and Management"; Volume 14, Issue 4, 2013; copyright Taylor & Francis; available online at: <http://dx.doi.org/10.3846/16111699.2012.668859>

# A multiobjective model for passive portfolio management: An application on the S&P 100 Index

**Fernando García**  
**Francisco Guijarro**  
**Ismael Moya**

*Universidad Politécnica de Valencia*  
*Camí de Vera s/n*  
*46022 – Valencia (Spain)*  
*fergarga@esp.upv.es*

## **Abstract**

Index tracking, a popular technique for passive portfolio management, seeks to minimize the unsystematic risk component by imitating the movements of a reference index. The partial index tracking only considers a subset of the stocks in the index, enabling a substantial cost reduction in comparison with full tracking. Two criteria are usually referred in the literature when constructing tracking portfolios: Tracking error variance and portfolio variance. By means of a multiobjective model, this paper considers a new parameter in the tracking error problem: The frontier curvature. This criterion is not defined for a particular portfolio, but for all the portfolios that define the tracking frontier. The main implication is that a manager can satisfy different investment profiles with the same subset of stocks. The new model is applied on the S&P 100 Index.

## **Keywords**

Index tracking; frontier curvature; tracking error variance; excess return; portfolio variance; mean-variance model; portfolio selection

## **1. Introduction**

The large number of academic papers published indicates that the analysis of the efficiency of investment funds remains a major area of research in the field of portfolio theory. The valuation of funds is still subject to analysis and comparison because of their crucial role in financial markets.

Jensen (1968) was among the first to point out the need to critically evaluate the performance of investment funds. The high number of research papers in this area has evolved in parallel with the growth in the number of managed funds and assets. Comparisons are often made between active and passive management funds (Elton *et al.*, 1993; Malkiel, 1995, 2003; Gruber 1996; Carhart, 1997; Edelen, 1999; Davis, 2001; among others).

These studies demonstrate how difficult it is for investment funds to outperform a benchmark. In the United States, the Standard & Poor's 500

Index has been between 65% and 85% more profitable than the funds analyzed over a long time window. Similar conclusions have been reached after considering the risk assumed.

The questionable success of many actively managed investment funds in outperforming the benchmark explains that index tracking is currently among the most popular techniques used by investment fund managers (Frino and Gallagher, 2001; Malkiel and Radisich, 2001, Coleman *et al.*, 2006). This technique has become even more popular after the appearance of exchange traded funds (ETFs).

Index tracking seeks to minimize the unsystematic risk component by imitating the movements of a reference benchmark – a stock index. Faced with active management techniques that endeavor to beat the underlying index, tracking portfolios in general and tracking indices in particular, are configured as a powerful passive strategy. Following this strategy, the manager does not necessarily pursue efficiency in the sense of mean-variance, but instead replicates the behavior of the market from a more conservative approach. Mixed approaches that search for consensus solutions between the two models can also be found in the literature (Burmeister *et al.*, 2005).

Index Tracking can be full or partial depending on the number of stocks that are considered.

In the case of full tracking, the portfolio includes the same stocks as the index, and an exact tracking is produced if these stocks are weighted in the same proportion as the index. It is also possible to generate other combinations of risk and return by varying the weights of the stocks in the tracking portfolio. However, in this case, the imitation of the stock index is not accurate, and it does not necessarily outperform the index in the mean-variance sense; while the greater or lesser required returns may lead to an increase or decrease in the proportional risk of the position (Roll, 1992). The disadvantages of full tracking include the high portfolio management and transaction costs, as well as the need to invest in all the stocks in the index. Another disadvantage is the cost associated with realigning the portfolio if the composition of the underlying index changes. This situation is particularly critical for stocks with low weighting in the index, and which often have little impact on the movement of the index. It is also critical for stocks with little liquidity, while various other drawbacks are mentioned in the literature (Ruiz-Torrubiano and Suarez, 2009). A restrictive view of the costs associated with tracking portfolios has also been discussed in numerous academic papers (Meade and Salkin, 1990; Adcock and Meade, 1994; Connor and Leland, 1995; Canakgoz and Beasley, 2008) and the drawbacks are usually addressed through mathematical programming models.

In partial tracking, which is the subject of this paper, a manager builds a portfolio from a subset of stocks contained in the underlying index and this process removes some of the drawbacks listed above. The counterpart is that an exact tracking of the index cannot be built. However, this does not necessarily imply a decline in the risk-return relationship – as will become evident later.

Three issues must be resolved when building a partial tracking.

Firstly, the number of stocks in the tracking must be chosen. An evaluation can be made using sensitivity analysis on the results to contrast the desirability of increasing or decreasing the cardinality of the portfolio (Tabata and Takeda, 1995).

After setting the number of stocks, the second question involves selecting the stocks among the available ones. This is the most complex problem when building a partial tracking. The simplest approach is to assess each potential stock, to measure the index tracking error, and then select those stocks that minimize this deviation. Unfortunately this approach is computationally difficult because it represents an NP-hard problem (Ruiz-Torrubiano and Suarez, 2009). For example, a number of 17,310,309,456,440 portfolios must be evaluated if the aim is to track the S&P 100 with a portfolio of ten stocks.

Finally, the third question involves the precise weight to be given to each stock in the tracking portfolio, depending on the desired return and the tracking error the manager is willing to assume.

The second issue of stock selection has received special attention from researchers and many methods for finding the local problem optimum have been proposed. These methods can be grouped into two broad families: those that make use of mathematical programming; and those using multivariate analysis techniques.

Without being exhaustive, authors using mathematical programming models for optimal local searches include: Tabata and Takeda (1995), whose approach is employed in this paper and discussed in a later section; Beasley *et al.* (2003), whose approach uses a population heuristic in which the cardinality of the portfolio is made explicit through the restriction  $\sum_{i=1}^N z_i = n$ ,  $n$  being the number of stocks in the tracking portfolio, and  $z_i$  a binary variable that indicates if the  $i$ -th stock is to be included in the portfolio. This approach supposes that the local optimum is conditioned by the whole problem resolution method being used; Derigs and Nickel (2004) use a procedure of Simulated Annealing; Ruiz-Torrubiano and Suárez (2009) combine a genetic algorithm with a model of quadratic programming in a more general formulation of the problem; Gaivoronoski *et al.* (2005) use different measures of risk in mathematical programming models, such as return variance, semi-variance, tracking error variance, or value at risk (VAR).

Works that make use of multivariate analysis techniques include: Focardi and Fabozzi (2004); and Dose and Cincotti (2005), who use cluster analysis on return time series so that the cardinality of the tracking portfolio is established from the number of clusters obtained – preferably by choosing a single stock as representative of each cluster in each tracking portfolio. Corielli and Marcellino (2006) propose the use of factor analysis so that stocks are grouped around various factors depending on their past returns, and the tracking portfolio contains those stocks that best explain the variability of these factors. In their results, several stocks are used as representatives of each factor, so the tracking portfolio can include stocks that explain the same parts of the variability in the performance of the index.

All these papers are characterized by the search for a single portfolio, characterized by a maximum of three possible parameters (Chow, 1995): tracking error variance, excess returns and volatility of returns. The stocks in the tracking portfolio are identified during this process and the given weighting complies with the constraints imposed on those parameters.

This paper proposes the addition of a new parameter: the curvature of the frontier. This criterion is not defined for a given portfolio, but for the set of portfolios that define the tracking frontier. The main advantage is that a fund manager can satisfy different investment profiles using the same subset of stocks – with all the portfolios on the frontier containing the same stocks and so reducing transaction costs.

Usually partial tracking portfolio models have attempted to obtain a single portfolio that will only satisfy those investors whose profile is perfectly aligned with the configuration chosen by the portfolio manager. It is noteworthy how this analysis has not pursued a parallel strategy to that followed in Markowitz's classical mean-variance model of 1952, in which the goal is the generation of the so-called efficient frontier – rather than the identification of a specific portfolio with a fixed risk and return.

The rest of the paper is structured as follows. The second section analytically presents the three key concepts for tracking indices: tracking error variance, excess return, and portfolio variance. The following section introduces a new criterion, the curvature of the tracking frontier, and discusses the benefits that arise from adding the concept of gradient to the previous ones. The fourth section presents a multiobjective programming model for generating tracking frontiers by simultaneously considering all these parameters. In addition, various other propositions regarding the curvature of the tracking frontier are discussed and demonstrated. In the fifth section, the above model is applied to the partial tracking of the S&P 100. A summary of the main conclusions is presented in the final section.

## 2. Parameters in the tracking portfolio problem: tracking error variance, excess return, and portfolio variance

Tracking error is defined as the difference between tracking portfolio returns and the returns produced by the index that is being tracked. Since the aim is for both portfolios to maintain a parallel evolution over time, the problem is posed as a minimization of the volatility in the tracking error. A reduction in the volatility of the tracking error means minimizing the variance in returns between the tracking portfolio and the stock index (Roll, 1992). In this way, a clear parallel with the mean-variance model (Markowitz, 1952, 1959) is established. However, with the difference that instead of looking for the portfolio with the least volatility for a given return, managers try to obtain the portfolio with the minimum tracking error variance for a given level of return in excess of the index. These are the foundations of the TEV (Tracking Error Variance) criterion: (1) minimize the TEV; (2) assume a certain TE (Tracking Error<sup>1</sup>). Both objectives are inherently conflicting, so the manager should look for consensus solutions. The TEV is given by the expression (1):

$$TEV = \mathbf{x}^t \mathbf{V} \mathbf{x} \quad (1)$$

where:

$\mathbf{x}$  = a vector of dimension  $N \times 1$ , contains the various weightings of the  $N$  stocks between the tracking portfolio and the index; that is,  $\mathbf{x} = \mathbf{x}_p - \mathbf{x}_b$ , where  $\mathbf{x}_p$  is the vector of weightings in the tracking portfolio and  $\mathbf{x}_b$  is the weighting vector in the index (subscript b for benchmark). A full tracking is obtained if all elements of  $\mathbf{x}$  are zero, while non-zero deviations can take risk-return positions that differ from the index. In the partial tracking, the vector  $\mathbf{x}_p$  will have the same number of non-zero elements as there are stocks included in the tracking,  $n$ , and the remaining weights will be left with a value of zero.

$\mathbf{V}$  = the variances-covariances matrix between the returns of the stocks.

The excess return  $G$  on the index is obtained as the difference between the returns of the tracking portfolio and the index (2):

$$G = \mathbf{x}^t \mathbf{R} = \mathbf{x}_p^t \mathbf{R} - \mathbf{x}_b^t \mathbf{R} = R_p - R_b \quad (2)$$

where:

$\mathbf{R}$  = vector of returns of  $N$  stocks.

---

<sup>1</sup> Alternatively, Rudolf *et al.* (1999) suggest using linear measurements of tracking error, and propose the use of goal programming for solving optimization models. This technique has also been recently used by Wu *et al.*, (2007).

$R_p (R_b)$  = returns of the tracking portfolio (index).

Unlike other models, in the tracking portfolio the return in excess  $G$  is obtained by subtracting the index return, and not the return of the risk-free asset. The full tracking can be easily resolved by using a quadratic mathematical model (3):

$$\begin{aligned} \text{Min} &= \mathbf{x}^t \mathbf{V} \mathbf{x} \\ \text{s. t.} \quad \mathbf{x}^t \mathbf{R} &= G \\ \mathbf{x}^t \mathbf{1} &= 0 \end{aligned} \quad (3)$$

where:

$\mathbf{1}$  = vector of dimension  $N \times 1$  with all the elements 1.

Note the need to explicitly include the constraint on  $G$ , since the profitability of the tracking portfolio and the index can differ by a constant, and the value of the TEV can then be zero. The second constraint ensures that the total investment in the tracking portfolio is the same as the index – and so the sum of positive and negative deviations is compensated. If the intention is to implement a partial tracking then an additional constraint should be added to  $n$ , although mathematical programming algorithms do not ensure the global optimum.

It is also worth to underline that the portfolios obtained with strictly positive values of  $G$  do not necessarily exceed the index. To exceed the index, in addition to having a better return than the index ( $G > 0$ ), it is necessary to obtain less variance in returns, something which is not guaranteed by model (3).

Some researchers (Canakgoz and Beasley, 2008) impose a restriction on the alpha and beta of the tracking portfolio, as estimated from the market model (4):

$$R_p = \alpha + \beta R_b \quad (4)$$

The exact imitation of the index supposes imposing restrictions (1)  $\alpha = 0$ , and (2)  $\beta = 1$ . The first restriction is equivalent to considering  $G = 0$ . The second restriction does not guarantee the achievement of efficient portfolios in the mean-variance sense. For proof of this statement, let us consider the decomposition of the total risk of a portfolio  $p$  in its systematic and unsystematic components (5):

$$\sigma_p^2 = \beta^2 \sigma_b^2 + \sigma_{e_p}^2 \quad (5)$$

where:

$\sigma_p^2 (\sigma_b^2)$  = return variance for tracking portfolio (index), with  $\sigma_p^2 = \mathbf{x}_p^t \mathbf{V} \mathbf{x}_p$  ( $\sigma_b^2 = \mathbf{x}_b^t \mathbf{V} \mathbf{x}_b$ ).

$\beta^2 \sigma_b^2$  = systematic risk of the tracking portfolio.

$\sigma_{ep}^2$  = unsystematic risk of the tracking portfolio: variance of the regression residuals between the index returns and the tracking portfolio returns.

$\sigma_{ep}^2 \geq 0$ , we have  $\sigma_p^2 \geq \sigma_b^2$  in order to impose  $\beta=1$ . That is, the tracking portfolio will offer the same return as the index ( $\alpha = 0$ ), but also with at least the same risk, which means it cannot outperform the index in the mean-variance space.

To overcome the agency problem that arises from this situation (Jorion, 2003), it is necessary to restrict the total portfolio risk rather than the systematic risk, which facilitates the generation of tracking portfolios that can exceed the index in the mean-variance space. Chow (1995) proposes a parametric model that in addition to considering the TEV and excess returns, also considers the third criterion set out in this section: return variance in the tracking portfolio.

### 3. An additional parameter: the curvature of the TEV frontier

Model (3) enables as many different portfolios to be obtained as there are different stocks with a profitability in excess  $G$ . These different portfolios are obtained by varying the weights of the stocks, and/or varying the stocks when the tracking is partial. Markowitz's minimum variance frontier model and TEV frontier appear in Figure 1. For the case of the full tracking, Roll (1992) shows that the distance in the axis of the returns variance between the two frontiers is constant,  $k$ , for any value of return  $R_p$ . Therefore, the TEV frontier is a simple shift of Markowitz's frontier in the variance axis, and the inefficiency of the index  $b$  can be quantified as  $k = \sigma_b^{2*} - \sigma_b^2$ , being constant for any portfolio on the tracking frontier.

The above property is not satisfied in the case of partial tracking. Figure 1 shows two TEV frontiers, each generated by removing a single stock from the tracking. The TEV frontier  $TEV_{-i}$  ( $TEV_{-j}$ ) results from the exclusion of the tracking of the  $i$ -th stock ( $j$ -th). Generally, the removal of one or more stocks from the tracking means a greater TEV without necessarily reducing the efficiency of the portfolios. In the example in the figure, the  $TEV_{-i}$  frontier and the  $TEV_{-j}$  frontier partially improve the efficiency of the original TEV in the mean-variance sense. Specifically, both frontiers generate better risk-returns in portfolios nearer to the  $R_b$  index than the TEV frontier in the full tracking. If the  $TEV_{-i}$  and  $TEV_{-j}$  frontiers are compared then different results will again be reached according to the considered return. However, it must always be remembered that Figure 1 only reflects risk and return, and not TEV.

[Insert Figure 1 here]



Figure 1 shows the different curvature of the  $TEV_{-i}$  and  $TEV_{-j}$  frontiers. It is precisely this characteristic that can be very useful for the fund manager. The  $TEV_{-j}$  frontier provides a better risk-return combination than the  $TEV_{-i}$  frontier for portfolios with a return of  $R_p \in [R_1, R_2]$ . However, for returns outside this range, the  $TEV_{-i}$  frontier generates returns that are clearly better than the portfolios on the  $TEV_{-j}$  frontier. In this situation, the manager must consider which of the two frontiers can best satisfy client profiles. For conservative profiles that intend to simply mimic the index, the  $TEV_{-j}$  frontier is the most suitable, and so the  $j$ -th stock is removed from the tracking. But if a return in excess  $G$  is required, then the  $TEV_{-i}$  frontier would be the best option. Therefore, not considering the curvature of the tracking portfolio frontier means that the proposed portfolios only satisfy specific values of risk, return and TEV, without considering the possibly varying risk profiles of the fund's clients.

When choosing between two tracking frontiers for a given value of  $G$  and with the same levels of risk-return and TEV, the manager must select the frontier with less curvature – because this enables more efficient options to be offered to investors. Examining the curvature of the tracking portfolio enables the manager to make a more global analysis of the offer presented to his/her clients. To achieve this, the entire TEV frontier should be examined and not just a specific point on it.

We can conclude that the manager will have the following preferences when evaluating tracking portfolios for the criteria presented:

**Assumption 1:** Investment fund manager preferences:

- a. Criteria concerning the tracking portfolio
  - a.1 Return: portfolios with higher returns are preferred, *ceteris paribus*.
  - a.2 Return variance: portfolios with less risk are preferred, *ceteris paribus*.
  - a.3 TEV: portfolios with less TEV are preferred, *ceteris paribus*.
- b. Criteria concerning the TEV frontier
  - b.1 Curvature of the TEV frontier: TEV frontiers with less curvature are preferred, *ceteris paribus*.

Note that for the curvature of the TEV frontier, Figure 1 only shows the frontier in the mean-variance plane. We will assume that TEV frontiers are preferred with less curvature in the mean-variance and mean-TEV spaces. The following section presents a multiobjective mathematical programming model that enables the simultaneous consideration of all these preferences. This methodology has been widely published in the field of operations research (Zeleny, 1974, 1982; Steuer, 1986), and is currently used in many financial applications (Zopounidis, 1998; Hallerbach and Spronk, 2002).

#### 4. A multiobjective approach to the problem of partially tracking portfolios

It is possible to consider the TEV frontier curvature, along with other criteria already referred to in the literature (excess return, return variance, and TEV) into the utility function (6):

$$U(p) = w_0 R_p - w_1 \sigma_p^2 - w_2 TEV_p - w_3 \kappa_f \quad (6)$$

where:

$\kappa_f$  = represents the curvature of the TEV frontier, of which portfolio  $p$  forms a part.

$w_i$  = weights of each criteria, with  $i = 0..3$ .

Note that the curvature is defined on a frontier  $f$ , and not on a given portfolio  $p$ , since the curvature is the same for all portfolios on the frontier (the returns variance and the TEV are quadratic functions).

Given that in the tracking portfolios the manager fixes a value for the parameter  $G$ , all of the portfolios evaluated with utility function (6) obtain the same return  $R_p = R_b + G$ . In this way, (6) can be simplified as (7):

$$U(p) = -w_1 \sigma_p^2 - w_2 TEV_p - w_3 \kappa_f \quad (7)$$

For convenience, the proposed model will be presented as a minimization problem (8):

$$\text{Max } U(p) \equiv \text{Min } (-U(p)) \equiv \text{Min } w_1 \sigma_p^2 + w_2 TEV_p + w_3 \kappa_f \quad (8)$$

The multiobjective mathematical programming model is (9):

$$\begin{aligned} \text{Min} &= w_1 \mathbf{x}_p^t \mathbf{V} \mathbf{x}_p + w_2 \mathbf{x}^t \mathbf{V} \mathbf{x} + w_3 \kappa_f \\ \text{s. t.} & \quad \mathbf{x}^t \mathbf{R} = G \\ & \quad \mathbf{x}^t \mathbf{1} = 0 \\ & \quad \mathbf{x}_p = \mathbf{x}_b + \mathbf{x} \end{aligned} \quad (9)$$

where the only unknown element is the weightings vector  $\mathbf{x}$ . Note that no restrictions are included on the cardinality of the tracking portfolio. For the application of model (9) it is necessary to address three issues. The first relates to how to find a good solution within the exponential number of portfolios that can be formed and limiting to  $n$  the number of stocks in the tracking portfolio. The objective of model (9) is to make a comparison between these portfolios using the utility function, and not to generate a frontier. The second question to address is how to calculate  $\kappa_f$ , the only parameter that has not yet been derived analytically. Finally, there remains

the determination of the  $w_i$  weights in the utility function. Each of these questions is discussed separately in the following subsections.

#### 4.1 Search for local optima

As mentioned in the introduction, the optimal solution to the problem of partially tracking portfolios is a difficult problem from a computational point of view. All optimal local search methodologies in the literature are consistent with model (9), and it is not the aim of this paper to propose new heuristic strategies. The greatest computational burden when solving an instance of model (9) is calculating the curvature of the TEV frontier, as shown in the following paragraph. In the example developed in a later section for the tracking of the S&P 100 an adaptation of the algorithm proposed by Tabata and Takeda (1995) has been used. This algorithm was chosen because it is simple to implement and generates good local optima. The algorithm ensures that the solution found cannot be improved unless two or more stocks are changed in the tracking portfolio. For a better understanding of the overall process, we present the adaptation of the algorithm<sup>2</sup> to the multiobjective mathematical programming model (9) (Algorithm 1). Tabata and Takeda (1995) have only considered a search for a portfolio with lesser TEV given a pre-determined  $n$  default cardinality; and so it is necessary to make an adaptation to look for the other two parameters of the objective function (9): portfolio return variance and frontier curvature. Moreover, Algorithm 1 can be easily adapted to the case of a single objective. It is only necessary to place a non-zero value for one of the  $w_i$  weights and leave the rest at zero.

Algorithm 1. Adaptation of the algorithm by Tabata and Takeda (1995)

Definitions:

$VAR_p(j, i)$  = change in return variance in tracking portfolio  $p$  after substituting the  $i$ -th stock for the  $j$ -th stock.

$TEV_x(j, i)$  = change in the TEV after substituting  $i$ -th stock for the  $j$ -th stock in the portfolio with  $x$  weighting vector differences.

$\kappa_f(j, i)$  = change in the curvature of the TEV frontier after substituting the  $i$ -th stock for the  $j$ -th stock.

$F(j, i)$  = function that evaluates the change in the objective function after substituting the  $i$ -th stock for the  $j$ -th stock in the tracking portfolio. Its value is calculated as  $F(j, i) = w_1 VAR_p(j, i) + w_2 TEV_x(j, i) + w_3 \kappa_f(j, i)$ .

---

<sup>2</sup> The adaptation of the Tabata and Takeda (1995) algorithm has been programmed in R version 2.2.0. The authors will provide the code on request.

$S^{(m)}(n)$  = set of stocks included in the tracking portfolio in the m-th iteration, where n represents the cardinality of the portfolio.

Pseudocode:

Step 0.  $s:=0$ . Let  $S^{(s)}(n)$  be the initial set of stocks, n cardinality.

Step 1. If an optimal solution has not been found for the  $\mathbf{x}_n^*$  weighting vector difference of  $S^{(s)}(n)$  and for the objective function  $F_x^*$ , it can be obtained using model (9) by considering only those stocks in the  $S^{(s)}(n)$  set. Set  $j:=n+1$ .

Step 2.  $\mathbf{x}_n^* = \mathbf{x}_n^{*(s)}$ . For  $S_{j,i}^{(s)}(n)$ , ( $i = 1 \dots n$ ) calculate  $F_x^* - F_x(j, i)$ . If  $F_x^* - F_x(j, q) = \max\{F_x^* - F_x(j, i)\} > 0$ , go to step 3. Otherwise,  $j:=j+1$ . If  $j > N$  then  $j:=n+1$ ,  $i:=1$ , go to step 4.

Step 3.  $s:=s+1$ ,  $S^{(s)}(n) := S_{j,q}^{(s-1)}(n)$ . Return to step 1.

Step 4. For  $S_{j,i}^{(s)}(n)$  calculate  $\mathbf{x}_n^{*(s)}$  and its corresponding  $F_x(j, i)$ . If  $F_x^* - F_x(j, i) > 0$ , then set  $s:=s+1$ ,  $S^{(s)}(n) := S_{j,i}^{(s-1)}(n)$  and return to step 1. Otherwise, perform  $i:=i+1$ . If  $i \leq n$ , then  $j:=j+1$ . If  $j \leq n$ , set  $i:=1$  and repeat step 4. If  $j > n$ , the current solution  $S^{(s)}(n)$  and  $\mathbf{x}_n^*$  is the optimal local solution for building a tracking portfolio with n stocks: STOP.

The algorithm requires the previous determination of the cardinality  $n$  of the tracking portfolio for the total  $N$  available stocks. Once this value is set, by following the steps defined in Algorithm 1, a local optimum of the problem that considers the three criteria defined in the objective function is obtained (9).

#### 4.2 The TEV frontier curvature

As Roll (1992) demonstrated, the full tracking TEV frontier is a shift of Markowitz's minimum variance frontier, and the curvatures of both frontiers necessarily coincide (Figure 1). This section sets out various propositions, including one that shows that the curvature of the TEV frontier generated from a subset of  $n$  stocks matches the curvature of the minimum variance frontier generated from the same  $n$  stocks.

The variance of a minimum variance portfolio  $p$  can be obtained by analytically solving Markowitz's mean-variance model (10).

$$\begin{aligned} \text{Min} &= \frac{1}{2} \mathbf{x}_p^t \mathbf{V} \mathbf{x}_p \\ \text{s. a.} \quad & \mathbf{x}_p^t \mathbf{R} = R_p \\ & \mathbf{x}_p^t \mathbf{1} = 1 \end{aligned} \tag{10}$$

Following Merton (1972), we propose using the Lagrangian (11) method on this model, deriving for the vector of portfolio weights  $\mathbf{x}_p$  and multipliers  $\lambda_1$  and  $\lambda_2$ , and equating to zero. The solution to the equation system appears in the expression (12).

$$\mathcal{L} = \frac{1}{2} \mathbf{x}_p^t \mathbf{V} \mathbf{x}_p + \lambda_1 (\mathbf{x}_p^t \mathbf{R} - R_p) + \lambda_2 (\mathbf{x}_p^t \mathbf{1} - 1) \quad (11)$$

$$\mathbf{x}_p = \mathbf{V}^{-1} [\mathbf{R} \quad \mathbf{1}] \mathbf{A}^{-1} \begin{bmatrix} R_p \\ 1 \end{bmatrix} \quad (12)$$

where  $\mathbf{A} = [\mathbf{R} \quad \mathbf{1}]^t \mathbf{V}^{-1} [\mathbf{R} \quad \mathbf{1}] = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ ,  $a = \mathbf{R}^t \mathbf{V}^{-1} \mathbf{R}$ ,  $b = \mathbf{R}^t \mathbf{V}^{-1} \mathbf{1}$  and  $c = \mathbf{1}^t \mathbf{V}^{-1} \mathbf{1}$ .

We can express the variance of the  $p$  portfolio using a weights vector (12) such as  $\sigma_p^2 = \mathbf{x}_p^t \mathbf{V} \mathbf{x}_p$ , and developing its expression to arrive at a result which depends on  $a$ ,  $b$  and  $c$  (13):

$$\begin{aligned} \sigma_p^2 = \mathbf{x}_p^t \mathbf{V} \mathbf{x}_p &= [R_p \quad 1] \mathbf{A}^{-1} \begin{bmatrix} \mathbf{R} \\ \mathbf{1} \end{bmatrix} \mathbf{V}^{-1} \mathbf{V} \mathbf{V}^{-1} [\mathbf{R} \quad \mathbf{1}] \mathbf{A}^{-1} \begin{bmatrix} R_p \\ 1 \end{bmatrix} = [R_p \quad 1] \mathbf{A}^{-1} \begin{bmatrix} R_p \\ 1 \end{bmatrix} \\ &= \frac{a - 2bR_p + cR_p^2}{ac - b^2} \end{aligned} \quad (13)$$

The  $\kappa_f$  curvature of the frontier of minimum variance is obtained as the second derivative of  $\sigma_p^2$  with respect to  $R_p$  (14):

$$\kappa_f = \frac{\partial^2 \sigma_p^2}{\partial R_p^2} = \frac{2c}{ac - b^2} \quad (14)$$

This curvature matches the curvature of the TEV frontier if the tracking is full. If the tracking is partial, the curvature cannot be calculated using the expression (14), as the values of  $a$ ,  $b$  and  $c$  are linked to the full set of stocks. Nevertheless, the following proposition shows how the computation is equivalent to the curvature of the minimum variance frontier generated using the same subset of stocks.

**Proposition 1.** The curvature of the TEV frontier generated from a subset of  $n$  stocks has the same curvature as the minimum variance frontier generated from the same subset of stocks.

Demonstration: See Annex 1.

Proposition 1 characterizes the case of a partial tracking that Roll (1992) demonstrated for the full tracking. In this way, to calculate the curvature of the TEV frontier in the partial tracking we can use the expression (14) derived from Markowitz's model. Nevertheless, it should be noted that  $a = \mathbf{R}^t \mathbf{V}^{-1} \mathbf{R}$ ,  $b = \mathbf{R}^t \mathbf{V}^{-1} \mathbf{1}$  and  $c = \mathbf{1}^t \mathbf{V}^{-1} \mathbf{1}$  must be calculated whenever a

change is made in the subset of stocks in the tracking – eliminating the corresponding rows and columns of the matrix  $\mathbf{V}$  – and with the inevitable computational cost in the mean-variance models (Phillips, 2008, p.185).

**Proposition 2.** The TEV frontier generated from a subset of  $n$  stocks ( $n < N$ ) is a shift of the minimum variance frontier obtained from the same subset of stocks.

Demonstration: See Annex A.2.

Proposition 2 presents an interesting difference between full and partial tracking. In full tracking, the TEV frontier is only a shift in the axis of variance of Markowitz's frontier. Therefore, all the tracking portfolios share the same inefficiency  $k$ , which is identical to the inefficiency of the index that it replicates (Figure 1). In addition to this shift, a deviation appears in the axis of returns in the partial tracking and this causes the inefficiency in the portfolios in the tracking frontier to vary according to the required level of return. In other words, the partial tracking minimum variance frontier and the tracking frontier are not parallel. This explains why dominance in the mean-variance sense sometimes alternates between the TEV frontier obtained with the partial tracking, and the frontier obtained with the full tracking.

Another characteristic to remember about the partial tracking is the relationship between the TEV frontier curvatures that depends on how stocks are excluded from the tracking.

**Proposition 3.** The TEV frontier curvature generated from a set of  $n$  stocks is less than the curvature of the TEV frontier obtained when excluding one or more of those stocks.

Demonstration: See Annex A.3.

Proposition 3 shows that the curvature increases when cardinality of the considered tracking decreases. Accordingly, each stock that is excluded supposes a worsening in the value of the curvature, although this does not necessarily mean a decrease in efficiency in the mean-variance sense. This differs significantly from Markowitz's classical model, where the progressive exclusion of stocks necessarily implies a decrease in efficiency.

### 4.3 Criteria weighting in the multiobjective utility function

The solution of the multiobjective programming model (9) depends on the  $w_i$  weights set for each of the three parameters considered in the objective function. This section proposes a solution for objectively quantifying these parameters:

Step 1. Apply Algorithm 1 with weights  $w_1 = 1$  and  $w_2 = w_3 = 0$ . Use the resulting vector  $\mathbf{x}_n^*$  to calculate the weight of the variance criteria

$w_1^* = 1/VAR_{\mathbf{x}_n^*}$ , being  $VAR_{\mathbf{x}_n^*}$  the variance of the tracking portfolio defined by weight vector  $\mathbf{x}_n^*$ .

Step 2. Apply Algorithm 1 with weights  $w_2 = 1$  and  $w_1 = w_3 = 0$ . Use  $w_2^* = 1/TEV_{\mathbf{x}_n^*}$ .

Step 3. Apply Algorithm 1 with weights  $w_3 = 1$  and  $w_1 = w_2 = 0$ . Use  $\mathbf{x}_n^*$  vector resulting to calculate the weights of the curvature criteria:  $w_3^* = 1/\kappa_{\mathbf{x}_n^*}$ , with  $\kappa_{\mathbf{x}_n^*}$  being the curvature of the TEV frontier generated with the stocks in the tracking portfolio.

The weight of each parameter is fixed in a way that is inversely proportional to the solution – the ideal value – that is obtained when applying Algorithm 1 to the corresponding monoobjective problem. The use of ideal values in the calculation of the  $w_i$  weights is common in multiobjective programming (Ballesteros and Romero, 1991) and, more specifically, in compromise programming (Yu, 1973; Zeleny, 1973).

If the multiobjective frontier does not satisfy the requirements of the investment fund manager, then the weights defined in Steps 1-3 can be changed until a solution is found that better fits the manager's preferences.

## 5. Application of the multiobjective model to the partial tracking of the S&P100 Index

This section develops an application of the multiobjective model (9) for obtaining tracking frontiers of the S&P 100. The data set was obtained from the OR-Library (Beasley, 1990), which has been used by various researchers for comparing tracking portfolio algorithms (Beasley *et al.*, 2003; Ruiz-Torrubiano and Suarez, 2009). The data includes the weekly returns of the index and 98 of its represented stocks during the period 1992-1997. Although the data is not recent, it remains equally valid for illustrating our proposal.

Before obtaining the tracking frontiers, two issues must be resolved prior to the implementation of the multiobjective model: first, the excess return required; and second, the number of stocks in the portfolio. For the first case, the possibility of allowing for negative returns on the underlying index ( $G < 0$ ) was dismissed, as this would assume that the investor is willing to receive a return below the index. We have conservatively assumed that the investor is content with the same return as the index ( $G = 0$ ). With respect to the cardinality of the portfolio, the results are presented considering 5, 7, 10, and 15 stocks. In this way, the robustness of the model can be tested for different sizes of portfolio, and the differences that unfold when increasing the  $n$  number of stocks can be observed.

The adapted Tabata and Takeda (1995) (Section 4.1) algorithm was used for the selection of the tracking portfolio stocks. The solution proposed in section 4.3 was used for the  $w_i$  weights. However, the frontiers obtained

were relatively close to the TEV frontier. Accordingly, the criterion of the return variance was over-weighted. Specifically, the weight was multiplied by  $\sqrt{n}$ . The square root of  $n$  was used because it is a function with a negative second derivative.

Following the implementation of Algorithm 1 as proposed in Section 4.1 various portfolios are generated according to the cardinality imposed. Table 1 shows the composition of the portfolios for the multiobjective case and the three monoobjective possibilities: minimize the variance of the tracking portfolio – optimization in the sense of Markowitz ( $w_1 = 1, w_2 = w_3 = 0$ ); minimize the TEV ( $w_2 = 1, w_1 = w_3 = 0$ ); and minimize the curvature of the tracking frontier ( $w_3 = 1, w_1 = w_2 = 0$ ). With the minimization of the variance, the portfolio with minimum variance and identical return to the index is obtained ( $G = 0$ ). Using these stocks it is possible to generate a frontier of minimum variance by changing the required return – following Markowitz’s classic mean-variance model. With the minimization of TEV the model selects the stocks that also produce the minimum TEV for the case  $G = 0$ , and with these same stocks the corresponding TEV frontier is also generated. Finally, in the model for minimizing the curvature, stocks are selected that minimize this expression and consider excessive returns to be null in the same way. In all cases, the number of stocks in the portfolio was limited to  $n$ .

Table 1 demonstrates how the composition of the portfolios varies as cardinality increases. Together with the stocks, the ratio between two numbers appears in brackets. The first is the number of stocks that are repeated with respect to the portfolio with immediately inferior cardinality. The second number is the cardinality. For example, for the multiobjective model with  $n = 15$ , there are 7 stocks that are repeated in the multiobjective portfolio with  $n = 10$ . Specifically, these are stocks are 05, 13, 33, 53, 57, 65 and 81. Therefore, the portfolio with  $n = 15$  has inherited 7 of the 10 stocks that made up the portfolio with  $n = 10$ , and so the ratio is 7/10. This offers an idea of the persistence with which stocks are held when cardinality increases.

The results show that the mean-variance monoobjective model is the most persistent in its stocks. The portfolio with  $n = 7$  selects 4 out of 5 stocks from  $n = 5$ ; and for  $n = 10$  it is 6 of the 7 possible stocks; while  $n = 15$  inherits 10 of the possible stocks in the  $n = 10$  portfolio. Of the four models suggested, the model that generates the most variable portfolios is the one that minimizes TEV.

[Insert Table 1 here]



It is interesting to analyze the graphical representation of the mean-variance and mean-TEV. Figure 2 shows the frontiers obtained for each model on the mean-variance plane according to the cardinalities considered in each case. The frontiers are generated from the stocks shown in Table 1 by simply varying the excess return required. For example, in the case  $n = 5$ , the frontier that minimizes the variance of the portfolio corresponds to Markowitz's classical model when only considering the stocks 33, 38, 52, 57 and 65. These stocks correspond to the minimum variance portfolio for  $G = 0$ , and so the portfolio at this point is less volatile than the other frontiers at the same  $G = 0$  point.

For nearly the entire spectrum of  $G$  values considered in the graph, the minimum variance frontier dominates the two frontiers generated with the monoobjective models: the TEV frontier and the frontier curvature. However, this does not happen with the multiobjective model frontier. For example, in the case  $n = 5$  it can be seen how the minimum variance frontier dominates the multiobjective frontier for weekly returns of between 0.22% and 0.37% (annual returns of 12.1% and 21.2% respectively). This implies that if the investor wants to achieve a return that is within this range, then he/she should select the minimum variance frontier. But if the investor will accept returns below 0.22%, or if the investor's risk profile requires returns greater than 0.37%, then the multiobjective frontier should be chosen.

The greater curvature of the minimum variance frontier implies that the distance between it and the multiobjective frontier grows rapidly when  $|G|$  increases. For example, if an investor wants a weekly return of 0.45%, the risk of his position on the minimum variance frontier would be 0.00046 when measured as the variance of return. The investor who chooses the multiobjective frontier would assume a variance of 0.00031. In other words, the variance recorded at the minimum variance frontier would be 50% higher than the variance in the multiobjective frontier.

Moreover, the greatest distance between the minimum variance frontier and the multiobjective frontier occurs at the weekly return point of 0.29%. The difference between the variances is 0.000047, while the relative advantage of investing in the minimum variance frontier instead of the multiobjective frontier is 24%.

Similar comments can be made for the remaining cardinalities. Figure 2 shows that as the cardinality of the portfolios increases, the frontier curvature decreases. This means the effect of including the curvature in the multiobjective model is dissipated, because the curvature of the minimum variance frontier is approaching the minimum curvature frontier. The range of returns in which the minimum variance frontier dominates the multiobjective frontier grows, albeit slowly. The difference between the two frontiers also decreases as cardinality increases. Therefore, the S&P100

can be efficiently tracked with 15 stocks with good results in the mean-variance plane. Adding more stocks to the tracking would not generate an improvement beyond that observed in Figure 2.

Figure 3 shows the frontiers in the mean-TEV plane when considering the same cardinalities as in Figure 2. It can be seen that the frontier that minimizes the TEV approaches the position of the index as cardinality in the portfolio increases. This frontier is preferred in the case of  $n = 15$  as it dominates the remaining frontiers in all the considered return rates. Something similar occurs with  $n = 10$ . However, if the number of stocks in the portfolio is restricted to 5, then the excessive curvature means that the multiobjective frontier dominates when returns are below 0.26% or are greater than 0.37%. This relationship of dominance only becomes clear in the case  $n = 5$  due to the already mentioned overweighting of the return variance in the multiobjective function (it has been multiplied by  $\sqrt{n}$ ). Similarly, the weight of the criteria can be varied so the multiobjective function tilts towards one in particular, depending on the strategy defined by the fund manager.

In any case, we can conclude that the multiobjective model finds consensus solutions among the monoobjective models. The model may therefore be a good alternative for managers who do not aim to optimize a particular criterion, but wish to offer clients a balanced solution and so satisfy a wider range of risk profiles.

[Insert Figure 2 here]

[Insert Figure 3 here]

## 6. Conclusions

Criticisms made about active investment fund management have boosted the success of passive strategies. Various studies have shown how difficult it is for active management to beat the results of passive management – even more so when transaction costs are considered. Tracking portfolios have become one of the most common passive management strategies – and the emergence of ETFs has heightened their popularity.

Many authors have suggested that costs can be reduced by employing heuristics for the partial tracking of portfolios. Unlike full tracking, partial tracking portfolios use only a subset of the stocks in the index. Researchers have made use of a limited number of parameters in the selection of these stocks: Tracking error variance (TEV) if the only objective is to imitate the behavior of an index; and return variance if the efficiency of the portfolio at the mean-variance plane is also under consideration. Both criteria are linked to the tracking portfolio, so the portfolio composition varies with the

level of return required. Accordingly, different returns can mean that different stocks are considered in the tracking. This represents an increase in transaction costs because fund managers who wish to satisfy clients with heterogeneous profiles are then forced to invest in many stocks. This practice reduces the advantages of passive management in comparison to active management.

This paper considers a new parameter for use with the above: Frontier curvature. This criterion is not defined for a given portfolio, but for the set of portfolios that define the tracking frontier. The main implication is that the manager can satisfy different investment profiles using the same subset of stocks, with all the portfolios containing the same stocks and so reducing transaction costs.

For the joint consideration of these criteria we propose the use of multiobjective mathematical programming. In this way the solution can generate a new frontier as a consensus between the frontiers obtained by separately considering each criterion.

The proposed model has been used for tracking the S&P 100. The results show how the multiobjective frontier is balanced between monoobjective frontiers. From a theoretical viewpoint, the generation of multiobjective solutions is justified for partial tracking portfolios for several reasons.

First, if only the TEV criterion is considered then naive solutions could be obtained in many cases, meaning solutions dominated by stocks with the highest market capitalizations. In such situations, the application of heuristics for building tracking portfolios would not offer a significant advantage with respect to a naive strategy of selecting stocks on the basis of market capitalization. This would occur mainly with composite indexes built from a relatively small number of stocks – such as the French CAC-40, the German DAX-30, or the Spanish IBEX-35.

Second, if only the variance of portfolio returns is considered, then portfolios would be obtained whose future behavior would not necessarily correspond with past behavior. This is one of the main problems with the mean-variance model in which returns and the covariance structure among stocks changes over time – negatively affecting the predictive ability of models. This does not occur with TEV models, where the recent history of stocks satisfactorily explains the evolution of the index. Moreover, these models tend to retain their explanatory power in the future. The reason is simple: there are many stocks that maintain their influence and weight in the composition of the index because of their substantial market capitalizations.

Finally, the inclusion of the curvature of the tracking frontier as a new criterion enables us to contemplate a wider range of investment profiles. With this criterion, it is possible to go beyond the objective of building a single tracking portfolio and to aim for a more general goal: to obtain a

tracking frontier that satisfies a larger number of investors by using the same subset of stocks.

## Annexes

### Annex A.1

Proposition 1. The curve of the TEV frontier generated from a subset of  $n$  stocks ( $n < N$ ) has the same curvature as a minimum variance frontier generated from the same subset of stocks.

This is demonstrated when  $n = N - 1$ , that is, the number of stocks is reduced by one.

Let us calculate the curve of a TEV frontier for a tracking portfolio built from  $N - 1$  stocks. Without loss of generality, let us suppose that the excluded  $\sigma_{TEV-i}^2$ , can be expressed as:

$$\sigma_{TEV-i}^2 = \mathbf{x}_p^t \mathbf{V} \mathbf{x}_p = (\mathbf{x}_b + \mathbf{x})^t \mathbf{V} (\mathbf{x}_b + \mathbf{x}) = \mathbf{x}_b^t \mathbf{V} \mathbf{x}_b + \mathbf{x}_{-i}^t \mathbf{V} \mathbf{x}_{-i} + 2\mathbf{x}_b^t \mathbf{V} \mathbf{x}_{-i} \quad (\text{A.1})$$

where:

$\mathbf{x}_{-i}$  = vector difference of weights between the index and the portfolio, excluding the portfolio's  $i$ -stock.

The vector  $\mathbf{x}_{-i}$  can be obtained by solving the following model:

$$\begin{aligned} \text{Min} \quad & \frac{1}{2} \mathbf{x}_{-i}^t \mathbf{V} \mathbf{x}_{-i} \\ \text{s. t.} \quad & \mathbf{x}_{-i}^t \mathbf{R} = G \\ & \mathbf{x}_{-i}^t \mathbf{1} = 0 \\ & \mathbf{x}_{-i}^t \mathbf{0}_i = -q_{b_i} \end{aligned} \quad (\text{A.2})$$

where:

$\mathbf{0}_i$  = is the weights vector sized  $N \times 1$  with all zero values except in the  $i$ -th position where it has a value of one.

$q_{b_i}$  = represents the weight of the  $i$ -th stock in the stock index.

By using the Lagrangian method for this model and solving the system, the solution for  $\mathbf{x}_{-i}$  is (A.3):

$$\mathbf{x}_{-i} = \mathbf{V}^{-1} [\mathbf{R} \quad \mathbf{1} \quad \mathbf{0}_i] \mathbf{A}_{-i}^{-1} \begin{bmatrix} G \\ 0 \\ -q_{b_i} \end{bmatrix} \quad (\text{A.3})$$

Below we will develop each of the terms in (A.1). Expressing the second term  $\mathbf{x}_{-i}^t \mathbf{V} \mathbf{x}_{-i}$  (A.4):

$$\begin{aligned}
\mathbf{x}_{-i}^t \mathbf{V} \mathbf{x}_{-i} &= [G \quad 0 \quad -q_{b_i}] \mathbf{A}_{-i}^{-1} \begin{bmatrix} \mathbf{R} \\ \mathbf{1} \\ \mathbf{0}_i \end{bmatrix} \mathbf{V}^{-1} \mathbf{V} \mathbf{V}^{-1} [\mathbf{R} \quad \mathbf{1} \quad \mathbf{0}_i] \mathbf{A}_{-i}^{-1} \begin{bmatrix} G \\ 0 \\ -q_{b_i} \end{bmatrix} = \\
&= [G \quad 0 \quad -q_{b_i}] \mathbf{A}_{-i}^{-1} \begin{bmatrix} G \\ 0 \\ -q_{b_i} \end{bmatrix} = \\
&= \frac{G^2(cf - e^2) + q_{b_i}^2(ac - b^2) - 2q_{b_i}G(be - cd)}{acf + 2bde - ae^2 - b^2f - cd^2}
\end{aligned} \tag{A.4}$$

Proceeding in the same way with the third variance term produces the expression (A.5):

$$\begin{aligned}
2\mathbf{x}_b^t \mathbf{V} \mathbf{x}_{-i} &= 2\mathbf{x}_b^t \mathbf{V} \mathbf{V}^{-1} [\mathbf{R} \quad \mathbf{1} \quad \mathbf{0}_i] \mathbf{A}_{-i}^{-1} \begin{bmatrix} G \\ 0 \\ -q_{b_i} \end{bmatrix} = 2[\mathbf{R}_b \quad \mathbf{1} \quad q_{b_i}] \mathbf{A}_{-i}^{-1} \begin{bmatrix} G \\ 0 \\ -q_{b_i} \end{bmatrix} = \\
&= 2 \frac{G(\mathbf{R}_b(cf - e^2) + q_{b_i}(be - cd) + de - bf) - q_{b_i}(\mathbf{R}_b(be - cd) + q_{b_i}(ac - b^2) + bd - ae)}{acf + 2bde - ae^2 - b^2f - cd^2}
\end{aligned} \tag{A.5}$$

By adding the three variance terms we obtain (A.6):

$$\begin{aligned}
\sigma_{TEV-i}^2 &= \sigma_b^2 + \frac{G^2(cf - e^2) + q_{b_i}^2(ac - b^2) - 2q_{b_i}G(be - cd)}{acf + 2bde - ae^2 - b^2f - cd^2} \\
&+ 2 \frac{G(\mathbf{R}_b(cf - e^2) + q_{b_i}(be - cd) + de - bf) - q_{b_i}(\mathbf{R}_b(be - cd) + q_{b_i}(ac - b^2) + bd - ae)}{acf + 2bde - ae^2 - b^2f - cd^2}
\end{aligned} \tag{A.6}$$

The second derivative with respect to  $G$  provides the frontier curve that we will term  $\kappa_{f1}$ :

$$\frac{\partial^2 \sigma_{TEV-i}^2}{\partial G^2} = \frac{2(cf - e^2)}{acf + 2bde - ae^2 - b^2f - cd^2} = \kappa_{f1} \tag{A.7}$$

To calculate the frontier curve of the minimum variance considering the same  $N - 1$  stocks, the next mathematical programming model can be used and the  $\mathbf{x}_{p/\{i\}}$  weight vector solved:

$$\begin{aligned}
&Min \quad \mathbf{x}_{p/\{i\}}^t \mathbf{V} \mathbf{x}_{p/\{i\}} \\
&s. t. \quad \mathbf{x}_{p/\{i\}}^t \mathbf{R} = R_p \\
&\quad \quad \mathbf{x}_{p/\{i\}}^t \mathbf{1} = 1 \\
&\quad \quad \mathbf{x}_{p/\{i\}}^t \mathbf{0}_i = 0
\end{aligned} \tag{A.8}$$

This is Markowitz's classic minimum variance model where only the constraint that the  $i$ -th weight must be zero has been added. The solution appears in (A.9):

$$\mathbf{x}_{p/\{i\}} = \mathbf{V}^{-1}[\mathbf{R} \ \mathbf{1} \ \mathbf{0}_i] \mathbf{A}_{-i}^{-1} \begin{bmatrix} R_p \\ 1 \\ 0 \end{bmatrix} \quad (\text{A.9})$$

where  $\mathbf{A}_{-i} = \begin{bmatrix} \mathbf{R}^t \mathbf{V}^{-1} \mathbf{R} & \mathbf{R}^t \mathbf{V}^{-1} \mathbf{1} & \mathbf{R}^t \mathbf{V}^{-1} \mathbf{0}_i \\ \mathbf{R}^t \mathbf{V}^{-1} \mathbf{1} & \mathbf{1}^t \mathbf{V}^{-1} \mathbf{1} & \mathbf{1}^t \mathbf{V}^{-1} \mathbf{0}_i \\ \mathbf{R}^t \mathbf{V}^{-1} \mathbf{0}_i & \mathbf{1}^t \mathbf{V}^{-1} \mathbf{0}_i & \mathbf{0}_i^t \mathbf{V}^{-1} \mathbf{0}_i \end{bmatrix}$ .

The variance of the portfolio with  $\mathbf{x}_{p/\{i\}}$  weights are calculated in (A.10):

$$\begin{aligned} \sigma_{p/\{i\}}^2 &= \mathbf{x}_{p/\{i\}}^t \mathbf{V} \mathbf{x}_{p/\{i\}} = [R_p \ 1 \ 0] \mathbf{A}_{-i}^{-1} \begin{bmatrix} \mathbf{R} \\ \mathbf{1} \\ \mathbf{0}_i \end{bmatrix} \mathbf{V}^{-1} \mathbf{V} \mathbf{V}^{-1} [\mathbf{R} \ \mathbf{1} \ \mathbf{0}_i] \mathbf{A}_{-i}^{-1} \begin{bmatrix} R_p \\ 1 \\ 0 \end{bmatrix} \\ &= [R_p \ 1 \ 0] \mathbf{A}_{-i}^{-1} \begin{bmatrix} R_p \\ 1 \\ 0 \end{bmatrix} \\ &= \frac{R_p^2 (cf - e^2) + 2R_p (de - bf) + af - d^2}{acf + 2bde - ae^2 - b^2f - cd^2} \end{aligned} \quad (\text{A.10})$$

Bearing in mind that the returns of the portfolio can be expressed as the returns of the stock index plus return relative to  $G$ :

$$\sigma_{p/\{i\}}^2 = \frac{(R_b + G)^2 (cf - e^2) + 2(R_b + G)(de - bf) + af - d^2}{acf + 2bde - ae^2 - b^2f - cd^2} \quad (\text{A.11})$$

The second derivative of  $\sigma_{p/\{i\}}^2$  with respect to  $G$  provides the curvature of the minimum variance frontier:

$$\frac{\partial^2 \sigma_{p/\{i\}}^2}{\partial G^2} = \frac{2(cf - e^2)}{acf + 2bde - ae^2 - b^2f - cd^2} \quad (\text{A.12})$$

This coincides with the  $\kappa_{f1}$  curvature of the TEV frontier.

In this way, the coincidence in the curvature of the efficient frontier built from a subset of  $n$  stocks and the TEV frontier obtained from the same  $n$  stocks is demonstrated. The generalization for  $2 \leq n < N - 1$  is immediate and so the proposition is demonstrated.

## Annex A.2

Proposition 2. The TEV frontier generated from a subset of  $n$  stocks ( $n < N$ ) is a shift of the minimum variance frontier obtained from the same subset of stocks.

This is firstly demonstrated for the case of a subset composed of  $N-1$  stocks. The extension to the general case is immediate.

With respect to the expression (A.6) of the variance of the TEV frontier, the slope of the curve for a given  $R_p$  is given by (A.13):

$$\frac{\partial \sigma_{TEV-i}^2}{\partial R_p} = \frac{2(R_p + G)(cf - e^2) - 2q_{b_i}(be - cd) + 2(de - bf)}{acf + 2bde - ae^2 - b^2f - cd^2} \quad (\text{A.13})$$

If we calculate the slope of the minimum variance frontier at a point with the same return we get (A.14):

$$\frac{\partial \sigma_{p/\{i\}}^2}{\partial R_p} = \frac{2R_p(cf - e^2) + 2(de - bf)}{acf + 2bde - ae^2 - b^2f - cd^2} \quad (\text{A.14})$$

The two curves only coincide under the improbable condition:

$$G(cf - e^2) = q_{b_i}(be - cd) \quad (\text{A.15})$$

Therefore, in the partial tracking of the TEV frontier, the axis of variance is displaced for the case of the full tracking, as already shown by Roll (1992) for full tracking. The slope of variance is also displaced for returns – since they have different slopes at a point with the same return. The coincidence in the slopes, which is the same as a non-shift in the returns axis, only occurs under condition (A.15).



### Annex A.3

Proposition 3. The curvature of a TEV frontier generated from a set of  $n$  stocks ( $n < N$ ) is less than the curvature of the TEV frontier generated excluding some of these stocks.

This is demonstrated by comparing the curvatures of a full tracking and a partial tracking after removing one stock. The generalization of the statement in the proposition is immediate.

The curvature of the TEV frontier in the full tracking coincides with the curvature of the minimum variance frontier (Roll, 1992), and so its expression coincides with (17), and is termed  $\kappa_{f2}$ :

$$\kappa_{f2} = \frac{2c}{ac - b^2} \quad (\text{A.16})$$

To calculate the curvature of the TEV frontier in the partial tracking we will suppose without loss of generality that the excluded stock occupies the  $i$ -th position. In this way, the curvature of the frontier takes the value  $\kappa_{f1}$  (A.7), which for convenience we will reproduce below:

$$\kappa_{f1} = \frac{2(cf - e^2)}{acf + 2bde - ae^2 - b^2f - cd^2} \quad (\text{A.17})$$

The aim is to compare the relationship  $\kappa_{f1} \geq \kappa_{f2}$ :

$$\begin{aligned} \frac{2(cf - e^2)}{acf + 2bde - ae^2 - b^2f - cd^2} &\geq \frac{2c}{ac - b^2} \\ ac^2f - ace^2 - b^2cf + b^2e^2 &\geq ac^2f + 2bcde - ace^2 - b^2cf - c^2d^2 \\ c^2d^2 + b^2e^2 - 2bcde &\geq 0 \\ (cd - be)^2 &\geq 0 \end{aligned} \quad (\text{A.18})$$

Given that any squared scale is greater than or equal to 0, it is shown that  $\kappa_{f1} \geq \kappa_{f2}$ .

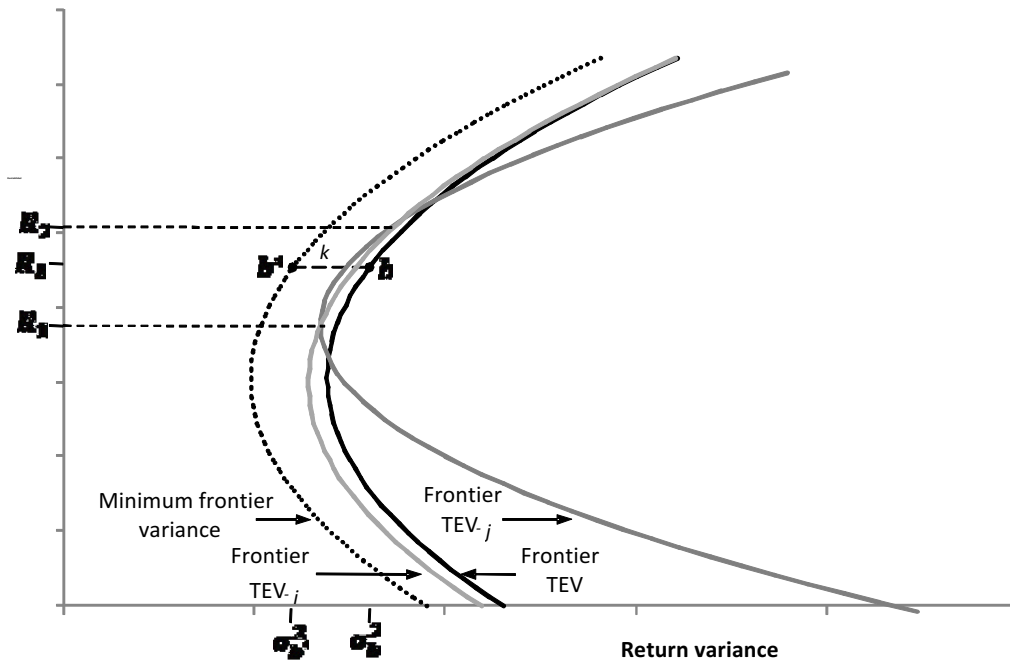
Because of similarity with the minimum variance frontier, it will be equal if the variance-covariance matrix is not invertible. Therefore, necessarily  $\kappa_{f1} > \kappa_{f2}$ .

## References

- Adcock, C.J.; Meade, N. 1994. A simple algorithm to incorporate transaction costs in quadratic optimization, *European Journal of Operational Research* 79: 85-94.
- Ballester, E.; Romero, C. 1991. A theorem connecting utility function optimization and compromise programming, *Operations Research Letters* 10: 421-427.
- Beasley, J. E. 1990. OR-Library: distributing test problems by electronic mail, *Journal of the Operational Research Society* 41(11): 1069-1072.
- Beasley, J.E.; Meade, N.; Chang, T.J. 2003. An evolutionary heuristic for the index tracking problem, *European Journal of Operational Research* 148: 621-643.
- Burmeister, C.; Mausser, H.; Mendoza, R. 2005. Actively managing tracking error, *Journal of Asset Management* 5(6): 410-422.
- Canakgoz, N.A.; Beasley, J.E. 2008. Mixed-integer programming approaches for index tracking and enhanced indexation, *European Journal of Operational Research* 196: 384-399.
- Carhart, M.M. 1997. On persistence in mutual fund performance, *Journal of Finance* 52: 57-82.
- Chow, G. 1995. Portfolio selection based on return, risk, and relative performance, *Financial Analysts Journal* Mar-Apr: 54-60.
- Coleman, T.F.; Li, Y.; Henniger, J. 2006. Minimizing tracking error while restricting the number of assets, *Journal of Risk* 8: 33-56.
- Connor, G.; Leland, H. 1995. Cash management for index tracking, *Financial Analysts Journal* Nov-Dec: 75-80.
- Corielli, F.; Marcellino, M. 2006. Factor based index tracking, *Journal of Banking & Finance* 30(8): 2215-2233.
- Davis, J. 2001. Mutual fund performance and manager style, *Financial Analysts Journal* 57: 19-26.
- Derigs, U.; Nickel, N.H. 2004. On a local-search heuristic for a class of tracking error minimization problems in portfolio management, *Annals of Operations Research* 131: 45-77.
- Dose, C.; Cincotti, S. 2005. Clustering of financial time series with application to index and enhanced index tracking portfolio, *Physica A* 335: 145-151.
- Edelen, R. 1999. Investor flows and the assessed performance of Open-end Mutual Funds. A comparison of benchmarks and benchmark comparisons, *Journal of Financial Economics* 53(3): 439-466.
- Elton, E.; Gruber, M.; Das, J.; Hlavka, M. 1993. Efficiency with costly information: A reinterpretation of the evidence for managed portfolios., *Review of Financial Studies* 6: 1-22.
- Focardi, S.M.; Fabozzi, F.J. 2004. A methodology for index tracking based on time-series clustering, *Quantitative Finance* 4: 417-425.
- Frino, A.; Gallagher, D.R. 2001. Tracking S&P500 index funds, *Journal of Portfolio Management* 28(1): 44-55.
- Gaivoronovski, A.A.; Krylov, S.; van der Wijst, N. 2005. Optimal portfolio selection and dynamic benchmark tracking, *European Journal of Operational Research* 163: 115-131.
- Gruber, M. 1996. Another puzzle: The growth in actively managed mutual funds, *The Journal of Finance* 51(3): 783-810.
- Hallerbach, W.G.; Spronk, J. 2002. The relevance of MCDM for financial decisions, *Journal of Multi-Criteria Decision Analysis* 11: 187-195.
- Jensen, M.C. 1968. The performance of mutual funds in the period 1945-1964, *Journal of Finance* 23(2): 389-416.
- Jorion, P. 2003. Portfolio optimization with tracking-error constraints, *Financial Analysts Journal* Sep-Oct: 70-82.

- Malkiel, B.G. 1995. Returns from investing in Equity Mutual Funds: 1971 to 1991, *Journal of Finance* 50(2): 549-572.
- Malkiel, B.G. 2003. Passive investment strategies and efficient markets, *European Financial Management* 9: 1-10.
- Malkiel, B.G.; Radisich, A. 2001. The Growth of Index Funds and the Pricing of Equity Stocks, *Journal of Portfolio Management* 27(2): 9-21.
- Markowitz, H.M. 1952. Portfolio Selection, *Journal of Finance* 7: 77-91.
- Markowitz, H.M. 1959. *Portfolio Selection: Efficient Diversification of Investments*, Wiley.
- Meade, N.; Salkin, G.R. 1990. Developing and maintaining an equity index fund, *Journal of the Operational Research Society* 41(7): 599-607.
- Merton, R.C. 1972. An analytic derivation of the efficient portfolio frontier, *The Journal of Financial and Quantitative Analysis* 7(4): 1851-1872.
- Phillips, H.E. 2008. *Deterministic Portfolio Selection Models, Selection Bias, and an Unlikely Hero*, in Lee C-F. (Ed.) *Advances in Quantitative Analysis of Finance and Accounting*, Vol. 6: 179-203. World Scientific.
- Roll, R. 1992. A mean/variance analysis of tracking error, *The Journal of Portfolio Management* 18: 13-22.
- Rudolf, M.; Wolter, H.-J.; Zimmermann, H. 1999. A linear model for tracking error minimization, *Journal of Banking & Finance* 23: 85-103.
- Ruiz-Torrobiano, R.; Suárez, A. 2009. A hybrid optimization approach to index tracking, *Annals of Operations Research* 166: 57-71.
- Steuer, R.E. 1986. *Multiple Criteria Optimization. Theory Computation and Application*, Wiley.
- Tabata, Y.; Takeda, E. 1995. Bicriteria optimization problem of designing an index fund, *Journal of the Operational Research Society* 46: 1023-1032.
- Wu, L.C.; Chou, S.C.; Yang, C.C.; Ong, C.S 2007. Enhanced index investing based on goal programming, *The Journal of Portfolio Management* Spring: 49-56.
- Yu, P.L. 1973. A class of solutions for group decision problems, *Management Science* 19: 936-946.
- Zeleny, M. 1973. Compromise Programming, in J.L. Cochrane and M. Zeleny (Eds.), *Multiple Criteria Decision Making*, University of South Carolina Press. 262-301.
- Zeleny, M. 1974. *Linear Multiobjective Programming*, Springer.
- Zeleny, M. 1982. *Multiple Criteria Decision Making*, McGraw-Hill.
- Zopounidis, C. 1998. *Operational Tools in the Management of Financial Risks*, Kluwer.

Figure 1. The minimum variance frontier and various TEV frontiers



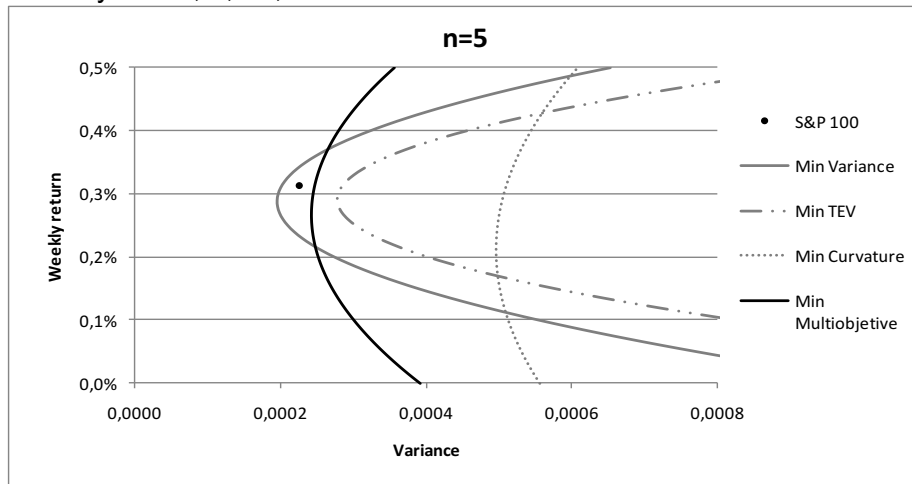
Key: - - - Minimum variance frontier; — TEV frontier; — TEV frontier excluding the  $j$ -th stock; — TEV frontier excluding the  $i$ -th stock;  $b$ : position of the index in the mean-variance plane;  $b^*$ : projection of the index on the minimum variance frontier;  $R_1$ : return of portfolio 1 (see Roll, 1992);  $R_2$ : return of portfolio 2 (see Roll, 1992);  $R_b$ : index return;  $\sigma_b^2$ : index return variance;  $\sigma_{b^*}^2$ : return variance of portfolio  $b^*$ .

Table 1. Composition of the portfolios in the solution of the multiobjective model and the three monoobjective models.

Cardinality of the portfolio	Monoobjective models	Multiobjective models		
	$w_1 = \sqrt{n}/VAR_{x_n^*}$ , $w_2 = 1/TEV_{x_n^*}$ , $w_3 = 1/C_{x_n^*}$	Min portfolio variance - Markowitz ( $w_1 = 1$ , $w_2 = w_3 = 0$ )	Min tracking error variance – TEV ( $w_2 = 1$ , $w_1 = w_3 = 0$ )	Min tracking frontier curvature ( $w_3 = 1$ , $w_1 = w_2 = 0$ )
$n = 5$	Stock06, Stock15, Stock 38, Stock60, Stock89	Stock33, Stock38, Stock 52, Stock57, Stock65	Stock05, Stock17, Stock 59, Stock79, Stock90	Stock01, Stock15, Stock 51, Stock84, Stock89
$n = 7$	Stock05, Stock19, Stock33, Stock50, Stock52, Stock53,	Stock13, Stock33, Stock38, Stock 53, Stock57, Stock65,	Stock05, Stock17, Stock26, Stock49, Stock 53, Stock74,	Stock08, Stock34, Stock50, Stock 51, Stock60,

	Stock65 (0/5)	Stock80 (4/5)	Stock90 (3/5)	Stock68, Stock89 (2/5)
$n = 10$	Stock05, Stock13, Stock19, Stock 33, Stock50, Stock52, Stock53, Stock57, Stock65, Stock81 (7/7)	Stock33, Stock38, Stock 52, Stock 53, Stock57, Stock61, Stock65, Stock75, Stock80, Stock97 (6/7)	Stock05, Stock12, Stock18, Stock19, Stock33, Stock43, Stock 53, Stock 59, Stock 65, Stock74 (3/7))	Stock08, Stock15, Stock34, Stock50, Stock 51, Stock60, Stock68, Stock75, Stock84, Stock89 (7/7)
$n = 15$	Stock05, Stock08, Stock13, Stock 33, Stock49, Stock53, Stock57, Stock65, Stock74, Stock80, Stock81, Stock84, Stock89, Stock90, Stock97 (7/10)	Stock08, Stock13, Stock33, Stock38, Stock 52, Stock 53, Stock57, Stock61, Stock65, Stock75, Stock80, Stock81, Stock83, Stock89, Stock97 (10/10)	Stock05, Stock14, Stock15, Stock18, Stock35, Stock38, Stock43, Stock47, Stock49, Stock 53, Stock 56, Stock 59, Stock 63, Stock 66, Stock96 (5/10)	Stock05, Stock08, Stock15, Stock16, Stock22, Stock25, Stock28, Stock33, Stock34, Stock53, Stock75, Stock78, Stock84, Stock89, Stock98 (6/10)

Figure 2. Graphical representation of the return variance versus the weekly returns for the multiobjective model and the three monoobjective models. Cardinality:  $n = 5, 7, 10, 15$



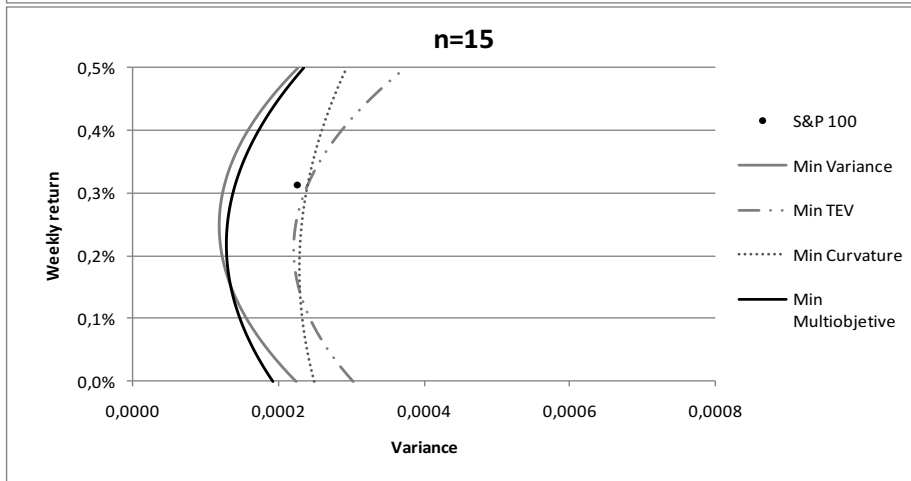
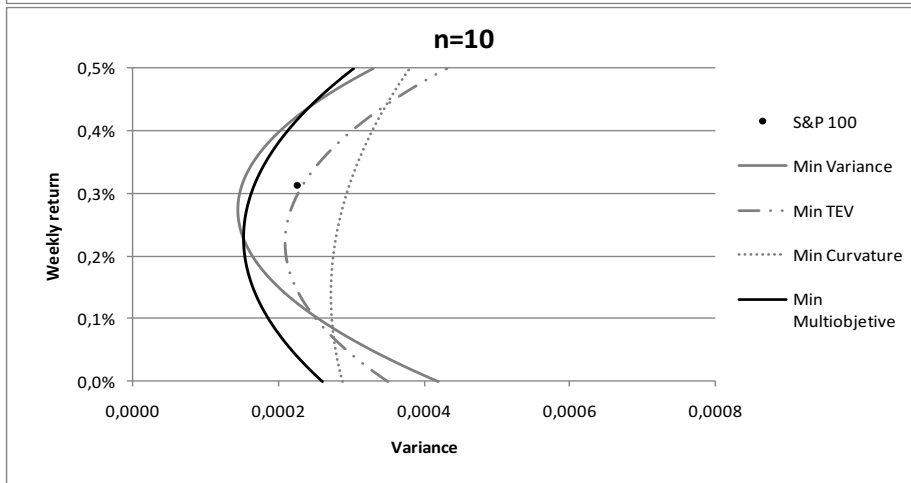
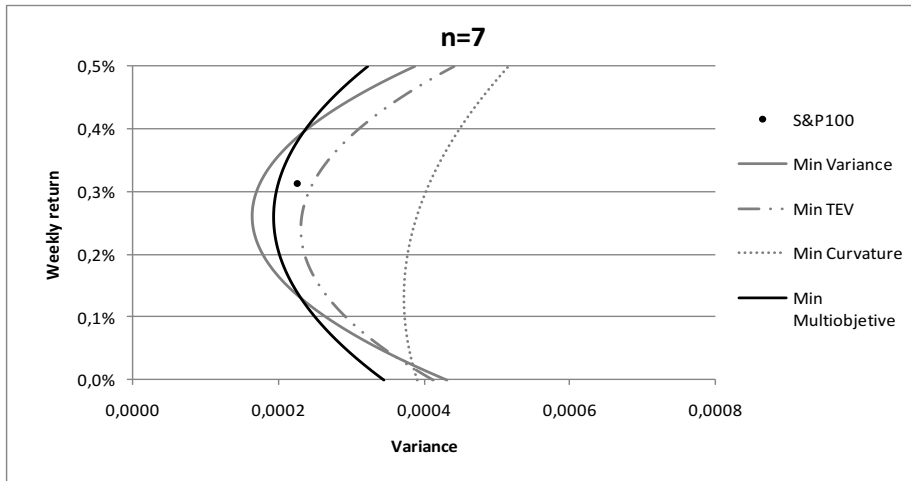


Figure 3. Graphical representation of the TEV versus the weekly returns for the multiobjective model and the three monoobjective models. Cardinality:  $n = 5, 7, 10, 15$

