Sparse Index Tracking: An $L_{1/2}$ Regularization Based Model and Solution

Xu Fengmin [*] Zongben Xu [†] Honggang Xue[‡]

Abstract

In recent years, sparse modeling has attracted extensive attention and successfully applied in mean-variance portfolio selection for promoting out-of-sample properties and decreasing transaction costs. The existing applications are, however, L_1 regularization based, which is invalid for the index tracking with budget and no-short selling constraints, and ineffective sometimes in promotion of sparsity and selection of regularization parameter. In this paper we propose an $L_{1/2}$ regularization based sparse index tracking model to cater for the index tracking with budget and no-short selling constraints. We formalize a hybrid half thresholding algorithm for fast solution of the model, by applying the half thresholding representation theory in $L_{1/2}$ regularization paradigm. The proposed model and algorithm are empirically tested on eight data sets from OR-library. The tests show that the new model can effectively solve the index tracking problems, and the resulted tracking portfolio is of higher sparsity, lower out-of-sample prediction error and higher consistency of in-sample and out-of-sample performance, as compared with other competitive methods. The algorithm suggested is parameter-free and allows the regularization parameter to be set adaptively in a nearly optimal way whenever the number of constituents of expected portfolio is fixed, the algorithm therefore provides a very convenient and fast solver for the index tracking, especially for the large scale problems.

Keywords: Index Tracking, $L_{1/2}$ regularization, Thresholding Algorithm.

---

[*]Department of Computational Mathematics, Xi'an Jiaotong University, P. R. China. E-mail: fengminxu@mail.xjtu.edu.cn. This author's work is supported partly by the NSFC project (10231060) and Chinese Postdoctoral Fund.

[†]Corresponding author, Institute for Information and system Science and MOE Key Lab for Intelligent Networks and Network Security, Xi'an Jiaotong University, P. R. China. E-mail: zbxu@mail.xjtu.edu.cn. This author's work is supported partly by the National 973 Project of China (2007CB311002)and NSFC project (61075054).

[‡]School of Economics and Finance, Xi'an Jiaotong University, Xi'an, 710049, China. E-mail: xhg@mail.xjtu.edu.cn. This author's work is supported by the Ministry of Education Humanities Social Science (No.09XJAZH005), the Ministry of Education New Century Elitist Supports Plan (No.NCET-10-0646).

1 Introduction

Stock index derivatives, such as index funds, index futures and index options, have developed very rapidly and become important tools in investment and risk management of global financial markets, especially in stabilizing the stock market in global fiance crisis. Index tracking (e.g., index replication) is the basis of stock index derivatives and it plays a core role in product design and risk management of index derivatives. It consists in constructing a tracking portfolio whose behavior is as similar as possible to a target index during a predefined period.

Broadly speaking, two different strategies can be used to track a given stock market index: the full replication and the non-full replication. The full replication consists in purchasing all constituent stocks of a given index. In practice, this strategy needs high transaction costs. An alternative way is the non-full replication, including the stratified sampling replication and the optimal replication. Since selection of the stocks in stratified sampling replication depends on the manager's experience, the tracking portfolio is non-optimal in general. We focus on the optimal replication strategy in this paper. The optimal replication aims to finding a portfolio that minimizes the tracking error by investing in only a subset of the assets using optimization technologies. This strategy involves much lower transaction costs, and can achieve acceptable and low tracking errors in principle.

Different quantitative approaches have been developed for index tracking in general and for the optimal replication in particular. Roll [27] first established mathematical models and conducted a mean-variance analysis on the basis of Markowitz's earlier study. Tabata and Takeda [31] discussed the index fund management based on the mean-variance model. Buckley and Korn [7] applied the optimal impulse control techniques to the index tracking with fixed and proportional transaction costs. Rudolf et al [28] suggested the piecewise linear measures of tracking error, and then solved the problem with linear programming. Alexander [2] proposed to construct the tracking portfolios by analyzing the coincidental structure between the time series of each of the assets and the time series of the tracked index. Ammann and Zimmermann [3] explored the relationship between statistical measures of tracking error and asset allocation restrictions based on admissible weight ranges. Gilli and Kellezi [18] suggested the use of the threshold accepting heuristic for solution of the problems with cardinality restrictions and transaction costs. Beasley et al. [5] addressed the problem through using the evolutionary heuristics with real-valued chromosome representations. Lobo et al. [22] studied the index tracking problem with transaction costs and proposed to solve the problem by a heuristic relaxation method which consists in solving a small number of convex optimization problems with fixed transaction costs. Canakgoz and Beasley [8] presented a nonlinear mixed-integer optimization model for the index tracking problem. Torrubiano et al [32] then designed a hybrid strategy that combined an evolutionary search with a quadratic programming solver to yield the optimal tracking portfolio of a problem in which invests are only in the selected assets.

Recently, the statistical regularization approaches have been applied to the mean-variance portfolio selections which successfully promote the identification of sparse portfolios with good out-of-sample properties and low transaction costs [13, 6, 16]. DeMiguel et al [13] analyzed the effect of constraints of an index tracking problem upon the covariance regularization, and then developed an L_p ($p = 1$ or 2) regularization based technique, as an extension of the method due to Jagannathan and Ma [20]. Brodie et al [6] emphasized the sparsity of the portfolio allocation and proposed an L_1 regularization based (or saying, LASSO based, [33]) algorithm for sparse solution of the index tracking. An important contribution of Fan et al [16] is the provision of deep mathematical insights to the utility approximations with the gross-exposure constraint. All these approaches, in essence, consist in solving the problems through constraining portfolio norms, say, the 2-norm in L_2 regularization based, as implied from the ridge regression ([19]), and the 1-norm in L_1 regularization based, as implied from sparse modeling [33, 21, 26]. Empirical results in a mean-variance framework support the use of the L_1 regularization based approaches when short selling is allowed. We observe, however, that the LASSO based approach can only validate for the index tracking with short selling constraints and it is invalid for the problems with budget and no-short selling constraints, since, the later cases lead to a constant value of the 1-norm of the asset weights, making no effect for constraining the portfolio norm. Moreover, the application results of those regularization approaches are seriously respective of the choice of regularization parameter, to which there is still no rule to follow (of course, many statistical criteria exist, see, e.g., [1, 29, 23]).

On the other hand, considering the index tracking problems with budget and no-short selling constraints is imperative [5]. The L_1 regularization based approach cannot be applied in this case. We develop an $L_{1/2}$ regularization based approach for the problems in the present paper.

The reasons why we apply the $L_{1/2}$ regularization modeling are as follows. First, a very recent study [17] showed that using L_q regularizations with $0 < q < 1$ can assuredly get the more sparse tracking portfolio than L_1 regularization, and, particularly, one can use the least stocks to track the target index by controlling the turnover, even it is known that the L_q regularization problems are nonconvex, non-smooth and hard to solve. Second, in spit of difficult for solution of a general L_q regularization problem, our recent studies [35, 36, 37, 38] revealed that among L_q regularizations, the $L_{1/2}$ regularization bears many specialities: It yields the most sparse solutions among L_q regularizations when $1/2 \leq q \leq 1$, and similar sparse solutions when $0 < q \leq 1/2$; The solutions of the $L_{1/2}$ regularization can be represented analytically as a forward-backward splitting form, and, consequently, a very powerful thresholding iterative procedure, called the $half$ algorithm, can be defined, corresponding to the well known $soft$ thresholding iterative method for L_1 regularization. Thus, $L_{1/2}$ regularization can be very fast solved as in the case of L_1 regularization. Third, the established $L_{1/2}$ regularization theory [37] provides a precise location of the optimal regularization parameter whenever the sparsity of the problem under consideration is known. This has lead to a nearly optimal regularization parameter setting strategy of the $half$ algorithm when applied to the sparsity-fixed problems. Such exclusive advantage meets the demand and is very desirable for the index tracking application, where managers are often requested to look for a tracking portfolio with a fixed number of stocks. This suggests that with the $L_{1/2}$ regularization framework, the index tracking problems might be very efficiently solved.

Thus, we focus, in the present paper, on the development of an $L_{1/2}$ regularization based new index tracking model as well as a $half$ algorithm based hybrid method for solution of the model. Different from the traditional way to find the optimal portfolio that best fits the historical evolution of assets, we are more interested in finding the sparse portfolio that predicts future well, in other words, has a good out-of-sample performance.

The remainder of the paper is organized as follows. In section 2 we briefly review the index tracking models and introduce the $L_{1/2}$ regularization framework with $half$ algorithm. In section 3 we present the $L_{1/2}$ regularization based sparse index tracking model and formalize the hybrid $half$ thresholding algorithm for fast solution of the model. The empirical tests of the model and the algorithm are provided in section 4 on the eight data sets in OR-library, together with several other competitive methods for index tracking. We then conclude the paper in section 5 with some useful remarks.

2 Preliminaries

In this section we review the index tracking models from the point of view of regression, and then, provide a general account of the $L_{1/2}$ regularization theory in sparse modeling.

2.1 Index Tracing Problem

Index tracking is a problem of determining a portfolio whose behavior is as close to a target index as possible during a predefined period. Different index tracking models can be formalized mathematically with different measures on tracking errors that measure quantitatively the difference of performance between the tracking portfolio and the target index [32].

The tracking error can be defined in many ways [3, 22, 30, 5, 7, 28]. Most of the definitions, however, are formulated either in terms of the correlations between the returns of tracking portfolio and the target index or with the estimations of the variance between the returns of the target index and the tracking portfolio [7, 28]. Beasley et al [5] argued against the use of variance as a measure of tracking error, based on an observation that the measure would be zero when the difference between the returns of the index and the tracking portfolio becomes constant. This latter property is unexpected because it has not taken the tracking bias into consideration. As an improvement, Beasley et al. [5] suggested a new definition of tracking error, the square of mean squared error of the return of the target index from the outcome of the tracking portfolio. We adopt this definition in the present study.

Let P_{it} $(i = 1, 2, ..., N)$ be the time series of stock prices for the N stocks that are included in a stock market index whose evolution we wish to replicate. Let $I(t)$ be the time series of this index. The series is assumed to be collected at the same series of times $t = 1, 2, ..., T$. According to [5], the tracking error is then defined by

$$TE = \frac{1}{T}\sum_{t=1}^{T}(\sum_{i=1}^{N} w_i r_{it} - R_t^I)^2 \tag{2.1}$$

where

r_{it} : the return rate of stock i at time t during single period, which is defined by

$$r_{it} = \frac{P_{it+1} - P_{it}}{P_{it}}; \quad i = 1, \cdots, N, \quad t = 1, 2, \cdots, T. \tag{2.2}$$

R_t^I : the return rate of the target index at time t during single period, that is,

$$R_t^I = \frac{I_{t+1} - I_t}{I_t}; \quad t = 1, 2, \cdots, T. \tag{2.3}$$

w_i : the weight attached to stock i, meaning the investment ratio of buying or selling the stock i.

Let

$$R^I = (R_1^I, R_2^I \cdots, R_T^I)^T \in R^{T\times 1}$$

and

$$R = (R_1, R_2, \cdots, R_N) = \begin{pmatrix} r_{11} & r_{21} & \cdots & r_{1N} \\ r_{12} & r_{22} & \cdots & r_{1N} \\ \vdots & \cdots & \cdots & \vdots \\ r_{1T} & r_{2T} & \cdots & r_{NT} \end{pmatrix}$$

where $R_i = (r_{i1}, r_{i2}, \cdots, r_{iT})^T$ is the column vector of return rate of the stock i , $i = 1, 2, \cdots, N$, $R \in R^{T\times N}$ is the matrix of all stock's return rate. Let $w = (w_1, w_2, \cdots, w_N)^T$ be the $N \times 1$ column vector of the stock weights. The tracking error (2.1) can be rewritten as

$$TE = \frac{1}{T}\|Rw - R^I\|_2^2. \tag{2.4}$$

The index tracking problem then aims to finding an optimal tracking portfolio through minimizing tracking error (2.4) under appropriate constraints.

The constraint we would be first confined is with the budget constraint

$$w_1 + w_2 + ..., +w_N = 1 \tag{2.5}$$

which ensures that all the capitals should be invested in the tracking portfolio. Here a negative w_i means the short selling is allowed, and the short selling is not allowed whenever all the w_i are nonnegative. The second constraint is concerned with the restriction of investment for each stock i :

$$\eta_i \le w_i \le \delta_i, \; i = 1, \cdots, N \tag{2.6}$$

where the lower bound, η_i, of the investment w_i aims to guarantee the lowest amount of investment and the upper bound, δ_i , is to control risk. Normally, η_i might be negative but δ_i must be positive. Other constraint is the cardinality constraint stated like

$$\sum_{i=1}^{N} z_i = \ k, \quad z_i = 0 \ \ or \ 1, \quad i = 1, 2, \cdots, N, \tag{2.7}$$

in which k, a preset positive integer, restricts the number of the stocks included in the tracking portfolio to be constructed, $z_i = 0$ means that asset i is not included in the portfolio, and $z_i = 1$ means the inclusion of asset i in the portfolio.

Thus, based on the definition of tracking error and constraints, a general index tracking model can be formulated as ([32])

$$\begin{array}{cc} \min_{w,z_i} & \frac{1}{T}\|Rw - R^I\|_2^2 \\ s.t. & w_1 + w_2 + ..., + w_N = 1 \\ & z_1 + z_2 + ... + z_N = k \\ & z_i\eta_i \leq w_i \leq z_i\delta_i \\ & z_i = 0 \ \ or \ 1, \ \ i = 1, 2, \cdots, N. \end{array} \tag{2.8}$$

The model (2.8) is a hybrid continuous-integer programming problem, which is highly nonlinear due to the existence of the cardinality constraint $z_1 + z_2 + ... + z_N = k$. Although solvable for not very large scale cases (see, e.g., [5, 8, 32]), it is generally very hard to have an optimal solution of (2.8). The model is seen also to be an ERM based regression model, which means that the optimal tracking portfolio yielded from the model can generally have a minimal in-sample tracking error, but, not necessarily, a highest out-of-sample performance.

Different from such ERM based approach, we will model the index tracking problem in the framework of regularization, which provides a natural trade-off between the in-sample performance and the out–of-sample performance. Furthermore, we model the problem as a sparsity problem which then brings a significant reduction of the solution-finding complexity. All these will be made possible when we apply the latest developed $L_{1/2}$ regularization theory in sparse modeling [35, 36, 37].

2.2 $L_{1/2}$ Regularization

We review the related $L_{1/2}$ regularization theory in this subsection. For the more details, we refer to [35, 36, 37].

Regularization is an approach for solution of ill-posed problems. Given an ill-posed problem, say, $Ax = b$, with $A \in R^{M \times N}$ and $M \ll N$, the regularization approach solves the problem through balancing the minimizations of the error $\|Ax - b\|_2^2$ and a penalty $p(x)$, in the form

$$\min_{x \in R^N} \left\{ \|Ax - b\|_2^2 + \lambda p(x) \right\} \tag{2.9}$$

where $\|.\|_2$ is the Euclidean norm, λ is the regularization parameter and $p(x)$ is a measure for the property we hope the solution of the problem to have. For example, $p(x) = \|x\|_2^2$ means the expectation we hope the solution is as smooth as possible, or, in machine learning terms, the resultant solution has the highest generalization capability, as demonstrated in support vector machine [34].

When we deal with a sparsity problem, i.e., to look for sparse solutions of a representation or an equation, the penalty $p(x)$ in (2.9) can be generally taken as the following L_q-norm form

$$p(x) = \sum_{i=1}^{N} |x_i|^q = \|x\|_q^q, \ 0 \le q \le 1$$

where $\|x\|_0^0 = |\sup\mathrm{pt}\{x\}|$ denotes the number of nonzero components of x. For each fixed $q \in [0,1]$, the regularization problem (2.9) is then referred to as L_q regularization. The L_0 regularization amounts to

$$\min_{x\in R^N} \left\{\|Ax-b\|_2^2 + \lambda \|x\|_0^0\right\}$$

which aims at finding the sparest solution of $Ax = b$ (in other words, the solution with the fewest nonzero components). Therefore, L_0 regularization is an idea model for sparsity problems. It is, however, combinatorial in nature and NP-hard to solve in general. A common practice is then to apply the L_1 regularization, as suggested independently by Tibshirani [33] and Chen, Donoho and Saunders [11], known respectively as LASSO ([33]) and Basis Pursuit ([11]). The L_1 regularization is a convex optimization problem and can be very efficiently solved. It also results in sparse solution of the considered problem, with a promise that, in many cases, the resultant solution coincides with one of the solutions of L_0 regularization [25, 12]. Because of this, the L_1 regularization gets its popularity and has been accepted as a most useful tool for solution of the sparsity problems.

Nevertheless, while L_1 regularization provides the best convex approximation to L_0 regularization and it is computationally efficient, the L_1 regularization can not yield the most sparse solution in most of application cases. In particular, it can not handle the collinearity and may yield inconsistent selection when applied to variable selection [24], and it can not recover a signal or image with the least measurements when applied to compressed sensing [36, 10, 9, 14]. Thus, a further modification is required. Among such efforts, the $L_{1/2}$ regularization

$$\min_{x\in R^N} \left\{\|Ax-b\|_2^2 + \lambda\|x\|_{1/2}^{1/2}\right\} \tag{2.10}$$

was highly recommended. The following theories and properties of $L_{1/2}$ regularization have been justified in [35]- [37]:

(i) The $L_{1/2}$ regularization can assuredly generate more sparse solution than L_1 regularization, and its sparsity-promotion ability is strongest among the L_q regularizations with all $q \in [1/2, 1]$, but similar in $q \in (0, 1/2]$.

(ii) For any $\mu \in (0, \|A\|_2^{-2})$, the solution of $L_{1/2}$ regularization problem (2.10) can be represented as

$$x^* = H_{\lambda\mu,1/2}(B_\mu(x^*)) \tag{2.11}$$

where $B_\mu(x) = x - \mu A^T(Ax - b)$, $H_{\lambda\mu,1/2}$ is the diagonally nonlinear thresholding operator defined by

$$H_{\lambda\mu,1/2}(x) = (h_{\lambda\mu,1/2}(x_1), h_{\lambda\mu,1/2}(x_2), \cdots, h_{\lambda\mu,1/2}(x_N)) \tag{2.12}$$

and $h_{\lambda\mu,1/2}(x_i)$ is the thresholding function, called the *half* thresholding function, defined by

$$h_{\lambda\mu,1/2}(x_i) = \begin{cases} \frac{2}{3}|x_i|\left(1 + \cos\left(\frac{2\pi}{3} - \frac{2\varphi_\lambda(x_i)}{3}\right)\right), & |x_i| > \frac{\sqrt[3]{54}}{4}(\lambda\mu)^{\frac{2}{3}} \\ 0, & otherwise \end{cases} \tag{2.13}$$

where

$$\cos\varphi_\lambda(x_i) = \frac{\lambda}{8}\left(\frac{|x_i|}{3}\right)^{-\frac{3}{2}}. \tag{2.14}$$

(iii) If $x^* = (x_1^*, x_2^*, ..., x_N^*)^T$ is a solution of $L_{1/2}$ regularization problem (2.10) and $\mu \in (0, \|A\|_2^{-2})$, then for any $i \in \{1, 2, ..., N\}$, either $x_i^* = 0$ or $|[B_\mu(x^*)]_i| \geq \frac{\sqrt[3]{54}}{4}(\lambda\mu)^{\frac{2}{3}}$.

(iv) If $L_{1/2}$ regularization problem (2.10) has $k-$sparsity solution (i.e., it has a solution x^* with k nonzero components), the regularization parameter λ must lie in

$$\lambda \in \left[\frac{\sqrt{96}}{9\mu}\left\{|B_\mu(x^*)|_{[k+1]}\right\}^{\frac{3}{2}}, \frac{\sqrt{96}}{9\mu}\left\{|B_\mu(x^*)|_{[k]}\right\}^{\frac{3}{2}}\right)$$

where $|B_\mu(x^*)|_{[k]}$ is the k-th largest component of vector $|B_\mu(x^*)|$.

(v) With operators $H_{\lambda\mu,1/2}$ and B_μ defined as in (2.11)-(2.14), the following iterations, called the *half* thresholding algorithm,

$$x_{n+1} = H_{\lambda_n\mu,1/2}(B_{\mu_n}(x_n)), x_0 \in R^N \tag{2.15}$$

defines a very efficient algorithm for solution of $L_{1/2}$ regularization problem (2.10), where $\mu \in (0, \|A\|_2^{-2})$ fixed and either

$$\lambda_n = \lambda \text{ or } \lambda_n = \frac{\sqrt{96}}{9\mu}\left\{|B_\mu(x_n)|_{[k+1]}\right\}^{\frac{3}{2}} \tag{2.16}$$

(vi) The *half* thresholding algorithm is convergent.

In the established theories, (i) shows the stronger sparsity-promotion ability of $L_{1/2}$ regularization over L_1 regularization; (ii) shows the analytic expressiveness of solutions of $L_{1/2}$ regularization, very exclusive in L_q regularizations when $0 < q < 1$; (iii) shows an alterative feature theorem of the solutions, which underlies the exact location on where the optimal regularization parameter λ should be ((iv)). This property is quite important, since setting an appropriate regularization parameter has been known a difficult issue, very crucial to the success of regularization methodology. It is the exact location information on regularization parameter ((iv)) that makes it possible to have the novel parameter setting strategy of λ_n, (2.16), in the $half$ thresholding algorithm. It is well known that there is a very efficient thresholding algorithm, called the $soft$ thresholding algorithm, for fast solution of L_1 regularization. The properties (v) and (vi) then show that there is still a fast thresholding algorithm, the $half$ thresholding algorithm, corresponding to the well developed $soft$ algorithm for L_1 regularization, even it is known that $L_{1/2}$ regularization (2.10) has lead to a nonconvex, non-smooth and non-Lipschitz optimization problem.

We observe that the problem of tracking a financial index using only a subset of stocks can be viewed as a sparsity problem. Particularly, whenever the number of stocks included in a tracking portfolio is fixed (say, k), the problem is a k-sparsity problem. Consequently, the $L_{1/2}$ regularization approach in general and the $half$ thresholding algorithm in particular cater for the need of solution of the index tracking problem. This explains partially why the $L_{1/2}$ regularization approach will be adopted in the present study.

3 The $L_{1/2}$ Regularization Based New Model And Algorithm

In this section, we formulate a new sparse index tracking model based on $L_{1/2}$ regularization, and, furthermore, we formalize a hybrid $half$ thresholding algorithm for fast solution of the model.

3.1 The New Model

We consider the index tracking problems with budget and no-short selling (thus $w_i \geq 0$ for all i). From constraints (2.7), if we ask $w_i = 0$ if and only if $z_i = 0$, the constraints can then be rewritten as

$$\|w\|_0^0 = k \tag{3.1}$$

and model (2.8) can be simplified into

$$\begin{array}{cc} \min_{w \in R^N} & \frac{1}{T}\|Rw - R^I\|_2^2 \\ s.t. & w_1 + w_2 + ..., +w_N = 1 \\ & \|w\|_0^0 = k \\ & \eta_i \leq w_i \leq \delta_i,\ i \in suppt(w). \end{array} \tag{3.2}$$

where $suppt(w) = \{i \in 1, 2, \cdots, N, w_i \neq 0\}$.

In the framework of L_0 regularization, this can be modeled as

$$\begin{array}{cc} \min_{w \in R^N} & \left\{ \|Rw - R^I\|_2^2 + \lambda \|w\|_0^0 \right\} \\ s.t. & w_1 + w_2 + ..., + w_N = 1 \\ & \eta_i \leq w_i \leq \delta_i, i \in suppt(w). \end{array} \tag{3.3}$$

where regularization parameter λ is introduced for balancing the two optimization terms $\|Rw - R^I\|_2^2$ and $\|w\|_0^0$, and for controlling the sparsity of the solution. As in sparse modeling of L_q regularization, the combinatorial term $\|w\|_0^0$ can be relaxed into the L_1 norm or $L_{1/2}$ norm so as to make model (3.3) computationally tractable, leading to the following L_1 regularization based model

$$\begin{array}{cc} \min_{w \in R^N} & \left\{ \|Rw - R^I\|_2^2 + \lambda \|w\|_1 \right\} \\ s.t. & w_1 + w_2 + ..., + w_N = 1 \\ & \eta_i \leq w_i \leq \delta_i, i \in suppt(w). \end{array} \tag{3.4}$$

and $L_{1/2}$ regularization based model

$$\begin{array}{cc} \min_{w \in R^N} & \left\{ \|Rw - R^I\|_2^2 + \lambda \|w\|_{1/2}^{1/2} \right\} \\ s.t. & w_1 + w_2 + ..., + w_N = 1 \\ & \eta_i \leq w_i \leq \delta_i, i \in suppt(w). \end{array} \tag{3.5}$$

We observe that the first constrain in (3.4) actually amounts to $\|w\|_1 = 1$. Bringing it into the objective function then leads to the following equivalent model of (3.4)

$$\begin{array}{cc} \min_{w \in R^N} & \|Rw - R^I\|_2^2 \\ s.t. & w_1 + w_2 + ..., + w_N = 1 \\ & \eta_i \leq w_i \leq \delta_i, i \in suppt(w). \end{array} \tag{3.6}$$

which comes back to the ERM model (2.8) and shows no effect of constraining the portfolio sparsity. This shows the invalidation of using the L_1 regularization based approach for the index tracking problems with budget and no-short selling. We therefore suggest the use of the $L_{1/2}$ regularization based index tracking model. This is also because, on one side, it can lead to a more sparse tracking portfolio, as concluded by the $L_{1/2}$ regularization theory introduced in the last section and, on the other side, the exact location of regularization parameter λ in $L_{1/2}$ regularization can naturally relate the regularization parameter with the sparsity of the solution, which can provide an efficient and intuitive way of controlling the number of stocks included in the tracking portfolio. We will explain this in more detail in the next section.

3.2 A Hybrid Half Thresholding Algorithm

The new model (3.5) is seen to be a constrained $L_{1/2}$ regularization problem, and therefore, the *half* thresholding algorithm introduced in the last section can not be directly applied. In this subsection we propose a hybrid $half$ thresholding algorithm for solution of the model.

The hybrid $half$ thresholding algorithm to be proposed will be divided into two steps. The first step aims to selecting a support set of the portfolio by solving a unconstrained $L_{1/2}$ regularization problem, and, the second step is to yield the optimal portfilo weights for the chosen stocks.

In the first step, we solve the k-sparsity problem with $L_{1/2}$ regularization form

$$\min_{w\in R^N} \quad \left\{\|Rw - R^I\|_2^2 + \lambda\|w\|_{1/2}^{1/2}\right\} \tag{3.7}$$

(namely, only minimizing the objective function) and result in a candidate portfolio $\widetilde{w}$ with $suppt(\widetilde{w}) = k$. In this step, the $half$ thresholding algorithm can be directly applied. With matrix A and vector b being replaced by R and R^I in (2.15)-(2.16), the algorithm can then be specified as

$$w_{n+1} = H_{\lambda_n\mu,1/2}(B_\mu(w_n)), \ w_0 \in R^N \tag{3.8}$$

where $B_\mu(w_n) = w_n + \mu R^T(R^I - Rw_n)$ and

$$\lambda_n = \frac{\sqrt{96}}{9}\|R\|^2\left\{|B_\mu(w_n)|_{[k+1]}\right\}^{\frac{3}{2}}, \mu = \frac{1}{\|R\|^2}. \tag{3.9}$$

Obviously, when so doing, the algorithm is an adaptive iteration, free from the choice of regularization parameter.

In the second step, we refine the tracking portfolio $\widetilde{w}$ and get an optimal portfolio w^*. The method is as follows. We let $w_i^* = 0$ when $i \notin suppt(\widetilde{w})$ and ask other components of w^* to be a solution of the following quadratic programming problem

$$\begin{array}{ll} min_{w\in R^N} & \|\widetilde{R}w - \widetilde{R}^I\|_2^2 \\ s.t. & \sum\limits_{i\in suppt\{\widetilde{w}\}} w_i = 1 \\ & \eta_i \le w_i \le \delta_i, \ i \in suppt\{\widetilde{w}\}. \end{array} \tag{3.10}$$

where $\widetilde{R} \in R^{K\times T}$ is the submatrix consisted of the k columns of R corresponding to the indices i in $suppt\{\widetilde{w}\}$, and $\widetilde{R}^I$ is the corresponding restriction of R^I. Many efficient algorithms exist for fast solution of (3.10), say, we can get a solution through simply applying the Matlab function (quadprog).

Remark 1 *It should be noted that though invalidated a direct use of the L_1 regularization based model (3.4), the two-step approach proposed above can apply to the L_1 regularization based model. Particularly, we can similarly suggest a hybrid L_1 regularization method for solution of the model. For example, we can apply the well developed Least Angle Regression (LARS) [15] method in the first step to get the support set, and then get the optimal portfolio as that in the above suggested hybrid $half$ thresholding algorithm. The corresponding hybrid algorithm could be named as the hybrid LARS* *algorithm. Different from controlling the regularization parameter by sparsity in the hybrid $half$ thresholding algorithm, the hybrid LARS seeks to solve the corresponding L_1 regularization problem for a range of values of regularization parameter λ, starting from a very large value, and gradually decreasing λ until the desired value is attainted. As λ evolves, the optimal solution moves on a piecewise affine path. As such, to find the needed tracking portfolio with k nonzero asset weights. We will compare such hybrid LARS with the hybrid half thresholding algorithms in the next section.*

4 Empirical Results

In this section, we evaluate the performance of the suggested $L_{1/2}$ regularization based sparse index tracking model (3.5) and the corresponding hybrid $half$ thresholding algorithm.

The evaluation was made by applying the hybrid half thresholding algorithm to conduct a series of optimal tracking portfolios for the standard market indexes in 1992-1997 (290 weeks). As recorded in the OR-Library ([4]), the weekly prices of the stocks included in Hang Seng (Hong Kong), DAX 100 (Germany), FTSE (Great Britain), Standard and Poor's 100 (USA), the Nikkei index (Japan), the Standard and Poor's 500 (USA), Russell 2000 (USA) and Russell 3000 (USA) are given. We are interested in finding portfolios to track each index in the predefined period. The performance is measured in terms of the out-of-sample error and the in-sample error.

For comparison purpose, two other competitive models and algorithms were simultaneously experimented, together with the new model (3.5) and the hybrid $half$ thresholding algorithm. These include the latest developed mixed integer programming model due to Torrubiano and Alberto [32] with the hybrid evolutionary algorithm developed in [32], and the L_1 regularization based sparse index tracking model (3.4) with the hybrid LARS algorithm. As that in Torrubiano and Alberto [32], we partitioned the data sets of weekly returns of the stocks into a training set containing the first half of the data (145 values) and a testing set with the rest of the data (145 values). The training sets were used to generate the optimal tracking portfolio, and the testing sets were used to estimate the out-of-sample error of the tracking portfolio. We simulated each model and algorithm 10 times with random initial values, and the average performance out of the 10 runs was taken as the final result of each model.

To facilitate cross comparison, given two models A and B, we let TEO_A and TEO_B respectively be the out-of-sample errors defined by models A and B, and let TEI_A be the in-sample error of model A. The consistency ($Cons_A$) between the in-sample performance and the out-of-sample performance of model A is defined by

$$Cons(A) = |TEI_A - TEO_A|$$

which measures how consistent a model and an algorithm when used to approximation and prediction. The higher consistency of a model and an algorithm, the more reliable the model and the algorithm in application. We also define superiority ($SupO(A, B)$)

of models A over B in out-of-sample performance by

$$SupO(A,B) = \frac{TEO_B - TEO_A}{TEO_B} \times 100\%$$

which measures the extent of performance improvement in terms of out-of-sample error when two models A and B are compared.

All experiments were conducted on a personal computer (2.67Ghz, 4Gb of RAM) with MATLAB 7.9 programming platform (R2009b). The lower and upper bound of the asset weight were set to $\eta_i = 0.01$ and $\delta_i = 0.5$, $i = 1, 2, \cdots, N$.

With so defined measures and parameter-settings, we report the experiment results one by one below.

4.1 Comparison with Mixed Integer Programming Model in [32]

We present experiments to compare the performance of $L_{1/2}$ regularization based model (3.5) with the latest developed mixed integer programming model of Torrubiano and Alberto [32]. The hybrid $half$ thresholding algorithm and the hybrid evolutionary algorithm developed in [32] were applied to select the portfolios consisted of k stocks ($k = 5, 6, 7, 8, 9, 10$), respectively tracking the indices Hang Seng, DAX 100, FTSE, Standard and Poor's 100 and the Nikkei index. The in-sample error and the out-of-sample error of the hybrid integer programming model are cited from [32]. The experimental results are then listed in Table 1.

Table 1 shows that both $L_{1/2}$ regularization based model (A) and the hybrid integer programming model (B) work well, with the consistency between the in-sample performance and the out-of-sample performance of A obviously higher than B. (Actually, it is seen from Table 1 that $Cons(A) = \bigcirc(10^{-5}) \sim \bigcirc(10^{-6})$ and $Cons(B) = \bigcirc(10^{-4}) \sim \bigcirc(10^{-5})$). This supports the philosophy of regularization approach that more emphasized the out-of-sample performance. Also, it is seen from Table 1 that the out-of-sample errors of the $L_{1/2}$ regularization based model are lower than those of the hybrid integer programming model at 80% cases, as indicated with $SupO(A,B) > 0$ (ranging from 2.87% *to* 61.22%) at 24/30 instances. This shows that the $L_{1/2}$ regularization based model has a higher out-of-sample performance as compared with the hybrid integer programming model.

As a meta-heuristic approach, the evolutionary algorithms are generally accepted as being able to yield globally optimal solutions. From this point of view, the hybrid evolutionary algorithm suggested by Torrubiano and Alberto [32] can find the optimal solutions with the lowest in-sample errors, as demonstrated in Table 2. This reveals that the higher out-of-sample performance of the $L_{1/2}$ regularization based model is at the cost of sacrifice of in-sample performance. Figure 1 shows the comparison of in-sample errors and out-of-sample errors of the two models when applied to FISE index tracking. A clear observation is that the $L_{1/2}$ regularization based model lies in the middle position both in-sample errors and out-of-sample errors, which supports that the $L_{1/2}$ regularization based model has provided a very smart trade-off between approximation and prediction.

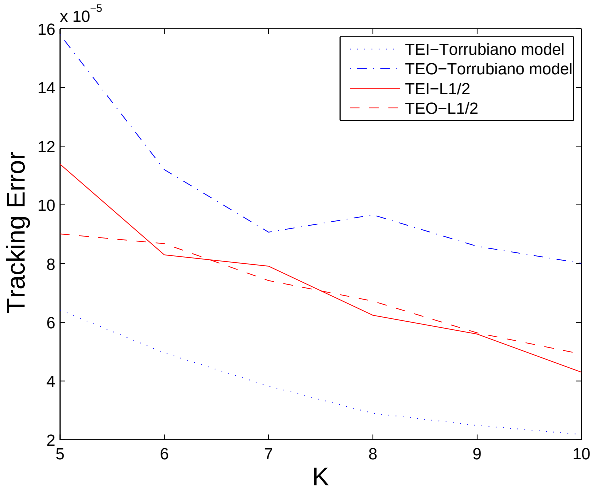

Figure 1: Performance comparsion between the $L_{1/2}$ regularization based model and the mixed integer programming model when applied to FISE index tracking.

Another advantage of the $L_{1/2}$ regularization based model over the hybrid integer programming model is the applicability of the use for large scale problems. Due to the diagonal nonlinearity of thresholding operator, the hybrid *half* thresholding algorithm can be effectively applied for very high dimensional problems, while the hybrid evolutionary algorithm only works for not large scale problems. For example, we consider the large scale cases: Standard and Poor's 500($N = 457$), Russell 2000 ($N = 1318$) and Russell 3000($N = 2151$). We selected the number of the tracking stocks $K = 80, 90, 100, 120, 150, 200$ to implement the two algorithms. The computational time was recorded in seconds. The simulation results are then shown in Table 2.

From Table 2 we see that the computational time of mixed integer programming model is much higher than $\mathrm{L}_{1/2}$ based model, In particulary, the latter model is 1581 times faster than the former model in average, observed from Table 2. Furthermore, it is seen also that the out-of-sample errors of the $L_{1/2}$ regularization based model are all lower than those of the hybrid integer programming model as indicated with $SupO(A, B) > 0$. This demonstrates that the $L_{1/2}$ regularization based model has a higher out-of-sample performance with lower computational cost as compared with the hybrid integer programming model.

4.2 Comparison with L_1 Regularization Based Model

This experiment was conducted to compare the performance of the $L_{1/2}$ regularization based model (A) with the L_1 regularization based model (C) through applying the hybrid *half* thresholding algorithm and the hybrid LARS algorithm to the index tracking. The experimental results are reported in Table 3.

From Table 3, we can see that both the $L_{1/2}$ regularization based and the L_1 regularization based models both fit and predict the data well, with an almost same consistency between the in-sample performance and the out-of-sample performance. Nevertheless, the $L_{1/2}$ regularization based model has a lower out-of-sample prediction error in most of the cases, as indicted in Table 3 where $SupO(A,C) > 0$ at 87%(= 26/30) instances. Moreover, we can find that the $L_{1/2}$ regularization based model can provide more sparse solutions to track the objective index. Actually, as shown in Figure 2 for Heng Seng index case, the out-of-sample prediction error of the L_1 regularization based model with $K = 10$ stocks is the same to that of the $L_{1/2}$ based model using $K = 5$ stocks.

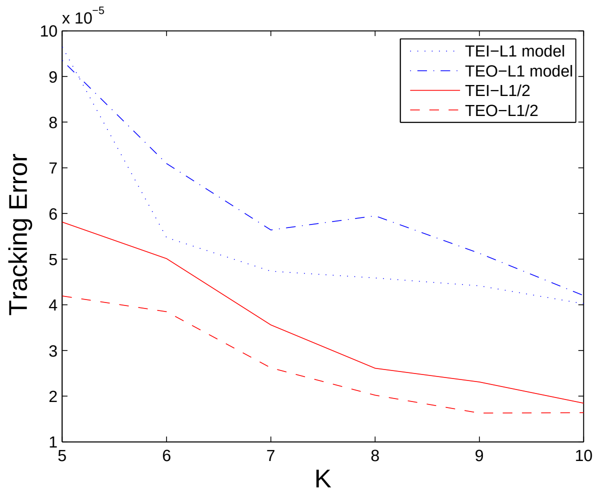

Figure 2: Performance comparsion between the $L_{1/2}$ regularization based model and the mixed integer programming model when applied to Heng Seng index tracking.

To be more precise, we can see from Table 3 (say, observe those data in boxes) that to track the objet index with accuracy $\bigcirc(10^{-5})$, the $L_{1/2}$ regularization based model needs 3 stocks(Hang Seng index), 6 stocks(DAX index), 5 stocks(FISE index),6 stocks (S&P100 index)and 8 stocks(Nikke225), while L_1 regularization based model needs 5 stocks(Hang Seng index), 6 stocks(DAX index), 9 stocks(FISE index),7 stocks(S&100 index) and 13 stocks(Nikke225). This demonstrate that the $L_{1/2}$ regularization based model has a stronger capability of yielding more sparse solutions than the L_1 regularization based model.

All the above experiments support that the proposed $L_{1/2}$ regularization based index tracking model and the hybrid *half* algorithm are feasible, effective and efficient. It outperforms other competitive approaches in yielding solutions with higher out-of-sample performance and sparser structure.

5 Conclusions

Index tracking is a passive financial strategy that aims at replicating the performance and risk-profile of a given index. One of the most common approaches for the index tracking consists in minimizing a given tracking error while limiting the maximum number of assets held in the portfolio. Having few active positions reduces the administrative and transaction costs and avoids detaining very small and illiquid positions, especially when the index has a large number of constituents. However, imposing an upper bound on the number of constituents of the tracking portfolio makes the optimization problem NP-hard. Although many evolutionary heuristic based algorithms have been suggested in recent years, for tackling such a NP-hard optimization problem, it is ineffective for large scale cases. The sparse modeling provides a novel approach for index tracking problems. The existing applications are, however, L_1 regularization based, which is ineffective sometimes in promotion of sparsity and selection of regularization parameter when applied to the index tracking with budget and no-short selling constraints.

In this paper we have proposed an $L_{1/2}$ regularization based sparse index tracking model to cater for the index tracking with budget and no-short selling constraints. Based on the newly developed $L_{1/2}$ regularization theory which, among others, shows that the $L_{1/2}$ regularization can assuredly have stronger sparsity-promotion capability, while it can still be very fast solved via a simple iterative thresholding procedure as that in L_1 regularization, we formalized a two-stage algorithm, called the hybrid $half$ thresholding algorithm, for fast solution of the model. The proposed algorithm solves a sparsity-fixed $L_{1/2}$ regularization problem in the first stage, and a convex quadratic programming problem in the second stage. Consequently, the algorithm is parameter-free and allows the regularization parameter λ to be set adaptively in a nearly optimal way, as suggested by $L_{1/2}$ regularization theory. The proposed model and algorithm are empirically tested on the eight data sets from OR-library. The tests show that the new model can effectively solve the index tracking problems, and the resulted tracking portfolio is of higher sparsity, lower out-of-sample prediction error and higher consistency of in-sample and out-of-sample performance, as compared with other competitive methods. The algorithm suggested provides an efficient solver for the index tracking, especially for the large scale problems whenever the number of constituents of expected portfolio is fixed. The obtained results can provide useful reference to the managers of index derivatives.

There are many problems deserving further research along the line of the present work. For example, to formulate an one-stage $L_{1/2}$ regularization based fast algorithm for the suggested sparse index tracking model, to justify the convergence of the proposed hybrid $half$ thresholding algorithm, and to extend the suggested model to the case of the index tracking with transaction costs. Some of those problems are under our current investigation.

References

[1] H. Akaike, Information theory and an extension of the maximum likelihood principle, In: B N Petrov , F Caki, eds. Second International Symposium on Information Theory. Budapest: Akademiai Kiado, 1973, pp. 267-281.

[2] C. Alexander, Optimal hedging using cointegration. Philosophical Transactions of the Royal Society of London, Series A. Mathematical, Physical and Engineering Sciences, 357 (1999), pp. 2039-2058.

[3] M. Ammann and H. Zimmermann, Tracking error and tactical asset allocation, Financial Analysts Journal, 57 (2001), pp. 32-43.

[4] J. E. Beasley, OR-Library: Distributing test problems by electronic mail, Journal of the Operational Research Society, 41 (1990), pp. 1069-1072.

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

[6] J. Brodie, I. Daubechies, C. DeMol, D. Giannone and D. Loris, Sparse and stable markowitz portfolios, (2008), ECB Working Paper Series.

[7] I. R. C. Buckley and R. Korn, Optimal index tracking under transaction costs and impulse control, International Journal of Theoretical and Applied Finance, 1 (1998), pp. 315-330.

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

[9] E. Candes and J. Romberg, Quantitative robust uncertainty principles and optimally sparse decompositions, Foundations of Computation Mathematics, 6 (2006), pp. 227-254.

[10] E. Candes, J. Romberg and T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Communications on Pure Applied Mathematics, 59 (2006), pp. 1207-1223.

[11] S. S. Chen, D. L. Donoho and M. A. Saunders, Atomic decomposition by basis pursuit,. SIAM Journal of Scientific Computing, 20 (1998), pp. 33-61.

[12] G. Davis, Adaptive Nonlinear Approximations, PhD thesis, New York University, 1994.

[13] V. DeMiguel, L. Garlappi, J. Nogales and R. Uppal, A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms, 2008, Working Paper.

[14] D. L. Donoho, Compressed sensing, IEEE Transactions on Information Theory, 52 (2006), pp. 1289-1306.

[15] B. Efron, T. Hastie, I. Johnstone and R. Tibshirani, Least angle regression, Annals of Statistics, 23 (2004), pp. 407-499.

[16] Jianqing Fan, Jingjin Zhang and Ke Yu, Asset allocation and risk assessment with gross exposure constraints for vast portfolios, 2008, Working paper.

[17] R. Fernholz, R. Garvy and J. Hannon, Diversity weighted indexing, Journal of Portfolio Management, 24 (1998), pp. 83-92.

[18] M. Gilli and E. Kellezi, Threshold accepting for index tracking, Working paper available from the first author at Department of Econometrics, University of Geneva, 1211 Geneva4, 2001, Switzerland.

[19] A. E. Hoerl and R. W. Kennard, Ridge regression: Biased estimation for nonorthogonal problems, Technometrics, 8 (1970), pp. 27-51.

[20] R. Jagannathan and T. Ma, Risk reduction in large portfolios:why imposing the wrong constraints helps, Journal of Finance, 58 (2003), pp. 1651-1683.

[21] Y. Karklin and M. S. Lewicki, Emergence of complex cell properties by learning to generalize in natural scenes, Nature, 457 (2009), pp. 83-86.

[22] M. Lobo, M. Fazel and S. Boyd, Portfolio optimization with linear and fixed transaction costs, Annals of Operations Research, 152 (2007), pp. 341-365.

[23] C. L. Mallows, Some comments on Cp, Technometrics,1973, 15 (1973), pp. 661-675.

[24] N. Meinshausen and P. Buhlmann, High dimensional graphs and variable selection with the lasso, Annals of Statistics, 34 (2006), pp. 1436-1462.

[25] B. K. Natarajan, Sparse approximate solutions to linear systems, SIAM Journal on Computing, 24 (1995), pp. 227-234.

[26] B. A. Olshausen and D. J. Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural images, Nature, 381 (1996), pp. 607-609.

[27] R. A. Roll, Mean-variance analysis of tracking error, The Journal of Portfolio Management, 18 (1992), pp. 13-22.

[28] M. Rudolf, H. J. Wolter and H. Zimmermann, A Linear model for tracking error minimization, Journal of Banking and Finance, 23 (1999), pp. 85-103.

[29] G. Schwarz, Estimating the dimension of a model. Annals of Statistics, 6 (1978), pp. 461-464.

[30] J. Shapcott, Index tracking: genetic algorithms for investment portfolio selection (Technical report, EPCC-SS92-24), Edinburgh, Parallel Computing Centre, 1992.

[31] Y. Tabata and E. Takeda, Bicriteria optimization problem of designing and index fund, Journal of the Operational Research Society, 46 (1995), pp. 1023-1032.

[32] P. R. Torrubiano and S. Alberto, A hybrid optimization approach to index tracking, Annals of Operation Research, 166 (2009), pp. 57-71.

[33] R. Tibshirani, Regression shrinkage and selection via the lasso, Journal of the Royal Statistical Society, 46 (1996), pp. 431-439.

[34] V. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, 1995.

[35] Z. B. Xu, H. Guo, Y. Wang and H. Zhang, The representation of L1/2 regularizer among Lq regularizer: an experimental study based on phase diagram, 2009, submitted.

[36] Z. B. Xu, H. Zhang, Y. Wang and X. Y. Chang, $L_{1/2}$ regularization, Science China Information Science, 53 (2010), pp. 1159-1169.

[37] Z. B. Xu, X. Y. Chang, F. M. Xu and H. Zhang, $L_{1/2}$ regularization: a thresholding representation theory and a fast solver, to appear in IEEE Transaction on Neural Networks, 2010.

[38] X. J. Chen, F. M. Xu and Y. Ye, Lower bound theory of nonzero entries in solution of L2-Lp minimization, SIAM Journal on Scientific Computing, 32 (2010), pp. 2832-2852.

Table 1: The experiment results of the $L_{1/2}$ regularization based model and the mixed integer programming model in [32].

Index	Sparsity	$L_{1/2}$ based model (A)			Mixed Integer Model (B)			$SupO(A,B)$
	K	TEI_A	TEO_A	$Cons_A$	TEI_B	TEO_B	$Cons_B$	
Hang Seng (N=31)	5	5.81e-5	4.19e-5	1.62e-5	4.14e-5	7.22e-5	3.08e-5	41.91
	6	5.01e-5	3.85e-5	1.16e-5	3.031e-5	4.76e-5	1.724e-5	19.03
	7	3.56e-5	2.62e-5	9.38e-6	2.37e-5	3.81e-5	1.44e-5	31.15
	8	2.61e-5	2.02e-5	5.92e-6	1.91e-5	2.90e-5	9.92e-6	30.36
	9	2.31e-5	1.63e-5	6.77e-6	1.62e-5	2.58e-5	9.59e-6	36.85
	10	1.84e-5	1.64e-5	2.07e-6	1.35e-5	2.06e-5	7.11e-6	20.36
DAX (N=85)	5	4.57e-5	1.20e-4	7.40e-5	2.21e-5	1.02e-4	7.97e-5	**-17.58**
	6	3.30e-5	8.78e-5	5.47e-5	1.76e-5	8.94e-5	7.17e-5	1.79
	7	2.41e-5	9.80e-5	7.39e-5	1.37e-5	8.46e-5	7.09e-5	**-15.83**
	8	2.14e-5	8.97e-5	6.83e-5	1.11e-5	7.93e-5	6.82e-5	**-13.08**
	9	1.94e-5	8.80e-5	6.86e-5	9.22e-6	7.78e-5	6.85e-5	**-13.14**
	10	2.96e-5	2.90e-5	5.68e-5	8.08e-6	7.48e-5	6.67e-5	61.22
FTSE N=89	5	1.14e-4	9.01e-5	2.37e-5	6.42e-5	1.58e-4	9.39e-5	43.00
	6	8.30e-5	8.68e-5	3.72e-6	4.96e-5	1.12e-4	6.23e-5	22.47
	7	7.91e-5	7.42e-5	4.87e-6	3.83e-5	9.07e-5	5.24e-5	18.15
	8	6.24e-5	6.72e-5	4.83e-6	2.90e-5	9.66e-5	6.76e-5	30.45
	9	5.60e-5	5.64e-5	6.19e-6	2.49e-5	8.59e-5	6.11e-5	34.41
	10	4.30e-5	4.92e-5	6.19e-6	2.18e-5	8.01e-5	5.82e-5	38.54
S&P N=98	5	1.21e-4	1.09e-4	1.06e-5	4.50e-5	1.14e-4	6.92e-5	3.72
	6	6.80e-5	8.30e-5	1.50e-5	3.37e-5	1.01e-4	6.70e-5	17.61
	7	8.72e-5	8.33e-5	3.88e-6	2.76e-5	7.80e-5	5.04e-5	**-6.80**
	8	3.89e-5	5.98e-5	2.08e-5	2.27e-5	6.76e-5	4.49e-5	11.66
	9	7.42e-5	4.90e-5	2.52e-5	1.94e-5	5.91e-5	3.97e-5	17.05
	10	3.99e-5	4.22e-5	2.25e-6	1.66e-5	5.55e-5	3.89e-5	23.96
Nikkei (N=225)	5	1.26e-4	1.58e-4	3.19e-5	5.46e-5	1.63e-4	1.08e-4	2.87
	6	1.15e-4	1.41e-4	2.58e-5	4.01e-5	1.47-4	1.07e-4	3.93
	7	8.81e-5	1.21e-4	3.38e-5	3.36e-5	1.32e-4	9.88e-5	7.93
	8	5.94e-5	9.34e-5	3.40e-5	2.60e-5	1.10e-4	8.40e-5	15.08
	9	5.96e-5	8.14e-5	2.18e-5	2.13e-5	9.80e-5	1.68e-5	17.01
	10	7.08e-5	6.96e-5	1.29e-6	1.80e-5	6.47e-5	4.67e-5	**-7.49**

Table 2: The performance comparison between the $L_{1/2}$ regularization based and the mixed integer programming models.

Index	Sparsity		$L_{1/2}$ Based Model (A)			Mixed Integer Model (B)		$SupO$
	K	TEI_A	TEO_A	$Time$	TEI_B	TEO_B	$Time$	
	80	6.07e-5	**8.74e-5**	2.7	1.13e-5	**9.61e-5**	4121	9.01
S&P	90	7.96e-5	**8.44e-5**	2.8	1.88e-5	1.03e-4	4963	17.97
(N=457)	100	7.16e-5	**9.48e-5**	2.8	1.90e-5	1.23e-4	6584	22.87
	120	9.79e-5	**9.52e-5**	3	2.68e-5	1.02e-4	6776	6.84
	150	5.77e-5	2.81e-4	5.9	2.77e-5	8.34e-4	13739	66.30
	200	5.43e-5	4.6e-4	9.5	$2.34e-8$	$6.58e-4$	139686	30.04
	80	1.31e-4	2.26e-4	18.3	3.06e-5	2.95e-4	3413	23.35
Russell2000	90	1.37e-4	2.16e-4	18.9	3.15e-5	2.75e-4	4639	21.49
(N=1318)	100	5.43e-5	2.45e-4	19.1	3.02e-5	4.20e-4	6256	41.80
	120	1.07e-4	2.08e-4	18.9	3.7e-5	2.41e-4	6817	13.70
	150	5.08e-6	3.96e-4	19.8	$3.23e-8$	$1.6e-3$	127105	52.63
	200	1.45e-5	3.1e-4	22.9	$6.76e-8$	$1.1e-3$	159855	71.99
	80	1.77e-4	1.65e-4	29.8	2.83e-5	1.95e-4	3975	29.8
Russell3000	90	4.04e-4	1.25e-4	29.9	4.26e-5	1.37e-4	6041	8.88
(N=2151)	100	1.24e-4	1.73e-4	30.1	4.41e-5	1.78e-4	6405	3.24
	120	2.94e-4	1.26e-4	29.6	6.07e-5	1.47e-4	6855	14.45
	150	8.37e-6	1.96e-4	36.2	$5.5e-5$	$2.3e-4$	10985	14.90
	200	6.41e-6	2.2e-4	41.1	$2.43e-7$	$2.0e-3$	136531	89.02

Table 3: The experiment results of the $L_{1/2}$ regularization based and the L_1 regularization based models

Index	Sparsity	$L_{1/2}$ Based Model (A)			L_1 Based Model (C)			*SupO*
	K	TEI_A	TEO_A	$Cons_A$	TEI_C	TEO_C	$Cons_C$	
	3	1.65e-4	**7.78e-5**	8.77e-5	2.70e-4	3.31e-4	4.31e-5	75.18
	4	9.65e-5	**6.81e-5**	2.84e-5	2.21e-4	2.26e-4	5.00e-6	69.87
Hang	5	5.81e-5	4.19e-5	1.62e-5	9.65e-5	**9.35e-5**	2.96e-6	55.15
Seng	6	5.01e-5	3.85e-5	1.16e-5	5.47e-5	**7.09e-5**	1.62e-5	45.73
(N=31)	7	3.56e-5	2.62e-5	9.38e-6	4.74e-5	5.64e-5	9.06e-6	53.50
	8	2.61e-5	2.02e-5	5.92e-6	4.59e-5	5.95e-5	1.36e-5	66.05
	9	2.31e-5	1.63e-5	6.77e-6	4.41e-5	5.13e-5	7.13e-6	68.20
	10	1.84e-5	1.64e-5	2.07e-6	4.02e-5	4.20e-5	1.79e-6	61.01
DAX	5	4.57e-5	1.20e-4	7.40e-5	3.26e-5	1.22e-4	8.94e-5	1.86
(N=85)	**6**	3.30e-5	**8.78e-5**	5.47e-5	2.25e-5	**9.89e-5**	7.64e-5	11.24
	7	2.41e-5	9.80e-5	7.39e-5	1.66e-5	8.57e-5	6.90e-5	**-14.37**
	8	2.14e-5	8.97e-5	6.83e-5	1.60e-5	8.04e-5	6.44e-5	**-11.48**
	9	1.94e-5	8.80e-5	6.86e-5	1.49e-5	7.81e-5	6.32e-5	**-12.62**
	10	2.96e-5	2.90e-5	5.68e-5	1.43e-5	7.81e-5	6.37e-5	62.82
FTSE	**5**	1.14e-4	**9.01e-5**	2.37e-5	1.06e-5	1.33e-4	2.67e-5	32.29
(N=89)	6	8.30e-5	**8.68e-5**	3.72e-6	9.94e-5	1.18e-4	1.83e-5	26.31
	7	7.91e-5	7.42e-5	4.87e-6	8.78e-5	1.14e-4	2.57e-5	34.63
	8	6.24e-5	6.72e-5	4.83e-6	7.61e-5	1.16e-4	4.02e-5	42.19
	9	5.60e-5	5.64e-5	6.19e-6	5.62e-5	**9.40e-5**	3.41e-5	39.95
	10	4.30e-5	4.92e-5	6.19e-6	5.34e-5	**8.75e-5**	3.41e-5	43.74
S&P	5	1.21e-4	1.09e-4	1.06e-5	1.01e-4	1.26e-4	2.44e-5	12.39
(N=98)	**6**	6.80e-5	**8.30e-5**	1.50e-5	8.15e-5	9.26e-4	1.10e-5	10.36
	7	8.72e-5	**8.33e-5**	3.88e-6	5.56e-5	**7.51e-5**	1.95e-5	**-10.95**
	8	3.89e-5	5.98e-5	2.08e-5	4.44e-5	**6.80e-5**	2.36e-5	12.13
	9	7.42e-5	4.90e-5	2.52e-5	4.27e-5	5.98e-5	1.71e-5	18.00
	10	3.99e-5	4.22e-5	2.25e-6	4.22e-5	5.73e-5	1.51e-5	26.39
Nikkei	5	1.26e-4	1.58e-4	3.19e-5	1.48e-5	2.10e-4	6.24e-5	24.72
(N=225)	6	1.15e-4	1.41e-4	2.58e-5	1.31e-5	2.20e-4	8.93e-5	35.87
	7	8.81e-5	1.21e-4	3.38e-5	1.18e-5	1.82e-4	6.39e-5	32.85
	8	5.94e-5	**9.34e-5**	3.40e-5	1.08e-5	1.66e-4	5.83e-5	43.69
	9	5.96e-5	**8.14e-5**	2.18e-5	9.89e-5	1.62e-4	6.29e-5	49.92
	10	7.08e-5	6.96e-5	1.29e-6	9.47e-5	1.59e-4	6.42e-5	56.25
	11	4.36e-5	7.15e-5	2.79e-5	8.27e-5	1.29e-4	4.63e-5	44.57
	12	3.87e-5	5.48e-5	1.61e-5	7.03e-5	1.06e-4	3.57e-5	48.30
	13	3.34e-5	5.20e-5	1.86e-5	6.62e-5	**7.83e-5**	1.24e-5	33.84