Expert Systems With Applications 201 (2022) 117110

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Robust enhanced index tracking problem with mixture of distributions

Zhilin Kang [a], Haixiang Yao [b], Xingyi Li [c,e], Zhongfei Li [d,c,*]

[a] School of Economics and Finance, Huaqiao University, Fujian 362021, China
[b] School of Finance, Guangdong University of Foreign Studies, Guangzhou 510006, China
[c] School of Business, Sun Yat-sen University, Guangzhou 510275, China
[d] School of Business, Southern University of Science and Technology, Shenzhen 518055, China
[e] Department of Management, Technology and Economics, ETH Zurich, Zurich, Switzerland

ARTICLE INFO

Keywords:
Enhanced indexing
Lower partial moment
Mixture model
Robust optimization
ϕ-divergence
Sortino index

ABSTRACT

Enhanced index tracking (EIT) is a popular form of investment strategy, which seeks to create a portfolio to generate excess return relative to a given benchmark index without purchasing all the index components. In this paper we develop an EIT model with mixture distribution under a lower partial moment (LPM) framework. Furthermore, we formulate the EIT problem as a robust and tractable model by integrating uncertain information on the proportions of Gaussian mixture distribution specified by the ϕ-divergence. By applying Lagrange duality theory, we demonstrate that the EIT problem on the basis of the worst-case LPMs of degree 1 and 2 can be transformed into a mathematically tractable optimization problem. Out-of-sample experiments using the FTSE100 and S&P500 data sets show that the portfolios based on our proposed model exhibit better performance than those from the benchmark index in most cases.

1. Introduction

In financial stock exchanges, market indices have become the standard benchmark to evaluate the investment styles and performance of fund managers. Index tracking (IT) is a type of passive investment strategy that aims to replicate the returns of a given benchmark index, but without purchasing all the constituent stocks that make up the index. The main characteristic of index tracking is mono-objective as the only goal of index tracking is to minimize the tracking error (Guastaroba & Speranza, 2012; Rudolf et al., 1999; Sant'Anna et al., 2017; Yu et al., 2006). For example, Sant'Anna et al. (2017) only minimize the tracking error computed as the average of the squared differences between the returns of the tracking portfolio and its benchmark index, and offer a hybrid method to solve this problem. The disadvantage of the IT investment is that if the benchmark index goes through a bearish phase, so does the constructed portfolio. EIT is an improved investment strategy in portfolio management, which focuses on outperforming, and not just tracking the given index or benchmark. In fact, tracking error and excess return relative to the benchmark index are two main elements in EIT (Beasley et al., 2003). A lot of pioneer works have been done on related problem from the perspective of modeling and solving methods. For instance, Guastaroba and Speranza (2012) introduce a mixed integer linear optimization model with the purpose to minimize the absolute deviation between the portfolio returns and the benchmark index returns. Bruni et al. (2015) model the EIT as a bi-objective linear program that maximizes the average excess return of the portfolio over the benchmark, and minimizes the maximum downside deviation of the portfolio return from the market index. Most of these studies apply evolutionary-based algorithms to solve EIT problems and rely on a backward perspective in that a portfolio that has tracked the index well in the past will also perform a good tracking in future.

As we know, there is substantial empirical evidence that the distribution of financial returns tends to be skewed and highly leptokurtic. For the purpose of expanding and enriching the IT and EIT models for tracking portfolios with heavy-tailed risk factors, we prefer to use mixture models to model the randomness of the parameters. As a convenient and flexible statistical tool, mixture models can not only fit better multivariate distributions, but also present a framework of constructing more complex distributions (Wang & Taaffe, 2015). Many popular distributions, such as elliptical distribution, generalized hyperbolic distribution, normal inverse Gaussian (NIG) distribution and skew-*t* distribution, can be regarded as a (infinite) mixture of normal distributions (McNeil et al., 2005). Among various mixture distributions, we are mainly interested in the Gaussian mixture distribution, which has been studied extensively in the literature (Hu et al., 2018; Kapsos et al., 2014; Zhu & Fukushima, 2009; Zhu et al., 2009). In these studies, the underlying distribution is a mixture of some predetermined component distributions, but there is no knowledge of the proportions of the mixture.

* Corresponding author.
E-mail addresses: zlk@hqu.edu.cn (Z. Kang), yaohaixiang@gdufs.edu.cn (H. Yao), lixy399@mail2.sysu.edu.cn (X. Li), lizf6@sustech.edu.cn (Z. Li).

https://doi.org/10.1016/j.eswa.2022.117110
Received 18 June 2021; Received in revised form 12 January 2022; Accepted 28 March 2022
Available online 9 April 2022

This study mainly focuses on data-driven robust EIT problem with a family of mixture distributions. To the best of our knowledge, this is the first application of the mixture distribution to an EIT problem. It is related to Zhu et al. (2014) and Hanasusanto et al. (2015) that the underlying distribution is also described as a mixture model. Our model in this paper is different from that in these two papers, in which they proposed a robust portfolio problem framework based on the left-tail risk measure CVaR and a distributionally robust approach to tackle the problem of CVaR-based risk averse multi-dimensional newsvendor, respectively. In this paper, we are primarily interested in considering EIT problems with a benchmark and describe the tracking error as an asymmetric measure (LPM) to capture the downside deviations from the index-plus-alpha returns, since upside deviations are not only allowable but also desirable in EIT. Different from the meta-heuristic and deep learning methods used in Li et al. (2011) and Kwak et al. (2021), this research aims to propose a mathematical programming framework to optimize the EIT problem.

The main contributions of this paper can be concluded as follows:

(1) We present a stochastic formulation for the EIT model to construct a tracking portfolio and achieve a specified excess return with respect to special benchmark index return (index-plus-alpha return). In the case of continuous random variables, a tractable deterministic equivalent reformulation can be developed under the assumption of a mixture of normals distributions. To the best of our knowledge, this is new to the EIT literature and the first paper using the Gaussian mixture method (GMM) to (enhanced) index tracking problems.

(2) Besides the EIT model, we propose another formulation based on the robust version of the LPMs. The formulation assumes the stocks and random benchmark return are specified by a mixture model where only certain inherent uncertainty (uncertainties of mixture proportions) is permitted. We analyze how to create tracking portfolios by minimizing worst-case LPMs under the mixture model and translate the robust EIT problem with a ϕ-uncertainty structure in mixture proportions into a deterministic non-linear programming. We show how it can be numerically calculated and optimized by a convex programming.

(3) In reality, investors concerns about transaction costs since they can affect the return of the investment. Apart from a budget constraint, we take into account proportional transaction cost all EIT models. The FTSE100 and S&P500 indices are used to evaluate the performances of the optimal strategies via backtesting procedures.

The remainder of this paper is organized as follows. Section 2 summaries the literature on EIT problems. In Section 3, we develop the EIT model by using index-plus approach and investigate the EIT problems with the mixture distribution. In Section 4, by incorporating uncertain information in the proportions of Gaussian mixture distribution specified by the ϕ-divergence, we reformulate the EIT problem as a robust and tractable model. Empirical analyses are presented in Section 5, which are aim to test the proposed EIT framework. Finally, we conclude in Section 6.

2. Literature review

In literature, many researchers have investigated problems regarding the EIT. The readers who are interested in overview of recent progress can see Canakgoz and Beasley (2008), Guastaroba et al. (2016) and Filippi et al. (2016). In the rest of this section, we concentrate our literature review on some of the most relevant papers to this paper.

Among the literature, the first mathematical formulation of the EIT problem is proposed by Beasley et al. (2003). In past decade, several formulations have been developed for the EIT problem (Bruni et al., 2015; Sehgal & Mehra, 2019; Xu et al., 2018). However, most of the researches use the standard deviation and absolute deviation as risk measures (de Paulo et al., 2016; Filippi et al., 2016; Huang et al., 2018; Kwon & Wu, 2017; Roll, 1992). For example, de Paulo et al. (2016) present an analytical solution for an uni-period EIT problem with limited number of assets held in the tracking portfolio, which is easier for numerical implementation. The drawback of these measures is that the tracking errors are symmetric distance measures, and cannot effectively control the tracking portfolio loss. Therefore, Some researchers propose conditional value at risk (CVaR) and downside deviation from the benchmark to capture the downside risk of a portfolio (Guastaroba et al., 2020; Li et al., 2011; Sehgal & Mehra, 2019). For example, Guastaroba et al. (2020) use mixed CVaR in EIT problem to build a risk-reward ratio optimization model, which is similar to the model proposed by Guastaroba et al. (2016). Li et al. (2011) formulate the enhanced index tracking problem of minimizing the downside deviation of order two from index return and maximizing the excess return with real-life features. Different from Li et al. (2011), we consider the τ-th moment ($\tau = 1, 2$) of the downside as an investor's risk measure and use index-plus-alpha portfolio (Koshizuka et al., 2009) to track the benchmark index return with an additional alpha return. Our paper focuses on (robust) EIT problem with a family of mixture distributions.

In recent years, some researchers use index-plus-alpha approach to track the returns of the benchmark index with an additional alpha return. Konno and Hatagi (2005) derive a mean absolute deviation model for index tracking, where their objective function is the difference between index values (scaled up by a factor alpha) and the tracking portfolio value. Koshizuka et al. (2009) build a convex minimization model with the aim of tracking an index-plus-alpha portfolio[1]. Two alternative measures of the tracking error have been proposed by the authors, namely the downside deviation and the absolute deviation between the portfolio return and the index-plus-alpha return. Guastaroba and Speranza (2012) develop linear programming formulations under framework of the IT and EIT models and use a heuristic approach (called Kernel Search) to solve mixed-integer linear programming models that also include cardinality, buy-in, and transaction costs constraints. More recently, Guastaroba et al. (2016) propose omega ratio with fixed and random targets as performance measure for the EIT.

Although there are various mathematical formulations for the IT and EIT problems in the existing studies, the IT with robust approach[2] has only recently been introduced in the literature, and the papers related to this approach are still very limited. Indeed, so far as we know, the earliest work in the robust index tracking literature is Chen and Kwon (2012) who propose a model to maximize pairwise similarities between selected assets and the assets of the target index when the similarity coefficients are uncertain. Another robust index tracking problem is formulated by Kwon and Wu (2017) and Wu and Wu (2019), they describe the uncertainty in returns through robust factor model. Their formulation is more straightforward to form index tracking portfolios: the objective is to maximize the expected portfolio return under the limitation of its tracking error. Lejeune (2012) formulates the EIT problem as a stochastic game theoretical framework where the probability, that the return of the portfolio is bigger than the benchmark, is maximized under the bound on the relative risk quantified by the downside deviation. Ling et al. (2014) study tracking error portfolios designed to follow specific benchmark portfolios such as indices. They develop a robust optimization approach based on a semidefinite programming for two downside risk measures, probabilities of loss and expected loss. Xu et al. (2018) follow the framework of Lejeune (2012) and develop a sparse EIT model with cardinality and chance constraints, in which the asset returns are modeled as random variables driven by various factor models. Recently, Khoshabar et al. (2020) study IT and EIT problems

[1] Index-plus-alpha portfolio is a term sometimes encountered in the literature on EIT, which refers to a portfolio that exceeds the benchmark by a given small quantity (Guastaroba et al., 2016; Sehgal & Mehra, 2019).

[2] In the last decades, robust optimization approach has been adopted as a powerful tool to model uncertain parameters with inexact distributions and seek optimal robust decisions. Useful guides for practitioners in this area can be found in Gorissen et al. (2015).

in portfolio optimization under interval uncertainty for returns and covariance matrix and test the models on EUROSTOXX 50 data set. Although the EIT problem is attracting a growing attention among the academics, the study of the robust EIT problem, which is the main topic of this paper, is a relatively new and immature research field.

3. Enhanced index tracking problem under mixture distribution

The EIT problem is the problem of managing a portfolio of assets that can beat a stock market with limited additional risk. A direct way to account for enhanced indexation is to simply construct an optimal portfolio that tracks an index-plus-alpha portfolio (Guastaroba et al., 2016; Koshizuka et al., 2009). Hence, we use the index-plus-alpha approach, proposed by Konno and Hatagi (2005) and Koshizuka et al. (2009) to track the benchmark index return with an additional alpha return. In contrast to the classical square deviations model, downside risk measure (especially the LPM measure) is a more general framework as it allows investors to consider different orders and to choose a favorable target. Hence, it motivates our interest in using LPM, that is better able to capture the left tail loss of portfolios, to establish our EIT model.

Consider an investor who selects a portfolio of n assets from an index of N assets ($n < N$). Denote the stochastic return vector of these assets by $\xi = (\xi_1, \dots, \xi_n)^T \in \mathbb{R}^n$. Let $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ denote the portfolio weight vectors and $\xi_b \in \mathbb{R}$ be the random variable representing the return of the benchmark index. The purpose of the investor is to choose a tracking portfolio x to solve the following τ-th LPM-based EIT model

$$(\mathrm{EIT}_\tau): \quad \min_{x\in\mathcal{X}} \ \mathrm{E}_{\mathrm{P}}[(\kappa + \xi_b - \sum_{i=1}^{n} \xi_i x_i)_+^\tau],$$

where $\mathcal{X}$ denotes the set of all admissible portfolios, κ is the desired excess return, τ is the order of moments and $(\cdot)_+ = max\{\cdot, 0\}$. In this model, we seek relative return (return relative to the benchmark index) rather than absolute return. We track the performance of the benchmark as much as possible, and to avoid poor performance, the performance is tracked by factor κ which is a constant.

Remark 1. In fact, in EIT, the investors may share a common interest in constructing an optimal portfolio that beats the return of the benchmark index over a given excess return. To achieve this goal, rather than simply use the return of the benchmark index ξ_b, one may use some reference random variable $\xi_b + \kappa$ to represent the portfolio return over the benchmark index return. When $\kappa \geq 0$, it means that the (EIT_τ) models are required to get excess return over benchmark. It is enough to take $\kappa = 0$ in tracking error models when the goal of investors is just to follow a market index ξ_b.

Remark 2. Suppose $\tau \geq 1$. For any distribution satisfying the regularity condition, i.e., $(\mathrm{E}_{\mathrm{P}}|\kappa + \xi_b - \xi^T x|^\tau)^{\frac{1}{\tau}} < \infty$, the τ-th LPM is a convex function of portfolio position x (Yao et al., 2021).

Denote the (future) return rates of the assets and decision vector

$$\tilde{\xi} = \begin{pmatrix} \xi \\ \xi_b \end{pmatrix} \in \mathbb{R}^{n+1}, \quad \tilde{x} = \begin{pmatrix} x \\ -1 \end{pmatrix} \in \mathbb{R}^{n+1},$$

respectively. Then, the problem (EIT_τ) can be formulated as

$$\min_{x\in\mathcal{X}} \ \mathrm{E}_{\mathrm{P}}[(\kappa - \tilde{\xi}^T \tilde{x})_+^\tau].$$

Since the τ-th moment of the downside risk $\mathrm{E}_{\mathrm{P}}(\kappa - \tilde{\xi}^T \tilde{x})_+^\tau$ ($\tau \geq 1$) is convex about x, this problem is a convex optimization problem and thus is easily solvable. In the following parts, we firstly introduce a multivariate mixture of distributions and model stochastic returns as mixture random variables. Then, for the cases $\tau = 1, 2$, we analytically derive a closed-form formula for the LPMs with a mixture of normal distributions.

3.1. Mixture distribution

When problems (EIT_τ), $\tau = 1, 2$ are being considered, the underlying distribution P is assumed as a given one. Many random phenomena can be described by normal distributions when their sample sizes are large enough. In reality, however, many market risk factors are heavy-tailed. In this paper, we firstly suppose that P is a mixture distribution in the following expression:

$$p(z) = \sum_{i=1}^{d} \lambda_i p_i(z), \tag{1}$$

where $p_i(z), i = 1, \dots, d$ are probability distributions which are also called component distributions, and $\lambda_i > 0, i = 1, \dots, d$ are the component proportions satisfying the equation $\sum_{i=1}^{d} \lambda_i = 1$. Obviously, a mixture distribution is defined as a convex combination of multiple distribution functions, commonly known as the component distributions. In particular, the weights corresponding to the mixture components are described as mixture proportions.

Mixture models are frequently used to analyze the characteristics of underlying states since the first mixture model was proposed (McLachlan & Peel, 2000). A most important and widespread instance of the mixture models in the field of statistical modeling is the mixture of Gaussian distribution. In general, Gaussian mixture distribution is a kind of useful distribution to describe the characteristics of heavy tails. An attractive feature of the Gaussian mixture model is that it is flexible enough to accommodate various shapes of continuous distributions, and it can capture skewed, leptokurtic and multimodal characteristics of financial time-series data. For example, Marron and Wand (1992) suggest that any distribution function can be arbitrarily approximated close by a mixture of Gaussian distribution functions. Bellalah and Lavielle (2002) also show that a three-Gaussian mixture can well approximate the empirical distribution for the main equity financial indices. One possible explanation for such a mixture is stated as follows. Let U be a discrete random variable taking values $i = 1, 2, \dots, d$ with the probability $\mathrm{P}(U = i) = \lambda_i$. Assume that the conditional distribution of the stochastic return vector $\tilde{\xi}$ knowing U is normal distribution and defined by

$$\tilde{\xi}|(U = i) \sim \mathcal{N}(\mu_i, \Sigma_i),$$

where μ_i is an $(n+1)$-dimensional vector and Σ_i is an $(n+1) \times (n+1)$ symmetric positive definite matrix. Then, the probability density function of $\tilde{\xi}$ satisfies the equation $p_{\tilde{\xi}}(\cdot) = \sum_{i=1}^{d} \lambda_i \varphi(\cdot|\mu_i, \Sigma_i)$, where each component $\varphi(\cdot|\mu_i, \Sigma_i)$ is a probability density function of (multivariate) normal distribution with mean vector μ_i and covariance matrix Σ_i.

Remark 3. Problem (EIT_τ) under a mixture distribution becomes

$$\min_{x\in\mathcal{X}} \ \sum_{i=1}^{d} \lambda_i \int_{\tilde{\xi}\in\mathbb{R}^{n+1}} (\kappa - \tilde{\xi}^T \tilde{x})_+^\tau p_i(\tilde{\xi}) d\tilde{\xi}, \tag{2}$$

where $\lambda \in \Lambda = \{\lambda \in \mathbb{R}^d \mid \lambda \geq 0, \sum_{i=1}^{d} \lambda_i = 1\}$. In the special case where $d = 1$, i.e., the distribution of $\tilde{\xi}$ is certainly known.

3.2. Closed-form expression for LPMs of mixed normal distribution

As mentioned, we suppose that the vector of asset returns and index return, denote by $\tilde{\xi} \in \mathbb{R}^{n+1}$, is modeled as a mixture of Gaussian distribution with mixture proportions λ_i, mean vector μ_i and the covariance matrix Σ_i which is positive definite $(i = 1, \dots, d)$, i.e.,

$$\tilde{\xi} \sim \sum_{i=1}^{d} \lambda_i \mathcal{N}(\mu_i, \Sigma_i). \tag{3}$$

For convenience, we define $Y = \tilde{\xi}^T \tilde{x}$, for any $\tilde{x} \in \mathbb{R}^{n+1}$, whose distribution as shown in Paolella (2015) is

$$f_Y(y) = \sum_{i=1}^{d} \lambda_i \varphi(y|v_i(x), \sigma_i^2(x)), \tag{4}$$

where $\varphi(y|v_i(x),\sigma_i^2(x))$ denotes the Gaussian distribution with mean $v_i(x)$ and variance $\sigma_i^2(x)$ assessed by the point $\tilde{x}$, $v_i(x)=\mu_i^T\tilde{x}$, and $\sigma_i^2(x)=\tilde{x}^T\Sigma_i\tilde{x}$, $i=1,\ldots,d$. The first two moments of Y are

$$\mu_Y=\mathrm{E_P}(Y)=\sum_{i=1}^d\lambda_i v_i(x),\quad \sigma_Y^2=\sum_{i=1}^d\lambda_i(\sigma_i^2(x)+v_i^2(x))-\mu_Y^2.$$

In the Gaussian mixture setting, we derive the following closed-form expressions for the LPMs

Proposition 1. *The closed-form expression for the LPMs in the Gaussian mixture setting can be presented as follows*

(1). For the case $\tau=1$, we have

$$\mathrm{E_P}[(\kappa-\tilde{\xi}^T\tilde{x})_+^1]=\sum_{i=1}^d\lambda_i[(\kappa-v_i(x))\Phi(\frac{\kappa-v_i(x)}{\sigma_i(x)})+\sigma_i(x)\varphi(\frac{\kappa-v_i(x)}{\sigma_i(x)})].$$

(2) . For the case $\tau=2$, we have

$$\mathrm{E_P}[(\kappa-\tilde{\xi}^T\tilde{x})_+^2]=\sum_{i=1}^d\lambda_i[(\sigma_i^2(x)+(\kappa-v_i(x))^2)\Phi(\frac{\kappa-v_i(x)}{\sigma_i(x)})+\sigma_i(x)(\kappa-v_i(x))\varphi(\frac{\kappa-v_i(x)}{\sigma_i(x)})],$$

where $\varphi(\cdot)$ and $\Phi(\cdot)$ are the standard normal probability density function and cumulative distribution function, respectively.

Proof. Using (4) we can pose the LPM of degree τ as

$$\begin{aligned}\mathrm{E_P}[(\kappa-\tilde{\xi}^T\tilde{x})_+^\tau]&=\mathrm{E}[(\kappa-Y)_+^\tau]\\&=\int_{y\in\mathbb{R}}(\kappa-y)_+^\tau\sum_{i=1}^d\lambda_i\varphi(y|v_i(x),\sigma_i^2(x))dy\\&=\sum_{i=1}^d\lambda_i\int_{y\in\mathbb{R}}(\kappa-y)_+^\tau\varphi(y|v_i(x),\sigma_i^2(x))dy.\end{aligned}$$

It is straightforward to obtain $\int_{-\infty}^a t\varphi(t)dt=-\varphi(a)$ and $\int_{-\infty}^a t^2\varphi(t)dt=\Phi(a)-a\varphi(a)$. Using integral calculation leads to

$$\begin{aligned}&\mathrm{E_P}[(\kappa-\tilde{\xi}^T\tilde{x})_+^1]\\&=\sum_{i=1}^d\lambda_i\int_{y\in\mathbb{R}}(\kappa-y)_+^1\varphi(y|v_i(x),\sigma_i^2(x))dy\\&=\sum_{i=1}^d\lambda_i\int_{-\infty}^\kappa(\kappa-y)\varphi(y|v_i(x),\sigma_i^2(x))dy\\&=\sum_{i=1}^d\lambda_i\int_{-\infty}^\kappa(\kappa-y)\frac{1}{\sqrt{2\pi}\sigma_i(x)}e^{-\frac{(y-v_i(x))^2}{2\sigma_i^2(x)}}dy\\&=\sum_{i=1}^d\lambda_i\{(\kappa-v_i(x))\int_{-\infty}^{\frac{\kappa-v_i(x)}{\sigma_i(x)}}\varphi(t)dt-\sigma_i(x)\int_{-\infty}^{\frac{\kappa-v_i(x)}{\sigma_i(x)}}t\varphi(t)dt\}\\&=\sum_{i=1}^d\lambda_i[(\kappa-v_i(x))\Phi(\frac{\kappa-v_i(x)}{\sigma_i(x)})+\sigma_i(x)\varphi(\frac{\kappa-v_i(x)}{\sigma_i(x)})]\end{aligned}$$

and

$$\begin{aligned}&\mathrm{E_P}[(\kappa-\tilde{\xi}^T\tilde{x})_+^2]\\&=\sum_{i=1}^d\lambda_i\int_{y\in\mathbb{R}}(\kappa-y)_+^2\varphi(y|v_i(x),\sigma_i^2(x))dy\\&=\sum_{i=1}^d\lambda_i\int_{-\infty}^\kappa(\kappa-y)^2\varphi(y|v_i(x),\sigma_i^2(x))dy\\&=\sum_{i=1}^d\lambda_i\int_{-\infty}^\kappa(\kappa-y)^2\frac{1}{\sqrt{2\pi}\sigma_i(x)}e^{-\frac{(y-v_i(x))^2}{2\sigma_i^2(x)}}dy\\&=\sum_{i=1}^d\lambda_i\sigma_i^2(x)\{(1+(\frac{\kappa-v_i(x)}{\sigma_i(x)})^2)\Phi(\frac{\kappa-v_i(x)}{\sigma_i(x)})+\frac{\kappa-v_i(x)}{\sigma_i(x)}\varphi(\frac{\kappa-v_i(x)}{\sigma_i(x)})\}\\&=\sum_{i=1}^d\lambda_i[(\sigma_i^2(x)+(\kappa-v_i(x))^2)\Phi(\frac{\kappa-v_i(x)}{\sigma_i(x)})+\sigma_i(x)(\kappa-v_i(x))\varphi(\frac{\kappa-v_i(x)}{\sigma_i(x)})],\end{aligned}$$

which yield the desired results. □

For convenience, define

$$R_i^1(x)=(\kappa-v_i(x))\Phi(\frac{\kappa-v_i(x)}{\sigma_i(x)})+\sigma_i(x)\varphi(\frac{\kappa-v_i(x)}{\sigma_i(x)}),i=1,\ldots,d,\tag{5}$$

$$\begin{aligned}R_i^2(x)=&(\sigma_i^2(x)+(\kappa-v_i(x))^2)\Phi(\frac{\kappa-v_i(x)}{\sigma_i(x)})\\&+\sigma_i(x)(\kappa-v_i(x))\varphi(\frac{\kappa-v_i(x)}{\sigma_i(x)}),\ i=1,\ldots,d.\end{aligned}\tag{6}$$

Given the specific form of LPMs in the Gaussian mixture setting described by Proposition 1, we can explicitly state problem (EIT$_\tau$) ($\tau=1,2$) under a mixture distribution as

$$\min_{x\in\mathcal{X}}\ \sum_{i=1}^d\lambda_i R_i^\tau(x),\tag{7}$$

where the mixture proportions $\lambda_i\in\mathbb{R}$, $\mu_i\in\mathbb{R}^{n+1}$ and $\Sigma_i\in\mathbb{R}^{(n+1)\times(n+1)}$ ($i=1,\ldots,d$) are directly estimated by a EM iteration process (see Section 5.1).

4. Robust enhanced index tracking problem under mixture distribution

As we know, stock values can be influenced by many different types of events such as industrial improvements, economic turmoil, and governmental crises. Due to the uncertainty nature of these events, it is difficult to accurately estimate the future value of stocks in a financial market. In the real world, it is often the case that only certain parts of the distribution of random variables are available. We suppose that P is ambiguous and only known to be an element of a given ambiguity set $\mathbb{D}$, which can be regarded as a confidence region in the space of distributions of stock return. Under this kind of distributional ambiguity the investor has to address the following variant of problem (EIT$_\tau$):

$$(\mathrm{DEIT}_\tau):\ \min_{x\in\mathcal{X}}\ \sup_{\mathrm{P}\in\mathbb{D}}\ \mathrm{E_P}[(\kappa-\tilde{\xi}^T\tilde{x})_+^\tau].$$

In general, it is not easy to derive the robust counterparts by using distributional ambiguity sets because of their complex structures. We suppose that only certain features of the distribution P are known as discussed in Section 3.1. Specifically, we assume that P is known to be a mixture of Gaussian distribution. However, some certain kind of inherent uncertainties (for example, uncertainties of mixture proportions) are assumed when modeling a mixture distribution. In particular, we suppose that the $(n+1)$-dimensional random vector $\tilde{\xi}$ follows a mixture distribution with a finite number of known components, and the precise mixing proportions are unknown. In other words, $\tilde{\xi}\sim\sum_{i=1}^d\lambda_i\mathcal{N}(\mu_i,\Sigma_i)$, where $\mathcal{N}(\mu_i,\Sigma_i)$ is the ith known mixture component with normal distribution, and λ_i is the ith unknown mixture proportion.

4.1. ϕ-Divergence

The ϕ-divergence quantifies how one probability distribution is different from a nominal probability distribution. The most popular class of uncertainty sets is based on ϕ-divergence according to Ben-Tal et al. (2013) and is used by many researchers. Such a divergence can be interpreted as "distance" between two distribution functions. Now, we start with the definition of ϕ-divergence.

Definition 1. Suppose that $\phi(t)$ is a convex function on $t\geq 0$. Then the ϕ-divergence between two vectors $p=(p_1,\ldots,p_d)^T\geq 0, q=(q_1,\ldots,q_d)^T\geq 0$ is defined by

$$I_\phi(p,q)=\sum_{i=1}^d q_i\phi(\frac{p_i}{q_i}),\tag{8}$$

where $\phi(1)=0$, $0\phi(a/0):=a\lim_{t\to\infty}\phi(t)/t$ for $a>0$ and $0\phi(0/0):=0$, p denotes the underlying true probability vector, and q denotes the empirical estimation of p.

Table 1
Definitions of some common ϕ-divergences, along with their conjugates $\phi^*(s)$.

Divergence	$\phi(t), t>0$	$I_\phi(p,q)$	$\phi^*(s)$
Kullback–Leibler	$t\log t - t + 1$	$\sum_i p_i \log(\frac{p_i}{q_i})$	$e^s - 1$
Burg entropy	$-\log t + t - 1$	$\sum_i q_i \log(\frac{q_i}{p_i})$	$-\log(1-s), s<1$
J-divergence	$(t-1)\log t$	$\sum_i (p_i - q_i)\log(\frac{p_i}{q_i})$	No-closed form
Likelihood	$-\log t$	$\sum_i q_i \log(\frac{q_i}{p_i})$	$-1-\log(-s), s<0$

The ϕ-divergence is always non-negative with $I_\phi(p,q)=0$ if and only if $p=q$. An important function related to the ϕ-divergence function is its convex conjugate, which is often used in problem reformulations. The conjugate $\phi^*: \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ of ϕ is defined by

$$\phi^*(s) = \sup_{t\geq 0}\{st - \phi(t)\}, \tag{9}$$

Using (8), Ben-Tal et al. (2013) define the uncertainty set V for the given probability vector q as follows

$$V = \{p \in \mathbb{R}^d | I_\phi(p,q) \leq \rho, p^T e = 1, p \geq 0\}, \tag{10}$$

where e denotes the unit vector. Note that V is a convex set since the ϕ-divergence function is convex. The smaller the size of sample data, the larger ρ; Conversely, the larger the size of sample data, the smaller ρ. Different choices of $\phi(\cdot)$ have been extensively investigated in the literature, for details we refer to Pardo (2006) and Ben-Tal et al. (2013). Table 1 shows common choices of the ϕ-divergence function $\phi(\cdot)$ together with the conjugate function $\phi^*(s)$.

4.2. Formulation with ϕ-divergence uncertainty on mixture

Since we have completely specified all the component distributions of the assumed mixture distribution, it is not hard to find that building a region of the original distribution is equivalent to creating a region of the mixture proportions. The most common way to construct an uncertainty set from a ϕ-divergence is to bound the distance to a nominal (e.g., estimated) mixture proportion vector $\hat{\lambda} \in \mathbb{R}^d_+$. Thus, a mixture distribution-based ambiguity set (all the mixture distributions with uncertainty of mixture proportions) can be defined as

$$\mathbb{D}^\phi_M = \{\sum_{i=1}^d \lambda_i \mathcal{N}(\mu_i, \Sigma_i) : \lambda \in \Lambda^\phi\}, \tag{11}$$

where the vector of mixture proportions λ is included in the uncertainty set Λ^ϕ (we call it "ϕ-divergence-type set") :

$$\Lambda^\phi = \{\lambda \in \mathbb{R}^d \mid \lambda \geq 0, \sum_{i=1}^d \lambda_i = 1, I_\phi(\lambda, \hat{\lambda}) \leq \rho\}, \tag{12}$$

where $\rho \geq 0$. We take the uncertainty of proportions into account and reflect the confidence in the estimated proportion vector via a parameter ρ, which controls the level of robustness. It corresponds to the overall amount of scaled deviations of the estimated proportions against which the investor would like to be protected. It is critical for investor to choose the magnitude of uncertainty, ρ, which cannot be set too large because the resulting portfolio will be too conservative. However, if it is too small, one may lose reliable protection. There is always a trade-off for the choice of uncertainty size in the financial context. Specially, when $\rho = 0$, we believe that the estimated proportion vector reflects the true one closely and the uncertainty in λ disappears. As mentioned before, the mixture distribution-based ambiguity set is $\mathbb{D}^\phi_M$ where it is known that the underlying distribution is a mixture distribution with pre-specified mixture components but unspecified mixture proportions. In this paper, hence, the robust strategy will be defined based on the unknown mixture proportions. Three 2-dimension examples for the ϕ-divergence-type set Λ^ϕ with $\rho = 0, 0.05, 0.1, 0.2$ are presented in Fig. 1, where ϕ is the Kullback–Leibler (KL) divergence ($\phi(t) = t\log t - t + 1$). It can also be found from Fig. 1 that Λ^ϕ is symmetric with respect to the straight line $\lambda_1 - \lambda_2 = 0$ when $\hat{\lambda}_1 = \hat{\lambda}_2 = 0.5$.

Based on the above analysis, we suppose that the distribution of $\tilde{\xi}$ is characterized by the mixture of a group of pre-specified distributions with unknown mixture proportions. Besides, we also assume that the credible region of mixture proportions is given as a ϕ-divergence-type uncertainty set Λ^ϕ. Then, the worst-case LPMs index tracking problem (DEIT$_\tau$) with mixture distribution-based ambiguity set $\mathbb{D}^\phi_M$ can be formulated as

$$\min_{x\in\mathcal{X}} \max_{\lambda\in\Lambda^\phi} \lambda^T R^\tau(x), \qquad \tau = 1,2, \tag{13}$$

where $R^\tau(x) := [R^\tau_1(x), R^\tau_2(x), \ldots, R^\tau_d(x)]^T, \tau = 1,2$ and Λ^ϕ is given by (12). Dualizing the inner problem of (13), the robust enhanced index tracking problem (DEIT$_\tau$) ($\tau = 1,2$) with partial information on the mixture distribution specified by (11), can be re-expressed as a deterministic optimization problem. We can derive the robust counterpart of problem (13) by the following theorem.

Theorem 1. *Suppose that $\mathbb{D}$ is defined by mixture-distribution-based ambiguity set $\mathbb{D}^\phi_M$. Then, the robust enhanced index tracking problem (DEIT$_\tau$) is equivalent to the following nonlinear optimization problem over $(x,\theta,\zeta) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_+$*

$$\begin{aligned}&\min_{x,\theta,\zeta} \theta + \zeta\rho + \zeta \sum_{i=1}^d \hat{\lambda}_i \phi^*(\frac{R^\tau_i(x) - \theta}{\zeta}),\\ &\text{s.t. } x \in \mathcal{X}, \zeta \geq 0,\end{aligned} \tag{14}$$

where θ and $\zeta \geq 0$ are the dual variables for constraints in the ϕ-divergence-type set Λ^ϕ, $\phi^*(\cdot)$ is the conjugate function given by (9) and $R^\tau_i (i = 1,\ldots,d, \tau = 1,2)$ are given by (5) and (6).

Proof. Taking the Lagrangian dual of the maximization problem

$$\max_{\lambda\in\Lambda^\phi} \lambda^T R^\tau(x) \tag{15}$$

with dual variables η and $\vartheta \geq 0$, we have

$$L(\lambda,\theta,\zeta) = \sum_{i=1}^d \lambda_i R^\tau_i(x) + \theta(1 - \sum_{i=1}^d \lambda_i) + \zeta(\rho - \sum_{i=1}^d \hat{\lambda}_i \phi(\frac{\lambda_i}{\hat{\lambda}_i})). \tag{16}$$

The dual function is

$$\begin{aligned}\psi(\theta,\zeta) &= \max_{\lambda\geq 0} L(\lambda,\theta,\zeta)\\ &= \theta + \rho\zeta + \max_{\lambda\geq 0}\{\sum_{i=1}^d \lambda_i[R^\tau_i(x) - \theta] - \zeta(\sum_{i=1}^d \hat{\lambda}_i \phi(\frac{\lambda_i}{\hat{\lambda}_i}))\}\\ &= \theta + \rho\zeta + \sum_{i=1}^d \hat{\lambda}_i \zeta \max_{t\geq 0}\{[\frac{R^\tau_i(x) - \theta}{\zeta}]t - \phi(t)\}\\ &= \theta + \rho\zeta + \sum_{i=1}^d \hat{\lambda}_i \zeta \phi^*(\frac{R^\tau_i(x) - \theta}{\zeta}),\end{aligned}$$

where $\phi^*(\cdot)$ is the conjugate of $\phi(\cdot)$ given by (9), with $0\phi^*(\frac{b}{0}) = 0$ if $b \leq 0$ and $0\phi^*(\frac{b}{0}) = 0$ if $b > 0$. So, we have dual form of problem (15)

$$\min_{\theta,\zeta\geq 0} \theta + \rho\zeta + \zeta \sum_{i=1}^d \hat{\lambda}_i \phi^*(\frac{R^\tau_i(x) - \theta}{\zeta}) \tag{17}$$

Similar to Ben-Tal et al. (2013), it can be shown that the problem (17) is a convex optimization problem. When $\rho > 0$, $\hat{\lambda}$ strictly satisfies the ϕ-divergence constraint $I_\phi(\hat{\lambda},\hat{\lambda}) = 0 < \rho$, i.e., the Slater condition holds, then we have strong duality for the Lagrange dual program. Therefore, we have

$$\begin{aligned}\max_{\lambda\in\Lambda_\phi} \lambda^T R^\tau(x) &= \min_{\theta,\zeta\geq 0} \varphi(\theta,\zeta)\\ &= \min_{\theta,\zeta\geq 0}\{\theta + \rho\zeta + \zeta \sum_{i=1}^d \hat{\lambda}_i \phi^*(\frac{R^\tau_i(x) - \theta}{\zeta})\}.\end{aligned} \tag{18}$$

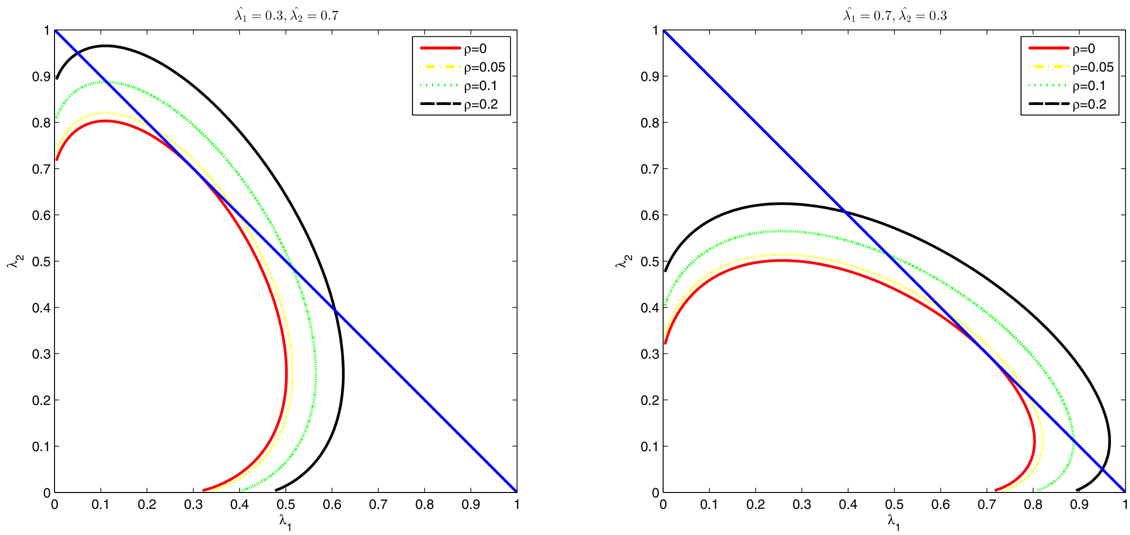

Fig. 1. A visualization of the ϕ-divergence-type set Λ^{ϕ} of the mixture proportions in two 2-dimensional examples with $(\hat{\lambda}_1, \hat{\lambda}_2) = (0.3, 0.7), (0.7, 0.3)$ and $\rho = 0, 0.05, 0.1, 0.2$. The set Λ^{ϕ} is the intersection of the ellipse and the blue line.

Substituting the inner maximization problem of (13) with (18), we derive (14). The proof is completed.

Remark 4. Kullback–Leibler (KL) divergence (also called information divergence or relative entropy), is the most famous special case of the ϕ-divergence measures family (Pardo, 2006). It is widely adopted because of the excellent properties of its conjugate function. Specifically, KL-divergence has closed-form representations of the conjugate ϕ^*. Thus, for KL divergence, the optimization problem (14) is computationally tractable and becomes

$$\min_{x,\theta,\zeta} \ \theta + \zeta\rho + \zeta \sum_{i=1}^{d} \hat{\lambda}_i [\exp(\frac{R_i^{\tau}(x) - \theta}{\zeta}) - 1] \quad (19)$$
$$\text{s.t. } \zeta \geq 0, x \in \mathcal{X}.$$

In our empirical analysis, we choose the commonly used KL-divergence and obtain the robust enhanced tracking portfolio by solving (19) with given ρ and d, where $\hat{\lambda}_i$, μ_i and Σ_i are directly estimated from the real market data by iteratively using the EM algorithm.

5. Empirical analysis

In this section, we carry out empirical analysis on the performance of EIT models by using real market data. The FTSE100 and S&P500 indices are chosen as benchmark index to construct tracking portfolios. We conduct out-of-sample analysis on the performance of the above models.

In real-life situations, it is indeed necessary to consider some trading requirements. For instance, it is important to consider transaction costs as they can affect the return of the investment. In this paper, we consider the proportional transaction cost function associated with a tracking portfolio[3]

$$c(x) = \sum_{i=1}^{n} c_i |x_i - x_i^0|,$$

where c_i is the unit transaction cost for the ith stock and $(x_1^0, \dots, x_n^0)^T$ is the initial position that the investor holds at the beginning. The feasible portfolio set is set as $\mathcal{X} = \{x \in \mathbb{R}^n | e^T x + c(x) = 1, x \geq 0\}$. When $c_i = 0$, our models become a special case that there is no transaction costs. In our empirical analysis, we assume that the unit transaction cost for the ith stocks is $c_i = 0.01$, $i = 1, \dots, n$ and $x_i^0 = 0$, $i = 1, \dots, n$ (i.e., no current portfolio is available).

5.1. Estimation of parameters

A common method of estimating of the parameters of the Gaussian mixture distribution are the maximum empirical likelihood principle. Unfortunately, the maximum likelihood equations for the Gaussian mixture distribution have no closed-form solution due to the composite operation of component-wise product and sum. Hence, estimations of parameters may be found via the usage of the iterative method. The expectation maximization (EM) method is a classical approach to obtain maximum likelihood estimator for Gaussian mixture model (McLachlan & Peel, 2000). Assume that we obtain an independently and identically distributed (i.i.d.) sequence from $\tilde{\xi}$, described as $\tilde{\xi}_1, \dots, \tilde{\xi}_N$. We introduce the following procedure to estimate the parameters of Gaussian mixture distribution.

EM Iteration Process (EMIP):

Step 1: Initialize the means $\mu_i^{(k)}$, covariances $\Sigma_i^{(k)}$ and mixing coefficient $\lambda_i^{(k)}$ $(i = 1, \dots, d)$ based on the sample, and evaluate the initial value of the log likelihood.

Step 2: Expectation Step. Compute

$$\gamma_{ji} = \frac{\lambda_i^{(k)} \varphi(\tilde{\xi}_j | \mu_i^{(k)}, \Sigma_i^{(k)})}{\sum_{i=1}^{d} \lambda_i^{(k)} \varphi(\tilde{\xi}_j | \mu_i^{(k)}, \Sigma_i^{(k)})}, \quad i = 1, \dots, d.$$

Step 3: Maximization Step. Re-estimate $\mu_i^{(k+1)}, \Sigma_i^{(k+1)}$ and $\lambda_i^{(k+1)}$ according to equations

$$\mu_i^{(k+1)} = \frac{1}{N_i} \sum_{j=1}^{N} \gamma_{ji} \tilde{\xi}_j, \quad i = 1, \dots, d,$$

$$\Sigma_i^{(k+1)} = \frac{1}{N_i} \sum_{j=1}^{N} \gamma_{ji} (\tilde{\xi}_j - \mu_i^{(k)})(\tilde{\xi}_j - \mu_i^{(k)})^T, \quad i = 1, \dots, d,$$

$$\lambda_i^{(k+1)} = \frac{N_i}{N}, \quad i = 1, \dots, d,$$

where $N_i = \sum_{j=1}^{N} \gamma_{ji}, i = 1, \dots, d.$

Step 4: Evaluate the logarithm of the likelihood function

$$log L^{(k+1)} = \sum_{j=1}^{N} ln\{\sum_{i=1}^{d} \lambda_i^{(k+1)} \varphi(\tilde{\xi}_j | \mu_i^{(k+1)}, \Sigma_i^{(k+1)})\}$$

[3] We consider the EIT models including proportional transaction costs but not fixed costs. To correctly model the fixed transaction costs, we should introduce binary variable which make the EIT problem becomes a complex mathematical programming problem. Interested readers can refer to Strub and Baumann (2018).

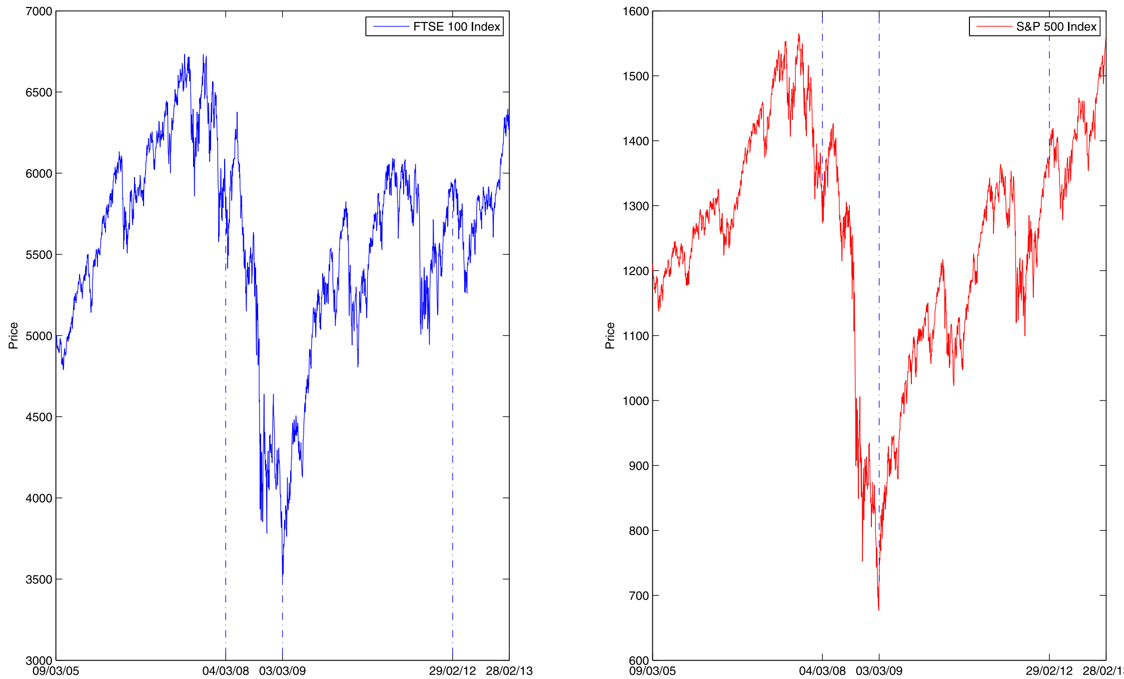

Fig. 2. The price of FTSE100 Index and S&P500 Index from 09/03/2005 to 28/02/2013.

and check for convergence of the log likelihood. Specifically, if the convergence criterion $\|log L^{(k+1)} - log L^{(k)}\| < \varepsilon$ (ε is pre-given tolerance parameter) is full-filled, set $k = k + 1$. If not, go back to Step 2.

The EM technique is a convenient yet powerful approach and will converge to some local maximum of the log likelihood function by repeating the Expectation and Maximization step. We estimate the parameters of the Gaussian mixture distribution by the EM algorithm after specifying the number of components. However, the above EM-based scheme is still a local algorithm which is very sensitive to the initial values of the parameters λ_i, μ_i and Σ_i. To reduce this kind of sensitivity, we utilize K-means algorithm to generate the initial values of the EM method.

Empirical results show that the Gaussian mixture model with three components is a reasonable method to statistically and economically capture the behavior of daily stock returns (Yan & Han, 2019). In subsequent experiments, for simplicity, we take $d = 3$ and presume that the component distributions represent three different market scenarios, i.e., the bear, oscillating and bull markets, respectively. In fact, testing for the number of components in a mixture model is known to be a difficult problem and the risk of overfitting still poses a real challenge. To limit the chances of overfitting as much as possible, we can further consider the Akaike information criterion (AIC) and Bayesian information criterion (BIC), which add to the log likelihood function a penalty proportional to the number of model parameters in order to create a measure of the goodness of fit of a model (Bishop, 2006). Alternatively, when the number of component is so large that overfitting problem is incurred, it is also possible to adopt the MCEM method of Sun et al. (2018) to automatically remove surplus components that provide insufficient contribution for explaining data.

5.2. Out-of-sample analysis

The discussion in this subsection is devoted to comparing and illustrating performance of the proposed portfolio allocation models. We provide some out-of-sample statistics for problems (DEIT$_\tau$) ($\tau = 1, 2$) and problems (EIT$_\tau$) ($\tau = 1, 2$). The robust tracking strategy is achieved by solving the DEIT problem (14) under the estimated distribution with KL-divergence-uncertainty, and the non-robust tracking strategy is the one achieved by solving the EIT problem (2) which uses nominal mixture proportions without considering any uncertainty. It can be regarded as a special case of the robust enhanced index tracking model in which the uncertainty set of the mixture proportions degenerates into a single.

5.2.1. Evaluation methodology

To compare the performances of the method under consideration, we make use of a rolling window procedure similar to Kang et al. (2019), which allows the rebalancing. This approach is more suitable for practical application since it can capture non-stationary market conditions. Specifically:

(1) Choose the estimation window size S ($S < L$), where L is the total number of samples in the data set.

(2) Employ the general EM algorithm to estimate the unknown mixture proportions λ_i ($i = 1, \dots, d$) and parameters of the component distributions, such as mean vectors μ_i and covariance matrices Σ_i ($i = 1, \dots, d$) for the Gaussian distributions.

(3) Compute the portfolio weight vectors x_t for each strategy at time t.

(4) Hold the portfolio x_t for one period and compute the following out-of-sample tracking portfolio return at time $t + 1$:

$$r_{p,t+1} = \xi_{t+1}^T x_t,$$

where ξ_{t+1} denotes the (realized) return rates of the assets.

(5) Repeat step 2 and 3 for the next period by including the next day and dropping the first data point of the estimation window and continue doing this until the end of the data set is reached (Since the data have been updated, the EM algorithm will be reused to further renew the estimation of parameters during the backtesting).

We use an estimation window composed of S sample returns, leaving the last L–S days for the out-of-sample analysis, to compare the different portfolios according to each strategy. The optimal portfolio is first determined by solving model (EIT$_\tau$) or model (DEIT$_\tau$) and using the in-sample S observations. Then, the optimal portfolios is evaluated by observing their behaviors over the $L - S$ days. At the end of this process, hence, we generate $L - S$ portfolio vectors for each strategy which is considered in the analysis. To evaluate the performance of the optimal portfolios, in our analysis, we use average returns (r_{ave} %),

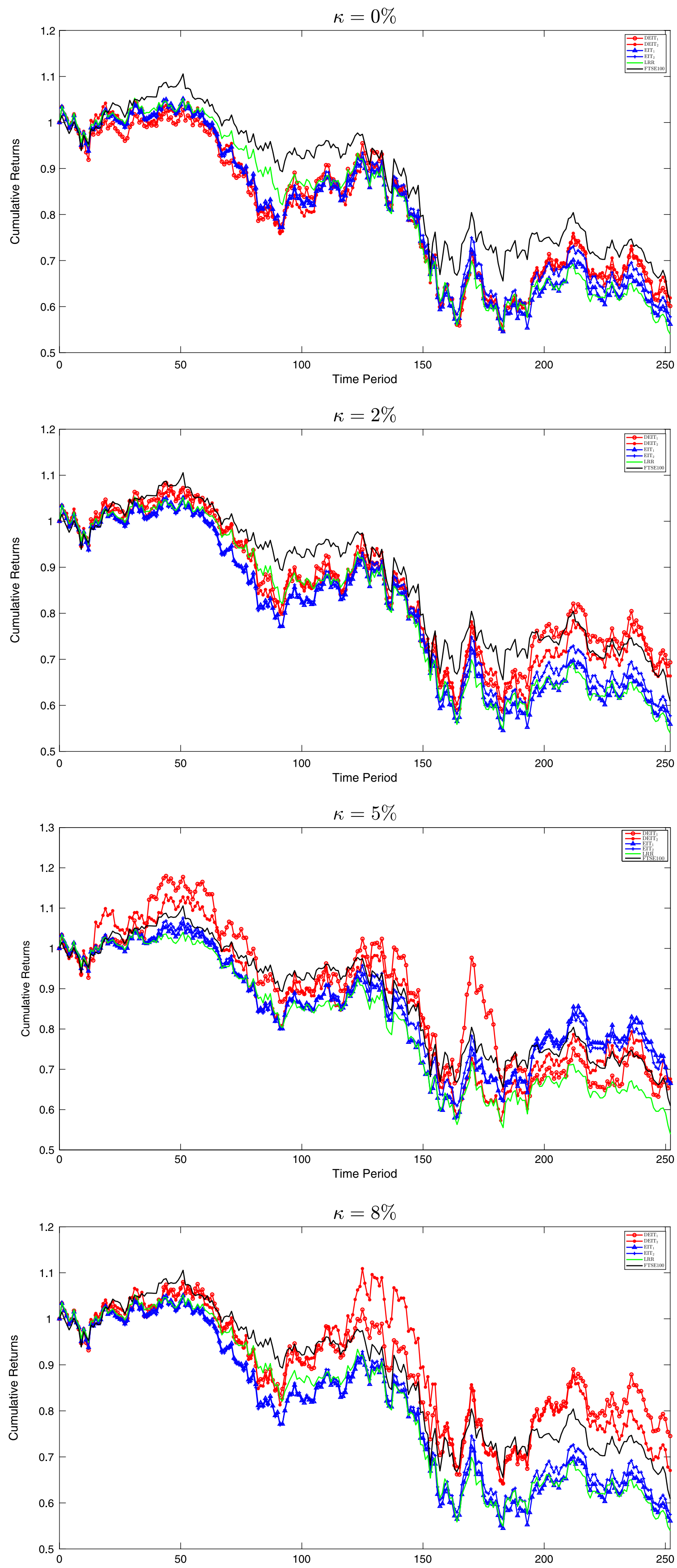

Fig. 3. The out-of-sample comparisons of cumulative returns of models DEIT$_1$, DEIT$_2$, EIT$_1$, EIT$_2$ and LRR in the investment subperiod, March 5th, 2008 to March 3rd, 2009, for the FTSE100 data set.

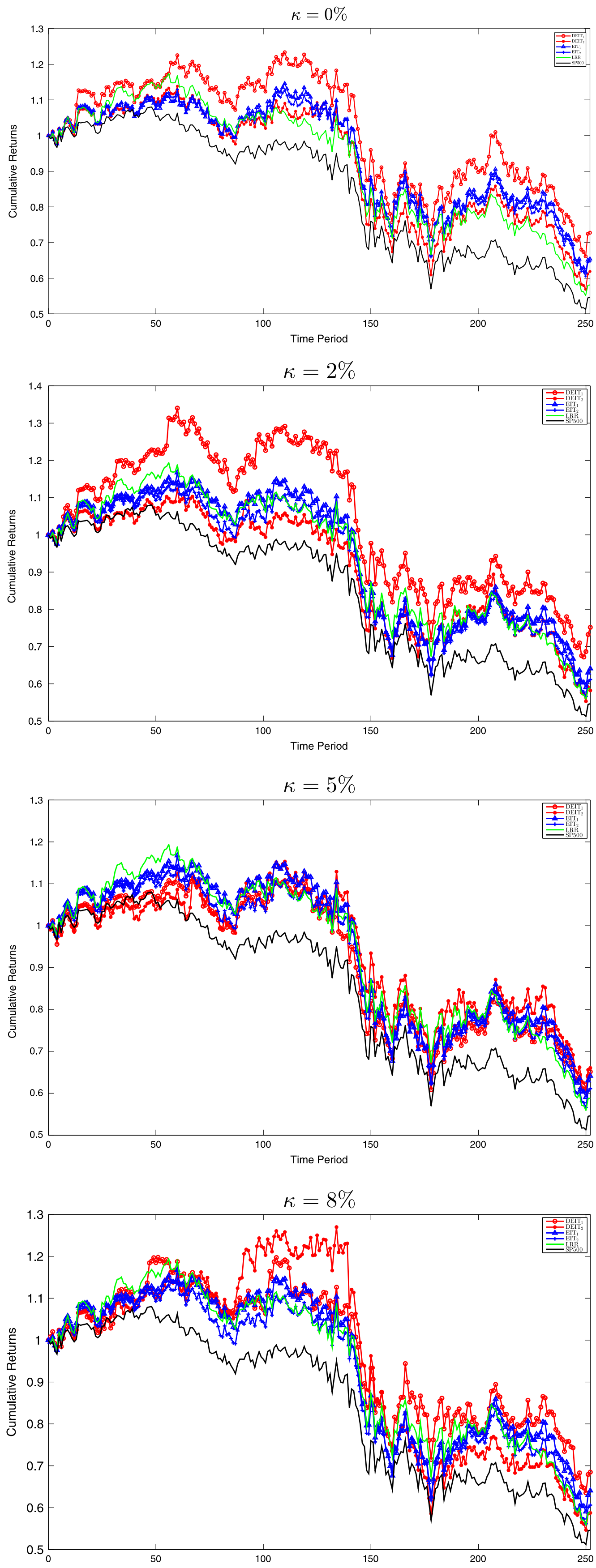

Fig. 4. The out-of-sample comparisons of cumulative returns of models DEIT_1, DEIT_2, EIT_1, EIT_2 and LRR in the investment subperiod, March 5th, 2008 to March 3rd, 2009, for the S&P500 data set.

excess returns (ER %), downside-standard deviation (Dstd), Sharpe ratio (SR) along with Sortino index (SI)[4]:

• $NoD(\%)$: the number of days, divided by $L-S$ and in percentage, that the portfolio returns above the benchmark in the out-of-sample period, i.e., $NoD = \frac{1}{L-S}\sum_{t=S}^{L-1}\mathbf{I}\{r_{p,t} > \xi_{b,t}\}$, where $\mathbf{I}(\cdot)$ is an indicator function.

• $r_{ave}(\%)$: the average return of the portfolio on yearly basis, i.e., $r_{ave} = \frac{1}{L-S}\sum_{t=S}^{L-1} r_{p,t}$;

• ER(Excess Return, %): the average out-of-sample returns achieved above and beyond the index return, in percentage and on yearly basis. It is computed as $ER = \frac{1}{L-S}\sum_{t=S}^{L-1}(r_{p,t} - \xi_{b,t})$;

• SR (Sharpe ratio): the average excess return divided by the portfolio's standard deviation of return Std, i.e., $SR = \frac{ER}{Std}$, where $Std = \sqrt{\frac{1}{L-S}\sum_{t=S}^{L-1}(\xi_{b,t} - r_{p,t})^2}$. The positive and negative deviations from the benchmark index are weighted in the same way since it quantifies return and risk through two-sided type measures;

• $DStd$: the portfolio's downside-standard deviation of return, i.e., $DStd = \sqrt{\frac{1}{L-S}\sum_{t=S}^{L-1}(\xi_{b,t} - r_{p,t})_+^2}$;

• SI (Sortino index or Sortino ratio): the average excess return divided by the downside-standard deviation $DStd$ (Sortino & Satchell, 2001), i.e., $SI = \frac{ER}{DStd}$. It measures the portfolio's average excess return per unit of downside risk incurred and is a modification of the Sharpe ratio (SR) which penalizes only negative deviation from the benchmark index. Compared to the original Sharpe ratio, this variant is a more attractive form to measure risk-adjusted performance.

The above statistics offer a clear assessment of out-of-sample performance of the optimized portfolios. Note that we have expressed the average portfolio return (r_{ave}) and excess return in percentage and on yearly basis for readability, even though they are expressed on daily basis. In theory, the parameter κ is exogenous and investors can set it freely. Here, we set κ equal to 0, 2%, 5% and 8% on a yearly basis. These correspond to values 0, $7.9365 * 10^{-5}$, $1.9841 * 10^{-4}$, $3.1746 * 10^{-4}$, on a daily basis, respectively.

Moreover, we compare the out-of-sample performance of the proposed models with the linear risk-return (LRR) model obtained by using the formulation introduced in Bruni et al. (2015), which maximizes the excess mean return under an upper bound constraint on the minimax risk measure. We use MATLAB R2012a on windows 64 bits Intel Core i5 CPU 2.40 GHz with 8 GB memory processor for solving all EIT problems and also report the running times of the algorithm.

5.2.2. Data sets

The following two data sets of FTSE100 and S&P500 are used in the out-of-sample analysis. Both of data sets are downloaded from Thomson Reuters Datastream and consisting of daily closing prices from 09/03/2005 to 28/02/2013, for a total of 2016 observations. We disregard the constituents with incomplete price data in our tests. Fig. 2 shows the evolution of the price and return of the benchmark for this time frame. Our sample covers the financial crisis of 2008. To capture different behaviors of the market, we divide the daily data into two periods with different starting points and denote them by falling market and rising market. The first sample period is from March. 9rd 2005 to March. 3rd 2009 (corresponding to the falling market time) with March. 9rd 2005–March. 4th 2008 as in-sample period and March. 5th 2008–March. 3rd 2009 as out-of-sample period. The second sample period is from March. 4th 2009 to Feb. 28th 2013 (corresponding to the rising market time) with March. 4th 2009–Feb. 29th 2012 as in-sample period and March. 1st 2012–Feb. 28th 2013 as out-of-sample period. The statistical properties of data of daily return of the FTSE100 Index

[4] Here some indicators similar to those in the literature Guastaroba et al. (2020) are used.

Table 2
Statistics of the daily return of FTSE100 Index and S&P500 Index, including sample size, mean return, Standard deviation (Std), Skewness (Skew), Kurtosis (Kurt). In this table, we divide the daily data into two periods with different starting points and denote them by falling market and rising market. The first 756 daily returns are training samples in the estimated period (Est-period) and the other 256 daily returns are test samples in the investment period (Inv-period). The mean return of FTSE100 Index and S&P500 Index are given in percentage (%). We disregard the constituents with incomplete price data in our tests.

Index	Period		Size	Mean	Std	Skew	Kurt
FTSE100 (falling)	Est-period	09/03/05–04/03/08	756	0.0231	0.0094	−0.3050	6.8340
	Inv-period	05/03/08–03/03/09	252	−0.1686	0.0237	0.2975	6.4748
FTSE100 (rising)	Est-period	04/03/09–29/02/12	756	0.0756	0.0124	0.0431	4.5446
	Inv-period	01/03/12–28/02/13	252	0.0355	0.0087	−0.1505	3.4751
S&P500 (falling)	Est-period	09/03/05–04/03/08	756	0.0155	0.0085	−0.3314	5.1969
	Inv-period	05/03/08–03/03/09	252	−0.2023	0.0274	0.1949	5.8645
S&P500 (rising)	Est-period	04/03/09–29/02/12	756	0.0936	0.0132	−0.1030	6.0190
	Inv-period	01/03/12–28/02/13	252	0.0532	0.0082	0.1013	3.9533

and S&P500 Index are shown in Table 2. They exhibit skewness and high excess kurtosis (Note: the sequences with a kurtosis greater than 3 is regarded as heavy-tailed). In both data sets, we estimate the mixture proportions λ_i and parameters μ_i and Σ_i by a EM iteration process by making use of the in-sample data.

When employing a solution approach for the EIT problems, we begin to construct a tracking portfolio by choosing a subset of stocks that are designed to statistically represent the index. As we known, the beta value for a tracking portfolio is a powerful parameter to explain and describe the risk-return relationship between the tracking portfolio and the benchmark. Thus it is desirable that the beta value of the target index for a tracking portfolio is 1. We follow Huang et al. (2018) and adopt an unbiased beta criterion to identify the optimal subset of assets, i.e., we compute the betas of constituent stocks in the estimation periods and then select stocks with betas close to 1. In our experiments, we choose $n = 10$ stocks from the FTSE100 constituents and $n = 30$ stocks from the S&P500 constituents. In addition, ρ is the same and fixed at 0.05 for all the out-of-sample analysis[5].

5.2.3. Empirical results

This subsection presents the main empirical results. We perform backtesting stated in Section 5.2.1 and test the models DEIT$_\tau$ ($\tau = 1, 2$), EIT$_\tau$ ($\tau = 1, 2$) and LRR along different periods to compare the behavior of the protection both in falling and rising markets. In Tables 1–4, we provide out-of-sample statistics obtained by solving all the optimization models (DEIT$_\tau$, EIT$_\tau$ and LRR) with data sets of the FTSE100 and S&P500 for different values κ. Tables 1 and 2 present results for a falling market and Tables 3 and 4 present results for a rising market.

First, we find that no matter the market is rising or falling, both models of (DEIT$_\tau$) and (EIT$_\tau$) perform better than the model (LRR) in most cases. Next, comparative analysis between models (DEIT$_\tau$) and (EIT$_\tau$) show that the out-of-sample average return from model (DEIT$_\tau$) exceeds that of model (EIT$_\tau$) at almost each κ. The portfolios from model (DEIT$_\tau$) yield better excess returns than those from model (EIT$_\tau$) in all phase for almost all κ, especially in falling market. For data set of the S&P500, the model (DEIT$_\tau$) obtains higher Sortino ratios in falling and rising markets. However, for data set of the FTSE100, the model (DEIT$_\tau$) achieves higher Sortino ratios in falling market, and model (EIT$_\tau$) generates better risk-adjusted measures in rising market in most of the cases. This phenomenon indicates the robust model has better tracking performance during period from March 5th, 2008 to March 3rd, 2009, a period in financial crisis. For data set of the FTSE100, the EIT$_\tau$ model have negative excess return in falling market when taking into account the transaction costs in most instances. As we expected, the introduction of robustness exerts a positive effect on the out-of-sample behavior of the tracking portfolios, i.e., the average values of the Sortino index from model (DEIT$_\tau$) in Tables 3 and 4 are almost better than that of model (EIT$_\tau$). However, there is no formulation that completely dominates the others according to all the computed measures and for all the test instances. Empirical results from out-of-sample using the FTSE100 index and S&P500 index as benchmark demonstrate that the robust counterpart has better tracking performance than that of portfolios determined by non-robust model. It is worth noting that no monotonic behavior seems to show up for the parameter κ and excess returns. Furthermore, there is no clear trend for the statistical values of all test instances with the increase of κ.

With respect to computational time, the models (DEIT$_\tau$) and (EIT$_\tau$) require more running time than the model (LRR) in the vast majority of cases. It can attribute this result to the facts that the models (DEIT$_\tau$) and (EIT$_\tau$) are nonlinear optimization problems. As for the LRR formulation that entails the solution of a LP problem, the computational times are much lower. However, the required computational time of these methods is less than 11 s, which should provide little difficulty in reality(see Tables 5 and 6).

Figs. 3–6 present the evolution of the cumulative returns of the models DEIT$_\tau$ ($\tau = 1, 2$), EIT$_\tau$ ($\tau = 1, 2$) and LRR throughout the out-of-sample period for the FTSE100 and S&P500 indices. Figs. 3 and 4 provide the accumulated returns trends in falling market and Figs. 5 and 6 present the accumulated returns in rising market. Although cumulative return is not one of the evaluation criterions adopted by investors to compare different investment strategies, they present an easy-to-understand description of the dynamic behavior of the portfolios during the whole out-of-sample period. We can observe that all the optimized tracking portfolios closely mimic the performance of the benchmark, i.e., cumulative returns for the portfolios and the benchmark index jointly decrease or increase with the exception of a few ex-post realizations.

Figs. 3 and 4 indicate that, for almost the entire out-of-sample period, the portfolios constructed by models DEIT$_\tau$ ($\tau = 1, 2$) and EIT$_\tau$ ($\tau = 1, 2$) mimic very closely the behavior of the benchmark index, often exceeding the cumulative returns generated by the LRR model (even if not for the whole out-of-sample period). When the market starts to decrease, the DEIT$_\tau$ ($\tau = 1, 2$) portfolio exhibits good performance and dominates the performance of the benchmark index during most of this period of market decline, especially for data set of the S&P500. Moreover, the cumulative returns generated by the EIT$_\tau$ ($\tau = 1, 2$) portfolios are lower than the ones of the benchmark index in some of the realizations. Compared with the model (EIT$_\tau$) and model (LRR), the performance of the portfolios constructed by model (DEIT$_\tau$) is more satisfying in general. From Figs. 5 and 6, during the out-of-sample period, the optimal portfolios from models (DEIT$_\tau$), (EIT$_\tau$) and LRR outperform the benchmark index, generating higher cumulative returns than the ones obtained by the benchmark in rising market. When the S&P500 index increases, both the model (DEIT$_\tau$) and (EIT$_\tau$) increase with the benchmark, but the model (DEIT$_\tau$) increases to a larger extent

[5] We should mention that there is one key parameter, ρ, in the formulation (12), which is crucial in practical implementation. It should be endogenously determined by the data (in a data-driven way, (Kang et al., 2019) based on some statistical principle, rather than arbitrarily exogenous. The choice of ρ is important and remains challenge.

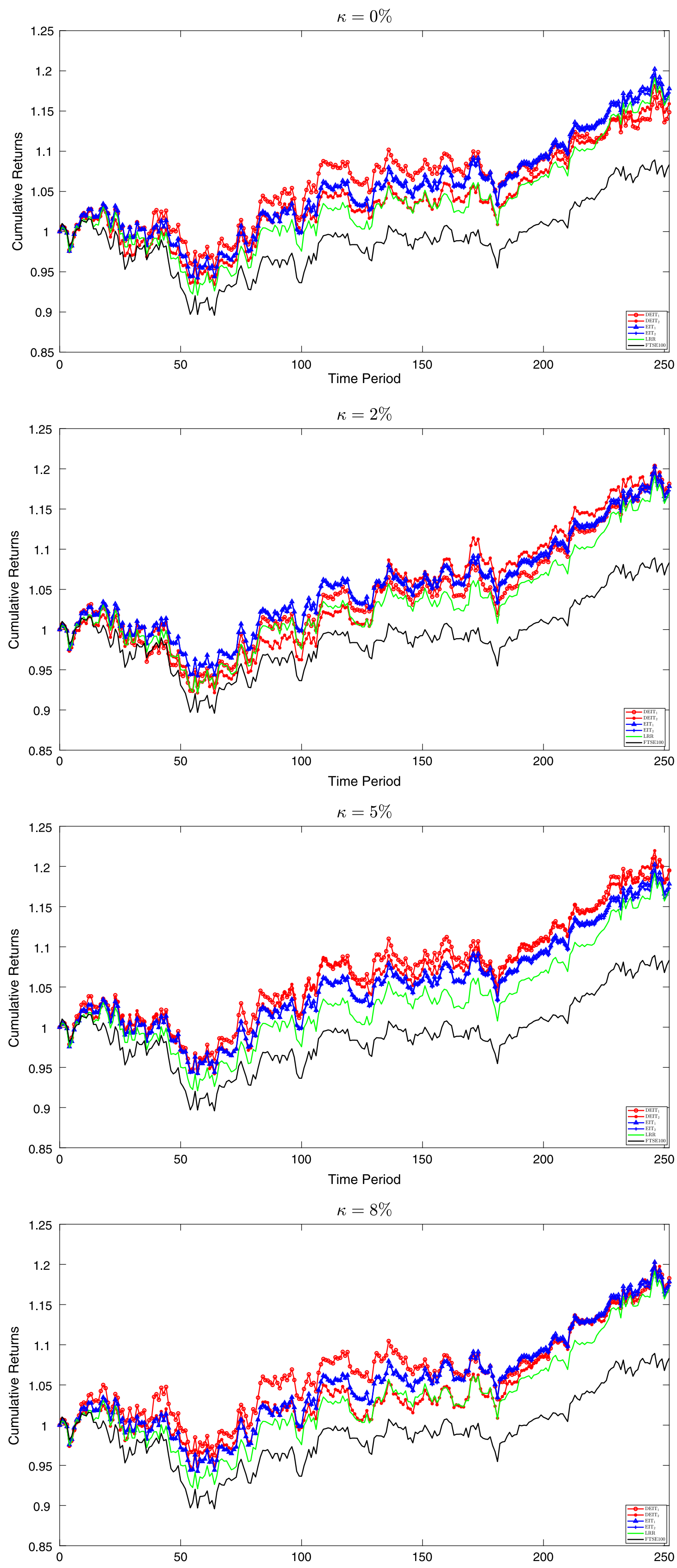

Fig. 5. The out-of-sample comparisons of cumulative returns of models DEIT$_1$, DEIT$_2$, EIT$_1$, EIT$_2$ and LRR in the investment subperiod, March 1st, 2012 to Feb 28th, 2013, for the FTSE100 data set.

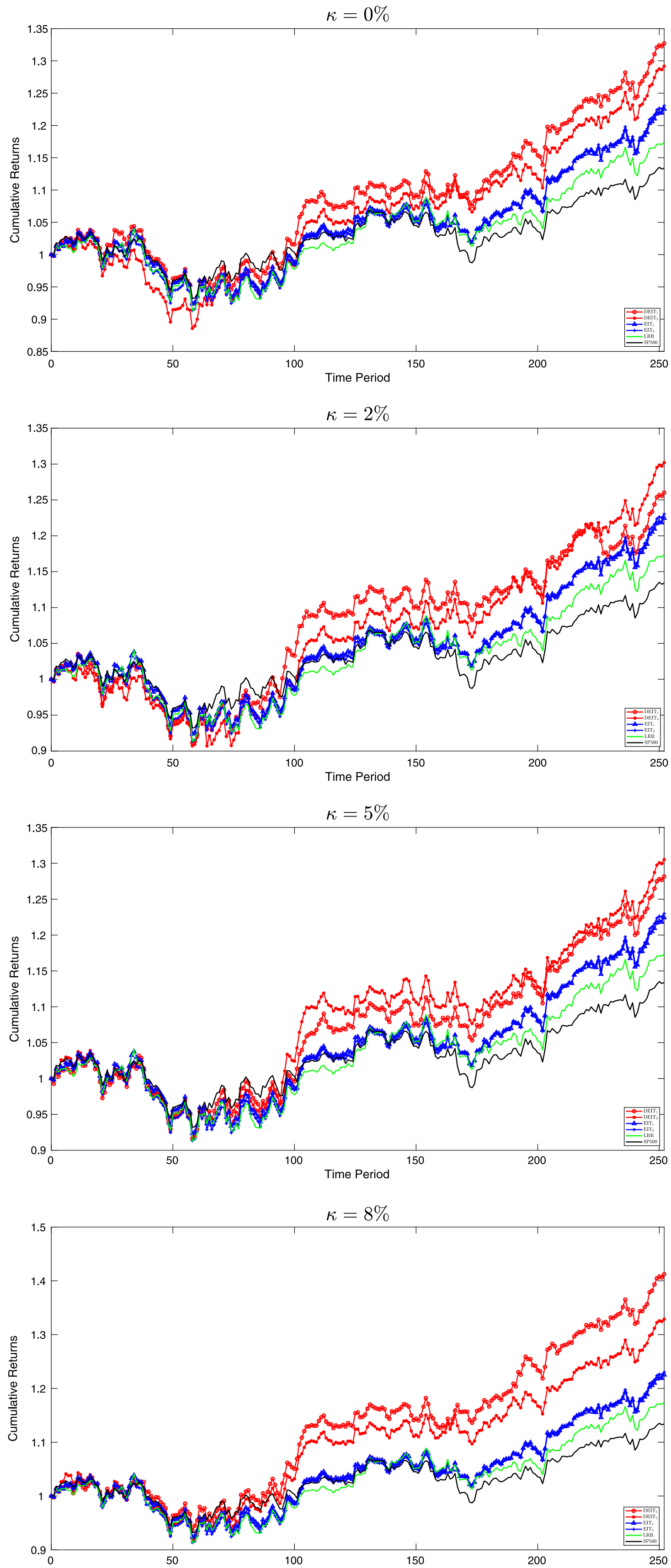

Fig. 6. The out-of-sample comparisons of cumulative returns of models $DEIT_1$, $DEIT_2$, EIT_1, EIT_2 and LRR in the investment subperiod, March 1st, 2012 to Feb 28th, 2013, for the S&P500 data set.

Table 3
Out-of-sample statistics of the different strategies $DEIT_1$, $DEIT_2$, EIT_1, EIT_2, LRR in the investment subperiod, March 5th, 2008 to March 3rd, 2009, for the FTSE100 data set. The average portfolio rate of return r_{ave}(%) and the average out-of-sample excess return of the portfolio over the index return ER(%) are given in percentage and on yearly basis.

FTSE100	Model	NoD(%)	r_{ave}(%)	ER(%)	Dstd	SR	SI	CPU(s)
$\kappa = 0\%$	$DEIT_1$	50.79	−41.76	0.72	0.0101	0.0011	0.0028	6.79
	$DEIT_2$	48.41	−39.29	3.20	0.0092	0.0048	0.0138	6.89
	EIT_1	49.20	−49.60	−7.11	0.0077	−0.0112	−0.0369	0.25
	EIT_2	48.01	−46.64	−4.15	0.0077	−0.0065	−0.0213	0.30
$\kappa = 2\%$	$DEIT_1$	48.41	−28.06	14.42	0.0104	0.0220	0.0552	8.00
	$DEIT_2$	46.82	−32.34	10.15	0.0103	0.0154	0.0389	7.93
	EIT_1	49.20	−50.08	−7.59	0.0077	−0.0120	−0.0393	0.31
	EIT_2	48.01	−46.72	−4.23	0.0077	−0.0066	−0.0217	0.42
$\kappa = 5\%$	$DEIT_1$	51.19	−29.75	12.73	0.0134	0.0173	0.0377	7.34
	$DEIT_2$	48.80	−28.60	13.88	0.0122	0.0190	0.0451	7.88
	EIT_1	53.17	−33.10	9.38	0.0088	0.0151	0.0424	2.48
	EIT_2	50.39	−32.96	9.53	0.0079	0.0151	0.0477	1.52
$\kappa = 8\%$	$DEIT_1$	47.61	−20.33	22.15	0.0102	0.0328	0.0864	7.75
	$DEIT_2$	48.41	−31.28	11.20	0.0098	0.0170	0.0454	8.08
	EIT_1	49.20	−49.76	−7.27	0.0077	−0.0115	−0.0377	0.29
	EIT_2	48.41	−47.16	−4.67	0.0077	−0.0073	−0.0241	0.40
	LRR	44.04	−54.05	−11.56	0.0067	−0.0187	−0.0686	0.05

Table 4
Out-of-sample statistics of the different strategies $DEIT_1$, $DEIT_2$, EIT_1, EIT_2, LRR in the investment subperiod, March 5th, 2008 to March 3rd, 2009, for the S&P500 data set. The average portfolio rate of return r_{ave}(%) and the average out-of-sample excess return of the portfolio over the index return ER(%) are given in percentage and on yearly basis.

S&P500	Model	NoD(%)	r_{ave}(%)	ER(%)	Dstd	SR	SI	CPU(s)
$\kappa = 0\%$	$DEIT_1$	55.15	−21.70	29.30	0.0079	0.0409	0.1464	4.23
	$DEIT_2$	52.38	−38.12	12.88	0.0059	0.0183	0.0860	4.36
	EIT_1	55.55	−32.70	18.30	0.0063	0.0261	0.1149	2.10
	EIT_2	55.55	−33.88	17.11	0.0056	0.0246	0.1223	0.22
$\kappa = 2\%$	$DEIT_1$	56.34	−18.25	32.75	0.0087	0.0454	0.1498	3.71
	$DEIT_2$	49.60	−43.53	7.45	0.0086	0.0102	0.0344	3.69
	EIT_1	54.36	−35.10	15.90	0.0061	0.0232	0.1028	0.43
	EIT_2	52.38	−39.81	11.16	0.0052	0.0163	0.0848	0.13
$\kappa = 5\%$	$DEIT_1$	54.36	−32.59	18.41	0.0069	0.0256	0.1052	3.21
	$DEIT_2$	53.57	−31.09	19.95	0.0096	0.0275	0.0830	3.52
	EIT_1	54.36	−35.08	15.97	0.0061	0.0233	0.1032	0.71
	EIT_2	52.38	−39.91	11.09	0.0052	0.0162	0.0843	0.12
$\kappa = 8\%$	$DEIT_1$	51.58	−26.54	24.45	0.0093	0.0323	0.1048	4.51
	$DEIT_2$	52.38	−42.70	8.29	0.0083	0.0115	0.0398	3.94
	EIT_1	54.36	−35.13	15.86	0.0061	0.0231	0.1025	0.81
	EIT_2	52.38	−40.68	10.32	0.0052	0.0150	0.0786	0.16
	LRR	51.58	−43.83	7.16	0.0068	0.0105	0.0420	0.95

Table 5
Out-of-sample statistics of the different strategies $DEIT_1$, $DEIT_2$, EIT_1, EIT_2, LRR in the investment subperiod, March 1st, 2012 to Feb 28th, 2013, for the FTSE100 data set. The average portfolio rate of return r_{ave}(%) and the average out-of-sample excess return of the portfolio over the index return ER(%) are given in percentage and on yearly basis.

FTSE100	Model	NoD(%)	r_{ave}(%)	ER(%)	Dstd	SR	SI	CPU(s)
$\kappa = 0\%$	$DEIT_1$	51.19	14.86	5.90	0.0030	0.0260	0.0774	9.78
	$DEIT_2$	55.55	15.74	6.78	0.0024	0.0305	0.0999	8.91
	EIT_1	55.55	17.35	8.39	0.0024	0.0379	0.1365	0.58
	EIT_2	54.36	16.88	7.91	0.0039	0.0359	0.1311	0.23
$\kappa = 2\%$	$DEIT_1$	55.55	17.67	8.71	0.0029	0.0388	0.1178	10.66
	$DEIT_2$	54.36	17.47	8.51	0.0029	0.0371	0.1150	10.18
	EIT_1	55.55	17.36	8.40	0.0024	0.0380	0.1368	0.66
	EIT_2	54.36	16.89	7.93	0.0024	0.0360	0.1313	0.26
$\kappa = 5\%$	$DEIT_1$	53.17	18.83	9.87	0.0028	0.0436	0.1409	10.40
	$DEIT_2$	55.95	18.85	9.89	0.0028	0.0437	0.1414	10.20
	EIT_1	55.55	17.38	8.41	0.0024	0.0380	0.1371	0.62
	EIT_2	54.36	16.89	7.92	0.0024	0.0360	0.1312	0.67
$\kappa = 8\%$	$DEIT_1$	55.95	17.81	8.84	0.0029	0.0389	0.1217	8.85
	$DEIT_2$	55.95	17.30	8.34	0.0029	0.0368	0.1152	7.15
	EIT_1	55.55	17.40	8.43	0.0024	0.0381	0.1375	0.62
	EIT_2	54.36	16.96	7.99	0.0024	0.0363	0.1323	0.27
	LRR	52.38	16.71	7.74	0.0030	0.0333	0.1032	0.06

Table 6
Out-of-sample statistics of the different strategies $DEIT_1$, $DEIT_2$, EIT_1, EIT_2, LRR in the investment subperiod, March 1st, 2012 to Feb 28th, 2013, for the S&P500 data set. The average portfolio rate of return $r_{ave}(\%)$ and the average out-of-sample excess return of the portfolio over the index return ER(%) are given in percentage and on yearly basis.

S&P500	Model	NoD(%)	$r_{ave}(\%)$	ER(%)	Dstd	SR	SI	CPU(s)
$\kappa = 0\%$	$DEIT_1$	56.74	29.42	16.01	0.0032	0.0677	0.1958	4.29
	$DEIT_2$	55.15	26.75	13.34	0.0031	0.0558	0.1706	4.43
	EIT_1	53.96	21.37	7.96	0.0024	0.0348	0.1332	1.63
	EIT_2	56.34	21.80	8.39	0.0022	0.0366	0.1485	0.26
$\kappa = 2\%$	$DEIT_1$	52.77	24.29	10.88	0.0035	0.0451	0.1222	4.35
	$DEIT_2$	55.15	27.60	14.19	0.0033	0.0575	0.1703	4.49
	EIT_1	53.57	21.30	7.89	0.0024	0.0345	0.1321	1.68
	EIT_2	56.34	21.76	8.35	0.0022	0.0364	0.1479	0.36
$\kappa = 5\%$	$DEIT_1$	55.15	25.95	12.54	0.0035	0.0525	0.1493	5.37
	$DEIT_2$	51.58	27.86	14.45	0.0038	0.0584	0.1509	4.40
	EIT_1	54.36	21.28	7.87	0.0024	0.0344	0.1320	1.97
	EIT_2	56.34	21.75	8.34	0.0024	0.0364	0.1477	0.39
$\kappa = 8\%$	$DEIT_1$	56.34	35.69	22.28	0.0045	0.0924	0.1963	3.20
	$DEIT_2$	54.76	29.57	16.16	0.0033	0.0675	0.1917	3.99
	EIT_1	55.15	21.35	7.94	0.0024	0.0346	0.1329	1.65
	EIT_2	56.34	21.77	8.36	0.0022	0.0365	0.1480	0.29
	LRR	52.38	17.14	3.73	0.0029	0.0158	0.0510	0.07

with better accumulated returns. It is worthy highlighting that there may not exist a significant difference between the performances of the robust and non-robust portfolios in the rising market for data set of the S&P500. We acknowledge that there are many other options for tracking the portfolios, and we are not going to indicate that our method is the best way for this issue. On the contrary, our purpose is to demonstrate that our approach may provide a valuable reference for (enhanced) index tracking.

6. Conclusions

Gaussian mixture models have been widely used in various areas. In this paper we use index-plus-alpha approach to develop the EIT model in LPM framework. Furthermore, we formulate robust EIT problems with ambiguity, in the sense that the stocks and random benchmark return are specified by a mixture distribution model where certain inherent uncertainty is permitted. Specifically, a region of mixture proportions is created by a ϕ-divergence-type set. Our empirical results with FTSE100 and S&P500 data sets show that the model ($DEIT_\tau$) obtain better excess returns and sortino indices, and the model (EIT_τ) has the lowest standard deviation among EIT strategies. Although we have obtained some preliminary conclusions about the (robust) EIT model with mixture distribution under a LPM framework, there is a lack of comprehensive comparative studies with most of the existing approaches for the EIT problem in terms of effectiveness of their models. It remains an important issue to be further studied.

CRediT authorship contribution statement

Zhilin Kang: Conceptualization, Methodology, Writing – original draft, Software, Writing – review & editing. **Haixiang Yao:** Methodology, Software, Formal analysis, Writing – review & editing. **Xingyi Li:** Formal analysis, Writing – original draft, Writing – review & editing. **Zhongfei Li:** Conceptualization, Methodology, Writing – original draft, Formal analysis, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China [grant numbers 71721001, 71991474, 71871071, 72071051]; the Ministry of Education of Humanities and Social Science Project, China [grant number 20YJA790011], the National Statistical Science Research Program, China [grant number 2021LY026], the Social Science Foundation of Fujian Province of China [grant number FJ2021BF009], the Natural Science Foundation of Guangdong Province of China [grant number 2018B030311004] and the State Scholarship Fund of China [grant number 202106380104]. We thank Prof. Binshan Lin, the associate editor and three anonymous referees for their effort and valuable comments/suggestions.

References

Beasley, J. E., Meade, N., & Chang, T. J. (2003). An evolutionary heuristic for the index tracking problem. *European Journal of Operational Research*, *148*(3), 621–643. http://dx.doi.org/10.1016/S0377-2217(02)00425-3.

Bellalah, M., & Lavielle, M. (2002). A decomposition of empirical distributions with applications to the valuation of derivative assets. *Multinational Finance Journal*, *6*(2), 99–130. http://dx.doi.org/10.17578/6-2-2.

Ben-Tal, A., Den Hertog, D., De Waegenaere, A., Melenberg, B., & Rennen, G. (2013). Robust solutions of optimization problems affected by uncertain probabilities. *Management Science*, *59*(2), 341–357. http://dx.doi.org/10.1287/mnsc.1120.1641.

Bishop, C. M. (2006). *Pattern recognition and machine learning* (pp. 32–33, 216–217). Springer.

Bruni, R., Cesarone, F., Scozzari, A., & Tardella, F. (2015). A linear risk-return model for enhanced indexation in portfolio optimization. *OR Spectrum*, *37*(3), 735–759. http://dx.doi.org/10.1007/s00291-014-0383-6.

Canakgoz, N. A., & Beasley, J. E. (2008). Mixed-integer programming approaches for index tracking and enhanced indexation. *European Journal of Operational Research*, *196*(1), 384–399. http://dx.doi.org/10.1016/j.ejor.2008.03.015.

Chen, C., & Kwon, R. H. (2012). Robust portfolio selection for index tracking. *Computers & Operations Research*, *39*(4), 829–837. http://dx.doi.org/10.1016/j.cor.2010.08.019.

de Paulo, W. L., de Oliveira, E. M., & do Valle Costa, O. L. (2016). Enhanced index tracking optimal portfolio selection. *Finance Research Letters*, *16*, 93–102. http://dx.doi.org/10.1016/j.frl.2015.10.005.

Filippi, C., Guastaroba, G., & Speranza, M. G. (2016). A heuristic framework for the bi-objective enhanced index tracking problem. *Omega*, *65*, 122–137. http://dx.doi.org/10.1016/j.omega.2016.01.004.

Gorissen, B., Yanikoglu, I., & Hertog, D. (2015). A practical guide to robust optimization. *Omega*, *53*, 124–137. http://dx.doi.org/10.1016/j.omega.2014.12.006.

Guastaroba, G., Mansini, R., Ogryczak, W., & Speranza, M. G. (2016). Linear programming models based on omega ratio for the enhanced index tracking problem. *European Journal of Operational Research*, *251*(3), 938–956. http://dx.doi.org/10.1016/j.ejor.2015.11.037.

Guastaroba, G., Mansini, R., Ogryczak, W., & Speranza, M. G. (2020). Enhanced index tracking with CVaR-based measures. *Annals of Operations Research*, *292*, 883–931. http://dx.doi.org/10.1007/s10479-020-03518-7.

Guastaroba, G., & Speranza, M. G. (2012). Kernel search: An application to the index tracking problem. *European Journal of Operational Research*, *217*(1), 54–68. http://dx.doi.org/10.1016/j.ejor.2011.09.004.

Hanasusanto, G. A., Kuhn, D., Wallace, S. W., & Zymler, S. (2015). Distributionally robust multi-item newsvendor problems with multimodal demand distributions. *Mathematical Programming*, *152*(1–2), 1–32. http://dx.doi.org/10.1007/s10107-014-0776-y.

Hu, Z., Sun, W., & Zhu, S. (2018). Chance constrained programs with mixture distributions. http://www.optimization-online.org/DB_FILE/2018/09/6798.pdf.

Huang, J., Li, Y., & Yao, H. (2018). Index tracking model, downside risk and non-parametric kernel estimation. *Journal of Economic Dynamics and Control*, *92*, 103–128. http://dx.doi.org/10.1016/j.jedc.2018.04.008.

Kang, Z. L., Li, X., Li, Z. F., & Zhu, S. (2019). Data-driven robust mean-CVaR portfolio selection under distribution ambiguity. *Quantitative Finance*, *19*(1), 105–121. http://dx.doi.org/10.1080/14697688.2018.1466057.

Kapsos, M., Christofides, N., & Rustem, B. (2014). Worst-case robust Omega ratio. *European Journal of Operational Research*, *234*(2), 499–507. http://dx.doi.org/10.1016/j.ejor.2013.04.025.

Khoshabar, N. A., Salahi, M., Lotfi, S., & Abdelouahed, H. (2020). Robust index-tracking and enhanced index-tracking in portfolio optimization. *Studies of Applied Economics*, *38*(1), http://dx.doi.org/10.25115/eae.v38i1.2955.

Konno, H., & Hatagi, T. (2005). Index-plus-alpha tracking under concave transaction cost. *Journal of Industrial & Management Optimization*, *1*(1), 87–98. http://dx.doi.org/10.3934/jimo.2005.1.87.

Koshizuka, T., Konno, H., & Yamamoto, R. (2009). Index-plus-alpha tracking subject to correlation constraint. *International Journal of Optimization: Theory, Methods and Applications*, *1*(2), 215–224.

Kwak, Y., Song, J., & Lee, H. (2021). Neural network with fixed noise for index-tracking portfolio optimization. *Expert Systems with Applications*, Article 115298. http://dx.doi.org/10.1016/j.eswa.2021.115298.

Kwon, R. H., & Wu, D. (2017). Factor-based robust index tracking. *Optimization and Engineering*, *18*(2), 443–466. http://dx.doi.org/10.1007/s11081-016-9314-5.

Lejeune, M. A. (2012). Game theoretical approach for reliable enhanced indexation. *Decision Analysis*, *9*(2), 146–155. http://dx.doi.org/10.1287/deca.1120.0239.

Li, Q., Sun, L., & Bao, L. (2011). Enhanced index tracking based on multi-objective immune algorithm. *Expert Systems with Applications*, *38*(5), 6101–6106. http://dx.doi.org/10.1016/j.eswa.2010.11.001.

Ling, A., Sun, J., & Yang, X. (2014). Robust tracking error portfolio selection with worst-case downside risk measures. *Journal of Economic Dynamics and Control*, *39*, 178–207. http://dx.doi.org/10.1016/j.jedc.2013.11.011.

Marron, J. S., & Wand, M. P. (1992). Exact mean integrated squared error. *The Annals of Statistics*, *20*(2), 712–736. http://dx.doi.org/10.1214/aos/1176348653.

McLachlan, G. J., & Peel, D. (2000). *Finite mixture models*. NewYork: John Wiley & Sons.

McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative risk management: Concepts, techniques and tools*. Princeton: Princeton University Press.

Paolella, M. S. (2015). Multivariate asset return prediction with mixture models. *The European Journal of Finance*, *21*(13–15), 1214–1252. http://dx.doi.org/10.1080/1351847X.2012.760167.

Pardo, L. (2006). *Statistical inference based on divergence measures*. Taylor & Francis Group, LLC.

Roll, R. (1992). A mean/variance analysis of tracking error. *The Journal of Portfolio Management*, *18*, 13–22. http://dx.doi.org/10.3905/jpm.1992.70192.

Rudolf, M., Wolter, H. J., & Zimmermann, H. (1999). A linear model for tracking error minimization. *Journal of Banking & Finance*, *23*(1), 85–103. http://dx.doi.org/10.1016/S0378-4266(98)00076-4.

Sant'Anna, L. R., Filomena, T. P., Guedes, P. C., & Borenstein, D. (2017). Index tracking with controlled number of assets using a hybrid heuristic combining genetic algorithm and non-linear programming. *Annals of Operations Research*, *258*(2), 849–867. http://dx.doi.org/10.1007/s10479-016-2111-x.

Sehgal, R., & Mehra, A. (2019). Enhanced indexing using weighted conditional value at risk. *Annals of Operations Research*, *280*(1–2), 211–240. http://dx.doi.org/10.1007/s10479-019-03132-2.

Sortino, F., & Satchell, S. (2001). *Managing downside risk in financial markets*. Oxford: Batterworth-Heinemann Finance, http://dx.doi.org/10.1016/B978-075064863-9.50001-3.

Strub, O., & Baumann, P. (2018). Optimal construction and rebalancing of index-tracking portfolios. *European Journal of Operational Research*, *264*(1), 370–387. http://dx.doi.org/10.1016/j.ejor.2017.06.055.

Sun, W., Hu, Z., & Hong, L. J. (2018). Gaussian mixture model-based random search for continuous optimization via simulation. In *2018 winter simulation conference* (pp. 2003–2014). IEEE, http://dx.doi.org/10.1109/WSC.2018.8632380.

Wang, J., & Taaffe, M. R. (2015). Multivariate mixtures of normal distributions: properties, random vector generation, fitting, and as models of market daily changes. *INFORMS Journal on Computing*, *27*(2), 193–203. http://dx.doi.org/10.1287/ijoc.2014.0616.

Wu, D., & Wu, D. D. (2019). An enhanced decision support approach for learning and tracking derivative index. *Omega*, *88*, 63–76. http://dx.doi.org/10.1016/j.omega.2018.10.021.

Xu, F., Wang, M., Dai, Y. H., & Xu, D. (2018). A sparse enhanced indexation model with chance and cardinality constraints. *Journal of Global Optimization*, *70*(1), 5–25. http://dx.doi.org/10.1007/s10898-017-0513-1.

Yan, H., & Han, L. (2019). Empirical distributions of stock returns: Mixed normal or kernel density? *Physica A: Statal Mechanics and Its Applications*, *514*, 473–486. http://dx.doi.org/10.1016/j.physa.2018.09.080.

Yao, H., Huang, J. B., Li, Y., & Humphrey, J. (2021). A general approach to smooth and convex portfolio optimization using lower partial moments. *Journal of Banking & Finance*, Article 106167.

Yu, L., Zhang, S. Z., & Zhou, X. Y. (2006). A downside risk analysis based on financial index tracking models. In *Stochastic finance, (Chapter 8)* (pp. 213–236). http://dx.doi.org/10.1007/0-387-28359-5.

Zhu, S., Fan, M., & Li, D. (2014). Portfolio management with robustness in both prediction and decision: A mixture model based learning approach. *Journal of Economic Dynamics and Control*, *48*, 1–25. http://dx.doi.org/10.1016/j.jedc.2014.08.015.

Zhu, S., & Fukushima, M. (2009). Worst-case conditional value-at-risk with application to robust portfolio management. *Operations Research*, *57*(5), 1155–1168. http://dx.doi.org/10.1287/opre.1080.0684.

Zhu, S., Li, D., & Wang, S. (2009). Robust portfolio selection under downside risk measures. *Quantitative Finance*, *9*(7), 869–885. http://dx.doi.org/10.1080/14697680902852746.