2014

Tracking-error models for multiple benchmarks: Theory and empirical performance

Yunchao XU
New York University

Zhichao ZHENG
Singapore Management University, DANIELZHENG@smu.edu.sg

Karthik NATARAJAN
Singapore University of Technology and Design

Chung-Piaw TEO
National University of Singapore

Tracking Error Models for Multiple Benchmarks

Yunchao Xu* Zhichao Zheng† Karthik Natarajan‡ Chung-Piaw Teo§

June 5, 2013

Abstract

We propose a new multiple-benchmark tracking-error model for portfolio selection problem. The tracking error of a portfolio from a set of benchmark portfolios is defined as the difference between its return and the highest return from the set of benchmarks. We derive closed-form solution of our portfolio strategy, whose main component is the sum of the benchmark portfolios weighted by their respective probabilities of attaining the highest return among the portfolios in the benchmark. These probabilities, also known as the persistency values, are less sensitive to estimation errors in the means and covariances. These features help to stabilize the computational performance of our portfolio strategy against estimation errors.

We use the proposed model to address several pertinent issues in active portfolio management: (1) What are the benefits in tracking performance of multiple benchmarks? We demonstrate that under suitable conditions, multiple benchmarks tracking error model can actually produce portfolio strategy that has less variability in portfolio returns, compared to the portfolio strategy constructed using single benchmark model, given a fixed target rate of returns. This addresses the agency issue in this problem, as portfolio managers are more concerned with variability of the excess returns above the benchmark, whereas the investors are more concerned with the variability of the total returns. (2) How and when to rebalance the portfolio allocation when prices and asset returns change over time, taking into account transaction cost? We show that our model can control for transaction cost by adding the buy-and-hold strategy into the set of benchmark portfolios. This approach reduces drastically the transaction volume of several popular static portfolio rules executed dynamically over time.

Last but not least, we perform comprehensive numerical experiments with various empirical data sets to demonstrate tha our approach can consistently provide higher net Sharpe ratio (after accounting for transaction cost), higher net aggregate return, and lower turnover rate, compared to ten different benchmark portfolios proposed in the literature, including the equally weighted portfolio (the $1/N$ strategy).

*Department of Decision Sciences, NUS Business School, National University of Singapore, Singapore. Email: bizxuy@nus.edu.sg. Part of this work was done when the author was at the Department of Mathematics, Faculty of Science, National University of Singapore, Singapore.

†Department of Decision Sciences, NUS Business School, National University of Singapore, Singapore. Email: zhichao@nus.edu.sg

‡Department of Engineering Systems and Design, Singapore University of Technology and Design, Singapore. Email: natarajan_karthik@sutd.edu.sg

§Department of Decision Sciences, NUS Business School, National University of Singapore, Singapore. Email: bizteocp@nus.edu.sg

1 Introduction

In practice, an institutional investor often evaluates the performance of a portfolio manager against a benchmark (e.g., an index fund). In active portfolio management, the portfolio manager makes specific investment with the goal of outperforming the benchmark, as his fees and compensations are directly linked to the excess returns above the benchmark. For a given target rate of returns, the portfolio manager would often seek to minimize the volatility of the deviation of the portfolio return from the benchmark return, i.e., the tracking-error volatility. The portfolio-selection models that minimize the tracking-error volatility are referred to as benchmark tracking-error models (cf. Roll (1992) and Jorion (2003)).

Roll (1992) investigated the benchmark tracking-error model that minimizes the tracking-error volatility subject to the full investment constraint and the constraint on target expected return, i.e.,

$$\min_{\boldsymbol{e}^T\boldsymbol{x}=1,\boldsymbol{\mu}^T\boldsymbol{x}=K} \mathbf{E}\left[\left(\tilde{\boldsymbol{r}}^T\boldsymbol{p} - \tilde{\boldsymbol{r}}^T\boldsymbol{x}\right)^2\right],$$

where $\tilde{\boldsymbol{r}} \in \mathbb{R}^n$ is the random return vector of the financial assets; $\boldsymbol{\mu}$ is the expected return of the assets, i.e., $\boldsymbol{\mu} = \mathbf{E}[\tilde{\boldsymbol{r}}]$; $\boldsymbol{e}$ is the column vector with all entries equal to 1; K is the target expected return; $\boldsymbol{p}$ is the benchmark portfolio; and $\boldsymbol{x} \in \mathbb{R}^n$ is the investment decision.

The deficiency with the tracking error approach is however well known. Roll (1992) observed that the optimal trading decision $(\boldsymbol{x} - \boldsymbol{p})$ does not depend on the benchmark at all. Furthermore, with this setup the portfolio manager will focus solely on the tracking-error volatility but ignore the total portfolio risk. This creates an agency problem, since the investor is more concerned with the latter. The tracking error model may thus produce seriously inefficient portfolios for the investor. To address these issues, Roll (1992) proposed to constrain the portfolio's beta; Jorion (2003) proposed to constrain the portfolio's total variance; and Alexander & Baptista (2008) proposed to constrain the portfolio's Value-at-Risk (VaR). However, most of these proposals are difficult to implement in practice, and do not address directly the connection with the benchmark based approach to portfolio management.

The choice of the proper benchmark is also a problem in practice. Poor active portfolio management could lead to less than perfectly diversified portfolio, and incur heavy transaction costs and assumes high total portfolio risk. El-Hassan & Kofman (2003) observed from their empirical analysis that in reality, the selected benchmark is often inefficient, and its expected return could fall below the expected return of the well-known minimum-variance portfolio. The immediate consequence is that during bear market conditions the benchmark tracking-error models will call for a huge amount of short selling, which can substantially increase the total portfolio risk. This problem is compounded by the fact that tracking error measurement does not differentitate between over-performing and under-performing vis-a-viz the benchmark portfolio, and hence the performance of the tracking error model can be adversely affected by a poorly selected benchmark.

To mitigate this problem of finding the right unique benchmark, one natural solution is to use multiple benchmarks to evaluate the performance of a portfolio manager. By choosing benchmarks that can counter-balance the performance of each other in different market environments, we can track the performance of our portfolio strategy in a more accurate and reliable manner. The literature on multiple-benchmark tracking-error strategy is however comparatively sparse. Wang (1999) extended the single-benchmark tracking-error model to track multiple benchmarks simultaneously. The tracking error of the portfolio with respect to a set of m benchmarks is defined by a weighted sum of single-benchmark tracking errors, i.e.,

$$\min_{\boldsymbol{e}^T\boldsymbol{x}=1,\boldsymbol{\mu}^T\boldsymbol{x}=K} \sum_{j=1}^{m} w_j \mathbf{E}\left[\left(\tilde{\boldsymbol{r}}^T\boldsymbol{p}^j - \tilde{\boldsymbol{r}}^T\boldsymbol{x}\right)^2\right],$$

where w_j is the weight on the tracking error of the j^{th} benchmark portfolio, $\boldsymbol{p}^j$, $j = 1, \ldots, m$. Rustem & Howe (2002) considered an alternative model. Their objective is to minimize the maximum tracking-error volatility across all benchmarks, i.e.,

$$\min_{\boldsymbol{e}^T\boldsymbol{x}=1,\boldsymbol{\mu}^T\boldsymbol{x}=K} \max_{j\in\{1,\ldots,m\}} \mathbf{E}\left[\left(\tilde{\boldsymbol{r}}^T\boldsymbol{p}^j - \tilde{\boldsymbol{r}}^T\boldsymbol{x}\right)^2\right].$$

However, it is not clear how a portfolio manager should choose the weights in Wang's model. The minimax approach, on the other hand, is often considered to be too conservative. A more critical issue is that these models still rely on the evaluation of single-benchmark tracking error and only combine them in the aggregate level. They fail to distinguish between over-performing and under-performing vis-a-viz the selected benchmarks, and that the performance of the different benchmarks may be correlated. Hence they do not fully capture the concerns arising from the real investment activities as discussed above.

In this paper, we propose a new class of tracking-error models for multiple benchmarks. This problem arises naturally when multiple natural benchmarks (e.g., risk-free returns, S&P 500 index etc.) are readily available in the market that can be used to evaluate the performance of the portfolio managers. It also arises when the portfolio manager is managing funds for different clients, each with a unique benchmark that will be used to evaluate the performance of the manager. Instead of managing different pools of funds, one for each client, we explore the possibility of pooling the funds and benchmarks together to derive a better portfolio strategy. Our target performance is to match the highest return among all the benchmarks, i.e.,

$$Z_B(\tilde{\boldsymbol{r}}) := \max_{j\in\{1,\ldots,m\}} \tilde{\boldsymbol{r}}^T\boldsymbol{p}^j.$$

Note that since the asset returns are random, the highest benchmark return is also random, and it depends on the realization of the asset returns. That is why we use the notation $Z_B(\tilde{\boldsymbol{r}})$ to represent the highest benchmark return. Our multiple-benchmark tracking error is defined as the difference between the portfolio return and the highest return induced from the benchmark portfolios, i.e., $\tilde{\boldsymbol{r}}^T\boldsymbol{x} - Z_B(\tilde{\boldsymbol{r}})$. We are interested in finding a portfolio $\boldsymbol{x}$ whose tracking-error volatility is minimized, i.e.,

$$\text{(T)} \qquad \min_{\boldsymbol{x}\in\mathcal{X}} \mathbf{E}\left[\left(Z_B(\tilde{\boldsymbol{r}}) - \tilde{\boldsymbol{r}}^T\boldsymbol{x}\right)^2\right],$$

where $\mathcal{X}$ is a set of feasible portfolios. The constraints in the set $\mathcal{X}$ includes the full investment constraint $\boldsymbol{e}^T\boldsymbol{x} = 1$. It is possible to capture additional constraints on the portfolio vector in the set $\mathcal{X}$, e.g., target expected return constraint, short-sale constraints, etc. Note that when there is only one benchmark portfolio, this reduces to the single-benchmark tracking-error model of Roll (1992).

Intuitively, as the financial asset returns are very volatile, it is almost impossible for a single benchmark to consistently perform well in every situation. Tracking the best return from a set of benchmarks appears to be a more attractive and practical objective, as it addresses partially the concern of a particularly bad benchmark dragging down the performance of the portfolio. The investor can also control the aggressiveness of the active investment by choosing an appropriate pool of benchmarks that suit the style and risk profile of the investor. Surprisingly, this approach can also be used to address the agency issue concerning the conflicting objectives between the investor and portfolio manager - the portfolio constructed using the multiple tracking error model may actually resulted in lower total returns variability, compared to the single benchmark approach.

Our main contributions in the paper are as follows:

1. Under the assumption of a normally distributed return vector, we derive the closed-form solution of our portfolio model without short-sale constraints, and show that the optimal multiple-benchmark tracking error portfolio relies on the probabilities that the benchmarks attain the highest return. This helps to stabilize the perofrmance of our portfolio strategy in numerical experiments, as those probabilities are generally less prone to estimation errors on means and covariances.

2. Using two suitably chosen benchmarks, we prove that one can generate the entire mean-variance efficient frontier using our model. This result is similar to the well known Two-Fund Theorem in classical portfolio theory.

3. We also compare the performance of our multiple benchmark tracking error model with the traditional single benchmark tracking error model, for fixed target return K. While the portfolio manager focuses on minimizing the variability of the excess return against the benchmark(s), we show that the total portfolio returns variability can be lower in the multiple benchmark environment. We identify the environments under which the multiple benchmarks portfolio strategy will dominate the performance of the single benchmark approach, i.e., lower total returns variability, at all target level of returns. This result exploits the fact that the variance of the returns of a linear combination of portfolio rules can be smaller than the variance of the returns of each indivudal portfolio rule.

4. We also show that the portfolio strategy constructed using our tracking error model will be preferred over using simple linear combination of the benchmark portfolios, in the environment when the portfolio managers have mean-variance utility functions with low risk aversion parameters.

5. More importantly, we show that our portfolio strategy performs well even with estimation errors and when transaction costs are properly accounted for[1]. We show that our model can be extended to penalize for transaction volumes. Alternately, we can also simply incorporate the buy-and-hold strategy into the set of benchmarks to reduce the transaction volumes. We show via extensive numerical experiments that this approach can significantly reduce transaction costs while not sacrificing the performance on returns. For instance, in the multi-period empirical tests, when we combine the partial minimum-variance (PARR) portfolio proposed by DeMiguel et al. (2009) with the buy-and-hold strategy as two benchmarks, our multiple-benchmark tracking-error portfolio incurs turnover rates that are less than half of those from the PARR portfolio. In terms of out-of-sample Sharpe ratio net of 50 basis point, our portfolio is significantly higher than the PARR portfolio. Our strategy also beats the equally weighted investment strategy (also known as the $1/N$ strategy) comprehensively when transaction costs are properly accounted for.

This paper is organized as follows. In the next section, we solve our multiple-benchmark tracking-error model and analyze the properties of its solution. We present and discuss the results of the numerical studies in Section 3 with a focus on including the buy-and-hold strategy as a benchmark to beat the other benchmark portfolios, especially the equally weighted portfolio. Finally, we provide some concluding remarks in Section 4.

2 Multiple-Benchmark Tracking-Error Portfolio

In this section, we will derive the solution to our model, i.e., Problem (T), and analytically investigate its features. Especially, we will compare our portfolio with the well-known Markowitz mean-variance efficient portfolio and the linear combination rule proposed by Tu & Zhou (2011). An extension of our model to penalize transaction costs is presented in the final part of this section.

To derive the closed-form solution of Problem (T), we first simplify the problem by linking it to the concept of persistency and Stein's identity.

2.1 Persistency and Stein's Identity

Bertsimas et al. (2006) define the persistency of a binary decision variable in a mixed zero-one linear optimization problem as the probability that the variable takes a value of one in an optimal solution. The persistency quantitatively captures the likelihood that a variable is a part of an optimal solution. It generalizes the definition of criticality index in project networks and choice probability in discrete choice models (c.f. Bertsimas et al. (2006), Natarajan et al. (2009), Mishra et al. (2012)). In the context of the benchmark tracking problem, we present the definition of persistency formally next.

[1]The most common approach in existing literature is to include either a penalty term in the objective function or a budget constraint in the portfolio models. For example, Brodie et al. (2009) proposed to add an additional penalty term to the classical Markowitz mean-variance framework, where the penalty is proportional to the sum of the absolute values of the portfolio weights.

Definition 1 *Define the* m *dimensional random vector*

$$\boldsymbol{p}(\tilde{\boldsymbol{r}}) = \left(\mathbb{I}_{Z_B(\tilde{\boldsymbol{r}})=\tilde{\boldsymbol{r}}^T\boldsymbol{p}^1}, \ldots, \mathbb{I}_{Z_B(\tilde{\boldsymbol{r}})=\tilde{\boldsymbol{r}}^T\boldsymbol{p}^m}\right)^T,$$

where the indicator function $\mathbb{I}_{Z_B(\tilde{\boldsymbol{r}})=\tilde{\boldsymbol{r}}^T\boldsymbol{p}^j}$ *takes a value of* 1 *if the* j^{th} *benchmark portfolio produces the maximum return in the set of benchmark portfolios and* 0 *otherwise. The persistency vector is an* m *dimensional vector whose* j^{th} *component is the probability that the* j^{th} *benchmark portfolio is the best performing portfolio in the set of benchmark portfolios, i.e.,*

$$\mathbf{E}[\boldsymbol{p}(\tilde{\boldsymbol{r}})] = \left(\mathbf{P}\left(Z_B(\tilde{\boldsymbol{r}}) = \tilde{\boldsymbol{r}}^T\boldsymbol{p}^1\right), \ldots, \mathbf{P}\left(Z_B(\tilde{\boldsymbol{r}}) = \tilde{\boldsymbol{r}}^T\boldsymbol{p}^m\right)\right)^T.$$

Definition 2 *Define the* $n \times m$ *benchmark portfolio matrix* $P = \left[\boldsymbol{p}^1, \ldots, \boldsymbol{p}^m\right]$. *The persistency weighted benchmark portfolio is defined as the* n *dimensional vector* $P\mathbf{E}[\boldsymbol{p}(\tilde{\boldsymbol{r}})]$.

Remark 1 *In this paper, we assume that* $\tilde{\boldsymbol{r}}$ *is a nondegenerate multivariate continuous random vector with a positive definite covariance matrix. The support of* $\tilde{\boldsymbol{r}}$ *over which more than one benchmark attains the maximum return has measure zero. Then* $\boldsymbol{p}(\boldsymbol{r})$ *is unique almost surely, and* $\mathbf{E}[\boldsymbol{p}(\tilde{\boldsymbol{r}})]$ *satisfies* $\sum_{j=1}^{m}\mathbf{E}[p_j(\tilde{\boldsymbol{r}})] = 1$.

As we will see later, the solution to the multiple-benchmark tracking-error minimization problem is related to the persistency when the return follows a multivariate normal distribution, i.e.,

(A) The random return vector, $\tilde{\boldsymbol{r}}$, follows a multivariate normal distribution with a finite mean, $\boldsymbol{\mu} \neq \mathbf{0}$[2], and a finite positive definite covariance matrix, Σ, denoted as $\tilde{\boldsymbol{r}} \sim N(\boldsymbol{\mu}, \Sigma)$.

This result is established by appealing to Stein's Identity in probability theory, which we will introduce next.

Lemma 1 *(Stein's Identity) Let the random vector* $\tilde{\boldsymbol{r}} = (\tilde{r}_1, ..., \tilde{r}_n)^T$ *be multivariate normally distributed with mean vector* $\boldsymbol{\mu}$, *and covariance matrix* Σ. *Consider a function* $h(r_1, ..., r_n) : \mathbb{R}^n \to \mathbb{R}$ *such that* $\partial h(r_1, ..., r_n)/\partial r_j$ *exists almost everywhere and* $\mathbf{E}[|\partial h(\tilde{\boldsymbol{r}})/\partial r_j|] < \infty$ *for all* $j = 1, \ldots, n$. *Denote*

$$\nabla h(\boldsymbol{r}) = (\partial h(\boldsymbol{r})/\partial r_1, ..., \partial h(\boldsymbol{r})/\partial r_n)^T.$$

Then $Cov(\tilde{\boldsymbol{r}}, h(\tilde{\boldsymbol{r}})) = \Sigma\mathbf{E}[\nabla h(\tilde{\boldsymbol{r}})]$ *or equivalently,*

$$Cov(\tilde{r}_l, h(\tilde{r}_1, ..., \tilde{r}_n)) = \sum_{j=1}^{n} Cov(\tilde{r}_l, \tilde{r}_j)\mathbf{E}\left[\frac{\partial}{\partial r_j}h(\tilde{r}_1, ..., \tilde{r}_n)\right] \quad \forall l = 1, ..., n.$$

For completeness, the proof of Lemma 1 is provided in Appendix A. Intuitively, if we treat $Z_B(\tilde{\boldsymbol{r}})$ as a function on $\tilde{\boldsymbol{r}}$, we can apply Stein's Identity to derive the covariance between the individual asset return and the highest benchmark return.

[2]Note that the assumption of $\boldsymbol{\mu} \neq \mathbf{0}$ is required only for the model analysis, especially on efficient frontiers. For our basic model, we can still obtain the solution when $\boldsymbol{\mu} = \mathbf{0}$.

2.2 Tracking Error Minimization

Applying Stein's Identity to $Cov(Z_B(\tilde{\boldsymbol{r}}), \tilde{\boldsymbol{r}}^T)$, we get

$$Cov(Z_B(\tilde{\boldsymbol{r}}), \tilde{\boldsymbol{r}}^T) = (\Sigma P\mathbf{E}[\boldsymbol{p}(\tilde{\boldsymbol{r}})])^T.$$

Our problem can be simplified as shown in the next proposition.

Proposition 1 *Under Assumption (A), the multiple benchmark tracking error portfolio in Model (T) can be found by solving the following convex quadratic minimization problem:*

$$(T') \qquad \min_{\boldsymbol{x}\in\mathcal{X}} \; \boldsymbol{x}^T(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T)\boldsymbol{x} - 2\Big(\Sigma P\mathbf{E}[\boldsymbol{p}(\tilde{\boldsymbol{r}})] + \mathbf{E}[Z_B(\tilde{\boldsymbol{r}})]\boldsymbol{\mu}\Big)^T\boldsymbol{x},$$

where $P\mathbf{E}[\boldsymbol{p}(\tilde{\boldsymbol{r}})]$ is the persistency weighted benchmark portfolio, and $\mathbf{E}[Z_B(\tilde{\boldsymbol{r}})]$ is the expected highest benchmark return.

To maintain the flow of the paper, all the proofs of our results in this section are relegated to Appendix B.

Remark 2 *Suppose that the random vector $\tilde{\boldsymbol{r}}$ is not normally distributed. It is still possible to find the multiple-benchmark tracking-error portfolio by solving the following convex quadratic programming problem:*

$$\min_{\boldsymbol{x}\in\mathcal{X}} \; \boldsymbol{x}^T(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T)\boldsymbol{x} - 2\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\tilde{\boldsymbol{r}}^T\right]\boldsymbol{x}.$$

This requires the estimation of $\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\tilde{\boldsymbol{r}}^T\right]$. The advantage of resorting to Stein's Identity for the multivariate normal distribution is twofold. First, by using Stein's Identity, we need to estimate (a) The persistency vector $\mathbf{E}[\boldsymbol{p}(\tilde{\boldsymbol{r}})]$, and (b) The expectation of the best benchmark return $\mathbf{E}[Z_B(\tilde{\boldsymbol{r}})]$. Estimation of the benchmark portfolio that different experts believe will outperform the rest is inherently easier to elicit from managers. Second, the transformed problem provides a simple characterization with a closed-form solution that allows for more in-depth analysis of the model. We elaborate on this issue in the next several subsections.

By re-writing the expression in Proposition 1, we can reinterpret our model as a variant of a single-benchmark tracking-error model: Problem (T′) is equivalent to

$$\min_{\boldsymbol{x}\in\mathcal{X}} \; (\boldsymbol{x} - P\mathbf{E}[\boldsymbol{p}(\tilde{\boldsymbol{r}})])^T \boldsymbol{\Sigma}(\boldsymbol{x} - P\mathbf{E}[\boldsymbol{p}(\tilde{\boldsymbol{r}})]) + \left(\mathbf{E}[Z_B(\tilde{\boldsymbol{r}})] - \boldsymbol{\mu}^T\boldsymbol{x}\right)^2.$$

The first term is essentially the variance of the tracking-error, measured against the persistency weighted portfolio. The second term penalizes the shortfall of the portfolio return from the highest benchmark return. When there is a constraint that fixes the expected portfolio return, the second term will vanish in the minimization problem, and our model reduces to a single-benchmark tracking-error model with the persistency weighted benchmark portfolio as the only benchmark. In general, our model anchors in the persistency weighted benchmark portfolio, and it is adjusted to recover the loss in the expected portfolio return from the highest benchmark return. This result shows that model (T) is related to the single-benchmark tracking-error literature in the following ways:

- If we fixed a target expected return, Problem (T) reduces to a tracking error minimization problem, where the benchmark tracked is $P\mathbf{E}\left[\boldsymbol{p}\left(\tilde{\boldsymbol{r}}\right)\right]$, the persistency strategy formed by the set of portfolio used as benchmarks.

- If we fixed a budget for the variance of the tracking-error, $\left(\boldsymbol{x} - P\mathbf{E}\left[\boldsymbol{p}\left(\tilde{\boldsymbol{r}}\right)\right]\right)^T \boldsymbol{\Sigma} \left(\boldsymbol{x} - P\mathbf{E}\left[\boldsymbol{p}\left(\tilde{\boldsymbol{r}}\right)\right]\right)$, then Problem (T) will find a strategy that has expected return as close as possible to the expected returns of the best strategy in the portfolio. Our model therefore uses $\mathbf{E}\left[Z_B\left(\tilde{\boldsymbol{r}}\right)\right]$ to anchor the selection of the portfolio strategy in the tracking-error model, to avoid excessive risk, instead of limiting the total risk (variance of the returns), as commonly used. The selection of the benchmarks used in our model is thus crucial to the performance of the portfolio strategy.

2.3 Closed-Form Solution

In this subsection, we present the closed-form expression of the multiple-benchmark tracking-error portfolio when the return vector $\tilde{\boldsymbol{r}}$ satisfies the multivariate normality assumption. To simplify the expression, we introduce three constants, $A = \boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu}$, $B = \boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{e}$, and $C = \boldsymbol{e}^T\Sigma^{-1}\boldsymbol{e}$. These constants are also used to describe the closed-form expression of the Markowitz mean-variance portfolio (cf. Steinbach (2001)). Note that by Assumption (A), $A > 0$ and $C > 0$.

Theorem 1 *Define the set of feasible portfolios as $\mathcal{X} = \left\{\boldsymbol{x} \in \mathbb{R}^n \,:\, \boldsymbol{e}^T\boldsymbol{x} = 1\right\}$. Under Assumption (A), the optimal multiple-benchmark tracking-error portfolio in (T) is given by*

$$P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] + \Sigma^{-1}\left(\frac{C\boldsymbol{\mu} - B\boldsymbol{e}}{(A+1)C - B^2}\right)\left(\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right] - \boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]\right). \tag{1}$$

Define $\mu_p := \boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]$, the mean return of the persistency weighted portfolio. It is well-known that if the returns of the portfolios in the set of benchmarks are negatively correlated, then it is possible for the variance of the persistency portfolio to be smaller than the variance of the individual portfolio. Our strategy builds on the persistency portfolio, and adjusts for higher mean returns through the term,

$$\Sigma^{-1}\left(\frac{C\boldsymbol{\mu} - B\boldsymbol{e}}{(A+1)C - B^2}\right)\left(\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right] - \boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]\right).$$

In this way, we can ensure that the mean returns of our strategy is at least as good as the persistency portfolio.

Proposition 2 *Under Assumption (A), the expected return of our multiple-benchmark tracking-error portfolio is not less than μ_p. In particular, when $\boldsymbol{\mu} \neq \boldsymbol{e}$, the portfolio has a strictly higher expected return than μ_p.*

2.4 Comparison with the Markowitz Mean-Variance Portfolio

The pioneering work of modern portfolio theory by Markowitz (1952) quantified the relationship between the expected return and risk of portfolios, which is measured by the variance in portfolio returns. Markowitz introduced the notion of an efficient portfolio as the portfolio with minimal variance at a given level of expected return. The continuum of such portfolios forms an efficient frontier in the mean-variance space of the portfolios.

In this subsection, we exploit the advantage of the closed-form solution and compare our portfolio with the Markowitz mean-variance portfolio.

2.4.1 Optimal Portfolio Weights

Consider the Markowitz portfolio optimization model of the following form:

$$\min_{\boldsymbol{e}^T\boldsymbol{x}=1} \frac{1}{2}\boldsymbol{x}^T\Sigma\boldsymbol{x} - \gamma\boldsymbol{\mu}^T\boldsymbol{x},$$

where γ is the risk aversion parameter. The closed-form solution (cf. Steinbach (2001)) is

$$\Sigma^{-1}\left(\frac{(1-\gamma B)\boldsymbol{e}+\gamma C\boldsymbol{\mu}}{C}\right) = \frac{1}{C}\Sigma^{-1}\boldsymbol{e} - \frac{\gamma}{C}\Sigma^{-1}\left(B\boldsymbol{e} - C\boldsymbol{\mu}\right).$$

Rearranging the closed-form expression of the multiple-benchmark tracking-error portfolio in Theorem 1 helps to make the comparison more explicit as follows:

$$P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] + \frac{\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right] - \boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]}{(A+1)C - B^2}\Sigma^{-1}\left(C\boldsymbol{\mu} - B\boldsymbol{e}\right).$$

From these formulas, it is clear that both portfolios consist of two components: (a) a baseline portfolio, and (b) an adjustment term with a common factor, $\Sigma^{-1}\left(B\boldsymbol{e} - C\boldsymbol{\mu}\right)$. The baseline portfolio of the Markowitz mean-variance portfolio is the minimum-variance portfolio, $\Sigma^{-1}\boldsymbol{e}/C$, and the adjustment is related to the risk aversion parameter, γ. For the multiple-benchmark tracking-error portfolio, the baseline portfolio is the persistency weighted benchmark portfolio. The adjustment term accounts for the impact of the random return on the performance of the benchmark portfolios, in other words, the selection of the best performer as the target expected return.

To better under the adjustment terms, consider the following optimization problem:

$$(\mathrm{M}_0) \qquad \min_{\boldsymbol{e}^T\boldsymbol{x}=0,\, \boldsymbol{\mu}^T\boldsymbol{x}=K'} \frac{1}{2}\boldsymbol{x}^T\Sigma\boldsymbol{x},$$

which attempts to find the minimum variance adjustment that gives the target expected return K'. The closed-form solution to this problem reads

$$K'\frac{\Sigma^{-1}\left(C\boldsymbol{\mu} - B\boldsymbol{e}\right)}{AC - B^2}.$$

The detailed derivation can be found in the proof of Proposition 5. Knowing this, the adjustment term of our multiple-benchmark tracking-error portfolio can be interpreted as the minimum variance adjustment that tries to bring the expected return of the portfolio close to $\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right]$. The resulting expected return of our portfolio is

$$\frac{C}{AC - B^2 + C}\boldsymbol{\mu}^T P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] + \frac{AC - B^2}{AC - B^2 + C}\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right].$$

which is the convex combination of the expected return of the persistency weighted benchmark portfolio and the expected highest benchmark return.

To provide some intuition on the difference between these two portfolios, consider a simple example of investment between a risk-free asset and a risky asset.

Example 1 *Suppose an investor has to decide a portfolio among two uncorrelated assets, one of which is risk-free with zero variance, and the other is risky with a variance of 1. Both assets have zero-mean returns. In this case, the investor who follows the Markowitz strategy will always choose the risk-free asset for any nonnegative risk aversion parameter. On the other hand, suppose we choose two extreme strategies as the benchmark portfolios – each strategy investing solely in one of the two assets. The multiple-benchmark tracking-error portfolio we obtained is one that divides the capital among the two assets with equal weights (under the normality assumption). This is simply the equally weighted investment strategy often used by practitioners.*

2.4.2 Volatility

It has been observed that the Markowitz mean-variance portfolio suffers from severe volatility in portfolio returns due to estimation errors in mean and covariance (cf. Michaud (1989) and Best & Grauer (1991)). Our multiple-benchmark tracking-error portfolio tends to exhibit less volatility since the risk of estimation errors is mitigated by the persistency values, which are more robust to estimate. This difference is indeed observed in the numerical studies we did using the real data in Section 3. In what follows, we use an experiment to illustrate such difference. We include the equally weighted portfolio (a.k.a. the $1/n$ portfolio) as a reference portfolio, since it is known to be effective in minimizing volatility, in particular, for a large pool of assets.

In the experiment, we simulate the monthly returns of 48 risky assets under multivariate normality assumption. We use the 48 Industry Portfolios from the Fama French online data library (in the period from 1981 to 2010) as the setup for this experiment. We adopt a rolling horizon method with an estimation window of 80 periods and investment horizon of 400 periods. We use the first 80 sample points to obtain sample mean, variance and covariance, of the returns parameters, based on which the portfolios are determined. The performance of the portfolios are evaluated using the 81^{st} sample point, which is an out-of-sample return. Next, the whole process moves one period forward, and we update the estimation of the sample mean, variance and covariance, using the last 80 sample points. We continue this experiment for 400 periods. The out-of-sample returns of the three portfolios over the whole investment horizon are plotted in Figure 1. In this experiment, there were 48 benchmark portfolios, each corresponding to an individual industry.

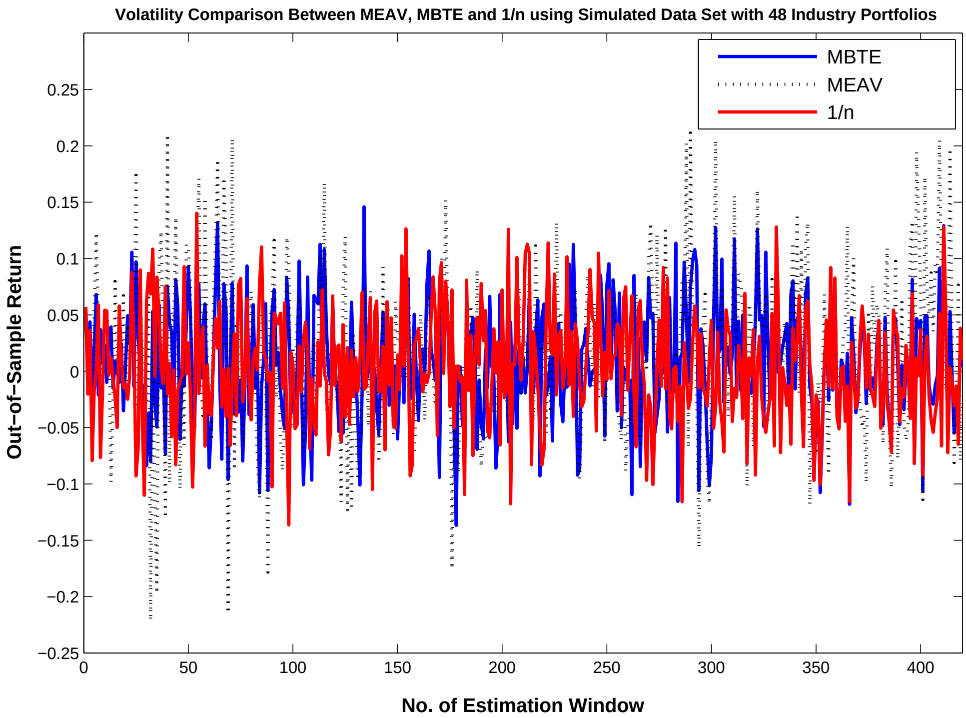


Figure 1: Out-of-sample returns of the $1/n$, Markowitz mean-variance (MEAV), and multiple-benchmark tracking-error (MBTE) portfolios over an investment horizon of 400 periods

The out-of-sample return of the $1/n$ portfolio is stable and shows only slight fluctuation over the course of the experiment. The Markowitz mean-variance portfolio however exhibits much larger volatility. In comparison, our multiple-benchmark tracking-error portfolio's performance is close to that of the $1/n$ portfolio, with significantly smaller fluctuations than the Markowitz mean-variance portfolio.

2.4.3 Efficient Frontier (In-Sample)

Next, we compare our multiple-benchmark tracking-error frontier with the mean-variance efficient frontier. To give an immediate idea, we first plot the multiple-benchmark tracking-error frontier and the efficient frontier assuming the full knowledge on the distributional parameters of the random returns. Here, we consider the following multiple-benchmark tracking-error model:

$$(\mathrm{T}'') \qquad \min_{\boldsymbol{e}^T\boldsymbol{x}=1,\, \boldsymbol{\mu}^T\boldsymbol{x}=K} \mathbf{E}\left[\left(Z_B(\tilde{\boldsymbol{r}}) - \tilde{\boldsymbol{r}}^T\boldsymbol{x}\right)^2\right],$$

which is a variant of Problem (T) with additional target expected return constraint. Similar to the case without target expected return constraint, we can derive the closed-form solution to Problem (T$''$) as follows:

$$P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] + \left(\boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - K\right)\Sigma^{-1}\left(\frac{B\boldsymbol{e} - C\boldsymbol{\mu}}{AC - B^2}\right). \tag{2}$$

To obtain the mean-variance efficient frontier, we consider the Markowitz model with a target expected return constraint as follows:

$$(\mathrm{M}) \qquad \min_{\boldsymbol{e}^T\boldsymbol{x}=1,\, \boldsymbol{\mu}^T\boldsymbol{x}=K} \frac{1}{2}\boldsymbol{x}^T\Sigma\boldsymbol{x}$$

We use an experiment to illustrate the different in the two frontiers. We simulate the monthly returns of 10 risky assets under multivariate normality assumption. We use the estimated mean, variance and covariance of the monthly returns of the 10 Industry Portfolios from the Fama French online data library (in the period from 1981 to 2010) as the underlying distributional parameters. We assume the complete knowledge of means and variances of returns when solving (T$''$) and (M). We consider a sequence of target expected returns, K, from 0 to 0.2 with a step size of 0.0001. For each K, we solve (T$''$) and (M) to obtain our multiple-benchmark tracking-error portfolio and the Markowitz mean-variance portfolio, respectively. Then we compute the variance of the two portfolios. The continuum of such K-variance pairs constitutes the frontier for each portfolio selection model, as plotted in Figure 2. Similarly, in this experiment, the benchmark portfolios are chosen to be all the extreme portfolios that invest solely in individual assets.

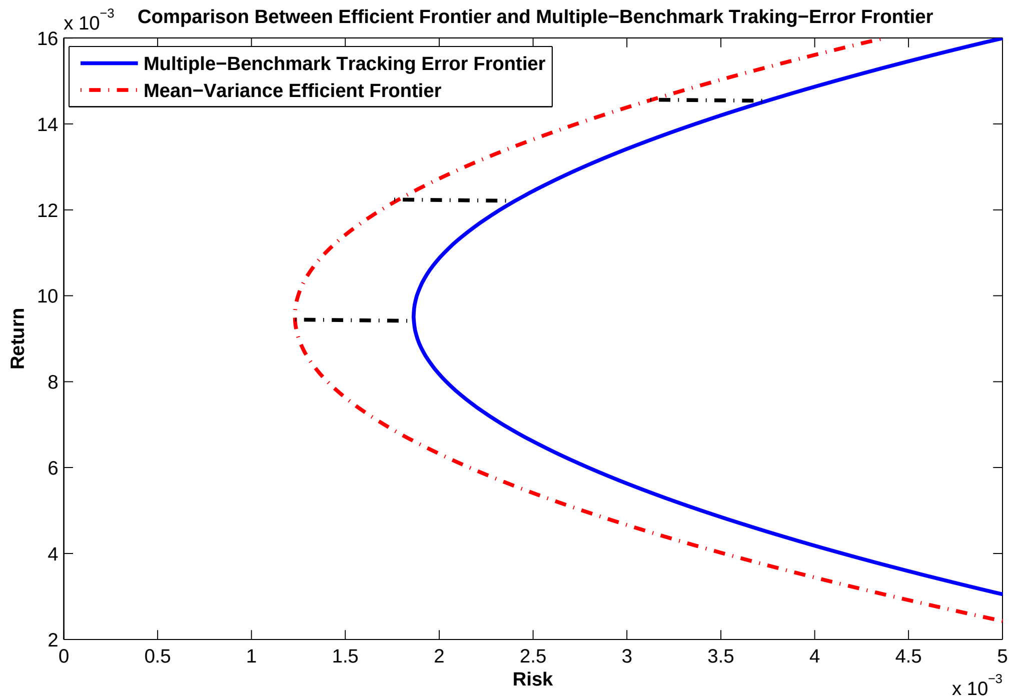


Figure 2: Risk and return with known distributional parameters and simulated data

Comparing the two frontiers in the risk-return plot, we observe a constant shift of the multiple-benchmark tracking-error frontier from the mean-variance efficient frontier. Note that a similar feature was observed for single-benchmark tracking-error portfolios by Roll (1992). The magnitude of the shift corresponds to the magnitude of the agency problem in this environment - when the portfolio manager focuses on minimizing the variability of the tracking error, the resulting portfolio is often inefficient and lies in the interior of the efficient frontier. We give an analytical expression for the constant shift next.

Proposition 3 *Under Assumption (A), the multiple-benchmark tracking-error frontier obtained from solving Problem (T''), is a constant shift from the mean-variance efficient frontier. More specifically, the distance between the two frontiers at each level of expected return is*

$$\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]^T P^T \Sigma P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \frac{C}{AC - B^2}\left(\boldsymbol{\mu}^T P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \frac{B}{C}\right)^2 - \frac{1}{C}.$$

Note that the first term in the summand corresponds to the variance of the returns of the persistency weighted portfolio strategy, and the term B/C corresponds to the mean returns of the minimum variance strategy. We can assume that the mean returns of each portfolio used in the set of benchmarks generate higher mean returns than the minimum variance strategy. Hence

$$\left(\boldsymbol{\mu}^T P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \frac{B}{C}\right)^2$$

corresponds to the square of the excess returns of the persistency weighted portfolio strategy above the minimum variance strategy. We can now use this result to rank the performance of the tracking error models using different benchmarks:

Theorem 2 *If*

$$Var(\tilde{\boldsymbol{r}}^T P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]) - \frac{C}{AC - B^2}\left(\boldsymbol{\mu}^T P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \frac{B}{C}\right)^2 \leq Var(\tilde{\boldsymbol{r}}^T P_j) - \frac{C}{AC - B^2}\left(\boldsymbol{\mu}^T P_j - \frac{B}{C}\right)^2,$$

then the frontier constructed from the multiple benchmark tracking error model (T'') dominates the frontier for the single benchmark tracking error model constructed using benchmark P_j.

This result can be used to identify complementary benchmark portfolio that can help improve the performance of the single benchmark tracking error model using P_j. For instance, if there exists portfolio Q_j such that

- $\boldsymbol{\mu}^T P_j = \boldsymbol{\mu}^T Q_j$,
- $\text{Var}(\tilde{\boldsymbol{r}}^T P_j) = \text{Var}(\tilde{\boldsymbol{r}}^T Q_j)$, and
- $\tilde{\boldsymbol{r}}^T P_j$ and $\tilde{\boldsymbol{r}}^T Q_j$ are independent or negatively correlated,

then

$$\text{Var}\left(\tilde{\boldsymbol{r}}^T(\kappa P_j + (1-\kappa)Q_j)\right) \leq \kappa^2 \text{Var}(\tilde{\boldsymbol{r}}^T P_j) + (1-\kappa)^2 \text{Var}(\tilde{\boldsymbol{r}}^T Q_j) \leq \text{Var}(\tilde{\boldsymbol{r}}^T P_j),$$

for any κ in $[0, 1]$. Thus Q_j can be used in our multiple benchmark model to improve the performance of the single benchmark tracking error model. This result shows the potential of the multiple benchmark tracking error model in reducing the impact of the agency problem for the investor, as it can bring the frontier of the tracking error model closer to the efficient frontier.

We can also show an interesting result similar to the famous Two-Fund Theorem, which says that any affine combination of two distinct mean-variance efficient portfolios is itself a mean-variance efficient portfolio.

Proposition 4 *Under Assumption (A), when the set of benchmark portfolios contains exactly two distinct efficient mean-variance portfolios, our multiple-benchmark tracking-error frontier coincides with the mean-variance efficient frontier. Consequently, the multiple-benchmark tracking-error portfolio obtained from solving Problem (T) falls on the mean-variance efficient frontier.*

This proposition shows that our multiple-benchmark tracking-error model has the flexibility to generate the entire mean-variance efficient frontier if the benchmark portfolios are chosen properly. Indeed, this result can be extended to the case with more than two mean-variance efficient benchmark portfolios.

2.4.4 Efficient Frontier (Out-of-Sample)

For the purpose of completeness, we conduct further numerical analysis by drawing the frontiers for both portfolios based on out-of-sample estimation in Figure 3. We simulate 130 samples from the underlying distributional parameters same as the previous experiment, and use the first 120 sample points to obtain sample mean, variance and covariance. We consider the sequence of values for K as before. For each K, we determine our multiple-benchmark tracking-error portfolio with the sample mean and covariance, and calculate the out-of-sample mean return and variance using the last 10 periods of data. By drawing such return-variance pairs for all K's, we get an out-of-sample multiple-benchmark tracking-error frontier. The out-of-sample mean-variance efficient frontier is obtained in a similar way.

Although the theoretical frontier of the Markowitz mean-variance portfolio could be more efficient in-sample, in the out-of-sample experiment, the estimation errors in mean and variance leads to much less efficient Markowitz portfolios.

2.5 Comparison with the Linear Combination Rule

To improve the performance of the Markowitz mean-variance portfolio under estimation errors, Tu & Zhou (2011) proposed to combine more sophisticated strategies with the naive $1/n$ rule. They found that the optimal affine combination of the estimated Markowitz portfolio and the $1/n$ portfolio often outperforms both portfolios in terms of expected mean-variance utility. In order to derive the desired result, the authors focused on the unconstrained version of the Markowitz model, i.e., without the requirement that the sum of portfolio weights equals to one. In a later study, Kirby & Ostdiek (2012) pointed out the importance of research design in driving the performance of the Markowitz portfolio. In particular, high target expected return will significantly inflate the estimation errors and result in extremely risky position for the Markowitz portfolio.

We investigate next the relationship between our multiple-benchmark tracking-error model and the linear combination rule, for different ranges of the target expected return. Let $\boldsymbol{p}^1$ and $\boldsymbol{p}^2$ be

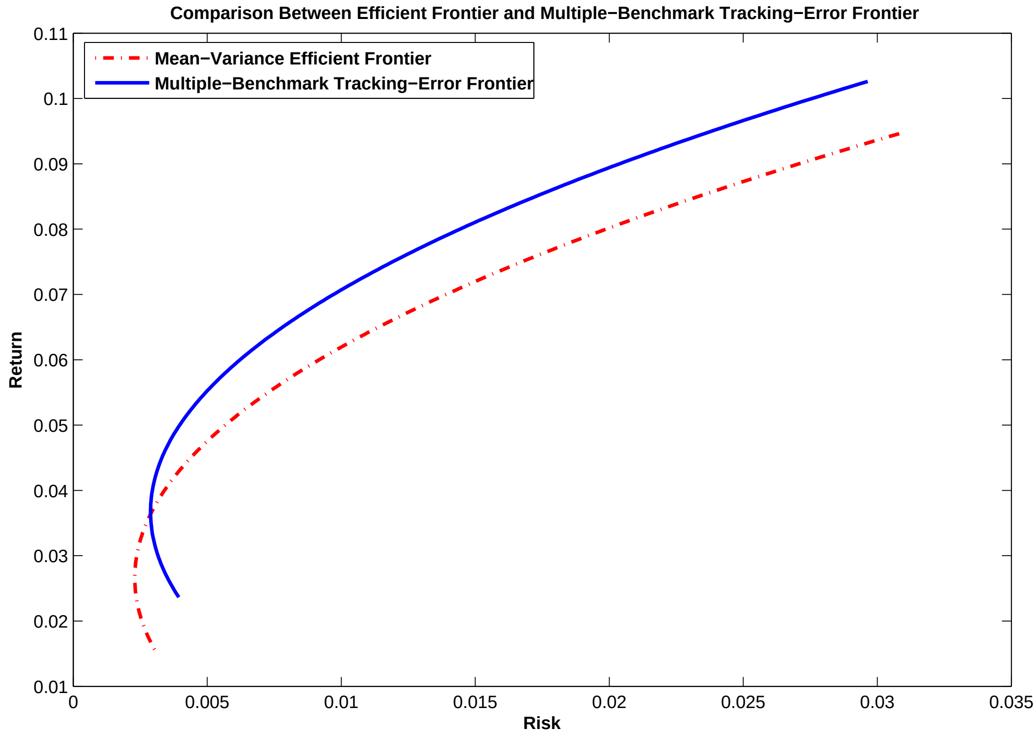


Figure 3: Risk and return of different portfolios using out-of-sample estimations of mean and variance

two distinct portfolios with different expected return, i.e., $\boldsymbol{\mu}^T \boldsymbol{p}^1 \neq \boldsymbol{\mu}^T \boldsymbol{p}^2$. The linear combination rule generates a series of portfolios of the form,

$$\delta \boldsymbol{p}^1 + (1 - \delta)\, \boldsymbol{p}^2,$$

where δ is the linear combination coefficient. To facilitate the comparison, for our model, we use the same two portfolios to construct our benchmark. Note that given a target expected return, the solution for the linear combination rule can be uniquely determined. Similarly, Equation (2) gives the closed-form solution of our multiple-benchmark tracking-error portfolio at the target expected return K.

In general, the variance of the linear combination portfolio will usually increase at a faster rate as the target expected return increases. On the other hand, as demonstrated earlier, the frontier of our multiple-benchmark tracking-error portfolio is a constant shift to the right from the mean-variance efficient frontier. As the target expected return increases, we expect our portfolio to be more efficient, i.e., having smaller variance than the solution produced by linear combination rule. We have the following result.

Proposition 5 *Under Assumption (A), the multiple-benchmark tracking-error frontier will dominate the linear combination rule frontier when the target expected return is high enough.*

The above result shows that there exists a threshold such that once the target expected return exceeds this threshold, the linear combination rule would be less efficient than the multiple-benchmark tracking-error portfolio. In fact, from our numerical tests, such threshold value is usually very small, and the performance of the linear combination rule deteriorates significantly when the target expected return increases.

Remark 3 *We can also interpret the above result from the perspective of utility theory. Suppose that the portfolio manager has the mean-variance utility of the form,*

$$\boldsymbol{\mu}^T \boldsymbol{x} - \frac{\gamma}{2} \boldsymbol{x}^T \Sigma \boldsymbol{x},$$

where γ is the risk aversion parameter. Then our multiple-benchmark tracking-error portfolio will be preferred over the linear combination rule for relatively small γ, i.e., less risk aversion. From the figures, the portfolio manager has to be extremely risk averse for the linear combination rule to be preferred to the multiple-benchmark tracking-error portfolio.

2.6 Transaction Cost

Transaction costs are often inevitable in real investment activities. As discussed before, we can explicitly include the buy-and-hold strategy as a benchmark portfolio to control the transaction volume. In this subsection, we show that our model is also capable of handling transaction cost in the conventional way by adding a penalty term into the objective function.

To facilitate the following exposition, we introduce some additional notation. Let W denote the wealth at the beginning of the current investment period, and define $\boldsymbol{x}^0$ as the starting portfolio, i.e., the initial weights of capital on each asset. As before, $\boldsymbol{x}$ represents the current investment decision. In this case, it can also be referred to as portfolio repositioning decision. The transaction volume is measured by $\sum_{i=1}^n W \left| x_i - x_i^0 \right| = W \boldsymbol{e}^T \left| \boldsymbol{x} - \boldsymbol{x}^0 \right|$. However, the problem becomes non-smooth if we directly work with the transaction cost that is linear in the transaction volume. To make the problem more tractable and emphasize on avoiding high transaction volume, we extend the basic model in Problem (T) by adding a penalty term on the quadratic transaction volume. The problem is formulated as follows:

$$\text{(TC)} \qquad \min_{\boldsymbol{e}^T \boldsymbol{x}=1} \mathbf{E} \left[\left(Z_B(\tilde{\boldsymbol{r}}) - \sum_{i=1}^n \tilde{r}_i x_i \right)^2 \right] + \nu (\boldsymbol{x} - \boldsymbol{x}^0)^T (\boldsymbol{x} - \boldsymbol{x}^0),$$

where $\nu \geq 0$ is a penalty parameter that captures the effect of the quadratic transaction volume, $W^2 (\boldsymbol{x} - \boldsymbol{x}^0)^T (\boldsymbol{x} - \boldsymbol{x}^0)$. Since ν is a constant, we can absorb W^2 into ν. The new objective can be interpreted as an adjusted disutility function of the investor with a penalty on the transaction volume, where ν characterizes the investor's aversion to high transaction volume. With such change, Problem (TC) remains a convex quadratic programming problem, and we are able to establish its closed-form solution as shown in the following proposition.

Proposition 6 *Under Assumption (A), the closed-form solution to Problem (TC) is given by*

$$\boldsymbol{x} = \frac{D\boldsymbol{e}}{\boldsymbol{e}^T D\boldsymbol{e}} + \left(I_n - \frac{DJ_n}{\boldsymbol{e}^T D\boldsymbol{e}}\right) D\left(\boldsymbol{\mu}\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right] + \Sigma P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] + \nu\boldsymbol{x}^0\right), \tag{3}$$

where $D = \left(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T + \nu I_n\right)^{-1}$*;* I_n *is the identity matrix; and* J_n *is the matrix in* $\mathbb{R}^{n\times n}$ *with all entries being 1.*

3 Numerical Studies

In this section, we will present some numerical studies based on real data to test the performance of the models proposed in this paper. We start by describing the setup of these studies, including data sets, comparison portfolio strategies, methodology and performance measures.

3.1 Data Sets

We evaluate the performance of our model in four empirical data sets listed in Table 1. The data sets we choose fall into two categories. The first three data sets are portfolios representing the U.S. stock market, and the last one is comprised of individual U.S. stocks. We present the analysis on the risk and return of these data sets in Figure 4. The first graph in each panel shows the annualized mean return and the second graph shows the annualized standard deviation of the returns. All the sample points are used in the calculation. From the figure, we observe that the four data sets demonstrate distinct risk and return characteristics, such as different spreads of mean returns within the same data set, and different risk levels, etc.

Abbreviation	Data Set and Description	n	Time Period	Source
10Ind	Ten industry portfolios representing the U.S. stock market	10	07/1963 –12/2011	K. French
48Ind	Forty-eight industry portfolios representing the U.S. stock market	48	07/1963 –12/2011	K. French
25FF	Twenty-five Fama and French portfolios of firms sorted by size and book-to-market	25	07/1963 –12/2011	K. French
8Stock	Eight U.S. stocks (Crude Oil, J.P. Morgan Funds, General Electric Company, The Coca-Cola Company, Johnson & Johnson, International Business Machines Corp., Gold Ounce, AT & T Inc.)	8	08/1980 –01/2013	Bloomberg

Notes: "K. French" refers to the Kenneth R. French data library available online at *http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html*

Table 1: Data sets used in empirical experiments

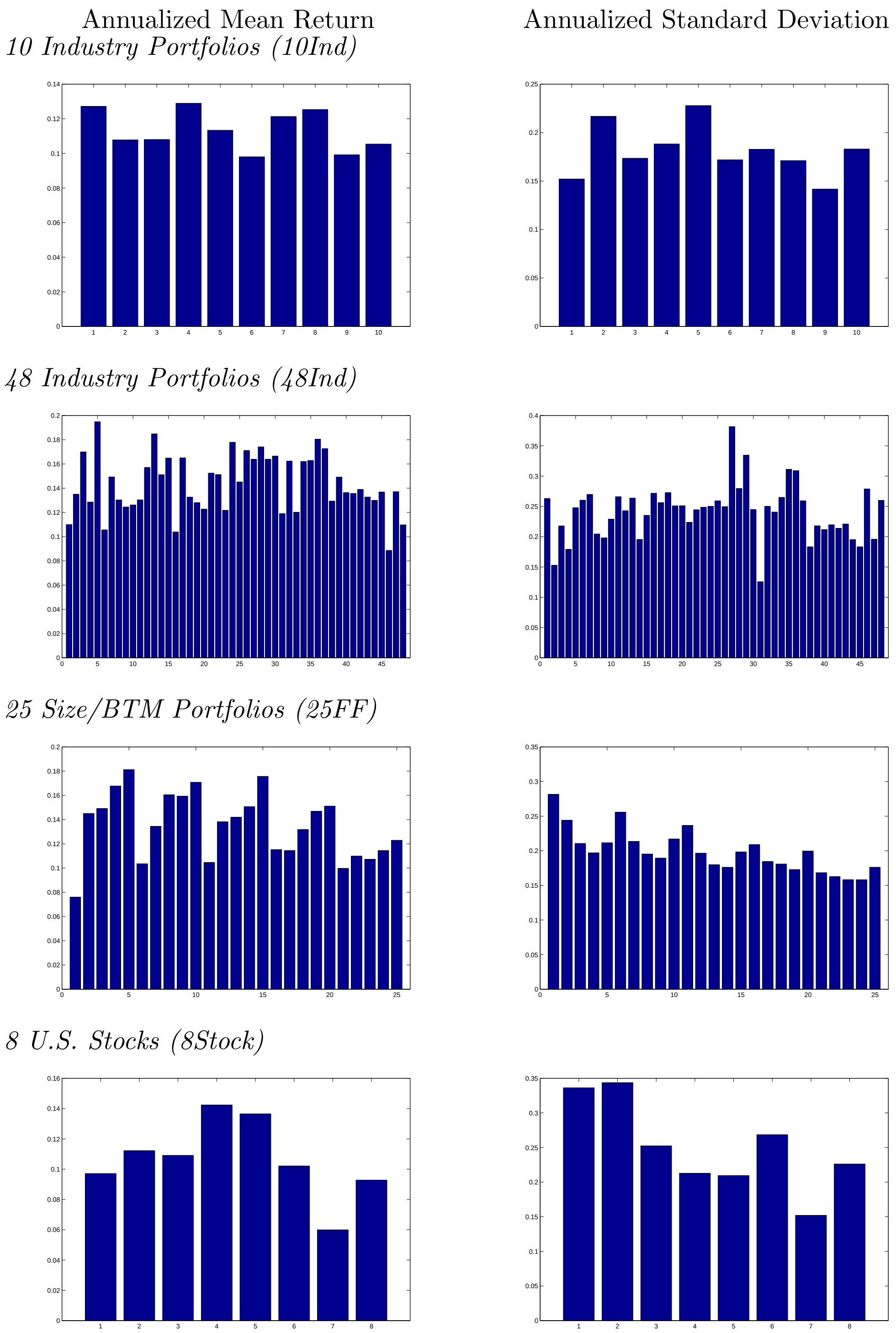


Figure 4: Risk and return characteristics of the data sets

3.2 Portfolio Models

The portfolio strategy developed in this chapter depends on the group of benchmark portfolios that are being tracked. To evaluate the performance of the multiple-benchmark tracking-error portfolio, we compare the performance of our portfolio against that of its benchmarks as well as that of the $1/n$ portfolio. We choose the buy-and-hold strategy in addition to another competitive portfolio as benchmarks. All the benchmark portfolios are listed in Table 2.

No.	Portfolio Selection Model	Abbreviation
0a	Multiple-benchmark tracking-error portfolio	MBTE
0b	Multiple-benchmark tracking-error portfolio with penalty on transaction volume	MBTEP
1	Equally-weighted ($1/n$) portfolio (DeMiguel et al. (2007))	1/n
2	Markowitz mean-variance portfolio with target expected return (Markowitz (1952))	MEAV
3	Minimum-variance portfolio without short-sale constraints	MINU
4	Minimum-variance portfolio with covariance matrix being a weighted average of sample covariance matrix and the single-index covariance matrix (Ledoit & Wolf (2003))	M1FAC
5	Minimum-variance portfolio with covariance matrix being a weighted average of sample covariance matrix and the diagonal covariance matrix (Ledoit & Wolf (2003))	MIND
6	Minimum-variance portfolio with covariance matrix being a weighted average of sample covariance matrix and the identity matrix (Ledoit & Wolf (2003))	M1PAR
7	Minimum-variance portfolio with generalized constraints (DeMiguel et al. (2007))	GMC
8	On-line portfolio using multiplicative updates (Helmbold et al. (1998))	MUL
9	Minimum CVaR (Conditional Value-at-Risk) portfolio (Rockafellar & Uryasev (2000), Rockafellar & Uryasev (2004))	CVAR
10	Partial minimum-variance portfolio with k calibrated by maximizing the portfolio return in previous period (DeMiguel et al. (2009))	PARR
11	Buy-and-hold strategy	B-N-H

Table 2: List of portfolio strategies considered

The multiple-benchmark tracking-error portfolio is listed as Portfolio 0a and the multiple-benchmark tracking-error portfolio with penalty on transaction volume is listed as Portfolio 0b in Table 2. We use the closed-from solutions in Equation (1) and Equation (3) to compute our multiple-benchmark tracking-error portfolios.

Portfolios 1–11 listed in Table 2 serve two purposes. First, a subset of these portfolios are used as benchmark portfolios to compute the multiple-benchmark tracking-error portfolios. Second, all of these portfolios serve as comparison portfolios to gauge the out-of-sample performance of our multiple-benchmark tracking-error portfolios.

The first comparison portfolio is the $1/n$ strategy, which simply assigns equal weights to all the assets in the data set. The Markowitz mean-variance portfolio (MEAV) relies on estimates of mean, variance and covariance of the returns, and is computed from Problem (M). The target expected return is set to be the expected return of the $1/n$ portfolio. Such a target has been observed to be more appropriate for the Markowitz model (cf. Kirby & Ostdiek (2012)). Then we consider the class of minimum-variance portfolios, the first of which is the minimum-variance portfolio without short-sale constraints (MINU). The next three minimum-variance portfolios (M1FAC, MIND, and M1PAR) are formed using different covariance estimation techniques as described in Ledoit & Wolf (2003) and Ledoit & Wolf (2004). The last portfolio in this set (GMC) is adopted from DeMiguel et al. (2007). It is a combination of the $1/n$ policy and the constrained minimum-variance portfolio. The additional constraint is $\boldsymbol{x} \geq a\boldsymbol{e}$ with $a = 1/(2n)$. We also consider the on-line portfolio (MUL) using multiplicative updates as studied in Helmbold et al. (1998). The portfolio that minimizes the Conditional Value-at-Risk (CVAR), a coherent risk measure, is also included in our study. This portfolio is supposed to be very conservative and would refrain from taking highly risky positions. For detailed discussion on CVaR, the reader is referred to Rockafellar & Uryasev (2000) and Rockafellar & Uryasev (2004). In addition, we consider the minimum-variance portfolio with a constraint on the portfolio norm developed in DeMiguel et al. (2009). In particular, PARR is the partial minimum-variance portfolio with k calibrated using cross-validation over portfolio variance, where k indicates which of the $n-1$ partial minimum-variance portfolios will yield the maximum last period portfolio return. Finally, we consider the buy-and-hold strategy, which makes no change in the allocation of capital in different assets. The initial portfolio for the buy-and-hold strategy is set to be the $1/n$ portfolio in all the experiments[3].

3.3 Methodology

In each data set, we apply the rolling-horizon procedure to conduct the empirical analysis. Consider the benchmark portfolios chosen from one of the two groups listed in Table 2. The details of the methodology are summarized as follows:

1. Denote the total number of returns in the data set to be $\hat{\tau}$. We choose a history of length τ with $\tau < \hat{\tau}$, over which we conduct the estimation. In our experiments, $\tau = 240$, which corresponds to 20 years of monthly data.

2. Using the data in the estimation window, we estimate the parameters $\boldsymbol{\mu}$, Σ, $\mathbf{E}[Z_B(\tilde{\boldsymbol{r}})]$, and $P\mathbf{E}[\boldsymbol{p}(\tilde{\boldsymbol{r}})]$, and compute the portfolios of investment strategies listed in Table 2.

3. The performance of the portfolios are then evaluated. The details of these measures are discussed in the next subsection.

4. Roll forward the time horizon by adding the next data point of the data set and dropping the first data point of the current estimation window.

5. By doing this repeatedly until the last time period, we obtain $\hat{\tau} - \tau$ portfolio-weight vectors for each strategy.

[3]We have tested various initial portfolio positions, and found that the results for the buy-and-hold strategy are not sensitive to this initial condition.

3.4 Performance Measures

Let t index the time periods with $t = 1, \ldots, \hat{\tau}$. We compute the portfolios at the beginning of period t for $t = \tau + 1, \ldots, \hat{\tau}$ using past information from the previous τ periods. Let $\tilde{\boldsymbol{r}}^t$ denote the return for period t. For a portfolio strategy, we use $\boldsymbol{x}^t$ to represent the investment decision made for period t, and $\boldsymbol{x}^{t0}$ to represent the portfolio position at the beginning of period t before the repositioning decision $\boldsymbol{x}^t$ is made. The performance measures are listed as follows:

1. *In-sample tracking error*:

$$\frac{1}{\hat{\tau} - \tau} \sum_{t=\tau+1}^{\hat{\tau}} \sum_{t'=t-\tau}^{t-1} \left[\left(\tilde{\boldsymbol{r}}^{t'} \right)^T \boldsymbol{x}^t - Z_B \left(\tilde{\boldsymbol{r}}^{t'} \right) \right]^2 .$$

2. *Turnover rate*:

$$\frac{1}{\hat{\tau} - \tau} \sum_{t=\tau+1}^{\hat{\tau}} \sum_{i=1}^{n} \left| x_i^t - x_i^{t0} \right| .$$

3. *Out-of-sample Sharpe ratio net of proportional transaction costs of 50 basis point (net Sharpe ratio)*:

$$\frac{\widehat{\mu}}{\widehat{\sigma}},$$

where

$$\widehat{\mu} = \frac{1}{\hat{\tau} - \tau} \sum_{t=\tau+1}^{\hat{\tau}} \left[\left(1 + (\tilde{\boldsymbol{r}}^t)^T \boldsymbol{x}^t \right) \left(1 - 0.005 \cdot \left| x_i^t - x_i^{t0} \right| \right) - 1 \right],$$

and

$$\widehat{\sigma} = \sqrt{ \frac{1}{\hat{\tau} - \tau - 1} \sum_{t=\tau+1}^{\hat{\tau}} \left[\left(1 + (\tilde{\boldsymbol{r}}^t)^T \boldsymbol{x}^t \right) \left(1 - 0.005 \cdot \left| x_i^t - x_i^{t0} \right| \right) - 1 - \widehat{\mu} \right]^2 }.$$

4. *Out-of-Sample Net Aggregate Return*:

$$\begin{cases} \prod_{t=\tau+1}^{T} \left(1 + \left(\tilde{\boldsymbol{r}}^t \right)^T \boldsymbol{x}^t \right) \left(1 - 0.005 \cdot \left| x_i^t - x_i^{t0} \right| \right), \\ \qquad \text{if } \left(1 + \left(\tilde{\boldsymbol{r}}^t \right)^T \boldsymbol{x}^t \right) \left(1 - 0.005 \cdot \left| x_i^t - x_i^{t0} \right| \right) > 0, \, \forall t = \tau + 1, \\ 0, \\ \qquad \text{otherwise.} \end{cases}$$

The out-of-sample net aggregate return measures the long-term wealth growth of the portfolio strategies, where the second situation represents bankruptcy. It is possible since the model allows short sales.

3.5 Normality Assumption

The closed-form solutions in the previous section are established under the assumption that the return follows a multivariate normal distribution. We first check the validity of the normality assumption by drawing the Quartile-Quartile plot (QQ plot) of the Mahalanobis distance of the data from the first estimation window against that of a multivariate normal distribution for each data set. For comparison, we use the sample mean and sample covariance in place of their respective true values for the multivariate normal distribution. The plots are presented in Figure 5.

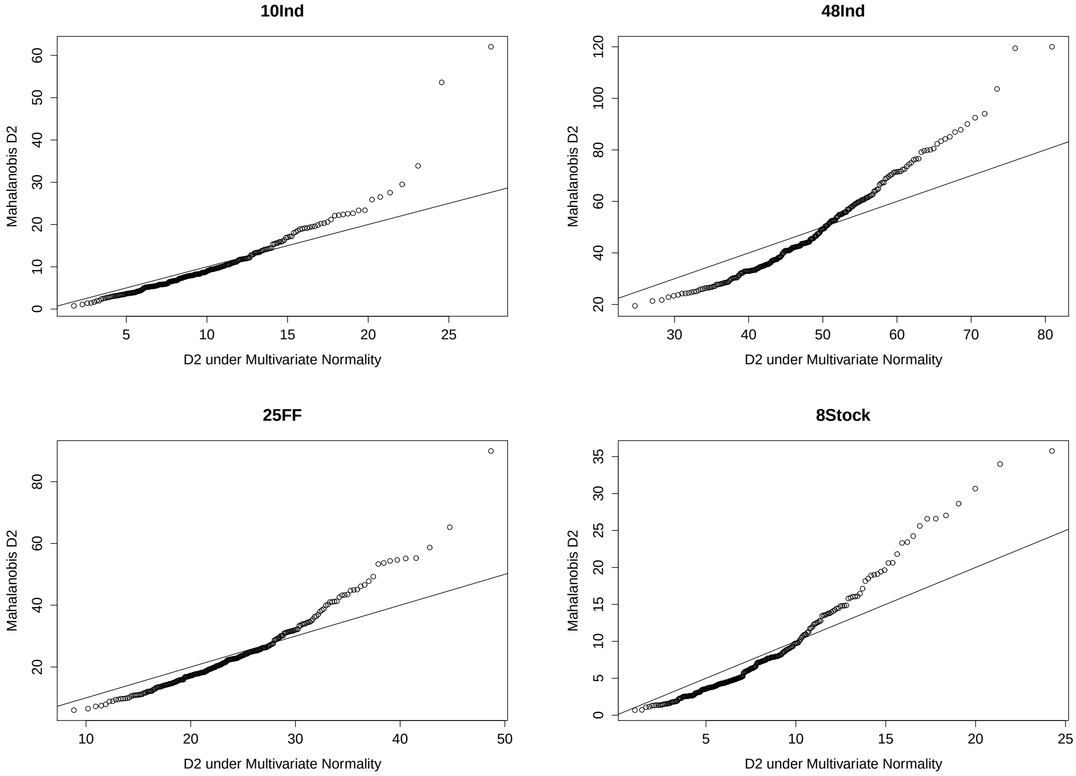


Figure 5: QQ plots of the distributions of asset returns against multivariate normal distribution

From these QQ plots, we observe that the sample Mahalanobis distance of the risky asset returns in all data sets demonstrate significant deviation from the normality assumption with heavy tails. However, as we will see later, the discrepancies shown in the QQ plots do not appear to be a major problem, and our multiple-benchmark tracking-error portfolio demonstrates greater superiority in the out-of-sample performance even though the normality assumption might not be completely satisfactory.

3.6 Results and Discussion

In this subsection, we first present results on the basic model, where the buy-and-hold strategy is included in the set of benchmark portfolios to control the transaction volume. Next, the results on the extended model, i.e., Problem (TC), are discussed.

3.6.1 Multiple-Benchmark Tracking-Error Portfolio with the Buy-and-Hold Strategy as a Benchmark

We first report results on the tracking error for each data set to show that our portfolio provides good tracking of the highest benchmark return. Next, we use all the other performance measures discussed before to evaluate the performance of our portfolio against its benchmarks. Finally, we discuss the results of the robustness tests on net aggregate returns, where we consider random starting times and random lengths of the investment horizon to gauge the out-of-sample performance of our portfolio strategy. Since all the performance measures except the tracking error are computed out-of-sample, we often drop the descriptive terms, "out-of-sample" and "in-sample", in the following discussion.

In all the experiments here, we use two benchmark portfolios to obtain our multiple-benchmark tracking-error portfolio, one from Portfolio 1–10 in Table 2, and the other is Portfolio 11 in Table 2, i.e., the buy-and-hold strategy. We report the results of the first benchmark portfolio and the corresponding multiple-benchmark tracking-error portfolio, as the buy-and-hold strategy for each single period only serves the purpose of controlling for transaction volume.

Tracking Error

As the portfolio is constructed to track a set of benchmark portfolios, the first step is to evaluate how closely our portfolio tracks the best return from the benchmark portfolios. The results on tracking errors are summarized in Table 3. Comparing the performance of the MBTE portfolio with that of its benchmarks, we observe that the tracking error of the MBTE portfolio is always smaller than those of its benchmarks. Note that the returns are not exactly normally distributed, so it is not guaranteed that the MBTE portfolio would be the best even in terms of in-sample mean squared tracking errors. However, the results in Table 3 provide partial justification that the MBTE portfolio might still perform well even when the normality assumption is not completely satisfied.

To visualize the tracking error, we plot the percentage decrease in in-sample mean squared tracking errors at every period from our portfolio compared to the PARR portfolio in Figure 6. Our multiple-benchmark tracking-error portfolio is constructed by using the buy-and-hold strategy and the PARR portfolio as benchmark portfolios. The figure shows a significant difference in tracking errors between the two portfolios. All the differences are positive, and on average, the in-sample tracking error of the MBTE portfolio demonstrates a 40% to 60% reduction from that of the PARR portfolio, which shows that the MBTE portfolio better tracks the highest return from the group of benchmark portfolios. Figures for the other scenarios are similar, so only one is reported here as an illustration.

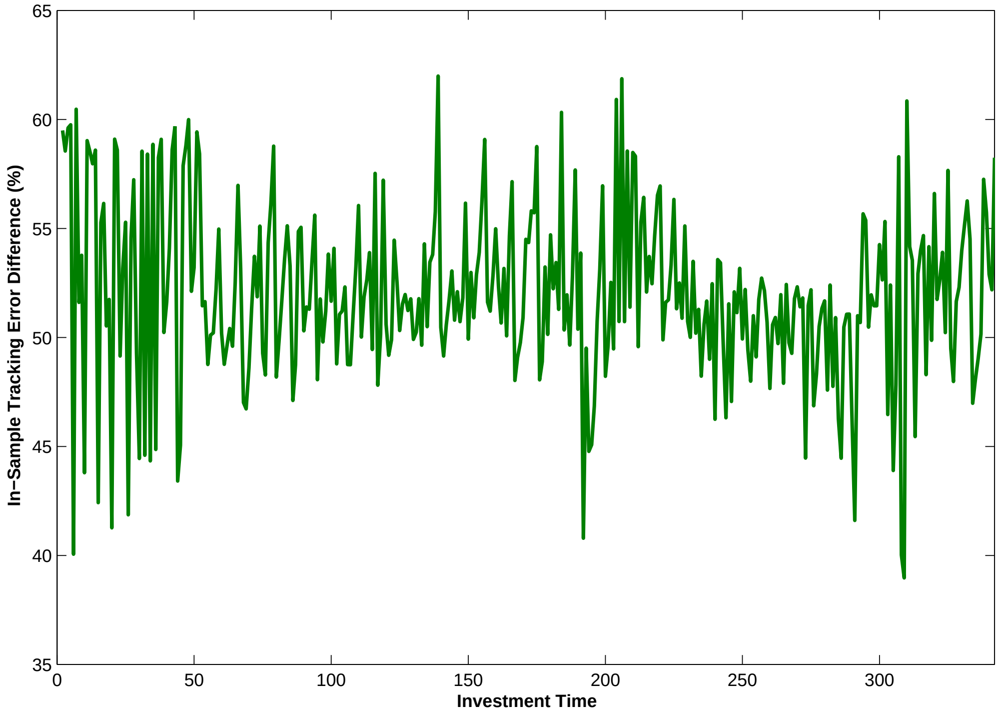
Out-of-Sample Net Sharpe Ratio

Table 4 reports the out-of-sample net Sharpe ratio and the corresponding p-value that the net Sharpe ratio of each benchmark strategy is smaller than that of the respective multiple-benchmark tracking-error portfolio. The one-sided p-values are computed based on the studentized circular block bootstrapping method used in Ledoit & Wolf (2008). From the table, we observe that the MBTE portfolio consistently dominates the benchmark portfolio used in its construction. In particular, the MBTE portfolio has higher net Sharpe ratios than the $1/n$ portfolio and all the differences are significant at 0.005 level for all except the "8Stock" data set. Moreover, the MBTE portfolio has higher net Sharpe ratios than the MUL portfolio across all the data sets, and all the differences are significant at 0.05 level. The MBTE portfolio also outperforms the PARR portfolio across all the data sets, and the differences are significant at 0.005 level in all but the "8Stock"

10Ind	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	3.772E-5	1.696E-5	2.527E-2	1.232E-2	6.447 E-4	2.939E-4	7.120E-4	2.962E-4	6.912E-4	2.863E-4
	(44.96%)		(48.74%)		(45.58%)		(41.60%)		(41.41%)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	6.102E-4	2.520E-4	5.898E-4	2.433E-4	1.778E-4	7.778E-5	3.596E-5	1.608E-05	1.823E-3	8.205E-4
	(41.30%)		(41.25%)		(43.75%)		(44.72%)		(45.00%)	
48Ind	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	1.319E-4	2.895E-5	1.270E-1	3.901E-2	4.850E-3	1.658E-3	5.328E-3	1.805E-3	4.668E-3	1.580E-3
	(21.95%)		(30.72%)		(34.18%)		(33.87%)		(33.84%)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	4.294E-3	1.403E-3	4.224E-3	1.388E-3	6.319E-4	1.746E-4	1.243E-4	2.749E-5	5.528E-2	1.515E-2
	(32.68%)		(32.86%)		(27.63%)		(22.12%)		(27.40%)	
25FF	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	3.626E-5	1.023E-5	7.647E-2	2.502E-2	2.630E-3	9.705E-4	2.949E-3	9.991E-4	2.684E-3	9.067E-4
	(28.23%)		(32.72%)		(36.91%)		(33.88%)		(33.78%)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	2.289E-3	7.833E-4	2.316E-4	7.915E-4	2.509E-4	8.129E-5	3.216E-5	9.357E-6	9.407E-3	3.304E-3
	(34.22%)		(34.17%)		(32.40%)		(29.09%)		(35.12%)	
8Stock	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	4.064E-4	1.945E-4	2.184E-2	1.082E-2	9.718E-4	4.655E-4	1.843E-3	9.255E-4	1.891E-3	9.527E-4
	(47.87%)		(49.53%)		(47.90%)		(50.22%)		(50.38%)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	1.762E-3	8.873E-4	1.439E-3	7.218E-4	1.641E-3	8.227E-4	3.964E-4	1.891E-4	2.035E-3	1.004E-3
	(50.36%)		(50.16%)		(50.14%)		(47.71%)		(49.31%)	

Note: This table reports the in-sample tracking error and the corresponding percentage reduction in in-sample tracking error by the MBTE portfolio (in brackets).

Table 3: Comparison on in-sample tracking error



Notes: This figure plots the percentage decrease in the in-sample mean squared tracking errors at every period from the MBTE portfolio compared to the PARR portfolio.

Figure 6: Tracking-error difference (in percentage) between the PARR portfolio and our multiple-benchmark tracking-error portfolio using the buy-and-hold strategy and the PARR portfolio as benchmarks in the "10Ind" data set

data set. Additionally, the MBTE portfolio shows significant difference from the MEAV, GMC, M1FAC, and CVAR portfolios in the "25FF" data set. It is worth noting that except the case when CVAR is used as one benchmark, all the other MBTE portfolios have much higher net Sharpe ratio than the $1/n$ portfolio, independent of the choice on the benchmark portfolio, in all the data sets we consider.

Turnover Rate

Table 5 reports the turnover rate and the corresponding percentage deduction in turnover rate by the MBTE portfolio. From the table, we observe that the MBTE portfolio has lower turnover rates than its respective benchmark portfolio in all cases except for the GMC portfolio in the "48Ind" data set, where the GMC portfolio has a slightly smaller turnover rate than the corresponding MBTE portfolio. This is exactly the desired effect of introducing the buy-and-hold strategy as one of the benchmark portfolios. In particular, the turnover rates of its corresponding MBTE portfolio are only half as large as the turnover rates of the PARR portfolio across all the data sets. The MBTE portfolio demonstrates a decrease in turnover rate of over 40% from the MUL portfolio across all the data sets. In addition, the turnover rates of the respective MBTE portfolios are at least 30% less than those of the $1/n$ and MEAV portfolios, and the turnover rates of the respective MBTE portfolios are over 20% less than those of the MIND, M1PAR, MINU,

10Ind	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	0.2260	0.2275	0.1941	0.2582	0.2809	0.2858	0.2915	0.2957	0.2929	0.2966
	(0.0050***)		(0.0005***)		(0.4785)		(0.2018)		(0.2483)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	0.2919	0.2950	0.2919	0.2948	0.2616	0.2626	0.2274	0.2285	0.2536	0.2593
	(0.3322)		(0.3222)		(0.1923)		(0.0290*)		(0.1818)	
48Ind	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	0.1805	0.1857	0.0759	0.2383	0.2560	0.2747	0.2277	0.2447	0.2474	0.2553
	(0.0005***)		(0.0015***)		(0.0790)		(0.3617)		(0.2922)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	0.2608	0.2678	0.2588	0.2640	0.2693	0.2704	0.1832	0.1877	0.1009	0.1450
	(0.3237)		(0.3477)		(0.4231)		(0.0025***)		(0.0015***)	
25FF	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	0.1999	0.2030	-0.02499	0.2530	0.3803	0.4046	0.3720	0.3965	0.3590	0.3791
	(0.0010**)		(0.0005***)		(0.0005***)		(0.0005***)		(0.0020***)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	0.3481	0.3615	0.3587	0.3716	0.2038	0.2064	0.2016	0.2039	0.3776	0.4115
	(0.0345*)		(0.0375*)		(0.0065**)		(0.0005***)		(0.0195*)	
8Stock	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	0.1807	0.1836	0.1982	0.2074	0.1773	0.1814	0.2551	0.2637	0.2567	0.2653
	(0.2253)		(0.3515)		(0.1514)		(0.1893)		(0.1963)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	0.2562	0.2649	0.2491	0.2574	0.2681	0.2756	0.1853	0.1854	0.2348	0.2443
	(0.1714)		(0.1504)		(0.1753)		(0.036*)		(0.0924)	

Note: This table reports the monthly out-of-sample Sharpe ratio net of proportional transaction costs of 50 basis points and the corresponding one-sided p-value (in brackets) that the net Sharpe ratio of each benchmark strategy is smaller than that of the respective MBTE portfolio. Star symbols are included for p-values: (.*) for significance at 0.05 level, (.**) for 0.01, and (.***) for 0.005.

Table 4: Comparison on net Sharpe ratio

M1FAC, and CVAR portfolios.

Net Aggregate Return

Table 6 reports the out-of-sample net aggregate return. From the table, we observe that the MBTE portfolios almost always dominate their respective benchmark portfolios over the whole investment period (07/1983–12/2011 for the "10Ind", "48Ind", "25FF" data sets, and 08/2000–01/2013 for the "8Stock" data set). In particular, the MBTE portfolio shows great superiority over the PARR portfolio. The net aggregate return of the MBTE portfolio is twice as large as that of the PARR portfolio in the "10Ind" data set, nearly 10 times larger in the "48Ind" portfolio, and 37 times larger in the "25FF" data set. Furthermore, the MBTE portfolio outperforms the CVAR portfolio by 88.6% in the "25FF" data set and the MINU portfolio by 38.6% in the same data set. However, the net aggregate return of the MBTE portfolio does not always outperform the GMC portfolio or the MUL portfolio, though the difference is small (of order 0.1%) in these two cases.

After all, these aggregate returns only provide partial information as we only consider one investment horizon. To demonstrate the robustness of the findings, we consider random starting times and random lengths of the investment horizon. Some results for the "48Ind" data set are provided in Figure 7 and 8[4].

In both Figure 7 and 8, we observe that the net aggregate returns of our multiple-benchmark tracking-error portfolios are constantly higher than or comparable to those of the $1/n$ portfolio. Figure 7 shows that the shorter the investment periods, the less difference in net aggregate returns between the MBTE portfolio and the $1/n$ portfolio. When the investment activity is conducted for the whole 28 years, we observe a clear difference between the net aggregate returns from these two portfolios.

Note that the MBTE portfolios in Figure 8 are obtained using the PARR and buy-and-hold portfolios as benchmarks, which do not contain the $1/n$ portfolio, but the performance of the $1/n$ portfolio at the same time periods are included for comparison. It is interesting to observe that although the PARR portfolio dominates the $1/n$ portfolio in terms of Sharpe ratio, when the transaction costs are considered, the resulting performance of the PARR portfolio is usually worse than the $1/n$ portfolio. However, if we put the PARR portfolio together with the buy-and-hold strategy in the set of benchmark portfolios, our model yields a new portfolio that combines the strength of both portfolios. The resulting portfolio provides a high level of return while requiring much less transaction, and the net aggregate returns clearly outperform both the $1/n$ portfolio and the PARR portfolio.

3.6.2 Multiple-Benchmark Tracking-Error portfolio with Penalty on Transaction Volume

We dedicate this part to evaluate the alternative way to control transaction volume as proposed in Section 2.6, in which we penalize the transaction volume directly in the objective function. In particular, we solve Problem (TC) to obtain the multiple-benchmark tracking-error portfolio with

[4]We have done similar tests on all the other data sets, and tried many other random starting points with random investment horizons. As the findings are similar, we only report part of the results here.

10Ind	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	0.0238	0.0137	0.9075	0.4094	0.1281	0.0820	0.0962	0.0627	0.0900	0.0593
	(42.44%)		(54.89%)		(35.99%)		(34.82%)		(34.11%)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	0.0801	0.0531	0.0773	0.0514	0.0262	0.0184	0.0238	0.0133	0.1852	0.1292
	(33.71%)		(33.51%)		(29.77%)		(45.72%)		(37.3%)	
48Ind	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	0.0306	0.0170	2.4015	0.9243	0.4632	0.2729	0.4459	0.2591	0.3026	0.1835
	(44.44%)		(61.51%)		(41.08%)		(41.89%)		(39.36%)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	0.3004	0.1809	0.2815	0.1705	0.0266	0.0267	0.0304	0.0165	1.4778	0.9257
	(39.78%)		(39.43%)		(−0.38%)		(45.72%)		(37.36%)	
25FF	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	0.0174	0.0120	3.1537	1.2923	0.4368	0.2699	0.4278	0.2633	0.3520	0.2193
	(31.03%)		(59.02%)		(38.21%)		(38.45%)		(37.70%)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	0.2569	0.1644	0.2532	0.1628	0.0347	0.0286	0.0173	0.0099	1.0400	0.6760
	(36.00%)		(35.70%)		(17.58%)		(42.77%)		(35.00%)	
8Stock	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	0.0406	0.0222	0.4034	0.1714	0.0540	0.0360	0.0381	0.0289	0.0373	0.0285
	(45.32%)		(57.51%)		(33.33%)		(24.15%)		(23.59%)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	0.0374	0.0282	0.0387	0.0278	0.0371	0.0274	0.0407	0.0217	0.0511	0.0371
	(24.60%)		(28.17%)		(26.15%)		(46.68%)		(27.40%)	

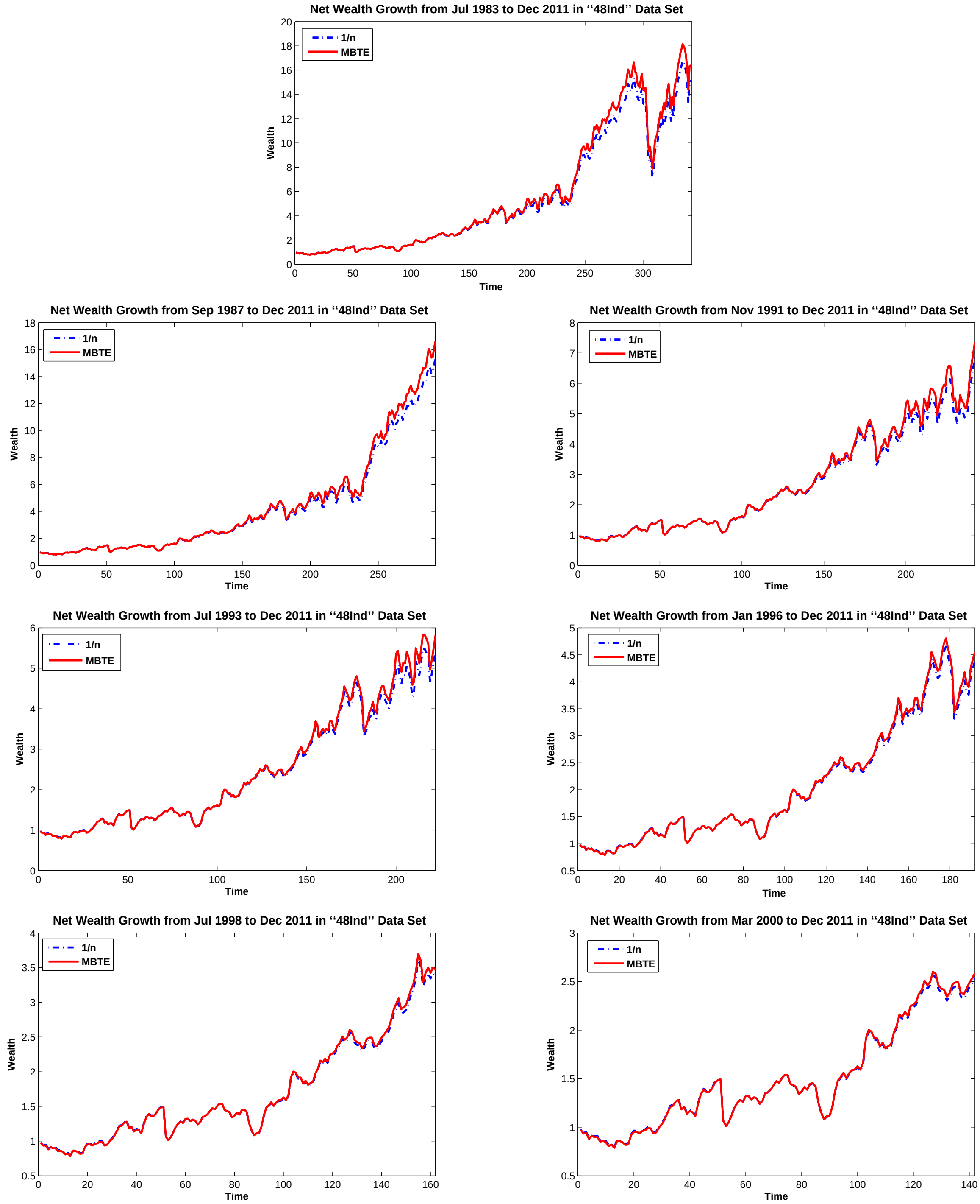
Note: This table reports the turnover rate and the corresponding percentage deduction in turnover rate by the MBTE portfolio (in brackets).

Table 5: Comparison on turnover rate

10Ind	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	19.2912	19.7050	8.8949	18.5599	26.7493	28.9076	31.5124	32.8691	31.4552	32.6069
	(2.15%)		(108.66%)		(8.07%)		(4.31%)		(3.66%)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	30.5331	31.5403	30.2300	31.1623	22.8364	23.2083	19.5630	19.8028	21.9975	22.8334
	(3.30%		(3.08%)		(1.63%)		(1.23%)		(3.80%)	
48Ind	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	15.0473	16.5383	2.0602	22.4944	22.6779	29.4095	18.5407	19.2705	17.7096	18.6572
	(9.91%)		(991.86%)		(29.68%)		(3.94%)		(5.35%)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	19.3608	20.3869	18.5407	19.2705	24.9412	24.8306	15.6788	16.8206	16.1426	19.2452
	(5.30%)		(3.94%)		(-0.44%)		(7.28%)		(19.22%)	
25FF	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	20.5105	21.8385	0.5317	20.2806	97.0287	134.3689	87.0998	120.7095	75.695	98.1391
	(6.47%)		(3714.29%)		(38.48%)		(38.59%)		(29.65%)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	65.7026	79.5398	73.8475	89.0208	17.4811	18.1652	20.9800	21.7274	234.5823	442.5228
	(21.06%)		(20.55%)		(3.91%)		(3.56%)		(88.64%)	
8Stock	$1/n$	MBTE	PARR	MBTE	MEAV	MBTE	MINU	MBTE	M1FAC	MBTE
	2.0784	2.0834	2.0900	2.1405	1.9623	1.9896	2.6711	2.7580	2.6878	2.7754
	(0.24%)		(2.42%)		(1.39%)		(3.25%)		(3.26%)	
	MIND	MBTE	M1PAR	MBTE	GMC	MBTE	MUL	MBTE	CVAR	MBTE
	2.6622	2.7484	2.5553	2.6308	2.7881	2.8705	2.1102	2.0995	2.5372	2.6398
	(3.24%)		(2.95%)		(2.96%)		(-0.51%)		(4.04%)	

Note: This table reports the out-of-sample net aggregate return and the corresponding percentage increment in net aggregate return by the MBTE portfolio (in brackets).

Table 6: Comparison on net aggregate return



Note. From the group of plots, we observe that the MBTE portfolio tracks its benchmark (the $1/n$ portfolio) very well while surpassing the benchmark return now and then.

Figure 7: Wealth growth of the multiple-benchmark tracking-error (MBTE) portfolio using the $1/n$ and buy-and-hold portfolios as benchmarks, and the $1/n$ portfolio with random starting times and evaluation periods in the “48Ind” data set

a penalty on transaction volume (MBTEP). We choose the $1/n$, M1FAC, and CVAR portfolios as benchmarks. In choosing the penalty parameter, ν, we use an in-sample calibration approach, where the transaction volume of the multiple-benchmark tracking-error portfolio is restricted to be less than that of the $1/n$ portfolio in the last period of estimation window. We fix $W = 1$

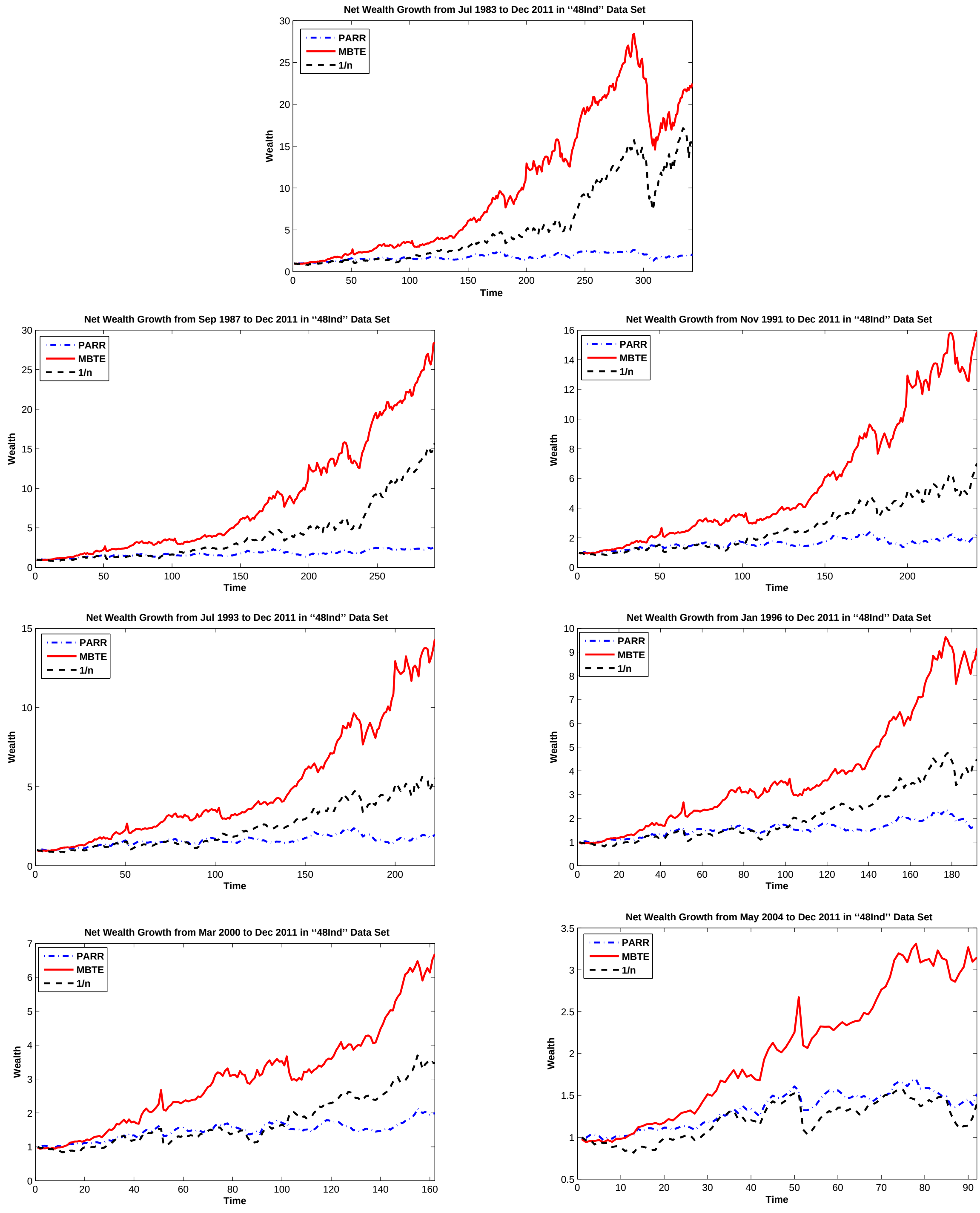


Figure 8: Wealth growth of the multiple-benchmark tracking-error (MBTE) portfolio using the PARR and buy-and-hold portfolios as benchmarks, the $1/n$ portfolio, and the PARR portfolio with random starting time for evaluation period in the “48Ind” data set

throughout the investment horizon to facilitate the search of ν.

Table 7 presents a comparison of the performance between our multiple-benchmark tracking-error portfolio (MBTEP) and the $1/n$ portfolio. From the table, we observe that the in-sample tracking errors of the MBTEP portfolio are consistently smaller than that of the $1/n$ portfolio across all the data sets. The net Sharpe ratios of our MBTEP portfolio outperform the $1/n$ portfolio across all the data sets, and the difference is statistically significant in all but the "8Stock" data set. Furthermore, over the whole investment period of the data sets, the MBTEP portfolio tends to yield higher net aggregate returns than the $1/n$ portfolio[5]. However, the turnover rates of the MBTEP portfolio are always higher than those of the $1/n$ portfolio. This is expected as the in-sample calibration of the penalty parameter might induce large out-of-sample turnovers.

Overall, the MBTEP portfolio provides better results than the $1/n$ portfolio. The flexibility in choosing a value for the penalty parameter could be either a bonus or a burden as determining the value is a judgment call of the portfolio manager. The desired performance can only be induced by appropriately chosen penalty values. Adding the buy-and-hold strategy to the set of benchmarks seems to be more natural and effective in controlling the transaction volume.

Data Set	Portfolio Model	Tracking Error	Net Sharpe Ratio	Turnover Rate	Net Aggregate Return
10Ind	$1/n$	0.1130	0.2260	0.0238	19.2912
	MBTEP	0.09597	0.2364	0.0370	19.3359
		(15.03%)	(0.0460*)	(−35.68%)	(0.23%)
48Ind	$1/n$	0.7323	0.1805	0.0306	15.0473
	MBTEP	0.6160	0.2224	0.0723	18.5023
		(15.89%)	(0.0050***)	(−57.68%)	(22.96%)
25FF	$1/n$	0.4768	0.1999	0.0174	20.5105
	MBTEP	0.4149	0.2317	0.0365	29.8953
		(12.97%)	(0.0040***)	(−52.33%)	(45.76%)
8Stock	$1/n$	0.02817	0.1807	0.0406	2.0784
	MBTEP	0.01712	0.2126	0.0493	2.2427
		(39.22%)	(0.1169)	(−17.65%)	(7.91%)

Note: The number inside the brackets under "Tracking Error" column is the corresponding percentage decrease in in-sample tracking error by the MBTE portfolio from the $1/n$ portfolio. The number inside the brackets under "Net Sharpe Ratio" column is the corresponding one-sided p-value that the net Sharpe ratio of the $1/n$ portfolio is smaller than that of the respective MBTE portfolio. Star symbols are included for p-values: (.*) for significance at 0.05 level, (.**) for 0.01, and (.***) for 0.005. The number inside the brackets under "Turnover Rate" column is the corresponding percentage deduction in turnover rate by the MBTE portfolio from the $1/n$ portfolio. Negative numbers indicate increased turnover rate. The number inside the brackets under "Net Aggregate Return" column is the corresponding percentage increment in net aggregate return by the MBTE portfolio from the $1/n$ portfolio.

Table 7: Comparison on the performance of the $1/n$ portfolio and the multiple-benchmark tracking-error portfolio with penalty on transaction volume (MBTEP)

4 Conclusion

We propose a new multiple-benchmark tracking-error model for portfolio selection. The target return being tracked is the highest return from a set of given benchmark portfolios. Our model

[5]We have also conducted a robustness test in this case. As the results are similar as before, we do not report it here.

differs from existing literature by directly capturing the concerns arising from real investment activities. By resorting to Stein's identity, we obtain the closed-form expression for the optimal portfolio weights under the assumption of normal return distribution. The closed-form solution reveals that persistency is the basic component of the optimal portfolio, which partially explains the robustness of our portfolio against the estimation errors, as the probability of one benchmark outperforming the rest are less prone to estimation errors on the expected returns of the financial assets.

The closed-form solution allows us to conduct more in-depth analysis of our model, especially, the comparison with the Markowitz mean-variance portfolio and the linear combination rule proposed by Tu & Zhou (2011). In particular, we showed that the linear combination rule would be inferior to our portfolio if the portfolio manager has a mean-variance utility with low risk aversion. This further strengthens the motivation of our multiple-benchmark tracking-error model. In addition, we prove that the entire mean-variance efficient frontier can be generated from our model when two distinct mean-variance portfolios are used as the benchmark portfolios, a result similar to the well-known Two-Fund Theorem.

To address the common problem of whether to reposition the portfolio, our modelling framework allows a natural solution by including the buy-and-hold strategy as one of the benchmark portfolios. Our numerical analysis showed that adding the buy-and-hold strategy as a benchmark can significantly reduce the turnover rate, which might be attractive to investors when transaction costs are considerable. When combining the buy-and-hold strategy with other benchmarks, we demonstrated using the real data sets that our portfolio has consistently provided higher net Sharpe ratio, higher net aggregate return, and lower turnover rate compared to the benchmark portfolios, in particular, the $1/n$ portfolio, a well-known tough benchmark to beat.

Although our theoretical analysis is built upon the assumption that the return distribution is multivariate normal, the results from the numerical analysis shows that the power of our model framework is strong enough to cover the violation of the assumption in real data. On the other hand, there have been some extensions of Stein's Identity to other probability distributions (cf. Adcock (2007), Barbour et al. (1992)). It would be interesting to extend some of the results in this paper to these cases.

References

Adcock, C. J. (2007) *Extensions of Stein's Lemma for the skew-normal distribution*, Communications in Statistics–Theory and Methods, **36**, pp. 1661–1671.

Alexander, G. J., A. M. Baptista (2008) *Active portfolio management with benchmarking: Adding a value-at-risk constraint*, Journal of Economic Dynamics and Control, **32**, pp. 779–820.

Barbour, A. D., L. H. Y. Chen, W. L. Loh (1992) *Compound poisson approximation for nonnegative random variables via stein's method*, Annals of Probability, **20**, pp. 1843–1866.

Bertsimas, D., K. Natarajan, C. P. Teo (2006) *Persistence in discrete optimization under data uncertainty*, Mathematical Programming, **108**, pp. 251–274.

Best, M. J., R. R. Grauer (1991) *On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results*, The Review of Financial Studies, **4**, pp. 315–342.

Brodie, J., I. Daubechies, C. D. Mol, D. Giannone, I. Loris (2009) *Sparse and stable Markowitz portfolios*, Proceedings of the National Academy of Sciences, **106**, pp. 12267–12272.

Chen H. H., H. T. Tsai, D. Lin et al. (2011) *Optimal mean-variance portfolio selection using Cauchy-Schwarz maximization*, Applied Economics, **43**, pp. 2795–2801.

Cornell, B., R. Roll (2005) *A delegated-agent asset-pricing model*, Financial Analysts Journal, **61**, pp. 57–69.

Cover, T. M. (1991) *Universal portfolios*, Mathematical Finance, **1**, pp. 1–29.

DeMiguel, V., L. Garlappi, F. J. Nogales, R. Uppal (2009) *A Generalized approach to portfolio optimization: Improving performance by constraining portfolio norms*, Management Science, **55**, pp. 798–812.

DeMiguel, V., L. Garlappi, R. Uppal (2007) *Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?*, The Review of Financial Studies, **22**, pp. 1915–1953.

El-Hassan, N., P. Kofman (2003) *Tracking error and active portfolio management*, Australian Journal of Management, **28**, pp. 183–207.

Helmbold, D. P., R. E. Schapire, Y. Singer, M. K. Warmuth (1998) *On-line portfolio selection using multiplicative updates*, Mathematical Finance, **8**, pp. 325–347.

Jorion, P. (2003) *Portfolio optimization with constraints on tracking error*, Financial Analysts Journal, **59**, pp. 70–82.

Jagannathan, R., T. Ma (2003) *Risk reduction in large portfolios: Why imposing the wrong constraints helps?*, Journal of Finance, **58**, pp. 1651–1684.

Kirby, C., B. Ostdiek (2012) *It's all in the timing: Simple active portfolio strategies that outperform naive diversification*, Journal of Financial and Quantitative Analysis, **47**, pp. 437–467.

Ledoit, O., M. Wolf (2003) *Improved estimation of the covariance matrix of stock returns with an application to portfolio selection*, Journal of Empirical Finance, **10**, pp. 603–621.

Ledoit, O., M. Wolf (2004) *A well-conditioned estimator for large-dimensional covariance matrices*, Journal of Multivariate Analysis, **88**, pp. 365–411.

Ledoit, O., M. Wolf (2008) *Robust performance hypothesis testing with the Sharpe ratio*, Journal of Empirical Finance, **15**, pp. 850–859.

Liu, J. S. (1994) *Siegel's formula via Stein's identities*, Statistics & Probability Letters, **21**, pp. 247–251.

Markowitz, H. M. (1952) *Portfolio selection*, Journal of Finance, **7**, pp. 77–91.

Michaud, R. O. (1989) *The Markowitz optimization enigma: Is 'optimized' optimal?*, Financial Analysts Journal, **45**, pp. 31–42.

Mishra, V. K., K. Natarajan, H. Tao, C. P. Teo (2012) *Choice prediction with semidefinite optimization when utilities are correlated*, IEEE Transactions on Automatic Control, **57**, pp. 2450–2463.

Natarajan, K., M. Song, C. P. Teo (2009) *Persistency model and its applications in choice modeling*, Management Science, **55**, pp. 453–469.

Rockafellar, R. T., S. Uryasev (2000) *Optimization of conditional value-at-risk*, Journal of Risk, **2**, pp. 21–42.

Rockafellar, R. T., S. Uryasev (2004) *Conditional value-at-risk for general loss distributions*, Journal of Banking and Finance, **26**, pp. 1443–1471.

Roll, R. (1992) *A mean/variance analysis of tracking error*, Journal of Portfolio Management, **18**, pp. 13–22.

Rudolf, M., H-J. Wolter, H. Zimmermann (1999) *A linear model for tracking error minimization*, Journal of Banking & Finance, **23**, pp. 85–103.

Rustem, B., Howe, M. (2002) *Algorithms for Worst-Case Design and Applications to Risk Management*, Princeton University Press, pp. 261–271.

Stein, C. M. (1972) *A bound for the error in the normal approximation to the distribution of a sum of dependent random variables*, Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, **2**, pp. 583–602.

Stein, C. M. (1981) *Estimation of the mean of a multivariate normal distribution*, The Annals of Statistics, **9**, pp. 1135–1151.

Steinbach, M. C. (2001) *Markowitz revisited: Mean-variance models in financial portfolio analysis*, SIAM Review, **43**, pp. 31–85.

Tu, J., G. Zhou (2011) *Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies*, Financial Analysts Journal, **55**, pp. 63–72.

Wang, M. Y. (1999) *Multiple-Benchmark and Multiple-Portfolio Optimization*, Journal of Financial Economics, **99**, pp. 204–215.

Appendix A. Proof of Stein's Identity

The proof is consolidated from Stein (1972), Stein (1981), and Liu (1994).

We begin by showing the uni-variate version of Lemma 1 (cf. Stein (1972) and Stein (1981)).

Let $\tilde{y}$ follow a standard normal distribution $N(0,1)$, and $\phi(y)$ denote the standard normal density with the derivative $\phi'(y) = -y\phi(y)$. For any function $g : \mathbb{R} \to \mathbb{R}$ such that g' exists almost everywhere and $\mathbf{E}[|g'(\tilde{y})|] < \infty$,

$$\begin{aligned}
\mathbf{E}\left[g'(\tilde{y})\right] &= \int_{-\infty}^{\infty} g'(y)\phi(y)\,dy \\
&= \int_{0}^{\infty} g'(y)\left[-\int_{y}^{\infty} -z\phi(z)\,dz\right]dy + \int_{-\infty}^{0} g'(y)\left[\int_{-\infty}^{y} -z\phi(z)\,dz\right]dy \\
&= \int_{0}^{\infty} z\phi(z)\left[\int_{0}^{z} g'(y)dy\right]dz - \int_{-\infty}^{0} z\phi(z)\left[\int_{z}^{0} g'(y)dy\right]dz \\
&= \left(\int_{0}^{\infty} + \int_{-\infty}^{0}\right)\left[z\phi(z)\left[g(z) - g(0)\right]\right]dz \\
&= \int_{-\infty}^{\infty} z\phi(z)\,g(z)dz \\
&= \mathbf{E}\left[\tilde{y}g(\tilde{y})\right],
\end{aligned}$$

where the third equality is justified by Fubini's Theorem. Note that since $\mathbf{E}[\tilde{y}] = 0$ and $Var(\tilde{y}) = 1$, the equality proved above is essentially

$$Cov\left(\tilde{y}, g(\tilde{y})\right) = Var(\tilde{y})\mathbf{E}\left[g'(\tilde{y})\right]. \tag{4}$$

Next, the result is generalized into the multivariate case (cf. Stein (1981) and Liu (1994)).

Let $\tilde{\boldsymbol{z}} = (\tilde{z}_1, \ldots, \tilde{z}_n)^T$, where $\tilde{z}_j$'s are independent and identically distributed standard normal random variables. It is straightforward to show by Equation (4) that for any function $\hat{h} : \mathbb{R}^n \to \mathbb{R}$ satisfying the same conditions as h,

$$\mathbf{E}\left[\tilde{z}_j\hat{h}(\tilde{\boldsymbol{z}})\middle|(\tilde{z}_2, \ldots, \tilde{z}_n)\right] = \mathbf{E}\left[\frac{\partial\hat{h}(\tilde{\boldsymbol{z}})}{\partial z_j}\middle|(\tilde{z}_2, \ldots, \tilde{z}_n)\right], \ \forall j = 1, \ldots, n.$$

Taking the expectation of both sides, we find that

$$\mathbf{E}\left[\tilde{z}_j\hat{h}(\tilde{\boldsymbol{z}})\right] = \mathbf{E}\left[\frac{\partial\hat{h}(\tilde{\boldsymbol{z}})}{\partial z_j}\right], \ \forall j = 1, \ldots, n, \ \text{i.e.,}\ Cov\left(\tilde{\boldsymbol{z}}, \hat{h}(\tilde{\boldsymbol{z}})\right) = \mathbf{E}\left[\nabla\hat{h}(\tilde{\boldsymbol{z}})\right].$$

Note that the random vector $\tilde{\boldsymbol{r}}$ can be written as $\tilde{\boldsymbol{r}} = \Sigma^{1/2}\tilde{\boldsymbol{z}} + \boldsymbol{\mu}$. Consider $\hat{h}(\tilde{\boldsymbol{z}}) = h\left(\Sigma^{1/2}\tilde{\boldsymbol{z}} + \boldsymbol{\mu}\right)$, then $\nabla\hat{h}(\tilde{\boldsymbol{z}}) = \Sigma^{1/2}\nabla h(\tilde{\boldsymbol{r}})$. Hence,

$$Cov\left(\tilde{\boldsymbol{r}}, h(\tilde{\boldsymbol{r}})\right) = Cov\left(\Sigma^{1/2}\tilde{\boldsymbol{z}}, \hat{h}(\tilde{\boldsymbol{z}})\right) = \Sigma^{1/2}\mathbf{E}\left[\nabla\hat{h}(\tilde{\boldsymbol{z}})\right] = \Sigma\mathbf{E}\left[\nabla h(\tilde{\boldsymbol{r}})\right].$$

Appendix B. Proofs of Results in Section 2

Proof of Proposition 1

Expanding the expectation term in (T), we get the equivalent formulation as follows:

$$\min_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{x}^T\left(\Sigma+\boldsymbol{\mu}\boldsymbol{\mu}^T\right)\boldsymbol{x} - 2\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\tilde{\boldsymbol{r}}^T\right]\boldsymbol{x} + \mathbf{E}\left[(Z_B(\tilde{\boldsymbol{r}}))^2\right].$$

Since the last term is independent of $\boldsymbol{x}$, we can exclude it from the minimization problem. Note that

$$\mathbf{E}[Z_B(\tilde{\boldsymbol{r}})\tilde{\boldsymbol{r}}^T] = Cov(Z_B(\tilde{\boldsymbol{r}}),\tilde{\boldsymbol{r}}^T) + \mathbf{E}[Z_B(\tilde{\boldsymbol{r}})]\mathbf{E}[\tilde{\boldsymbol{r}}^T] = Cov(Z_B(\tilde{\boldsymbol{r}}),\tilde{\boldsymbol{r}}^T) + \mathbf{E}[Z_B(\tilde{\boldsymbol{r}})]\boldsymbol{\mu}^T.$$

Using differentiation by parts, we get

$$\begin{aligned}
\mathbf{E}\left[\frac{\partial Z_B(\tilde{\boldsymbol{r}})}{\partial \tilde{r}_l}\right] &= \mathbf{E}\left[\frac{\partial}{\partial \tilde{r}_l}\left(\sum_{i=1}^{n}\sum_{j=1}^{m}\tilde{r}_i P_{i,j} p_j(\tilde{\boldsymbol{r}})\right)\right] \\
&= \mathbf{E}\left[\sum_{j=1}^{m} P_{l,j} p_j(\tilde{\boldsymbol{r}}) + \sum_{i=1}^{n}\sum_{j=1}^{m}\tilde{r}_i P_{i,j}\frac{\partial p_j(\tilde{\boldsymbol{r}})}{\partial \tilde{r}_l}\right] \\
&= P_{(l)}\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right],
\end{aligned}$$

where $P_{(l)}$ denotes the lth row of P and $P_{i,j} = p_i^j$, $\forall i = 1,\ldots,n,\ j = 1,\ldots,m$. The last equality follows from our assumption on $\tilde{\boldsymbol{r}}$ so that $\partial p_j(\tilde{\boldsymbol{r}})/\partial \tilde{r}_l$ exists almost everywhere and equals zero wherever it exists. Applying Stein's Identity to $Cov(Z_B(\tilde{\boldsymbol{r}}),\tilde{\boldsymbol{r}}^T)$, we get

$$Cov(Z_B(\tilde{\boldsymbol{r}}),\tilde{\boldsymbol{r}}^T) = (\Sigma P\mathbf{E}[\boldsymbol{p}(\tilde{\boldsymbol{r}})])^T,$$

and thus Problem (T′).

Proof of Theorem 1

Since Problem (T′) is strictly convex, the first-order optimality conditions are both necessary and sufficient. In particular, the Lagrangian of Problem (T′) is given by

$$\mathcal{L}(\boldsymbol{x},\pi) = \boldsymbol{x}^T\left(\Sigma+\boldsymbol{\mu}\boldsymbol{\mu}^T\right)\boldsymbol{x} - 2\left(\Sigma P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] + \mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right]\boldsymbol{\mu}\right)^T\boldsymbol{x} + 2\pi\left(1-\boldsymbol{e}^T\boldsymbol{x}\right).$$

The first-order conditions yield

$$2\left(\Sigma+\boldsymbol{\mu}\boldsymbol{\mu}^T\right)_{(l)}\boldsymbol{x}^* - 2\left(\Sigma P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] + \mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right]\boldsymbol{\mu}\right)_{(l)} - 2\pi = 0,\ l = 1,...,n,$$

and

$$\sum_{i=1}^{n} x_i^* = 1.$$

Multiplying Σ^{-1} to both sides of the first set of equalities, we get

$$\begin{pmatrix} \Sigma^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^T + I_n & -\Sigma^{-1}\boldsymbol{e} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}^* \\ \pi \end{pmatrix} = \Sigma^{-1}\boldsymbol{\mu}\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right] + P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right], \tag{5}$$

where I_n denotes the identity matrix of dimension $n \times n$. Multiplying $\boldsymbol{\mu}^T$ to both sides, we have

$$\begin{pmatrix} \boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^T + \boldsymbol{\mu}^T I_n & -\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{e} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}^* \\ \pi \end{pmatrix} = \boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu}\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right] + \boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right].$$

Making the substitution with A, B, and C and dividing both sides by $(A+1)$, we get

$$\begin{pmatrix} \boldsymbol{\mu}^T & -\frac{B}{A+1} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}^* \\ \pi \end{pmatrix} = \frac{A}{A+1}\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right] + \frac{\boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]}{A+1}.$$

Subtracting $\mathbf{E}[Z_B(\tilde{\boldsymbol{r}})]$ from both sides gives

$$\begin{pmatrix} \boldsymbol{\mu}^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{x}^* \\ \pi \end{pmatrix} - \mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right] = \frac{\boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right] + \pi B}{A+1}. \tag{6}$$

Back to Equation (5), we can rewrite it as follows:

$$\Sigma^{-1}\boldsymbol{\mu}\left[\begin{pmatrix} \boldsymbol{\mu}^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{x}^* \\ \pi \end{pmatrix} - \mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right]\right] + \begin{pmatrix} I_n & -\Sigma^{-1}\boldsymbol{e} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}^* \\ \pi \end{pmatrix} = P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]. \tag{7}$$

Substituting Equation (6) into (7), we have

$$\begin{pmatrix} I_n & \frac{\Sigma^{-1}\boldsymbol{\mu}B}{A+1} - \Sigma^{-1}\boldsymbol{e} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}^* \\ \pi \end{pmatrix} = P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \frac{\boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right]}{A+1}\Sigma^{-1}\boldsymbol{\mu}. \tag{8}$$

Multiplying $\boldsymbol{e}^T$ to both sides of the above equation, we get

$$\boldsymbol{e}^T\boldsymbol{x}^* + \left(\frac{B^2}{A+1} - C\right)\pi = \boldsymbol{e}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \frac{B}{A+1}\left(\boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right]\right).$$

Note that $\boldsymbol{e}^T\boldsymbol{x}^* = \boldsymbol{e}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] = 1$. Canceling these two terms from both side, we have

$$\pi = \frac{B}{(A+1)C - B^2}\left(\boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right]\right).$$

Substituting the above formula for π into Equation (8), we get

$$\boldsymbol{x}^* = P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] + \left(\boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right]\right)\Sigma^{-1}\left(\frac{B\boldsymbol{e} - C\boldsymbol{\mu}}{(A+1)C - B^2}\right),$$

which is the closed-form solution as shown in the theorem.

Proof of Proposition 2

From the closed-form solution, the expected return of our multiple-benchmark tracking-error portfolio is

$$\boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] + \left(1 - \frac{C}{(A+1)C - B^2}\right)\left(\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right] - \boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]\right).$$

Note that

$$\begin{aligned}
& \mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right] - \boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] \\
= \; & \mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}}) - \tilde{\boldsymbol{r}}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]\right]^T \\
= \; & \mathbf{E}\left[\max_{j\in\{1,\ldots,m\}} \tilde{\boldsymbol{r}}^T\boldsymbol{p}^j - \sum_{j=1}^m \mathbf{E}\left[p_j(\tilde{\boldsymbol{r}})\right]\left(\tilde{\boldsymbol{r}}^T\boldsymbol{p}^j\right)\right]^T \\
\geq \; & 0.
\end{aligned}$$

If

$$1 \geq \frac{C}{(A+1)C - B^2}, \tag{9}$$

then

$$\boldsymbol{\mu}^T\boldsymbol{x} \geq \boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] = \mu_p.$$

Now we shall show (9) holds. Let $\boldsymbol{\alpha} = \Sigma^{-1/2}\boldsymbol{e}$, and $\boldsymbol{\beta} = \Sigma^{-1/2}\boldsymbol{\mu}$. By Cauchy-Schwartz Inequality,

$$B^2 = \left(\boldsymbol{e}^T\Sigma^{-1}\boldsymbol{\mu}\right)^2 = \left(\boldsymbol{\alpha}^{\boldsymbol{T}}\boldsymbol{\beta}\right)^2 \leq \left(\boldsymbol{\alpha}^T\boldsymbol{\alpha}\right)\left(\boldsymbol{\beta}^T\boldsymbol{\beta}\right) = \left(\boldsymbol{e}^T\Sigma^{-1}\boldsymbol{e}\right)\left(\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu}\right) = AC.$$

We have

$$\frac{1}{1 + \dfrac{AC - B^2}{C}} \leq 1,$$

i.e.,

$$\frac{C}{(A+1)C - B^2} \leq 1.$$

The equality holds if and only if $\boldsymbol{\alpha} = \boldsymbol{\beta}$ or $\boldsymbol{\mu} = \boldsymbol{e}$, since Σ is positive definite and so is $\Sigma^{-1/2}$.

Proof of Proposition 3

As shown in Equation (2), the closed-form solution to Problem (T″) is

$$\boldsymbol{x}_{MBTE} = P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] + \left(\boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - K\right)\Sigma^{-1}\left(\frac{B\boldsymbol{e} - C\boldsymbol{\mu}}{AC - B^2}\right). \tag{10}$$

The corresponding portfolio variance is

$$\begin{aligned}
\sigma^2_{MBTE} = \; & \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]^T P^T\Sigma P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \frac{C}{AC - B^2}\left(\boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]\right)^2 \\
& + \frac{2B}{AC - B^2}\boldsymbol{\mu}^T P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] + \frac{CK^2 - 2KB}{AC - B^2}.
\end{aligned}$$

At the return level of K, the optimal Markowitz mean-variance portfolio is given by

$$\boldsymbol{x}_{MEAV} = \frac{(A - BK)\,\Sigma^{-1}\boldsymbol{e} + (CK - B)\,\Sigma^{-1}\boldsymbol{\mu}}{AC - B^2}.$$

Its variance is

$$\sigma^2_{MEAV} = \frac{1}{C} + \frac{(KC - B)^2}{C\,(AC - B^2)}. \tag{11}$$

At the expected return level of K, the difference in portfolio variances of the two models is given by

$$\sigma^2_{MBTE} - \sigma^2_{MEAV} = \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]^T P^T \Sigma P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \frac{C}{AC - B^2}\left(\boldsymbol{\mu}^T P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \frac{B}{C}\right)^2 - \frac{1}{C}.$$

Note that the above difference is independent of K, which indicates that the multiple-benchmark tracking-error frontier is a constant shift from the mean-variance efficient frontier.

Proof of Proposition 4

Proof. In order to prove this result, suffice it to show that the gap between the two frontiers is zero.

Let $\boldsymbol{p}^1$ and $\boldsymbol{p}^2$ be two distinct portfolios on the mean-variance efficient frontier, and they serve as the benchmark portfolios for our multiple-benchmark tracking-error model. Their persistency values satisfy

$$\mathbf{E}\left[p_1(\tilde{\boldsymbol{r}})\right] + \mathbf{E}\left[p_2(\tilde{\boldsymbol{r}})\right] = 1.$$

From the Two-Fund Theorem, we know that the persistency weighted benchmark portfolio,

$$\mathbf{E}\left[p_1(\tilde{\boldsymbol{r}})\right]\boldsymbol{p}^1 + \mathbf{E}\left[p_2(\tilde{\boldsymbol{r}})\right]\boldsymbol{p}^2 = P\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right],$$

is also a mean-variance portfolio with expected return of $\boldsymbol{\mu}^T P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]$. From Equation (11), the variance of this portfolio is

$$\mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]^T P^T \Sigma P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] = \frac{1}{C} + \frac{\left(\boldsymbol{\mu}^T P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] C - B\right)^2}{C\,(AC - B^2)}.$$

By Proposition 3, the gap between the multiple-benchmark tracking-error frontier and the mean-variance efficient frontier is

$$\begin{aligned}
& \sigma^2_{MBTE} - \sigma^2_{MEAV} \\
= \; & \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]^T P^T \Sigma P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \frac{C}{AC - B^2}\left(\boldsymbol{\mu}^T P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \frac{B}{C}\right)^2 - \frac{1}{C} \\
= \; & \frac{1}{C} + \frac{\left(\boldsymbol{\mu}^T P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] C - B\right)^2}{C\,(AC - B^2)} - \frac{C}{AC - B^2}\left(\boldsymbol{\mu}^T P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \frac{B}{C}\right)^2 - \frac{1}{C} \\
= \; & 0.
\end{aligned}$$

Therefore, we have completed the proof.

Proof of Proposition 5

At the target expected return K, the linear combination coefficient is given by

$$\delta = \frac{K - \boldsymbol{\mu}^T \boldsymbol{p}^2}{\boldsymbol{\mu}^T \left(\boldsymbol{p}^1 - \boldsymbol{p}^2\right)}.$$

Then the linear combination portfolio has the closed-form expression as follows:

$$\boldsymbol{x}_{LCR} = \frac{K}{\boldsymbol{\mu}^T \left(\boldsymbol{p}^1 - \boldsymbol{p}^2\right)} \left(\boldsymbol{p}^1 - \boldsymbol{p}^2\right) + \frac{\left(\boldsymbol{\mu}^T \boldsymbol{p}^1\right) \boldsymbol{p}^2 - \left(\boldsymbol{\mu}^T \boldsymbol{p}^2\right) \boldsymbol{p}^1}{\boldsymbol{\mu}^T \left(\boldsymbol{p}^1 - \boldsymbol{p}^2\right)}.$$

To emphasize the portfolio variance's dependence on the target expected return, we denote the variance of the linear combination portfolio as $\sigma^2_{LCR}(K)$. Then

$$\begin{aligned} \sigma^2_{LCR}(K) = & \ \boldsymbol{x}^T_{LCR} \Sigma \boldsymbol{x}_{LCR} \\ = & \ a_{LCR} K^2 + b_{LCR} K + c_{LCR}, \end{aligned}$$

where

$$\begin{aligned} a_{LCR} = & \ \tfrac{1}{\left[\boldsymbol{\mu}^T \left(\boldsymbol{p}^1 - \boldsymbol{p}^2\right)\right]^2} \left(\boldsymbol{p}^1 - \boldsymbol{p}^2\right)^T \Sigma \left(\boldsymbol{p}^1 - \boldsymbol{p}^2\right), \\ b_{LCR} = & \ \tfrac{1}{\left[\boldsymbol{\mu}^T \left(\boldsymbol{p}^1 - \boldsymbol{p}^2\right)\right]^2} \left(\boldsymbol{p}^1 - \boldsymbol{p}^2\right)^T \Sigma \left(\left(\boldsymbol{\mu}^T \boldsymbol{p}^1\right) \boldsymbol{p}^2 - \left(\boldsymbol{\mu}^T \boldsymbol{p}^2\right) \boldsymbol{p}^1\right) \\ & + \tfrac{1}{\left[\boldsymbol{\mu}^T \left(\boldsymbol{p}^1 - \boldsymbol{p}^2\right)\right]^2} \left(\left(\boldsymbol{\mu}^T \boldsymbol{p}^1\right) \boldsymbol{p}^2 - \left(\boldsymbol{\mu}^T \boldsymbol{p}^2\right) \boldsymbol{p}^1\right)^T \Sigma \left(\boldsymbol{p}^1 - \boldsymbol{p}^2\right), \text{ and} \\ c_{LCR} = & \ \tfrac{1}{\left[\boldsymbol{\mu}^T \left(\boldsymbol{p}^1 - \boldsymbol{p}^2\right)\right]^2} \left(\left(\boldsymbol{\mu}^T \boldsymbol{p}^1\right) \boldsymbol{p}^2 - \left(\boldsymbol{\mu}^T \boldsymbol{p}^2\right) \boldsymbol{p}^1\right)^T \Sigma \left(\left(\boldsymbol{\mu}^T \boldsymbol{p}^1\right) \boldsymbol{p}^2 - \left(\boldsymbol{\mu}^T \boldsymbol{p}^2\right) \boldsymbol{p}^1\right). \end{aligned}$$

At the target expected return K, the multiple-benchmark tracking-error portfolio is given by Equation (2), and its variance is

$$\sigma^2_{MBTE}(K) = a_{MBTE} K^2 + b_{MBTE} K + c_{MBTE},$$

where

$$\begin{aligned} a_{MBTE} = & \ \frac{C}{AC - B^2}, \\ b_{MBTE} = & \ -\frac{2B}{AC - B^2}, \text{ and} \\ c_{MBTE} = & \ \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]^T P^T \Sigma P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right] - \frac{C}{AC - B^2} \left(\boldsymbol{\mu}^T P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]\right)^2 \\ & + \frac{2B}{AC - B^2} \boldsymbol{\mu}^T P \mathbf{E}\left[\boldsymbol{p}(\tilde{\boldsymbol{r}})\right]. \end{aligned}$$

From Equation (11), the variance of the Markowitz mean-variance portfolio at target expected return K is

$$\sigma^2_{MEAV}(K) = a_{MEAV} K^2 + b_{MEAV} K + c_{MEAV},$$

where

$$\begin{aligned} a_{MEAV} = & \ \frac{C}{AC - B^2}, \\ b_{MEAV} = & \ -\frac{2B}{AC - B^2}, \text{ and} \\ c_{MEAV} = & \ \frac{1}{C}. \end{aligned}$$

Observe that all three variances are quadratic functions of the target expected return , and

$$a_{MBTE} = a_{MEAV}, \ b_{MBTE} = b_{MEAV}.$$

Before proving the main result, we will first establish two claims.

Claim 1. *The multiple-benchmark tracking-error frontier intersects the linear combination rule frontier at the target expected return equal to the expected return of the persistency weighted benchmark portfolio, i.e., when* $K = \boldsymbol{\mu}^T \left(\mathbf{E}\left[p_1(\tilde{\boldsymbol{r}})\right] \boldsymbol{p}^1 + \mathbf{E}\left[p_2(\tilde{\boldsymbol{r}})\right] \boldsymbol{p}^2\right)$.

Proof. When $K = \boldsymbol{\mu}^T \left(\mathbf{E}\left[p_1(\tilde{\boldsymbol{r}})\right] \boldsymbol{p}^1 + \mathbf{E}\left[p_2(\tilde{\boldsymbol{r}})\right] \boldsymbol{p}^2\right)$, the linear combination portfolio is exactly the persistency weighted benchmark portfolio, i.e., $\mathbf{E}\left[p_1(\tilde{\boldsymbol{r}})\right] \boldsymbol{p}^1 + \mathbf{E}\left[p_2(\tilde{\boldsymbol{r}})\right] \boldsymbol{p}^2$. From Equation (2), the multiple-benchmark tracking-error portfolio is

$$\begin{aligned} \boldsymbol{x}_{MBTE} = &\ \mathbf{E}\left[p_1(\tilde{\boldsymbol{r}})\right] \boldsymbol{p}^1 + \mathbf{E}\left[p_2(\tilde{\boldsymbol{r}})\right] \boldsymbol{p}^2 \\ &+ \left(\boldsymbol{\mu}^T \left(\mathbf{E}\left[p_1(\tilde{\boldsymbol{r}})\right] \boldsymbol{p}^1 + \mathbf{E}\left[p_2(\tilde{\boldsymbol{r}})\right] \boldsymbol{p}^2\right) - K\right) \Sigma^{-1} \left(\frac{B\boldsymbol{e} - C\boldsymbol{\mu}}{AC - B^2}\right) \\ = &\ \mathbf{E}\left[p_1(\tilde{\boldsymbol{r}})\right] \boldsymbol{p}^1 + \mathbf{E}\left[p_2(\tilde{\boldsymbol{r}})\right] \boldsymbol{p}^2, \end{aligned}$$

which is the same as the linear combination portfolio. Thus, Claim 1 is proved. ■

Claim 2. *The quadratic coefficient in* $\sigma^2_{MBTE}(K)$ *is less than or equal to the quadratic coefficient in* $\sigma^2_{LCR}(K)$*, i.e.,* $a_{MEAV} \leq a_{LCR}$.

Proof. Consider the following optimization problem:

$$(\mathrm{M}_0) \qquad \min_{\boldsymbol{e}^T\boldsymbol{x}=0,\, \boldsymbol{\mu}^T\boldsymbol{x}=K'} \frac{1}{2}\boldsymbol{x}^T\Sigma\boldsymbol{x}.$$

The system of first-order optimality conditions reads

$$\begin{cases} \Sigma\boldsymbol{x}^* - \lambda_1\boldsymbol{e} - \lambda_2\boldsymbol{\mu} = 0, \\ \boldsymbol{e}^T\boldsymbol{x}^* = 0, \\ \boldsymbol{\mu}^T\boldsymbol{x}^* = K', \end{cases}$$

where λ_1 and λ_2 are Lagrange multipliers. From the first equation, we get

$$\boldsymbol{x}^* = \lambda_1\Sigma^{-1}\boldsymbol{e} + \lambda_2\Sigma^{-1}\boldsymbol{\mu}.$$

Substituting the above expression of $\boldsymbol{x}$ into the last two equations of the optimality conditions, we have

$$\begin{cases} \boldsymbol{e}^T\boldsymbol{x}^* = \lambda_1\boldsymbol{e}^T\Sigma^{-1}\boldsymbol{e} + \lambda_2\boldsymbol{e}^T\Sigma^{-1}\boldsymbol{\mu} = \lambda_1 C + \lambda_2 B = 0, \\ \boldsymbol{\mu}^T\boldsymbol{x}^* = \lambda_1\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{e} + \lambda_2\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu} = \lambda_1 B + \lambda_2 A = K', \end{cases}$$

which yields

$$\lambda_1 = -\frac{K'B}{AC - B^2}, \text{ and } \lambda_2 = \frac{K'C}{AC - B^2}.$$

Therefore, the optimal solution is

$$\boldsymbol{x}^* = \frac{K^{'}C}{AC-B^2}\Sigma^{-1}\boldsymbol{\mu} - \frac{K^{'}B}{AC-B^2}\Sigma^{-1}\boldsymbol{e},$$

and the minimum objective value scaled by 2 is

$$\begin{aligned}
\boldsymbol{x}^{*T}\Sigma\boldsymbol{x}^* = & \left(\tfrac{K^{'}C}{AC-B^2}\right)^2 \boldsymbol{\mu}^T\Sigma^{-1}\Sigma\Sigma^{-1}\boldsymbol{\mu} + \left(\tfrac{K^{'}B}{AC-B^2}\right)^2 \boldsymbol{e}^T\Sigma^{-1}\Sigma\Sigma^{-1}\boldsymbol{e} \\
& -\tfrac{K^{'2}CB}{(AC-B^2)^2}\boldsymbol{\mu}^T\Sigma^{-1}\Sigma\Sigma^{-1}\boldsymbol{e} - \tfrac{K^{'2}BC}{(AC-B^2)^2}\boldsymbol{e}^T\Sigma^{-1}\Sigma\Sigma^{-1}\boldsymbol{\mu} \\
= & \left(\tfrac{K^{'}C}{AC-B^2}\right)^2 A + \left(\tfrac{K^{'}B}{AC-B^2}\right)^2 C - \tfrac{2K^{'2}CB^2}{(AC-B^2)^2} \\
= & \tfrac{K^{'2}C^2A-K^{'2}CB^2}{(AC-B^2)^2} \\
= & \tfrac{C}{AC-B^2}K^{'2}.
\end{aligned}$$

Observe that $\left(\boldsymbol{p}^1-\boldsymbol{p}^2\right)$ is a feasible solution to Problem (M_0) with $K^{'} = \boldsymbol{\mu}^T\left(\boldsymbol{p}^1-\boldsymbol{p}^2\right)$, then it must satisfy

$$\begin{aligned}
\left(\boldsymbol{p}^1-\boldsymbol{p}^2\right)^T\Sigma\left(\boldsymbol{p}^1-\boldsymbol{p}^2\right) \geq & \ \boldsymbol{x}^{*T}\Sigma\boldsymbol{x}^* \\
= & \ \tfrac{C}{AC-B^2}K^{'2} \\
= & \ \tfrac{C}{AC-B^2}\left[\boldsymbol{\mu}^T\left(\boldsymbol{p}^1-\boldsymbol{p}^2\right)\right]^2.
\end{aligned}$$

Rearrange the terms, we get

$$\frac{1}{\left[\boldsymbol{\mu}^T\left(\boldsymbol{p}^1-\boldsymbol{p}^2\right)\right]^2}\left(\boldsymbol{p}^1-\boldsymbol{p}^2\right)^T\Sigma\left(\boldsymbol{p}^1-\boldsymbol{p}^2\right) \geq \frac{C}{AC-B^2},$$

which is exactly $a_{LCR} \geq a_{MEAV}$. Therefore, we have proved Claim 2. ■

Now in order to prove the proposition, we only need to discuss two cases following Claim 2: $a_{LCR} = a_{MEAV}$ and $a_{LCR} > a_{MEAV}$.

Case 1. $a_{LCR} = a_{MEAV}$.

By the definition of mean-variance efficient frontier, $\sigma^2_{MEAV}(K) \leq \sigma^2_{LCR}(K)$, for any K, i.e.,

$$b_{MEAV}K + c_{MEAV} \leq b_{LCR}K + c_{LCR},\ \forall K.$$

Then we must have $b_{MEAV} = b_{LCR}$. Otherwise, the above inequality will be violated as $K \to +\infty$ if $b_{MEAV} > b_{LCR}$, or $K \to -\infty$ if $b_{MEAV} < b_{LCR}$. Consequently,

$$a_{MBTE} = a_{MEAV} = a_{LCR}, \text{ and } b_{MBTE} = b_{MEAV} = b_{LCR}.$$

Furthermore, since the multiple-benchmark tracking-error frontier has an intersection point with the linear combination rule frontier, it must be the case that

$$c_{MBTE} = c_{LCR},$$

which implies that the multiple-benchmark tracking-error frontier coincides with the linear combination rule frontier. i.e.,

$$\sigma^2_{LCR}(K) = \sigma^2_{MBTE}(K), \forall K.$$

Case 2. $a_{LCR} > a_{MEAV}$.

Recall that $a_{MBTE} = a_{MEAV}$. Then $a_{LCR} > a_{MBTE}$. In this case, it is obvious that there exists a constant $\bar{K}$ such that

$$\sigma^2_{LCR}(K) > \sigma^2_{MBTE}(K), \forall K \geq \bar{K}.$$

Combining Case 1 and Case 2, we complete the proof.

Proof of Proposition 6

Since the convexity is preserved in Problem (P) by adding the quadratic penalty term, the first-order optimality conditions are both necessary and sufficient to characterize the solution. The Lagrangian of Problem (TC) is given by

$$\mathcal{L}(\boldsymbol{x}, \pi) = \mathbf{E}\left[(Z_B(\tilde{\boldsymbol{r}}) - \tilde{\boldsymbol{r}}^T\boldsymbol{x})^2\right] + \nu(\boldsymbol{x} - \boldsymbol{x}^0)^T(\boldsymbol{x} - \boldsymbol{x}^0) + 2\pi\left(1 - \boldsymbol{e}^T\boldsymbol{x}\right).$$

The first-order conditions yield

$$\frac{\partial \mathcal{L}(\boldsymbol{x}, \pi)}{\partial x_l} = 2\mathbf{E}\left[\left(Z_B(\tilde{\boldsymbol{r}}) - \tilde{\boldsymbol{r}}^T\boldsymbol{x}\right)(-\tilde{r}_l)\right] + 2\nu(x_l - x_l^0) - 2\pi = 0,\ l = 1, ..., n,\ \text{and}\ \sum_{i=1}^n x_i = 1.$$

The first set of equalities can be rewritten as

$$\begin{pmatrix} \Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T + \nu I_n & -\boldsymbol{e} \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \pi \end{pmatrix} = \mathbf{E}[Z_B(\tilde{\boldsymbol{r}})\tilde{\boldsymbol{r}}] + \nu\boldsymbol{x}^0.$$

Applying Lemma 1, we have

$$\begin{pmatrix} \Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T + \nu I_n & -\boldsymbol{e} \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \pi \end{pmatrix} = \boldsymbol{\mu}\mathbf{E}[Z_B(\tilde{\boldsymbol{r}})] + \Sigma P\mathbf{E}[\boldsymbol{c}(\tilde{\boldsymbol{r}})] + \nu\boldsymbol{x}^0.$$

Since $\nu \geq 0$ and $(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T)$ is positive definite, $(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T + \nu I_n)$ is also positive definite. In particular, it has an inverse. The we can multiple both sides of above equation by $(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T + \nu I_n)^{-1}$ and obtain

$$\begin{aligned} &\begin{pmatrix} I_n & -\left(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T + \nu I_n\right)^{-1}\boldsymbol{e} \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \pi \end{pmatrix} \\ &= \left(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T + \nu I_n\right)^{-1}\left(\boldsymbol{\mu}\mathbf{E}[Z_B(\tilde{\boldsymbol{r}})] + \Sigma P\mathbf{E}[\boldsymbol{c}(\tilde{\boldsymbol{r}})] + \nu\boldsymbol{x}^0\right). \end{aligned} \tag{12}$$

Multiplying $\boldsymbol{e}^T$ to the above equality and simplifying using the fact that $\boldsymbol{e}^T\boldsymbol{x} = 1$, we get

$$\pi = \frac{1 - \boldsymbol{e}^T\left(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T + \nu I_n\right)^{-1}\left(\boldsymbol{\mu}\mathbf{E}[Z_B(\tilde{\boldsymbol{r}})] + \Sigma P\mathbf{E}[\boldsymbol{c}(\tilde{\boldsymbol{r}})] + \nu\boldsymbol{x}^0\right)}{\boldsymbol{e}^T\left(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T + \nu I_n\right)^{-1}\boldsymbol{e}}.$$

Substituting this expression of π into Equation (12), simplifying the resulting expression with a substitution of $D = \left(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T + \nu I_n\right)^{-1}$, we have

$$\boldsymbol{x} = \frac{D\boldsymbol{e}}{\boldsymbol{e}^T D\boldsymbol{e}} + \left(I_n - \frac{DJ_n}{\boldsymbol{e}^T D\boldsymbol{e}}\right) D \left(\boldsymbol{\mu}\mathbf{E}\left[Z_B(\tilde{\boldsymbol{r}})\right] + \Sigma P\mathbf{E}\left[\boldsymbol{c}(\tilde{\boldsymbol{r}})\right] + \nu\boldsymbol{x}^0\right),$$

where J_n denotes the matrix in $\mathbb{R}^{n\times n}$ with all entries being 1. Thus, we obtain the closed-form solution to Problem (TC).