A Nonparametric Approach to Portfolio Shrinkage

Chulwoo Han*

Abstract

This paper develops a shrinkage model for portfolio choice. It places a layer on a conventional portfolio problem where the optimal portfolio is shrunk towards a reference portfolio. The model can accommodate a wide range of portfolio problems with various objectives and constraints, and its implementation is simple and straightforward. A data-driven method to determine the shrinkage level is offered. A comprehensive comparative study suggests the proposed model substantially enhances the performance of its underlying model and outperforms existing shrinkage models as well as the naïve strategy. The naïve strategy serves better as the reference portfolio than the current portfolio.

JEL Classification: G11

Keywords: Turnover minimization; Shrinkage estimator; Parameter uncertainty; Portfolio optimization

*Chulwoo Han is with Durham University. Durham Business School, Mill Hill Lane, Durham, DH1 3LB UK; Tel: +44 1913345892; E-mail: chulwoo.han@durham.ac.uk.

1. Introduction

There has been a long debate on the effectiveness of optimal portfolios and their competitive advantages against the naïve, equal-weight portfolio (*a.k.a.* 1/N rule). In their seminal work, DeMiguel et al. (2009) test fourteen portfolio strategies on seven datasets and find none of them consistently outperforms the naïve strategy. They further show that when returns are *i.i.d.* normal, an unrealistically long sample period is required for the Markowitz (1952) mean-variance portfolio to outperform the equal-weight portfolio. Although their results are somewhat exaggerated (*e.g.*, see discussions in Kirby and Ostdiek (2012) and Kan et al. (2016)), the sheer number of citations of their paper reflects the impact it has brought to academia and industry.[1] Undoubtedly, there has been a backlash. Kirby and Ostdiek (2012), using similar sets of data, find that a mean-variance strategy constrained to invest only in risky assets outperforms the naïve strategy. Bessler et al. (2014) show that a strategy based on the Black and Litterman (1992) framework outperforms the naïve strategy when applied to a multi-asset dataset. Branger et al. (2019) develop a grouping strategy in which portfolio optimization is performed on groups of equally weighted stocks, and show that their strategy outperforms many existing strategies that aim to address estimation risks. Han (2019) develops a shrinkage model that improves upon the shrinkage models of Kan and Zhou (2007) and Tu and Zhou (2011) and finds that the proposed model outperforms the existing shrinkage models as well as the naïve strategy.

The race between the naïve and optimal strategies essentially depends upon the predictability of input parameters, *i.e.*, expected returns and covariance matrix. If both input parameters are unknown, it would be reasonable to assume that all the assets have the same expected return and variance. Alternatively, based on asset pricing models, the same return-risk ratio could be assumed for all assets. Both assumptions lead to the equal-weight portfolio as the optimal portfolio.[2] If only the variances are known, the covariances of all

[1]Based on the Google Scholar search, their paper has been cited 2,383 times at the time of writing (December 2019).

[2]Pflug et al. (2012) also show that the naïve portfolio is optimal when estimation errors are significant.

asset pairs may be assumed to be equal, and so are the expected returns. This would lead to the volatility timing strategy of Kirby and Ostdiek (2012). If the covariance matrix is known, the expected returns could be assumed the same across assets, in which case, the minimum-variance portfolio would be optimal. If all the input parameters are known, the Markowitz (1952) mean-variance portfolio should be the choice. In reality, input parameters will be estimated with errors, and a portfolio strategy that takes estimation errors into account, *e.g.*, a Bayesian method or robust optimization, would be preferred.

From this perspective, the remarkable performance of the naïve strategy merely reaffirms the difficulty of reliable input parameter estimation. Even when the input parameters can be predicted within a certain accuracy, the classical mean-variance strategy could result in a ruinous allocation due to its high parameter sensitivity and error-maximizing property (Michaud, 1989), and it is crucial to address estimation errors for successful utilization of portfolio optimization.

There has been a considerable amount of effort dedicated to addressing input parameter uncertainty and portfolio sensitivity. One pillar has been formed by the Bayesian approach: *e.g.*, Klein and Bawa (1976), Brown (1976, 1978), Jorion (1986), Black and Litterman (1992), Pástor (2000), Pástor and Stambaugh (2000). For a review of Bayesian models, the reader is referred to Avramov and Zhou (2010). More recently, the robust optimization that finds an optimal portfolio under a worst-case scenario became popular: *e.g.*, Goldfarb and Iyengar (2003), Fabozzi et al. (2007), Ceria and Stubbs (2016). The shrinkage estimator, first proposed by Kan and Zhou (2007), optimally combines two or more portfolios so that the expected utility loss is minimized. This approach has been adopted later by Tu and Zhou (2011), DeMiguel et al. (2015), Kan et al. (2016), and Han (2019), among others. Other approaches include imposing weight constraints (Jagannathan and Ma, 2003) or using a shrinkage method for parameter estimation (Ledoit and Wolf, 2004). Brandt et al. (2009) skip input parameter estimation completely by specifying portfolio weights as a function of firm characteristics.

Although these models have shown some degree of success, they are also subject to limitations. Above all, many models assume the knowledge of estimation error distribution, altough its estimation can be as challenging as input parameter estimation. The distribution of estimation errors is a crucial determinant of asset allocation in these models, and its misspecification can result in poor portfolio performance.

Bayesian approaches typically assume that the covariance matrix is precisely known and focus on the estimation error of the expected returns. While the covariance matrix can be estimated more accurately, Kan and Zhou (2007) show that its estimation error can have a nontrivial impact on asset allocation and portfolio performance when combined with the estimation error of the expected returns. Furthermore, Bayesian updates are carried out at the input parameter level, which is not necessarily optimal from a portfolio perspective.

In contrast, shrinkage estimators recognize the uncertainty of both input parameters and find an optimal combination of multiple portfolios from a portfolio perspective by minimizing the expected utility loss (or equivalently, maximizing the expected out-of-sample utility). Nevertheless, shrinkage estimators suffer from shrinkage parameter uncertainty: the optimal shrinkage parameters (the coefficients on the portfolios) are functions of unknown input parameters and therefore inherit input parameter uncertainty (Han, 2019). The uncertainty of the shrinkage parameters can lead to a nontrivial utility loss. Shrinkage estimators also lack practicality. As they maximize the expected out-of-sample utility, the risk aversion parameter needs to be specified, which is not always straightforward, especially for institutional investors. It is also difficult to incorporate constraints such as the short-sale constraint into these models.

This paper proposes a new shrinkage method for portfolio choice, **turnover minimization**. The main idea of the turnover minimization is to minimize the distance between an optimal portfolio and a reference portfolio subject to return or risk constraints.[3] In contrast to maximizing utility, this mitigates the error maximizing property of the classical

[3] Minimizing the distance from a reference portfolio usually leads to a lower turnover even when the reference portfolio is not the currently held portfolio, hence the name – turnover minimization.

mean-variance portfolio. This approach is also consistent with the decision-making process of institutional investors: they often prefer to have a stable portfolio that meets their return/risk targets rather than a portfolio that maximizes return or minimizes risk. As detailed in the next section, the turnover minimization has several advantages compared to existing models that account for estimation errors: it does not require an explicit assumption of error distribution and can be easily incorporated into conventional portfolio problems with any type of constraints.

The turnover minimization is motivated by the observation that, while a classical optimal portfolio tends to contain extreme weights, there often exists a near-optimal portfolio with considerably more balanced weights. Consider a two-asset allocation problem with the expected returns and the covariance matrix:

$$\mu = \begin{bmatrix} 0.10 \\ 0.15 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 0.16 & 0.15 \\ 0.15 & 0.25 \end{bmatrix}.$$

If we maximize the expected return while constraining the variance under 0.2, the optimal portfolio becomes $w^* = [0.30\ 0.70]'$ with the expected return of 0.135. Now if we only require 95% of the expected return of the optimal portfolio, *i.e.*, $\mu_p = 0.95 \cdot 0.135 = 0.128$, and make the portfolio as close to the equal-weight portfolio as possible, we obtain $w^* = [0.43\ 0.57]'$. That is, more balanced asset allocation can be achieved with a little sacrifice of optimality. As it turns out, this property holds in a wide range of portfolio optimization problems.

Several versions of the turnover minimization are evaluated through a comprehensive comparative study that involves various portfolio models. In particular, models that incorporate the equal-weight portfolio are chosen as a benchmark as well as classical portfolio rules. With the desirable characteristics of the equal-weight portfolio such as low turnover and no short positions, and its recent role as a benchmark in portfolio studies, it was no surprise to witness the emergence of portfolio models incorporating it. Tu and Zhou (2011) combine the equal-weight portfolio with an optimal portfolio so that the expected utility loss is minimized: Markowitz (1952) rule, Jorion (1986) rule, Kan and Zhou (2007) rule, and MacKinlay and Pástor (2000) rule are considered for the optimal portfolio. Bessler et al. (2014) derive the equilibrium return from the equal-weight portfolio in the Black-Litterman framework. Branger et al. (2019) develop a strategy in which assets are grouped to form equal-weight portfolios before fed into optimization. Besides these models, a new model based on the work of Treynor and Black (1973) is also considered, in which the equal-weight portfolio is used as a proxy for the market portfolio.

The empirical study suggests that minimum-turnover portfolios outperform their underlying portfolios and other benchmarks both before and after transaction costs. In particular, the mean-variance and global minimum-variance portfolios with the short-sale constraint perform superior when augmented with the turnover minimization. These results are robust to datasets, test periods, and other variations.

The rest of the paper is organized as follows. Section 2 develops the turnover minimization framework, where a method to calibrate the model is also proposed. Section 3 carries out an empirical analysis, and Section 4 concludes. The implementation details of the test models can be found in the appendix, and additional empirical results are provided in the accompanying internet appendix (IA).

2. Turnover Minimization

The turnover minimization aims to minimize the distance from a reference portfolio, subject to return/risk constraints. Roughly, the problem can be written in the form

$$\begin{aligned} &\min_{w}\ (w-w_0)'(w-w_0) \\ &\text{subject to } \textit{return/risk constraints} \\ &\qquad\qquad\quad \textit{and other constraints}, \end{aligned}$$

where w_0 is the reference portfolio, which can be any known portfolio at the time of rebalancing. This paper mainly considers the equal-weight portfolio, w_{ew}, and the current portfolio, w_{t-}, for the reference portfolio, but also discusses other possibilities. As illustrated later in this section, the return and risk constraints are not only given exogenously but also determined endogenously to maximize portfolio performance.

The rationale behind the turnover minimization is at least twofold. First, minimizing turnover mitigates the error maximizing property of the classical portfolio optimization and yields a more robust portfolio. Second, investors are not necessarily return/risk optimizers: they often prefer a more robust portfolio as long as it meets their return/risk targets.

The turnover minimization is formulated as a two-stage optimization problem: classical portfolio optimization and turnover minimization. Consider the following return maximization problem of N assets subject to a variance constraint:

$$\begin{aligned} &\max_{w} w'\mu \\ &\text{subject to } w'\Sigma w \leq \sigma_T^2 \\ &w \in \mathcal{D}, \end{aligned} \tag{1}$$

where $\mu \in \mathbb{R}^N$ and $\Sigma \in \mathbb{R}^{N \times N}$ are the mean and covariance matrix of N asset returns in excess of the risk-free rate, $w \in \mathbb{R}^N$ is the portfolio weights, σ_T^2 is a risk tolerance (target variance), and $\mathcal{D}$ denotes the feasible set of w defined by other constraints such as the budget constraint or short-sale constraint. Denoting the optimal portfolio that solves (1) by w^*, the expected return of w^* is given by $\mu_p^* = w^{*\prime}\mu$. The minimum-turnover portfolio, w_{tm}, is then obtained by solving the second stage problem:

$$\begin{aligned} w_{tm} = \underset{w}{\operatorname{argmin}}\ & (w - w_0)'(w - w_0) \\ \text{subject to } & w'\Sigma w \leq \sigma_T^2 \\ & w \in \mathcal{D} \\ & w'\mu \geq (1-\tau)\mu_p^*, \end{aligned} \tag{2}$$

where $\tau > 0$ denotes the proportion of the optimal value the investor is willing to sacrifice in exchange for the gain in robustness.[4] When $\tau = 0$, the minimum-turnover portfolio is the same as the optimal portfolio from the first stage, whereas when $\tau \to \infty$, it becomes the reference portfolio unless other constraints are binding.

The turnover minimization is intuitive in that it first finds the optimal portfolio for the underlying portfolio problem and then moves it towards the reference portfolio by tolerating sub-optimality, while satisfying all the constraints imposed in the first stage. The turnover minimization can be easily incorporated into any portfolio optimization problems such as variance minimization or Sharpe ratio maximization. Below are some examples.

Variance Minimization-Turnover Minimization

$$\begin{aligned} {\sigma_P^2}^* = \min_w\ & w'\Sigma w \\ \text{subject to } & w \in \mathcal{D} \end{aligned} \quad \Rightarrow \quad \begin{aligned} w_{tm} = \underset{w}{\operatorname{argmin}}\ & (w - w_0)'(w - w_0) \\ \text{subject to } & w \in \mathcal{D} \\ & w'\Sigma w \leq (1+\tau)^2 {\sigma_P^2}^* \end{aligned} \tag{3}$$

[4] This two-step approach is equivalent to adding a penalty term, $\lambda(w - w_0)'(w - w_0)$ for some constant λ, to the objective function in (1). It is a special case of the model of Olivares-Nadal and DeMiguel (2018) when w_0 is the current portfolio. It is also equivalent to DeMiguel et al. (2009)'s norm-constrained model if w_0 is the risk-free asset. However, in this formulation, λ can have any nonnegative value making its calibration challenging. In contrast, τ will typically be chosen from $[0, 1]$, allowing more efficient calibration.

Sharpe Ratio Maximization-Turnover Minimization

$$
\begin{aligned}
SR^* = \max_w \frac{w'\mu}{\sqrt{w'\Sigma w}} \qquad & \qquad w_{tm} = \underset{w}{\operatorname{argmin}} \ (w - w_0)'(w - w_0) \\
\text{subject to } w \in \mathcal{D} \qquad \Rightarrow & \qquad \text{subject to } w \in \mathcal{D} \\
& \qquad \frac{w'\mu}{\sqrt{w'\Sigma w}} \geq (1-\tau)SR^*
\end{aligned}
\tag{4}
$$

Utility Maximization-Turnover Minimization

$$
\begin{aligned}
U^* = \max_w w'\mu - \frac{\gamma}{2}w'\Sigma w \qquad & \qquad w_{tm} = \underset{w}{\operatorname{argmin}} \ (w - w_0)'(w - w_0) \\
\text{subject to } w \in \mathcal{D} \qquad \Rightarrow & \qquad \text{subject to } w \in \mathcal{D} \\
& \qquad w'\mu - \frac{\gamma}{2}w'\Sigma w \geq (1-\tau)U^*
\end{aligned}
\tag{5}
$$

γ : risk aversion coefficient

If the first-stage problem can be formulated as a convex programming problem, the second-stage problem also becomes a convex programming problem and can be solved efficiently using a specialized software package, such as CVX, Gurobi, or MOSEK.

2.1. A Closer Look at the Turnover Minimization

The turnover minimization can be viewed as a shrinkage estimator as it shrinks the optimal portfolio towards the reference portfolio. While it is generally impossible to obtain an analytic solution for a turnover minimization problem, the following special case provides insights into the model and its connection to existing models. Consider the utility maximization-turnover minimization problem, now with a new distance function, $(w - w_0)'\Sigma(w - w_0)$.[5]

[5] This is for analytical tractability. With Σ, an asset with a larger variance will be penalized more severely for the deviation from w_0, whereas all assets are penalized equally in the original specification. Shrinking volatile assets more severely makes sense because they are likely to have larger estimation errors. Also, when the reference portfolio is the current portfolio, the quadratic term can be interpreted as the (scaled) transaction cost: see Gârleanu and Pedersen (2013), Olivares-Nadal and DeMiguel (2018).

The Lagrangian of the problem has the form

$$\mathcal{L} = \frac{1}{2}(w - w_0)'\Sigma(w - w_0) - \lambda\left(w'\mu - \frac{\gamma}{2}w'\Sigma w - (1-\tau)U^*\right). \tag{6}$$

From the first-order condition, $\frac{\partial \mathcal{L}}{\partial w} = 0$, the minimum-turnover portfolio is given by (see Appendix A for proof)

$$w_{tm} = \frac{1}{1+\lambda\gamma}w_0 + \frac{\lambda\gamma}{1+\lambda\gamma}w_{mk}, \tag{7}$$

where $w_{mk} = \frac{1}{\gamma}\Sigma^{-1}\mu$ is the optimal portfolio from the first stage, *i.e.*, the utility-maximizing portfolio. The minimum-turnover portfolio is a linear combination of the optimal portfolio and the reference portfolio. A portfolio of this form is equivalent to the shrinkage estimator of Tu and Zhou (2011) when $w_0 := w_{ew}$, and the shrinkage estimator of Kan and Zhou (2007) when w_0 is the global minimum variance portfolio. In this regard, the turnover minimization can be considered a generalized shrinkage estimator that encompasses existing models. The main difference, however, is that the turnover minimization does not make any particular assumption for the estimation errors and can easily accommodate various types of objective functions and constraints. This flexibility comes at a cost of analytical tractability, and τ needs to be calibrated from data.[6]

When the constraint is binding, *i.e.*, $w'\mu - \frac{\gamma}{2}w'\Sigma w = (1-\tau)U^*$, it can be shown that

$$\lambda = \frac{1}{\gamma}\left(\sqrt{\frac{U^* - U_0}{\tau U^*}} - 1\right), \tag{8}$$

and

$$w_{tm} = \sqrt{\frac{\tau U^*}{U^* - U_0}}w_0 + \left(1 - \sqrt{\frac{\tau U^*}{U^* - U_0}}\right)w_{mk}, \tag{9}$$

[6] While a closed form is always preferred, Han (2019) shows that, due to model parameter uncertainty, the closed-form solutions offered by Kan and Zhou (2007) and Tu and Zhou (2011) are sub-optimal even when all the assumptions are correct. He argues that the optimal shrinkage level should be higher than that suggested by these models, and cannot be determined analytically.

where $U_0 = w_0'\mu - \frac{\gamma}{2}w_0'\Sigma w_0$ is the utility of the reference portfolio. Note that the loading on w_0 is 0 when $\tau = 0$ and increases with τ, and w_{tm} converges to w_0 when $\tau = 1 - U_0/U^*$. Since $U_0 \le U^*$, $\tau \le 1$. Equation (9) can be rearranged as follows:

$$(w_{tm} - w_0) = \left(1 - \sqrt{\frac{\tau}{1-k}}\right)(w_{mk} - w_0), \tag{10}$$

where $k := U_0/U^*$. The distance between w_{tm} and w_0, normalized by the distance between w_{mk} and w_0, is given by

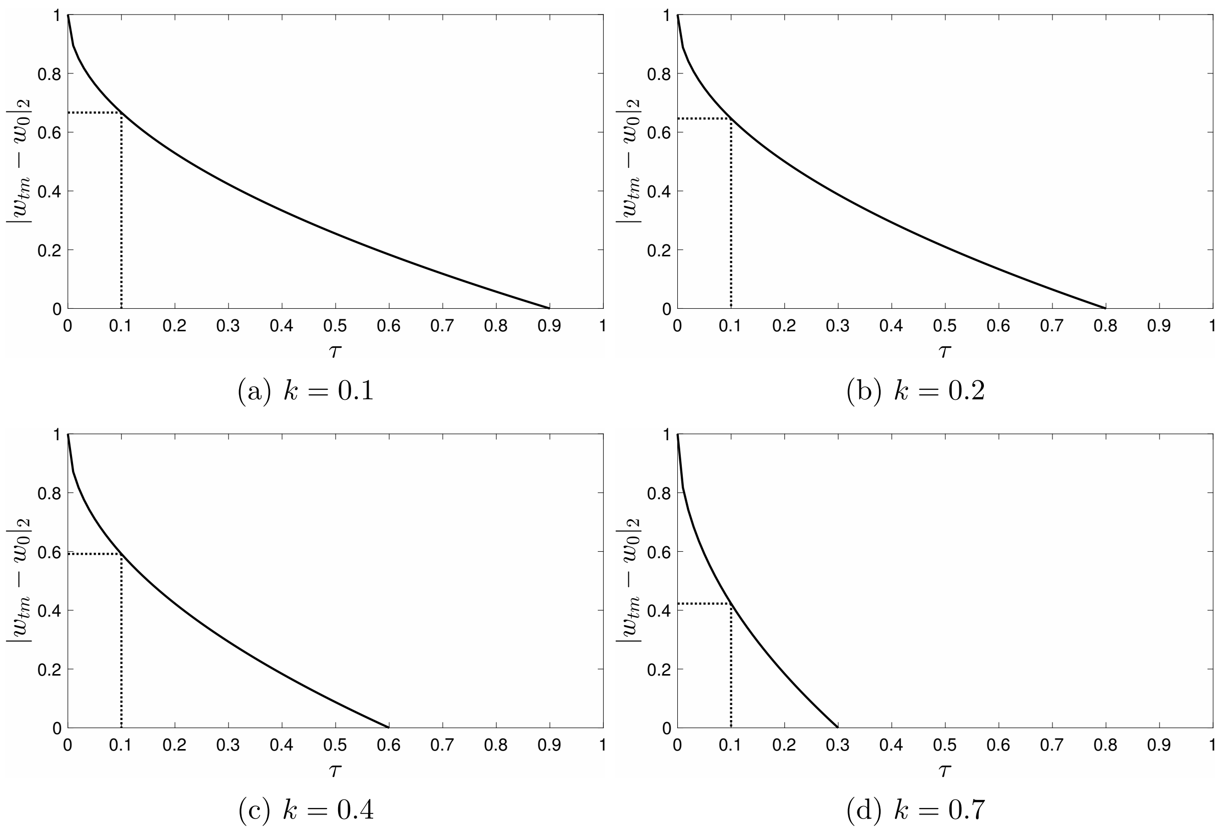
$$\frac{|w_{tm} - w_0|_2}{|w_{mk} - w_0|_2} = \frac{\lambda\gamma}{1+\lambda\gamma} = 1 - \sqrt{\frac{\tau}{1-k}}, \tag{11}$$

where $|x|_2$ is the l^2-norm. The equation shows that, for a given τ, the normalized distance depends only on the relative utility U_0/U^*, and is independent of input parameters. Moreover, when U_0 is closer to U^*, the optimal portfolio shrinks towards w_0 more rapidly.

Figure 1 shows the relationship between the distance to w_0 and the tolerance τ for different values of k. The vertical axis is the normalized distance, $1 - \sqrt{\tau/(1-k)}$. The figure suggests that, even when the utility of the reference portfolio is considerably lower than that of the Markowitz portfolio, a robust portfolio (*i.e.*, a portfolio close to w_0) can be obtained without any significant loss of utility. For instance, when $U_0 = 0.1U^*$ ($k = 0.1$), 10% tolerance results in 33% reduction of the distance, whereas the reduction increases to 41% when $U_0 = 0.4U^*$ ($k = 0.4$). Note that U^* is a hypothetical maximum utility that can be achieved in the absence of estimation errors. The actual utility of the Markowitz portfolio will be significantly lower, and so is the utility loss caused by turnover minimization.

2.2. Optimal τ

The tolerance level τ determines the degree of shrinkage and the choice of τ is critical for the performance of the minimum-turnover portfolio. While a closed-form formula for the optimal τ does not exist in general, it can be obtained in some special cases. In particular, if



(a) $k = 0.1$

(b) $k = 0.2$

(c) $k = 0.4$

(d) $k = 0.7$

Figure 1: Shrinkage by Turnover Minimization

This figure demonstrates the distance between the minimum-turnover portfolio, w_{tm}, and the reference portfolio, w_0, as a function of τ. The curves are obtained from Equation (11). The distance is normalized by $|w_{mk} - w_0|_2$.

the returns are *i.i.d.* normal, an optimal τ can be obtained in a closed-from for the problem stated above.

Under the normality assumption, the maximum likelihood (ML) estimator is efficient, and Equation (7) can be implemented using the maximum likelihood estimates of μ and Σ, $\hat{\mu}$ and $\hat{\Sigma}$:

$$w_{tm} = aw_0 + (1-a)\hat{w}_{mk}, \tag{12}$$

where $a := \frac{1}{1+\lambda\gamma}$ and $\hat{w}_{mk} := \frac{1}{\gamma}\hat{\Sigma}^{-1}\hat{\mu}$. The optimal a can be determined so that the expected out-of-sample utility of w_{tm} is maximized. The optimal a is given by (see Appendix A for proof)

$$a^* = \frac{\pi_2}{\pi_1 + \pi_2}, \tag{13}$$

$$\pi_1 = 2(U^* - U_0), \tag{14}$$

$$\pi_2 = 2(c_1 - 1)U^* + \frac{c_1}{\gamma}\frac{N}{T}, \tag{15}$$

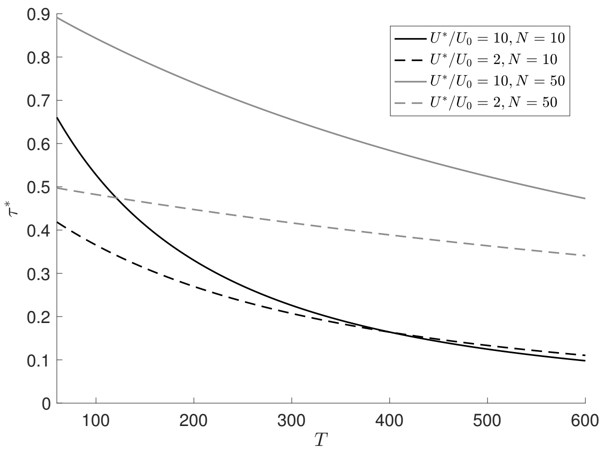
$$c_1 = \frac{(T-2)(T-N-2)}{(T-N-1)(T-N-4)}. \tag{16}$$

From (8), the optimal τ has the form

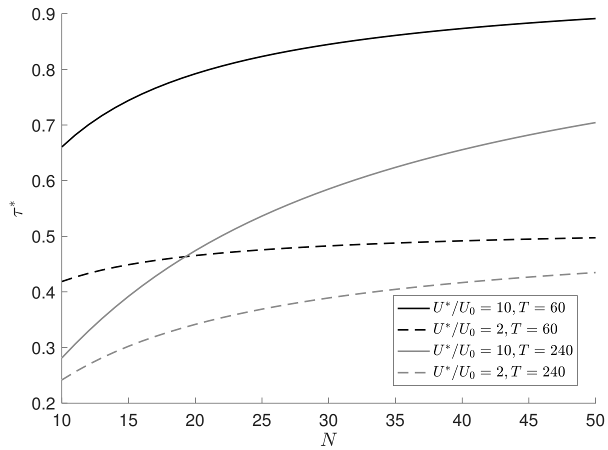
$$\tau^* = a^{*2}\frac{(U^* - U_0)}{U^*}. \tag{17}$$

The above result reveals that the optimal τ depends only on four factors: the utilities U^* and U_0, the number of assets N, and the size of the estimation window T. When the number of assets increases or the estimation window size decreases, both a^* and τ^* increases. This result is expected since a greater estimation error (larger N or smaller T) would require higher tolerance for the portfolio to become robust. The optimal τ is inversely related to U_0, *i.e.*, the closer U_0 to U^*, the lower the required level of tolerance. In contrast, τ^* is not a monotonic function of U^*: τ^* is zero when $U^* = U_0$, increases with U^* to a certain point and converges to $(1 - 1/c_1)$ as $U^* \to \infty$. When the gap between U^* and U_0 starts to increase, it is optimal to shrink more towards the reference portfolio, but when U^* is significantly higher than U_0, the loss of utility dominates, and the optimal shrinkage level decreases.

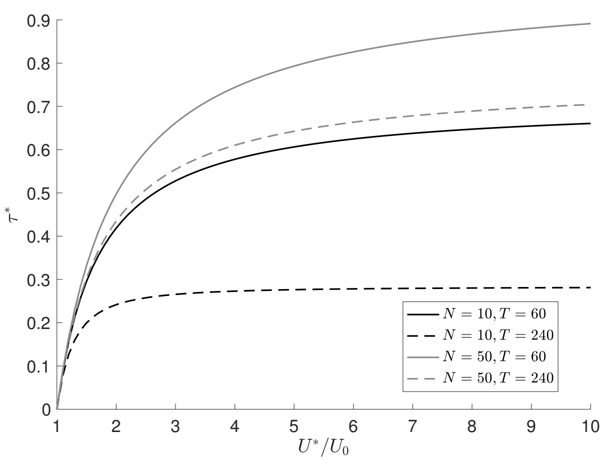
Figure 2 illustrates the relationship between the optimal τ and the four variables. As mentioned earlier, the optimal τ increases as T decreases, N increases, or U_0 decreases, and it increases and then decreases as U^* increases.



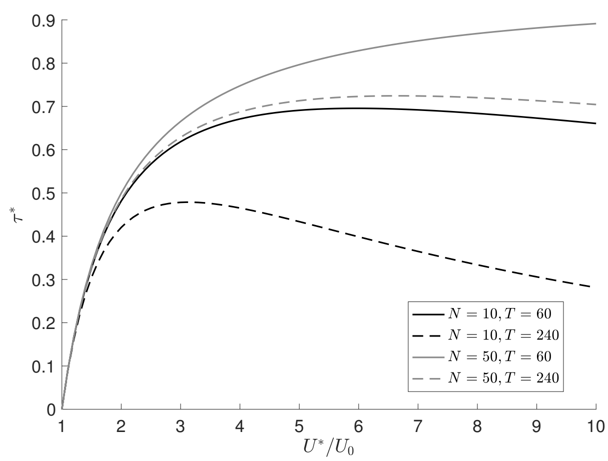
(a) τ^* vs. T, $U^* = 0.02$



(b) τ^* vs. N, $U^* = 0.02$



(c) τ^* vs. U_0, $U^* = 0.02$



(d) τ^* vs. U^*, $U_0 = 0.002$

Figure 2: Optimal tolerance level, τ^*

This figure illustrates the relationship between optimal τ and (a) the estimation window size T, (b) the number of assets N, (c) the utility of the reference portfolio U_0, and (d) the utility of the Markowitz portfolio U^*. All graphs are obtained from Equation (17).

2.3. Calibration of τ

For a general turnover minimization problem, τ needs to be determined numerically. The following data-driven calibration method is proposed.

1. For the first ten months of the evaluation period, τ is set to 0.05.

2. When the month $t > 10$, τ is calibrated every month so that the Sharpe ratio during $1, \ldots, t-1$ is maximized.[7] The optimal τ is found via line search spanning the range $[0, 1]$ in its log space.

3. In the presence of transaction costs, the same procedure is repeated using the Sharpe ratio net of transaction costs.

The first step is required only for the empirical study, as there is no data for calibration at the beginning of the evaluation period. In practice, one may set a period in the past so that τ can be calibrated from the first month of portfolio construction, and recalibrate τ either by rolling the estimation window or by accumulating the sample as the portfolio evolves. Another calibration method based on the last period return is also examined later in Section 3.3.

2.4. Simulation Studies

2.4.1. Distance to the Reference Portfolio

As illustrated in Section 2.1, shrinking towards a reference portfolio does not necessarily involve a considerable loss of optimality. This property is further investigated via simulation. Using the sample moments of the four datasets D1, D2, D5, and D8 in Table 2 as the true μ and Σ, the utility maximization-turnover minimization problem in (5) is solved for different values of τ. The equal-weight portfolio is chosen for the reference portfolio. Figure 3

[7] The Sharpe ratio is used as the calibration criterion regardless of the objective function of the underlying problem as the performance of a portfolio is evaluated by the Sharpe ratio.

displays the relationship between τ and the distance between the reference portfolio and the minimum-turnover portfolio, normalized by the distance between the Markowitz portfolio and the reference portfolio.

Remarkably, the distance to the equal-weight portfolio is reduced by more than half only for 10% utility loss ($\tau = 0.1$), even when the equal-weight portfolio's utility is substantially lower than that of the optimal portfolio.[8] This result confirms that the turnover minimization can yield a substantially more robust portfolio at the cost of a small fraction of optimality. Again, it is worth emphasizing that the utility loss is being measured against the hypothetical maximum utility, and the actual utility of w_{mk} is likely to be considerably lower due to estimation errors.

2.4.2. Expected Out-of-sample Utility

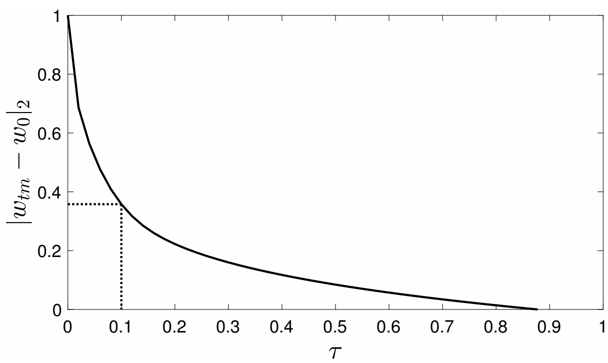
This section examines the expected out-of-sample utility of the minimum-turnover portfolio. Using the sample moments from the datasets D1, D2, D5, and D8 as true moments, the maximum likelihood estimates of the input parameters are randomly generated from

$$\hat{\mu} \sim \mathcal{N}\left(\mu, \frac{\Sigma}{T}\right), \quad \hat{\Sigma} \sim \mathcal{W}_N(T-1, \Sigma)\frac{1}{T}, \tag{18}$$

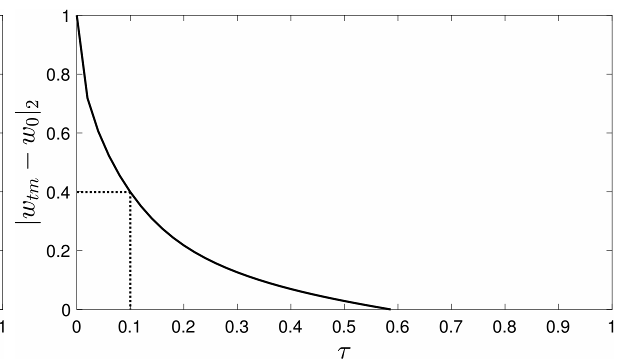
where T is the estimation window size, and $\mathcal{N}$ and $\mathcal{W}_N$ respectively denote the normal and the N-dimensional Wishart distributions. The minimum-turnover portfolio is then constructed from Equation (12) with the optimal τ defined in (17).[9] This procedure is iterated

[8] The relative utility of the equal-weight portfolio can be measured by $(1-\tau)$ on the x-axis, *i.e.*, when $|w_{tm} - w_0|_2 = 0$: *e.g.*, in D1, the utility of the equal-weight portfolio is about 12% of the optimal portfolio's utility.

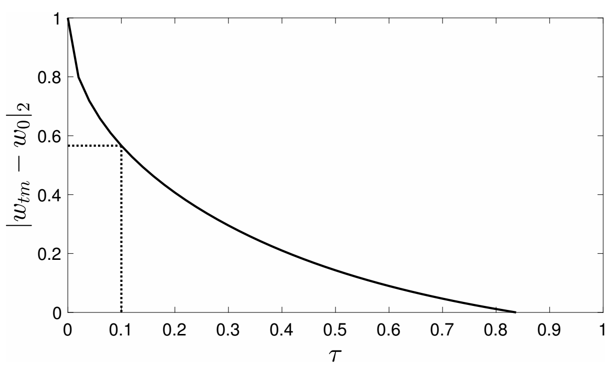
[9] For the calculation of τ^*, U^* and U_0 are estimated from the input parameter estimates via the method of Kan and Zhou (2007).



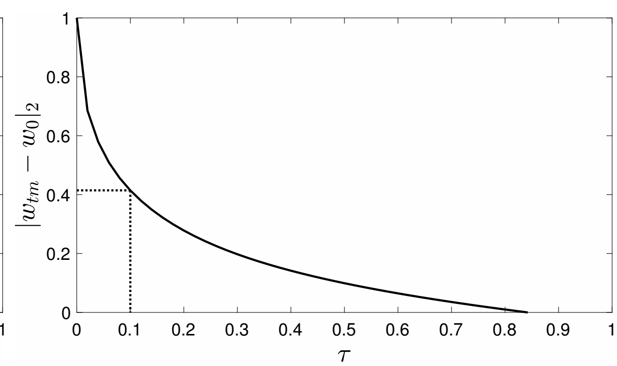
(a) D1. International



(b) D2. Industry



(c) D5. Fama-French



(d) D8. Momentum

Figure 3: Minimum-Turnover Portfolios

This figure demonstrates the relationship between the tolerance level τ and the distance from the minimum-turnover portfolio to the equal-weight reference portfolio. The distance is normalized by $|w_{mk} - w_0|_2$, *i.e.*, the distance of the underlying optimal portfolio. The datasets are described in Table 2.

10,000 times to obtain the expected out-of-sample utility[10]

$$EU = \mu_p - \frac{\gamma}{2}\sigma_p^2, \tag{19}$$

$$\mu_p = E[w'r] = E[w'\mu], \tag{20}$$

$$\sigma_p^2 = V[w'r] = E[w'\Sigma w] + E[(w'\mu)^2] - E[w'\mu]^2, \tag{21}$$

where the expectations are estimated from the sample averages.

Figure 4 compares the expected out-of-sample utility of the minimum-turnover portfolio (TMKE) with those of the underlying Markowitz portfolio (MK), the reference portfolio (EW), and the ex-post optimal portfolio (W*), for different estimation window sizes. The effectiveness of turnover minimization is evident. The minimum-turnover portfolio consistently outperforms its underlying portfolio in all datasets, and the improvement is particularly significant when T is small, *i.e.*, when the estimation error is significant. Even when the utility of EW is lower than that of MK, a higher utility is achieved by shrinking MK towards EW via turnover minimization.

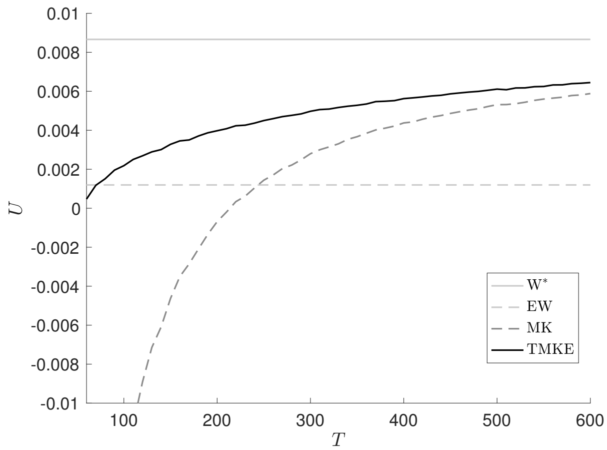
3. Empirical Analysis

This section evaluates the turnover minimization model against other portfolio models using real market data. Section 3.1 and 3.2 respectively describe the models and datasets, and the remainder of the section discusses empirical results.

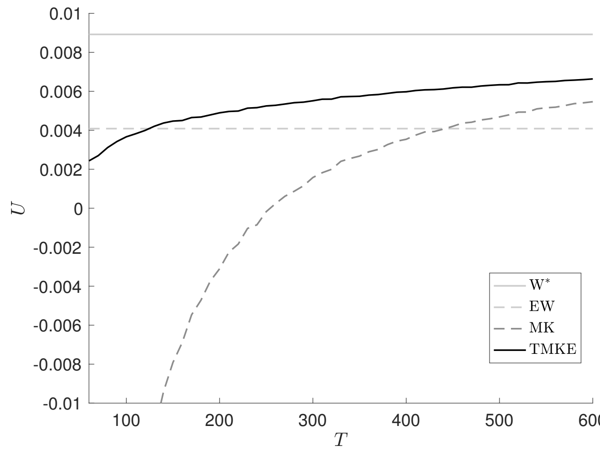
3.1. Portfolio Models

The portfolio models are listed in Table 1. Their implementation details are provided in Appendix B. W* is the *ex-post* mean-variance optimal portfolio obtained from the sample moments during the evaluation period. It represents the performance of the Markowitz

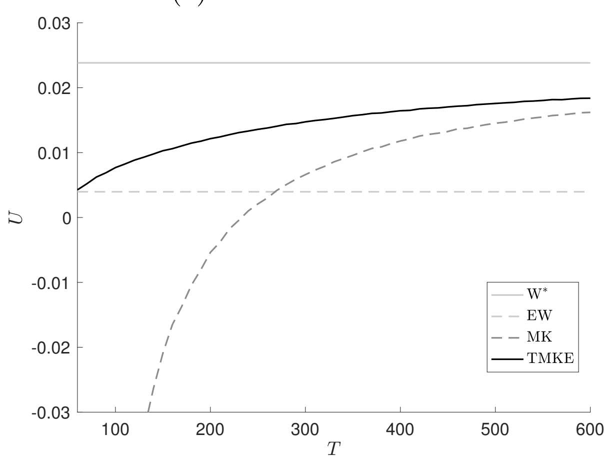
[10]Following Kan and Wang (2016), the unconditional variance is used.



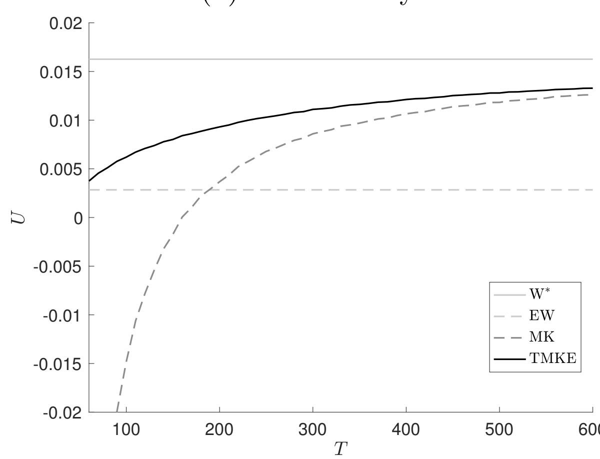
(a) D1. International



(b) D2. Industry



(c) D5. Fama-French



(d) D8. Momentum

Figure 4: Expected Out-of-sample Utility

This Figure displays the expected out-of-sample utilities of the minimum-turnover portfolio (TMKE), the underlying Markowitz portfolio (MK), the reference equal-weight portfolio (EW), and the ex-post optimal portfolio (W*). The horizontal axis represents the estimation window size.

portfolio when no estimation error is present. EW is the equal-weight portfolio. Both W* and EW are rebalanced back to their original weights every month.

The Markowitz mean-variance portfolio (MK), the global minimum-variance portfolio (MV), and their short-sale constrained versions (MK+, MV+) are also tested. OC(+) and VT are respectively the optimal constrained portfolio (with the short-sale constraint) and the volatility timing strategy of Kirby and Ostdiek (2012). TZMK and TZKZ are the shrinkage estimators of Tu and Zhou (2011), which respectively combine the Markowitz rule and the Kan and Zhou rule with the 1/N rule. TB+ is an extension of the active portfolio model of Treynor and Black (1973). In TB+, the equal-weight portfolio is employed in lieu of the market portfolio, and active assets are identified by regressing their returns on the return of the equal-weight portfolio. BL+ is an extension of the Black-Litterman model by Bessler et al. (2014), in which the prior is derived from the equal-weight portfolio instead of the market portfolio.

Eight versions of the turnover minimization model are tested. TMKE(+) and TMK0(+) are the turnover minimization incorporated with utility maximization (with the risk aversion parameter $\gamma = 3$), and TMVE(+) and TMV0(+) are the turnover minimization incorporated with variance minimization. The last letter indicates the reference portfolio: 'E' for the equal-weight portfolio and '0' for the current portfolio. '+' denotes the short-sale constraint.

3.2. Data

The portfolio models are evaluated on the eleven datasets described in Table 2. These are similar to the datasets used in DeMiguel et al. (2009) and Kirby and Ostdiek (2012), but more comprehensive. In the table, the evaluation period refers to the out-of-sample period, during which portfolios are rebalanced monthly.

The input parameters are estimated monthly from a rolling estimation window, $T = 60$, 120, or 240 months. The same evaluation period is used regardless of the estimation window size so that the empirical results can be compared across window sizes. For instance, when

Table 1: The Portfolio Models

This table lists the portfolio models considered in the empirical analysis. The '+' in the abbreviation denotes a model with the short-sale constraint (applied only to the risky assets). The details of the models are described in Section 2 and Appendix B.

Abbreviation	Description
W*	*Ex-post* mean-variance optimal portfolio
EW	Equal-weight portfolio
Classical models	
MK, MK+	Markowitz (1952) mean-variance portfolio
MV, MV+	Global minimum-variance portfolio
Kirby and Ostdiek (2012)	
OC, OC+	Optimal constrained portfolio: MK(+) without the risk-free asset
VT	Volatility timing strategy
Tu and Zhou (2011)	
TZMK	Combination of MK and EW
TZKZ	Combination of Kan and Zhou (2007) three-fund rule and EW
Incorporating the 1/N rule (in place of the market portfolio)	
TB+	Treynor and Black (1973)
BL+	Black and Litterman (1992)
Turnover Minimization	
TMKE(τ), TMKE(τ)+	Utility maximization-Turnover Minimization, $w_0 = w_{ew}$
TMK0(τ), TMK0(τ)+	Utility maximization-Turnover Minimization, $w_0 = w_{t-}$
TMVE(τ), TMVE(τ)+	Variance minimization-Turnover Minimization, $w_0 = w_{ew}$
TMV0(τ), TMV0(τ)+	Variance minimization-Turnover Minimization, $w_0 = w_{t-}$
τ: tolerance factor	

$T = 240$, the parameters are estimated using the sample in 1931.01-1950.12 in the first month, and when $T = 120$, they are estimated using the sample in 1941.01-1950.12.

Table 2: The Datasets

This table lists the datasets used in the empirical analysis. The eight international indices in D1 are the gross returns on large/mid-cap stocks from eight countries: Canada, France, Germany, Italy, Japan, Switzerland, United Kingdom, and the USA. All the other datasets consist of US stocks. The 20 size-sort portfolios (D4, 8, 9, 10) are obtained from the corresponding 25 portfolios by removing the five largest portfolios. The 50 large-cap stocks are the 50 largest stocks as of 1970.01 that still exist at the end of 2018.12. The dataset D1 is from the MSCI website (`https://www.msci.com/end-of-day-data-country`), D11 from the CRSP, and all other datasets are from K. French website (`http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`).

Dataset	Desciption	N	Evaluation Period
D1	8 International + World Indices	9	1990.01 - 2018.12
D2	10 Industry Portfolios + Market	11	1951.01 - 2018.12
D3	30 Industry Portfolios + Market	31	1951.01 - 2018.12
D4	20 FF Portfolios + Market	21	1951.01 - 2018.12
D5	10 Momentum Portfolios + Market	11	1951.01 - 2018.12
D6	10 Short-Term Reversal Portfolios + Market	11	1951.01 - 2018.12
D7	10 Long-Term Reversal Portfolios + Market	11	1951.01 - 2018.12
D8	20 Size/Momentum Portfolios + Market	21	1951.01 - 2018.12
D9	20 Size/Short-Term Reversal Portfolios + Market	21	1951.01 - 2018.12
D10	20 Size/Long-Term Reversal Portfolios + Market	21	1951.01 - 2018.12
D11	50 Large-cap Stocks in the US market	50	1990.01 - 2018.12

3.3. Empirical Results

3.3.1. Portfolio Construction and Evaluation

The input parameters are estimated every month during the evaluation period via the maximum likelihood estimator. These input parameters are used to rebalance the portfolios in Table 1. The minimum-turnover portfolios are constructed using two calibrated τ's, τ_b and τ_a, respectively for before and after transaction costs.

For the out-of-sample performance evaluation, the Sharpe ratio (SR), the certainty equivalent (CE), the cumulative return (CR), the skewness (SK), and the turnover (TO) are employed. These are defined as follows:

$$SR = \frac{\bar{r}_p}{s_p}, \tag{22}$$

$$CE = \bar{r}_p - \frac{\gamma}{2} s_p^2, \tag{23}$$

$$CR = \prod_{t=1}^{K} (1 + r_{p,t}), \tag{24}$$

$$SK = \frac{1}{K} \sum_{t=1}^{K} (r_{p,t} - \bar{r}_p)^3, \tag{25}$$

$$TO = \frac{1}{KN} \sum_{t=1}^{K} \sum_{i=1}^{N} |w_{i,t} - w_{i,t^-}|, \tag{26}$$

where $\bar{r}_p$ and s_p are respectively the mean and the standard deviation of the portfolio returns $r_{p,t}$ over the evaluation period, K and N are the number of months in the evaluation period and the number of assets, and $w_{i,t-}$ and $w_{i,t}$ are the weights of asset i immediately before and after rebalancing at time t. Following Balduzzi and Lynch (1999) and Olivares-Nadal and DeMiguel (2018), transaction costs are assumed to be 50 basis points for both buying and selling risky assets and 0 for the risk-free asset. The actual transaction costs of institutional investors are likely to be lower than this, and this assumption adversely affects the performance of optimal strategies, which carry higher turnover than the naïve strategy. For the statistical inference of the Sharpe ratio, the p-value for the Sharpe ratio difference between an optimal portfolio and the equal-weight portfolio is calculated using the method of Ledoit and Wolf (2008).

Since different portfolio models have different criteria and some portfolios are constrained to invest only in the risky assets, comparing models on a level playing field is not trivial. To mitigate the effects of these discrepancies, all the models are constrained to have the same variance. Variance targeting can be accomplished by adjusting portfolio weights as follows:

$$w := w \frac{\hat{\sigma}_p}{\sigma_T}, \tag{27}$$

where $\hat{\sigma}_p^2 = w'\hat{\Sigma}w$ is the *ex-ante* variance of the optimal portfolio, and σ_T^2 is the target variance. For W*, the true covariance matrix is used in $\hat{\sigma}_p^2$ instead of $\hat{\Sigma}$. σ_T^2 is set to be the variance of the equal-weight portfolio over the evaluation period. In order to ensure that the performance is not driven by the variance constraint, the models are also evaluated under the standard utility maximization objective.[11]

This section evaluates the portfolios mainly based on the results from the 120-month estimation window, and the results from other estimation window sizes as well as those from additional datasets are discussed in the robustness check in Section 3.4. These results are reported in the internet appendix (IA). The findings from the 120-month estimation window largely remain valid in other settings.

3.3.2. Performance of Turnover Minimization

Table 3 and 4 report the Sharpe ratios before and after transaction costs, and Table 5 and 6 report the turnovers, respectively under variance targeting and utility maximization. The difference between the turnover minimization models and their underlying models are also highlighted in Figure 5 and 6. The certainty equivalent, cumulative return, and skewness are reported in the IA. The empirical results are consistent with the findings from the simulation and reaffirm the effectiveness of turnover minimization.

Compared to their underlying models, the turnover minimization models incorporating the equal-weight portfolio present a higher Sharpe ratio and lower turnover. The results from variance targeting (Table 3) reveal that the average Sharpe ratio of TMKE and TMKE+ are respectively 0.250 and 0.174 before transaction costs and 0.187 and 0.161 after transaction costs, whereas the corresponding values of MK and MK+ are 0.200 and 0.167 before transaction costs and 0.039 and 0.149 after transaction costs. Similarly, the Sharpe ratios of TMVE(+) are 0.201 (0.168) and 0.168 (0.158) respectively before and after transaction

[11] Exceptions are EW, MV(+), and VT. These portfolios are constructed without exploiting the mean returns, whereas maximizing utility involves the mean returns as the adjustment formula is given by $w := w\frac{1}{\gamma}\frac{\hat{\mu}_p}{\hat{\sigma}_p^2}$, where $\hat{\mu}_p = w'\hat{\mu}$, $\hat{\sigma}_p^2 = w'\hat{\Sigma}w$. In order to preserve their nature, these portfolios are not adjusted.

Table 3: Sharpe Ratio under Variance Targeting

This table reports the Sharpe ratios of the portfolios in Table 1 under variance targeting. Input parameters are estimated from a rolling window of size $T = 120$, and transaction costs are assumed to be 50 basis points for risky assets and 0 for the risk-free asset. The columns represent the datasets described in Table 2. For the turnover minimization models, τ_b (τ_a) denotes the τ calibrated without (with) transaction costs. The Sharpe ratios statistically significantly higher than that of EW at 10% are marked by *.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	Mean
Before Transaction Cost												
W*	0.228	0.231	0.285	0.378	0.312	0.241	0.207	0.455	0.445	0.331	0.418	0.321
EW	0.083	0.156	0.146	0.149	0.123	0.134	0.148	0.142	0.138	0.157	0.172	0.141
MK	0.056	0.124	0.104	0.264*	0.283*	0.157	0.122	0.445*	0.385*	0.196	0.060	0.200
MK+	0.100	0.152	0.151	0.182*	0.200*	0.161*	0.153	0.223*	0.195*	0.163	0.156	0.167
MV	0.128	0.189	0.169	0.237*	0.180*	0.167	0.185	0.232*	0.168	0.203	0.135	0.181
MV+	0.114*	0.184	0.180	0.159	0.136	0.139	0.170*	0.162*	0.151	0.164	0.161	0.156
OC	0.115	0.149	0.129	0.282*	0.284*	0.177	0.147	0.441*	0.386*	0.205	0.080	0.218
OC+	0.111	0.150	0.144	0.168	0.193*	0.152	0.159	0.200*	0.184*	0.160	0.144	0.161
VT	0.093*	0.171*	0.160*	0.158*	0.133*	0.137	0.152*	0.150*	0.140	0.161*	0.206*	0.151
TB+	0.091	0.160	0.171	0.177*	0.191*	0.153*	0.155	0.222*	0.198*	0.161	0.136	0.165
BL+	0.095	0.158	0.160	0.178*	0.195*	0.161*	0.157	0.216*	0.185*	0.167	0.143	0.165
TZMK	0.057	0.155	0.147	0.286*	0.289*	0.167	0.155	0.447*	0.387*	0.235*	0.191	0.229
TZKZ	0.077	0.180	0.168	0.289*	0.286*	0.181	0.170	0.445*	0.388*	0.250*	0.182	0.238
TMKE(τ_b^*)	0.131	0.182	0.185	0.298*	0.305*	0.182	0.162	0.479*	0.401*	0.237*	0.184	0.250
TMKE+(τ_b^*)	0.097	0.159	0.167	0.186*	0.201*	0.160*	0.161	0.225*	0.197*	0.170*	0.186	0.174
TMK0(τ_b^*)	-0.042	0.083	-0.010	0.231*	0.265*	0.139	0.129	0.442*	0.368*	0.167	0.016	0.163
TMK0+(τ_b^*)	0.079	0.160	0.139	0.182*	0.196*	0.159*	0.151	0.221*	0.195*	0.159	0.178	0.166
TMVE(τ_b^*)	0.150*	0.202*	0.208*	0.254*	0.199*	0.173*	0.188	0.252*	0.172	0.232*	0.186	0.201
TMVE+(τ_b^*)	0.115*	0.195*	0.197*	0.170*	0.143*	0.146*	0.172*	0.170*	0.153	0.171*	0.213*	0.168
TMV0(τ_b^*)	0.121	0.180	0.154	0.234*	0.177*	0.156	0.187	0.228*	0.160	0.181	0.132	0.174
TMV0+(τ_b^*)	0.110*	0.178	0.191*	0.153	0.125	0.131	0.161	0.166*	0.146	0.154	0.181	0.154
After Transaction Cost												
W*	0.206	0.205	0.244	0.330	0.287	0.220	0.189	0.417	0.399	0.297	0.391	0.290
EW	0.081	0.153	0.143	0.147	0.121	0.132	0.146	0.140	0.137	0.156	0.166	0.138
MK	-0.083	-0.103	-0.134	0.092	0.169	0.028	-0.038	0.291*	0.216	0.026	-0.032	0.039
MK+	0.088	0.133	0.130	0.159	0.188*	0.146	0.133	0.208*	0.182*	0.139	0.131	0.149
MV	0.085	0.114	0.056	0.143	0.129	0.113	0.126	0.130	0.074	0.113	0.068	0.105
MV+	0.107	0.177	0.170	0.151	0.124	0.128	0.160	0.154	0.145	0.156	0.144	0.147
OC	0.021	-0.023	-0.078	0.123	0.194*	0.077	0.015	0.296*	0.239*	0.041	-0.010	0.081
OC+	0.097	0.130	0.124	0.145	0.181*	0.132	0.141	0.184*	0.172*	0.138	0.121	0.142
VT	0.090*	0.168*	0.156*	0.156*	0.130*	0.135	0.150*	0.148*	0.138	0.159*	0.199*	0.148
TB+	0.082	0.146	0.149	0.152	0.177*	0.134	0.141	0.209*	0.185*	0.136	0.108	0.147
BL+	0.083	0.145	0.145	0.165	0.185*	0.151*	0.145	0.205*	0.172*	0.152	0.119	0.151
TZMK	-0.046	-0.068	-0.072	0.151	0.187	0.064	0.046	0.304*	0.234*	0.119	0.166	0.099
TZKZ	-0.009	-0.021	-0.040	0.160	0.187	0.092	0.078	0.306*	0.236*	0.145	0.139	0.116
TMKE(τ_a^*)	0.080	0.133	0.122	0.207*	0.238*	0.139	0.148	0.376*	0.307*	0.154	0.158	0.187
TMKE+(τ_a^*)	0.087	0.154	0.152	0.169*	0.188*	0.147	0.146	0.210*	0.182*	0.159	0.176	0.161
TMK0(τ_a^*)	-0.138	-0.083	-0.160	0.181	0.224*	0.117	0.028	0.341*	0.283*	0.123	-0.018	0.082
TMK0+(τ_a^*)	0.068	0.146	0.116	0.154	0.189*	0.150*	0.141	0.217*	0.189*	0.159	0.173	0.155
TMVE(τ_a^*)	0.122*	0.183*	0.180*	0.199*	0.160*	0.141	0.153	0.206*	0.134	0.190	0.179	0.168
TMVE+(τ_a^*)	0.109*	0.184*	0.186*	0.162*	0.137*	0.133	0.162*	0.158*	0.146	0.164	0.195*	0.158
TMV0(τ_a^*)	0.106	0.182	0.163	0.177	0.128	0.142	0.120	0.183	0.069	0.136	0.105	0.137
TMV0+(τ_a^*)	0.103	0.173	0.182*	0.136	0.113	0.139	0.160*	0.154	0.148	0.142	0.175	0.148

Table 4: Sharpe Ratio under Utility Maximization

This table reports the Sharpe ratios of the portfolios in Table 1 under utility maximization. Input parameters are estimated from a rolling window of size $T = 120$, and transaction costs are assumed to be 50 basis points for risky assets and 0 for the risk-free asset. The columns represent the datasets described in Table 2. For the turnover minimization models, τ_b (τ_a) denotes the τ calibrated without (with) transaction costs. The Sharpe ratios statistically significantly higher than that of EW at 10% are marked by *.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	Mean
Before Transaction Cost												
W*	0.228	0.231	0.285	0.378	0.312	0.241	0.207	0.455	0.445	0.331	0.418	0.321
EW	0.094	0.162	0.150	0.154	0.130	0.140	0.157	0.148	0.143	0.164	0.186	0.148
MK	0.043	0.128	0.107	0.276*	0.277*	0.165	0.136	0.429*	0.390*	0.197	0.056	0.200
MK+	0.055	0.154	0.149	0.167	0.181*	0.152	0.142	0.203*	0.178	0.145	0.136	0.151
MV	0.129	0.179	0.156	0.243*	0.182*	0.172	0.181	0.234*	0.180	0.206	0.136	0.182
MV+	0.125*	0.174	0.170	0.165	0.143	0.146	0.175*	0.168*	0.157	0.168	0.167	0.160
OC	0.054	0.062	0.063	0.215	0.258*	0.123	0.075	0.388*	0.374*	0.150	0.016	0.162
OC+	0.082	0.131	0.128	0.179*	0.203*	0.155	0.139	0.212*	0.204*	0.162	0.114	0.155
VT	0.103*	0.171*	0.161*	0.164*	0.141*	0.144	0.160*	0.157*	0.146	0.169*	0.215	0.157
TB+	0.032	0.146	0.153	0.163	0.171	0.147	0.145	0.200*	0.181	0.142	0.117	0.145
BL+	0.058	0.163	0.157	0.160	0.179	0.146	0.144	0.193	0.163	0.147	0.126	0.149
TZMK	0.059	0.158	0.144	0.284*	0.270*	0.180	0.167	0.416*	0.388*	0.219	0.201	0.226
TZKZ	0.065	0.174	0.164	0.290*	0.264*	0.190	0.182	0.411*	0.387*	0.236*	0.188	0.232
TMKE(τ_b^*)	0.105	0.181	0.177	0.296*	0.295*	0.195	0.188	0.467*	0.400*	0.209*	0.209*	0.247
TMKE+(τ_b^*)	0.092	0.182	0.190*	0.176	0.180*	0.159	0.154	0.206*	0.175*	0.171	0.190	0.170
TMK0(τ_b^*)	0.120	0.178	0.160	0.259*	0.250*	0.155	0.112	0.424*	0.374*	0.170	0.148	0.214
TMK0+(τ_b^*)	0.101	0.171	0.170	0.168	0.177*	0.152	0.174*	0.200*	0.174	0.158	0.195	0.167
TMVE(τ_b^*)	0.156*	0.197*	0.200*	0.259*	0.201*	0.182*	0.188	0.260*	0.199*	0.237*	0.198	0.207
TMVE+(τ_b^*)	0.125*	0.175	0.186*	0.177*	0.150*	0.150	0.178*	0.173*	0.160	0.176	0.214*	0.169
TMV0(τ_b^*)	0.124	0.173	0.172	0.241*	0.169	0.154	0.158	0.225*	0.180	0.191	0.137	0.175
TMV0+(τ_b^*)	0.121*	0.167	0.170	0.162	0.140	0.137	0.163	0.171*	0.152	0.148	0.190	0.156
After Transaction Cost												
W*	0.196	0.188	0.214	0.283	0.265	0.210	0.181	0.366	0.346	0.267	0.364	0.262
EW	0.091	0.159	0.147	0.153	0.129	0.139	0.155	0.146	0.142	0.163	0.180	0.146
MK	-0.140	-0.491	-0.861	-0.056	0.076	-0.022	-0.121	0.063	-0.021	-0.093	-0.123	-0.163
MK+	0.040	0.130	0.124	0.143	0.165	0.135	0.120	0.186	0.163	0.118	0.106	0.130
MV	0.088	0.108	0.051	0.153	0.132	0.120	0.124	0.139	0.089	0.121	0.073	0.109
MV+	0.118	0.167	0.160	0.159	0.132	0.136	0.165	0.161	0.153	0.161	0.149	0.151
OC	-0.136	-0.501	-0.886	-0.117	0.061	-0.071	-0.186	0.040	-0.007	-0.141	-0.157	-0.191
OC+	0.067	0.114	0.111	0.158	0.194*	0.139	0.119	0.199*	0.195*	0.138	0.095	0.139
VT	0.101*	0.168*	0.158*	0.162*	0.139*	0.142	0.159	0.156*	0.145	0.167*	0.208	0.155
TB+	0.020	0.126	0.129	0.139	0.152	0.125	0.125	0.185	0.165	0.112	0.089	0.124
BL+	0.045	0.148	0.139	0.145	0.168	0.135	0.130	0.182	0.150	0.132	0.099	0.134
TZMK	-0.062	-0.215	-0.206	0.105	0.127	0.055	0.028	0.206	0.161	0.077	0.172	0.041
TZKZ	-0.025	-0.112	-0.125	0.135	0.140	0.090	0.074	0.221	0.179	0.110	0.144	0.076
TMKE(τ_a^*)	0.088	0.155	0.146	0.201*	0.222*	0.149	0.154	0.340*	0.298*	0.173	0.201	0.193
TMKE+(τ_a^*)	0.090	0.159	0.156	0.154	0.164	0.140	0.153	0.187	0.162	0.164	0.182	0.156
TMK0(τ_a^*)	0.118*	0.141	0.149	0.101	0.175	0.097	0.119	0.224	0.230	0.105	0.130	0.144
TMK0+(τ_a^*)	0.101	0.170	0.163	0.138	0.164	0.137	0.162	0.193*	0.176*	0.148	0.189	0.158
TMVE(τ_a^*)	0.125*	0.171	0.179*	0.204*	0.166*	0.150	0.156	0.219*	0.143	0.197	0.195	0.173
TMVE+(τ_a^*)	0.119*	0.169	0.175	0.168*	0.146*	0.143	0.167*	0.163*	0.153	0.170	0.203*	0.161
TMV0(τ_a^*)	0.109	0.170	0.153	0.160	0.134	0.149	0.109	0.198	0.109	0.151	0.112	0.141
TMV0+(τ_a^*)	0.113	0.165	0.172	0.146	0.124	0.147	0.166	0.162	0.153	0.152	0.184	0.153

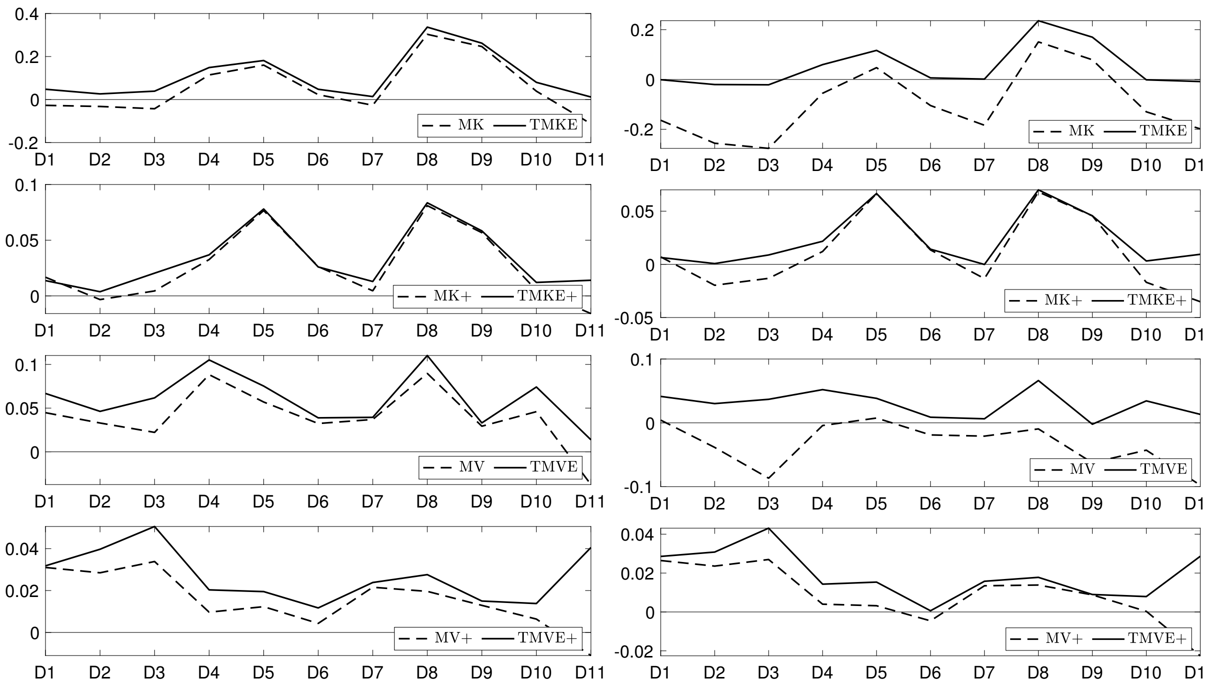
costs, whereas the corresponding values of MV(+) are 0.181 (0.156) and 0.105 (0.147). The improvement is consistent across datasets and more prominent after transaction costs, owing to the significantly lower turnover of the minimum-turnover portfolios. There are only few cases where the turnover minimization models are outperformed by their underlying models before transaction costs, *e.g.*, TMKE+ in D9 under utility maximization, and none of them is outperformed after transaction costs.

All minimum-turnover portfolios shrunk towards the equal-weight portfolio, *i.e.*, TMKE, TMKE+, TMVE, and TMVE+, outperform EW in all datasets before transaction costs under variance targeting.[12] If only statistically significant cases are counted, they are respectively 5, 6, 8, and 10. Among the other models, only VT outperforms EW in all datasets (9 times statistically significant). The superior performance of the turnover minimization is maintained even after the conservatively set transaction costs are applied. In particular, the short-sale constrained models (TMKE+ and TMVE+) outperform EW in all datasets after transaction costs (4 and 8 times statistically significant, respectively). VT also outperforms EW in all datasets after transaction costs but yields a lower Sharpe ratio compared to TMKE+ or TMVE+.

The turnover minimization models continue to perform superior under utility maximization, but the performances of TMKE and TMKE+ become less pronounced. In contrast, TMVE and TMVE+ maintain a similar level of performance, outperforming EW in all datasets both before and after transaction costs. The performance difference between variance targeting and utility maximization can be attributed to the fact that variance targeting is less susceptible to the estimation error of the mean, as it shifts the portfolio to have the target variance.

When short-sale is not allowed, the underlying portfolios are closer to EW, and the benefit from turnover minimization becomes limited. Nevertheless, incorporating turnover

[12]The market portfolio (obtained from the K. French website), another potential benchmark, slightly underperforms the EW of the 20 FF portfolios + Market, with the Sharpe ratio 0.142, cumulative return 6.968, certainty equivalent 0.333, and skewness -0.542.



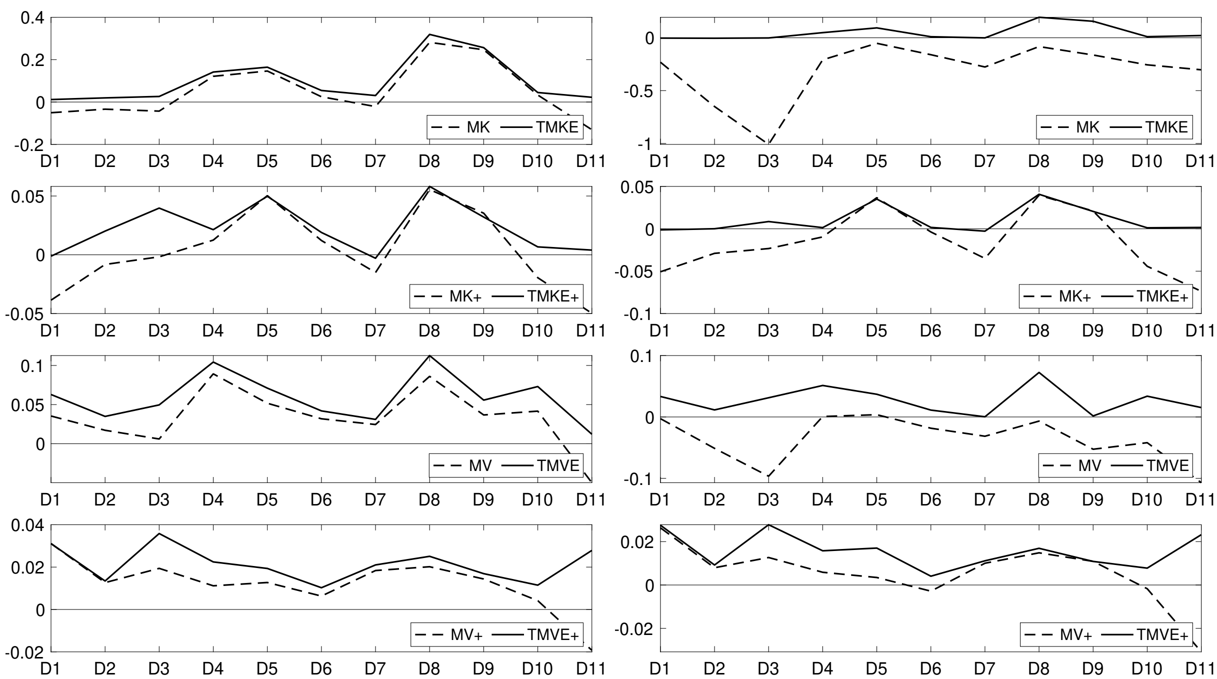
(a) Before transaction cost (b) After transaction cost

Figure 5: Sharpe Ratio under Variance Targeting

This figure compares the turnover minimization models with their underlying models under variance targeting when $T = 120$. The vertical axis represents the Sharpe ratio difference from EW, and the horizontal axis represents the datasets.

minimization improves the performance of the underlying models consistently across datasets and optimization criteria, and TMKE+ and TMVE+ are the best performing long-only portfolios in terms of the mean Sharpe ratio, respectively under variance targeting and utility maximization.

As evidenced by the superior performance of the short-sale constrained minimum-turnover portfolios, it is often beneficial to constrain portfolio weights to reduce turnover and leverage. Besides, many financial institutions do not allow short positions in their portfolios. The turnover minimization admits the flexibility of adding these constraints while accounting for parameter uncertainty.



(a) Before transaction cost (b) After transaction cost

Figure 6: Sharpe Ratio under Utility Maximization

This figure compares the turnover minimization models with their underlying models under utility maximizatoin when $T = 120$. The vertical axis represents the Sharpe ratio difference from EW, and the horizontal axis represents the datasets.

As to the reference portfolio, the current portfolio appears to be an ineffective target. TMK0(+) and TMV0(+) perform considerably weaker than their equal-weight counterparts, TMKE(+) and TMVE(+). These models usually underperform their underlying models

Table 5: Turnover under Variance Targeting

This table reports the turnover of the portfolio models in Table 1 under variance targeting when $T = 120$. Turnover is defined by the formula in (26). The columns represent the datasets described in Table 2. For the turnover minimization models, τ_b (τ_a) denotes the τ calibrated without (with) transaction costs.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	Mean
W*	0.022	0.019	0.012	0.023	0.020	0.017	0.014	0.018	0.023	0.016	0.004	0.017
EW	0.003	0.002	0.001	0.001	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.001
MK	0.128	0.178	0.097	0.109	0.111	0.124	0.149	0.107	0.111	0.104	0.033	0.114
MK+	0.013	0.016	0.007	0.011	0.010	0.012	0.016	0.008	0.007	0.012	0.005	0.011
MV	0.047	0.068	0.051	0.062	0.050	0.051	0.058	0.068	0.063	0.058	0.021	0.054
MV+	0.007	0.006	0.003	0.004	0.010	0.009	0.008	0.004	0.003	0.004	0.004	0.006
OC	0.093	0.137	0.085	0.101	0.086	0.097	0.125	0.101	0.098	0.100	0.031	0.096
OC+	0.015	0.016	0.006	0.011	0.011	0.016	0.014	0.008	0.006	0.011	0.004	0.011
VT	0.003	0.002	0.001	0.001	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.002
TB+	0.009	0.012	0.007	0.012	0.012	0.015	0.011	0.007	0.007	0.013	0.005	0.010
BL+	0.012	0.010	0.005	0.007	0.008	0.008	0.010	0.005	0.007	0.007	0.004	0.008
TZMK	0.099	0.167	0.079	0.082	0.097	0.095	0.094	0.098	0.099	0.066	0.004	0.089
TZKZ	0.085	0.157	0.080	0.079	0.094	0.082	0.082	0.095	0.099	0.062	0.009	0.084
TMKE(τ_b^*)	0.060	0.149	0.061	0.060	0.064	0.066	0.099	0.075	0.067	0.055	0.003	0.069
TMKE+(τ_b^*)	0.011	0.017	0.007	0.011	0.009	0.012	0.012	0.007	0.006	0.007	0.002	0.009
TMK0(τ_b^*)	0.068	0.160	0.076	0.064	0.048	0.058	0.095	0.061	0.058	0.052	0.008	0.068
TMK0+(τ_b^*)	0.005	0.009	0.005	0.011	0.007	0.012	0.004	0.004	0.002	0.004	0.001	0.006
TMVE(τ_b^*)	0.049	0.016	0.010	0.038	0.053	0.046	0.054	0.052	0.063	0.033	0.004	0.038
TMVE+(τ_b^*)	0.009	0.007	0.004	0.005	0.004	0.011	0.009	0.006	0.004	0.004	0.002	0.006
TMV0(τ_b^*)	0.027	0.022	0.012	0.054	0.046	0.017	0.039	0.036	0.049	0.050	0.006	0.033
TMV0+(τ_b^*)	0.006	0.004	0.003	0.003	0.005	0.003	0.005	0.003	0.002	0.002	0.001	0.003
TMKE(τ_a^*)	0.024	0.014	0.004	0.035	0.051	0.045	0.023	0.058	0.049	0.040	0.002	0.032
TMKE+(τ_a^*)	0.010	0.010	0.005	0.008	0.009	0.010	0.006	0.007	0.006	0.003	0.002	0.007
TMK0(τ_a^*)	0.059	0.158	0.068	0.030	0.031	0.026	0.053	0.041	0.035	0.030	0.008	0.049
TMK0+(τ_a^*)	0.005	0.005	0.002	0.003	0.003	0.003	0.002	0.002	0.002	0.001	0.001	0.003
TMVE(τ_a^*)	0.026	0.008	0.005	0.015	0.023	0.016	0.037	0.026	0.005	0.019	0.004	0.017
TMVE+(τ_a^*)	0.011	0.008	0.004	0.004	0.005	0.004	0.007	0.004	0.005	0.004	0.002	0.005
TMV0(τ_a^*)	0.009	0.006	0.006	0.015	0.023	0.013	0.022	0.031	0.020	0.025	0.006	0.016
TMV0+(τ_a^*)	0.004	0.004	0.002	0.002	0.002	0.002	0.003	0.002	0.001	0.002	0.001	0.002

Table 6: Turnover under Utility Maximization

This table reports the turnover of the portfolio models in Table 1 under utility maximization when $T = 120$. Turnover is defined by the formula in (26). The columns represent the datasets described in Table 2. For the turnover minimization models, τ_b (τ_a) denotes the τ calibrated without (with) transaction costs.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	Mean
W*	0.055	0.061	0.044	0.114	0.088	0.046	0.032	0.126	0.138	0.068	0.030	0.073
EW	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
MK	0.439	3.922	4.338	1.022	0.805	0.630	0.824	1.537	1.556	0.759	0.397	1.475
MK+	0.024	0.047	0.021	0.021	0.026	0.027	0.034	0.018	0.014	0.023	0.017	0.025
MV	0.034	0.043	0.024	0.032	0.035	0.037	0.040	0.032	0.033	0.028	0.011	0.032
MV+	0.006	0.004	0.002	0.003	0.008	0.008	0.007	0.003	0.002	0.003	0.002	0.004
OC	0.401	3.597	4.210	0.847	0.741	0.535	0.630	1.348	1.407	0.689	0.359	1.342
OC+	0.016	0.015	0.006	0.010	0.009	0.014	0.017	0.007	0.005	0.012	0.004	0.011
VT	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
TB+	0.018	0.038	0.021	0.022	0.032	0.032	0.028	0.016	0.014	0.024	0.014	0.024
BL+	0.014	0.022	0.011	0.010	0.013	0.014	0.017	0.009	0.009	0.010	0.010	0.013
TZMK	0.139	0.636	0.236	0.236	0.315	0.200	0.208	0.448	0.422	0.130	0.004	0.270
TZKZ	0.097	0.421	0.144	0.176	0.227	0.145	0.155	0.337	0.311	0.095	0.008	0.192
TMKE(τ_b^*)	0.012	0.012	0.007	0.245	0.272	0.160	0.036	0.511	0.358	0.008	0.001	0.147
TMKE+(τ_b^*)	0.002	0.006	0.003	0.021	0.019	0.022	0.003	0.014	0.011	0.004	0.001	0.010
TMK0(τ_b^*)	0.012	0.030	0.014	0.799	0.405	0.448	0.031	1.418	0.582	0.014	0.003	0.341
TMK0+(τ_b^*)	0.002	0.004	0.002	0.019	0.018	0.022	0.002	0.015	0.005	0.002	0.001	0.008
TMVE(τ_b^*)	0.042	0.021	0.006	0.020	0.031	0.040	0.035	0.023	0.032	0.017	0.003	0.024
TMVE+(τ_b^*)	0.008	0.006	0.002	0.004	0.003	0.008	0.009	0.004	0.003	0.003	0.002	0.005
TMV0(τ_b^*)	0.019	0.012	0.004	0.028	0.026	0.011	0.023	0.015	0.024	0.022	0.005	0.017
TMV0+(τ_b^*)	0.005	0.002	0.001	0.002	0.005	0.003	0.003	0.002	0.001	0.001	0.001	0.002
TMKE(τ_a^*)	0.008	0.002	0.001	0.061	0.124	0.082	0.004	0.163	0.141	0.004	0.001	0.054
TMKE+(τ_a^*)	0.002	0.004	0.002	0.011	0.018	0.017	0.003	0.014	0.010	0.001	0.001	0.008
TMK0(τ_a^*)	0.003	0.003	0.003	0.179	0.168	0.132	0.012	0.330	0.283	0.008	0.003	0.102
TMK0+(τ_a^*)	0.002	0.002	0.001	0.014	0.011	0.012	0.001	0.008	0.003	0.000	0.001	0.005
TMVE(τ_a^*)	0.021	0.006	0.004	0.009	0.015	0.015	0.032	0.014	0.003	0.011	0.003	0.012
TMVE+(τ_a^*)	0.011	0.006	0.002	0.003	0.004	0.005	0.005	0.003	0.003	0.003	0.002	0.004
TMV0(τ_a^*)	0.006	0.004	0.003	0.009	0.014	0.009	0.015	0.013	0.010	0.012	0.005	0.009
TMV0+(τ_a^*)	0.004	0.002	0.001	0.001	0.001	0.001	0.002	0.001	0.000	0.001	0.001	0.001

before transaction costs and perform marginally better only after transaction costs. A similar observation has been made by Han (2019), who analytically shows that the equal-weight portfolio is a more effective shrinkage target.

The calibration of τ proves to be effective. In the presence of transaction costs, high turnover is harmful and a higher degree of tolerance is desired. The calibration results reported in Table 7 are in line with this statement: the tolerance levels calibrated under transaction costs (τ_a) are greater than those calibrated without transaction costs (τ_b), and result in substantially lower turnover as reported in Table 5 and 6. Consequently, the minimum-turnover portfolios from τ_b perform better before transaction costs, and those from τ_a perform better after transaction costs.

3.3.3. Performance of the Other Models

Among the other models, Tu and Zhou (2011) shrinkage estimators, TZMK and TZKZ, perform comparably to TMKE before transaction costs. Their average Sharpe ratios before transaction costs are 0.229 and 0.238 under variance targeting and 0.226 and 0.232 under utility maximization, whereas the corresponding values of TMKE are 0.250 and 0.226. Between the two, TZKZ appears to perform better than TZMK. Nevertheless, their performance is significantly deteriorated once transaction costs are taken into account due to high turnover. While TZMK and TZKZ successfully enhance the Sharpe ratio and reduce turnover in comparison to the Markowitz model, they still suffer from costly portfolio rebalancing and underperform EW after transaction costs in the majority of the datasets. This result further emphasizes the importance of the ability to incorporate constraints in shrinkage models.

Among the strategies that incorporate the 1/N rule, the variant of the Black-Litterman model (BL+) performs best in terms of the mean Sharpe ratio when transaction costs are taken into account. Nevertheless, it hardly outperforms EW statistically significantly, and is generally outperformed by the turnover minimization models.

Table 7: Calibration of τ

This table reports the mean of the calibrated τ's for each turnover minimization model when $T = 120$. $\bar{\tau}_b$ ($\bar{\tau}_a$) denotes the mean of the τ's calibrated without (with) transaction costs over the sample period. Transaction costs are assumed to be 50 basis points for risky assets and 0 for the risk-free asset.

		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11
Variance Targeting												
TMKE	$\bar{\tau}_b^*$	0.574	0.615	0.647	0.340	0.162	0.223	0.392	0.152	0.241	0.363	0.954
	$\bar{\tau}_a^*$	0.814	0.978	0.989	0.612	0.290	0.517	0.890	0.336	0.420	0.636	0.955
TMKE+	$\bar{\tau}_b^*$	0.335	0.717	0.103	0.021	0.034	0.001	0.290	0.020	0.040	0.230	0.808
	$\bar{\tau}_a^*$	0.473	0.904	0.518	0.102	0.050	0.070	0.589	0.040	0.059	0.366	0.821
TMK0	$\bar{\tau}_b^*$	0.614	0.191	0.691	0.347	0.319	0.407	0.448	0.135	0.277	0.465	0.954
	$\bar{\tau}_a^*$	0.826	0.233	0.987	0.698	0.415	0.579	0.737	0.423	0.484	0.737	0.954
TMK0+	$\bar{\tau}_b^*$	0.304	0.056	0.264	0.015	0.034	0.009	0.403	0.021	0.069	0.287	0.804
	$\bar{\tau}_a^*$	0.335	0.067	0.641	0.186	0.056	0.127	0.436	0.045	0.080	0.463	0.805
TMVE	$\bar{\tau}_b^*$	0.082	0.448	0.293	0.116	0.042	0.108	0.004	0.036	0.143	0.078	0.790
	$\bar{\tau}_a^*$	0.125	0.977	0.417	0.279	0.108	0.288	0.108	0.148	0.817	0.210	0.814
TMVE+	$\bar{\tau}_b^*$	0.008	0.087	0.046	0.081	0.049	0.042	0.008	0.040	0.143	0.065	0.355
	$\bar{\tau}_a^*$	0.011	0.638	0.053	0.093	0.069	0.108	0.018	0.061	0.167	0.079	0.375
TMV0	$\bar{\tau}_b^*$	0.074	0.592	0.440	0.123	0.052	0.116	0.026	0.037	0.133	0.078	0.659
	$\bar{\tau}_a^*$	0.106	0.974	0.586	0.475	0.219	0.173	0.213	0.103	0.718	0.341	0.670
TMV0+	$\bar{\tau}_b^*$	0.009	0.317	0.063	0.165	0.080	0.057	0.030	0.121	0.111	0.168	0.273
	$\bar{\tau}_a^*$	0.012	0.414	0.112	0.208	0.118	0.064	0.038	0.143	0.118	0.202	0.273
Utility Maximization												
TMKE	$\bar{\tau}_b^*$	0.696	0.490	0.616	0.341	0.162	0.243	0.391	0.185	0.265	0.433	0.957
	$\bar{\tau}_a^*$	0.844	0.977	0.990	0.672	0.358	0.448	0.867	0.480	0.513	0.741	0.963
TMKE+	$\bar{\tau}_b^*$	0.715	0.128	0.210	0.034	0.047	0.084	0.405	0.027	0.044	0.471	0.837
	$\bar{\tau}_a^*$	0.745	0.404	0.545	0.255	0.073	0.186	0.557	0.040	0.076	0.614	0.845
TMK0	$\bar{\tau}_b^*$	0.702	0.331	0.574	0.339	0.271	0.409	0.430	0.111	0.290	0.497	0.954
	$\bar{\tau}_a^*$	0.860	0.985	0.987	0.763	0.498	0.584	0.793	0.593	0.605	0.772	0.958
TMK0+	$\bar{\tau}_b^*$	0.658	0.084	0.158	0.029	0.032	0.058	0.346	0.022	0.135	0.443	0.828
	$\bar{\tau}_a^*$	0.670	0.357	0.472	0.248	0.071	0.195	0.437	0.051	0.159	0.620	0.831
TMVE	$\bar{\tau}_b^*$	0.088	0.335	0.267	0.114	0.052	0.077	0.007	0.042	0.150	0.083	0.816
	$\bar{\tau}_a^*$	0.150	0.460	0.387	0.272	0.130	0.260	0.153	0.147	0.776	0.215	0.850
TMVE+	$\bar{\tau}_b^*$	0.008	0.184	0.044	0.079	0.044	0.038	0.009	0.053	0.113	0.085	0.386
	$\bar{\tau}_a^*$	0.014	0.207	0.049	0.089	0.060	0.086	0.021	0.080	0.126	0.096	0.394
TMV0	$\bar{\tau}_b^*$	0.084	0.557	0.566	0.114	0.098	0.103	0.077	0.047	0.103	0.093	0.709
	$\bar{\tau}_a^*$	0.117	0.650	0.715	0.454	0.208	0.141	0.243	0.099	0.543	0.305	0.720
TMV0+	$\bar{\tau}_b^*$	0.009	0.332	0.263	0.151	0.067	0.046	0.048	0.113	0.108	0.184	0.289
	$\bar{\tau}_a^*$	0.012	0.338	0.275	0.191	0.114	0.055	0.062	0.134	0.107	0.184	0.290

Another model that is worth noting is the volatility timing (VT) of Kirby and Ostdiek (2012). Although it outperforms EW only marginally (the average Sharpe ratios of VT and EW before (after) transaction costs are respectively 0.151 (0.148) and 0.141 (0.138) under variance targeting, and 0.157 (0.155) and 0.148 (0.146) under utility maximization), the difference is statistically significant in nine datasets after transaction costs under variance targeting. This is because the portfolio weights of VT are entirely determined by the cross-sectional variation of the variances of the asset returns, which are stable over time. The stable cross-sectional variation leads to a stable VT portfolio and consequently a small standard deviation of the Sharpe ratio difference between VT and EW. Therefore, even a marginal outperformance becomes statistically significant. With the superior performance of the volatility timing, it could be that using VT as the shrinkage target enhances the performance of the turnover minimization. This conjecture is examined in Section 3.3.6.

All in all, the turnover minimization with the equal-weight reference portfolio performs superior in comparison to the other strategies. The unconstrained models (TMKE and TMVE) perform best before transaction costs, while the short-sale constrained counterparts (TMKE+ and TMVE+) perform better when subject to transaction costs. The results from certainty equivalents (reported in the IA) are qualitatively similar to those from Sharpe ratios and reaffirm the superiority of the turnover minimization models. The turnover minimization allows us to enjoy the benefits of the shrinkage estimator without sacrificing modeling flexibility.

3.3.4. Comparison with the Other Models

To examine whether the turnover minimization models significantly outperform the other models, Table 8 computes the p-value of the Sharpe ratio difference between TMKE and an alternative model. TMKE is chosen among the turnover minimization models as it has the highest mean Sharpe ratio.

The table reveals that TMKE significantly outperforms EW, MK, MK+, MV+, OC+, VT, TB+, and BL+ in at least five datasets before transaction costs, and MK, MV, OC, TZMK, and TZKZ after transaction costs. TMKE performs particularly well in datasets D4, D5, D8, D9, and D10. Notably, a large Sharpe ratio difference does not always imply a significant difference.

3.3.5. Alternative Calibration Method

DeMiguel et al. (2009) develop portfolio models that generalize norm constraints and calibrate them using two methods; cross-validation and last period return maximization. The latter leads to superior portfolio performance in their study and is examined as an alternative method to calibrate τ. The procedure is simple: choose τ that maximizes the last period return. The results are reported in Table 9.

The new calibration method improves portfolio performance in certain datasets, D2 and D3 in particular, and the minimum-turnover portfolios statistically significantly outperform EW in more datasets before transaction costs. However, it yields lower average Sharpe ratios and there is no discernible improvement across all datasets. Moreover, maximizing the last period return accompanies high turnover, and the portfolios perform significantly worse after accounting for transaction costs compared to those using the calibration method in Section 2.3. Although DeMiguel et al. (2009) do not take transaction costs into account, they also recognize that maximizing the last period return involves higher turnover.

3.3.6. Alternative Reference Portfolio

The volatility timing portfolio (VT) marginally but consistently outperforms the equal-weight portfolio. Hence, it may serve better as the reference portfolio than the equal-weight portfolio. This section tests the turnover minimization using VT as the reference portfolio. Table 10 reports the Sharpe ratios of these models under variance targeting. The turnover minimization models incorporating VT are denoted by TMKVT(+) when the underlying strategy is utility maximization, and TMVVT(+) when it is variance minimization.

Table 8: Sharp ratio difference between TMKE and an alternative model

This table reports the Sharpe ratio difference between TMKE and a model and its p-value (in parentheses). A positive value means TMKE has a larger Share ratio. p-values below 0.1 are marked by *.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	Mean
Before Transaction Cost												
EW	0.048	0.026	0.039	0.149*	0.181*	0.048	0.014	0.337*	0.262*	0.080*	0.012	0.109
	(0.334)	(0.488)	(0.254)	(0.001)	(0.000)	(0.245)	(0.670)	(0.000)	(0.000)	(0.031)	(0.683)	
MK	0.075	0.058*	0.082*	0.034*	0.021	0.025	0.040	0.033*	0.016	0.041*	0.125	0.050
	(0.123)	(0.070)	(0.056)	(0.084)	(0.141)	(0.127)	(0.110)	(0.009)	(0.340)	(0.035)	(0.181)	
MK+	0.031	0.030	0.035	0.116*	0.105*	0.021	0.009	0.256*	0.205*	0.075*	0.028	0.083
	(0.508)	(0.357)	(0.270)	(0.002)	(0.010)	(0.551)	(0.754)	(0.000)	(0.001)	(0.026)	(0.531)	
MV	0.003	-0.007	0.017	0.061*	0.124*	0.015	-0.023	0.247*	0.233*	0.034	0.049	0.069
	(0.954)	(0.835)	(0.694)	(0.092)	(0.006)	(0.705)	(0.443)	(0.000)	(0.000)	(0.326)	(0.452)	
MV+	0.017	-0.002	0.005	0.139*	0.169*	0.043	-0.008	0.317*	0.249*	0.074*	0.023	0.093
	(0.741)	(0.951)	(0.877)	(0.001)	(0.001)	(0.303)	(0.804)	(0.000)	(0.000)	(0.052)	(0.613)	
OC	0.016	0.033	0.056	0.016	0.021	0.005	0.015	0.037*	0.014	0.033	0.104	0.032
	(0.763)	(0.331)	(0.158)	(0.465)	(0.386)	(0.863)	(0.526)	(0.018)	(0.531)	(0.131)	(0.263)	
OC+	0.020	0.032	0.041	0.130*	0.111*	0.030	0.004	0.279*	0.217*	0.077*	0.040	0.089
	(0.687)	(0.336)	(0.208)	(0.002)	(0.011)	(0.423)	(0.909)	(0.000)	(0.000)	(0.028)	(0.373)	
VT	0.038	0.011	0.025	0.140*	0.172*	0.045	0.010	0.328*	0.260*	0.076*	-0.022	0.099
	(0.444)	(0.769)	(0.441)	(0.002)	(0.001)	(0.260)	(0.743)	(0.000)	(0.000)	(0.043)	(0.502)	
TB+	0.041	0.022	0.015	0.121*	0.113*	0.029	0.007	0.256*	0.202*	0.076*	0.048	0.085
	(0.393)	(0.525)	(0.661)	(0.002)	(0.007)	(0.432)	(0.825)	(0.000)	(0.000)	(0.027)	(0.240)	
BL+	0.036	0.025	0.025	0.120*	0.110*	0.021	0.005	0.263*	0.216*	0.071*	0.041	0.085
	(0.440)	(0.454)	(0.420)	(0.003)	(0.009)	(0.555)	(0.859)	(0.000)	(0.000)	(0.038)	(0.281)	
TZMK	0.074*	0.027	0.038	0.012	0.015	0.015	0.007	0.031*	0.013	0.002	-0.007	0.021
	(0.076)	(0.279)	(0.191)	(0.591)	(0.424)	(0.465)	(0.757)	(0.026)	(0.446)	(0.897)	(0.845)	
TZKZ	0.054	0.002	0.017	0.009	0.019	0.001	-0.008	0.034*	0.012	-0.012	0.002	0.012
	(0.206)	(0.927)	(0.558)	(0.676)	(0.304)	(0.952)	(0.726)	(0.036)	(0.459)	(0.523)	(0.953)	
After Transaction Cost												
EW	-0.001	-0.020	-0.021	0.060	0.117*	0.006	0.002	0.236*	0.170*	-0.001	-0.008	0.049
	(0.948)	(0.200)	(0.137)	(0.119)	(0.012)	(0.858)	(0.909)	(0.000)	(0.002)	(0.969)	(0.785)	
MK	0.163*	0.236*	0.255*	0.115*	0.069*	0.111*	0.185*	0.085*	0.091*	0.129*	0.190*	0.148
	(0.006)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.043)	
MK+	-0.008	-0.001	-0.008	0.048	0.050	-0.007	0.015	0.168*	0.124*	0.015	0.027	0.039
	(0.716)	(0.975)	(0.732)	(0.128)	(0.188)	(0.835)	(0.329)	(0.001)	(0.016)	(0.582)	(0.535)	
MV	-0.005	0.018	0.065*	0.064*	0.109*	0.025	0.022	0.246*	0.232*	0.041	0.089	0.083
	(0.878)	(0.582)	(0.090)	(0.063)	(0.014)	(0.553)	(0.392)	(0.000)	(0.000)	(0.233)	(0.149)	
MV+	-0.027	-0.044	-0.048	0.056	0.114*	0.011	-0.012	0.222*	0.161*	-0.002	0.014	0.040
	(0.260)	(0.080)	(0.052)	(0.114)	(0.014)	(0.771)	(0.426)	(0.000)	(0.004)	(0.960)	(0.737)	
OC	0.058	0.156*	0.200*	0.084*	0.044*	0.061*	0.133*	0.080*	0.068*	0.114*	0.167*	0.106
	(0.116)	(0.001)	(0.000)	(0.009)	(0.090)	(0.094)	(0.000)	(0.001)	(0.009)	(0.000)	(0.062)	
OC+	-0.017	0.002	-0.003	0.062*	0.057	0.007	0.007	0.192*	0.134*	0.016	0.037	0.045
	(0.455)	(0.912)	(0.906)	(0.081)	(0.164)	(0.834)	(0.670)	(0.000)	(0.008)	(0.590)	(0.426)	
VT	-0.010	-0.035	-0.035	0.051	0.108*	0.004	-0.002	0.228*	0.168*	-0.005	-0.041	0.039
	(0.555)	(0.082)	(0.052)	(0.159)	(0.016)	(0.918)	(0.873)	(0.000)	(0.002)	(0.867)	(0.136)	
TB+	-0.002	-0.014	-0.027	0.055*	0.061	0.005	0.007	0.167*	0.122*	0.018	0.049	0.040
	(0.919)	(0.448)	(0.274)	(0.088)	(0.135)	(0.893)	(0.637)	(0.001)	(0.014)	(0.522)	(0.257)	
BL+	-0.004	-0.012	-0.023	0.042	0.053	-0.012	0.003	0.171*	0.135*	0.002	0.039	0.036
	(0.820)	(0.545)	(0.280)	(0.177)	(0.164)	(0.701)	(0.813)	(0.000)	(0.009)	(0.943)	(0.316)	
TZMK	0.125*	0.200*	0.194*	0.056*	0.051*	0.075*	0.102*	0.072*	0.073*	0.035*	-0.008	0.089
	(0.000)	(0.000)	(0.000)	(0.020)	(0.020)	(0.005)	(0.000)	(0.000)	(0.001)	(0.072)	(0.832)	
TZKZ	0.088*	0.154*	0.162*	0.047*	0.051*	0.047*	0.070*	0.070*	0.070*	0.009	0.019	0.072
	(0.027)	(0.000)	(0.001)	(0.041)	(0.012)	(0.032)	(0.002)	(0.001)	(0.001)	(0.643)	(0.665)	

Table 9: Sharpe Ratio under Variance Targeting: Last Period Return Maximization

This table reports the Sharpe ratios of the turnover minimization models when τ is calibrated so that the last period return is maximized. Input parameters are estimated from a rolling window of size $T = 120$, and transaction costs are assumed to be 50 basis points for risky assets and 0 for the risk-free asset. The Sharpe ratios statistically significantly higher than that of EW at 10% are marked by *.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	Mean
Before Transaction Cost												
TMKE(τ_b^*)	0.141*	0.207*	0.219*	0.263*	0.257*	0.163	0.180	0.419*	0.337*	0.188	0.138	0.228
TMKE+(τ_b^*)	0.078	0.167	0.188*	0.172*	0.180*	0.149*	0.159*	0.202*	0.173*	0.162	0.193	0.166
TMK0(τ_b^*)	0.006	0.110	0.081	0.270*	0.256*	0.163	0.108	0.431*	0.329*	0.182	0.065	0.182
TMK0+(τ_b^*)	0.109	0.148	0.141	0.181*	0.193*	0.156*	0.143	0.224*	0.196*	0.164	0.159	0.165
TMVE(τ_b^*)	0.099	0.214*	0.216*	0.217*	0.201*	0.175*	0.200*	0.272*	0.175*	0.215*	0.206	0.199
TMVE+(τ_b^*)	0.093	0.194*	0.200*	0.171*	0.139*	0.139	0.164*	0.164*	0.159*	0.169*	0.214*	0.164
TMV0(τ_b^*)	0.128*	0.188	0.162	0.234*	0.189*	0.163*	0.186*	0.231*	0.170	0.210*	0.133	0.181
TMV0+(τ_b^*)	0.114*	0.183*	0.178*	0.158	0.134*	0.138	0.167*	0.162*	0.153*	0.165	0.159	0.156
After Transaction Cost												
TMKE(τ_a^*)	-0.325	-0.594	-0.611	-0.392	-0.204	-0.336	-0.451	-0.294	-0.330	-0.508	-0.164	-0.383
TMKE+(τ_a^*)	-0.011	0.080	0.094	0.074	0.086	0.055	0.065	0.111	0.082	0.061	0.075	0.070
TMK0(τ_a^*)	-0.165	-0.103	-0.232	0.066	0.116	-0.020	-0.040	0.309*	0.092	0.045	-0.009	0.005
TMK0+(τ_a^*)	0.100	0.136	0.128	0.165*	0.188*	0.147*	0.132	0.215*	0.191*	0.152	0.144	0.154
TMVE(τ_a^*)	-0.207	-0.112	-0.236	-0.343	-0.158	-0.159	-0.188	-0.265	-0.285	-0.303	-0.052	-0.210
TMVE+(τ_a^*)	0.019	0.100	0.085	0.071	0.056	0.056	0.074	0.064	0.053	0.067	0.110	0.069
TMV0(τ_a^*)	0.098	0.136	0.096	0.166	0.147	0.129	0.141	0.152	0.103	0.134	0.086	0.126
TMV0+(τ_a^*)	0.108*	0.177	0.171*	0.153	0.128	0.131	0.161*	0.156*	0.149	0.160	0.146	0.149

TMKVT(+) perform comparably to TMKE(+) in most datasets before transaction costs. However, they tend to bear lower transaction costs and outperform TMKE(+) after transaction costs. More importantly, they statistically significantly outperform EW in more datasets. TMVVT(+) also perform slightly better than TMVE(+), but the difference is marginal.

Overall, VT appears to serve better as the reference portfolio than EW, but the evidence is weak. One point to consider in choosing between EW and VT is that the performance of VT depends on the accuracy of variance estimation, whereas EW is immune to estimation errors.

Table 10: Sharpe Ratio under Variance Targeting: Alternative w_0

This table reports the Sharpe ratios of the turnover minimization models incorporating VT. TMKVT(+) and TMVVT(+) respectively denote TMK and TMV with VT as the reference portfolio. Input parameters are estimated from a rolling window of size $T = 120$, and transaction costs are assumed to be 50 basis points for risky assets and 0 for the risk-free asset. The Sharpe ratios statistically significantly higher than that of EW at 10% are marked by *.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	Mean
Before Transaction Cost												
TMKVT(τ_b^*)	0.123	0.178	0.171	0.293*	0.305*	0.182	0.163	0.478*	0.400*	0.240*	0.202	0.249
TMKVT+(τ_b^*)	0.098	0.175*	0.173*	0.185*	0.201*	0.160*	0.167*	0.225*	0.196*	0.171*	0.212*	0.178
TMVVT(τ_b^*)	0.153*	0.200*	0.205*	0.255*	0.197*	0.174*	0.187	0.252*	0.173	0.231*	0.191	0.202
TMVVT+(τ_b^*)	0.120*	0.194*	0.197*	0.170*	0.142*	0.145*	0.172*	0.170*	0.153	0.171*	0.216*	0.168
After Transaction Cost												
TMKVT(τ_a^*)	0.095	0.144	0.133	0.210*	0.238*	0.140	0.146	0.375*	0.308*	0.153	0.193	0.194
TMKVT+(τ_a^*)	0.087	0.159	0.157*	0.167*	0.188*	0.146	0.147	0.210*	0.182*	0.160*	0.202*	0.164
TMVVT(τ_a^*)	0.122*	0.182*	0.180*	0.200*	0.160*	0.138	0.154	0.207*	0.137	0.191	0.182	0.168
TMVVT+(τ_a^*)	0.115*	0.184*	0.185*	0.162*	0.136*	0.136	0.161*	0.158*	0.145	0.163	0.206*	0.159

3.4. Robustness Check

Comprehensive robustness tests further extend the empirical study and verify its findings. Firstly, the effect of the estimation window size is examined by estimating input parameters from $T = 60$ and 240. The empirical analysis is also extended by adding ten additional datasets. These datasets are the same as those in Table 2 but exclude the market portfolio. The turnover minimization strategy is also compared with recently developed strategies.

Finally, the robustness of the findings over time is investigated by evaluating sub-period performance.

3.4.1. Effects of Estimation Window Size

The turnover minimization remains superior across different estimation windows. With the smaller estimation window ($T = 60$), optimal strategies tend to perform poorer, especially after transaction costs. This is because a smaller sample size leads to larger estimation errors and higher turnover. Notwithstanding, the turnover minimization successfully calibrates τ and maintains superior performance distancing itself further from the underlying model.

The comparison of the performances from $T = 120$ and $T = 240$ suggests that increasing the estimation window size does not necessarily reduce estimation errors: many optimal strategies indeed perform poorer when $T = 240$. This signifies that using a parametric approach for the estimation error can be potentially dangerous. For instance, TZMK and TZKZ assume that the estimation errors are reduced when T increases, and assign more weight to the optimal portfolio while reducing the weight on the equal-weight portfolio. However, if this assumption does not hold true, these models will overweight the optimal portfolio and perform poorly. In contrast, the turnover minimization chooses the shrinkage level nonparametrically from the observed data and is robust to the violation of the assumptions.

3.4.2. Alternative Datasets

The results from the alternative datasets that exclude the market portfolio are qualitatively similar to those presented in Section 3.3.2. The ranking of the portfolios remains largely unchanged, and the turnover minimization models continue to perform superior. One difference is that the portfolios involving high turnover perform better in the alternative datasets when subject to transaction costs. This is because the exclusion of the market portfolio reduces the feasible set size and mitigates leverage and turnover as the market portfolio cannot be sold short to buy other assets.

3.4.3. Comparison with More Models

Brandt et al. (2009) employs a no-trade region and find it helps reduce turnover and transaction costs. The turnover minimization models are compared with their underlying models equipped with a no-trade region. Table 11 reports the Sharpe ratios of the underlying models with a no-trade region defined by the hypersphere radius $k = 0.05$ or 0.15.[13] The models that significantly underperform the corresponding turnover minimization model are marked by *.

The results show that the no-trade region improves the performance of the underlying models, but they still underperform the turnover minimization models. The reason for the unsatisfactory performance of the no-trade region strategy can be attributed to the fact that it uses the current portfolio as the reference portfolio. If the first portfolio is poorly chosen, the subsequent portfolios can also perform poorly as they are shrunk towards the current portfolio. As reported earlier, the turnover minimization models that use the current portfolio also suffer the same problem and perform unsatisfactorily.

The turnover minimization models are also compared to the grouping strategy of Branger et al. (2019). Table 3 in Branger et al. (2019) provides the performance of the grouping strategy in multiple datasets, of which three overlap with those in this paper: 10 industry, 30 industry, and 25 Fama-French. Although the sample periods in their study are not exactly the same as those in this paper, the performances of EW and MK are similar. Therefore, comparing the models based on their relative performance to EW is expected to provide a reasonably accurate assessment.[14]

Table 12 compares the Sharpe ratios of the turnover minimization models with those of the grouping strategy. The relative Sharpe ratios (rSR) suggest that the grouping strat-

[13] The radius $k = 0.10$ and 0.20 were also tested and the results were similar.

[14] It would be better to evaluate the turnover minimization during the same periods, but the out-of-sample periods in Branger et al. (2019) are not clear defined.

Table 11: Underlying Models with No-Trade Region

This table reports the Sharpe ratios of the underlying models equipped with the no-trade region strategy (Brandt et al., 2009). Two no-trade region sizes are considered: hypersphere radius $k = 0.05$ and 0.15. The models that significantly underperform their turnover minimization version at 10% are marked by *.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	Mean
Before Transaction Cost												
$k = 0.05$												
MK	0.059	0.124*	0.105*	0.260*	0.280*	0.155*	0.119*	0.446*	0.379	0.199*	0.072	0.200
MK+	0.102	0.146*	0.141*	0.178*	0.192*	0.158	0.147*	0.220*	0.197	0.164	0.178	0.166
MV	0.128	0.186	0.165*	0.230	0.171*	0.165	0.183	0.222*	0.165	0.198*	0.148	0.178
MV+	0.114	0.181	0.180*	0.155	0.133	0.137	0.169	0.159	0.149	0.162	0.197	0.158
$k = 0.15$												
MK	0.046*	0.123*	0.100*	0.255*	0.273*	0.153*	0.115*	0.447*	0.363*	0.200*	0.036*	0.192
MK+	0.083	0.154	0.152	0.175	0.188	0.155	0.144*	0.211	0.183	0.159	0.177	0.162
MV	0.126	0.185	0.165*	0.216*	0.162*	0.162	0.185	0.202*	0.159	0.193*	0.131	0.171
MV+	0.118	0.170*	0.170	0.150	0.132	0.136	0.168	0.158	0.148	0.170	0.166	0.153
TMKE(τ_b^*)	0.131	0.182	0.185	0.298	0.305	0.182	0.162	0.479	0.401	0.237	0.184	0.250
TMKE+(τ_b^*)	0.097	0.159	0.167	0.186	0.201	0.160	0.161	0.225	0.197	0.170	0.186	0.174
TMVE(τ_b^*)	0.150	0.202	0.208	0.254	0.199	0.173	0.188	0.252	0.172	0.232	0.186	0.201
TMVE+(τ_b^*)	0.115	0.195	0.197	0.170	0.143	0.146	0.172	0.170	0.153	0.171	0.213	0.168
After Transaction Cost												
$k = 0.05$												
MK	-0.049*	-0.069*	-0.079*	0.137*	0.194*	0.058*	-0.009*	0.333*	0.253*	0.079*	0.035	0.080
MK+	0.098	0.139	0.137	0.169	0.187	0.152	0.138	0.216*	0.192*	0.155	0.177	0.160
MV	0.103	0.136*	0.095*	0.173	0.140	0.134	0.145	0.157*	0.109	0.145*	0.128	0.133
MV+	0.112	0.178	0.178	0.153	0.129	0.134	0.166	0.156	0.148	0.160	0.196	0.155
$k = 0.15$												
MK	-0.031*	-0.029*	-0.030*	0.182	0.218	0.092*	0.023*	0.376	0.280	0.131	0.022*	0.112
MK+	0.082	0.152	0.152	0.173	0.187	0.153	0.141	0.211	0.182	0.157	0.177	0.160
MV	0.112	0.154	0.124*	0.185	0.144	0.144	0.163	0.164	0.128	0.166	0.126	0.146
MV+	0.117	0.169	0.170	0.149	0.131	0.135	0.167	0.157	0.148	0.169	0.166	0.153
TMKE(τ_a^*)	0.080	0.133	0.122	0.207	0.238	0.139	0.148	0.376	0.307	0.154	0.158	0.187
TMKE+(τ_a^*)	0.087	0.154	0.152	0.169	0.188	0.147	0.146	0.210	0.182	0.159	0.176	0.161
TMVE(τ_a^*)	0.122	0.183	0.180	0.199	0.160	0.141	0.153	0.206	0.134	0.190	0.179	0.168
TMVE+(τ_a^*)	0.109	0.184	0.186	0.162	0.137	0.133	0.162	0.158	0.146	0.164	0.195	0.158

egy slightly outperforms the utility maximization-turnover minimization models (TMKE, TMKE+), but is outperformed by the variance minimization-turnover minimization models (TMVE, TMVE+). Table 3 of Branger et al. (2019) also reports the performances of alternative models, *e.g.*, Li (2015), Chen and Yuan (2016). These models underperform the grouping strategy and are expected to underperform the turnover minimization as well. Although the comparison is rather crude, it still suggests that the turnover minimization performs at least as good as latest models.

Table 12: Comparison with the Grouping Strategy

This table compares the Sharpe ratios of the turnover minimization models with those of the grouping strategy (Branger et al., 2019). The rSR columns report the Sharpe ratios divided by the Sharpe ratio of EW.

	10 Industry		30 Industry		25 FF	
	SR	rSR	SR	rSR	SR	rSR
Branger et al. (2019)						
EW	0.165		0.153		0.161	
Grouping	0.197	1.194	0.196	1.281	0.176	1.093
This paper						
EW	0.158		0.147		0.151	
TMKE(τ_b^*)	0.161	1.019	0.163	1.109	0.313	2.073
TMKE+(τ_b^*)	0.161	1.019	0.166	1.129	0.188	1.245
TMVE(τ_b^*)	0.195	1.234	0.209	1.422	0.258	1.709
TMVE+(τ_b^*)	0.196	1.241	0.198	1.347	0.168	1.113

3.4.4. Sub-period Performance

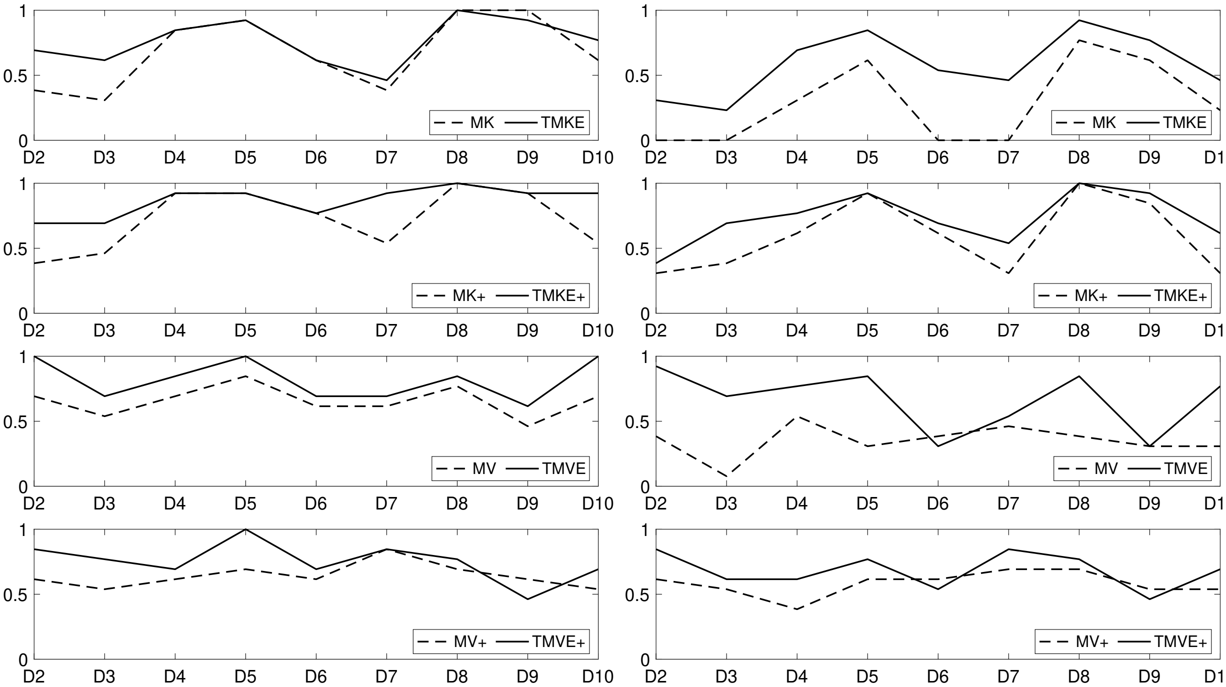
Figure 7 portrays the sub-period performance of the turnover minimization models. Each sub-period is ten-year-long and five-year apart from each other except for the last sub-period, which is shorter due to the size of the whole sample period. The datasets D1 and D11 are omitted as their sample periods are shorter. The y-axis is the percentage of the sub-periods in which a strategy outperforms EW (outperformance ratio). The solid line represents a turnover minimization model, and the dotted line represents its underlying model.

In line with the findings in Section 3.3.2, the turnover minimization models generally outperform EW as well as their underlying models in the sub-periods. They outperform EW in more than half of the sub-periods in most datasets. In particular, TMVE and TMVE+ exhibit robust performance across sub-periods and datasets. Overall, the sub-period analysis confirms the robustness of the findings in Section 3.3.2.

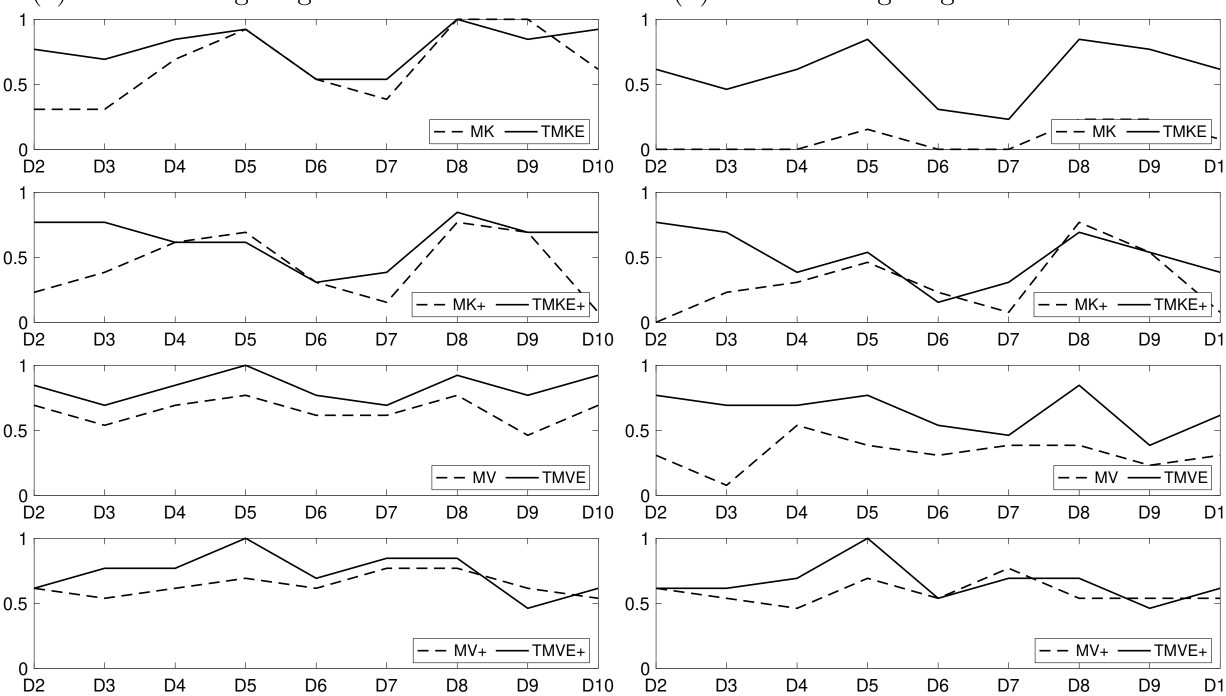
4. Concluding Remarks

This paper develops a new shrinkage portfolio estimator, turnover minimization. It places an additional layer on a conventional portfolio problem, in which the optimal portfolio found in the original problem is shrunk towards a reference portfolio. Unlike existing shrinkage models, the proposed model does not assume the distribution of estimation errors but determines the optimal shrinkage level from observed data. This nonparametric approach makes the model better suited to the real-world data whose distribution is unknown and difficult to estimate. Another advantage of the model is that it can be easily tailored to accommodate a wide range of portfolio problems with various objectives and constraints. This flexibility is particularly beneficial to practitioners who often encounter various constraints imposed by internal policy or regulation. The implementation is straightforward, while the gain from the added layer is substantial.

The proposed model is evaluated against various benchmarks, including well-known classical models and existing shrinkage models. The simulation and empirical studies reveal that minimum-turnover portfolios outperform their underlying portfolios and other benchmarks. They are characterized by low turnover owing to the way they are shrunk and demonstrate superior performance after transaction costs are taken into account. While unconstrained minimum-turnover portfolios perform superior without transaction costs, they are outperformed by their short-sale constrained counterparts when subject to transaction costs. This highlights the importance of the ability to incorporate constraints in shrinkage models. Another important finding is that the equal-weight portfolio serves better as the shrinkage target compared to the current portfolio. The effectiveness of the turnover minimization is



(a) Variance targeting: no transaction costs (b) Variance targeting: transaction costs



(c) Utility maximization: no transaction costs (d) Utility maximization: transaction costs

Figure 7: Sub-Period Performance

This figure demonstrates the sub-period performance of the turnover minimization models when $T = 120$. The y-axis is the percentage of the sub-periods in which EW underperforms. The solid line represents a turnover minimization model, and the dotted line represents its underlying model. There are 13 sub-periods (ten-year-long and five-year apart from each other).

reaffirmed through a comprehensive robustness check.

A. Proofs

Proof of Equation (7)

The Lagrangian of the problem has the form

$$\mathcal{L} = \frac{1}{2}(w - w_0)'\Sigma(w - w_0) - \lambda\left(w'\mu - \frac{\gamma}{2}w'\Sigma w - (1-\tau)U^*\right), \tag{A.1}$$

and the first-order condition is given by

$$\frac{\partial \mathcal{L}}{\partial w} = \Sigma w - \Sigma w_0 - \lambda\mu + \lambda\gamma\Sigma w = 0. \tag{A.2}$$

Solving the first-order condition for w, the optimal portfolio is given by

$$w_{tm} = \frac{1}{1+\lambda\gamma}w_0 + \frac{\lambda}{1+\lambda\gamma}\Sigma^{-1}\mu = \frac{1}{1+\lambda\gamma}w_0 + \frac{\lambda\gamma}{1+\lambda\gamma}w_{mk}. \tag{A.3}$$

Proof of Equation (8)

When the constraint is binding, we have

$$w'\mu - \frac{\gamma}{2}w'\Sigma w = (1-\tau)U^*. \tag{A.4}$$

By substituting w_{tm} in (A.3) for w and solving for λ, we obtain

$$\lambda = \frac{1}{\gamma}\left(\sqrt{\frac{U^* - U_0}{\tau U^*}} - 1\right), \tag{A.5}$$

where $U_0 = w_0'\mu - \frac{\gamma}{2}w_0'\Sigma w_0$ is the utility of the reference portfolio.

Proof of Equation (17)

The optimal a can be obtained by maximizing the expected out-of-sample utility of the portfolio w_{tm} with respect to a:

$$\max_a E[U(a)] = \max_a E\left[w_{tm}'\mu - \frac{\gamma}{2} w_{tm}'\Sigma w_{tm}\right]. \tag{A.6}$$

Using $w_{tm} = aw_0 + (1-a)\hat{w}_{mk}$,

$$\begin{aligned} E[U(a)] = {} & aw_0'\mu + (1-a)E[\hat{w}_{mk}]'\mu - \frac{\gamma}{2}a^2 w_0'\Sigma w_0 \\ & - \frac{\gamma}{2}(1-a)^2 E[\hat{w}_{mk}'\Sigma\hat{w}_{mk}] - \gamma a(1-a)E[\hat{w}_{mk}]'\Sigma w_0. \end{aligned} \tag{A.7}$$

The first-order condition, $\frac{\partial E[U(a)]}{\partial a} = 0$, yields the optimal a of the form

$$a^* = \frac{1}{\gamma} \frac{w_0'\mu - E[\hat{w}_{mk}]'\mu + \gamma E[\hat{w}_{mk}'\Sigma\hat{w}_{mk}] - \gamma E[\hat{w}_{mk}]'\Sigma w_0}{w_0'\Sigma w_0 + E[\hat{w}_{mk}'\Sigma\hat{w}_{mk}] - 2E[\hat{w}_{mk}]'\Sigma w_0}. \tag{A.8}$$

When the asset returns are *i.i.d.* normal, $\hat{\mu}$ and $\hat{\Sigma}$ are independent of each other and

$$\hat{\mu} \sim N\left(\mu, \frac{\Sigma}{K}\right), \quad \hat{\Sigma} \sim \mathcal{W}_N(T-1, \Sigma)\frac{1}{T}. \tag{A.9}$$

It follows that (see Kan and Zhou (2007) and the references therein):

$$E[\hat{w}_{mk}] = \frac{1}{\gamma}\Sigma^{-1}\mu, \tag{A.10}$$

$$E[\hat{w}_{mk}'\Sigma\hat{w}_{mk}] = \frac{c_1}{\gamma^2}\left(\frac{N}{K} + 2\gamma U^*\right). \tag{A.11}$$

where $c_1 = \frac{(T-2)(T-N-2)}{(T-N-1)(T-N-4)}$. Combining (A.8) with (A.10) and (A.11), we have

$$a^* = \frac{\pi_2}{\pi_1 + \pi_2}, \tag{A.12}$$

$$\pi_1 = 2(U^* - U_0), \tag{A.13}$$

$$\pi_2 = 2(c_1 - 1)U^* + \frac{c_1}{\gamma}\frac{N}{T}, \tag{A.14}$$

$$c_1 = \frac{(T-2)(T-N-2)}{(T-N-1)(T-N-4)}, \tag{A.15}$$

and from Equation (A.5)

$$\tau^* = a^{*2}\frac{(U^* - U_0)}{U^*}. \tag{A.16}$$

B. Implementation of the Models

This section describes the implementation details of the models in Table 1. For the full details of each model, the reader is referred to the original papers.

Under variance targeting, a portfolio, w, is adjusted as follows to meet the variance target, σ_T^2:

$$w := w\frac{\sigma_T}{\sqrt{w'\hat{\Sigma}w}}. \tag{B.1}$$

B.1. *Ex-post* Optimal Portfolio, W*

The *ex-post* optimal portfolio maximizes the utility using the true μ and Σ:

$$w^* = \frac{1}{\gamma}\Sigma^{-1}\mu. \tag{B.2}$$

When w^* is adjusted to meet the variance target, the true covariance matrix Σ is used instead of its estimate, $\hat{\Sigma}$. W* is rebalanced back to the optimal portfolio every month.

B.2. Equal-Weight Portfolio, EW

The equal-weight portfolio allocates the wealth equally to the risky assets:

$$w_{ew} = \frac{1}{N} 1_N, \tag{B.3}$$

where 1_N is an N-dimensional vector of ones. EW is rebalanced monthly.

B.3. Markowitz (1952) Mean-Variance Portfolio, MK(+)

The mean-variance portfolio can be obtained from the formulas below.

- Unconstrained Utility Maximization (MK)

$$w_{mk} = \frac{1}{\gamma} \hat{\Sigma}^{-1} \hat{\mu}. \tag{B.4}$$

- Short-sale Constrained Utility Maximization (MK+)

$$\begin{gathered} w_{mk+} = \underset{w}{\operatorname{argmax}}\ w'\hat{\mu} - \frac{\gamma}{2} w' \hat{\Sigma} w \\ \text{subject to } w_i \geq 0, \quad i = 1, \dots, N. \end{gathered} \tag{B.5}$$

B.4. Global Minimum-Variance Portfolio, MV(+)

The global minimum-variance portfolio can be obtained from the formulas below.

- Unconstrained Variance Minimization (MV)

$$w_{mv} = \frac{\hat{\Sigma}^{-1} 1_N}{1'_N \hat{\Sigma}^{-1} 1_N}. \tag{B.6}$$

- Short-sale Constrained Variance Minimization (MV+)

$$\begin{aligned} w_{mv+} = \underset{w}{\operatorname{argmin}}\ & w'\hat{\Sigma}w \\ \text{subject to } & w'1_N = 1 \\ & w_i \geq 0, \quad i = 1, \ldots, N. \end{aligned} \tag{B.7}$$

B.5. Optimal Constrained Portfolio (Kirby and Ostdiek, 2012), OC(+)

Kirby and Ostdiek (2012) show that a mean-variance portfolio constrained to invest only in the risky assets and have the same expected return as the naïve portfolio outperforms the naïve portfolio. The same strategy is considered here, but to be consistent with other strategies, the variance is constrained to meet the target rather than the expected return.

The OC portfolio under variance targeting can be obtained from

$$\begin{aligned} w_{oc} = \underset{w}{\operatorname{argmax}}\ & w'\hat{\mu} \\ \text{subject to } & w'1_N = 1 \\ & w'\hat{\Sigma}w \leq \sigma_T^2. \end{aligned} \tag{B.8}$$

As OC invests only in the risky assets, the variance constraint is imposed during optimization. Other OC optimization problems are similarly defined.

B.6. Volatility Timing (Kirby and Ostdiek, 2012), VT

Kirby and Ostdiek (2012) also introduce portfolio strategies based on volatility. One of their volatility-based strategies that utilizes only the variance is considered. The portfolio from the volatility timing strategy is determined by the formula

$$w_{vt,i} = \frac{\left(1/\hat{\sigma}_i^2\right)^m}{\sum_{i=1}^{N} \left(1/\hat{\sigma}_i^2\right)^m}, \quad i = 1, \ldots, N, \tag{B.9}$$

where m is a tuning parameter which determines the aggressiveness of the weight adjustment in response to changes in the volatility of the asset, and $\hat{\sigma}_i^2$ is the sample variance of the i-th asset return. In the empirical analysis, m is set to 1.

B.7. Black and Litterman (1992) Model, BL+

Black and Litterman (1992) introduce a Bayesian asset allocation model where subjective investor views can be incorporated in the market portfolio. Their framework is adopted by Bessler et al. (2014), who use the naïve portfolio as a proxy for the market portfolio and the sample mean as investor views. The implementation procedure is as follows.

The equilibrium return implied by the equal-weight portfolio is given by

$$\bar{\mu} = \gamma \Sigma \mu_{ew}, \tag{B.10}$$

where γ is the risk aversion parameter. The equilibrium return is assumed to be an unbiased estimate of the true mean:

$$\bar{\mu} = \mu + \eta, \quad \eta \sim N(0, \bar{\Sigma}), \tag{B.11}$$

where $\bar{\Sigma} = \kappa \hat{\Sigma}$ for some constant κ. The investor view is defined as the sample mean, $\hat{\mu}$, and is assumed to be an unbiased estimator of μ:

$$\hat{\mu} = \mu + \epsilon, \quad \epsilon \sim N(0, \Omega). \tag{B.12}$$

Ω represents the uncertainty of the view and is assumed to be of the form $\kappa \text{diag}(\hat{\Sigma})$, where $\text{diag}(\hat{\Sigma})$ is a diagonal matrix derived from $\hat{\Sigma}$. From (B.11) and (B.12), the mean and covariance matrix of the asset returns can be estimated via the generalized least squares and are given by:

$$\tilde{\mu} = \bar{\mu} + \bar{\Sigma}(\bar{\Sigma} + \Omega)^{-1}(\hat{\mu} - \bar{\mu}), \tag{B.13}$$

$$\tilde{\Sigma} = \hat{\Sigma} + \bar{\Sigma} - \bar{\Sigma}(\bar{\Sigma} + \Omega)^{-1}\bar{\Sigma}. \tag{B.14}$$

γ and κ are respectively set to 3 and 0.1. The Black-Litterman optimal portfolio is obtained by solving the constrained mean-variance problem in (B.5) with the mean and covariance estimates defined above.

B.8. Treynor and Black (1973) Model, TB+

Treynor and Black (1973) develop an active portfolio strategy for an optimal allocation between active assets (assets with abnormal excess returns) and the market portfolio. We adopt their model and use the naïve portfolio as a proxy for the market portfolio.

The "active assets" are first identified by regressing asset returns on the naïve portfolio returns:

$$r_{it} = \alpha_i + \beta_i r_{ew,t} + e_{it}, \quad t = 1, \ldots, T, \tag{B.15}$$

where r_{it} and $r_{ew,t}$ are respectively the returns of asset i and the naïve portfolio at time t in excess of the risk-free rate. The assets with a significant α_i at 5% are identified as active assets. The regression is carried out and active assets are identified every month.

The optimal portfolio from the active assets and the equal-weight portfolio can be obtained by solving the usual mean-variance problem with the equal-weight portfolio added in the asset pool. Treynor and Black (1973) provide a closed-form solution for an unconstrained problem, but it needs to be solved numerically when subject to the short-sale constraint:

$$\begin{aligned} w_{tb+} = \underset{w}{\operatorname{argmax}}\ & w'\hat{\mu} - \frac{\gamma}{2} w'\hat{\Sigma} w \\ \text{subject to } & w_i + \frac{1}{M} w_{M+1} \geq 0, \quad i = 1, \ldots, M, \end{aligned} \tag{B.16}$$

where M denotes the number of active assets, and $\hat{\mu} \in \mathbb{R}^{M+1}$ and $\hat{\Sigma} \in \mathbb{R}^{(M+1)\times(M+1)}$ are the input parameter estimates of the M active assets and the equal-weight portfolio (($M+1$)-th asset). Note that since the ($M+1$)-th asset is the equal-weight portfolio, the short-sale constraint has a different form.

B.9. Tu and Zhou (2011) Models, TZMK and TZKZ

Tu and Zhou (2011) develop a shrinkage portfolio model that combines an optimal portfolio with the naïve portfolio. They specifically consider an optimal mix of the naïve portfolio with the Markowitz (1952) rule, Jorion (1986) rule, Kan and Zhou (2007) rule, and MacKinlay and Pástor (2000) rule. In this paper, the Markowitz (1952) rule and Kan and Zhou (2007) rule are considered.

- MK+EW (TZMK)

  The Tu and Zhou (2011) portfolio that combines the equal-weight portfolio with the Markowitz portfolio is given by

$$w_{tzmk} = \hat{a} w_{mk} + (1-\hat{a}) w_{ew}, \tag{B.17}$$

  where

$$\hat{a} = \frac{\pi_2}{\pi_1+\pi_2}, \tag{B.18}$$

$$\pi_1 = w_{ew}'\hat{\Sigma} w_{ew} - \frac{2}{\gamma} w_{ew}'\hat{\mu} + \frac{1}{\gamma^2}\tilde{\theta}^2, \tag{B.19}$$

$$\pi_2 = \frac{1}{\gamma^2}(c_1-1)\tilde{\theta}^2 + \frac{c_1}{\gamma^2}\frac{N}{T} \tag{B.20}$$

$$c_1 = \frac{(T-2)(T-N-2)}{(T-N-1)(T-N-4)}, \tag{B.21}$$

  where $\tilde{\theta}^2$ is an estimate of $\theta = \mu\Sigma^{-1}\mu$. Tu and Zhou (2011) suggest to use the estimator of Kan and Zhou (2007):

$$\tilde{\theta}^2 = \frac{(T-N-2)\hat{\theta}^2 - N}{T} + \frac{2(\hat{\theta}^2)^{N/2}(1+\hat{\theta}^2)^{-(T-2)/2}}{TB_{\hat{\theta}^2/(1+\hat{\theta}^2)}(N/2,(T-N)/2)}, \tag{B.22}$$

where $\hat{\theta}^2 = \hat{\mu}'\hat{\Sigma}^-1\hat{\mu}$, and $B_x(a,b) = \int_0^x y^{a-1}(1-y)^{b-1}dy$ is an incomplete beta function.

- KZ+EW (TZKZ)

The Tu and Zhou (2011) portfolio that combines the equal-weight portfolio with the Kan and Zhou (2007) three-fund rule is given by

$$w_{tzkz} = \hat{a}w_{kz} + (1-\hat{a})w_{ew}, \tag{B.23}$$

where w_{kz} is the Kan-Zhou rule defined below, and

$$\hat{a} = \frac{\pi_1 - \pi_{13}}{\pi_1 - 2\pi_{13} + \pi_3}, \tag{B.24}$$

$$\begin{aligned}\pi_{13} = {} & \frac{1}{\gamma^2}\tilde{\theta}^2 - \frac{1}{\gamma}w_{ew}'\hat{\mu} + \frac{1}{\gamma c_1}\left(\hat{\eta}w_{ew}'\hat{\mu} + (1-\hat{\eta})\hat{\mu}_g w_{ew}' 1_N\right) \\ & - \frac{1}{\gamma^2 c_1}\left(\hat{\eta}\hat{\mu}'\tilde{\Sigma}^{-1}\hat{\mu} + (1-\hat{\eta})\hat{\mu}_g\hat{\mu}'\tilde{\Sigma}^{-1}1_N\right),\end{aligned} \tag{B.25}$$

$$\pi_3 = \frac{1}{\gamma^2}\tilde{\theta}^2 - \frac{1}{\gamma^2 c_1}\left(\tilde{\theta}^2 - \frac{N}{T}\hat{\eta}\right). \tag{B.26}$$

π_1 is as defined above, and $\hat{\eta}$ and $\hat{\mu}_g$ are as given below.

- Kan and Zhou (2007) Three-Fund Rule

The three-fund rule of Kan and Zhou (2007) is given by

$$w_{kz} = \hat{a}w_{mk} + \hat{b}\tilde{w}_{mv}, \tag{B.27}$$

where $\tilde{w}_{mv} = \frac{1}{\gamma}\hat{\Sigma}^{-1}1_N$, and

$$\hat{a} = \frac{1}{c_1}\hat{\eta}, \quad \hat{b} = \frac{1}{c_1}(1-\hat{\eta}), \quad \hat{\eta} = \frac{\tilde{\phi}^2}{\tilde{\phi}^2 + N/T}. \tag{B.28}$$

$\tilde{\phi}^2$ is given by

$$\tilde{\phi}^2 = \frac{(T-N-1)\hat{\phi}^2 - (N-1)}{T} + \frac{2(\tilde{\phi}^2)^{(N-1)/2}(1+\tilde{\phi}^2)^{-(T-2)/2}}{TB_{\tilde{\phi}^2/(1+\tilde{\phi}^2)}((N-1)/2,(T-N+1)/2)}, \tag{B.29}$$

where

$$\hat{\phi}^2 = (\hat{\mu} - \hat{\mu}_g 1_N)' \, \hat{\Sigma}^{-1} \, (\hat{\mu} - \hat{\mu}_g 1_N), \quad \hat{\mu}_g = \frac{\hat{\mu}'\hat{\Sigma}^{-1}1_N}{1_N'\hat{\Sigma}^{-1}1_N}. \tag{B.30}$$

References

Avramov, D., Zhou, G., 2010. Bayesian portfolio analysis. The Annual Review of Financial Economics 2, 25–47.

Balduzzi, P., Lynch, A.W., 1999. Transaction costs and predictability: Some utility cost calculations. Journal of Financial Economics 52, 47–78.

Bessler, W., Opfer, H., Wolff, D., 2014. Multi-asset portfolio optimization and out-of-sample performance: an evaluation of Black–Litterman, mean-variance, and naïve diversification approaches. The European Journal of Finance , 1–30.

Black, F., Litterman, R., 1992. Global portfolio optimization. Financial Analysts Journal 48, 28–43.

Brandt, M.W., Santa-Clara, P., Valkanov, R., 2009. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. The Review of Financial Studies 22, 3411–3447.

Branger, N., Lučivjanská, K., Weissensteiner, A., 2019. Optimal granularity for portfolio choice. Journal of Empirical Finance 50, 125–146.

Brown, S.J., 1976. Optimal portfolio choice under uncertainty: A Bayesian approach. Ph.D. thesis. University of Chicago, Graduate School of Business.

Brown, S.J., 1978. The portfolio choice problem: Comparison of certainty equivalence and optimal Bayes portfolios. Communications in Statistics-Simulation and Computation 7, 321–334.

Ceria, S., Stubbs, R.A., 2016. Incorporating estimation errors into portfolio selection: Robust portfolio construction, in: Asset Management. Springer, pp. 270–294.

Chen, J., Yuan, M., 2016. Efficient portfolio selection in a large market. Journal of Financial Econometrics 14, 496–524.

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

DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? Review of Financial Studies 22, 1915–1953.

DeMiguel, V., Martín-Utrera, A., Nogales, F.J., 2015. Parameter uncertainty in multiperiod portfolio optimization with transaction costs. Journal of Financial and Quantitative Analysis 50, 1443–1471.

Fabozzi, F.J., Kolm, P.N., Pachamanova, D.A., Focardi, S.M., 2007. Robust portfolio optimization and management. John Wiley & Sons.

Gârleanu, N., Pedersen, L.H., 2013. Dynamic trading with predictable returns and transaction costs. The Journal of Finance 68, 2309–2340.

Goldfarb, D., Iyengar, G., 2003. Robust portfolio selection problems. Mathematics of Operations Research 28, 1–38.

Han, C., 2019. How much should portfolios shrink? Financial Management .

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

Jorion, P., 1986. Bayes-Stein estimation for portfolio analysis. The Journal of Financial and Quantitative Analysis 21, 279–292.

Kan, R., Wang, X., 2016. On the economic value of alphas. Working Paper .

Kan, R., Wang, X., Zhou, G., 2016. On the value of portfolio optimization in the presence of estimation risk: The case with and without risk-free asset. Working Paper .

Kan, R., Zhou, G., 2007. Optimal portfolio choice with parameter uncertainty. Journal of Financial and Quantitative Analysis 42, 621–656.

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

Klein, R.W., Bawa, V.S., 1976. The effect of estimation risk on optimal portfolio choice. Journal of Financial Economics 3, 215–231.

Ledoit, O., Wolf, M., 2004. Honey, I shrunk the sample covariance matrix. The Journal of Portfolio Management 30, 110–119.

Ledoit, O., Wolf, M., 2008. Robust performance hypothesis testing with the sharpe ratio. Journal of Empirical Finance 15, 850–859.

Li, J., 2015. Sparse and stable portfolio selection with parameter uncertainty. Journal of Business & Economic Statistics 33, 381–392.

MacKinlay, A.C., Pástor, Ľ., 2000. Asset pricing models: Implications for expected returns and portfolio selection. Review of Financial Studies 13, 883–916.

Markowitz, H., 1952. Portfolio selection. The Journal of Finance 7, 77–91.

Michaud, R.O., 1989. The markowitz optimization enigma: Is optimized optimal? Financial Analysts Journal 45, 31–42.

Olivares-Nadal, A.V., DeMiguel, V., 2018. A robust perspective on transaction costs in portfolio optimization. Operations Research 66, 733–739.

Pástor, L., 2000. Portfolio selection and asset pricing models. The Journal of Finance 55, 179–223.

Pástor, L., Stambaugh, R.F., 2000. Comparing asset pricing models: An investment perspective. Journal of Financial Economics 56, 335–381.

Pflug, G.C., Pichler, A., Wozabal, D., 2012. The 1/N investment strategy is optimal under high model ambiguity. Journal of Banking and Finance 36, 410–417.

Treynor, J.L., Black, F., 1973. How to use security analysis to improve portfolio selection. The Journal of Business 46, 66–86.

Tu, J., Zhou, G., 2011. Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. Journal of Financial Economics 99, 204–215.