ON THE OPTIMAL COMBINATION OF NAIVE AND MEAN-VARIANCE PORTFOLIO STRATEGIES

Nathan Lassance, Rodolphe Vanderveken, Frédéric Vrins

LIDAM Discussion Paper LFIN
2022 / 06

UCLouvain

LFIN | LIDAM
Louvain Institute of Data Analysis and Modeling in economics and statistics

LFIN

Voie du Roman Pays 34, L1.03.01 B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: lidam-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/lfin/publications.html

On the optimal combination of naive and mean-variance portfolio strategies*

Nathan Lassance[1] Rodolphe Vanderveken*,[1] Frédéric Vrins[1,2]

July 13, 2022

Abstract

A disheartening fact in portfolio choice is that the naive equally weighted portfolio often outperforms the estimated optimal mean-variance portfolio out of sample. In an influential paper, Tu and Zhou (2011) reaffirm the value of portfolio theory by combining the two portfolios to optimize out-of-sample performance. They achieve this under a seemingly natural convexity constraint: the two combination coefficients must sum to one. We show that this constraint is unnecessary in theory and has several undesirable consequences relative to the unconstrained portfolio combination we derive. In particular, it leads to an overinvestment in the sample mean-variance portfolio, and a worse performance than the risk-free asset for sufficiently risk-averse investors. However, although wrong in theory, we demonstrate that the convexity constraint acts as a bound constraint on combination coefficients and thus can help improve performance when they are estimated. Our empirical analysis shows that the Tu and Zhou rule performs well for investors with small risk aversion, but quickly deteriorates as risk aversion increases. In contrast, our portfolio rules perform consistently well. Finally, we show theoretically and empirically that there are larger out-of-sample diversification gains from combining the sample mean-variance portfolio with the equally weighted portfolio instead of the minimum-variance portfolio.

Keywords: portfolio optimization, parameter uncertainty, estimation risk, equally weighted portfolio, portfolio constraints.
JEL Classification: G11, G12.

*Corresponding author. Affiliations: (1) LIDAM/LFIN, UCLouvain, Belgium, (2) Decision science department, HEC Montréal, Canada. Email addresses: nathan.lassance@uclouvain.be, rodolphe.vanderveken@uclouvain.be, frederic.vrins@uclouvain.be. This work is supported by the Fonds de la Recherche Scientifique (F.R.S.-FNRS) under Grant Number J.0115.22 (N. Lassance) and T.0221.22 (R. Vanderveken and F. Vrins), and by the Belgian Federal Science Policy Office under Grant Number ARC 18-23/089 (F. Vrins).

1 Introduction

The portfolio theory of Markowitz (1952) is often criticized because estimated optimal mean-variance portfolios typically perform very poorly out of sample. This is because mean-variance portfolios are highly sensitive to estimation errors in the sample mean and sample covariance matrix of asset returns (Best and Grauer, 1991; DeMiguel, Garlappi, Nogales, and Uppal, 2009; Barroso and Saxena, 2021). In a thought-provoking paper, DeMiguel, Garlappi, and Uppal (2009) show that this problem is severe because *none* of the 13 robust extensions to the sample mean-variance (SMV) model they consider is able to consistently outperform the naive equally weighted (EW) portfolio. As Tu and Zhou (2011, p.205) put it, "*these findings raise a serious doubt on the usefulness of the investment theory.*" However, in their influential paper, Tu and Zhou (2011) reinstill hope in the value of portfolio optimization because they show that by cleverly *combining* the SMV portfolio with the naive EW portfolio, one can outperform both in most scenarios. This is made possible by balancing between the two strategies according to the degree of estimation error, measured by the ratio N/T where N is the number of assets and T is the sample size. In this paper, we raise some concerns about Tu and Zhou's methodology and explain how to alleviate them to further improve out-of-sample performance for mean-variance investors.

Tu and Zhou (2011) exploit the analytical framework introduced by Kan and Zhou (2007) and combine the SMV portfolio with the EW portfolio to optimize the expected out-of-sample utility. The way they achieve this is by finding the two optimal combination coefficients attached to each portfolio. Tu and Zhou optimize them subject to a *convexity constraint*: the two combination coefficients must sum to one. This constraint is sensible when shrinking covariance matrices as in Ledoit and Wolf (2004) or when combining portfolios that are fully invested in risky assets as in Kan, Wang, and Zhou (2021). However, Tu and Zhou (2011) consider the SMV portfolio that invests in the risk-free asset as well, and therefore we show that the convexity constraint in unwarranted.

Our first contribution is to derive the *optimal* combination of the SMV and EW portfolios that relaxes the convexity constraint and to demonstrate that relative to the *optimal strategy*, the *constrained strategy* of Tu and Zhou (2011) presents five main weaknesses. Moreover,

these weaknesses become more severe as estimation error increases, that is, the constrained strategy fails most precisely when it is most needed.

The first weakness is that the constrained strategy does not provide the optimal exposure to the three funds underlying the combination, that is, the sample tangent portfolio, the EW portfolio, and the risk-free asset. In particular, we prove the striking result that *the constrained strategy always overinvests in the sample tangent portfolio*, which is a portfolio known to be problematic in practice. It also overinvests in the EW portfolio for most risk aversion, and underinvests in the risk-free asset under mild conditions. These differences are often substantial. For example, consider the case in which the mean and covariance matrix of excess returns are calibrated to a dataset of $N = 25$ portfolios sorted on size and book-to-market (25BSTM), the sample size is $T = 120$ months, and the risk-aversion coefficient is five. Then, the constrained strategy allocates 23% to the tangent portfolio, 71% to the EW portfolio, and 6% to the risk-free asset. In contrast, the optimal strategy invests less in the tangent and EW portfolios (16% and 30%) and much more in the risk-free asset (54%).

Second, the constrained strategy does not provide the theoretically optimal expected out-of-sample utility attained by the optimal strategy, except for one specific value of the risk-aversion coefficient. Moreover, this utility loss can be large. For the case of the 25SBTM dataset and a sample size of $T = 120$ months, the annualized utility loss can go up to 6% approximately as the risk-aversion coefficient varies between 1 and 20.

Third, under mild conditions the constrained strategy *underperforms the risk-free asset* once the risk-aversion coefficient is above a given threshold. This result can be explained because due to the convexity constraint, the constrained strategy can only invest in the risk-free asset via the SMV portfolio, which is subject to high estimation risk. The threshold is not particularly large; it is equal to 5.4 for the 25SBTM dataset and $T = 120$ months. In contrast, the optimal strategy *always* outperforms the risk-free asset.

Fourth, the constrained strategy does not fully profit from adding the EW portfolio in the combination because we show that, beyond a risk-aversion coefficient threshold, it underperforms the optimal two-fund portfolio of Kan and Zhou (2007) that only invests in the sample tangent portfolio and the risk-free asset. The threshold is generally small; it is equal to 3.8 for the 25SBTM dataset. This result means that many investors are better off

relying on the optimal two-fund rule that does not invest in the EW portfolio rather than adding the EW portfolio via the convexity constraint.

Fifth, the optimal strategy has another desirable property: it delivers *less extreme weights* on the risky assets. Specifically, we show that under mild conditions the ℓ_2-norm of portfolio weights of the optimal strategy is smaller than that of the constrained strategy if the risk-aversion coefficient belongs to a specific wide interval. This effect is particularly prominent when the risk-aversion coefficient is large. For the 25SBTM dataset, $T = 120$ months, and a risk aversion of 10, the ℓ_2-norm of the constrained strategy is 0.72 while that of the optimal strategy is only 0.26. This smaller ℓ_2-norm has two main benefits. On the one hand, DeMiguel et al. (2009) show that constraining portfolio norms helps improve performance. On the other hand, we find empirically that the optimal strategy delivers lower turnover and transaction costs than the constrained strategy due to its less extreme and thus less variable weights.

These five weaknesses of the constrained strategy relative to the optimal strategy all result from the convexity constraint. Although intuitive at first sight, it is actually unnecessary and renders the combination theoretically suboptimal by removing a degree of freedom.

Our second contribution is to treat the impact of estimation errors in combination coefficients, which is important because Kan and Wang (2021) and Lassance, Martín-Utrera, and Simaan (2022) show that such errors have a non-negligible impact on out-of-sample performance. In particular, we show analytically that, although wrong in theory, *imposing the convexity constraint can help*, as in Jagannathan and Ma (2003). Indeed, this constraint acts as a bound constraint because both constrained coefficients are between zero and one. In contrast, the optimal coefficient on the EW portfolio is unbounded and thus much more sensitive to estimation errors in mean returns. As a result, when the optimal strategy only delivers a small theoretical gain relative to the constrained strategy, the latter may actually outperform when combination coefficients are estimated. We formalize this insight by deriving the interval of risk-aversion coefficients for which the constrained strategy outperforms the optimal strategy. The length of this interval is not negligible. For the 25SBTM dataset and a sample size of $T = 120$ months, it is equal to $[1.90 \pm 1.65]$. We then exploit this result to propose a *mixed strategy* that invests in the constrained strategy when the risk aversion belongs to the interval, and in the optimal strategy otherwise.

The results on the usefulness of the convexity constraint are far-reaching. In Appendix IA.2.1, we derive a similar mixed strategy for the three-fund combination considered by Kan and Zhou (2007) that combines the SMV portfolio with the sample global-minimum-variance (SGMV) portfolio. In Appendix IA.2.2, we show that the convexity constraint implies that the optimal combination of the SMV and SGMV portfolios without a risk-free asset in Kan, Wang, and Zhou (2021) actually outperforms the three-fund rule of Kan and Zhou (2007) for some risk-aversion levels; that is, *investing in the risk-free asset can hurt performance*.

Our third contribution is to test empirically the out-of-sample performance of the two main portfolio strategies we introduce: the optimal three-fund rule and the mixed strategy that invests in either the optimal or the constrained three-fund rule. We compare the two strategies to the constrained three-fund rule of Tu and Zhou (2011), the combination rules of Kan and Zhou (2007) and Kan, Wang, and Zhou (2021), and the EW and SGMV portfolios. We compare the strategies in terms of out-of-sample utility net of transaction costs, across four datasets of characteristic- and industry-sorted portfolios, and for a range of risk aversion similar to that in DeMiguel, Garlappi, and Uppal (2009) and Martellini and Ziemann (2010).

The main comparison concerns the three-fund rules that invest in the EW portfolio. The empirical observations match our theoretical predictions detailed above. Specifically, the constrained strategy of Tu and Zhou (2011) outperforms the optimal strategy for small degrees of risk aversion but quickly deteriorates as risk aversion increases, in which case it often delivers negative out-of-sample utilities and underperforms the two-fund rule of Kan and Zhou (2007) that does not invest in the EW portfolio. In contrast, the optimal strategy performs consistently well and always delivers positive utilities that are larger than those of the two-fund rule. We also confirm that, due to its less extreme weights, the optimal strategy yields less turnover and transaction costs than the constrained strategy. Finally, the mixed strategy achieves its objective by relying mostly on the constrained strategy when risk aversion is small, and mostly on the optimal strategy when risk aversion is high. Therefore, the mixed strategy is safer than relying exclusively on the optimal or constrained strategy.

The comparison to the other benchmarks delivers two more insights. First, even though the SGMV portfolio largely outperforms the EW portfolio in our datasets, we find that combining the SMV portfolio with the EW portfolio as we consider can be preferable to

combining with the SGMV portfolio as in Kan and Zhou (2007). In particular, combining with EW delivers better performance when the sample size is rather small ($T = 120$ months), and only a marginally worse performance when it increases to $T = 240$. We explain this puzzling result theoretically and empirically from a smaller correlation between the out-of-sample return of the SMV portfolio and the EW portfolio, and thus larger out-of-sample diversification gains.[1] Second, combining the SMV portfolio with the EW portfolio is useful only when EW is not, by chance, close to optimal. Indeed, for the three datasets of characteristic-sorted portfolios we find that EW is largely inefficient, and consequently the combination with SMV largely outperforms EW alone. However, the EW portfolio is much closer to being efficient for industry-sorted portfolios (Kirby and Ostdiek, 2012). Therefore, we find as in Tu and Zhou (2011) that combining the SMV and EW portfolios cannot outperform the EW portfolio alone for industry-sorted portfolios because the small potential gain cannot compensate for the increased estimation risk.

Our work is related to three strands of the portfolio optimization literature. First, we contribute to the literature on portfolio combinations. While many papers advocate using the SGMV portfolio over any other SMV portfolio, Kan and Zhou (2007) show in their pioneering work that it is possible to outperform both out of sample by optimally *combining* them. Following Kan and Zhou (2007), many scholars propose combination rules that optimize out-of-sample performance; see, for instance, Frahm and Memmel (2010), Tu and Zhou (2011), DeMiguel, Martín-Utrera, and Nogales (2013, 2015), Branger, Lučivjanská, and Weissensteiner (2019), Kan, Wang, and Zhou (2021), Kan and Wang (2021), and Lassance, Martín-Utrera, and Simaan (2022). In this paper, we follow the work of Tu and Zhou (2011) who combine the SMV portfolio with the EW portfolio to optimize expected out-of-sample utility under a convexity constraint on the combination coefficients. We derive the *unconstrained* combination and demonstrate several undesirable consequences of this constraint relative to the optimal solution. We also bring new insights to the portfolio combination literature by showing analytically how and when the convexity constraint actually helps *improve*

[1] In Appendix IA.2.4, we also derive the optimal four-fund portfolio that invests in the SMV portfolio and both the SGMV and EW portfolios. However, this optimal four-fund rule is generally outperformed by the optimal three-fund rule that does not invest in SGMV, which can be explained because the theoretical gain is small and thus is offset by the estimation risk coming from the additional combination coefficient.

out-of-sample utility because of reduced estimation error in combination coefficients.

Second, our work is related to other papers that study the value of the naive EW portfolio. Bloomfield, Leftwich, and Long (1977) show in an early study that SMV portfolios do not outperform the EW portfolio. DeMiguel, Garlappi, and Uppal (2009) find that none of the SMV portfolio and 13 robust extensions can outperform the EW portfolio consistently. Pflug, Pichler, and Wozabal (2012) and Yan and Zhang (2017) explain that the EW portfolio is not so naive, because it is nearly optimal under high model ambiguity and in the absence of mispricing, respectively. Tu and Zhou (2011) propose a method to nearly consistently outperform the EW portfolio by combining it with the SMV portfolio. Similarly, Frahm and Memmel (2010), Simaan and Simaan (2019), and Bodnar, Okhrin, and Parolya (2021) combine the EW portfolio with mean-variance or minimum-variance portfolios. In this paper, we show how to combine the SMV portfolio with the EW portfolio optimally when investing in the risk-free asset is allowed. Our results confirm the value of the EW portfolio because we show theoretically and empirically that our optimal three-fund rule consistently outperforms the optimal two-fund rule of Kan and Zhou (2007) that does not invest in the EW portfolio. We also demonstrate the larger out-of-sample diversification gains one can obtain from combining the SMV portfolio with EW rather than SGMV.

Third, we contribute to a large literature on the role of constraints in portfolio selection. Frost and Savarino (1988), Jagannathan and Ma (2003), and Behr, Guettler, and Miebs (2013) show that imposing lower and upper bound constraints on portfolio weights helps improve performance even when these constraints are wrong. Levy and Levy (2014) introduce a class of differential variance-based constraints that significantly outperforms homogeneous bound constraints. DeMiguel et al. (2009), Yen (2016), Ao, Li, and Zheng (2019), and Zhao, Ledoit, and Jiang (2021) show the large benefits of norm constraints. Ban, El Karoui, and Lim (2018) introduce performance-based constraints inspired from machine learning. A key differentiating feature between our work and the aforementioned papers is that we show the benefits of constraints for combining *sample portfolios*, instead of assets. In particular, we show theoretically and empirically in several settings that constraining combination coefficients on sample portfolios to sum to one, although wrong in theory, acts as a bound constraint and thus helps improve out-of-sample utility for some degrees of risk aversion.

2 The portfolio combination problem

In Section 2.1, we introduce the setting and the assumptions we use throughout the paper. In Section 2.2, we recall the notion of expected out-of-sample utility. In Section 2.3, we review the methodology of Tu and Zhou (2011) to combine the SMV and EW portfolios.

2.1 Setting and assumptions

We consider the classical portfolio choice problem in which an investor allocates her wealth among $N \geq 1$ risky assets and a risk-free asset. We denote $\boldsymbol{r}_t$ the $N \times 1$ vector of asset excess returns at time t, which has mean $\boldsymbol{\mu}$ and positive-definite covariance matrix $\boldsymbol{\Sigma}$. We assume that asset returns are normally distributed, which is a common assumption made for tractability (Kan and Zhou, 2007; Tu and Zhou, 2011; Ao, Li, and Zheng, 2019).

At time T, the investor collects a sample of T return observations, $(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_T)$, which we assume independent and identically distributed. The investor then chooses her optimal portfolio $\boldsymbol{w} = (w_1, ..., w_N)'$, the weight on the risk-free asset being $w_0 = 1 - \mathbf{1}'\boldsymbol{w}$ with $\mathbf{1}$ a conformable vector of ones. We assume throughout the paper that $N + 4 < T < \infty$.

As in Markowitz (1952) and Tu and Zhou (2011), we assume that the investor has mean-variance preferences and maximizes her utility. If the moments of asset returns are known, this amounts to maximize

$$U(\boldsymbol{w}) = \boldsymbol{w}'\boldsymbol{\mu} - \frac{\gamma}{2}\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}, \tag{1}$$

where $\gamma > 0$ is the investor's risk-aversion coefficient. Note that utility is in excess of the risk-free rate, and thus the risk-free asset has a utility of zero. It is well known that the optimal *mean-variance portfolio* maximizing (1) is

$$\boldsymbol{w}^\star = \frac{1}{\gamma}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}, \tag{2}$$

which provides the maximum utility

$$U^\star = U(\boldsymbol{w}^\star) = \frac{\theta^2}{2\gamma}, \tag{3}$$

where $\theta = \sqrt{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}$ is the maximum Sharpe ratio. For further reference, note that the mean-variance portfolio can be decomposed as a linear combination of the fully invested tangent portfolio $\boldsymbol{w}_{tan} = \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}/(\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})$ and the risk-free asset:

$$\boldsymbol{w}^\star = \frac{\gamma_{tan}}{\gamma}\boldsymbol{w}_{tan} \quad \text{and} \quad w_0^\star = 1 - \frac{\gamma_{tan}}{\gamma}, \tag{4}$$

where $w_0^\star = 1 - \mathbf{1}'\boldsymbol{w}^\star$ is the weight on the risk-free asset in the mean-variance strategy and

$$\gamma_{tan} = \mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \tag{5}$$

is the value of the risk-aversion coefficient γ for which $\boldsymbol{w}_{tan}$ maximizes $U(\boldsymbol{w})$.

2.2 Parameter uncertainty and expected out-of-sample utility

The mean-variance portfolio $\boldsymbol{w}^\star$ in (2) and its maximum utility $U^\star$ in (3) are unfeasible in practice because the moments of asset returns, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, are unknown and must be estimated. As in Kan and Zhou (2007), Tu and Zhou (2011), and Kan, Wang, and Zhou (2021), among others, we rely on sample estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ for tractability, which are given by

$$\hat{\boldsymbol{\mu}} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{r}_t \quad \text{and} \quad \widehat{\boldsymbol{\Sigma}} = \frac{1}{T-N-2}\sum_{t=1}^{T}(\boldsymbol{r}_t - \hat{\boldsymbol{\mu}})(\boldsymbol{r}_t - \hat{\boldsymbol{\mu}})'. \tag{6}$$

The coefficient $1/(T-N-2)$ ensures that the inverse sample covariance matrix is unbiased, $\mathbb{E}[\widehat{\boldsymbol{\Sigma}}^{-1}] = \boldsymbol{\Sigma}^{-1}$. Employing the sample covariance matrix is a conversative approach given that more accurate estimators exist, such as shrinkage estimators (Ledoit and Wolf, 2004, 2017). We also consider shrinkage estimators of $\boldsymbol{\Sigma}$ in our empirical analysis of Section 5, and our theoretical results in Sections 3 and 4 can easily be applied to the case where $\boldsymbol{\Sigma}$ is known, as in Garlappi, Uppal, and Wang (2007), by letting $c = 1$ in (14).

The *sample mean-variance (SMV)* portfolio, which exploits $\hat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$, is defined as

$$\hat{\boldsymbol{w}}^\star = \frac{1}{\gamma}\widehat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}. \tag{7}$$

Due to estimation errors, the SMV portfolio does not perform best in practice (Jobson

and Korkie, 1980; Michaud, 1989; DeMiguel, Garlappi, and Uppal, 2009). To quantify the impact of estimation errors on performance, we follow Kan and Zhou (2007) and measure the performance of an estimated portfolio $\hat{\boldsymbol{w}}$ via its *expected out-of-sample utility* (EU):

$$EU(\hat{\boldsymbol{w}}) = \mathbb{E}[U(\hat{\boldsymbol{w}})] = \mathbb{E}[\hat{\boldsymbol{w}}'\boldsymbol{\mu}] - \frac{\gamma}{2}\mathbb{E}[\hat{\boldsymbol{w}}'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}]. \tag{8}$$

In this expression, the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ come from (1) and are the population moments of next-period excess returns, while $\hat{\boldsymbol{w}}$ is a random vector estimated from historical samples.

Kan and Zhou (2007), Tu and Zhou (2011), and Kan, Wang, and Zhou (2021), among others, show how to exploit the EU criterion to *optimally combine* portfolios and improve out-of-sample performance relative to the SMV portfolio. In the next section, we review how Tu and Zhou (2011) exploit the EU criterion to combine the SMV and EW portfolios.

2.3 Constrained portfolio combination

Tu and Zhou (2011) consider the set of portfolios that combine the SMV portfolio $\hat{\boldsymbol{w}}^\star$ and the equally weighted (EW) portfolio $\boldsymbol{w}_{ew} = \mathbf{1}/N$:

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \kappa_1\hat{\boldsymbol{w}}^\star + \kappa_2\boldsymbol{w}_{ew}, \tag{9}$$

where $\boldsymbol{\kappa} = (\kappa_1, \kappa_2)$ is the vector of *combination coefficients*. We denote $\mu_{ew} = \boldsymbol{w}_{ew}'\boldsymbol{\mu}$ and $\sigma_{ew}^2 = \boldsymbol{w}_{ew}'\boldsymbol{\Sigma}\boldsymbol{w}_{ew}$ the mean excess return and variance of the EW portfolio, respectively. Moreover, we define γ_{ew} as[2]

$$\gamma_{ew} = \mu_{ew}/\sigma_{ew}^2, \tag{10}$$

and ψ^2 as the difference between the maximum squared Sharpe ratio, θ^2, and that of the EW portfolio, $\theta_{ew}^2 = \mu_{ew}^2/\sigma_{ew}^2$:

$$\psi^2 = \theta^2 - \theta_{ew}^2 \geq 0. \tag{11}$$

The methodology of Tu and Zhou (2011) consists in finding the combination coefficients $\boldsymbol{\kappa}$

[2]The parameter γ_{ew} has the following interpretation: an investor with risk aversion γ investing in the EW portfolio and the risk-free asset fully allocates her wealth to the EW portfolio if and only if $\gamma = \gamma_{ew}$.

maximizing the EU in (8). They achieve this under the following convexity constraint:

$$\kappa_1 + \kappa_2 = 1. \tag{12}$$

We review their results in the next proposition, which we present in a different form.

Proposition 1. *The following holds concerning the combination of the sample mean-variance and equally weighted portfolios in (9):*

1. *The expected out-of-sample utility is*

$$EU(\hat{\boldsymbol{w}}(\boldsymbol{\kappa})) = \frac{\kappa_1}{\gamma}\theta^2 + \kappa_2\mu_{ew} - \frac{\gamma}{2}\left(\frac{\kappa_1^2}{\gamma^2}\Big(\theta^2 + d\Big) + \kappa_2^2\sigma_{ew}^2 + \frac{2\kappa_1\kappa_2}{\gamma}\mu_{ew}\right), \tag{13}$$

   where

$$d = cN/T + (c-1)\theta^2 \quad \textit{with} \quad c = \frac{(T-N-2)(T-2)}{(T-N-1)(T-N-4)} \geq 1. \tag{14}$$

2. *The combination coefficients maximizing (13) under the constraint $\kappa_1 + \kappa_2 = 1$ are*

$$\boldsymbol{\kappa}^{tz} = (\varphi(\gamma), 1 - \varphi(\gamma)), \tag{15}$$

   where

$$\varphi(\gamma) = \frac{\psi^2 + \sigma_{ew}^2(\gamma - \gamma_{ew})^2}{\psi^2 + \sigma_{ew}^2(\gamma - \gamma_{ew})^2 + d} \in [0, 1]. \tag{16}$$

Note that the EU of the SMV portfolio is $EU(\hat{\boldsymbol{w}}^\star) = (\theta^2 - d)/(2\gamma)$. Therefore, the SMV portfolio delivers a negative EU and underperforms the risk-free asset if $d > \theta^2$, which happens already for $T < 325$ for the 25SBTM dataset. This result shows the necessity of regularizing the SMV portfolio to improve performance in the presence of estimation risk.

In the next section, we discuss why imposing the convexity constraint (12) is actually unnecessary. Moreover, we derive the optimal *unconstrained* combination coefficients and demonstrate that the optimal portfolio combination offers several important benefits relative to the constrained portfolio combination of Tu and Zhou (2011). Finally, we show in Section 4 that the convexity constraint, although theoretically suboptimal, can actually help improve the EU for some degrees of risk aversion γ when combination coefficients are estimated.

3 Optimal versus constrained portfolio combination

Imposing the convexity constraint (12) is needed when considering combinations of fully invested portfolios, $\boldsymbol{w} = \kappa_1 \boldsymbol{w}^1 + \kappa_2 \boldsymbol{w}^2$ with $\mathbf{1}'\boldsymbol{w}^1 = \mathbf{1}'\boldsymbol{w}^2 = 1$.[3] The reason is that constraint (12) is then equivalent to the full-investment constraint on the combined portfolio:

$$\mathbf{1}'\boldsymbol{w} = 1 \quad \Leftrightarrow \quad \kappa_1 \mathbf{1}'\boldsymbol{w}^1 + \kappa_2 \mathbf{1}'\boldsymbol{w}^2 = 1 \quad \Leftrightarrow \quad \kappa_1 + \kappa_2 = 1.$$

However, the situation is different when combining the SMV and EW portfolios because, as we recall in Section 2.1, SMV is not fully invested: $\mathbf{1}'\hat{\boldsymbol{w}}^\star = (\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}})/\gamma \neq 1$, in general. In fact, combining the SMV and EW portfolios amounts to invest in three funds: the sample tangent portfolio and the EW portfolio, which are fully invested in risky assets, and the risk-free asset. The problem is that, under constraint (12), a *single* coefficient κ_1 controls the weight allocated to the three funds simultaneously, the second combination coefficient being $\kappa_2 = 1 - \kappa_1$. In particular, there is no way to disentangle the weights allocated to the sample tangent portfolio and the risk-free asset by playing with κ_1.

To circumvent this issue, we propose to relax constraint (12) and use two different combination coefficients κ_1 and κ_2 while still respecting the full-investment constraint. We derive the optimal unconstrained combination coefficients in the next proposition.[4]

Proposition 2. *The optimal combination coefficients maximizing* (13) *are*

$$\boldsymbol{\kappa}^{opt} = \left(\varphi(\gamma_{ew}), \frac{\gamma_{ew}}{\gamma}(1 - \varphi(\gamma_{ew}))\right) = \left(\frac{\psi^2}{\psi^2 + d}, \frac{\gamma_{ew}}{\gamma}\frac{d}{\psi^2 + d}\right). \tag{17}$$

In the sequel, we call the *constrained strategy* and the *optimal strategy* the portfolio combinations (9) exploiting the constrained coefficients (15) and the optimal coefficients (17), respectively. We denote them as $\hat{\boldsymbol{w}}^{tz} = \hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{tz})$ and $\hat{\boldsymbol{w}}^{opt} = \hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{opt})$.

[3] For example, Kan, Wang, and Zhou (2021) combine the fully invested SMV and SGMV portfolios and impose this constraint; see Appendix IA.2.2. The convexity constraint is also sensible when shrinking the sample covariance matrix to estimate the GMV portfolio (Ledoit and Wolf, 2004).

[4] Kan and Wang (2021, Section 4.2) use similar coefficients in a different context, that of combining a benchmark factor model with a set of test assets. In Appendix IA.2.4, we also derive the optimal combination of the SMV, SGMV, and EW portfolios.

We now proceed by showing that the optimal strategy delivers five key benefits relative to the constrained strategy of Tu and Zhou (2011): 1) optimal exposure to the three funds, 2) outperformance in expected out-of-sample utility, 3) outperformance of the risk-free asset, 4) outperformance of the optimal two-fund rule, 5) less extreme weights on risky assets. Our theory in this section assumes that we know the combination coefficients $\boldsymbol{\kappa}$ perfectly; we study the impact of estimation errors in these coefficients in Section 4.

3.1 Optimal exposure to the three funds

We first compare the constrained and optimal combination coefficients, $\boldsymbol{\kappa}^{tz}$ in (15) and $\boldsymbol{\kappa}^{opt}$ in (17). The two combination strategies are not always different: (i) they fully invest in the SMV portfolio in the absence of estimation risk ($N/T \to 0$), (ii) they fully invest in the risk-free asset as $\gamma \to \infty$, and (iii) they coincide when $\gamma = \gamma_{ew}$.

Otherwise, the constrained strategy has a striking difference: it *always* overinvests in the SMV portfolio. In comparison, the weight on the EW portfolio can be larger or smaller, depending on μ_{ew} and γ. We also study the average weight allocated to the risk-free asset,[5]

$$\pi_{rf}^{opt} = 1 - \pi_{tan}^{opt} - \pi_{ew}^{opt} \quad \text{and} \quad \pi_{rf}^{tz} = 1 - \pi_{tan}^{tz} - \pi_{ew}^{tz}, \tag{18}$$

where

$$\pi_{tan} = \frac{\gamma_{tan}}{\gamma}\kappa_1 \quad \text{and} \quad \pi_{ew} = \kappa_2 \tag{19}$$

are the average weight allocated to the sample tangent and EW portfolios, respectively, with γ_{tan} defined in (5). We show that under mild conditions the constrained strategy underexploits the risk-free asset. We summarize these results in the next proposition.

Proposition 3. *The following holds regarding the constrained and optimal strategies:*

1. *The constrained strategy overinvests in the SMV portfolio:*

$$0 \leq \kappa_1^{opt} \leq \kappa_1^{tz} \leq 1, \tag{20}$$

[5]We study the average weight because the weight depends on $\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}$, which is sample-dependent.

and $\kappa_1^{opt} = \kappa_1^{tz}$ *if and only if* $\gamma = \gamma_{ew}$.

2. *The investment in the EW portfolio depends on its mean excess return* μ_{ew}*:*

   - *If* $\mu_{ew} = 0$*, the constrained strategy is long the EW portfolio and the optimal strategy discards it:* $0 = \kappa_2^{opt} \leq \kappa_2^{tz}$*, with equality if and only if* $\gamma = \infty$.
   - *If* $\mu_{ew} < 0$*, the constrained (optimal) strategy is long (short) the EW portfolio:* $\kappa_2^{opt} \leq 0 \leq \kappa_2^{tz}$.
   - *If* $\mu_{ew} > 0$*, both strategies invest are long the EW portfolio:*

$$0 \leq \kappa_2^{opt} \leq \kappa_2^{tz} \quad \textit{if} \quad \gamma \in \left[\gamma_{ew}, \frac{\theta^2 + d}{\mu_{ew}}\right] \quad \textit{and} \quad \kappa_2^{opt} \geq \kappa_2^{tz} \quad \textit{otherwise.} \tag{21}$$

3. *Let* $0 < \gamma_{ew} < \gamma_{tan} < (\theta^2 + d)/\mu_{ew}$.[6] *Then, the constrained strategy underexploits the risk-free asset:*

$$\pi_{rf}^{opt} \leq \pi_{rf}^{tz} \leq 0 \quad \textit{if} \quad \gamma \leq \gamma_{ew} \quad \textit{and} \quad \pi_{rf}^{opt} \geq \pi_{rf}^{tz} \quad \textit{otherwise.} \tag{22}$$

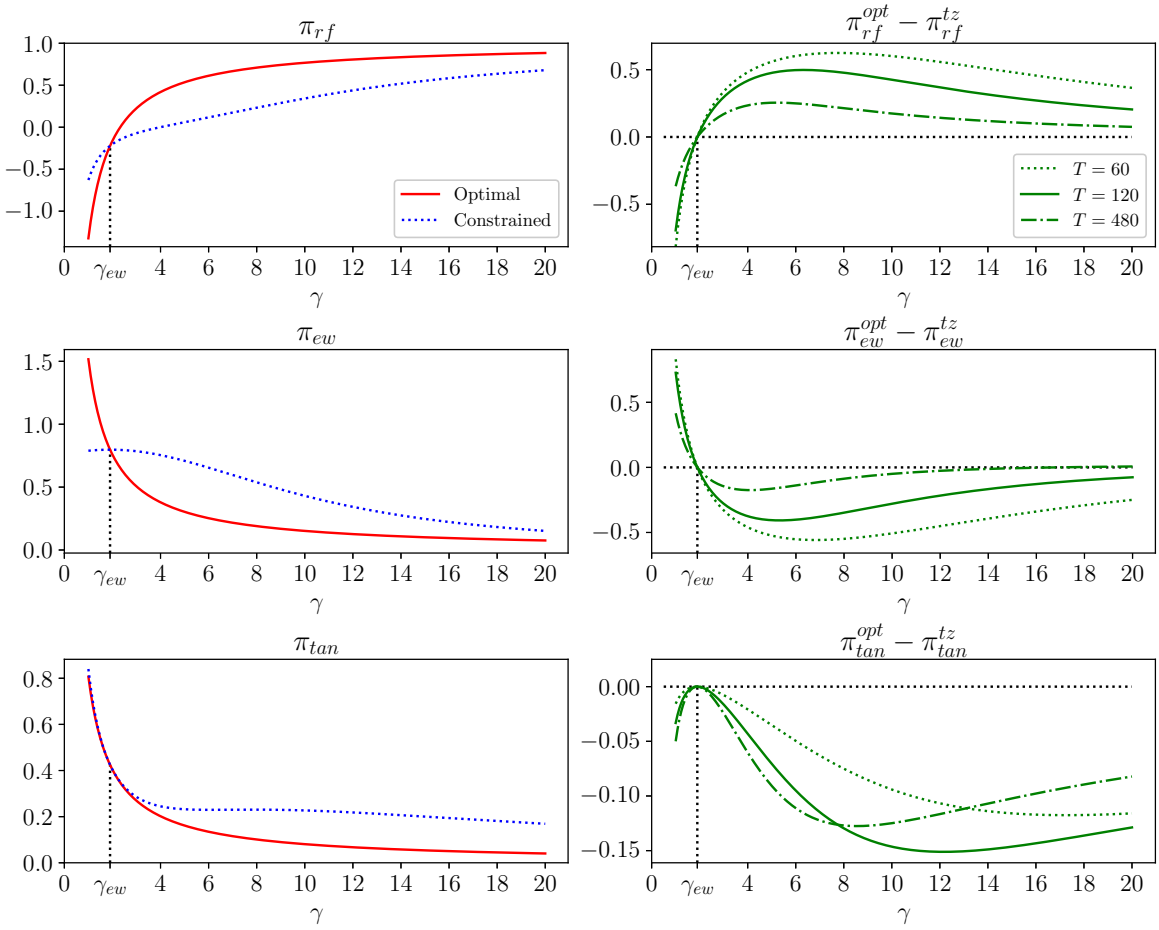
We illustrate Proposition 3 in Figure 1 by calibrating the moments of asset excess returns, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, to a dataset of $N = 25$ portfolios of firms sorted on size and book-to-market spanning July 1926 to December 2021 (25SBTM dataset hereafter). The left panels use $T = 120$, and the right panels use $T = 60, 120$, and 480.

The figure shows that the allocation to the three funds in the constrained strategy is substantially different from that in the optimal strategy. Consider an investor with $\gamma = 5$, which is larger than $\gamma_{ew} = 1.9$. When $T = 120$, the constrained strategy commands a weight of 23% on the tangent portfolio, 71% on the EW portfolio, and 6% on the risk-free asset. In contrast, the optimal strategy invests much more in the risk-free asset (54%) and much less in the tangent and EW portfolios (16% and 30%, respectively).

A remarkable result in Proposition 3 is that the constrained strategy overinvests in the tangent portfolio and underexploits the risk-free asset. This result can be understood because, due to the $\kappa_1 + \kappa_2 = 1$ constraint, the only way the investor can adjust her exposure to the

[6] This is a mild assumption for typical datasets. In particular, all datasets we use in the empirical analysis of Section 5 fulfill this assumption.

Figure 1: Weights of the three funds in the constrained and optimal strategies



Notes. This figure compares the average weight allocated to the three funds by the constrained and optimal combination strategies defined in (15) and (17), respectively. The three funds (weights) are the risk-free asset (π_{rf}), the EW portfolio (π_{ew}), and the sample tangent portfolio (π_{tan}). The formulas for the average weights are given by (18)–(19). The figure is constructed by calibrating the population vector of means and covariance matrix of excess asset returns from monthly returns on the 25 portfolios of firms sorted on size and book-to-market spanning July 1926 to December 2021. The left panels depict the weight allocated to the three funds as a function of γ when the sample size $T = 120$. The right panels depict the difference in weights between the optimal and constrained strategies as a function of γ when $T = 60, 120$, and 480. The figure illustrates the results in Proposition 3: (i) $\pi_{rf}^{opt} \leq \pi_{rf}^{tz} \leq 0$ if $\gamma \leq \gamma_{ew}$ and $\pi_{rf}^{opt} \geq \pi_{rf}^{tz}$ otherwise, (ii) $\pi_{tan}^{tz} \geq \pi_{tan}^{opt}$, and (iii) $\pi_{ew}^{tz} \geq \pi_{ew}^{opt}$ for $\gamma \in [\gamma_{ew}, (\theta^2 + d)/\mu_{ew}] = [1.9, 42.4]$ when $T = 120$.

risk-free asset is by investing in the SMV portfolio. However, the SMV portfolio is risky, and thus being optimally exposed to the risk-free asset would deliver a too risky portfolio. This forces the constrained strategy to reduce the exposure to the risk-free asset, but up to a point where the exposure to the SMV portfolio is still larger than in the optimal strategy. Figure 1 shows that this effect is substantial, particularly when γ is high. For instance, when $\gamma = 10$ and $T = 120$, we have $(\pi_{rf}^{opt}, \pi_{tan}^{opt}) = (77\%, 8\%)$ and $(\pi_{rf}^{tz}, \pi_{tan}^{tz}) = (34\%, 23\%)$, that is, the constrained weight on the risk-free asset is 56% smaller than the optimal one, and the constrained weight on the tangent portfolio is 180% larger.

Proposition 3 also implies that for a wide range of risk-aversion coefficients the constrained strategy overinvests in the EW portfolio, which is because the interval in (21) is wide.[7] Therefore, while investing in the EW portfolio helps alleviate estimation errors and improve performance, the convexity constraint (12) often forces the investor to be overexposed to this portfolio. For example, when $\gamma = 5$ and $T = 120$, $\pi_{ew}^{opt} = 30\%$ while $\pi_{ew}^{tz} = 71\%$.

Finally, the right panels of Figure 1 show that the overexposure to risky assets and the underexposure to the risk-free asset in the constrained strategy is particularly pronounced when T is rather small. That is, it is when combining portfolios is most needed because estimation errors are large that the convexity constraint hurts the most.

3.2 Outperformance in expected out-of-sample utility

Just like there is no particular reason in theory to impose the convexity constraint on the combination coefficients, there is no reason for the maximum EU found along the line $\kappa_1 + \kappa_2 = 1$ to coincide with the global optimum, which we show in the next proposition.

Proposition 4. *The optimal strategy delivers a larger expected out-of sample utility than the constrained strategy:*

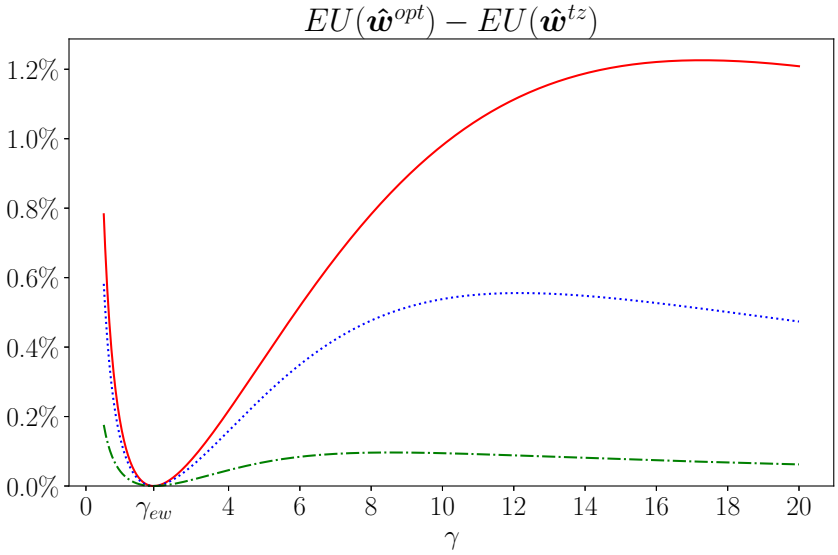
$$EU(\hat{\boldsymbol{w}}^{opt}) - EU(\hat{\boldsymbol{w}}^{tz}) = \frac{d}{2\gamma}\left(\kappa_1^{tz} - \kappa_1^{opt}\right) \geq 0, \tag{23}$$

with equality if and only if $\gamma = \gamma_{ew}$ or $\gamma = \infty$. Moreover, the utility gain in (23) increases

[7] For the data used to construct Figure 1, we have $\gamma_{ew} = 1.90$ and $(\theta^2 + d)/\mu_{ew} = 42.4$ when $T = 120$. Other datasets yield similar values.

with estimation risk N/T.

Figure 2: Expected out-of-sample utility of optimal and constrained strategy



Notes. This figure depicts the gain in expected out-of-sample utility delivered by the optimal combination strategy $\hat{\boldsymbol{w}}^{opt}$ relative to the constrained combination strategy $\hat{\boldsymbol{w}}^{tz}$ as a function of the risk-aversion coefficient γ between 0.5 and 20 for $N = 25$ and $T = 60$ (red solid), $T = 120$ (blue dotted), and $T = 480$ (green dash-dotted). The figure is constructed by calibrating the population vector of means and covariance matrix of excess asset returns from monthly returns on the 25 portfolios of firms sorted on size and book-to-market spanning July 1926 to December 2021. The gain in expected out-of-sample utility is given in Proposition 4. Both strategies are equivalent when $\gamma = \gamma_{ew} = 1.90$ or $\gamma = \infty$, and the gain is always positive.

Proposition 4 shows that it is never beneficial to choose the constrained strategy over the optimal strategy in terms of EU.[8] Moreover, the gain delivered by the optimal strategy increases with estimation risk, N/T. This result can be explained by Proposition 3: the constrained strategy overinvests in the SMV portfolio relative to the optimal strategy, which is the only portfolio in the combination that is exposed to estimation risk. We illustrate Proposition 4 in Figure 2 using the 25SBTM dataset as in Figure 1. The figure shows that the EU gain allowed by the optimal strategy can be substantial when γ departs from $\gamma_{ew} = 1.9$. For example, when γ varies between 0.5 and 20 and $T = 120$, the EU gain in (23) goes up to around 0.5-0.6%, that is, around 6-7% annually. Moreover, we find in the empirical analysis of Section 5 that the higher exposure to the tangent portfolio in the constrained strategy

[8] In Appendix IA.2.6, we show similarly that the optimal strategy always delivers a theoretically larger expected out-of sample Sharpe ratio than that of the constrained strategy.

increases turnover, which makes the EU gain even more substantial net of transaction costs.

3.3 Outperformance relative to the risk-free asset

In the next proposition, we show another important deficiency of the constrained combination of Tu and Zhou (2011): it can *underperform the risk-free asset* when there is sufficient estimation risk N/T and risk aversion γ. In contrast, the optimal unconstrained strategy always outperforms the risk-free asset. This result is well corroborated by the empirical results presented in Section 5.

Proposition 5. *The following holds in relation to the risk-free asset:*

1. *The constrained strategy $\hat{\boldsymbol{w}}^{tz}$ delivers a smaller expected out-of-sample utility than the risk-free asset, $EU(\hat{\boldsymbol{w}}^{tz}) < 0$, if and only if i) $\theta^2 < d$ and ii)*

$$\gamma > \gamma_{ew}\left(1 + \sqrt{1 + \frac{\theta^4/\theta_{ew}^2}{d - \theta^2}}\right) = \gamma_{neg}. \tag{24}$$

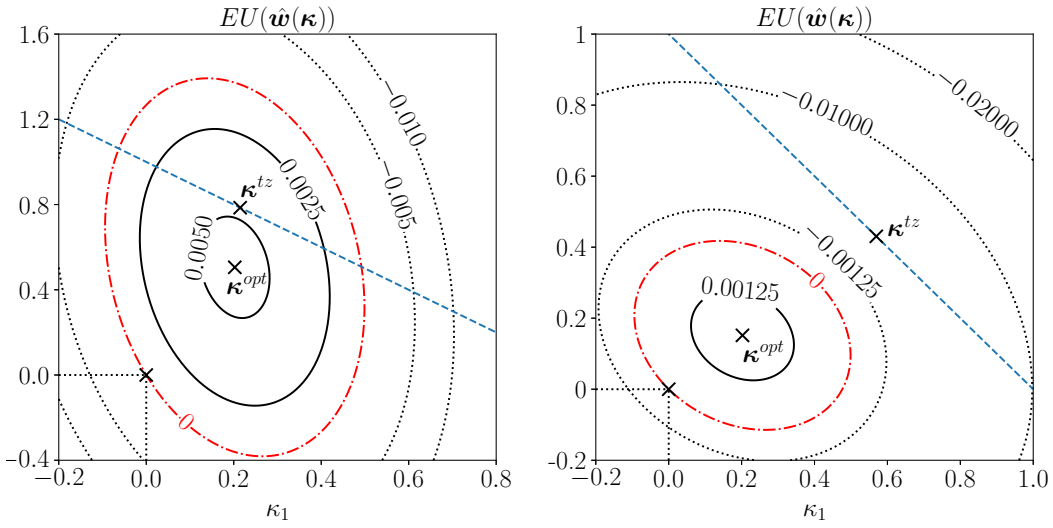
   Moreover, γ_{neg} decreases with estimation risk N/T.

2. *The optimal strategy $\hat{\boldsymbol{w}}^{opt}$ always delivers a larger expected out-of-sample utility than the risk-free asset, $EU(\hat{\boldsymbol{w}}^{opt}) \geq 0$, with equality if and only if $\gamma = +\infty$.*

Proposition 5 shows that because the constrained strategy lacks a degree of freedom to optimally invest in the three funds, it might be less desirable than simply holding the risk-free asset for investors who are sufficiently risk averse, $\gamma > \gamma_{neg}$. This is because, as explained above, the only way to invest in the risk-free asset under the convexity constraint is via the SMV portfolio. However, as estimation risk increases, the constrained strategy allocates more weight to the EW strategy (κ_2^{tz} increases) and, because of the convexity constraint, less weight to the SMV strategy (κ_1^{tz} decreases). By doing so, it does not sufficiently invest in the risk-free asset, whose weight is proportional to κ_1^{tz}, and thus may underperform. This problem does not arise in the optimal strategy which, by using two unconstrained combination coefficients κ_1 and κ_2, renders a full investment in the risk-free asset attainable by setting $\boldsymbol{\kappa} = \mathbf{0}$.

This deficiency of the constrained portfolio combination is particularly notable because the two conditions under which it is outperformed by the risk-free asset can easily be met. As

Figure 3: Optimal and constrained strategy versus the risk-free asset



(a) $\gamma = 3 < \gamma_{neg}$ (b) $\gamma = 10 > \gamma_{neg}$

Notes. This figure depicts the contours of expected out-of-sample utility (EU) as a function of the combination coefficients $\boldsymbol{\kappa} = (\kappa_1, \kappa_2)$. The figure is constructed by calibrating the population vector of means and covariance matrix of excess asset returns from monthly returns on the 25 portfolios of firms sorted on size and book-to-market spanning July 1926 to December 2021, and using a sample size $T = 120$. The black solid and dotted contours represent positive and negative EU values, respectively. The red dash-dotted contour is the set of $\boldsymbol{\kappa}$ delivering zero EU. The pair located at the origin, identified by a black marker, corresponds to the risk-free asset. The blue dashed line represents the convexity constraint $\kappa_1 + \kappa_2 = 1$, on which is located the constrained combination coefficients $\boldsymbol{\kappa}^{tz}$ in (15). The optimal coefficients $\boldsymbol{\kappa}^{opt}$ in (17) deliver the global EU optimum. The two panels depict the contours for risk-aversion coefficients $\gamma = 3$ and 10. As shown in Proposition 5, because $\theta^2 = 0.092 < d = 0.29$, $\boldsymbol{\kappa}^{tz}$ delivers a positive EU if $\gamma \leq \gamma_{neg} = 5.41$ and a negative EU otherwise. In contrast, the optimal coefficients $\boldsymbol{\kappa}^{opt}$ always deliver a positive EU.

discussed in Section 2.3, the condition $\theta^2 < d$ is equivalent to the SMV portfolio $\hat{\boldsymbol{w}}^\star$ having a negative EU. For typical values of $\theta = 0.3$ and $T = 120$, this condition is met as soon as $N > 8$. Moreover, the lower bound γ_{neg} is typically not large. For example, $\gamma_{neg} = 5.41$ for the 25SBTM dataset and $T = 120$.[9]

We illustrate Proposition 5 in Figure 3 for the 25SBTM dataset and $T = 120$, in which case $\gamma_{neg} = 5.41$. We depict the EU contour lines in the (κ_1, κ_2) plane for two levels of risk aversion: $\gamma = 3 < \gamma_{neg}$ and $\gamma = 10 > \gamma_{neg}$. The figure shows that not only the $\kappa_1 + \kappa_2 = 1$ line does not pass by the global optimum $\boldsymbol{\kappa}^{opt}$, but it also only passes by negative EU values

[9]In their experiments, Tu and Zhou (2011) only consider relatively small values of $\gamma = 1$ and 3. Already for $\gamma = 3$, they find in Panel A of their Table 6 that some of the constrained combination strategies they introduce can deliver negative utilities and thus underperform the risk-free asset.

when $\gamma > \gamma_{neg}$, in which case the constrained strategy underperforms the risk-free asset. In comparison, the optimal strategy always delivers positive utilities.

3.4 Outperformance relative to the optimal two-fund rule

In this section, we show that the constrained strategy does not fully profit from adding the EW portfolio in the three-fund combination. Indeed, it does not systematically outperform the optimal two-fund rule of Kan and Zhou (2007), which drops the EW portfolio from the combination in (9). This optimal two-fund portfolio combination is

$$\hat{\boldsymbol{w}}^{2f} = \kappa^{2f}\hat{\boldsymbol{w}}^{\star} \quad \text{with} \quad \kappa^{2f} = \frac{\theta^2}{\theta^2 + d}. \tag{25}$$

In the next proposition, we show that while the optimal three-fund rule $\hat{\boldsymbol{w}}^{opt}$ always outperforms the optimal two-fund rule $\hat{\boldsymbol{w}}^{2f}$, the constrained three-fund rule $\hat{\boldsymbol{w}}^{tz}$ only outperforms if γ is small enough, that is, $\gamma \leq 2\gamma_{ew}$.

Proposition 6. *The optimal three-fund strategy $\hat{\boldsymbol{w}}^{opt}$ always delivers a larger expected out-of-sample utility than the optimal two-fund strategy $\hat{\boldsymbol{w}}^{2f}$. In contrast, the constrained three-fund strategy $\hat{\boldsymbol{w}}^{tz}$ does so if and only if $\gamma \leq 2\gamma_{ew}$.*

Considering that the value of γ_{ew} is typically quite small, the result in Proposition 6 means that many investors are better off relying on the optimal two-fund rule that does not invest in the EW portfolio rather than adding the EW portfolio via the convexity constraint (12).

Moreover, we show in Appendix IA.2.3 that even though the additional combination coefficient to estimate in the optimal three-fund rule, κ_2^{opt}, is sensitive to estimation errors, the required sample size for the estimated optimal three-fund rule to deliver a larger EU than the perfectly estimated optimal two-fund rule is not large. For example, a sample size $T \geq 75$ is required for the 25SBTM dataset.

3.5 Less extreme portfolio weights

A final desirable property of the optimal strategy relative to the constrained one is that the weights allocated to the risky assets are less extreme, in the sense of having a smaller

norm. Large weights are undesirable in practice because they are less easily implementable and often result in larger turnover and transaction costs, which we confirm in our empirical analysis in Section 5. Moreover, DeMiguel et al. (2009) and Zhao, Ledoit, and Jiang (2021), among others, show that constraining portfolio norms helps improve performance. We use the ℓ_2-norm of the vector of weights on the risky assets,

$$||\boldsymbol{w}||_2 = \sqrt{\sum_{i=1}^{N} w_i^2}. \tag{26}$$

In the next proposition, we show that for a large range of values of the risk-aversion coefficient γ, the optimal strategy has smaller ℓ_2-norm than that of the constrained strategy.

Proposition 7. *Let $\gamma_{tan} \geq 0$, $\mu_{ew} \geq 0$,[10] and $\gamma \in [\gamma_{ew}, (\theta^2 + d)/\mu_{ew}]$. Then, the ℓ_2-norm of the vector of weights on the risky assets in (26) is smaller for the optimal strategy than for the constrained strategy,*

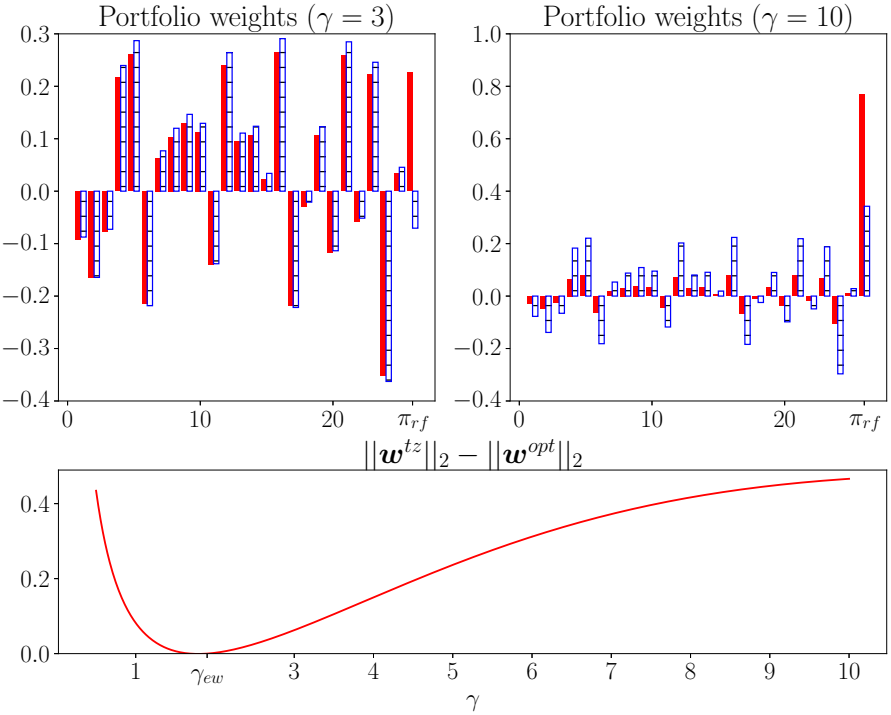
$$||\hat{\boldsymbol{w}}^{opt}||_2 \leq ||\hat{\boldsymbol{w}}^{tz}||_2. \tag{27}$$

This result follows from Proposition 3, which shows that relative to the optimal strategy, the constrained strategy always allocates more weight to the tangent portfolio, and also to the EW portfolio when $\gamma \in [\gamma_{ew}, (\theta^2 + d)/\mu_{ew}]$. This interval is wide in practice; it is equal to $[1.9, 42.4]$ for the data used in Figure 1 when $T = 120$. Moreover, it is only a sufficient condition, and the optimal strategy may still deliver a smaller ℓ_2-norm than the constrained strategy when γ is outside of this interval, as we find in the following illustration.

We illustrate Proposition 7 in Figure 4 for the 25SBTM dataset and $T = 120$. We depict the average portfolio weights on the 25 risky assets and the risk-free asset for $\gamma = 3$ and 10, and the difference in ℓ_2-norm, $||\hat{\boldsymbol{w}}^{tz}||_2 - ||\hat{\boldsymbol{w}}^{opt}||_2$, as a function of γ between 0.5 and 10. When $\gamma = 3$, the optimal strategy allocates a positive weight to the risk-free asset, whereas the constrained strategy leverages the risk-free asset with a negative weight. When $\gamma = 10$, the difference between the optimal and constrained strategies is particularly striking: the ℓ_2-norm is much smaller with the optimal strategy, $||\boldsymbol{w}^{opt}||_2 = 0.26 < ||\boldsymbol{w}^{tz}||_2 = 0.72$, and consequently it allocates much more weight to the risk-free asset, $\pi_{rf}^{opt} = 0.77 > \pi_{rf}^{tz} = 0.34$.

[10]The two assumptions $\gamma_{tan} \geq 0$ and $\mu_{ew} \geq 0$ are met when the GMV portfolio and the EW portfolio have a positive mean excess return, which is typically the case.

Figure 4: Comparison of portfolio weights and norms



Notes. The two upper panels depict the average portfolio weights allocated to the 25 risky assets and the risk-free asset (π_{rf}), for risk-aversion coefficients of $\gamma = 3$ and $\gamma = 10$, respectively. The red solid bars represent the weights in the optimal strategy while the blue striped bars represent the weights in the constrained strategy. The bottom panel depicts the difference between the ℓ_2-norm of the optimal and constrained strategies as a function of γ. The figure is constructed by calibrating the population vector of means and covariance matrix of excess asset returns from monthly returns on the 25 portfolios of firms sorted on size and book-to-market spanning July 1926 to December 2021, and using a sample size $T = 120$. The figure shows that the constrained strategy delivers a larger ℓ_2-norm for all risk-aversion coefficients.

Finally, the figure shows that the constrained strategy delivers a larger ℓ_2-norm for any γ.

4 When the convexity constraint can help

The theoretical results in Section 3 show the benefits of dropping the convexity constraint (12) to optimally combine the sample tangent portfolio, the EW portfolio, and the risk-free asset. However, these results rely on the crucial assumption that the combination coefficients are known. Because they depend on the population moments of asset returns, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, this assumption is not realistic: both $\boldsymbol{\kappa}^{tz}$ and $\boldsymbol{\kappa}^{opt}$ must be estimated from the data. Therefore, they are themselves subject to estimation errors and, as Kan and Wang

(2021) and Lassance, Martín-Utrera, and Simaan (2022) show, the impact of these errors on out-of-sample performance can be substantial.

In this section, we show that although the optimal strategy outperforms the constrained strategy in theory, it can underperform for a certain range of risk-aversion coefficients when we account for estimation errors in combination coefficients. The reason is that, as Jagannathan and Ma (2003) put it, *imposing the wrong constraint can help*. This effect is well known for portfolio weights: imposing bound constraints (Frost and Savarino, 1988; Jagannathan and Ma, 2003) or norm constraints (DeMiguel et al., 2009; Ao, Li, and Zheng, 2019) is theoretically suboptimal but has a regularization effect by preventing the weights to explode, which can help improve out-of-sample performance. A similar phenomenon happens when imposing the convexity constraint on combination coefficients: the resulting constrained coefficients in (15) are always between zero and one. In comparison, κ_1^{opt} is also between zero and one, but κ_2^{opt} is *unbounded* because it is proportional to $\gamma_{ew} = \mu_{ew}/\sigma_{ew}^2$. This suggests that when γ is close to γ_{ew} and the theoretical gain of the optimal strategy over the constrained strategy is small, the constrained strategy might actually perform better out of sample.

This intuition is confirmed in Appendix IA.1. We simulate normally distributed asset returns with parameters $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and repeatedly estimate the combination coefficients. We find that $\kappa_1^{tz} = 1 - \kappa_2^{tz}$ and κ_1^{opt} are not very sensitive to estimation errors. In particular, estimation errors in mean returns $\boldsymbol{\mu}$ have a limited effect because they are partially offset in the numerator and denominator of the coefficients. However, this result is no longer true for κ_2^{opt} that is much more impacted by estimation errors in $\boldsymbol{\mu}$, which is problematic because mean returns are notoriously difficult to estimate (Merton, 1980). This behavior can be explained because $\kappa_2^{opt} = \frac{1}{\gamma}\frac{\mu_{ew}}{\sigma_{ew}^2}(1 - \kappa_1^{opt})$ and therefore is directly proportional to mean returns $\boldsymbol{\mu}$ via μ_{ew}. This suggests that the additional impact of estimation errors in the optimal strategy relative to the constrained strategy mainly results from μ_{ew} in κ_2^{opt}.

Therefore, in the next proposition, we account for the relative impact of estimation errors in combination coefficients by deriving the EU of the optimal strategy when the parameter μ_{ew} is estimated in κ_2^{opt}:

$$\hat{\kappa}_2^{opt} = \frac{1}{\gamma}\frac{\hat{\mu}_{ew}}{\sigma_{ew}^2}(1 - \kappa_1^{opt}), \tag{28}$$

where $\hat{\mu}_{ew} = \boldsymbol{w}_{ew}'\hat{\boldsymbol{\mu}}$. Moreover, we show that the resulting estimated optimal strategy is outperformed by the constrained strategy when γ belongs to an interval centered around γ_{ew}.

Proposition 8. *Let the combination coefficient κ_2^{opt} be estimated by $\hat{\kappa}_2^{opt}$ in (28). Then,*

1. *The expected out-of-sample utility of the estimated optimal strategy is*

$$EU\Big(\hat{\boldsymbol{w}}(\kappa_1^{opt}, \hat{\kappa}_2^{opt})\Big) = EU(\hat{\boldsymbol{w}}^{opt}) - \frac{1 - (\kappa_1^{opt})^2}{2\gamma T}, \tag{29}$$

   where $EU(\hat{\boldsymbol{w}}^{opt}) = U^\star - \frac{d}{2\gamma}\kappa_1^{opt}$ is the utility when κ_2^{opt} is known.

2. *Let $N \geq 2$.[11] Then, the constrained strategy delivers a larger expected out-of-sample utility than the estimated optimal strategy, $EU(\hat{\boldsymbol{w}}^{tz}) \geq EU\Big(\hat{\boldsymbol{w}}(\kappa_1^{opt}, \hat{\kappa}_2^{opt})\Big)$, if and only if γ belongs to the following interval:*

$$\gamma \in \left[\gamma_{ew} \pm \sigma_{ew}^{-1}\sqrt{\frac{(\psi^2 + d)(2\psi^2 + d)}{dT(\psi^2 + d) - (2\psi^2 + d)}}\right] = \Big[\underline{\gamma}_{ew}, \overline{\gamma}_{ew}\Big]. \tag{30}$$

   Moreover, the length of this interval decreases with N.[12]

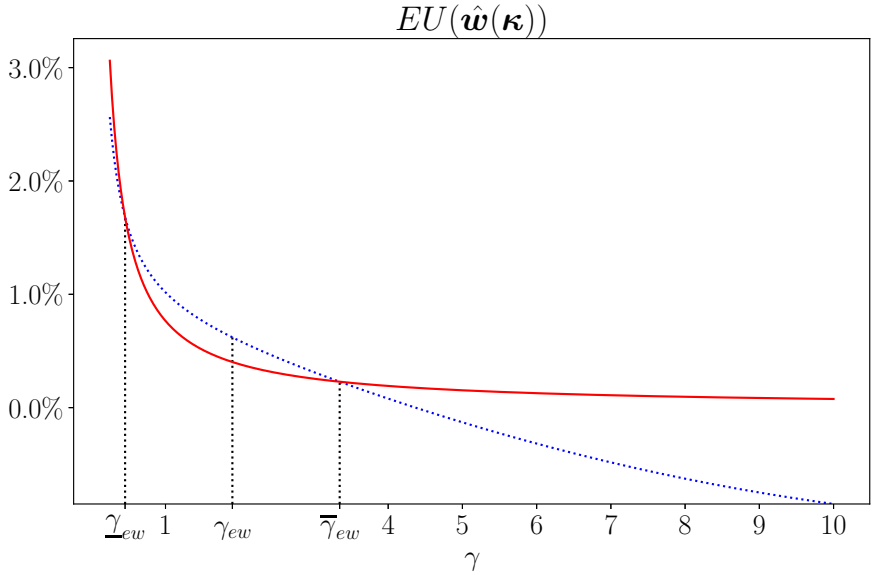
Proposition 8 shows that when γ is sufficiently close to γ_{ew} according to the interval (30), and thus that the theoretical gain provided by the optimal strategy in (23) is not large enough, the constrained strategy can deliver a *larger* EU due to the impact of estimation errors in μ_{ew} on κ_2^{opt}. However, as the number of assets N increases and the constrained strategy becomes more suboptimal in theory, the length of the interval (30) decreases. We illustrate Proposition 8 in Figure 5, in which we compare the EU of the estimated optimal strategy and that of the constrained strategy for the 25SBTM dataset and $T = 120$.

The length of this interval in (30) is not negligible. For the 25SBTM dataset example, it is equal to $\gamma \in [1.90 \pm 1.65]$ when $T = 120$ and $\gamma \in [1.90 \pm 1.40]$ when $T = 240$. Therefore, it is more desirable to rely on the constrained strategy than the optimal one for most small values of γ, but the optimal strategy remains preferable when γ is medium or

[11] When $N = 1$, the quantity in the square root in the interval (30) turns negative if T is large enough, in which case the estimated optimal strategy delivers a larger EU for all γ.

[12] As the sample size $T \to \infty$, both the optimal and constrained strategy deliver the maximum EU, $U^\star$ in (3), and thus the interval of γ for which one does better than the other is ill-defined. This explains why the interval in (30) does not vanish as $T \to \infty$ but, instead, converges to $\gamma_{ew} \pm \sigma_{ew}^{-1}\sqrt{2\psi^2/(N-2)}$.

Figure 5: Impact of estimation errors in combination coefficients on performance



Notes. This figure depicts the expected out-of-sample utility (EU) as a function of the risk-aversion coefficient γ for two different strategies: the estimated optimal strategy in (29) (solid red) and the constrained strategy in (15) (dotted blue). The figure is constructed by calibrating the population vector of means and covariance matrix of excess asset returns from monthly returns on the 25 portfolios of firms sorted on size and book-to-market spanning July 1926 to December 2021, and using a sample size $T = 120$ and number of assets $N = 50$. As shown in Proposition 8, the constrained strategy delivers a larger EU than the estimated optimal strategy when $\gamma \in [\underline{\gamma}_{ew}, \overline{\gamma}_{ew}] = [1.90 \pm 1.45]$.

large. This is expected because it is particularly when γ gets large enough that the convexity constraint (12) hurts performance by preventing enough investment in the risk-free asset, as we discuss in Section 3.3.

In the empirical analysis of Section 5, we exploit Proposition 8 to implement a *mixed strategy* that invests in either the optimal or constrained strategy according to γ:[13]

$$\hat{\boldsymbol{w}}^{mix} = \begin{cases} \hat{\boldsymbol{w}}^{tz} & \text{if } \gamma \in \left[\underline{\gamma}_{ew}, \overline{\gamma}_{ew}\right], \\ \hat{\boldsymbol{w}}^{opt} & \text{otherwise.} \end{cases} \tag{31}$$

The usefulness of the convexity constraint we demonstrate in this section is not restricted to the combination of the SMV and EW portfolios. In Appendix IA.2.1, we derive the con-

[13]Similarly, in Appendix IA.2.3 we exploit Proposition 8 to propose a mixed strategy that invests in the optimal two-fund rule of Kan and Zhou (2007), the constrained strategy, or our estimated optimal strategy.

strained portfolio combination for the three-fund rule of Kan and Zhou (2007) that combines the SMV and SGMV portfolios, which we show outperforms the optimal combination for some degrees of risk aversion. We exploit this result to construct a mixed strategy similar to that in (31). In Appendix IA.2.2, we show that the optimal combination of the SMV and SGMV portfolios without a risk-free asset in Kan, Wang, and Zhou (2021) actually outperforms the three-fund rule of Kan and Zhou (2007) for some risk-aversion levels due to the convexity constraint; that is, *investing in the risk-free asset can hurt performance*.

5 Empirical analysis

In this section, we evaluate the out-of-sample performance of the portfolio strategies we introduce in this paper. We discuss the datasets and portfolio strategies in Section 5.1, the out-of-sample performance measure in Section 5.2, and the results in Section 5.3. We provide additional empirical results in Appendix IA.3.

5.1 Datasets and portfolio strategies

We consider four commonly used datasets of monthly excess returns that we list in Table 1: three datasets of characteristic-sorted portfolios and one of industry-sorted portfolios.

We evaluate the out-of-sample performance of eight portfolio strategies that we list in Table 2 with their abbreviation. Among them, two are novel strategies we introduce in this paper. First, the optimal three-fund combination strategy in Section 3 (OPT3F). Second, the mixed combination strategy in Section 4 (MIX3F).[14]

The remaining six strategies are alternative portfolio combination rules. First and most relevant to us, the constrained three-fund combination strategy of Tu and Zhou (2011) in Section 2.3 (TZ3F). Second and third, the optimal two-fund and three-fund rules in Kan and Zhou (2007) (KZ2F and KZ3F). Fourth, the optimal two-fund combination in Kan, Wang, and Zhou (2021) without a risk-free asset (KWZ). Fifth and sixth, the optimal combination

[14] In theory, for a fixed risk-aversion coefficient γ, the mixed strategy is equal to either the optimal or the constrained strategy, depending on whether $\gamma \in \left[\underline{\gamma}_{ew}, \overline{\gamma}_{ew}\right]$ in (30) or not. However, this interval is estimated using rolling windows, and thus MIX3F switches between the two strategies over time, even for a fixed γ.

of the EW or SGMV portfolio with the risk-free asset, which correspond to $(\gamma_{ew}/\gamma)\boldsymbol{w}_{ew}$ and $(\mu_g/(c\gamma))\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}$, respectively (EWRF and GMVRF). The last two combinations generally outperform their more traditional fully invested counterpart and are more consistent with the other strategies that also invest in the risk-free asset, except for KWZ.

Although we use the sample covariance matrix (6) in our theory, we implement the portfolio strategies both with the sample estimate and the linear shrinkage estimate of Ledoit and Wolf (2004).[15] This is because Kan, Wang, and Zhou (2021) and Lassance, Martín-Utrera, and Simaan (2022) find that using both optimal combination coefficients and shrinkage estimates of the covariance matrix delivers better performance than using each in isolation.

We need to estimate several parameters to determine the optimal combination coefficients in the eight strategies. We estimate the parameter ψ^2 via the adjusted estimator in Kan and Wang (2021) and the parameters θ^2 and $\psi_g^2 = \theta^2 - \mu_g^2/\sigma_g^2$ via the adjusted estimators in Kan and Zhou (2007), where μ_g and σ_g^2 are the mean excess return and variance of the GMV portfolio, respectively. Regarding μ_{ew}, σ_{ew}^2, μ_g, and σ_g^2, we estimate them by plugging in the sample estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, as in Kan and Zhou (2007) and Tu and Zhou (2011).

Finally, we consider a wide range of values for the risk-aversion coefficient γ to evaluate how the portfolio strategies perform and compare for different types of investors. Specifically, consistent with the values used by DeMiguel, Garlappi, and Uppal (2009) and Martellini and Ziemann (2010), we consider $\gamma = 3, 5, 10$, and 15.

5.2 Out-of-sample performance measure

To measure the out-of-sample performance of the eight portfolio strategies in Section 5.1, we implement a classical rebalancing methodology as in DeMiguel, Garlappi, and Uppal (2009) and Tu and Zhou (2011). Specifically, at the end of month t we estimate the portfolio strategy k using an estimation window composed of the T previous months, and we compute its out-of-sample return using the return data in month $t+1$. We consider $T = 120$ and 240 months. This procedure is repeated iteratively until the end of the sample, which gives a time series of $\tau - T$ out-of-sample gross returns $r_{gross,k,t}$ for each strategy k, where τ is the

[15] We also consider the nonlinear shrinkage estimate of Ledoit and Wolf (2020) and find the results are similar to those obtained with linear shrinkage. Therefore, we do not report these results for conciseness.

total number of months in the dataset. We then compute the net out-of-sample returns as $r_{net,k,t} = r_{gross,k,t}$ if $t = T + 1$ and

$$r_{net,k,t} = (1 + r_{gross,k,t})(1 - p \times \text{turnover}_{k,t-1}) - 1 \quad \text{if} \quad t = T + 2, \ldots, \tau, \tag{32}$$

where p is the proportional cost required to rebalance the portfolio and

$$\text{turnover}_{k,t} = \sum_{i=1}^{N} |w_{i,k,t} - w_{i,k,(t-1)^+}|, \quad t = T + 1, \ldots, \tau, \tag{33}$$

with $w_{i,k,t}$ the weight of asset i in month t and $w_{i,k,(t-1)^+}$ the prior-month weight before rebalancing in month t. We set $p = 10$ basis points as in Ao, Li, and Zheng (2019), which is in line with Engle, Ferstenberg, and Russell (2012) who find that the average cost level for NYSE stocks is 8.8 basis points. Finally, we compare the different portfolio strategies in terms of *annualized out-of-sample utility net of transaction costs*,

$$U_k = 12 \times \left(\hat{\mu}_k - \frac{\gamma}{2}\hat{\sigma}_k^2\right), \tag{34}$$

where $\hat{\mu}_k$ and $\hat{\sigma}_k^2$ are the sample mean and variance of $r_{net,k,t}$ in (32). We consider the out-of-sample utility because this is the criterion that the eight combination strategies we consider are designed to optimize. In Appendix IA.3.1, we also report the out-of-sample Sharpe ratio.

5.3 Discussion of results

The out-of-sample performance of the portfolio strategies listed in Table 2 is reported in Table 3 for the sample covariance matrix and Table 4 for the shrinkage covariance matrix. We also report the average monthly turnover in Table 5, only when $T = 120$ for conciseness.

In the vast majority of cases, we find that exploiting the shrinkage covariance matrix of Ledoit and Wolf (2004) to estimate the portfolio strategies in Table 4 delivers better performance than exploiting the sample covariance matrix $\widehat{\mathbf{\Sigma}}$, defined in (6), in Table 3. Therefore, we recommend investors to rely on a shrinkage covariance matrix, even if the sample one underlies the theory. However, the ranking of the different strategies remains similar, and therefore the main conclusions we discuss below are applicable to both cases.

The main comparison concerns the three-fund rules that combine the sample tangent portfolio, the EW portfolio, and the risk-free asset. The results are consistent with our theoretical predictions in Sections 3 and 4. Specifically, when γ is rather small and thus close to the estimated values of γ_{ew}, the constrained strategy TZ3F of Tu and Zhou (2011) outperforms the optimal strategy OPT3F, in line with Proposition 8. This happens when $\gamma = 3$ when using the sample covariance matrix, and $\gamma = 3$ and 5 when using the shrinkage covariance matrix. This result shows that the convexity constraint can indeed help performance in practice by alleviating the impact of estimation errors on combination coefficients. Nonetheless, the utility loss with OPT3F is typically small.

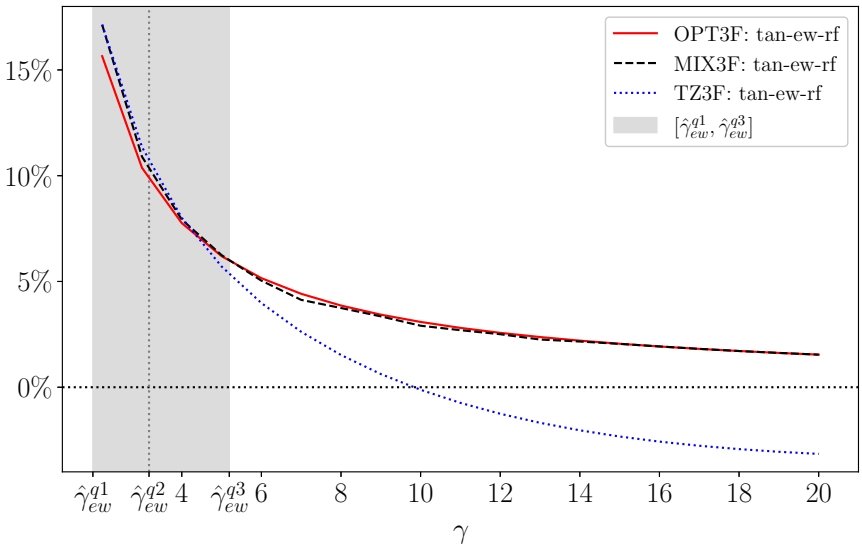
Although TZ3F performs well when γ is small, it quickly deteriorates as γ increases, in which case it is largely outperformed by OPT3F. Take, for instance, the sample covariance matrix, $T = 120$, and $\gamma = 10$. Then, TZ3F delivers an annualized net utility of -0.12, 1.66, -3.02, and -0.33 for the four datasets, versus 3.09, 2.61, 0.67, and 1.91 for OPT3F.

In line with Propositions 5 and 6 we also note that as γ increases TZ3F often delivers negative utility values, meaning that it is outperformed by the risk-free asset, and it is often outperformed by the optimal two-fund rule of Kan and Zhou (2007) (KZ2F). The out-of-sample utility with OPT3F is on the contrary *always* positive and larger than that of KZ2F. This means that the OPT3F strategy is consistently valuable to investors on a risk-adjusted basis, and profits from adding the EW portfolio in the combination.

Moreover, we find in Table 5 that OPT3F delivers a smaller turnover than that of TZ3F. This can be explained from Proposition 7, which shows that TZ3F delivers more extreme weights for many degrees of risk aversion. This larger turnover hurts the performance net of proportional transaction costs. Moreover, OPT3F delivers a lower turnover than KZ2F thanks to the investment in the low-turnover EW portfolio.

The mixed strategy introduced in Section 4 (MIX3F) is designed to take the best out of the OPT3F and TZ3F strategies, and our results show that this objective is accomplished. Indeed, MIX3F is close to and generally better than OPT3F when γ is rather small ($\gamma = 3$ and 5), and still outperforms TZ3F when γ gets larger. This means that MIX3F is a safer approach than relying exclusively on either OPT3F or TZ3F. However, because of estimation error in the interval of γ that defines when TZ3F outperforms OPT3F, $\left[\underline{\gamma}_{ew}, \overline{\gamma}_{ew}\right]$ in (30),

Figure 6: Annualized net out-of-sample utility of three-fund strategies



Notes. This figure depicts the annualized net out-of-sample utility as a function of the risk-aversion coefficient γ for three different three-fund strategies described in Table 2: optimal strategy (OPT3F, solid red), mixed strategy (MIX3F, dashed black), and constrained strategy (TZ3F, dotted blue). The figure is constructed following the empirical methodology described in Section 5, for the 25SBTM dataset, a sample size $T = 120$, and under the sample covariance matrix. The net out-of-sample utility is computed using proportional transaction costs of 10 basis points as in Ao, Li, and Zheng (2019). The grey region depicts the first, second, and third quartiles of the estimated parameter γ_{ew} defined in (10).

MIX3F is not able to perfectly replicate the performance of the best-performing strategy depending on γ. We illustrate the performance of the OPT3F, TZ3F, and MIX3F strategies in Figure 6 for the 25SBTM dataset, $T = 120$, and the sample covariance matrix. The figure shows that the performance of MIX3F switches from that of TZ3F to that of OPT3F for γ around the middle of the first and third quartiles of estimated values of γ_{ew}.

We turn next to the comparison with the optimal three-fund rule of Kan and Zhou (2007) (KZ3F) and its fully invested counterpart in Kan, Wang, and Zhou (2021) (KWZ). When the risk aversion γ is rather large ($\gamma = 10$ and 15), KZ3F nearly systematically outperforms KWZ, because KWZ cannot invest in the risk-free asset. However, the contrary often happens when $\gamma = 3$ and 5, and even systematically so when $T = 240$. We demonstrate in Appendix IA.2.2 that this result can be explained with a similar reasoning to that in Section 4. Specifically, Kan, Wang, and Zhou (2021) impose a convexity constraint on the combination coefficients to ensure that KWZ is fully invested in risky assets, which alleviates

the impact of estimation errors in combination coefficients relative to KZ3F. As a result, KWZ can deliver better performance than KZ3F for some degrees of risk aversion.[16]

The comparison of the optimal combination of the SMV and EW portfolios (OPT3F) with the optimal combination of the SMV and SGMV portfolios in Kan and Zhou (2007) (KZ3F) delivers another surprising insight. When looking at the EWRF and GMVRF strategies, we see that GMVRF performs largely better overall. Specifically, it outperforms for all datasets when $T = 240$ and for all datasets except 30IND when $T = 120$. Therefore, one would expect that KZ3F should also largely outperform OPT3F. However, this is not the case. When $T = 240$, KZ3F does better but their performances are quite close. When $T = 120$, OPT3F is actually even better than KZ3F. To give a concrete example, consider the shrinkage covariance matrix, $T = 120$, $\gamma = 3$, and the 25SBTM dataset. Then, GMVRF delivers an EU of 5.79 and EWRF only 3.29, but still, OPT3F delivers an EU of 13.25 while KZ3F only delivers 12.13. Intuitively, this result should be explained by different diversification properties that arise when combining SMV with EW or SGMV. To confirm this intuition, we derive in Appendix IA.2.5 a closed-form expression for the correlation between the out-of-sample return of the SMV portfolio and that of either the SGMV portfolio or the EW portfolio. By calibrating this correlation to our four datasets, we find that the correlation with the EW portfolio is systematically much smaller, which is in line with the empirically observed correlations.[17] Therefore, there are larger out-of-sample diversification gains from combining the SMV and EW portfolios, which helps explain the puzzling results above.[18]

Finally, we compare our strategies with the two naive benchmarks, that is, the combination of the EW and SGMV portfolios with the risk-free asset, EWRF and GMVRF, respectively. For the three datasets of characteristic-sorted portfolios, these two strategies are largely outperformed by the portfolios that also invest in the SMV portfolio. In particu-

[16] Specifically, KWZ outperforms KZ3F when γ belongs to an interval centered around $\frac{T-N-1}{T-2}\gamma_{tan}$.

[17] For example, for the 25SBTM dataset, the empirical correlation between the out-of-sample returns of the SMV and EW portfolios is 0.18 when $T = 120$ and 0.23 when $T = 240$. In contrast, the empirical correlation between SMV and SGMV is twice larger: 0.37 for $T = 120$ and 0.46 for $T = 240$.

[18] One could hope to profit from both the EW and SGMV portfolios by combining them both with the SMV portfolio. However, we derive such an optimal four-fund portfolio in Appendix IA.2.4 and find that it is generally outperformed by our optimal three-fund rule that does not invest in SGMV. This can be explained because the theoretical gain from adding SGMV is small and thus is offset by the estimation risk coming from having to estimate an additional combination coefficient.

lar, when using the shrinkage covariance matrix, our OPT3F strategy outperforms the two naive benchmarks for all risk aversion γ and sample size T we consider, and in a substantial manner. However, combining with SMV is generally detrimental to out-of-sample performance for the 30IND dataset. The explanation for these results is that while the GMVRF and EWRF portfolios are far from being efficient for the characteristic-sorted datasets, there is only little cross-sectional variation in sample mean returns for industry-sorted portfolios (Kirby and Ostdiek, 2012) and thus GMVRF and EWRF are close to being efficient for the 30IND dataset. In that case, the small potential gain from combining with SMV cannot compensate for the increased estimation risk. To elaborate on this explanation, we depict in Figure 7 the time series of parameter ψ^2, defined in (11) as the difference between the maximum squared Sharpe ratio and that of the EW portfolio, which we estimate using the adjusted estimator of Kan and Wang (2021).[19] The figure shows that it is for the 30IND dataset that ψ^2 is smallest and thus EWRF most efficient. Similarly, we find in unreported results that GMVRF is more efficient for the 30IND dataset than for the other three datasets.

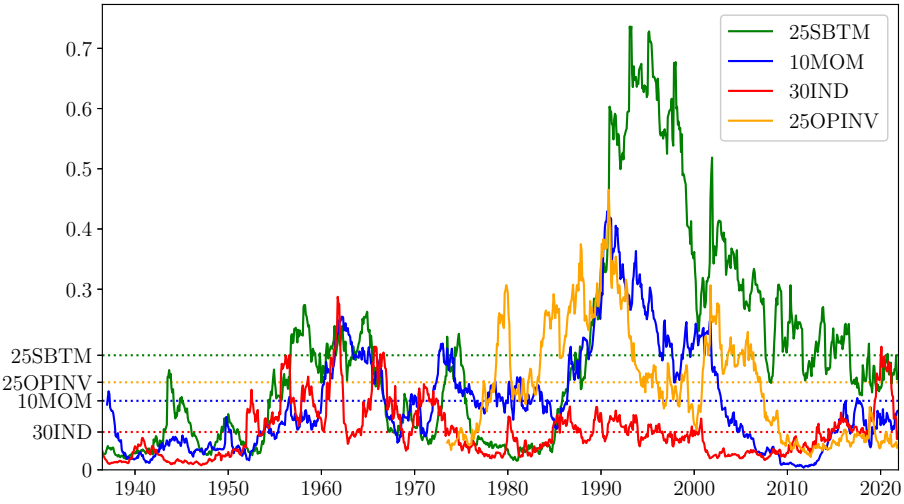
6 Conclusion

The sample mean-variance (SMV) portfolio is notoriously deficient and typically underperforms the equally weighted (EW) portfolio out of sample. Nonetheless, Tu and Zhou (2011) show that investors should not completely abandon the SMV portfolio because it is possible to obtain consistently good performance by *combining* the SMV and EW portfolios together.

In this paper, we explore the consequences of a seemingly natural *convexity constraint* that Tu and Zhou impose on the combination coefficients: they have to sum to one. We show that this constraint is unnecessary because the risk-free asset is part of the investment set, and it has several undesirable consequences relative to the optimal unconstrained portfolio combination we derive. In particular, it leads to an overinvestment in the SMV portfolio, a worse performance than the risk-free asset for sufficiently risk-averse investors, and more extreme weights on the risky assets. Our empirical analysis corroborates these results.

[19] The parameter ψ^2 is a sensible measure of inefficiency of the EWRF portfolio because one can show that the expected out-of-sample utility loss of the EWRF portfolio, $U^\star - EU\big(\frac{\gamma_{ew}}{\gamma}\boldsymbol{w}_{ew}\big)$, is equal to $\psi^2/(2\gamma)$.

Figure 7: Estimated inefficiency of the equally weighted portfolio



Notes. This figure depicts the time series of the estimated values of ψ^2, defined in (11) as the difference between the maximum squared Sharpe ratio and that of the equally weighted portfolio. This parameter is estimated for each dataset listed in Table 1 using the adjusted estimator of Kan and Wang (2021). The estimate for a given month is obtained from the previous $T = 120$ monthly returns. The dotted lines represent the mean of the time series for each dataset.

However, the convexity constraint points to a new way of improving out-of-sample performance when combination coefficients are estimated, in which case *imposing the wrong constraint can help*. Indeed, the convexity constraint acts as a bound constraint on combination coefficients and thus the constrained coefficients are less sensitive to estimation errors, which helps improve performance relative to the optimal combination of the SMV and EW portfolios for some degrees of risk aversion. This novel insight is quite general. We show that it can be used to improve performance when combining the SMV and global-minimum-variance (GMV) portfolios, and to explain why investing in the risk-free asset hurts performance for some investors. There is a large literature on the benefits of weight constraints for combining assets and, given our results, an interesting avenue for future research is to exploit them to combine estimated portfolios and alleviate estimation errors in combination coefficients.

Finally, the GMV portfolio is often preferred to the EW portfolio in the literature as a naive investment choice because it is located on the sample efficient frontier and generally performs better. However, we demonstrate theoretically and empirically that the correlation between the out-of-sample return of the SMV portfolio and the EW portfolio is substantially

smaller than that with the GMV portfolio. Therefore, the EW portfolio can be preferable for portfolio combinations due to larger out-of-sample diversification gains.

Tables

Table 1: List of datasets considered in the empirical analysis

Dataset	N	Time period	Abbreviation
25 portfolios formed on size and book-to-market	25	07/1926 - 12/2021	25SBTM
10 momentum portfolios	10	01/1927 - 12/2021	10MOM
30 industry portfolios	30	07/1926 - 12/2021	30IND
25 portfolios formed on operating profitability and investment	25	07/1963 - 12/2021	25OPINV

Notes. This table lists the datasets of monthly excess returns we use in the empirical analysis of Section 5. All data are downloaded from Kenneth French's website.

Table 2: List of portfolio strategies considered in the empirical analysis

Abbreviation	Description
Portfolio strategies introduced in the paper	
OPT3F: tan-ew-rf	Optimal combination of the sample tangent portfolio, equally weighted portfolio, and risk-free asset. See Proposition 2.
MIX3F: tan-ew-rf	Mixed strategy combining the constrained and optimal three-fund portfolio combinations. See Equation (31).
Benchmark portfolio strategies	
TZ3F: tan-ew-rf	Constrained combination of the sample tangent portfolio, equally weighted portfolio, and risk-free asset in Tu and Zhou (2011).
KZ2F: tan-rf	Optimal combination of the sample tangent portfolio and risk-free asset in Kan and Zhou (2007).
KZ3F: tan-gmv-rf	Optimal combination of the sample tangent portfolio, sample global-minimum-variance portfolio, and risk-free asset in Kan and Zhou (2007).
KWZ: tan-gmv	Optimal combination of the sample tangent portfolio and sample global-minimum-variance portfolio with no-risk free asset in Kan, Wang, and Zhou (2021).
EWRF	Optimal combination of the equally weighted portfolio and the risk-free asset.
GMVRF	Optimal combination of the sample global-minimum-variance portfolio and the risk-free asset.

Notes. This table lists the portfolio strategies we use in the empirical analysis of Section 5. All strategies are estimated with the sample covariance matrix in (6) or the shrinkage estimator of Ledoit and Wolf (2004).

Table 3: Annualized net out-of-sample utility (sample covariance matrix)

Dataset	$T = 120$				$T = 240$			
	25SBTM	10MOM	30IND	25OPINV	25SBTM	10MOM	30IND	25OPINV
Panel (a): $\gamma = 3$								
Proposed portfolios								
OPT3F: tan-ew-rf	10.37	8.71	2.25	6.38	8.29	9.30	0.27	8.54
MIX3F: tan-ew-rf	10.90	8.74	2.81	5.81	8.79	9.63	1.03	9.53
Benchmark portfolios								
TZ3F: tan-ew-rf	11.36	8.99	3.25	8.47	9.22	10.71	2.43	8.57
KZ2F: tan-rf	9.13	8.02	0.03	5.35	7.48	9.00	−1.84	6.29
KZ3F: tan-gmv-rf	8.85	9.81	1.15	6.43	8.98	10.75	0.63	9.47
KWZ: tan-gmv	9.54	9.16	1.97	6.19	10.57	12.53	2.83	9.58
EWRF	3.29	2.71	3.78	2.09	2.34	1.30	1.78	2.45
GMVRF	5.80	4.53	2.18	7.74	6.59	3.33	2.08	9.26
Panel (b): $\gamma = 5$								
Proposed portfolios								
OPT3F: tan-ew-rf	6.20	5.22	1.34	3.82	4.94	5.58	0.16	5.12
MIX3F: tan-ew-rf	6.25	5.52	1.21	3.65	4.71	5.33	0.16	4.67
Benchmark portfolios								
TZ3F: tan-ew-rf	5.71	5.16	1.08	4.71	5.16	6.61	0.88	4.86
KZ2F: tan-rf	5.45	4.80	0.02	3.20	4.45	5.40	−1.11	3.76
KZ3F: tan-gmv-rf	5.28	5.88	0.68	3.85	5.36	6.45	0.38	5.68
KWZ: tan-gmv	6.45	5.78	1.28	4.60	7.90	8.51	2.21	7.41
EWRF	1.97	1.62	2.27	1.25	1.41	0.78	1.07	1.47
GMVRF	3.47	2.72	1.31	4.64	3.95	2.00	1.24	5.55
Panel (c): $\gamma = 10$								
Proposed portfolios								
OPT3F: tan-ew-rf	3.09	2.61	0.67	1.91	2.45	2.79	0.08	2.56
MIX3F: tan-ew-rf	2.91	2.47	0.71	1.73	2.34	2.70	0.07	2.55
Benchmark portfolios								
TZ3F: tan-ew-rf	−0.12	1.66	−3.02	−0.33	1.06	2.74	−2.61	0.53
KZ2F: tan-rf	2.71	2.40	0.01	1.60	2.21	2.70	−0.55	1.88
KZ3F: tan-gmv-rf	2.63	2.94	0.34	1.92	2.67	3.22	0.19	2.84
KWZ: tan-gmv	0.64	−0.74	−1.97	−0.22	2.96	2.04	−0.70	2.38
EWRF	0.99	0.81	1.13	0.63	0.70	0.39	0.53	0.74
GMVRF	1.73	1.36	0.65	2.32	1.97	1.00	0.62	2.77
Panel (d): $\gamma = 15$								
Proposed portfolios								
OPT3F: tan-ew-rf	2.06	1.74	0.45	1.27	1.63	1.86	0.05	1.70
MIX3F: tan-ew-rf	2.05	1.67	0.38	1.27	1.63	1.86	0.05	1.70
Benchmark portfolios								
TZ3F: tan-ew-rf	−2.33	0.38	−5.13	−2.70	−0.45	1.41	−3.80	−1.10
KZ2F: tan-rf	1.81	1.60	0.00	1.06	1.47	1.80	−0.37	1.25
KZ3F: tan-gmv-rf	1.75	1.96	0.23	1.28	1.78	2.15	0.12	1.89
KWZ: tan-gmv	−4.58	−6.70	−5.65	−5.26	−1.48	−3.42	−3.98	−2.52
EWRF	0.66	0.54	0.76	0.42	0.47	0.26	0.36	0.49
GMVRF	1.15	0.90	0.43	1.54	1.31	0.67	0.41	1.85

Notes. This table reports the annualized net out-of-sample utility in percentage points for the portfolio strategies described in Table 2 when using the sample covariance matrix, according to the methodology in Section 5. The net out-of-sample utility is computed using proportional transaction costs of 10 basis points as in Ao, Li, and Zheng (2019). We consider two sample sizes, $T = 120$ and $T = 240$, and risk-aversion coefficients $\gamma = 3, 5, 10$, and 15.

Table 4: Annualized net out-of-sample utility (shrinkage covariance matrix)

Dataset	$T = 120$				$T = 240$			
	25SBTM	10MOM	30IND	25OPINV	25SBTM	10MOM	30IND	25OPINV
Panel (a): $\gamma = 3$								
Proposed portfolios								
OPT3F: tan-ew-rf	13.25	11.72	1.74	9.28	12.40	11.00	0.01	9.09
MIX3F: tan-ew-rf	14.02	11.94	2.34	9.01	12.93	11.35	0.77	10.16
Benchmark portfolios								
TZ3F: tan-ew-rf	15.23	12.78	3.05	12.07	13.64	12.42	2.22	9.31
KZ2F: tan-rf	13.33	11.60	0.12	9.90	12.20	10.94	−2.00	7.40
KZ3F: tan-gmv-rf	12.13	12.71	0.85	9.51	11.99	12.30	0.51	9.23
KWZ: tan-gmv	12.35	12.45	2.64	8.91	13.19	14.08	3.14	10.04
EWRF	3.29	2.71	3.78	2.09	2.34	1.30	1.78	2.45
GMVRF	5.79	4.25	2.15	8.66	5.27	3.20	2.18	8.44
Panel (b): $\gamma = 5$								
Proposed portfolios								
OPT3F: tan-ew-rf	7.94	7.03	1.04	5.56	7.43	6.60	0.01	5.45
MIX3F: tan-ew-rf	8.02	7.48	0.92	5.51	7.26	6.43	0.00	5.01
Benchmark portfolios								
TZ3F: tan-ew-rf	8.61	7.67	1.10	6.66	8.05	7.70	0.79	5.16
KZ2F: tan-rf	8.00	6.96	0.07	5.93	7.31	6.56	−1.20	4.43
KZ3F: tan-gmv-rf	7.26	7.63	0.50	5.69	7.18	7.38	0.30	5.53
KWZ: tan-gmv	8.57	7.99	2.22	6.81	9.52	9.50	2.65	7.76
EWRF	1.97	1.62	2.27	1.25	1.41	0.78	1.07	1.47
GMVRF	3.46	2.55	1.29	5.19	3.15	1.92	1.31	5.06
Panel (c): $\gamma = 10$								
Proposed portfolios								
OPT3F: tan-ew-rf	3.97	3.51	0.52	2.78	3.71	3.30	0.00	2.72
MIX3F: tan-ew-rf	3.78	3.36	0.55	2.55	3.59	3.20	−0.01	2.72
Benchmark portfolios								
TZ3F: tan-ew-rf	2.47	3.32	−2.81	0.69	3.07	3.48	−2.72	0.55
KZ2F: tan-rf	4.00	3.48	0.03	2.97	3.65	3.28	−0.60	2.21
KZ3F: tan-gmv-rf	3.63	3.81	0.24	2.84	3.59	3.69	0.15	2.76
KWZ: tan-gmv	2.72	0.92	−0.54	2.00	3.97	2.72	−0.02	2.78
EWRF	0.99	0.81	1.13	0.63	0.70	0.39	0.53	0.74
GMVRF	1.73	1.27	0.64	2.59	1.57	0.96	0.65	2.53
Panel (d): $\gamma = 15$								
Proposed portfolios								
OPT3F: tan-ew-rf	2.65	2.34	0.35	1.85	2.47	2.20	0.00	1.81
MIX3F: tan-ew-rf	2.64	2.27	0.26	1.85	2.47	2.20	0.00	1.81
Benchmark portfolios								
TZ3F: tan-ew-rf	0.30	1.78	−5.02	−1.80	1.34	2.01	−3.97	−1.12
KZ2F: tan-rf	2.67	2.32	0.02	1.98	2.43	2.18	−0.40	1.47
KZ3F: tan-gmv-rf	2.42	2.54	0.16	1.89	2.39	2.46	0.10	1.84
KWZ: tan-gmv	−2.07	−4.98	−3.77	−2.66	−0.55	−2.72	−3.10	−1.98
EWRF	0.66	0.54	0.76	0.42	0.47	0.26	0.36	0.49
GMVRF	1.15	0.85	0.43	1.73	1.05	0.64	0.43	1.68

Notes. This table reports the annualized net out-of-sample utility in percentage points for the portfolio strategies described in Table 2 when using the shrinkage covariance matrix of Ledoit and Wolf (2004), according to the methodology in Section 5. The net out-of-sample utility is computed using proportional transaction costs of 10 basis points as in Ao, Li, and Zheng (2019). We consider two sample sizes, $T = 120$ and $T = 240$, and risk-aversion coefficients $\gamma = 3, 5, 10$, and 15.

Table 5: Average monthly turnover ($T = 120$)

Dataset	Sample $\boldsymbol{\Sigma}$				Linear shrinkage $\boldsymbol{\Sigma}$			
	25SBTM	10MOM	30IND	25OPINV	25SBTM	10MOM	30IND	25OPINV
Panel (a): $\gamma = 3$								
Proposed portfolios								
OPT3F: tan-ew-rf	3.07	1.36	0.94	1.91	1.62	0.86	0.79	1.47
MIX3F: tan-ew-rf	3.12	1.40	0.96	1.96	1.65	0.89	0.82	1.50
Benchmark portfolios								
TZ3F: tan-ew-rf	3.25	1.45	1.08	2.00	1.71	0.92	0.91	1.55
KZ2F: tan-rf	3.26	1.44	1.10	2.05	1.75	0.92	0.94	1.58
KZ3F: tan-gmv-rf	3.35	1.39	1.32	2.11	1.77	0.88	1.09	1.61
KWZ: tan-gmv	2.75	1.35	1.08	1.61	1.46	0.85	0.88	1.25
EWRF	0.06	0.06	0.07	0.07	0.06	0.06	0.07	0.07
GMVRF	2.00	0.55	0.89	1.41	1.06	0.37	0.71	1.05
Panel (b): $\gamma = 5$								
Proposed portfolios								
OPT3F: tan-ew-rf	1.84	0.82	0.56	1.15	0.97	0.52	0.48	0.88
MIX3F: tan-ew-rf	1.87	0.84	0.59	1.17	0.99	0.54	0.50	0.90
Benchmark portfolios								
TZ3F: tan-ew-rf	2.07	0.91	0.71	1.25	1.06	0.57	0.58	0.97
KZ2F: tan-rf	1.96	0.86	0.66	1.23	1.05	0.55	0.56	0.95
KZ3F: tan-gmv-rf	2.01	0.84	0.79	1.26	1.06	0.53	0.65	0.96
KWZ: tan-gmv	1.79	0.85	0.76	1.08	0.94	0.53	0.59	0.81
EWRF	0.03	0.04	0.04	0.04	0.03	0.04	0.04	0.04
GMVRF	1.20	0.33	0.53	0.85	0.64	0.22	0.42	0.63
Panel (c): $\gamma = 10$								
Proposed portfolios								
OPT3F: tan-ew-rf	0.92	0.41	0.28	0.57	0.49	0.26	0.24	0.44
MIX3F: tan-ew-rf	0.93	0.42	0.29	0.58	0.49	0.26	0.24	0.44
Benchmark portfolios								
TZ3F: tan-ew-rf	1.31	0.53	0.55	0.79	0.65	0.33	0.42	0.62
KZ2F: tan-rf	0.98	0.43	0.33	0.62	0.52	0.28	0.28	0.47
KZ3F: tan-gmv-rf	1.00	0.42	0.40	0.63	0.53	0.26	0.33	0.48
KWZ: tan-gmv	1.15	0.51	0.55	0.74	0.58	0.31	0.39	0.52
EWRF	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Panel (d): $\gamma = 15$								
Proposed portfolios								
OPT3F: tan-ew-rf	0.61	0.27	0.19	0.38	0.32	0.17	0.16	0.29
MIX3F: tan-ew-rf	0.62	0.27	0.19	0.38	0.32	0.17	0.16	0.29
Benchmark portfolios								
TZ3F: tan-ew-rf	1.06	0.40	0.52	0.67	0.54	0.25	0.39	0.53
KZ2F: tan-rf	0.65	0.29	0.22	0.41	0.35	0.18	0.19	0.32
KZ3F: tan-gmv-rf	0.67	0.28	0.26	0.42	0.35	0.18	0.22	0.32
KWZ: tan-gmv	0.97	0.41	0.50	0.66	0.48	0.25	0.34	0.45
EWRF	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
GMVRF	0.40	0.11	0.18	0.28	0.21	0.07	0.14	0.21

Notes. This table reports the average monthly turnover for the portfolio strategies described in Table 2 when using either the sample covariance matrix or the shrinkage covariance matrix of Ledoit and Wolf (2004), according to the methodology in Section 5. The average turnover is computed as the average of the monthly turnover values in (33). We consider a sample size $T = 120$ and risk-aversion coefficients $\gamma = 3, 5, 10$, and 15.

References

Ao, M., Y. Li, and X. Zheng. 2019. Approaching mean-variance efficiency for large portfolios. *The Review of Financial Studies* 32:2890–919.

Ban, G.-Y., N. El Karoui, and A. Lim. 2018. Machine learning and portfolio optimization. *Management Science* 64:1136–54.

Barroso, P., and K. Saxena. 2021. Lest we forget: using out-of-sample forecast errors in portfolio optimization. *The Review of Financial Studies* 35:1222–78.

Behr, P., A. Guettler, and F. Miebs. 2013. On portfolio optimization: Imposing the right constraints. *Journal of Banking and Finance* 37:1232–42.

Best, M., and R. Grauer. 1991. On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results. *The Review of Financial Studies* 4:315–42.

Bloomfield, T., R. Leftwich, and J. Long. 1977. Portfolio strategies and performance. *Journal of Financial Economics* 5:201–18.

Bodnar, T., Y. Okhrin, and N. Parolya. 2021. Optimal shrinkage-based portfolio selection in high dimensions. *Journal of Business and Economic Statistics* (forthcoming).

Branger, N., K. Lučivjanská, and A. Weissensteiner. 2019. Optimal granularity for portfolio choice. *Journal of Empirical Finance* 50:125–46.

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

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies* 22:1915–53.

DeMiguel, V., A. Martín-Utrera, and F. Nogales. 2013. Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking and Finance* 37:3018–34.

———. 2015. Parameter uncertainty in multiperiod portfolio optimization with transaction costs. *Journal of Financial and Quantitative Analysis* 50:1443–71.

Engle, R., R. Ferstenberg, and J. Russell. 2012. Measuring and modeling execution cost and risk. *The Journal of Portfolio Management* 38:14–28.

Frahm, G., and C. Memmel. 2010. Dominating estimators for minimum-variance portfolios. *Journal of Econometrics* 159:289–302.

Frost, P., and J. Savarino. 1988. For better performance: Constrain portfolio weights. *The Journal of Portfolio Management* 15:29–34.

Garlappi, L., R. Uppal, and T. Wang. 2007. Portfolio selection with parameter and model uncertainty: A multi-prior approach. *The Review of Financial Studies* 20:41–81.

Jagannathan, R., and T. Ma. 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance* 58:1651–84.

Jobson, J., and B. Korkie. 1980. Estimation for markowitz efficient portfolios. *Journal of the American Statistical Association* 75:544–54.

Kan, R., and X. Wang. 2021. Optimal portfolio choice with benchmark. *Rotman School of Management Working Paper No. 3760640*.

Kan, R., X. Wang, and G. Zhou. 2021. Optimal portfolio choice with estimation risk: No risk-free asset case. *Management Science* 68:2047–68.

Kan, R., and G. Zhou. 2007. Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis* 42:621–56.

Kirby, C., and B. Ostdiek. 2012. It's all in the timing: Simple active portfolio strategies that outperform naïve diversification. *Journal of Financial and Quantitative Analysis* 47:437–67.

Lassance, N., A. Martín-Utrera, and M. Simaan. 2022. The risk of out-of-sample performance. *Available at SSRN 3855546*.

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

———. 2017. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies* 30:4349–88.

———. 2020. Analytical nonlinear shrinkage of large-dimensional covariance matrices. *The Annals of Statistics* 48:3043–65.

Levy, H., and M. Levy. 2014. The benefits of differential variance-based constraints in portfolio optimization. *European Journal of Operational Research* 234:372–81.

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

Martellini, L., and V. Ziemann. 2010. Improved estimates of higher-order comoments and implications for portfolio selection. *The Review of Financial Studies* 23:1467–502.

Merton, R. 1980. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics* 8:323–61.

Michaud, R. 1989. The markowitz optimization enigma: Is 'optimized' optimal? *The Review of Financial Studies* 45:31–42.

Pflug, G., A. Pichler, and D. Wozabal. 2012. The 1/N investment strategy is optimal under high model ambiguity. *Journal of Banking and Finance* 36:410–7.

Simaan, M., and Y. Simaan. 2019. Rational explanation for rule-of-thumb practices in asset allocation. *Quantitative Finance* 19:2095–109.

Tu, J., and G. Zhou. 2011. Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics* 99:204–15.

Yan, C., and H. Zhang. 2017. Mean-variance versus naïve diversification: The role of mispricing. *Journal of International Financial Markets, Institutions and Money* 48:61–81.

Yen, Y.-M. 2016. Sparse weighted-norm minimum variance portfolios. *Review of Finance* 20:1259–87.

Zhao, Z., O. Ledoit, and H. Jiang. 2021. Risk reduction and efficiency increase in large portfolios: Gross-exposure constraints and shrinkage of the covariance matrix. *Journal of Financial Econometrics* 1–33.

Internet Appendix to

On the optimal combination of naive and mean-variance portfolio strategies

This internet appendix is divided in four sections. In Section IA.1, we describe an experiment to compare the impact of estimation errors on optimal and constrained combination coefficients. In Section IA.2, we provide additional theoretical results. In Section IA.3, we discuss additional empirical results. Finally, in Section IA.4, we detail the proofs of all theoretical results in the main body of the paper and in this appendix.

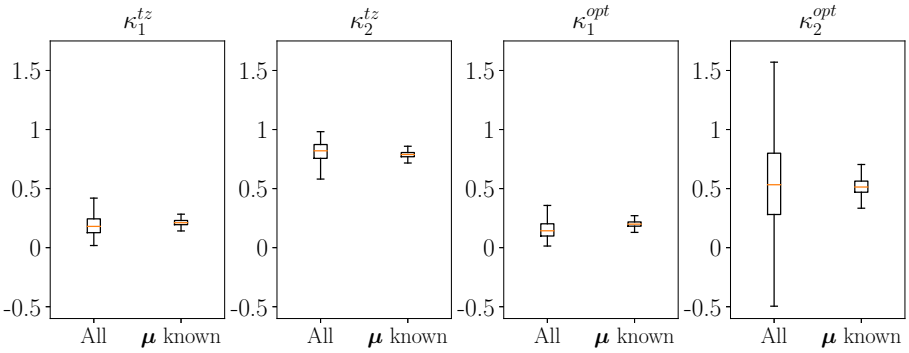
IA.1 Estimation errors in combination coefficients

To establish more precisely the effect of estimation errors in mean returns $\boldsymbol{\mu}$ on combination coefficients discussed in Section 4, consider the experiment in Figure 1(a). We calibrate the mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ to the 25SBTM dataset and simulate 100,000 times $T = 120$ multivariate normal excess returns. For each simulation, we estimate the optimal and constrained combination coefficients with a risk-aversion coefficient $\gamma = 3$, in two different cases. In the first case, all parameters in the coefficients are estimated. We estimate the parameters as in our empirical analysis; see Section 5.1. In the second case, we consider that mean returns $\boldsymbol{\mu}$ are known, and we only plug in the sample estimate of $\boldsymbol{\Sigma}$.

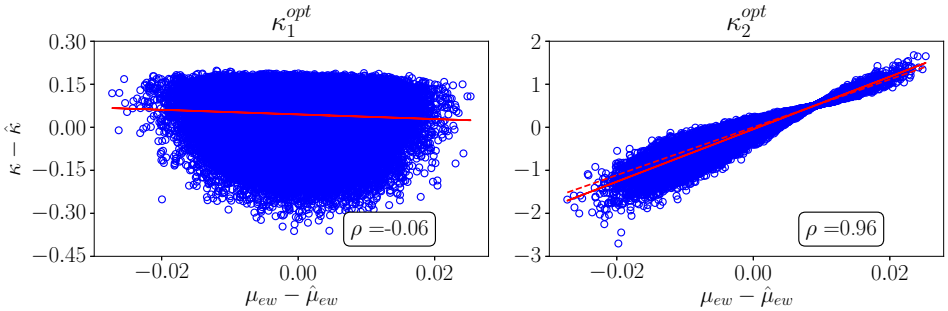
We make two main observations from Figure 1(a). First, mostly errors in $\boldsymbol{\mu}$ matter, as the boxplots are much thinner when only $\boldsymbol{\Sigma}$ is estimated. Second, the boxplots are similar and quite thin for $\kappa_1^{tz} = 1 - \kappa_2^{tz}$ and κ_1^{opt}, which is because they are both bounded between zero and one and estimation errors compensate in their numerator and denominator. In contrast, when $\boldsymbol{\mu}$ is estimated, the boxplot for κ_2^{opt} is substantially wider. Given the formula for $\kappa_2^{opt} = \frac{1}{\gamma}\frac{\mu_{ew}}{\sigma_{ew}^2}(1 - \kappa_1^{opt})$ and that κ_1^{opt} is moderately sensitive to $\boldsymbol{\mu}$, this difference is explained by the proportionality to $\mu_{ew} = \boldsymbol{w}_{ew}'\boldsymbol{\mu}$. Indeed, in Figure 1(b) we depict scatterplots of estimation errors in κ_1^{opt} and κ_2^{opt} as a function of estimation errors in μ_{ew} and we find a near-zero correlation (-6%) for κ_1^{opt} versus a near-perfect correlation (96%) for κ_2^{opt}. Moreover, the best linear fit for estimation errors in κ_2^{opt} is nearly identical to the derivative of κ_2^{opt} with respect to μ_{ew} assuming that κ_1^{opt} is independent of μ_{ew}.

These results confirm that it is justified to focus on the impact of estimation errors in μ_{ew} on κ_2^{opt} to capture the difference in estimation risk between optimal and constrained combination coefficients.

Figure IA.1: Estimation errors in optimal and constrained combination coefficients



(a) Boxplots of estimation errors in all combination coefficients



(b) Scatterplots of estimation errors in optimal combination coefficients

Notes. This figure depicts estimation errors in constrained combination coefficients, κ_1^{tz} and κ_2^{tz} in (15), and optimal combination coefficients, κ_1^{opt} and κ_2^{opt} in (17), when the risk-aversion coefficient $\gamma = 3$. The results are obtained by simulating 100,000 times $T = 120$ asset excess returns from a multivariate normal distribution whose moments are calibrated from a dataset of 25 portfolios of firms sorted on size and book-to-market spanning July 1926 to December 2021. In the boxplots of panel a), we consider two cases to estimate the coefficients: 1) all parameters are estimated as in our empirical analysis (see Section 5.1), and 2) mean returns $\boldsymbol{\mu}$ are known and we only plug in the sample estimate of $\boldsymbol{\Sigma}$. In panel b), we depict how estimation errors in optimal combination coefficients depend on estimation errors in the parameter $\mu_{ew} = \boldsymbol{w}_{ew}'\boldsymbol{\mu}$. The best linear fit in each scatterplot is depicted in solid red. The dashed red line in the right plot of panel b) is the derivative of κ_2^{opt} with respect to μ_{ew} assuming that κ_1^{opt} is independent of μ_{ew}.

IA.2 Additional theoretical results

In this section, we provide theoretical results that complement those in the main body of the paper. First, we provide a mixed strategy for the combination of the SMV and SGMV portfolios. Second, we show that the convexity constraint helps explain why combining portfolios with the risk-free asset can hurt performance relative to fully invested combinations. Third, we introduce a mixed strategy that switches between two-fund and three-fund rules. Fourth, we derive an optimal four-fund portfolio that combines the SMV portfolio with both the SGMV and EW portfolios. Fifth, we compare the correlation between the out-of-sample return of the SMV portfolio and that of the SGMV and EW portfolios. Sixth, we compare the expected out-of-sample Sharpe ratio of the optimal and constrained strategies.

IA.2.1 Mixed strategy for three-fund rule of Kan and Zhou

In Section 4, we introduce a mixed strategy that switches between the optimal and constrained combination of the SMV and EW portfolios. In this section, we derive a similar mixed strategy for the three-fund rule of Kan and Zhou (2007) that combines the SMV and SGMV portfolios. Specifically, we consider the three-fund portfolio combination

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \kappa_1 \hat{\boldsymbol{w}}^{\star} + \kappa_2 \hat{\boldsymbol{w}}_g, \tag{IA1}$$

where $\hat{\boldsymbol{w}}_g = \widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}/(\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1})$ is the SGMV portfolio and $\hat{\boldsymbol{w}}^{\star}$ is the SMV portfolio in (7). This combination is comparable to that in Tu and Zhou (2011) because, just like $\boldsymbol{w}_{ew}$, $\hat{\boldsymbol{w}}_g$ is fully invested in risky assets. Kan and Zhou (2007) consider a similar portfolio combination but with unnormalized weights for the SGMV portfolio.

We introduce the notation

$$\mu_g = \boldsymbol{w}_g'\boldsymbol{\mu}, \quad \sigma_g^2 = \boldsymbol{w}_g'\boldsymbol{\Sigma}\boldsymbol{w}_g, \quad \theta_g = \mu_g/\sigma_g, \quad \text{and} \quad \psi_g^2 = \theta^2 - \theta_g^2 \geq 0. \tag{IA2}$$

In the following proposition, we derive the EU of the portfolio combination in (IA1) as well as that of the optimal and constrained combination coefficients.

Proposition IA.1. *The following holds concerning the three-fund portfolio combination in* (IA1)*:*

1. *The expected out-of-sample utility is*

$$EU(\hat{\mathbf{w}}(\boldsymbol{\kappa})) = \frac{\kappa_1}{\gamma}\theta^2 + \kappa_2\mu_g - \frac{\gamma}{2}\left(\frac{\kappa_1^2}{\gamma^2}(\theta^2+d) + c_1\kappa_2^2\sigma_g^2 + \frac{2c_1\kappa_1\kappa_2}{\gamma}\mu_g\right), \tag{IA3}$$

where

$$c_1 = c\frac{T-N-4}{T-N-2} = \frac{T-2}{T-N-1} \geq 1. \tag{IA4}$$

2. *The combination coefficients maximizing* (IA3) *are*

$$\kappa_1^{opt} = \frac{1}{c}\frac{\psi_g^2}{\psi_g^2 + N/T + \frac{2\theta_g^2}{T-N-2}} \in [0,1] \quad \textit{and} \quad \kappa_2^{opt} = \frac{\gamma_{tan}}{\gamma}(1/c_1 - \kappa_1^{opt}). \tag{IA5}$$

3. *The combination coefficients maximizing* (IA3) *under the constraint* $\kappa_1 + \kappa_2 = 1$ *are*

$$\kappa_1^{const} = \frac{\psi_g^2 + c_1\sigma_g^2(\gamma-\gamma_{tan})^2 + (c_1-1)\mu_g(\gamma-\gamma_{tan})}{c\left(\psi_g^2 + N/T + \frac{2\theta_g^2}{T-N-2}\right) + c_1\sigma_g^2(\gamma-\gamma_{tan})^2} \tag{IA6}$$

and $\kappa_2^{const} = 1 - \kappa_1^{const}$.

4. *The expected out-of-sample utility delivered by the optimal and constrained combinations coefficients is*

$$EU(\hat{\mathbf{w}}(\boldsymbol{\kappa}^{opt})) = U^\star - \frac{d}{2\gamma}\kappa_1^{opt} - \frac{1}{2}(c_1-1)\mu_g\kappa_2^{opt} \tag{IA7}$$

$$EU(\hat{\mathbf{w}}(\boldsymbol{\kappa}^{const})) = U^\star - \frac{d}{2\gamma}\kappa_1^{const} - \frac{1}{2}(c_1-1)\mu_g\kappa_2^{const}. \tag{IA8}$$

The constrained combination coefficients correspond to the optimal ones only when $\kappa_1^{opt} + \kappa_2^{opt} = 1$, which happens when

$$\gamma = \gamma_{tan}\frac{1/c_1 - \kappa_1^{opt}}{1-\kappa_1^{opt}} < \gamma_{tan}. \tag{IA9}$$

For the 25SBTM dataset and $T = 120$, we have $\gamma_{tan} = 3.99$ and $\gamma_{tan}\frac{1/c_1-\kappa_1^{opt}}{1-\kappa_1^{opt}} = 3.00$.

The EU of the optimal combination is always larger than that of the constrained combination when combination coefficients are known. However, similar to the combination with the EW portfolio in Section 4, κ_2^{opt} is unbounded and proportional to γ_{tan}, and therefore is more sensitive to estimation errors in mean returns $\boldsymbol{\mu}$. Therefore, in the next proposition we derive the EU of the optimal strategy when κ_2^{opt} is estimated as

$$\hat{\kappa}_2^{opt} = \frac{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}{\gamma}(1/c_1 - \kappa_1^{opt}), \tag{IA10}$$

and we identify the range of risk aversion γ for which the constrained strategy delivers a larger EU than that of the estimated optimal strategy.

Proposition IA.2. *Let the combination coefficient κ_2^{opt} be estimated by $\hat{\kappa}_2^{opt}$ in (IA10). Then,*

1. *The expected out-of-sample utility of the estimated optimal strategy is*

$$EU(\hat{\boldsymbol{w}}(\kappa_1^{opt}, \hat{\kappa}_2^{opt})) = EU(\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{opt})) - \frac{c_1}{2\gamma T}(1/c_1^2 - (\kappa_1^{opt})^2), \tag{IA11}$$

   where $EU(\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{opt}))$ in (IA7) is the utility when κ_2^{opt} is known.

2. *The constrained strategy delivers a larger expected out-of-sample utility than the estimated optimal strategy, $EU(\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{const})) \geq EU\Big(\hat{\boldsymbol{w}}(\kappa_1^{opt}, \hat{\kappa}_2^{opt})\Big)$, if and only if γ belongs to the following interval:*

$$\gamma \in \left[\gamma_{mid} \pm \frac{1}{\sigma_g}\sqrt{\frac{\gamma_{mid}}{\gamma_{tan}}}\sqrt{\mu_g\gamma_{mid} + \frac{\theta^2(a-d)+ad}{d-c_1a}}\right] = \left[\underline{\gamma}_{mid}, \overline{\gamma}_{mid}\right], \tag{IA12}$$

   where

$$\gamma_{mid} = \gamma_{tan}\frac{d - c_1a}{c_1(d-a) - \theta_g^2(c_1-1)^2}, \tag{IA13}$$

$$a = d\kappa_1^{opt} + (c_1-1)\theta_g^2(1/c_1 - \kappa_1^{opt}) + \frac{c_1}{T}(1/c_1^2 - (\kappa_1^{opt})^2). \tag{IA14}$$

We find empirically that the interval of γ in (IA12) is typically wider than that for the combination with the EW portfolio in (30). For the 25SBTM dataset with $T = 120$, we

find that $[\underline{\gamma}_{mid}, \overline{\gamma}_{mid}] = [2.98 \pm 2.01]$ while $\left[\underline{\gamma}_{ew}, \overline{\gamma}_{ew}\right] = [1.90 \pm 1.65]$. Therefore, the mixed strategy is theoretically preferable for a wider range of investors for this particular portfolio combination. This can be explained because estimation errors in $\hat{\boldsymbol{\mu}}$ are amplified by $\boldsymbol{\Sigma}^{-1}$ in (IA10). Indeed, the standard deviation of $\hat{\gamma}_{ew} = \boldsymbol{w}_{ew}'\hat{\boldsymbol{\mu}}/\sigma_{ew}^2$ is equal to $1/\sqrt{\sigma_{ew}^2 T}$, which is smaller than the standard deviation of $\hat{\gamma}_{tan} = \mathbf{1}'\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\mu}}$ that is equal to $1/\sqrt{\sigma_g^2 T}$.

In Section IA.3, we investigate the empirical performance of the optimal, constrained, and mixed strategies presented in this section.

IA.2.2 Investing in the risk-free asset can hurt performance

We observe in our empirical results in Tables 3 and 4 that the optimal combination of the SMV and SGMV portfolios without a risk-free asset in Kan, Wang, and Zhou (2021) often outperforms the combination with a risk-free asset in Kan and Zhou (2007) when the risk-aversion coefficient γ is rather small ($\gamma = 3$ and 5). In this section, we show this can be explained due to the convexity constraint, similar to the results in Sections 4 and IA.2.1.

Kan, Wang, and Zhou (2021) combine the fully invested SMV and SGMV portfolios, which corresponds to the portfolio combination

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \kappa_1 \hat{\boldsymbol{w}}_g + \kappa_2(\hat{\boldsymbol{w}}_g + \hat{\boldsymbol{w}}_z), \tag{IA15}$$

where $\hat{\boldsymbol{w}}_g = \widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}/(\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1})$ is the SGMV portfolio and

$$\hat{\boldsymbol{w}}_z = \frac{1}{\gamma}\widehat{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\mu}} - \hat{\mu}_g \mathbf{1}). \tag{IA16}$$

is a zero-cost portfolio ($\mathbf{1}'\hat{\boldsymbol{w}}_z = 0$). To ensure that $\hat{\boldsymbol{w}}(\boldsymbol{\kappa})$ is fully invested in risky assets, the authors impose the convexity constraint $\kappa_1 + \kappa_2 = 1$, so that the combination depends on a single combination coefficient κ:

$$\hat{\boldsymbol{w}}(\kappa) = \hat{\boldsymbol{w}}_g + \kappa \hat{\boldsymbol{w}}_z. \tag{IA17}$$

Kan, Wang, and Zhou (2021) show that the optimal combination coefficient κ is

$$\kappa^{kwz} = \frac{(T-N)(T-N-3)}{(T-2)(T-N-2)} \frac{\psi_g^2}{\psi_g^2 + (N-1)/T}. \tag{IA18}$$

Relaxing the convexity constraint would amount to combine the SMV and SGMV portfolios with the risk-free asset, which is what the three-fund rule of Kan and Zhou (2007) is designed to do. Their optimal portfolio combination is

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{kz}) = \frac{\kappa_1^{kz}}{\gamma} \widehat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\mu}} + \frac{\kappa_2^{kz}}{\gamma} \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{1} \tag{IA19}$$

with

$$\kappa_1^{kz} = \frac{1}{c} \frac{\psi_g^2}{\psi_g^2 + N/T} \quad \text{and} \quad \kappa_2^{kz} = \mu_g (1/c - \kappa_1^{kz}). \tag{IA20}$$

In the next proposition, we derive the EU of the optimal portfolio rules of Kan and Zhou (2007) and Kan, Wang, and Zhou (2021).

Proposition IA.3. *The expected out-of-sample utility of the optimal portfolio combinations of Kan and Zhou (2007) and Kan, Wang, and Zhou (2021) are, respectively,*

$$EU(\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{kz})) = U^\star - \frac{d}{2\gamma} \kappa_1^{kz} - \frac{c-1}{2\gamma} \gamma_{tan} \kappa_2^{kz}, \tag{IA21}$$

$$EU(\hat{\boldsymbol{w}}(\kappa^{kwz})) = \mu_g - \frac{\gamma}{2} c_1 \sigma_g^2 + \frac{T-N-2}{T-N-1} \frac{\psi_g^2}{2\gamma} \kappa^{kwz}, \tag{IA22}$$

where c_1 is defined in (IA4).

Similar to the combination with the EW portfolio in Section 4, κ_2^{kz} in (IA20) is unbounded and proportional to μ_g, and therefore is more sensitive to estimation errors in mean returns $\boldsymbol{\mu}$ relative to the other coefficients. Therefore, in the next proposition we derive the EU of the Kan-Zhou three-fund strategy when κ_2^{kz} is estimated as

$$\hat{\kappa}_2^{kz} = \hat{\mu}_g (1/c - \kappa_1^{kz}) = \frac{\mathbf{1}' \boldsymbol{\Sigma}^{-1} \hat{\boldsymbol{\mu}}}{\mathbf{1}' \boldsymbol{\Sigma}^{-1} \mathbf{1}} (1/c - \kappa_1^{kz}). \tag{IA23}$$

Moreover, we identify the range of risk aversion γ for which the Kan-Wang-Zhou fully invested

rule delivers a larger EU than that of the estimated Kan-Zhou three-fund rule that also invests in the risk-free asset.

Proposition IA.4. *Let the combination coefficient κ_2^{kz} be estimated by $\hat{\kappa}_2^{kz}$ in (IA23). Then,*

1. *The expected out-of-sample utility of the estimated optimal three-fund strategy of Kan and Zhou (2007) is*

$$EU(\hat{\mathbf{w}}(\kappa_1^{kz}, \hat{\kappa}_2^{kz})) = EU(\hat{\mathbf{w}}(\boldsymbol{\kappa}^{kz})) - \frac{c}{2\gamma T}(1/c^2 - (\kappa_1^{kz})^2), \tag{IA24}$$

   where $EU(\hat{\mathbf{w}}(\boldsymbol{\kappa}^{kz}))$ in (IA21) is the utility when κ_2^{kz} is known.

2. *The fully invested strategy of Kan, Wang, and Zhou (2021) delivers a larger expected out-of-sample utility than the estimated optimal three-fund strategy of Kan and Zhou (2007), $EU(\hat{\mathbf{w}}(\kappa^{kwz})) \geq EU\Big(\hat{\mathbf{w}}(\kappa_1^{kz}, \hat{\kappa}_2^{kz})\Big)$, if and only if $b \geq 0$ and γ belongs to the following interval:*

$$\gamma \in \left[\frac{\gamma_{tan}}{c_1} \pm \frac{\sqrt{b}}{c_1 \sigma_g}\right], \tag{IA25}$$

   where

$$b = \theta_g^2 + c_1\left(-\theta^2 + \psi_g^2 \frac{T-N-2}{T-N-1}\kappa^{kwz} + d\kappa_1^{kz} + (c-1)\gamma_{tan}\kappa_2^{kz} + \frac{c}{T}(1/c^2 - (\kappa_1^{kz})^2)\right). \tag{IA26}$$

For the 25SBTM dataset and $T = 120$, we find that the interval is $\gamma \in [3.18 \pm 1.73]$. Those investors are better off not investing in the risk-free asset besides the fully invested SMV and SGMV portfolios.

IA.2.3 Mixing two-fund and three-fund rules

Proposition 6 shows that it is always optimal to combine the sample tangent portfolio not just with the risk-free asset as in Kan and Zhou (2007), but also with the EW portfolio, because the optimal three-fund rule delivers a larger utility than the optimal two-fund rule,

$EU(\hat{\boldsymbol{w}}^{opt}) \geq EU(\hat{\boldsymbol{w}}^{2f})$. However, similar to the insight in Section 4, $\hat{\boldsymbol{w}}^{2f}$ requires estimating only one coefficient, κ^{2f} in (25), that is bounded between zero and one, while $\hat{\boldsymbol{w}}^{opt}$ also requires estimating κ_2^{opt} that is unbounded because it is proportional to γ_{ew}. As shown in Proposition 8, when κ_2^{opt} is estimated by $\hat{\kappa}_2^{opt}$ in (28), the EU of the optimal three-fund rule is reduced, and thus can get smaller than that of the optimal two-fund rule. In the next proposition, we identify the required sample size for the estimated optimal three-fund rule to outperform the optimal two-fund rule.

Proposition IA.5. *Let the combination coefficient κ_2^{opt} be unknown and estimated by $\hat{\kappa}_2^{opt}$ in (28). Then, the estimated optimal three-fund combination strategy delivers a larger expected out-of-sample utility than that of the optimal two-fund strategy, $EU(\hat{\boldsymbol{w}}(\kappa_1^{opt}, \hat{\kappa}_2^{opt})) \geq EU(\hat{\boldsymbol{w}}^{2f})$, if the sample size is large enough:*

$$(\theta^2 T + N)(\psi^4 - dT\psi^2(\psi^2 + d)) + (\psi^2 + d)^2(\theta^2 T(dT/c - 1) - N) \geq 0. \tag{IA27}$$

For the 25SBTM dataset the required sample size is not large, $T \geq 75$, and similarly for the other datasets. Armed with this result, as well as the result in Proposition 6 that the constrained three-fund strategy $\hat{\boldsymbol{w}}^{tz}$ delivers a larger EU than $\hat{\boldsymbol{w}}^{2f}$ if and only if $\gamma \leq 2\gamma_{ew}$, we can optimally mix the two-fund rule $\hat{\boldsymbol{w}}^{2f}$ with the mixed strategy $\hat{\boldsymbol{w}}^{mix}$ in (31) that combines the optimal and constrained three-fund rules, $\hat{\boldsymbol{w}}^{opt}$ and $\hat{\boldsymbol{w}}^{tz}$, respectively. Denoting T_{2f} the smallest value of the sample size T for which (IA27) holds, the mixed strategy is

$$\begin{cases} \hat{\boldsymbol{w}}^{mix} & \text{if } \ T \geq T_{2f}, \\ \hat{\boldsymbol{w}}^{tz} & \text{if } \ T \leq T_{2f}, \ \gamma \in \left[\underline{\gamma}_{ew}, \overline{\gamma}_{ew}\right] \text{ and } \gamma \leq 2\gamma_{ew}, \\ \hat{\boldsymbol{w}}^{2f} & \text{otherwise.} \end{cases} \tag{IA28}$$

We evaluate the empirical performance of this mixed strategy in Section IA.3.

IA.2.4 The optimal four-fund combination rule

Kan and Zhou (2007) combine the SMV portfolio with the SGMV portfolio, while as in Tu and Zhou (2011) we combine the SMV portfolio with the EW portfolio. As discussed in

Section 5.3, we find empirically in Tables 3 and 4 that it is generally preferable to combine with the EW portfolio when the sample size $T = 120$, and with the SGMV portfolio when the sample size $T = 240$. In this section, we derive the optimal four-fund portfolio that combines the SMV portfolio with both the SGMV and EW portfolios.

The resulting four-fund portfolio combination is

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \frac{\kappa_1}{\gamma}\widehat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}} + \kappa_2 \boldsymbol{w}_{ew} + \frac{\kappa_3}{\gamma}\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}. \tag{IA29}$$

Tu and Zhou (2011) combine the optimal three-fund portfolio of Kan and Zhou (2007) with the EW portfolio using the convexity constraint (12). In the next proposition instead, we derive the EU of the four-fund portfolio in (IA29) as well as the optimal combination coefficients. These are novel results in the literature.

Proposition IA.6. *The following holds concerning the four-fund portfolio combination in (IA29):*

1. *The expected out-of-sample utility is*

$$\begin{aligned} EU(\hat{\boldsymbol{w}}(\boldsymbol{\kappa})) &= \frac{\kappa_1}{\gamma}\theta^2 + \kappa_2\mu_{ew} + \frac{\kappa_3}{\gamma}\gamma_{tan} \\ &- \frac{\gamma}{2}\left(c\left(\frac{\kappa_1^2}{\gamma^2}\left(\theta^2 + \frac{N}{T}\right) + \frac{\kappa_3^2}{\gamma^2}\frac{1}{\sigma_g^2} + \frac{2\kappa_1\kappa_3}{\gamma^2}\gamma_{tan}\right) + \kappa_2^2\sigma_{ew}^2 + \frac{2\kappa_1\kappa_2}{\gamma}\mu_{ew} + \frac{2\kappa_2\kappa_3}{\gamma}\right). \end{aligned} \tag{IA30}$$

2. *The combination coefficients maximizing (IA30) are*

$$\kappa_1^{opt} = \frac{1}{f}\left[\psi^2 - \theta_g^2 - \frac{\theta^2}{c}\frac{\sigma_g^2}{\sigma_{ew}^2} + \left(1 + \frac{1}{c}\right)\mu_g\gamma_{ew}\right] \in [0,1], \tag{IA31}$$

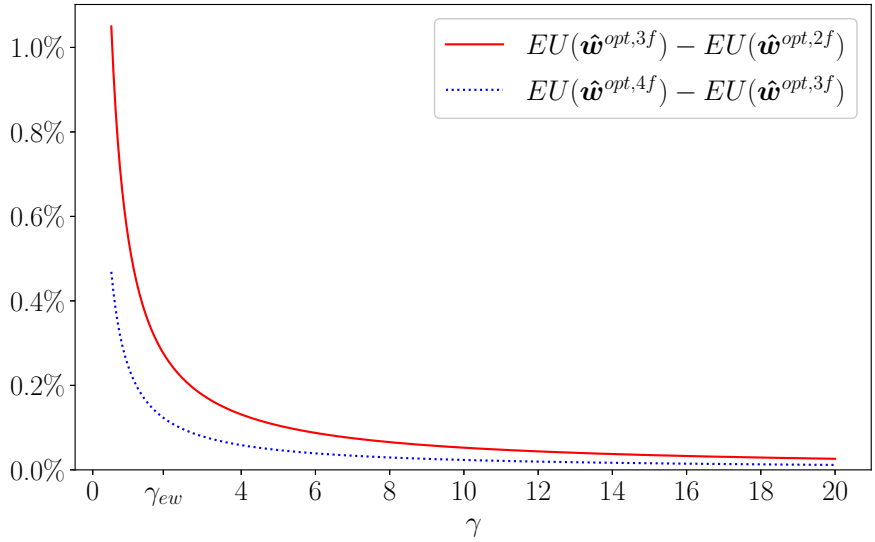
$$\kappa_2^{opt} = \frac{\gamma_{ew}}{\gamma}\frac{1}{f}\left[\frac{N}{T}\left(c - \frac{\mu_g}{\mu_{ew}}\right) + (c-1)\psi_g^2\right], \tag{IA32}$$

$$\kappa_3^{opt} = \frac{\mu_g}{f}\left[\frac{d}{c}\left(1 - \frac{\gamma_{ew}}{\gamma_{tan}}\right) - \left(1 - \frac{1}{c}\right)\psi^2\right], \tag{IA33}$$

where

$$f = \psi^2 - c\theta_g^2 + d - \frac{\sigma_g^2}{\sigma_{ew}^2}\left(\theta^2 + \frac{N}{T}\right) + 2\mu_g\gamma_{ew}. \tag{IA34}$$

Figure IA.2: Theoretical gain in expected out-of-sample utility



Notes. This figure depicts the difference between the expected out-of-sample utility of: (i) the optimal three-fund and two-fund rules (solid red), and (ii) the optimal four-fund and three-fund rules (dotted blue). The two-fund rule is that in Kan and Zhou (2007) that combines the sample tangent portfolio with the risk-free asset. The three-fund rule adds the equally weighted portfolio, and the four-fund rule adds the sample global-minimum-variance portfolio. The figure is constructed by calibrating the population vector of means and covariance matrix of stock excess returns from monthly returns on the 25 portfolios of stocks sorted on size and book-to-market spanning July 1926 to December 2021, and using a sample size $T = 120$. The differences are depicted as a function of the risk-aversion coefficient γ between 0.5 and 20.

3. The expected out-of-sample utility of the optimal four-fund portfolio is

$$EU(\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{opt})) = U^{\star} - \frac{d}{2\gamma}\kappa_1^{opt} - \frac{c-1}{2\gamma}\gamma_{tan}\kappa_3^{opt}. \tag{IA35}$$

In Figure IA.2, we evaluate how much EU can be gained in theory by opting for the optimal four-fund rule rather than the optimal three-fund rule we consider in the main body of the paper. That is, how much can be gained by adding the SGMV portfolio in the portfolio mix besides the SMV and EW portfolios. We depict this theoretical gain for the 25SBTM dataset and $T = 120$, and we compare it to the theoretical gain when going from the optimal two-fund rule of Kan and Zhou (2007) to the optimal three-fund rule that also invests in the EW portfolio. The figure shows that the incremental gain from adding the SGMV portfolio and relying on the optimal four-fund portfolio is quite small relative to the gain obtained when going from two-fund to three-fund. Combined with the additional

coefficient to estimate in the four-fund rule relative to the three-fund rule, this means that the optimal four-fund rule may not outperform in practical settings.

We evaluate the empirical performance of the optimal four-fund combination strategy in Section IA.3.

IA.2.5 Out-of-sample return correlations

In Section 5.3, we discuss the puzzling result that even though in Tables 3 and 4 the GMVRF portfolio largely outperforms the EWRF portfolio, combining the SMV portfolio with the EW portfolio performs quite closely, and even better when $T = 120$, to the combination of the SMV portfolio with the SGMV portfolio in Kan and Zhou (2007). We conjectured that this result might be due to more favorable diversification properties that arise when combining with the EW portfolio. To test this conjecture, in the next proposition we derive the correlation between the out-of-sample return of the SMV portfolio and that of either the SGMV portfolio or the EW portfolio.

Proposition IA.7. *Let $\boldsymbol{r}_{T+1} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ be the out-of-sample asset excess returns, $\hat{\boldsymbol{w}}^\star = \frac{1}{\gamma}\widehat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}$ the sample mean-variance portfolio, $\boldsymbol{w}_{ew} = \mathbf{1}/N$ the equally weighted portfolio, and $\hat{\boldsymbol{w}}_g^\star = \frac{1}{\gamma}\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}$ the sample global-minimum-variance portfolio. Then,*

$$\mathbb{C}\text{orr}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}, \boldsymbol{w}_{ew}'\boldsymbol{r}_{T+1}] = \frac{\theta_{ew}}{\sqrt{\frac{c}{T}(\theta^2(T+1)+N) + \frac{2\theta^4}{T-N-4}}} \xrightarrow{T\to\infty} \frac{\theta_{ew}}{\theta}, \tag{IA36}$$

$$\mathbb{C}\text{orr}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}, \hat{\boldsymbol{w}}_g'\boldsymbol{r}_{T+1}] = \frac{\theta_g(c+(c-1)\theta^2)}{\sqrt{\left(c+(c-1)\theta_g^2\right)\left(\frac{c}{T}(\theta^2(T+1)+N) + \frac{2\theta^4}{T-N-4}\right)}} \xrightarrow{T\to\infty} \frac{\theta_g}{\theta}. \tag{IA37}$$

Assuming that θ_{ew} and θ_g are positive, it holds that $\mathbb{C}\text{orr}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}, \boldsymbol{w}_{ew}'\boldsymbol{r}_{T+1}]$ is smaller than $\mathbb{C}\text{orr}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}, \hat{\boldsymbol{w}}_g'\boldsymbol{r}_{T+1}]$ if

$$\theta_{ew} \leq \theta_g \frac{c+(c-1)\theta^2}{\sqrt{c+(c-1)\theta_g^2}} \approx \theta_g\sqrt{c}. \tag{IA38}$$

In Figure IA.3, we calibrate Equations (IA36)–(IA37) to the four datasets we use in

our empirical analysis (see Table 1) and we depict the two correlations as a function of the sample size. The figure shows as conjectured that, for all datasets and sample size, the correlation is much smaller between the SMV portfolio and the EW portfolio than between the SMV portfolio and the SGMV portfolio. Moreover, we also depict the empirical value of the correlations obtained from our empirical analysis in Section 5 for $T = 120$ and 240 months, and we find that these empirical correlations are overall in line with the theoretical ones. We depict the empirical correlations using net out-of-sample returns, which we find are essentially the same as those obtained using gross returns.

IA.2.6 Expected out-of-sample Sharpe ratio

In Proposition 4, we show that the optimal strategy delivers, by design, a larger EU than the constrained strategy. In this section, we show moreover that the optimal strategy delivers the largest expected out-of-sample Sharpe ratio (ESR), and thus, a larger one than the constrained strategy. However, this theoretical gain is relatively small and mostly disappears when accounting for the fact that optimal combination coefficients are more sensitive to estimation errors in mean returns, as discussed in Section 4.

Just like the maximum utility in (3) is unattainable in the presence of parameter uncertainty, the maximum Sharpe ratio $\theta = \sqrt{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}$ is also unattainable. To quantify the impact of parameter uncertainty on the out-of-sample Sharpe ratio of an estimated portfolio $\hat{\boldsymbol{w}}$, we define the expected out-of-sample Sharpe ratio (ESR) as in DeMiguel, Martín-Utrera, and Nogales (2013):

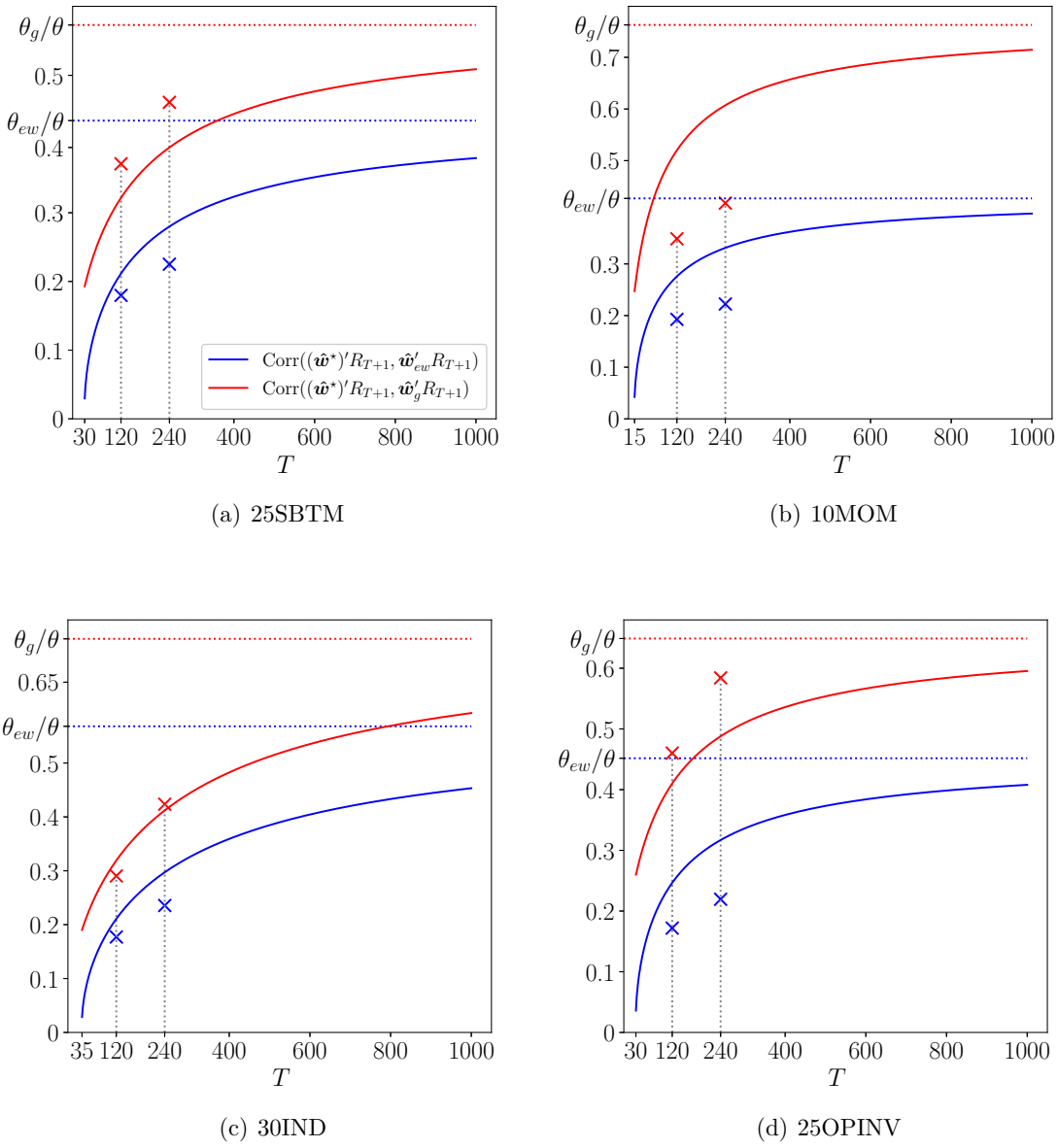
$$ESR(\hat{\boldsymbol{w}}) = \frac{\mathbb{E}[\hat{\boldsymbol{w}}'\boldsymbol{\mu}]}{\sqrt{\mathbb{E}[\hat{\boldsymbol{w}}'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}]}}. \tag{IA39}$$

From the proof of Proposition 1, we have that the ESR of the combination between the SMV and EW portfolios in (9) is

$$ESR(\hat{\boldsymbol{w}}(\boldsymbol{\kappa})) = \frac{\frac{\kappa_1}{\gamma}\theta^2 + \kappa_2\mu_{ew}}{\sqrt{\frac{\kappa_1^2}{\gamma^2}(\theta^2 + d) + \kappa_2^2\sigma_{ew}^2 + \frac{2\kappa_1\kappa_2}{\gamma}\mu_{ew}}}. \tag{IA40}$$

In the following proposition we show that the optimal combination strategy $\hat{\boldsymbol{w}}^{opt}$ delivers the maximum ESR, which is independent of the risk-aversion coefficient γ. In contrast, the

Figure IA.3: Out-of-sample return correlations



(a) 25SBTM

(b) 10MOM

(c) 30IND

(d) 25OPINV

Notes. This figure depicts the correlation between the out-of-sample return of the sample mean-variance (SMV) portfolio and that of the equally weighted (EW) portfolio (in blue) and sample global-minimum-variance (SGMV) portfolio (in red). The correlation is depicted for the four datasets listed in Table 1 as a function of the sample size T between 50 and 1,000 months. The solid lines depict the theoretical value of the correlations obtained from Proposition IA.7. The dotted horizontal lines depict the value as the sample size T goes to infinity. The crosses depict the empirical value of the correlations, which we obtain from the time series of net out-of-sample portfolio returns in the empirical analysis of Section 5 for $T = 120$ and 240.

constrained strategy $\hat{\boldsymbol{w}}^{tz}$ does not deliver the maximum ESR, except for two specific values of γ, and generally performs worst in ESR as γ tends to zero and infinity.

Proposition IA.8. *The following holds concerning the expected out-of-sample Sharpe ratio (ESR):*

1. *The optimal combination strategy $\hat{\boldsymbol{w}}^{opt}$ achieves the maximum ESR of all combination strategies, which is bounded by the Sharpe ratios of the EW and optimal mean-variance portfolio:*

$$\theta_{ew} \leq \max_{\boldsymbol{\kappa}} ESR(\hat{\boldsymbol{w}}(\boldsymbol{\kappa})) = ESR(\hat{\boldsymbol{w}}^{opt}) = \sqrt{\theta_{ew}^2 + (\theta^2 - \theta_{ew}^2)\kappa_1^{opt}} \leq \theta. \tag{IA41}$$

2. *The constrained combination strategy $\hat{\boldsymbol{w}}^{tz}$ delivers the maximum ESR if and only if $\gamma = \gamma_{ew}$ or $\gamma = \theta^2/\mu_{ew}$. Moreover, if and only if $d > \frac{\theta^2}{\theta_{ew}^2}(\theta - 3\theta_{ew})(\theta - \theta_{ew})$, $\hat{\boldsymbol{w}}^{tz}$ achieves its lowest ESR when $\gamma = 0$ or ∞:*[20]

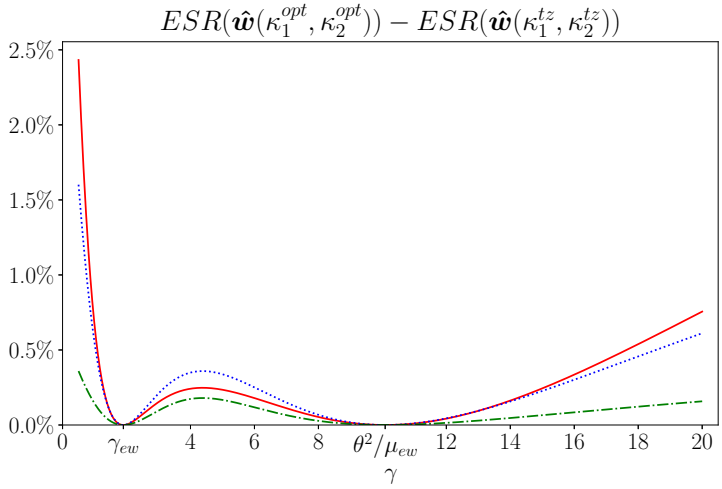
$$\min_{\gamma} ESR(\hat{\boldsymbol{w}}^{tz}) = \lim_{\gamma \to 0} ESR(\hat{\boldsymbol{w}}^{tz}) = \lim_{\gamma \to \infty} ESR(\hat{\boldsymbol{w}}^{tz}) = \frac{\theta^2}{\sqrt{\theta^2 + d}}. \tag{IA42}$$

The intuition behind part 1 of Proposition IA.8 is similar to the classical result from portfolio theory that any portfolio maximizing utility in (2) also achieves the maximum Sharpe ratio. Similarly, any portfolio combination maximizing the EU also achieves the maximum ESR. However, it is no longer the case under the convexity constraint (12), apart from two specific values of γ, and the loss in ESR is generally largest for investors with small and large degrees of risk aversion.

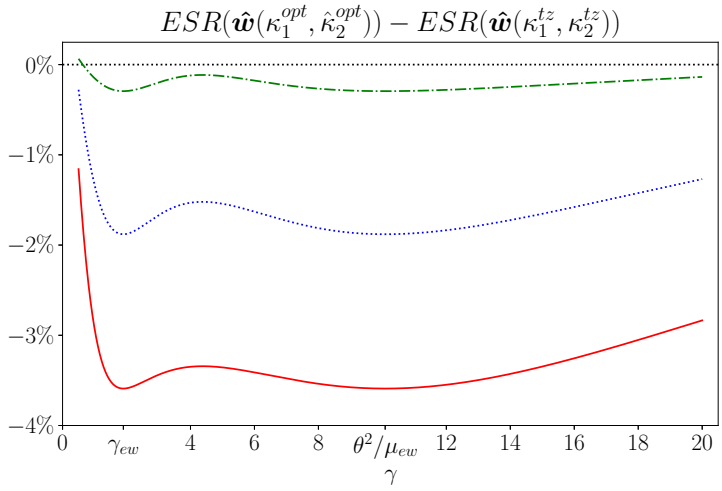
We illustrate Proposition IA.8 in panel a) of Figure **??** for the 25SBTM dataset. We depict the difference between the ESR of the optimal strategy $\boldsymbol{\kappa}^{opt}$ and that of the constrained strategy $\boldsymbol{\kappa}^{tz}$ as a function of γ for different samples sizes: $T = 60, 120$ and 480. The figure shows that the difference is typically not large, expect for small and large values of γ. Nonetheless, it is always positive and gets larger as estimation risk N/T increases.

[20] Because $d > 0$, a sufficient condition for $d > \frac{\theta^2}{\theta_{ew}^2}(\theta - 3\theta_{ew})(\theta - \theta_{ew})$ to hold is $\theta_{ew} > \theta/3$, which is often the case.

Figure IA.4: Expected out-of-sample Sharpe ratio of optimal and constrained strategies



(a) Known combination coefficients



(b) κ_2^{opt} is estimated by $\hat{\kappa}_2^{opt}$

Notes. This figure depicts the difference between the expected out-of-sample Sharpe ratio (ESR), in percentage points, of the optimal combination strategy $\hat{\boldsymbol{w}}^{opt}$ and that of the constrained combination strategy $\hat{\boldsymbol{w}}^{tz}$ in two different settings. In panel (a), all combination coefficients are known without error. In panel (b), κ_2^{opt} is estimated by $\hat{\kappa}_2^{opt}$ in (28). The figure is constructed by calibrating the population vector of means and covariance matrix of stock excess returns from monthly returns on the 25 portfolios of stocks sorted on size and book-to-market spanning July 1926 to December 2021. The ESR difference is depicted as a function of the risk-aversion coefficient γ between 0.5 and 20 for a sample size $T = 60$ (red solid), $T = 120$ (blue dotted), and $T = 480$ (green dash-dotted).

The consistent outperformance in ESR delivered by the optimal strategy in Proposition IA.8 assumes however that combination coefficients are known. As discussed in Section 4, the main difference in estimation errors between the optimal and constrained coefficients is that while the constrained ones are bounded, κ_2^{opt} is very sensitive to errors in mean returns because it is unbounded and proportional to μ_{ew}. When κ_2^{opt} is estimated by $\hat{\kappa}_2^{opt}$ in (28), we have from the proof of Proposition 8 that the ESR of the estimated optimal strategy becomes

$$ESR(\hat{\boldsymbol{w}}(\kappa_1^{opt}, \hat{\kappa}_2^{opt})) = \frac{\mathbb{E}[(\hat{\boldsymbol{w}}^{opt})'\boldsymbol{\mu}]}{\sqrt{\mathbb{E}[(\hat{\boldsymbol{w}}^{opt})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}^{opt}] + \frac{1-(\kappa_1^{opt})^2}{\gamma^2 T}}}, \tag{IA43}$$

where

$$\mathbb{E}[(\hat{\boldsymbol{w}}^{opt})'\boldsymbol{\mu}] = \frac{\kappa_1^{opt}}{\gamma}\theta^2 + \kappa_2^{opt}\mu_{ew}, \tag{IA44}$$

$$\mathbb{E}\Big[(\hat{\boldsymbol{w}}^{opt})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}^{opt}\Big] = \frac{(\kappa_1^{opt})^2}{\gamma^2}(\theta^2 + d) + (\kappa_2^{opt})^2\sigma_{ew}^2 + \frac{2\kappa_1^{opt}\kappa_2^{opt}}{\gamma}\mu_{ew} \tag{IA45}$$

are the expected out-of-sample mean return and variance of the optimal strategy when κ_2^{opt} is known, respectively. In panel b) of Figure ??, we replicate panel a) but taking into account the estimation error in κ_2^{opt}. The figure shows that the theoretical gain in ESR completely disappears due to the additional estimation risk in optimal combination coefficients.

In Section IA.3, we report the empirical out-of-sample Sharpe ratio of the different portfolio strategies we consider in the main body of the paper.

IA.3 Additional empirical results

In this section, we report additional empirical results that are related to the additional theoretical results we report in Section IA.2 of this internet appendix. First, we report the out-of-sample Sharpe ratio. Second, we discuss the performance of the optimal, constrained, and mixed strategies constructed for the Kan-Zhou three-fund rule. Third, we discuss the performance of the mixed strategy that combines two-fund and three-fund rules. Fourth, we document that the optimal four-fund rule is outperformed by the optimal three-fund rule.

IA.3.1 Out-of-sample Sharpe ratio

In this section, we report and discuss the out-of-sample Sharpe ratio of the portfolio strategies listed in Table 2. We follow the methodology of Section 5 and define the annualized out-of-sample net Sharpe ratio of portfolio strategy k as

$$SR_k = \sqrt{12} \times \frac{\hat{\mu}_k}{\hat{\sigma}_k}, \tag{IA46}$$

where $\hat{\mu}_k$ and $\hat{\sigma}_k$ are the sample mean and standard deviation of the out-of-sample portfolio returns net of proportional transaction costs for strategy k.

Table IA.1 reports the annualized net out-of-sample Sharpe Ratio for all datasets listed in Table 1 and portfolio strategies listed in Table 2. We only report the results when using the sample covariance matrix for conciseness; the conclusions are consistent when using the shrinkage covariance matrix of Ledoit and Wolf (2004). In line with the theoretical predictions in Section IA.2.6, we observe that the constrained strategy of Tu and Zhou (2011) (TZ3F) outperforms the optimal strategy (OPT3F) in most cases, although the difference is small. The mixed strategy (MIX3F) approaches the Sharpe ratio of the constrained strategy, but not quite due to estimation errors in the interval (30). Just like for the EU in Table 3, the proposed OPT3F strategy consistently outperforms the optimal two-fund rule of Kan and Zhou (2007) (KZ2F). The optimal combination of the SMV and SGMV portfolios in Kan and Zhou (2007) (KZ3F) performs similarly to OPT3F when the sample size $T = 120$, but consistently outperforms for $T = 240$. This means that, in terms of Sharpe ratio, combining SMV with SGMV is preferable to combining SMV with EW. Finally, the different combination strategies largely outperform the two naive benchmarks, EWRF and GMVRF, except for the 30IND dataset where EWRF achieves the best performance for $T = 120$ and the second-best performance for $T = 240$.

IA.3.2 Performance of Kan-Zhou mixed strategy

In this section, we discuss the out-of-sample performance of the optimal, constrained, and mixed strategies introduced in Section IA.2.1 for the three-fund combination of Kan and

Table IA.1: Annualized net out-of-sample Sharpe Ratio

Dataset	$T = 120$				$T = 240$			
	25SBTM	10MOM	30IND	25OPINV	25SBTM	10MOM	30IND	25OPINV
Panel (a): $\gamma = 3$								
Proposed portfolios								
OPT3F: tan-ew-rf	88.87	80.40	50.39	72.87	86.96	81.96	34.96	83.03
MIX3F: tan-ew-rf	89.67	80.60	51.85	71.10	88.35	83.10	38.40	84.89
Benchmark portfolios								
TZ3F: tan-ew-rf	90.90	81.09	50.50	77.23	88.90	85.66	41.83	83.21
KZ2F: tan-rf	83.32	77.13	30.45	67.27	84.18	80.81	10.78	78.35
KZ3F: tan-gmv-rf	87.75	84.10	48.45	78.04	89.37	86.30	39.56	88.30
KWZ: tan-gmv	76.82	78.61	38.72	62.44	81.71	88.71	41.23	75.94
EWRF	47.99	44.16	51.12	45.85	42.08	34.11	40.41	44.65
GMVRF	70.87	57.47	48.26	75.34	71.30	50.61	44.07	83.15
Panel (b): $\gamma = 5$								
Proposed portfolios								
OPT3F: tan-ew-rf	88.85	80.39	50.39	72.87	86.93	81.96	34.95	83.03
MIX3F: tan-ew-rf	89.34	81.98	50.23	72.23	87.00	81.73	36.17	82.11
Benchmark portfolios								
TZ3F: tan-ew-rf	90.60	82.35	52.05	78.75	90.52	87.87	44.99	84.20
KZ2F: tan-rf	83.30	77.12	30.45	67.27	84.14	80.80	10.78	78.35
KZ3F: tan-gmv-rf	87.74	84.10	48.45	78.05	89.34	86.29	39.55	88.30
KWZ: tan-gmv	83.44	84.09	45.89	71.90	91.39	95.59	49.09	86.20
EWRF	47.99	44.16	51.12	45.85	42.08	34.10	40.41	44.65
GMVRF	70.85	57.47	48.27	75.35	71.29	50.60	44.07	83.14
Panel (c): $\gamma = 10$								
Proposed portfolios								
OPT3F: tan-ew-rf	88.83	80.39	50.39	72.87	86.90	81.95	34.95	83.03
MIX3F: tan-ew-rf	88.13	80.29	51.75	72.61	86.29	81.67	36.13	82.98
Benchmark portfolios								
TZ3F: tan-ew-rf	87.58	83.71	48.51	76.25	91.93	88.40	41.41	82.94
KZ2F: tan-rf	83.28	77.12	30.46	67.27	84.12	80.79	10.77	78.34
KZ3F: tan-gmv-rf	87.73	84.09	48.46	78.05	89.31	86.28	39.54	88.30
KWZ: tan-gmv	84.46	83.64	49.71	77.25	97.95	96.90	53.86	91.39
EWRF	47.99	44.16	51.12	45.85	42.08	34.10	40.41	44.65
GMVRF	70.84	57.46	48.27	75.35	71.28	50.60	44.06	83.14
Panel (d): $\gamma = 15$								
Proposed portfolios								
OPT3F: tan-ew-rf	88.82	80.38	50.38	72.87	86.89	81.95	34.95	83.03
MIX3F: tan-ew-rf	88.77	79.81	50.06	72.87	86.89	81.95	34.95	83.03
Benchmark portfolios								
TZ3F: tan-ew-rf	86.81	83.92	45.34	74.62	91.84	87.70	36.74	82.18
KZ2F: tan-rf	83.28	77.12	30.46	67.27	84.11	80.79	10.77	78.34
KZ3F: tan-gmv-rf	87.72	84.09	48.46	78.05	89.31	86.28	39.54	88.30
KWZ: tan-gmv	81.45	79.56	49.95	77.28	96.77	92.80	54.96	91.15
EWRF	47.99	44.16	51.12	45.85	42.08	34.10	40.41	44.65
GMVRF	70.84	57.46	48.27	75.35	71.28	50.60	44.06	83.14

Notes. This table reports the annualized net out-of-sample Sharpe ratio in percentage points for the portfolio strategies described in Table 2 when using the sample covariance matrix, according to the methodology in Sections 5 and IA.3.1. The net out-of-sample Sharpe ratio is computed using proportional transaction costs of 10 basis points as in Ao, Li, and Zheng (2019). We consider two sample sizes, $T = 120$ and $T = 240$, and risk-aversion coefficients $\gamma = 3, 5, 10$, and 15.

Table IA.2: Annualized net out-of-sample utility of Kan-Zhou mixed strategy

Dataset	$T = 120$				$T = 240$			
	25SBTM	10MOM	30IND	25OPINV	25SBTM	10MOM	30IND	25OPINV
Panel (a): $\gamma = 3$								
OPT3F	4.92	9.51	−2.35	2.76	8.23	10.75	−0.27	7.82
CONST3F	10.16	10.39	1.50	7.08	9.77	11.41	1.62	8.92
MIX3F	3.28	9.14	−1.47	5.68	8.11	10.72	0.79	7.36
Panel (b): $\gamma = 5$								
OPT3F	2.90	5.70	−1.43	1.63	4.91	6.45	−0.17	4.68
CONST3F	6.04	6.31	0.75	4.26	6.28	7.48	1.53	6.13
MIX3F	3.51	5.90	−0.42	2.96	5.07	6.91	0.96	5.38
Panel (c): $\gamma = 10$								
OPT3F	1.43	2.85	−0.72	0.81	2.44	3.22	−0.08	2.34
CONST3F	1.88	2.35	−1.79	0.10	2.76	3.20	−0.84	2.24
MIX3F	2.93	2.98	−0.82	1.02	2.72	2.98	−0.68	2.69
Panel (d): $\gamma = 15$								
OPT3F	0.95	1.90	−0.48	0.54	1.62	2.15	−0.06	1.56
CONST3F	−0.47	0.62	−3.80	−2.69	0.85	1.47	−2.69	−0.44
MIX3F	0.74	1.74	−0.65	−0.27	1.32	2.06	−0.59	1.29

Notes. This table reports the annualized net out-of-sample utility in percentage points for three portfolio strategies derived in Section IA.2.1 when using the sample covariance matrix, according to the methodology in Section 5. The first strategy (OPT3F) is the optimal combination of the sample tangent portfolio, the fully invested sample global-minimum-variance portfolio, and the risk-free asset. The second strategy (CONST3F) is the combination of these three funds under the convexity constraint (12). The third strategy (MIX3F) is a mixed strategy that combines the optimal and constrained strategies. The net out-of-sample utility is computed using proportional transaction costs of 10 basis points as in Ao, Li, and Zheng (2019). We consider two sample sizes, $T = 120$ and $T = 240$, and risk-aversion coefficients $\gamma = 3, 5, 10$, and 15.

Zhou (2007), that is, the combination of the sample tangent portfolio, the fully invested SGMV portfolio, and the risk-free asset. We report the net out-of-sample utility of the three strategies in Table IA.2.

We observe that for rather small values of $\gamma = 3$ and 5, the performance gain obtained by relying on the constrained strategy rather than the optimal strategy is substantial. This result confirms the observations in the main body of the paper that imposing the convexity constraint can help performance. Moreover, the performance gain is larger than that obtained for the combination of the SMV and EW portfolios in Table 3. As explained in Section IA.2.1, this is because κ_2^{opt} is proportional to γ_{tan} instead of γ_{ew}, and the estimated γ_{tan} has a larger standard deviation than that of the estimated γ_{ew}. However, as γ increases to 10 and 15, the convexity constraint hurts performance by preventing enough investment in the risk-free asset, and thus the optimal strategy outperforms the constrained strategy.

Table IA.3: Annualized net out-of-sample utility of mix of two-fund and three-fund rules

Dataset	$T = 120$				$T = 240$			
	25SBTM	10MOM	30IND	25OPINV	25SBTM	10MOM	30IND	25OPINV
Panel (a): $\gamma = 3$								
MIX3F	10.90	8.74	2.81	5.81	8.79	9.63	1.03	9.53
MIX2F3F	10.72	9.52	2.79	5.51	8.58	8.94	0.68	9.42
Panel (b): $\gamma = 5$								
MIX3F	6.25	5.52	1.21	3.65	4.71	5.33	0.16	4.67
MIX2F3F	5.73	5.21	0.31	3.86	4.04	5.29	−1.01	4.07
Panel (c): $\gamma = 10$								
MIX3F	2.91	2.47	0.71	1.73	2.34	2.70	0.07	2.55
MIX2F3F	2.50	2.23	0.19	1.82	2.00	2.51	−0.51	2.31
Panel (d): $\gamma = 15$								
MIX3F	2.05	1.67	0.38	1.27	1.63	1.86	0.05	1.70
MIX2F3F	1.78	1.40	0.03	1.33	1.41	1.73	−0.34	1.54

Notes. This table reports the annualized net out-of-sample utility in percentage points for two portfolio strategies when using the sample covariance matrix, according to the methodology in Section 5. The first strategy (MIX3F) is the mixed strategy in Section 4 that combines the optimal and constrained three-fund rules. The second strategy (MIX2F3F) introduced in Section IA.2.3 adds the optimal two-fund rule. The net out-of-sample utility is computed using proportional transaction costs of 10 basis points as in Ao, Li, and Zheng (2019). We consider two sample sizes, $T = 120$ and $T = 240$, and risk-aversion coefficients $\gamma = 3, 5, 10,$ and 15.

Finally, as expected, the mixed strategy trades off between the performance of the optimal and constrained strategies, and therefore can be a safer approach because it is not known beforehand for a fixed value of γ which of the optimal or constrained strategy is preferable.

IA.3.3 Performance of mix of two-fund and three-fund rules

In this section, we discuss the out-of-sample performance of the mixed strategy introduced in Section IA.2.3 that combines the optimal two-fund rule with the optimal and constrained three-fund rules. In Table IA.3, we compare the performance of this strategy with that of the mixed strategy introduced in Section 4 that does not invest in the optimal two-fund rule.

The table shows that adding the two-fund rule does not deliver any gain in performance. This can be explained because we use sample sizes that are large enough for the optimal three-fund rule to consistently outperform the optimal two-fund rule, as we find in Table 3 in the main body of the paper.

Table IA.4: Annualized net out-of-sample utility of optimal three-fund and four-fund rules

Dataset	$T = 120$				$T = 240$			
	25SBTM	10MOM	30IND	25OPINV	25SBTM	10MOM	30IND	25OPINV
Panel (a): $\gamma = 3$								
OPT3F: tan-ew-rf	10.37	8.71	2.25	6.38	8.29	9.30	0.27	8.54
OPT4F: tan-ew-gmv-rf	8.57	8.33	−0.23	3.86	9.00	10.08	−0.76	7.10
Panel (b): $\gamma = 5$								
OPT3F: tan-ew-rf	6.20	5.22	1.34	3.82	4.94	5.58	0.16	5.12
OPT4F: tan-ew-gmv-rf	5.10	4.99	−0.15	2.30	5.36	6.04	−0.46	4.25
Panel (c): $\gamma = 10$								
OPT3F: tan-ew-rf	3.09	2.61	0.67	1.91	2.45	2.79	0.08	2.56
OPT4F: tan-ew-gmv-rf	2.54	2.49	−0.08	1.14	2.67	3.02	−0.23	2.12
Panel (d): $\gamma = 15$								
OPT3F: tan-ew-rf	2.06	1.74	0.45	1.27	1.63	1.86	0.05	1.70
OPT4F: tan-ew-gmv-rf	1.69	1.66	−0.05	0.76	1.78	2.01	−0.16	1.41

Notes. This table reports the annualized net out-of-sample utility in percentage points for two portfolio strategies when using the sample covariance matrix, according to the methodology in Section 5. The first strategy is the optimal three-fund rule in the main body of the paper, which combines the sample tangent portfolio, the equally weighted portfolio, and the risk-free asset. The second strategy is the optimal four-fund rule in Section IA.2.4, which adds the sample global-minimum-variance portfolio. The net out-of-sample utility is computed using proportional transaction costs of 10 basis points as in Ao, Li, and Zheng (2019). We consider two sample sizes, $T = 120$ and $T = 240$, and risk-aversion coefficients $\gamma = 3, 5, 10$, and 15.

IA.3.4 Performance of optimal four-fund rule

In this section, we compare the out-of-sample performance of the optimal four-fund rule introduced in Section IA.2.4 with that of the optimal three-fund rule in the main body of the paper that does not invest in the SGMV portfolio. We report the net out-of-sample utility of the two strategies in Table IA.4.

The table shows that the three-fund rule outperforms the four-fund rule, with the exception of the 25SBTM and 10MOM datasets when $T = 240$, in which case the four-fund portfolio does slightly better. This result is consistent with Section IA.2.4, in which we show that the theoretical gain from adding the SGMV portfolio is small and thus may disappear in practical settings due to estimation errors in the additional combination coefficient to estimate.

IA.4 Proofs of all results

We make extensive use of three properties throughout the proofs. First, because returns are assumed iid normal, the sample mean is distributed as $\hat{\boldsymbol{\mu}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/T)$, the sample covariance matrix is distributed as $(T-N-2)\widehat{\boldsymbol{\Sigma}} \sim \mathcal{W}_N(T-1, \boldsymbol{\Sigma})$, and they are independent of each other. Second, the inverse sample covariance matrix is unbiased because of the $1/(T-N-2)$ coefficient in (6), $\mathbb{E}[\widehat{\boldsymbol{\Sigma}}^{-1}] = \boldsymbol{\Sigma}^{-1}$. Third, the expectation of a quadratic form in the random variable $\boldsymbol{x}$ is

$$\mathbb{E}[\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}] = \mathbb{E}[\boldsymbol{x}]'\boldsymbol{A}\mathbb{E}[\boldsymbol{x}] + \text{Trace}(\boldsymbol{A}\mathbb{V}[\boldsymbol{x}]), \tag{IA47}$$

where $\boldsymbol{A}$ is a constant matrix (Rencher and Schaalje, 2008).

Proof of Proposition 1

Part 1. Because the inverse sample covariance matrix is unbiased, the SMV portfolio $\boldsymbol{w}^\star$ is unbiased as well, and the expected out-of-sample mean return of the portfolio combination $\hat{\boldsymbol{w}}(\boldsymbol{\kappa})$ in (9) is

$$\mathbb{E}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})'\boldsymbol{\mu}] = \frac{\kappa_1}{\gamma}\theta^2 + \kappa_2\mu_{ew}. \tag{IA48}$$

We can then obtain the expected out-of-sample variance, $\mathbb{E}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\boldsymbol{\kappa})]$, from Equation (IA47). This requires knowing the covariance matrix of the SMV portfolio $\hat{\boldsymbol{w}}^\star$, which from Javed, Mazur, and Ngailo (2021, Corollary 2.2) is given by

$$\mathbb{V}[\hat{\boldsymbol{w}}^\star] = \frac{(T-N-2)(\theta^2 + (T-2)/T)\boldsymbol{\Sigma}^{-1} + (T-N)\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}}{\gamma^2(T-N-1)(T-N-4)}. \tag{IA49}$$

Using $\mathbb{E}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})] = \boldsymbol{w}(\boldsymbol{\kappa})$ and $\mathbb{V}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})] = \kappa_1^2\mathbb{V}[\hat{\boldsymbol{w}}^\star]$, we find from (IA47) that the expected out-of-sample variance of $\hat{\boldsymbol{w}}(\boldsymbol{\kappa})$ is

$$\mathbb{E}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\boldsymbol{\kappa})] = \frac{\kappa_1^2}{\gamma^2}(\theta^2 + d) + \kappa_2^2\sigma_{ew}^2 + \frac{2\kappa_1\kappa_2}{\gamma}\mu_{ew}. \tag{IA50}$$

Combining (IA48) and (IA50) results in the EU formula in (13), which completes the proof.

Part 2. We obtain the combination coefficients maximizing the EU under the convexity constraint (12) by setting the derivative of (13) with respect to κ_1 equal to zero, taking into account that $\kappa_2 = 1 - \kappa_1$.

Proof of Proposition 2

The optimal combination coefficients are found by setting the derivative of (13) with respect to κ_1 and κ_2 equal to zero and solving the resulting linear system of two equations with two unknowns.

Proof of Proposition 3

Part 1. The proof is direct by inspecting the formula for κ_1^{tz} in (15) and κ_1^{opt} in (17).

Part 2. It is clear from the formula for κ_2^{tz} in (15) and κ_2^{opt} in (17) that $\kappa_2^{tz} \in [0, 1]$ and the sign of κ_2^{opt} is equal to that of μ_{ew}. This proves the first two statements. Regarding the third statement when $\mu_{ew} > 0$, we can show that the inequality $\kappa_2^{opt} \geq \kappa_2^{tz}$ is equivalent to

$$(\gamma - \gamma_{ew})\left(\gamma - \frac{\theta^2 + d}{\mu_{ew}}\right) \geq 0. \tag{IA51}$$

Since $(\theta^2 + d)/\mu_{ew} \geq \gamma_{ew}$ for $\mu_{ew} > 0$, it directly follows that $0 \leq \kappa_2^{opt} \leq \kappa_2^{tz}$ if $\gamma \in \left[\gamma_{ew}, \frac{\theta^2+d}{\mu_{ew}}\right]$ and $\kappa_2^{opt} \geq \kappa_2^{tz}$ otherwise, which completes the proof.

Part 3. After some developments, the inequality $\pi_{rf}^{opt} \geq \pi_{rf}^{tz}$ is equivalent to

$$\gamma^2\Big[\sigma_{ew}^2(\gamma_{tan} - \gamma_{ew})\Big] + \gamma\Big[2\mu_{ew}(\gamma_{ew} - \gamma_{tan}) + \psi^2 + d\Big] + \Big[\gamma_{tan}\theta_{ew}^2 - \gamma_{ew}(\theta^2 + d)\Big] \geq 0. \tag{IA52}$$

Because we assume $\gamma_{tan} > \gamma_{ew}$, this polynomial has two roots:

$$\gamma = \gamma_{ew} \quad \text{and} \quad \gamma = \frac{\gamma_{tan} - (\theta^2 + d)/\mu_{ew}}{\gamma_{tan}/\gamma_{ew} - 1}. \tag{IA53}$$

Let $0 < \gamma_{ew} < \gamma_{tan}$, we can then write the following about inequality (IA52):

$$\pi_{rf}^{opt} \leq \pi_{rf}^{tz} \quad \text{if} \quad \gamma \in \left[\frac{\gamma_{tan} - (\theta^2 + d)/\mu_{ew}}{\gamma_{tan}/\gamma_{ew} - 1}, \gamma_{ew}\right] \quad \text{and} \quad \pi_{rf}^{opt} \geq \pi_{rf}^{tz} \quad \text{otherwise.} \tag{IA54}$$

The lower bound of the interval is negative under the assumption $\gamma_{tan} < (\theta^2 + d)/\mu_{ew}$. Therefore, under this assumption the inequality $\pi_{rf}^{opt} \leq \pi_{rf}^{tz}$ holds for $\gamma \leq \gamma_{ew}$ because any risk-aversion coefficient $\gamma > 0$. Finally, when $\gamma \leq \gamma_{ew}$, we have that $\pi_{rf}^{tz} \leq 0$ because $\pi_{rf}^{tz} = \frac{1}{\gamma}\kappa_1^{tz}(\gamma - \gamma_{tan})$ and $\gamma_{tan} > \gamma_{ew}$ by assumption, which completes the proof.

Proof of Proposition 4

The EU of the optimal and constrained strategies is found by replacing the combination coefficients (κ_1, κ_2) by (17) for the optimal strategy and (15) for the constrained strategy in the EU formula (13). After some developments, we find that

$$EU(\hat{\boldsymbol{w}}^{opt}) = U^\star - \frac{d}{2\gamma}\kappa_1^{opt}, \tag{IA55}$$

$$EU(\hat{\boldsymbol{w}}^{tz}) = U^\star - \frac{d}{2\gamma}\kappa_1^{tz}. \tag{IA56}$$

Therefore, the EU gain of the optimal strategy relative to the constrained strategy is

$$EU(\hat{\boldsymbol{w}}^{opt}) - EU(\hat{\boldsymbol{w}}^{tz}) = \frac{d}{2\gamma}\left(\kappa_1^{tz} - \kappa_1^{opt}\right) \geq 0, \tag{IA57}$$

which is positive because $\kappa_1^{tz} \geq \kappa_1^{opt}$ as shown in Proposition 2. Inequality (IA57) holds with equality if and only if $\gamma = \infty$ or γ_{ew}, because in the latter case we have $\kappa_1^{tz} = \kappa_1^{opt}$. Finally, because d increases with N/T we can show that the EU gain in (IA57) increases with N/T if $d(\kappa_1^{tz} - \kappa_1^{opt})$ increases with d. This is the case because

$$d\left(\kappa_1^{tz} - \kappa_1^{opt}\right) = \frac{\sigma_{ew}^2(\gamma - \gamma_{ew})^2}{(\psi^2/d + 1)((\psi^2 + \sigma_{ew}^2(\gamma - \gamma_{ew})^2)/d + 1)} \tag{IA58}$$

is an increasing function of d, which completes the proof.

Proof of Proposition 5

Part 1. Using Equation (IA56), the constrained strategy $\hat{\boldsymbol{w}}^{tz}$ delivers a negative EU when

$$U^{\star} - \frac{d}{2\gamma}\kappa_1^{tz} < 0. \tag{IA59}$$

After some developments, we find that inequality (IA59) is equivalent to

$$\gamma^2[\sigma_{ew}^2(d-\theta^2)] + \gamma[2\mu_{ew}(\theta^2-d)] - \theta^4 > 0. \tag{IA60}$$

The discriminant of this second-degree polynomial is

$$\Delta = 4\sigma_{ew}^2(d-\theta^2)(\theta_{ew}^2 d + \theta^2\psi^2). \tag{IA61}$$

It follows from (IA61) that $\Delta \geq 0$ and real roots to the polynomial exist if and only if $\theta^2 < d$. In that case, (IA60) holds if and only if γ is not between the two roots:

$$\gamma \notin \left[\gamma_{ew}\left(1 \pm \sqrt{1 + \frac{\theta^4/\theta_{ew}^2}{d-\theta^2}}\right)\right]. \tag{IA62}$$

The negative sign gives a negative root, which we can ignore because any risk-aversion coefficient $\gamma > 0$. Therefore, the constrained strategy delivers a negative EU if and only if (5) holds. The threshold γ_{neg} decreases with N/T because d increases with N/T and γ_{neg} decreases with d, which completes the proof.

Part 2. The optimal strategy necessarily delivers a positive EU because $\boldsymbol{\kappa} = \mathbf{0}$, which delivers zero EU, is part of the search space.

Proof of Proposition 6

The optimal three-fund strategy always delivers a larger EU than that of the optimal two-fund strategy by construction because it adds one more fund, the EW portfolio, to the portfolio combination. To determine when the constrained three-fund strategy outperforms the optimal two-fund one, we need to determine the EU of the optimal two-fund rule. Plugging

$\kappa_1 = \kappa^{2f} = \theta^2/(\theta^2 + d)$ and $\kappa_2 = 0$ in Equation (13), this EU is

$$EU(\hat{\boldsymbol{w}}^{2f}) = U^\star - \frac{d}{2\gamma}\kappa^{2f}. \tag{IA63}$$

Therefore, given the formula for the EU of the constrained three-fund strategy in (IA56), the constrained three-fund strategy delivers a larger EU than that of the optimal two-fund strategy when $\kappa_1^{tz} \leq \kappa^{2f}$, which is equivalent to $\gamma \leq 2\gamma_{ew}$ and thus completes the proof.

Proof of Proposition 7

We want to find when the ℓ_2-norm of the optimal strategy is smaller than that of the constrained strategy, which is equivalent to showing the same result for the squared ℓ_2-norm because the ℓ_2-norm is positive. We define $f(\gamma)$ the function of γ corresponding to the difference between the squared ℓ_2-norms:

$$f(\gamma) = g(\boldsymbol{\kappa}^{tz}) - g(\boldsymbol{\kappa}^{opt}), \tag{IA64}$$

where

$$g(\boldsymbol{\kappa}) = ||\kappa_1 \boldsymbol{w}^\star + \kappa_2 \boldsymbol{w}_{ew}||_2^2 = \frac{\kappa_1^2}{\gamma^2}\|\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\|_2^2 + \frac{\kappa_2^2}{N} + \frac{2\kappa_1\kappa_2}{N\gamma}\gamma_{tan}. \tag{IA65}$$

We want to find the values of γ for which $f(\gamma) \geq 0$ holds, which can be rewritten as

$$f(\gamma) = \frac{(\kappa_1^{tz})^2 - (\kappa_1^{opt})^2}{\gamma^2}\|\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\|_2^2 + \frac{(\kappa_2^{tz})^2 - (\kappa_1^{opt})^2}{N} + 2\frac{\gamma_{tan}}{N\gamma}\left(\kappa_1^{tz}\kappa_2^{tz} - \kappa_1^{opt}\kappa_2^{opt}\right) \geq 0. \tag{IA66}$$

Proposition 3 shows that $\kappa_1^{tz} \geq \kappa_1^{opt} \geq 0$, and $\|\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\|_2^2 \geq 0$ by definition. Therefore, under the assumption $\gamma_{tan} \geq 0$, a sufficient condition for (IA66) to hold is:

$$(\kappa_2^{tz})^2 - (\kappa_2^{opt})^2 + \frac{2\gamma_{tan}}{\gamma}\left(\kappa_1^{tz}\kappa_2^{tz} - \kappa_1^{opt}\kappa_2^{opt}\right) \geq 0. \tag{IA67}$$

Inequality (IA67) holds for all values of γ_{tan} if the following two inequalities hold:

$$(\kappa_2^{tz})^2 \geq (\kappa_2^{opt})^2 \quad \text{and} \quad \kappa_1^{tz}\kappa_2^{tz} \geq \kappa_1^{opt}\kappa_2^{opt}. \tag{IA68}$$

Both inequalities hold whenever $0 \leq \kappa_2^{opt} \leq \kappa_2^{tz}$ which, from Proposition 3, holds when $\mu_{ew} \geq 0$ and $\gamma \in \left[\gamma_{ew}, \frac{\theta^2+d}{\mu_{ew}}\right]$. Therefore, $f(\gamma) \geq 0$ in (IA64) if $\gamma_{tan} \geq 0$, $\mu_{ew} \geq 0$, and $\gamma \in [\gamma_{ew}, (\theta^2+d)/\mu_{ew}]$, which completes the proof.

Proof of Proposition 8

Part 1. The two combination coefficients are κ_1^{opt} in (17) and $\hat{\kappa}_2^{opt}$ in (28). Because $\hat{\mu}_{ew}$ is unbiased, $\hat{\kappa}_2^{opt}$ is unbiased as well, and the expected out-of-sample mean return of the estimated optimal strategy is the same as that when κ_2^{opt} is known:

$$\mathbb{E}\left[\hat{\boldsymbol{w}}(\kappa_1^{opt}, \hat{\kappa}_2^{opt})'\boldsymbol{\mu}\right] = \frac{\kappa_1^{opt}}{\gamma}\theta^2 + \kappa_2^{opt}\mu_{ew}. \tag{IA69}$$

In comparison, the expected out-of-sample variance is larger than that when κ_2^{opt} is known. Specifically,

$$\begin{aligned}\mathbb{E}\left[\hat{\boldsymbol{w}}(\kappa_1^{opt}, \hat{\kappa}_2^{opt})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\kappa_1^{opt}, \hat{\kappa}_2^{opt})\right] &= \mathbb{E}\left[\left(\frac{\kappa_1^{opt}}{\gamma}\hat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma} + \hat{\kappa}_2^{opt}\boldsymbol{w}_{ew}'\boldsymbol{\Sigma}\right)'\left(\frac{\kappa_1^{opt}}{\gamma}\widehat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}} + \hat{\kappa}_2^{opt}\boldsymbol{w}_{ew}\right)\right] \\ &= \frac{(\kappa_1^{opt})^2}{\gamma^2}(\theta^2+d) + \mathbb{E}\left[(\hat{\kappa}_2^{opt})^2\right]\sigma_{ew}^2 + \frac{2\kappa_1^{opt}}{\gamma}\mathbb{E}\left[\hat{\kappa}_2^{opt}\hat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\boldsymbol{w}_{ew}\right].\end{aligned} \tag{IA70}$$

From the definition of $\hat{\kappa}_2^{opt}$ in (28) and the identity $\mathbb{E}[\hat{\mu}_{ew}^2] = \mu_{ew}^2 + \sigma_{ew}^2/T$, we find that

$$\sigma_{ew}^2\mathbb{E}\left[(\hat{\kappa}_2^{opt})^2\right] = (\kappa_2^{opt})^2\sigma_{ew}^2 + \frac{(1-\kappa_1^{opt})^2}{\gamma^2 T} \tag{IA71}$$

and

$$\frac{2\kappa_1^{opt}}{\gamma}\mathbb{E}\left[\hat{\kappa}_2^{opt}\hat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\boldsymbol{w}_{ew}\right] = \frac{2\kappa_1^{opt}\kappa_2^{opt}}{\gamma}\mu_{ew} + \frac{2\kappa_1^{opt}(1-\kappa_1^{opt})}{\gamma^2 T}. \tag{IA72}$$

Therefore, the expected out-of-sample variance is

$$\mathbb{E}\Big[\hat{\boldsymbol{w}}(\kappa_1^{opt},\hat{\kappa}_2^{opt})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\kappa_1^{opt},\hat{\kappa}_2^{opt})\Big] = \mathbb{E}\Big[(\hat{\boldsymbol{w}}^{opt})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}^{opt}\Big] + \frac{1-(\kappa_1^{opt})^2}{\gamma^2 T}, \tag{IA73}$$

where $\mathbb{E}[(\hat{\boldsymbol{w}}^{opt})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}^{opt}] = \frac{(\kappa_1^{opt})^2}{\gamma^2}(\theta^2+d) + (\kappa_2^{opt})^2\sigma_{ew}^2 + \frac{2\kappa_1^{opt}\kappa_2^{opt}}{\gamma}\mu_{ew}$ is the expected out-of-sample variance when κ_2^{opt} is known. Finally, using Equation (IA55), the EU loss relative to $\hat{\boldsymbol{w}}^{opt}$ when κ_2^{opt} is estimated by $\hat{\kappa}_2^{opt}$ is given by (29), which completes the proof.

Part 2. Using Equation (IA55), the constrained strategy delivers a larger EU than that of the estimated optimal strategy, $EU(\hat{\boldsymbol{w}}^{tz}) \geq EU\Big(\hat{\boldsymbol{w}}(\kappa_1^{opt},\hat{\kappa}_2^{opt})\Big)$, when

$$U^\star - \frac{d}{2\gamma}\kappa_1^{tz} \geq U^\star - \frac{d}{2\gamma}\kappa_1^{opt} - \frac{1-(\kappa_1^{opt})^2}{2\gamma T}. \tag{IA74}$$

After some developments, we find that inequality (IA74) is equivalent to

$$\gamma^2[-\sigma_{ew}^2] + \gamma[2\mu_{ew}] + k \geq 0, \tag{IA75}$$

where

$$k = \frac{(\psi^2+d)(2\psi^2+d)}{dT(\psi^2+d)-(2\psi^2+d)} - \theta_{ew}^2. \tag{IA76}$$

The discriminant of the second-degree polynomial in (IA75) is $\Delta = 4\sigma_{ew}^2(k+\theta_{ew}^2)$, which is positive because we assume $N \geq 2$. Therefore, inequality (IA75) holds when the risk-aversion coefficient γ belongs to the following interval:

$$\gamma \in \left[\gamma_{ew} \pm \sigma_{ew}^{-1}\sqrt{\frac{(\psi^2+d)(2\psi^2+d)}{dT(\psi^2+d)-(2\psi^2+d)}}\right], \tag{IA77}$$

which proves the result in Equation (30). Finally, because d is increasing in N, we can prove that the length of the interval (IA77) decreases with N if

$$\frac{\partial}{\partial d}\left[\frac{(\psi^2+d)(2\psi^2+d)}{dT(\psi^2+d)-(2\psi^2+d)}\right] \leq 0 \quad \Longleftrightarrow \quad \frac{2\psi^6 T + 4\psi^4(dT+1) + 2\psi^2 d(dT+2) + d^2}{(dT(\psi^2+d)-(2\psi^2+d))^2} \geq 0, \tag{IA78}$$

which always holds and thus completes the proof.

Proof of Proposition IA.1

Part 1. The SMV portfolio $\hat{\boldsymbol{w}}^\star$ is unbiased, and the SGMV portfolio $\hat{\boldsymbol{w}}_g$ too (Okhrin and Schmid, 2006). Therefore, the expected out-of-sample mean return of the portfolio combination (IA1) is

$$\mathbb{E}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})] = \frac{\kappa_1}{\gamma}\theta^2 + \kappa_2\mu_g. \tag{IA79}$$

The expected out-of-sample variance decomposes as

$$\mathbb{E}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\boldsymbol{\kappa})] = \frac{\kappa_1^2}{\gamma^2}(\theta^2 + d) + \kappa_2^2\mathbb{E}\Big[\hat{\boldsymbol{w}}_g'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_g\Big] + \frac{2\kappa_1\kappa_2}{\gamma}\mathbb{E}\Big[\hat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_g\Big]. \tag{IA80}$$

Fom Kan, Wang, and Zhou (2021, Lemma 1), it holds that

$$\mathbb{E}\Big[\hat{\boldsymbol{w}}_g'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_g\Big] = c_1\sigma_g^2, \tag{IA81}$$

where the constant c_1 is defined in (IA4). Therefore, it remains to evaluate the expectation

$$\mathbb{E}\Big[\hat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_g\Big] = \mathbb{E}\left[\frac{\hat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}}{\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}}\right]. \tag{IA82}$$

In the following lemma, we prove that this expectation is equal to $c_1\mu_g$.

Lemma 1. $\mathbb{E}\left[\frac{\hat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}}{\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}}\right] = c_1\mu_g$, *where* c_1 *is defined in* (IA4).

We prove this lemma in the next subsection.[21] Combining Equations (IA80)–(IA81) and Lemma 1, the expected out-of-sample variance is

$$\mathbb{E}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\boldsymbol{\kappa})] = \frac{\kappa_1^2}{\gamma^2}(\theta^2 + d) + c_1\kappa_2^2\sigma_g^2 + \frac{2c_1\kappa_1\kappa_2}{\gamma}\mu_g, \tag{IA83}$$

which results in the EU formula in (IA3) as desired.

[21] We thank Raymond Kan for his help on proving this lemma.

Part 2. The optimal combination coefficients are found by setting the derivative of (IA3) with respect to κ_1 and κ_2 equal to zero and solving the resulting linear system of two equations with two unknowns.

Part 3. We obtain the combination coefficients maximizing the EU under the convexity constraint (12) by setting the derivative of (IA3) with respect to κ_1 equal to zero, taking into account that $\kappa_2 = 1 - \kappa_1$.

Part 4. The EU of the optimal and constrained strategies are found by replacing the combination coefficients (κ_1, κ_2) by (IA5) for the optimal strategy and (IA6) for the constrained strategy in the EU formula (IA3).

Proof of Lemma 1

Proving the lemma amounts to show that

$$\mathbb{E}\left[\frac{\widehat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}_{ml}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}_{ml}^{-1}\mathbf{1}}{\mathbf{1}'\widehat{\boldsymbol{\Sigma}}_{ml}^{-1}\mathbf{1}}\right] = \frac{T(T-2)\mu_g}{(T-N-1)(T-N-2)}, \tag{IA84}$$

where $\widehat{\boldsymbol{\Sigma}}_{ml} = \frac{T-N-2}{T}\widehat{\boldsymbol{\Sigma}}$ is the maximum-likelihood estimator of $\boldsymbol{\Sigma}$.

Let $\boldsymbol{P}$ be an $N \times N$ orthonormal matrix with the first two columns being

$$\boldsymbol{\nu} = \sigma_g \boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{1}, \tag{IA85}$$

$$\boldsymbol{\eta} = \frac{\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{\mu} - \mu_g\mathbf{1})}{\psi_g}, \tag{IA86}$$

where ψ_g is defined in (IA2). Let

$$\boldsymbol{z} = \sqrt{T}\boldsymbol{P}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\widehat{\boldsymbol{\mu}} \sim \mathcal{N}(\boldsymbol{\mu}_z, \boldsymbol{I}_N), \tag{IA87}$$

$$\boldsymbol{W} = T\boldsymbol{P}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\widehat{\boldsymbol{\Sigma}}_{ml}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P} \sim \mathcal{W}_N(T-1, \boldsymbol{I}_N), \tag{IA88}$$

and they are independent of each other. We can show that

$$\boldsymbol{\mu}_z = \sqrt{T}\boldsymbol{P}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu} = \sqrt{T}\theta_g\boldsymbol{e}_1 + \sqrt{T}\psi\boldsymbol{e}_2, \tag{IA89}$$

where $\boldsymbol{e}_i$ is the ith column of the identity matrix $\boldsymbol{I}_N$. Using $\boldsymbol{W}$ and $\boldsymbol{z}$, we can write

$$\mathbf{1}'\widehat{\boldsymbol{\Sigma}}_{ml}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}_{ml}^{-1}\hat{\boldsymbol{\mu}} = \frac{T^{\frac{3}{2}}}{\sigma_g}\boldsymbol{e}_1'\boldsymbol{W}^{-2}\boldsymbol{z}, \tag{IA90}$$

$$\mathbf{1}'\widehat{\boldsymbol{\Sigma}}_{ml}^{-1}\mathbf{1} = \frac{T\boldsymbol{e}_1'\boldsymbol{W}^{-1}\boldsymbol{e}_1}{\sigma_g^2}. \tag{IA91}$$

It follows that

$$\frac{\hat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}_{ml}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}_{ml}^{-1}\mathbf{1}}{\mathbf{1}'\widehat{\boldsymbol{\Sigma}}_{ml}^{-1}\mathbf{1}} = \frac{\sqrt{T}\sigma_g\boldsymbol{e}_1'\boldsymbol{W}^{-2}\boldsymbol{z}}{\boldsymbol{e}_1'\boldsymbol{W}^{-1}\boldsymbol{e}_1} =: q. \tag{IA92}$$

We are interested in obtaining the exact mean of q, which is equal to

$$\mathbb{E}[q] = T\mu_g\mathbb{E}\left[\frac{\boldsymbol{e}_1'\boldsymbol{W}^{-2}\boldsymbol{e}_1}{\boldsymbol{e}_1'\boldsymbol{W}^{-1}\boldsymbol{e}_1}\right] + T\sigma_g\psi_g\mathbb{E}\left[\frac{\boldsymbol{e}_1'\boldsymbol{W}^{-2}\boldsymbol{e}_2}{\boldsymbol{e}_1'\boldsymbol{W}^{-1}\boldsymbol{e}_1}\right]. \tag{IA93}$$

Partition $\boldsymbol{W}$ into four blocks such that $\boldsymbol{W}_{11}$ is its (1,1)th element. Using Muirhead (1982, Theorem 3.2.10), we can show that

$$x := (\boldsymbol{e}_1'\boldsymbol{W}^{-1}\boldsymbol{e}_1)^{-1} = \boldsymbol{W}_{11} - \boldsymbol{W}_{12}\boldsymbol{W}_{22}^{-1}\boldsymbol{W}_{21} \sim \chi^2_{T-N}, \tag{IA94}$$

$$\boldsymbol{y} := -\boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{W}_{21} \sim \mathcal{N}(\mathbf{0}_{N-1}, \boldsymbol{I}_{N-1}), \tag{IA95}$$

$$\boldsymbol{W}_{22} \sim \mathcal{W}_{N-1}(T-1, \boldsymbol{I}_{N-1}), \tag{IA96}$$

and they are independent of each other. Let $\boldsymbol{Q} = [\boldsymbol{e}_2, \ldots, \boldsymbol{e}_N]$. From the partition matrix inverse formula, we can verify that

$$\boldsymbol{Q}'\boldsymbol{W}^{-1}\boldsymbol{e}_1 = \frac{\boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{y}}{x}, \tag{IA97}$$

$$\boldsymbol{Q}'\boldsymbol{W}^{-1}\boldsymbol{Q} = \boldsymbol{W}_{22}^{-1} + \frac{\boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{y}\boldsymbol{y}'\boldsymbol{W}_{22}^{-\frac{1}{2}}}{x}, \tag{IA98}$$

so we have

$$\boldsymbol{e}_2'\boldsymbol{W}^{-1}\boldsymbol{e}_1 = \frac{[1,\ \mathbf{0}_{N-2}']\boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{y}}{x}. \tag{IA99}$$

It follows that

$$\boldsymbol{e}_1'\boldsymbol{W}^{-2}\boldsymbol{e}_1 = \boldsymbol{e}_1'\boldsymbol{W}^{-1}[\boldsymbol{e}_1,\ \boldsymbol{Q}][\boldsymbol{e}_1,\ \boldsymbol{Q}]'\boldsymbol{W}^{-1}\boldsymbol{e}_1$$

$$
\begin{aligned}
&= (\boldsymbol{e}_1'\boldsymbol{W}^{-1}\boldsymbol{e}_1)^2 + \boldsymbol{e}_1'\boldsymbol{W}^{-1}\boldsymbol{Q}\boldsymbol{Q}'\boldsymbol{W}^{-1}\boldsymbol{e}_1 \\
&= \frac{1+\boldsymbol{y}'\boldsymbol{W}_{22}^{-1}\boldsymbol{y}}{x^2}, \qquad \text{(IA100)} \\
\boldsymbol{e}_1'\boldsymbol{W}^{-2}\boldsymbol{e}_2 &= \boldsymbol{e}_1'\boldsymbol{W}^{-1}[\boldsymbol{e}_1,\ \boldsymbol{Q}][\boldsymbol{e}_1,\ \boldsymbol{Q}]'\boldsymbol{W}^{-1}\boldsymbol{e}_2 \\
&= (\boldsymbol{e}_1'\boldsymbol{W}^{-1}\boldsymbol{e}_1)(\boldsymbol{e}_1'\boldsymbol{W}^{-1}\boldsymbol{e}_2) + \boldsymbol{e}_1'\boldsymbol{W}^{-1}\boldsymbol{Q}\boldsymbol{Q}'\boldsymbol{W}^{-1}\boldsymbol{e}_2 \\
&= \frac{[1,\ \boldsymbol{0}_{N-2}']\boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{y}}{x^2} + \frac{\boldsymbol{y}'\boldsymbol{W}_{22}^{-\frac{1}{2}}}{x}\left(\boldsymbol{W}_{22}^{-1} + \frac{\boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{y}\boldsymbol{y}'\boldsymbol{W}_{22}^{-\frac{1}{2}}}{x}\right)\begin{bmatrix}1\\ \boldsymbol{0}_{N-2}\end{bmatrix}. \qquad \text{(IA101)}
\end{aligned}
$$

Using the above expressions, we obtain

$$
\begin{aligned}
\mathbb{E}\left[\frac{\boldsymbol{e}_1'\boldsymbol{W}^{-2}\boldsymbol{e}_1}{\boldsymbol{e}_1'\boldsymbol{W}^{-1}\boldsymbol{e}_1}\right] &= \mathbb{E}\left[\frac{1+\boldsymbol{y}'\boldsymbol{W}_{22}^{-1}\boldsymbol{y}}{x}\right] \\
&= \mathbb{E}[1+\boldsymbol{y}'\boldsymbol{W}_{22}^{-1}\boldsymbol{y}]\mathbb{E}[x^{-1}] \\
&= \frac{1+\frac{N-1}{T-N-1}}{T-N-2} \\
&= \frac{T-2}{(T-N-1)(T-N-2)}, \qquad \text{(IA102)} \\
E\left[\frac{\boldsymbol{e}_1'\boldsymbol{W}^{-2}\boldsymbol{e}_2}{\boldsymbol{e}_1'\boldsymbol{W}^{-1}\boldsymbol{e}_1}\right] &= 0, \qquad \text{(IA103)}
\end{aligned}
$$

and therefore we have from (IA93) that

$$
\mathbb{E}[q] = \frac{T(T-2)\mu_g}{(T-N-1)(T-N-2)}, \qquad \text{(IA104)}
$$

which completes the proof.

Proof of Proposition IA.2

The two combination coefficients are κ_1^{opt} in (IA5) and $\hat{\kappa}_2^{opt}$ in (IA10). Because $\mathbf{1}'\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\mu}}$ is unbiased, $\hat{\kappa}_2^{opt}$ is unbiased as well, and the expected out-of-sample mean return of the estimated optimal strategy is the same as that when κ_2^{opt} is known:

$$
\mathbb{E}\left[\hat{\boldsymbol{w}}(\kappa_1^{opt}, \hat{\kappa}_2^{opt})'\boldsymbol{\mu}\right] = \frac{\kappa_1^{opt}}{\gamma}\theta^2 + \kappa_2^{opt}\mu_g. \qquad \text{(IA105)}
$$

In comparison, the expected out-of-sample variance is larger than that when κ_2^{opt} is known. Specifically,

$$\mathbb{E}\Big[\hat{\boldsymbol{w}}(\kappa_1^{opt},\hat{\kappa}_2^{opt})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\kappa_1^{opt},\hat{\kappa}_2^{opt})\Big] = \frac{(\kappa_1^{opt})^2}{\gamma^2}(\theta^2+d) + \mathbb{E}\Big[(\hat{\kappa}_2^{opt})^2\hat{\boldsymbol{w}}_g'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_g\Big] + \frac{2\kappa_1^{opt}}{\gamma}\mathbb{E}\Big[\hat{\kappa}_2^{opt}\hat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_g\Big]. \tag{IA106}$$

Applying (IA47), $\mathbb{E}\Big[\hat{\boldsymbol{w}}_g'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_g\Big] = c_1\sigma_g^2$ in (IA81), and Lemma 1, we have that

$$\begin{aligned}\mathbb{E}\Big[(\hat{\kappa}_2^{opt})^2\hat{\boldsymbol{w}}_g'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_g\Big] &= \frac{(1/c_1-\kappa_1^{opt})^2}{\gamma^2}\mathbb{E}\left[\frac{\hat{\boldsymbol{\mu}}'\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}\mathbf{1}'\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}{(\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1})^2}\right]\\ &= c_1\sigma_g^2(\kappa_2^{opt})^2 + \frac{c_1(1/c_1-\kappa_1^{opt})^2}{\gamma^2 T}\end{aligned} \tag{IA107}$$

and

$$\begin{aligned}\frac{2\kappa_1^{opt}}{\gamma}\mathbb{E}\Big[\hat{\kappa}_2^{opt}\hat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_g\Big] &= \frac{2\kappa_1^{opt}(1/c_1-\kappa_1^{opt})}{\gamma^2}\mathbb{E}\left[\frac{\hat{\boldsymbol{\mu}}'\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}}{\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}}\right]\\ &= \frac{2c_1\kappa_1^{opt}\kappa_2^{opt}}{\gamma}\mu_g + \frac{2c_1\kappa_1^{opt}(1/c_1-\kappa_1^{opt})}{T\gamma^2}.\end{aligned} \tag{IA108}$$

Therefore, the expected out-of-sample variance is

$$\mathbb{E}\Big[\hat{\boldsymbol{w}}(\kappa_1^{opt},\hat{\kappa}_2^{opt})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\kappa_1^{opt},\hat{\kappa}_2^{opt})\Big] = \mathbb{E}\Big[\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{opt})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{opt})\Big] + \frac{c_1(1/c_1^2-(\kappa_1^{opt})^2)}{\gamma^2 T}, \tag{IA109}$$

where $\mathbb{E}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{opt})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{opt})] = \frac{(\kappa_1^{opt})^2}{\gamma^2}(\theta^2+d) + c_1\sigma_g^2(\kappa_2^{opt})^2 + \frac{2c_1\kappa_1^{opt}\kappa_2^{opt}}{\gamma}\mu_g$ is the expected out-of-sample variance when κ_2^{opt} is known. Finally, the loss in EU relative to $\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{opt})$ when κ_2^{opt} is estimated by $\hat{\kappa}_2^{opt}$ is given by (IA11), which completes the proof.

Part 2. Using Equations (IA8) and (IA11), the constrained strategy delivers a larger EU than that of the estimated optimal strategy, $EU(\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{const})) \geq EU\Big(\hat{\boldsymbol{w}}(\kappa_1^{opt},\hat{\kappa}_2^{opt})\Big)$, when

$$U^\star - \frac{d}{2\gamma}\kappa_1^{const} - \frac{1}{2}(c_1-1)\mu_g\kappa_2^{const} \geq U^\star - \frac{d}{2\gamma}\kappa_1^{opt} - \frac{1}{2}(c_1-1)\mu_g\kappa_2^{opt} - \frac{c_1(1/c_1^2-(\kappa_1^{opt})^2)}{2\gamma T}. \tag{IA110}$$

After some developments, we find that inequality (IA110) is equivalent to

$$\gamma^2\Big[\sigma_g^2((c_1-1)^2\theta_g^2+c_1a)\Big]+\gamma[\mu_g((1-c_1)d-2c_1a)]+d(a-1)+a\theta^2\geq 0, \tag{IA111}$$

where a is defined in (IA14). The discriminant of this second-degree polynomial is

$$\Delta=4\sigma_g^2\Big(\theta_g^2(d(1-c_1)-2c_1a)^2+4(d(1-a)-a\theta^2)((c_1-1)^2\theta_g^2+c_1a)\Big), \tag{IA112}$$

and therefore after some developments we can show that the two roots of the polynomial (IA111) are $\left[\underline{\gamma}_{mid},\overline{\gamma}_{mid}\right]$ in (IA12), which completes the proof.

Proof of Proposition IA.3

The EU of the combination of the SMV and SGMV portfolios,

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa})=\frac{\kappa_1}{\gamma}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\mu}+\frac{\kappa_2}{\gamma}\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}, \tag{IA113}$$

is obtained by plugging $\kappa_2=0$ in (IA30), which gives

$$EU(\hat{\boldsymbol{w}}(\boldsymbol{\kappa}))=\frac{\kappa_1}{\gamma}\theta^2+\frac{\kappa_2}{\gamma}\gamma_{tan}-\frac{c\gamma}{2}\left(\frac{\kappa_1^2}{\gamma^2}\left(\theta^2+\frac{N}{T}\right)+\frac{\kappa_2^2}{\gamma^2}\frac{1}{\sigma_g^2}+\frac{2\kappa_1\kappa_2}{\gamma^2}\gamma_{tan}\right). \tag{IA114}$$

The EU of the optimal portfolio combination of Kan and Zhou (2007) in (IA21) is then obtained by plugging κ_1^{kz} and κ_2^{kz} given by (IA20) in the EU formula (IA114).

Concerning the fully invested combination of the SMV and SGMV portfolios, $\hat{\boldsymbol{w}}(\kappa)$ in (IA17), we have from Kan, Wang, and Zhou (2021) that the EU depends on κ as

$$EU(\hat{\boldsymbol{w}}(\kappa))=\mu_g-\frac{\gamma}{2}c_1\sigma_g^2+\frac{1}{\gamma}\frac{T-N-2}{T-N-1}\left(\kappa\psi_g^2-\frac{\kappa^2}{2}\left(\psi_g^2+\frac{N-1}{T}\right)\frac{(T-2)(T-N-2)}{(T-N)(T-N-3)}\right). \tag{IA115}$$

The EU of the optimal portfolio combination is then obtained by plugging κ^{kwz} given by (IA18) in the EU formula (IA115), which completes the proof.

Proof of Proposition IA.4

Part 1. The two combination coefficients are κ_1^{kz} in (IA20) and $\hat{\kappa}_2^{kz}$ in (IA23). Because $\hat{\mu}_g$ is unbiased, $\hat{\kappa}_2^{kz}$ is unbiased as well, and the expected out-of-sample mean return of the estimated optimal strategy is the same as that when κ_2^{kz} is known in (IA114):

$$\mathbb{E}\Big[\hat{\boldsymbol{w}}(\kappa_1^{kz},\hat{\kappa}_2^{kz})'\boldsymbol{\mu}\Big]=\frac{\kappa_1^{kz}}{\gamma}\theta^2+\kappa_2^{kz}\gamma_{tan}. \tag{IA116}$$

In comparison, the expected out-of-sample variance is larger than that when κ_2^{kz} is known. Specifically,

$$\begin{aligned}\mathbb{E}\Big[\hat{\boldsymbol{w}}(\kappa_1^{kz},\hat{\kappa}_2^{kz})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\kappa_1^{kz},\hat{\kappa}_2^{kz})\Big]&=\frac{(\kappa_1^{kz})^2}{\gamma^2}(\theta^2+d)+\frac{1}{\gamma^2}\mathbb{E}\Big[(\hat{\kappa}_2^{kz})^2\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}\Big]\\&\quad+\frac{2\kappa_1^{kz}}{\gamma^2}\mathbb{E}\Big[\hat{\kappa}_2^{kz}\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}\Big].\end{aligned} \tag{IA117}$$

Applying (IA47) and $\mathbb{E}\Big[\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}^{-1}\Big]=c\boldsymbol{\Sigma}^{-1}$, we have that

$$\begin{aligned}\frac{1}{\gamma^2}\mathbb{E}\Big[(\hat{\kappa}_2^{kz})^2\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}\Big]&=\frac{(1/c-\kappa_1^{kz})^2}{\gamma^2}\mathbb{E}\left[\frac{\hat{\boldsymbol{\mu}}'\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}\mathbf{1}'\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}{(\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1})^2}\right]\\&=\frac{c(\kappa_2^{kz})^2}{\gamma^2\sigma_g^2}+\frac{c(1/c-\kappa_1^{kz})^2}{\gamma^2T}\end{aligned} \tag{IA118}$$

and

$$\begin{aligned}\frac{2\kappa_1^{kz}}{\gamma^2}\mathbb{E}\Big[\hat{\kappa}_2^{kz}\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}\Big]&=\frac{2\kappa_1^{kz}(1/c-\kappa_1^{kz})}{\gamma^2}\mathbb{E}\left[\frac{\hat{\boldsymbol{\mu}}'\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}\right]\\&=\frac{2c\kappa_1^{kz}\kappa_2^{kz}}{\gamma^2}\gamma_{tan}+\frac{2c\kappa_1^{kz}(1/c-\kappa_1^{kz})}{\gamma^2T}.\end{aligned} \tag{IA119}$$

Therefore, the expected out-of-sample variance is

$$\mathbb{E}\Big[\hat{\boldsymbol{w}}(\kappa_1^{kz},\hat{\kappa}_2^{kz})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\kappa_1^{kz},\hat{\kappa}_2^{kz})\Big]=\mathbb{E}\Big[\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{kz})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{kz})\Big]+\frac{c(1/c^2-(\kappa_1^{kz})^2)}{\gamma^2T}, \tag{IA120}$$

where $\mathbb{E}\Big[\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{kz})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{kz})\Big]=c\Big(\frac{(\kappa_1^{kz})^2}{\gamma^2}(\theta^2+\frac{N}{T})+\frac{(\kappa_2^{kz})^2}{\gamma^2}\frac{1}{\sigma_g^2}+\frac{2\kappa_1^{kz}\kappa_2^{kz}}{\gamma^2}\gamma_{tan}\Big)$ is the expected out-of-sample variance when κ_2^{kz} is known. Finally, the loss in EU relative to $\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{kz})$ when κ_2^{kz}

is estimated by $\hat{\kappa}_2^{kz}$ is given by (IA24), which completes the proof.

Part 2. Using Equations (IA22) and (IA24), the fully invested strategy of Kan, Wang, and Zhou (2021) delivers a larger EU than the estimated optimal three-fund strategy of Kan and Zhou (2007), $EU(\hat{\boldsymbol{w}}(\kappa^{kwz})) \geq EU\Big(\hat{\boldsymbol{w}}(\kappa_1^{kz}, \hat{\kappa}_2^{kz})\Big)$, when

$$\mu_g - \frac{\gamma}{2} c_1 \sigma_g^2 + \frac{T-N-2}{T-N-1} \frac{\psi_g^2}{2\gamma} \kappa^{kwz} \geq U^\star - \frac{d}{2\gamma} \kappa_1^{kz} - \frac{c-1}{2\gamma} \gamma_{tan} \kappa_2^{kz} - \frac{c(1/c^2 - (\kappa_1^{kz})^2)}{2\gamma T}. \tag{IA121}$$

After some developments, we find that inequality (IA121) is equivalent to

$$\gamma^2[-c_1\sigma_g^2] + \gamma[2\mu_g] + \frac{b - \theta_g^2}{c_1} \geq 0, \tag{IA122}$$

where b is defined in (IA26). The discriminant of this second-degree polynomial is

$$\Delta = 4\sigma_g^2 b. \tag{IA123}$$

If $b < 0$, the polynomial (IA122) has no real roots and the Kan-Zhou three-fund strategy outperforms for any γ. Otherwise, the two real roots of the polynomial (IA122) are those in in (IA25), which completes the proof.

Proof of Proposition IA.5

From Equation (IA63), the EU of the optimal two-fund rule is $EU(\hat{\boldsymbol{w}}^{2f}) = U^\star - \frac{d}{2\gamma}\kappa^{2f}$, where $\kappa^{2f} = \theta^2/(\theta^2 + d)$. Therefore, the EU of the estimated three-fund portfolio in (29) is larger than that of the optimal two-fund portfolio if

$$U^\star - \frac{d}{2\gamma}\kappa_1^{opt} - \frac{1 - (\kappa_1^{opt})^2}{2\gamma T} \geq U^\star - \frac{d}{2\gamma}\kappa^{2f}, \tag{IA124}$$

which after some developments simplifies to inequality (IA27). Whether inequality (IA27) holds is independent of γ, and thus, only depends on whether the quantity $(1 - (\kappa_1^{opt})^2)/T$ in (IA124) is small enough. That is, because this quantity decreases with the sample size T, inequality (IA27) holds if the sample size T is large enough, which completes the proof.

Proof of Proposition IA.6

Part 1. Because the four-fund portfolio combination in (IA29) is unbiased, the out-of-sample mean return is unbiased as well,

$$\mathbb{E}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})'\boldsymbol{\mu}] = \frac{\kappa_1}{\gamma}\theta^2 + \kappa_2\mu_{ew} + \frac{\kappa_3}{\gamma}\gamma_{tan}. \tag{IA125}$$

To find the expected out-of-sample variance in (IA47), we need the covariance matrix of $\hat{\boldsymbol{w}}(\boldsymbol{\kappa})$. Using the result in Javed, Mazur, and Ngailo (2021, Corollary 2.2), it is given by

$$\mathbb{V}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})] = c\frac{\kappa_1^2}{\gamma^2}\left(\frac{\theta^2}{T-2} + \frac{1}{T}\right)\boldsymbol{\Sigma}^{-1} \tag{IA126}$$

$$+ c_2\left(\frac{\kappa_1^2}{\gamma^2}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1} + \frac{\kappa_3^2}{\gamma^2}\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}'\boldsymbol{\Sigma}^{-1} + \frac{\kappa_1\kappa_3}{\gamma^2}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}\mathbf{1}' + \mathbf{1}\boldsymbol{\mu}')\boldsymbol{\Sigma}^{-1}\right), \tag{IA127}$$

where

$$c_2 = \frac{c(T-N)}{(T-2)(T-N-2)} = \frac{T-N}{(T-N-1)(T-N-4)}. \tag{IA128}$$

The expected out-of-sample variance is then given by (IA47), where

$$\mathbb{E}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})]'\boldsymbol{\Sigma}\mathbb{E}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})] = \frac{\kappa_1^2}{\gamma^2}\theta^2 + \kappa_2^2\sigma_{ew}^2 + \frac{\kappa_3^2}{\gamma^2}\frac{1}{\sigma_g^2} + \frac{2\kappa_1\kappa_2}{\gamma}\mu_{ew} + \frac{2\kappa_1\kappa_3}{\gamma^2}\gamma_{tan} + \frac{2\kappa_2\kappa_3}{\gamma}, \tag{IA129}$$

$$\text{Trace}(\boldsymbol{\Sigma}\mathbb{V}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})]) = c\frac{\kappa_1^2}{\gamma^2}\frac{N}{T} + \left(c_2 + \frac{cN}{T-2}\right)\left(\frac{\kappa_1^2}{\gamma^2}\theta^2 + \frac{\kappa_3^2}{\gamma^2}\frac{1}{\sigma_g^2} + \frac{2\kappa_1\kappa_3}{\gamma^2}\gamma_{tan}\right). \tag{IA130}$$

Noticing that $1 + c_2 + \frac{cN}{T-2} = c$, the expected out-of-sample variance is

$$\mathbb{E}[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\boldsymbol{\kappa})] = c\left(\frac{\kappa_1^2}{\gamma^2}\left(\theta^2 + \frac{N}{T}\right) + \frac{\kappa_3^2}{\gamma^2}\frac{1}{\sigma_g^2} + \frac{2\kappa_1\kappa_3}{\gamma^2}\gamma_{tan}\right) + \kappa_2^2\sigma_{ew}^2 + \frac{2\kappa_1\kappa_2}{\gamma}\mu_{ew} + \frac{2\kappa_2\kappa_3}{\gamma}, \tag{IA131}$$

and thus the EU is given by (IA30), which completes the proof.

Part 2. The EU of the four-fund portfolio in (IA30) can be rewritten in matrix format as

$$EU(\hat{\boldsymbol{w}}(\boldsymbol{\kappa})) = \boldsymbol{\kappa}'\boldsymbol{\eta} - \frac{\gamma}{2}\boldsymbol{\kappa}'\boldsymbol{S}\boldsymbol{\kappa}, \tag{IA132}$$

where

$$\boldsymbol{\eta} = \begin{pmatrix} \theta^2/\gamma \\ \mu_{ew} \\ \gamma_{tan}/\gamma \end{pmatrix} \quad \text{and} \quad \boldsymbol{S} = \begin{pmatrix} \frac{c}{\gamma^2}\left(\theta^2 + \frac{N}{T}\right) & \frac{\mu_{ew}}{\gamma} & \frac{c\gamma_{tan}}{\gamma^2} \\ \frac{\mu_{ew}}{\gamma} & \sigma^2_{ew} & \frac{1}{\gamma} \\ \frac{c\gamma_{tan}}{\gamma^2} & \frac{1}{\gamma} & \frac{c}{\gamma^2\sigma_g^2} \end{pmatrix}. \tag{IA133}$$

The matrix $\boldsymbol{S}$ is positive definite and thus invertible. This is because the covariance matrix $\boldsymbol{\Sigma}$ is assumed positive definite, and thus any expected out-of-sample variance is strictly positive, $\boldsymbol{\kappa}'\boldsymbol{S}\boldsymbol{\kappa} > 0$ for any $\boldsymbol{\kappa} \neq \boldsymbol{0}$. As a result, the combination coefficients $\boldsymbol{\kappa}$ maximizing the EU in (IA30) are

$$\boldsymbol{\kappa}^{opt} = \arg\max_{\boldsymbol{\kappa}} \ \boldsymbol{\kappa}'\boldsymbol{\eta} - \frac{\gamma}{2}\boldsymbol{\kappa}'\boldsymbol{S}\boldsymbol{\kappa} = \frac{1}{\gamma}\boldsymbol{S}^{-1}\boldsymbol{\eta}, \tag{IA134}$$

which from (IA133) simplify to (IA31)–(IA33), thus completing the proof.

Part 3. Just like the utility of the optimal mean-variance portfolio $\boldsymbol{w}^\star$ is $U^\star = \theta^2/(2\gamma) = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}/(2\gamma)$, the EU achieved by the optimal four-fund portfolio combination $\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{opt})$ obtained from the combination coefficients (IA134) is

$$EU(\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{opt})) = \frac{\boldsymbol{\eta}'\boldsymbol{S}^{-1}\boldsymbol{\eta}}{2\gamma}, \tag{IA135}$$

where $\boldsymbol{\eta}$ and $\boldsymbol{S}$ are defined in (IA133). After some developments, we can show that (IA135) corresponds to (IA35).

Proof of Proposition IA.7

Let us begin with the correlation between the out-of-sample return of the SMV and EW portfolios, which is given by

$$\mathbb{C}\text{orr}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}, \boldsymbol{w}_{ew}'\boldsymbol{r}_{T+1}] = \frac{\mathbb{C}\text{ov}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}, \boldsymbol{w}_{ew}'\boldsymbol{r}_{T+1}]}{\sqrt{\mathbb{V}\text{ar}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}]\mathbb{V}\text{ar}[\boldsymbol{w}_{ew}'\boldsymbol{r}_{T+1}]}}. \tag{IA136}$$

We must evaluate the three terms appearing in (IA136). It is clear that $\mathbb{V}\text{ar}[\boldsymbol{w}_{ew}'R_{T+1}] = \sigma_{ew}^2$. Moreover, applying (IA47),

$$\mathbb{C}\text{ov}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}, \boldsymbol{w}_{ew}'\boldsymbol{r}_{T+1}] = \mathbb{E}[\boldsymbol{r}_{T+1}'\boldsymbol{w}_{ew}\hat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{r}_{T+1}] - \mathbb{E}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}]\mathbb{E}[\boldsymbol{w}_{ew}'\boldsymbol{r}_{T+1}] = \mu_{ew}/\gamma. \tag{IA137}$$

Finally, from the law of total variance,

$$\mathbb{V}\text{ar}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}] = \mathbb{E}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}^\star] + \boldsymbol{\mu}'\mathbb{V}\text{ar}[\hat{\boldsymbol{w}}^\star]\boldsymbol{\mu}, \tag{IA138}$$

where $\mathbb{E}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}^\star] = \frac{\theta^2+d}{\gamma^2}$ from the proof of Proposition 1 and $\mathbb{V}\text{ar}[\hat{\boldsymbol{w}}^\star]$ is given by (IA49). This gives

$$\mathbb{V}\text{ar}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}] = \frac{1}{\gamma^2}\left(\frac{c}{T}(\theta^2(T+1)+N) + \frac{2\theta^4}{T-N-4}\right). \tag{IA139}$$

Plugging $\mathbb{C}\text{ov}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}, \boldsymbol{w}_{ew}'\boldsymbol{r}_{T+1}]$, $\mathbb{V}\text{ar}[\boldsymbol{w}_{ew}'R_{T+1}]$, and $\mathbb{V}\text{ar}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}]$ into (IA136) results in (IA36). It is straightforward to check that this correlation tends to θ_{ew}/θ as $T \to \infty$.

Let us now consider the correlation between the out-of-sample return of the SMV and SGMV portfolios, which is given by

$$\mathbb{C}\text{orr}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}, \hat{\boldsymbol{w}}_g'\boldsymbol{r}_{T+1}] = \frac{\mathbb{C}\text{ov}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}, \hat{\boldsymbol{w}}_g'\boldsymbol{r}_{T+1}]}{\sqrt{\mathbb{V}\text{ar}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}]\mathbb{V}\text{ar}[\hat{\boldsymbol{w}}_g'\boldsymbol{r}_{T+1}]}}. \tag{IA140}$$

Applying (IA47) and $\mathbb{E}[\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}\widehat{\boldsymbol{\Sigma}}^{-1}] = c\boldsymbol{\Sigma}^{-1}\boldsymbol{A}\boldsymbol{\Sigma}^{-1}$ with $\boldsymbol{A}$ a constant matrix, we have that

$$\mathbb{C}\text{ov}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}, \hat{\boldsymbol{w}}_g'\boldsymbol{r}_{T+1}] = \mathbb{E}[\boldsymbol{r}_{T+1}'\hat{\boldsymbol{w}}_g\hat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{r}_{T+1}] - \mathbb{E}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}]\mathbb{E}[\hat{\boldsymbol{w}}_g'\boldsymbol{r}_{T+1}]$$

$$= \frac{\gamma_{tan}}{\gamma^2}(c + (c-1)\theta^2). \tag{IA141}$$

Similarly, we can show that

$$\mathbb{V}\text{ar}[\hat{\boldsymbol{w}}_g' \boldsymbol{r}_{T+1}] = \frac{1}{\gamma^2}\Big(\mathbb{E}[\boldsymbol{r}_{T+1}' \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}\mathbf{1}' \widehat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{r}_{T+1}] - \gamma_{tan}^2\Big) = \frac{1}{\gamma^2}\Big(c/\sigma_g^2 + (c-1)\gamma_{tan}^2\Big). \tag{IA142}$$

Finally, plugging $\mathbb{C}\text{ov}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}, \hat{\boldsymbol{w}}_g'\boldsymbol{r}_{T+1}]$, $\mathbb{V}\text{ar}[\hat{\boldsymbol{w}}_g'\boldsymbol{r}_{T+1}]$, and $\mathbb{V}\text{ar}[(\hat{\boldsymbol{w}}^\star)'\boldsymbol{r}_{T+1}]$ into (IA140) results in (IA37). It is straightforward to check that this correlation tends to θ_g/θ as $T \to \infty$, which completes the proof.

Proof of Proposition IA.8

Part 1. Given Equation (13), the ESR of the three-fund portfolio in (9) can be rewritten in matrix format as

$$ESR(\hat{\boldsymbol{w}}(\boldsymbol{\kappa})) = \frac{\boldsymbol{\kappa}'\boldsymbol{\eta}}{\sqrt{\boldsymbol{\kappa}'\boldsymbol{S}\boldsymbol{\kappa}}}, \tag{IA143}$$

where

$$\boldsymbol{\eta} = \begin{pmatrix} \theta^2/\gamma \\ \mu_{ew} \end{pmatrix} \quad \text{and} \quad \boldsymbol{S} = \begin{pmatrix} \frac{c}{\gamma^2}\left(\theta^2 + \frac{N}{T}\right) & \frac{\mu_{ew}}{\gamma} \\ \frac{\mu_{ew}}{\gamma} & \sigma_{ew}^2 \end{pmatrix}. \tag{IA144}$$

The matrix $\boldsymbol{S}$ is positive definite and thus invertible. This is because the covariance matrix $\boldsymbol{\Sigma}$ is assumed positive definite, and thus any expected out-of-sample variance is strictly positive, $\boldsymbol{\kappa}'\boldsymbol{S}\boldsymbol{\kappa} > 0$ for any $\boldsymbol{\kappa} \neq \mathbf{0}$. As a result, just like any maximum-utility portfolio $\frac{1}{\gamma}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ provides the maximum Sharpe ratio $\sqrt{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}$, any vector of optimal combination coefficients $\boldsymbol{\kappa}^{opt} = \frac{1}{\gamma}\boldsymbol{S}^{-1}\boldsymbol{\eta}$ provides the maximum ESR:

$$\max_{\boldsymbol{\kappa}} \; ESR(\hat{\boldsymbol{w}}(\boldsymbol{\kappa})) = ESR(\hat{\boldsymbol{w}}(\boldsymbol{\kappa}^{opt})) = \sqrt{\boldsymbol{\eta}'\boldsymbol{S}^{-1}\boldsymbol{\eta}}. \tag{IA145}$$

After some developments, we find that

$$\sqrt{\boldsymbol{\eta}'\boldsymbol{S}^{-1}\boldsymbol{\eta}} = \sqrt{\theta_{ew}^2 + (\theta^2 - \theta_{ew}^2)\kappa_1^{opt}}, \tag{IA146}$$

which is in between θ_{ew}^2 and θ^2 because $\kappa_1^{opt} \in [0,1]$, and thus completes the proof.

Part 2. The combination coefficients maximizing the ESR in (IA40) under the convexity constraint (12), are

$$\boldsymbol{\kappa} = \arg\max_{\boldsymbol{\kappa}'\mathbf{1}=1} \frac{\boldsymbol{\kappa}'\boldsymbol{\eta}}{\sqrt{\boldsymbol{\kappa}'\boldsymbol{S}\boldsymbol{\kappa}}} = \frac{\boldsymbol{S}^{-1}\boldsymbol{\eta}}{\mathbf{1}'\boldsymbol{S}^{-1}\boldsymbol{\eta}} = \left(\frac{\psi^2}{\psi^2 + \frac{\gamma_{ew}}{\gamma}d}, \frac{\frac{\gamma_{ew}}{\gamma}d}{\psi^2 + \frac{\gamma_{ew}}{\gamma}d}\right), \tag{IA147}$$

and they correspond to the constrained combination coefficients in (15) for two values of γ: $\gamma = \gamma_{ew}$ and $\gamma = \theta^2/\mu_{ew}$. Therefore, the constrained strategy delivers the maximum ESR for these two values of γ.

To find when the ESR of the constrained strategy $\hat{\boldsymbol{w}}^{tz}$ is minimized, we first simplify its ESR by plugging (15) in (IA40), which gives

$$ESR(\hat{\boldsymbol{w}}^{tz}) = \frac{d\mu_{ew}\gamma + \theta^2(\psi^2 + \sigma_{ew}^2(\gamma - \gamma_{ew})^2)}{\sqrt{(\psi^2 + \sigma_{ew}^2(\gamma - \gamma_{ew})^2 + d)(d\sigma_{ew}^2\gamma^2 + \theta^2(\psi^2 + \sigma_{ew}^2(\gamma - \gamma_{ew})^2)}}. \tag{IA148}$$

From (IA148), it is easy to check that

$$\lim_{\gamma\to 0} ESR(\hat{\boldsymbol{w}}^{tz}) = \lim_{\gamma\to\infty} ESR(\hat{\boldsymbol{w}}^{tz}) = \frac{\theta^2}{\sqrt{\theta^2 + d}}. \tag{IA149}$$

To see under which condition no value of γ between 0 and ∞ can give a smaller ESR than (IA149), we differentiate $ESR(\hat{\boldsymbol{w}}^{tz})$ with respect to γ, which gives

$$\frac{\partial}{\partial\gamma} ESR(\hat{\boldsymbol{w}}^{tz}) = \frac{d^2(\theta^2 - \mu_{ew}\gamma)(\mu_{ew} - \sigma_{ew}^2\gamma)(\theta^2 - \sigma_{ew}^2\gamma^2)}{\left((\psi^2 + \sigma_{ew}^2(\gamma - \gamma_{ew})^2 + d)(d\sigma_{ew}^2\gamma^2 + \theta^2(\psi^2 + \sigma_{ew}^2(\gamma - \gamma_{ew})^2))\right)^{3/2}}. \tag{IA150}$$

This derivative is equal to zero when $\gamma = \gamma_{ew}$ and $\gamma = \theta^2/\mu_{ew}$, which correspond to the two maxima identified above, and when $\gamma = \theta/\sigma_{ew}$, which corresponds to a local minimum. To conclude the proof, we must thus find when the value of $ESR(\hat{\boldsymbol{w}}^{tz})$ with $\gamma = \theta/\sigma_{ew}$ is larger than $\theta^2/\sqrt{\theta^2 + d}$ in (IA149). By plugging $\gamma = \theta/\sigma_{ew}$ in (IA148), we find that

$$ESR(\hat{\boldsymbol{w}}^{tz}) = \theta\frac{\theta_{ew}d + 2\theta^2(\theta - \theta_{ew})}{\theta d + 2\theta^2(\theta - \theta_{ew})} \quad \text{when} \quad \gamma = \theta/\sigma_{ew}, \tag{IA151}$$

and after some developments this value is larger than $\theta^2/\sqrt{\theta^2+d}$ when

$$d > \frac{\theta^2}{\theta_{ew}^2}(\theta - 3\theta_{ew})(\theta - \theta_{ew}), \tag{IA152}$$

which corresponds to the desired condition and thus completes the proof.

References in Internet Appendix

Ao, M., Y. Li, and X. Zheng. 2019. Approaching mean-variance efficiency for large portfolios. *The Review of Financial Studies* 32:2890–919.

DeMiguel, V., A. Martín-Utrera, and F. Nogales. 2013. Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking and Finance* 37:3018–34.

Javed, F., S. Mazur, and E. Ngailo. 2021. Higher order moments of the estimated tangency portfolio weights. *Journal of Applied Statistics* 48:517–35.

Kan, R., X. Wang, and G. Zhou. 2021. Optimal portfolio choice with estimation risk: No risk-free asset case. *Management Science* 68:2047–68.

Kan, R., and G. Zhou. 2007. Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis* 42:621–56.

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

Muirhead, R. 1982. *Aspects of Multivariate Statistical Theory.* New Jersey: John Wiley & Sons.

Okhrin, Y., and W. Schmid. 2006. Distributional properties of portfolio weights. *Journal of Econometrics* 134:235–56.

Rencher, A., and G. B. Schaalje. 2008. *Linear Models in Statistics (2nd ed.).* New Jersey: John Wiley & Sons.

Tu, J., and G. Zhou. 2011. Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics* 99:204–15.