RECONCILING MEAN-VARIANCE PORTFOLIO THEORY WITH NON-GAUSSIAN RETURNS

Nathan Lassance

REPRINT • 2021 / 13

UCLouvain

LIDAM
LFIN | 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

Reconciling mean-variance portfolio theory with non-Gaussian returns

Nathan Lassance
UCLouvain (BE), Louvain Finance, Louvain Institute of Data Analysis and Modeling
Chaussée de Binche 151, 7000 Mons, Belgium
nathan.lassance@uclouvain.be

Forthcoming in *European Journal of Operational Research* (2021)

Abstract

Mean-variance portfolio theory remains frequently used as an investment rationale because of its simplicity, its closed-form solution, and the availability of well-performing robust estimators. At the same time, it is also frequently rejected on the grounds that it ignores the higher moments of non-Gaussian returns. However, higher-moment portfolios are associated with many different objective functions, are numerically more complex, and exacerbate estimation risk. In this paper, we reconcile mean-variance portfolio theory with non-Gaussian returns by identifying, among all portfolios on the mean-variance efficient frontier, the one that optimizes a chosen higher-moment criterion. Numerical simulations and an empirical analysis show, for three higher-moment objective functions and adjusting for transaction costs, that the proposed portfolio strikes a favorable tradeoff between specification and estimation error. Specifically, in terms of out-of-sample Sharpe ratio and higher moments, it outperforms the global-optimal portfolio, and also the global-minimum-variance portfolio except when using monthly returns for which estimation error is more prominent.

Keywords

Finance, mean-variance portfolio, higher moments, estimation risk.

1. Introduction

The mean-variance portfolio theory introduced by Markowitz (1952) remains the most important framework in investment science, for several reasons. First, it has intuitively clear implications in terms of portfolio strategy (choose a portfolio on the mean-variance efficient (MVE) frontier) and asset pricing (the capital market line and CAPM). Second, it has simple and efficient mathematical and numerical solutions because its objective function is convex. Third, a vast literature exists on robust estimators of mean-variance portfolios such as, to mention a few, shrinkage estimation (Ledoit and Wolf 2004, 2017), Bayesian statistics (Jorion 1986, Bauder et al. 2018), dimension reduction (Chen and Yuan 2016), robust optimization (Goldfarb and Iyengar 2003), portfolio-weight constraints (DeMiguel et al. 2009, Levy and Levy 2014), and portfolio combination (Kan and Zhou 2007, Kan et al. 2021).

However, mean-variance portfolio theory is often rejected based on numerous evidence that asset returns are non-Gaussian/elliptical (Mandelbrot 1963, Gormsen and Jensen 2020) and that investors have preferences for higher moments (Scott and Horvath 1980, Ang et al. 2006). This has pushed researchers to propose higher-moment portfolio strategies as alternatives to mean-variance portfolios; see Briec et al. (2013) and Section 1.2 of Lassance (2020) for reviews. However, in departing from the MVE frontier, one abandons its aforementioned benefits. First, while the mean-variance objective is clear (maximize quadratic utility), it is not obvious for the investor which alternative to select because there are dozens of different higher-moment approaches corresponding, for example, to performance ratios (Cogneau and Hubner 2014 list many of them), non-quadratic utility functions (Levy 1969, Martellini and Ziemann 2010, Dahlquist et al. 2017), moment constraints (Athayde and Flores 2004, Briec et al. 2007), or downside-risk criteria such as the Value-at-Risk (VaR), expected shortfall, and lower-partial-moment (LPM) measures (Price et al. 1982, El Ghaoui et al. 2003, Alexander and Baptista 2004, Leòn and Moreno 2017, Lwin et al. 2017, Brandtner et al. 2020). Second, such alternative objective functions are mathematically and numerically more complex, particularly in high dimension, because they require for example the identification of a mean-variance-skewness(-kurtosis) efficient surface or the global optimum of a nonconvex objective function. Third, although a recent literature proposes robust estimation methods for higher-moment portfolio strategies (Harvey et al. 2010, Martellini and Ziemann 2010, Boudt et al. 2020, Lassance and Vrins 2021), optimizing higher moments severely exacerbates estimation risk compared to mean-variance portfolios, which is problematic out-of-sample (OOS).

Let us suppose however that those three issues can be ignored: a single higher-moment objective function has been identified, the global optimum has been found, and the parameters are known without estimation errors. Even in this ideal scenario, a whole body of literature provides evidence that the investor is not much worse off relying on a portfolio on the MVE frontier nonetheless because, for a wide class of utility functions, MVE portfolios provide good approximations of maximum-expected-utility portfolios; see Levy and Markowitz (1979), Pulley (1979), Simaan (1993, 2014), Kroll et al. (1984), and Markowitz (2014).

Putting all these elements together, we propose to reconcile mean-variance portfolio theory with non-Gaussian returns by identifying, among the continuum of all MVE portfolios, the one that optimizes a chosen higher-moment criterion. Indeed, while all portfolios on the MVE frontier are optimal in terms of quadratic utility for some choice of risk-tolerance coefficient, they do not behave similarly in terms of higher moments when asset returns are not Gaussian. We achieve this by identifying the quadratic-utility risk-tolerance coefficient that maximizes the chosen higher-moment criterion, which is a simple single-variable optimization problem. With this approach, one keeps the aforementioned benefits of Markowitz mean-variance framework while improving higher moments and keeping an efficient mean-variance tradeoff. This idea of calibrating the risk-tolerance coefficient to find the optimal portfolio on the MVE frontier is similar in spirit to the framework introduced by Ter Horst et al. (2006) and Kan and Zhou (2007), who search for the risk-aversion coefficient maximizing the expected OOS quadratic utility under parameter uncertainty. Our objective is instead to calibrate the coefficient to maximize a chosen higher-moment criterion.

The paper is structured as follows. In Section 2, we introduce and analyze the proposed method. In particular, we consider for tractability the specific case of the CARA utility and a Gaussian-mixture distribution for the asset returns similar to that used in DeMiguel and Nogales (2009) and Khashanah et al. (2020). In that setting, we derive an approximate closed-form expression for the proposed optimal MVE (OMVE) portfolio, we discuss how it differs from the global-optimal (GOPT) portfolio, and we show that the OMVE portfolio and its three-fund extension outperform the GOPT portfolio OOS.

In Section 3, we conduct an empirical analysis using daily returns from three characteristic portfolio datasets, and monthly returns from five MSCI international indices and five S&P asset class indices. We consider three higher-moment objective functions that represent popular criteria, are well distinct from one another, and are expressed in terms of the first four moments for which robust estimates have been proposed in Martellini and Ziemann (2010) and Boudt et al. (2020). These are the Cornish-Fisher expansion of the VaR (Favre and Galeano 2002, Eling and Schuhmacher 2007), the Gram-Charlier expansion of the LPM (Leòn and Moreno 2017), and the Taylor expansion of the CRRA utility (Guidolin and Timmermann 2008, Martellini and Ziemann 2010). Looking at OOS performance adjusted for transaction costs, we obtain compelling evidence that, for the three daily return datasets, the proposed strategy outperforms the GOPT portfolio in terms of the Sharpe ratio and the considered higher-moment criteria, while reducing portfolio turnover. Therefore, there is no OOS benefit in moving away from the MVE frontier. Moreover, the

OMVE portfolio is a better choice on the MVE frontier relative to the global-minimum-variance (GMV) portfolio and the optimal portfolio under parameter uncertainty proposed by Kan et al. (2021) because it often achieves an improved Sharpe ratio and higher-moment risk. The conclusions are robust to considering the two datasets with monthly returns. The main difference is that, consistent with Martellini and Ziemann (2010), the OMVE portfolio no longer outperforms the GMV portfolio in terms of the higher-moment criteria, but it still performs closely and has an improved mean-variance tradeoff. This is because the GMV portfolio has smaller estimation error, which compensates for the increased specification error when using monthly returns. Therefore, we conclude that a misspecified portfolio, which strikes a better balance between the two types of errors, can outperform the correctly specified portfolio. Sometimes, the best misspecified portfolio can be the OMVE portfolio, sometimes it can be the GMV portfolio.

These results are in line with Simaan (2014) who, for the log, exponential, and power utilities, finds a positive OOS opportunity cost of relying on the maximum-expected-utility portfolio rather than an MVE portfolio. He concludes that "[...] the mean–variance model extracts more information from sample data because it uses the covariance matrix of returns. The expected utility model may reach its optimal solution without using information from the covariance matrix." In contrast, the objective functions we consider depend directly on the portfolio return moments, and thus, the GOPT portfolios make use of the covariance matrix (and higher-comoment matrices). Therefore, our results indicate that the OOS superiority of MVE portfolios goes beyond the specific setting of Simaan (2014) based on sample estimates of maximum-expected-utility portfolios.

In Section 4, we rely on Monte Carlo simulations and the VaR objective function to compare the OOS performance of the GMV portfolio, our OMVE strategy, and the GOPT portfolio. As in Ter Horst et al. (2006) and Kan and Zhou (2007), we evaluate the OOS performance via the expected-performance-loss metric, which quantifies the tradeoff between specification and estimation error. We find that, except for cases of small estimation risk with 10 assets and a large sample size, our strategy always achieves a smaller expected VaR loss than the GOPT portfolio, and also outperforms the GMV portfolio. Thus, it has lower estimation error than the GOPT portfolio, and lower specification error than the GMV portfolio. Moreover, by staying on the MVE frontier, our portfolio benefits from the smallest expected Sharpe ratio loss in all considered cases.

We close by emphasizing how our work differs from Simaan (2014). First, whereas Simaan only considers expected utility as an objective function, which is not always the most natural choice in practice when asset returns are non-Gaussian, our framework allows for any objective function desired by the investor. In particular, the log, exponential and power utility functions are often qualified as “locally quadratic”, so that Simaan's results are not very surprising. In contrast, we use three distinct higher-moment objective functions as mentioned above. Second, Simaan's results are based on monthly returns, and the corresponding small sample size is disadvantageous for optimal portfolio rules based on higher moments (Martellini and Ziemann 2010), especially since he relies on sample estimates. In contrast, we consider both daily and monthly returns, and both sample and robust shrinkage estimates of moments. Third, Simaan only experiments his method using data from the Dow Jones index, whereas we consider five real datasets in Section 3 and simulated data in Section 4. Finally, Simaan does not consider the GMV portfolio as benchmark even though it is often the best MVE portfolio OOS (Jagannathan and Ma 2003). We include it in our results and find indeed that it performs well compared to the GOPT portfolios. Still, our alternative OMVE portfolio strategy often does better than GMV in terms of Sharpe ratio and higher moments even though we adjust for transaction costs.

Finally, Section 5 concludes the paper. There are supplementary materials to the paper in which we provide the proofs of results and additional empirical results.

2. Theoretical framework

We clarify the notation in Section 2.1, review MVE portfolios in Section 2.2, introduce our optimal MVE strategy in Section 2.3, and analyze its theoretical properties in a specific setting in Section 2.4.

2.1. Notation

We consider a standard static one-period framework and adopt the following notation. $\boldsymbol{X} \in \mathbb{R}^N$ is a random vector of asset returns with mean $\boldsymbol{\mu}$ and positive-definite covariance matrix $\boldsymbol{\Sigma}$. We denote by $\boldsymbol{w} \in \mathbb{R}^N$ the vector of portfolio weights. When a risk-free asset r_f is available, the portfolio return is $P = \boldsymbol{w}'\boldsymbol{X} + (1 - \boldsymbol{w}'\mathbf{1})r_f$, and we denote by $\boldsymbol{\mu}_e = \boldsymbol{\mu} - r_f\mathbf{1}$ the excess returns. When only risky assets are considered, the portfolio return is $P = \boldsymbol{w}'\boldsymbol{X}$ with $\boldsymbol{w}'\mathbf{1} = 1$. All vectors are column vectors. The mean and variance of P are μ_P and σ_P^2, the third and fourth central moments are $m_{3,P} := \mathbb{E}[(P - \mu_P)^3]$ and $m_{4,P} := \mathbb{E}[(P - \mu_P)^4]$, and the skewness and excess kurtosis are $\zeta_P := m_{3,P}/\sigma_P^3$ and $\kappa_P := m_{4,P}/\sigma_P^4 - 3$. Although our framework can easily accommodate constraints on portfolio weights such as short-selling constraints, we assume throughout that portfolio weights are unconstrained.[1] That is, $\boldsymbol{w} \in \mathbb{R}^N$ when a risk-free asset is available and $\boldsymbol{w} \in \mathcal{W}$ otherwise, where

$$\mathcal{W} := \left\{\boldsymbol{w} \in \mathbb{R}^N \;\middle|\; \boldsymbol{w}'\mathbf{1} = 1\right\}. \tag{1}$$

2.2. Mean-variance-efficient portfolios

Our objective is to identify the optimal portfolio on the MVE frontier according to a specified criterion. It is well-known that this frontier is different in the case with and without a risk-free asset, and thus, we distinguish between the two cases. In the empirical analysis and Monte Carlo simulations of Sections 3 and 4, we consider the case with only risky assets as commonly done in the literature (Simaan 1993a, Jagannathan and Ma 2003, DeMiguel et al 2009, Martellini and Ziemann 2010, Levy and Levy 2014). However, the case with a risk-free asset is mathematically more tractable, which is useful for the theoretical analysis of Section 2.4.

2.2.1. The case with a risk-free asset

Given a risk-tolerance coefficient $\lambda \in \mathbb{R}^+$, the MVE portfolio maximizes quadratic utility:

$$\boldsymbol{w}(\lambda) := \underset{\boldsymbol{w}\in\mathbb{R}^N}{\operatorname{argmax}}\ \mu_P - \frac{1}{2\lambda}\sigma_P^2 = \underset{\boldsymbol{w}\in\mathbb{R}^N}{\operatorname{argmax}}\ \boldsymbol{w}'\boldsymbol{\mu} + (1 - \boldsymbol{w}'\mathbf{1})r_f - \frac{1}{2\lambda}\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}.$$

The MVE portfolio $\boldsymbol{w}(\lambda)$ has the well-known solution

$$\boldsymbol{w}(\lambda) = \lambda\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_e. \tag{2}$$

At this stage, it is important to realize that there are an infinite number of MVE portfolios depending on the coefficient λ, which correspond to different combinations between the risk-free asset r_f and the MSR portfolio that is obtained for $\lambda = (\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_e)^{-1}$.

2.2.2. The case without a risk-free asset

Keeping the same notation as for the case with a risk-free asset for simplicity, the MVE portfolio with no risk-free asset is given by

$$\boldsymbol{w}(\lambda) := \underset{\boldsymbol{w}\in\mathcal{W}}{\operatorname{argmax}}\ \mu_P - \frac{1}{2\lambda}\sigma_P^2 = \underset{\boldsymbol{w}\in\mathcal{W}}{\operatorname{argmax}}\ \boldsymbol{w}'\boldsymbol{\mu} - \frac{1}{2\lambda}\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}.$$

[1] When short-selling is allowed, the investor's wealth may become negative. Although this rarely happens with daily returns, this is inconsistent with the CRRA utility as long as it can happen with a non-zero probability. However, we still allow for negative weights because the four-moment Taylor expansion in (22) does not suffer as much from the negative-wealth problem, and because, as reviewed in Section 2.2, it is commonly done in the literature. We refer to e.g. Jeon and Shin (2019) for the imposition of a negative-wealth constraint under the CRRA utility.

Defining $\boldsymbol{Q} := \boldsymbol{\Sigma}^{-1} - (\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1})^{-1}\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}'\boldsymbol{\Sigma}^{-1}$, $\boldsymbol{w}(\lambda)$ has the well-known solution

$$\boldsymbol{w}(\lambda) = \boldsymbol{w}_{GMV} + \lambda \boldsymbol{Q}\boldsymbol{\mu}, \tag{3}$$

where

$$\boldsymbol{w}_{GMV} := \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}} = \boldsymbol{w}(0) \tag{4}$$

is the GMV portfolio. Another notable MVE portfolio is the MSR portfolio,

$$\boldsymbol{w}_{MSR} := \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} = \boldsymbol{w}(\lambda_{MSR}), \tag{5}$$

where $\lambda_{MSR} = (\boldsymbol{w}_{MSR} - \boldsymbol{w}_{GMV})'\boldsymbol{\mu}/(\boldsymbol{\mu}'\boldsymbol{Q}\boldsymbol{\mu})$.

Just like for the case with a risk-free asset, there are an infinite number of MVE portfolios depending on the coefficient λ. For example, $\lambda = 0$ corresponds to the GMV portfolio, $\lambda = \infty$ to the maximum-mean-return portfolio, and $\lambda = \lambda_{MSR}$ to the MSR portfolio.

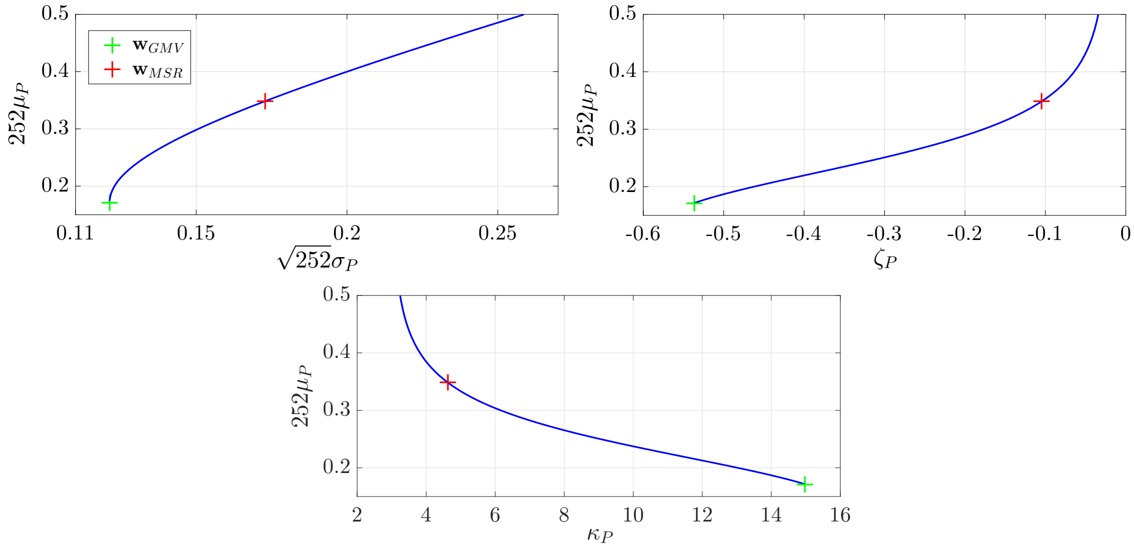
2.3. Optimal MVE portfolio with non-Gaussian returns

While all these portfolios are MVE, they do behave similarly in terms of higher moments, as we illustrate in the next example for the case without a risk-free asset.

Example 1. We collect value-weighted daily returns on the 25 size-and-book-to-market portfolios dataset spanning January 1970 to May 2020. We consider all MVE portfolios with annualized mean return between $\mu_{GMV} = 0.1715$ and 0.5, corresponding to $\lambda \in [0, 0.1593]$ in (3). In Figure 1, we plot the mean-variance, mean-skewness and mean-excess-kurtosis frontiers of all these MVE portfolios. Depending on the MVE portfolio chosen, one can obtain different moment profiles. For example, the GMV portfolio $\boldsymbol{w}_{GMV}$, corresponding to $\lambda = 0$, has $(252\mu_P, \sqrt{252}\sigma_P, \zeta_P, \kappa_P) = (0.1715, 0.1206, -0.5369, 14.98)$, and the MSR portfolio $\boldsymbol{w}_{MSR}$, corresponding to $\lambda_{MSR} = 0.0858$, has $(252\mu_P, \sqrt{252}\sigma_P, \zeta_P, \kappa_P) = (0.3528, 0.1730, -0.1047, 4.628)$. Thus, one can already reap in-sample higher-moment improvements by going from the GMV to the MSR portfolio.

Because investors care about higher moments, it is theoretically no longer optimal to rely on the quadratic utility as an objective function when asset returns are non-Gaussian. We suppose that the investor has chosen an alternative criterion $\mathcal{C}(\boldsymbol{w})$. As reviewed in Section 1, many higher-moment criteria have been put forward in the literature. Our objective in this paper is not to propose a new one or argue in favor of an existing one, hence we take $\mathcal{C}(\boldsymbol{w})$ as a given. We denote the GOPT

Figure 1 First four moments of mean-variance-efficient portfolios



Notes. The figure is based on value-weighted daily returns from the 25 size-and-book-to-market portfolios dataset from Kenneth French's library, for the time period January 1970 to May 2020. The three plots depict the mean-variance (top left), mean-skewness (top right) and mean-excess-kurtosis (bottom) frontiers of all mean-variance-efficient portfolios with mean return between that of the minimum-variance portfolio ($252\mu_P = 0.1715$) and $252\mu_P = 0.5$. The green cross is the global-minimum-variance portfolio $\boldsymbol{w}_{GMV}$ and the red cross is the maximum-Sharpe-ratio portfolio $\boldsymbol{w}_{MSR}$.

portfolio maximizing $\mathcal{C}(\boldsymbol{w})$ as

$$\boldsymbol{w}_{\mathcal{C}} := \underset{\boldsymbol{w}\in\mathcal{W}}{\text{argmax}}\ \mathcal{C}(\boldsymbol{w}) \quad \text{or} \quad \boldsymbol{w}_{\mathcal{C}} := \underset{\boldsymbol{w}\in\mathbb{R}^N}{\text{argmax}}\ \mathcal{C}(\boldsymbol{w}) \tag{6}$$

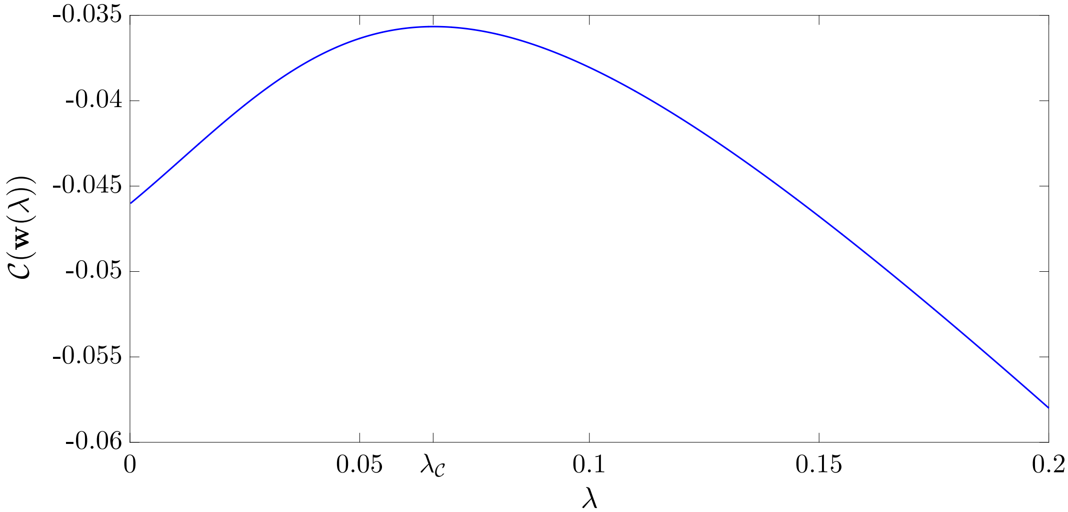
depending on whether a risk-free asset is available. Instead of relying on $\boldsymbol{w}_{\mathcal{C}}$ that may be difficult to find numerically and may be highly prone to estimation risk, we propose to find the MVE portfolio that maximizes $\mathcal{C}(\boldsymbol{w})$. In doing so, one can obtain improvements in terms of higher moments as illustrated in Example 1 while keeping the appeal of the MVE frontier such as its clear objective function, its closed-form solution, its mean-variance efficiency, and its reduced estimation risk compared to higher-moment portfolio strategies.

Mathematically, we select the portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}}) = \lambda_{\mathcal{C}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_e$ or $\boldsymbol{w}(\lambda_{\mathcal{C}}) = \boldsymbol{w}_{GMV} + \lambda_{\mathcal{C}}\boldsymbol{Q}\boldsymbol{\mu}$, where

$$\lambda_{\mathcal{C}} := \underset{\lambda\in\mathbb{R}^+}{\text{argmax}}\ \mathcal{C}(\boldsymbol{w}(\lambda)). \tag{7}$$

Contrary to the identification of $\boldsymbol{w}_{\mathcal{C}}$, we are pretty much guaranteed to find the OMVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ because $\lambda_{\mathcal{C}}$ is the result of a single-variable optimization problem. The OMVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ only requires estimating $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ and $\lambda_{\mathcal{C}}$, and all higher-moment estimation risk is transferred to the latter parameter. This may be beneficial OOS given how challenging it is to estimate higher

Figure 2 Cornish-Fisher Value-at-Risk of mean-variance-efficient portfolios



Notes. The figure is based on value-weighted daily returns from the 25 size-and-book-to-market portfolios dataset spanning January 1970 to May 2020. The figure depicts the Cornish-Fisher Value-at-Risk $\mathcal{C}(\boldsymbol{w})$ in (8), with $\alpha = 1\%$, of the MVE portfolio $\boldsymbol{w}(\lambda)$ with a risk-tolerance coefficient $\lambda \in [0, 0.2]$ in (3). The value of λ maximizing $\mathcal{C}(\boldsymbol{w}(\lambda))$ is $\lambda_{\mathcal{C}} = 0.066$, corresponding to $\mathcal{C}(\boldsymbol{w}(\lambda_{\mathcal{C}})) = -3.57\%$

moments in a robust manner. Our proposal is similar in spirit to the framework introduced by Ter Horst et al. (2006) and Kan and Zhou (2007), which consists in shrinking the MSR portfolio toward the GMV portfolio to optimize the expected OOS quadratic utility. Our objective is instead to optimize a given higher-moment criterion that is different from quadratic utility.

Example 1 (Continued)**.** Consider as higher-moment objective function $\mathcal{C}(\boldsymbol{w})$ the Cornish-Fisher VaR (Favre and Galeano 2002, Eling and Schuhmacher 2007), which is a four-moment expansion of the VaR corresponding to

$$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR} := \mu_P - \sigma_P \left(\frac{1}{36}(2z_\alpha^3 - 5z_\alpha)\zeta_P^2 - \frac{1}{6}(z_\alpha^2 - 1)\zeta_P - \frac{1}{24}(z_\alpha^3 - 3z_\alpha)\kappa_P - z_\alpha \right), \tag{8}$$

where z_α is the standard Gaussian quantile at confidence level α. For the data of Example 1, Figure 2 depicts how $\mathcal{C}(\boldsymbol{w}(\lambda))$ evolves as a function of $\lambda \in [0, 0.2]$ in (3) for $\alpha = 1\%$. The optimal $\lambda_{\mathcal{C}} = 0.066$, corresponding to the MVE portfolio with annualized mean $\mu_P = 0.3076$ in Figure 1. It has a 1% CFVaR of $\mathcal{C}(\boldsymbol{w}(\lambda_{\mathcal{C}})) = -3.57\%$, versus -4.60% for the GMV portfolio for example.

2.4. Optimal MVE portfolio versus global-optimal portfolio

In this section, we analyze theoretically the proposed OMVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ and the GOPT portfolio $\boldsymbol{w}_{\mathcal{C}}$. In Section 2.4.1, we describe the specific objective function $\mathcal{C}(\boldsymbol{w})$ and distributional model for the asset returns $\boldsymbol{X}$ that we consider. In Section 2.4.2, we derive analytical approximations of the GOPT and OMVE portfolios. In Section 2.4.3, we evaluate their expected OOS performance.

2.4.1. Performance criterion and distributional model

We consider the setting with a risk-free asset as in Section 2.2.1 because it is analytically more tractable. Concerning the criterion $\mathcal{C}(\boldsymbol{w})$, we adopt the well-known CARA utility, which for an initial wealth normalized to one is given by

$$\mathcal{C}(\boldsymbol{w}) = -\lambda \mathbb{E}\left[e^{-P(\boldsymbol{w})/\lambda}\right], \tag{9}$$

where λ is the CARA risk-tolerance coefficient.

Concerning the distribution of asset returns, we use a tractable model that identifies higher moments in a parsimonious way. Specifically, we adopt a mixture model similar to that used in Simaan (1993), DeMiguel and Nogales (2009), and Khashanah et al. (2020), under which the asset returns are elliptical in a first regime but are subject to jumps in a second regime, which induces higher-moment risk. We refer to Guidolin and Timmermann (2008), for instance, for evidence on jumps and regimes in asset returns. We follow Khashanah et al. (2020) and consider a setting with a risk-free asset in which the asset excess returns are

$$\boldsymbol{X} = \boldsymbol{G} + \boldsymbol{\gamma} J, \tag{10}$$

where $\boldsymbol{G} \sim \mathcal{N}(\boldsymbol{\mu_G}, \boldsymbol{\Sigma_G})$ is the Gaussian component, $J = \alpha$ with probability π and zero otherwise is the jump component, and $\boldsymbol{\gamma}$ is the vector of exposures to the jump component. As explained by Simaan (1993), in such a model the higher comoments of asset returns are determined by the jump component $\boldsymbol{\gamma} J$.

2.4.2. Derivation of optimal portfolios

We now recall the main steps followed by Khashanah et al. (2020) to derive an approximation of the maximum-CARA utility portfolio $\boldsymbol{w}_{\mathcal{C}}$ for $\mathcal{C}(\boldsymbol{w})$ in (9), and then provide a new result identifying the optimal MVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ under a similar approximation.

Khashanah et al. (2020) show that the CARA utility is given by

$$\mathcal{C}(\boldsymbol{w}) = -\lambda \exp\left(-\frac{\boldsymbol{w}'\boldsymbol{\mu_G}}{\lambda} + \frac{\boldsymbol{w}'\boldsymbol{\Sigma_G}\boldsymbol{w}}{2\lambda^2}\right)\left(1 - \pi + \pi e^{-\gamma_P/\lambda}\right), \tag{11}$$

where $\gamma_P = \boldsymbol{w}'\tilde{\boldsymbol{\gamma}}$ with $\tilde{\boldsymbol{\gamma}} = \alpha\boldsymbol{\gamma}$. Because α is typically negative, one can see that increasing the portfolio exposure to the jump component, γ_P, decreases utility because it decreases skewness as well. Looking for the portfolio weights $\boldsymbol{w}$ maximizing the log of (11) leads to the first order condition

$$\boldsymbol{\mu_G} - \frac{1}{\gamma}\boldsymbol{\Sigma_G w} + \phi(\gamma_P)\tilde{\boldsymbol{\gamma}} = 0,$$

where

$$\phi(x) := \frac{\pi}{\pi + (1-\pi)e^{x/\lambda}}. \tag{12}$$

Using then the linear approximation

$$\phi(x) \approx \pi - \frac{\pi(1-\pi)}{\lambda}x, \tag{13}$$

we find by solving for $\boldsymbol{w}$ that the GOPT portfolio is approximately

$$\boldsymbol{w}_{\mathcal{C}} \approx \lambda\boldsymbol{\Sigma}_{\boldsymbol{G}}^{-1}\left(\boldsymbol{\mu_G} + \phi_1\tilde{\boldsymbol{\gamma}}\right), \tag{14}$$

where

$$\phi_1 = \pi - \pi(1-\pi)\frac{\boldsymbol{\mu}_{\boldsymbol{G}}'\boldsymbol{\Sigma}_{\boldsymbol{G}}^{-1}\tilde{\boldsymbol{\gamma}} + \pi\tilde{\boldsymbol{\gamma}}'\boldsymbol{\Sigma}_{\boldsymbol{G}}^{-1}\tilde{\boldsymbol{\gamma}}}{1 + \pi(1-\pi)\tilde{\boldsymbol{\gamma}}'\boldsymbol{\Sigma}_{\boldsymbol{G}}^{-1}\tilde{\boldsymbol{\gamma}}}. \tag{15}$$

Consistent with results in Simaan (1993), Adcock (2014), and Bernard et al. (2020), the GOPT portfolio in (14) is characterized by a three-fund separation rule. First, the risk-free asset. Second, the portfolio $\frac{\boldsymbol{\Sigma}_{\boldsymbol{G}}^{-1}\boldsymbol{\mu_G}}{\boldsymbol{1}'\boldsymbol{\Sigma}_{\boldsymbol{G}}^{-1}\boldsymbol{\mu_G}}$, which is the MSR portfolio for the Gaussian component. Third, the additional portfolio $\frac{\boldsymbol{\Sigma}_{\boldsymbol{G}}^{-1}\tilde{\boldsymbol{\gamma}}}{\boldsymbol{1}'\boldsymbol{\Sigma}_{\boldsymbol{G}}^{-1}\tilde{\boldsymbol{\gamma}}}$, which controls for skewness via the exposure to the jump component.

Let us now derive the proposed OMVE portfolio. The mean and variance of asset excess returns $\boldsymbol{X} = \boldsymbol{G} + \boldsymbol{\gamma}J$ are easily found to be

$$\begin{aligned} \boldsymbol{\mu}_e &= \boldsymbol{\mu_G} + \pi\tilde{\boldsymbol{\gamma}}, \\ \boldsymbol{\Sigma} &= \boldsymbol{\Sigma_G} + \pi(1-\pi)\tilde{\boldsymbol{\gamma}}\tilde{\boldsymbol{\gamma}}', \end{aligned} \tag{16}$$

and the MVE portfolio for the risk-tolerance coefficient λ is $\boldsymbol{w}(\lambda) = \lambda\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_e$ as in (2). In the next proposition, we identify the optimal portfolio on the MVE frontier maximizing the CARA utility.

Proposition 1. *Let the asset returns follow the mixture model in* (10)*. When a risk-free asset is available, and under the approximation of* $\phi(x)$ *in* (13)*, the OMVE portfolio* $\boldsymbol{w}(\lambda_{\mathcal{C}})$ *for* $\mathcal{C}(\boldsymbol{w})$ *being* *the CARA utility in* (11) *is given by*

$$\boldsymbol{w}(\lambda_{\mathcal{C}}) = \phi_2 \boldsymbol{w}(\lambda), \tag{17}$$

where $\boldsymbol{w}(\lambda) = \lambda \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_e$ *is the MVE portfolio optimizing quadratic utility and*

$$\phi_2 = \frac{\overbrace{\boldsymbol{\mu}_e' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_e}^{\text{Squared Sharpe ratio of } \boldsymbol{w}(\lambda)}}{\underbrace{\boldsymbol{\mu}_e' \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{\boldsymbol{G}} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_e}_{\substack{\text{Variance of } \boldsymbol{w}(\lambda) \text{ from} \\ \text{Gaussian component}}} + \pi(1-\pi)(\underbrace{\boldsymbol{\mu}_e' \boldsymbol{\Sigma}^{-1} \tilde{\boldsymbol{\gamma}}}_{\substack{\text{Exposure of } \boldsymbol{w}(\lambda) \\ \text{to jump component}}})^2}. \tag{18}$$

Three comments are in order. First, Proposition 1 shows that, by design, the OMVE portfolio is a two-fund portfolio that invests in the risk-free asset and the MSR portfolio. The way portfolio return skewness, coming from the exposure to the jump component, is accounted for is via the parameter ϕ_2 in (18) that acts as a weight on the MVE portfolio $\boldsymbol{w}(\lambda)$ maximizing quadratic utility. In particular, the higher the exposure of $\boldsymbol{w}(\lambda)$ to the jump component, the lower is ϕ_2 and the more the investor should shrink toward the risk-free asset to maximize the CARA utility.

Second, the parameter ϕ_2 is independent of λ, and thus, all investors should adjust their MVE portfolio $\boldsymbol{w}(\lambda)$ in the same way.

Third, the way skewness is accounted for is different for the OMVE versus the GOPT portfolio. The GOPT portfolio requires the identification of two funds in addition to the risk-free asset, one related to the Gaussian component and the other to the jump component. Instead, the OMVE portfolio needs the classical MSR portfolio in addition to the risk-free asset, and adjusts the combination of the two funds according to ϕ_2. Khashanah et al. (2020) find, similarly to Simaan (2014), that the two-fund MVE portfolio $\boldsymbol{w}(\lambda)$ outperforms the GOPT portfolio $\boldsymbol{w}_{\mathcal{C}}$ OOS. Our proposal is thus to stay on the efficient frontier, but to choose the OMVE portfolio on this frontier, which translates to the portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}}) = \phi_2 \boldsymbol{w}(\lambda)$ in this specific setting.

2.4.3. Comparison of expected out-of-sample CARA utility

We now evaluate the expected OOS CARA utility of several portfolios. First, the OMVE portfolio, considering both the approximation in in Proposition 1 and the true one found numerically. Second, the GOPT portfolio. To make sure we truly use the global optimal, we compute it numerically instead of using the approximation in (14). Third, the three-fund portfolio of Kan and Zhou (2007), called KZ, which is designed to maximize the expected OOS utility when asset returns are Gaussian (i.e., when $\alpha = 0$). Fourth, inspired by the KZ portfolio, we consider a three-fund version of the proposed OMVE portfolio, which invests in the risk-free asset, the GMV portfolio, and the MSR portfolio. It is defined as

$$\boldsymbol{w}(\lambda_{\mathcal{C}}^{1}, \lambda_{\mathcal{C}}^{2}) = \lambda_{\mathcal{C}}^{1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_{e} + \lambda_{\mathcal{C}}^{2} \boldsymbol{\Sigma}^{-1} \mathbf{1}, \tag{19}$$

where $(\lambda_{\mathcal{C}}^{1}, \lambda_{\mathcal{C}}^{2}) > 0$ are computed numerically so as to maximize the CARA utility in (11).

We proceed as follows. First, we fix the population values of the parameters. We calibrate $\boldsymbol{\mu}_{\boldsymbol{G}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{G}}$ as the mean and covariance matrix of the daily returns on the 25 size-and-book-to-market portfolios as in Example 1. We also set $\alpha = -2\%$, $\pi = 5\%$, and $\boldsymbol{\gamma}$ takes equally spaced values between 1/25 and 1 and is ordered according to the diagonal of $\boldsymbol{\Sigma}_{\boldsymbol{G}}$; that is, the highest-volatility assets are also subject to larger jumps. We take a risk-tolerance coefficient $\lambda = 1/3$ as in Kan and Zhou (2007). Second, we proceed via Monte Carlo simulations and generate $M = 10,000$ times T returns from the Gaussian-mixture distribution in (10) with the above parameters. For each draw, we estimate $(\boldsymbol{\mu}_{\boldsymbol{G}}, \boldsymbol{\Sigma}_{\boldsymbol{G}}, \tilde{\boldsymbol{\gamma}}, \pi)$ via the expectation-maximization algorithm using the Matlab fitgmdist function, and then estimate the four portfolios above. Third, we evaluate the expected OOS CARA utility of the estimated portfolios as $M^{-1} \sum_{m=1}^{M} \mathcal{C}(\hat{\boldsymbol{w}}_m)$, where $\hat{\boldsymbol{w}}_m$ is the estimated portfolio at the mth simulation.

In Table 1, we report the difference between the expected OOS CARA utility of the considered portfolios and that of the GOPT portfolio as a benchmark, using a sample size T between 0.5 and 10 years of daily returns (assuming 252 days per year). While the difference is negative when using the analytical approximation of the OMVE portfolio, it is always positive when considering the true OMVE portfolio. This establishes theoretically that the $\boldsymbol{w}(\lambda_{\mathcal{C}})$ portfolio can outperform the $\boldsymbol{w}_{\mathcal{C}}$ portfolio, and thus, that estimation error can dominate specification error. Also, the KZ portfolio leads to a large loss in performance compared to the GOPT portfolio. Therefore, when combining the risk-free asset, the GMV portfolio, and the MSR portfolio, ignoring non-Gaussianity comes at a cost. Finally, the three-fund OMVE portfolio outperforms all considered portfolios for all sample sizes. Therefore, investing in the GMV portfolio improves OOS performance compared to the OMVE portfolio that only invests in the risk-free asset and the MSR portfolio.

3. Empirical analysis

We now test the OOS performance of the proposed portfolio strategy on real data. We explain the methodology in Section 3.1, describe the data in Section 3.2, and discuss the results in Section 3.3.

Table 1 Expected out-of-sample CARA utility relative to the global-optimal portfolio

	$T = 0.5$y	$T = 1$y	$T = 2$y	$T = 5$y	$T = 10$y
Analytical approximation OMVE	0.0011	0.0002	0.0001	0.0002	0.0003
Numerical true OMVE	0.0009	-0.0003	-0.0006	-0.0010	-0.0011
Kan-Zhou three fund	-0.0697	-0.0315	-0.0165	-0.0076	-0.0044
Numerical three-fund OMVE	0.0038	0.0057	0.0047	0.0035	0.0012

Notes. The table reports the difference between the expected out-of-sample CARA utility of several portfolios and that of the global-optimal portfolio $\boldsymbol{w}_{\mathcal{C}}$. The first portfolio is the proposed optimal MVE (OMVE) portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$, computed first with the analytical approximation in Proposition 1 and, second, numerically by solving (7). The second portfolio is the three-fund portfolio of Kan and Zhou (2007). The third portfolio is a three-fund version of the proposed OMVE portfolio, in Equation (19). The assumed distribution for the asset returns and the formula for the CARA utility are available in Section 2.4.2, and the specific value of the population parameters considered in Section 2.4.3. The expected out-of-sample CARA utility is computed via 10,000 Monte Carlo simulations.

3.1. Methodology

Data. We use three daily return characteristic portfolio datasets, and two monthly return datasets made of international and asset class indices. The three daily return datasets are collected from Kenneth French's library: 10 industry portfolios (*10IND*), 25 size-and-book-to-market portfolios (*25BTM*), and 25 size-and-operating-profitability portfolios (*25Prof*). The sample period spans January 1970 to May 2020. The two monthly return indices datasets are as follows. First, following Guidolin and Nicodano (2009) and Ghalanos et al. (2015) who consider investing across MSCI indices, we consider five total return MSCI international indices: MSCI Europe, MSCI North America, MSCI Asia Pacific, MSCI Emerging Markets, and MSCI World. The sample period spans March 2001 to May 2020. Total returns are expressed in USD. Second, we consider a dataset with different US asset class indices, including real estate as in Guidolin et al. (2020) who find economic value in holding such an asset class. We use five total return S&P asset class indices: S&P 500 investment grade corporate bond, S&P 500 high yield corporate bond, S&P GSCI, S&P US REIT, and S&P500. The sample period spans January 1995 to May 2020.[2]

Portfolio strategies. We consider six portfolio strategies. First, the equally weighted (EW) portfolio $\mathbf{1}/N$. Second, the GMV portfolio $\boldsymbol{w}_{GMV}$. Third, the MSR portfolio $\boldsymbol{w}_{MSR}$.[3] Fourth, the optimal MVE portfolio $\boldsymbol{w}_{KWZ}$ of Kan et al. (2021) that maximizes the expected OOS quadratic utility assuming iid Gaussian asset returns, using a risk-aversion coefficient $\gamma = 10$.[4] Fifth, the proposed OMVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$. Sixth, the GOPT portfolio $\boldsymbol{w}_{\mathcal{C}}$. We use three objective functions $\mathcal{C}(\boldsymbol{w})$ that

[2]The MSCI data are downloaded from Thomson Reuters Eikon and the S&P data from Bloomberg.

[3]The MSR portfolio is almost always the worst-performing portfolio, and thus, we do not report its performance.

[4]We consider $\gamma = 10$ because a smaller value deteriorates OOS performance due to a larger exposure to the mean return, and a higher value makes it increasingly similar to the GMV portfolio.

are listed in the next paragraph. $\lambda_{\mathcal{C}}$ and $\boldsymbol{w}_{\mathcal{C}}$ are found using Matlab *GlobalSearch* optimizer. We consider the case with no risk-free asset as in Simaan (2014), and we allow short-selling.[5]

Objective functions. The first objective is the CFVaR in (8). We set the CFVaR confidence level at $\alpha = 1\%$ as in Bali et al. (2013). The second objective is the LPM; see, for example, Price et al. (1982) and Sortino and van der Meer (1991). In all generality, the LPM is defined as

$$\mathrm{LPM}(\tau, m) := \int_{-\infty}^{\tau} (\tau - x)^m f_P(x) \mathrm{d}x, \tag{20}$$

where f_P is the portfolio return density. We consider the case $m = 2$ and $\tau = \mu_P$, which corresponds to the well-known semivariance. To express the LPM in terms of moments, we rely on Leòn and Moreno (2017) who derive a four-moment Gram-Charlier expansion of this measure. For the general case $\mathrm{LPM}(\tau, m)$, the expansion takes a complex form. However, we show in the supplementary materials that, for the special case $m = 2$ and $\tau = \mu_P$, the expansion reduces to maximizing the objective function

$$\mathcal{C}(\boldsymbol{w}) = -\mathrm{GCLPM} := -\sigma_P^2 \big(1/2 - (18\pi)^{-1/2} \zeta_P\big). \tag{21}$$

The third objective is the four-moment Taylor-series expansion of the CRRA utility (Guidolin and Timmermann 2008, Martellini and Ziemann 2010):

$$\mathcal{C}(\boldsymbol{w}) = \mathrm{CRRA} := \mu_P - \frac{\gamma}{2}\sigma_P^2 + \frac{\gamma(\gamma+1)}{6} m_{3,P} - \frac{\gamma(\gamma+1)(\gamma+2)}{24} m_{4,P}. \tag{22}$$

Because $m_{3,P}$ and $m_{4,P}$ are small in magnitude compared to the mean and variance, we set $\gamma = 100$ to ensure that higher moments matter. This also ensures a minor sensitivity to estimation error in the mean return μ_P, just like Martellini and Ziemann (2010) who remove μ_P from (22).

Rolling-window approach. All results are OOS using an estimation window of five years, on which the portfolios are estimated, and a subsequent testing window of one month, on which the OOS returns are computed. The windows are rolled over by one month over time, and the performance metrics below are computed on all OOS returns.

Performance metrics and transaction costs. We evaluate the OOS performance in terms of the mean, volatility, Sharpe ratio, CFVaR, square root of GCLPM, and CRRA. To evaluate the level

[5] The case with no risk-free asset is better suited to the objective functions we consider that, to avoid the large estimation errors in sample mean returns, mostly amount to minimize risk. Indeed, unreported experiments show that, when including the risk-free asset, the portfolios are nearly entirely exposed to the risk-free asset.

of transaction cost, we compute the portfolio turnover defined in DeMiguel et al. (2009) as

$$\text{Turnover} = \frac{1}{\tau - 1} \sum_{t=1}^{\tau-1} \sum_{i=1}^{N} |w_{i,t+1} - w_{i,t^+}|, \tag{23}$$

where τ is the length of the OOS period, $w_{i,t+1}$ is the desired weight on asset i at time $t+1$, and w_{i,t^+} the weight before rebalancing at time $t+1$. The returns are computed net of transaction costs according to DeMiguel et al. (2009a, p.1930), and we use a proportional transaction cost of 30 basis points that corresponds to the middle value considered by Barroso and Saxena (2021). Results without transaction costs are available in the supplementary materials.

Moment estimation. We consider two types of moment estimators. First, sample moment estimators. Second, shrinkage estimators using a target derived under the assumption that asset returns are iid. Such estimators have been derived by Ledoit and Wolf (2004) for the covariance matrix, and Martellini and Ziemann (2010) and Boudt et al. (2020) for the coskewness and cokurtosis matrices. We use the R package *PerformanceAnalytics* to compute the shrinkage estimators. For the portfolio mean return, we use the estimator of Jorion (1986). It is consistent with the iid target because it shrinks the sample mean return toward the mean return of the GMV portfolio, and the GMV portfolio is obtained by assuming identical asset mean returns. We observe that shrinkage estimators are not useful for datasets with daily returns. This is because, when using daily returns, introducing additional bias may hurt OOS performance (Jagannathan and Ma 2003). However, they largely improve performance for datasets with monthly returns because sample estimators are inefficient for such small sample sizes. Shrinkage estimators also have the effect of reducing turnover, making the portfolios more feasible in practice. For these reasons, we report the results using shrinkage estimators; results for sample estimators are available in the supplementary materials.

3.2. Descriptive statistics

Before discussing the results, we report in Table 2 some descriptive statistics about the first four moments and correlation of asset returns for each dataset. In terms of correlation, the four equity datasets all feature a high correlation, whereas the different asset classes composing the S&P dataset are less correlated. The mean returns for *10IND* are similar, hence the GMV portfolio is close to optimal in terms of Sharpe ratio and the OMVE portfolios are expected to perform similarly in terms of mean-variance tradeoff. However, there is an important dispersion in terms of skewness and kurtosis across the industry portfolios. The statistics are similar for *25BTM* and

Table 2 Descriptive summary statistics for the five datasets

		Daily returns			Monthly returns	
		10IND	*25BTM*	*25Prof*	*MSCI*	*S&P*
Correlation	Min	0.49	0.67	0.69	0.80	-0.01
	Median	0.70	0.85	0.87	0.86	0.15
	Max	0.89	0.95	0.96	0.98	0.55
Mean (annual)	Min	10.43%	4.63%	7.96%	2.52%	1.44%
	Median	11.59%	12.73%	12.40%	4.56%	8.76%
	Max	12.55%	15.34%	14.11%	7.44%	11.16%
Std dev (annual)	Min	14.60%	16.19%	16.67%	15.10%	5.09%
	Median	17.94%	17.62%	17.94%	16.45%	14.96%
	Max	23.34%	22.22%	21.75%	21.41%	22.55%
Skewness	Min	-0.67	-0.85	-0.79	-0.74	-1.70
	Median	-0.31	-0.53	-0.52	-0.60	-0.68
	Max	-0.01	-0.26	-0.33	-0.56	-0.55
Kurtosis	Min	12.20	10.76	10.19	4.19	4.25
	Median	17.37	15.97	17.15	4.62	8.54
	Max	28.09	27.07	22.03	4.78	15.56

Notes. The table reports, for the five datasets used in the empirical analysis, the minimum, median and maximum of the correlation and first four moments across all assets. The first three datasets are made of daily returns on the 10 industry portfolios and 25 size-and-book-to-market and size-and-operating-profitability portfolio datasets spanning January 1970 to May 2020. The last two datasets are made of monthly returns from five MSCI international indices (March 2001 to May 2020) and five S&P asset class indices (January 1995 to May 2020).

25Prof, except that the *25Prof* dataset depicts a larger dispersion in terms of mean return, which offers more opportunity along the MVE frontier. The *MSCI* dataset is peculiar because, except for mean returns, the correlation and second to fourth moments are similar across assets. Hence, it does not offer much optimization benefits a priori. In contrast, the *S&P* dataset is made of several asset classes that depict important differences for all moments.

3.3. Results

Daily return datasets. We begin by discussing the results for the three daily return datasets, reported in Table 3. The results indicate that the OMVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ is a better portfolio on the MVE frontier than the GMV and KWZ portfolios and, moreover, outperforms the GOPT portfolio $\boldsymbol{w}_{\mathcal{C}}$ both in terms of Sharpe ratio and higher moments, with a reduced turnover.

In terms of Sharpe ratio, the OMVE portfolios largely outperform the GOPT portfolios as well as the EW portfolio. Moreover, they have a larger Sharpe ratio than the KWZ portfolio in all cases except one and, for the GCLPM and CRRA objective functions, achieve a larger Sharpe ratio than the GMV portfolio. Thus, by staying on the MVE frontier, we enjoy a strong mean-variance tradeoff

Table 3 Out-of-sample empirical performance for the three daily return datasets

(a) *10IND* dataset				$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$		$\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$		$\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$	
	$\mathbf{1}/N$	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$252 \times \mu_P$	12.76%	12.99%	11.60%	12.78%	12.22%	13.78%	9.24%	13.39%	13.61%
$\sqrt{252} \times \sigma_P$	16.13%	12.29%	14.01%	12.52%	13.85%	12.74%	13.44%	12.55%	12.69%
$\sqrt{252} \times \mu_P/\sigma_P$	0.79	1.06	0.83	1.02	0.88	**1.08**	0.69	1.07	1.07
CFVaR	-7.35%	-6.57%	-6.01%	-6.43%	**-5.79%**	-6.25%	-6.01%	-6.36%	-6.35%
$\sqrt{252\text{GCLPM}}$	12.41%	**9.13%**	10.45%	9.31%	10.43%	9.44%	9.76%	9.36%	9.47%
CRRA×1000	-16.29	**-7.20**	-9.48	-7.44	-8.97	-7.57	-8.50	-7.43	-7.63
Turnover	**0.44%**	1.22%	3.08%	1.69%	2.65%	1.69%	2.81%	1.57%	1.38%
(b) *25BTM* dataset				$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$		$\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$		$\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$	
	$\mathbf{1}/N$	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$252 \times \mu_P$	14.06%	16.62%	**20.17%**	16.46%	14.20%	17.07%	16.41%	17.13%	16.22%
$\sqrt{252} \times \sigma_P$	17.21%	**10.63%**	13.36%	10.64%	11.78%	**10.63%**	11.10%	**10.63%**	10.97%
$\sqrt{252} \times \mu_P/\sigma_P$	0.82	1.56	1.51	1.55	1.21	**1.61**	1.48	**1.61**	1.48
CFVaR	-6.66%	-5.50%	**-3.65%**	-5.50%	-5.51%	-5.49%	-5.28%	-5.34%	-5.25%
$\sqrt{252\text{GCLPM}}$	13.33%	8.09%	9.94%	8.10%	9.00%	**8.02%**	8.39%	**8.02%**	8.17%
CRRA×1000	-17.65	-4.24	-5.51	-4.26	-5.71	-4.20	-4.62	**-4.11**	-4.42
Turnover	**0.32%**	4.23%	11.62%	4.53%	9.06%	4.43%	5.38%	4.53%	5.87%
(c) *25Prof* dataset				$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$		$\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$		$\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$	
	$\mathbf{1}/N$	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$252 \times \mu_P$	13.80%	13.48%	**17.15%**	13.23%	10.87%	13.76%	12.79%	13.76%	12.05%
$\sqrt{252} \times \sigma_P$	17.21%	11.39%	14.47%	11.40%	11.92%	**11.38%**	12.06%	11.39%	12.65%
$\sqrt{252} \times \mu_P/\sigma_P$	0.80	1.18	1.19	1.16	0.91	**1.21**	1.06	**1.21**	0.95
CFVaR	-6.47%	-5.97%	**-4.36%**	-5.87%	-6.28%	-5.94%	-5.62%	-5.84%	-5.57%
$\sqrt{252\text{GCLPM}}$	13.29%	8.43%	10.34%	8.43%	8.99%	8.43%	8.90%	**8.42%**	9.57%
CRRA×1000	-17.16	-5.45	-7.60	-5.40	-6.56	-5.42	-6.07	**-5.36**	-6.96
Turnover	**0.30%**	4.34%	11.35%	4.81%	7.65%	4.43%	6.37%	4.67%	7.03%

Notes. The table reports, using daily returns from the 10 industry portfolios and 25 size-and-book-to-market and size-and-operating-profitability portfolios datasets spanning January 1970 to May 2020, the out-of-sample performance of the EW portfolio $\mathbf{1}/N$, GMV portfolio $\boldsymbol{w}_{GMV}$, optimal MVE portfolio $\boldsymbol{w}_{KWZ}$ of Kan et al. (2021), proposed optimal MVE portfolios $\boldsymbol{w}(\lambda_{\mathcal{C}})$, and global-optimal portfolios $\boldsymbol{w}_{\mathcal{C}}$. We use three higher-moment objective functions $\mathcal{C}(\boldsymbol{w})$: the Cornish-Fisher Value-at-Risk (CFVaR) in (8) with $\alpha = 1\%$, the Gram-Charlier lower-partial-moment (GCLPM) in (21), and the four-moment CRRA utility in (22) with $\gamma = 100$. The out-of-sample performance is evaluated via an estimation window of five years and a testing window of one month that are rolled over time by one month. The portfolio returns are net of transaction costs using a proportional transaction cost of 30 basis points. Moments are estimated via robust shrinkage estimators. More details on the methodology are provided in Section 3.1. The bold figures indicate the best portfolio(s) for each criterion.

Table 4 Out-of-sample empirical performance for the two monthly return datasets

(a) *MSCI* dataset				$\mathcal{C}(\boldsymbol{w})=$ CFVaR		$\mathcal{C}(\boldsymbol{w})=-$GCLPM		$\mathcal{C}(\boldsymbol{w})=$ CRRA	
	$\mathbf{1}/N$	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$12\times\mu_P$	**5.70%**	1.90%	2.37%	3.13%	4.01%	1.93%	-0.52%	2.79%	3.61%
$\sqrt{12}\times\sigma_P$	17.01%	**13.33%**	13.48%	13.55%	14.08%	**13.33%**	14.26%	13.43%	13.91%
$\sqrt{12}\times\mu_P/\sigma_P$	**0.34**	0.14	0.18	0.23	0.28	0.15	-0.04	0.21	0.26
CFVaR	-14.90%	**-10.93%**	-11.23%	-11.07%	-12.41%	-11.03%	-11.49%	-11.00%	-12.24%
$\sqrt{12\text{GCLPM}}$	13.25%	**10.22%**	10.46%	10.45%	11.01%	10.25%	10.89%	10.34%	10.86%
CRRA×10	-14.97	**-5.08**	-5.53	-5.44	-7.47	-5.18	-5.97	-5.27	-7.08
Turnover	**1.40%**	11.34%	18.26%	13.49%	11.38%	11.40%	20.03%	11.95%	7.90%
(b) *S&P* dataset				$\mathcal{C}(\boldsymbol{w})=$ CFVaR		$\mathcal{C}(\boldsymbol{w})=-$GCLPM		$\mathcal{C}(\boldsymbol{w})=$ CRRA	
	$\mathbf{1}/N$	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$12\times\mu_P$	8.28%	7.39%	**8.63%**	8.17%	8.12%	7.57%	7.21%	7.84%	8.00%
$\sqrt{12}\times\sigma_P$	8.53%	**5.66%**	7.62%	6.40%	7.30%	5.71%	5.85%	5.97%	6.45%
$\sqrt{12}\times\mu_P/\sigma_P$	0.97	1.31	1.13	1.28	1.11	**1.33**	1.23	1.31	1.24
CFVaR	-11.70%	-7.30%	-8.03%	-7.60%	-8.98%	-7.16%	**-6.86%**	-7.51%	-7.60%
$\sqrt{12\text{GCLPM}}$	7.41%	**4.91%**	5.68%	5.02%	6.13%	4.94%	4.96%	4.96%	5.40%
CRRA×10	-2.97	-0.62	-1.14	-0.72	-1.37	-0.62	**-0.60**	-0.66	-0.83
Turnover	**2.69%**	3.63%	11.65%	6.49%	4.58%	3.84%	5.92%	4.35%	3.08%

Notes. The table reports, using monthly returns from the MSCI international indices (March 2001 to May 2020) and S&P asset class indices (January 1995 to May 2020) datasets, the out-of-sample performance of the EW portfolio $\mathbf{1}/N$, GMV portfolio $\boldsymbol{w}_{GMV}$, optimal MVE portfolio $\boldsymbol{w}_{KWZ}$ of Kan et al. (2021), proposed optimal MVE portfolios $\boldsymbol{w}(\lambda_{\mathcal{C}})$, and global-optimal portfolios $\boldsymbol{w}_{\mathcal{C}}$. We use three higher-moment objective functions $\mathcal{C}(\boldsymbol{w})$: the Cornish-Fisher Value-at-Risk (CFVaR) in (8) with $\alpha=1\%$, the Gram-Charlier lower-partial-moment (GCLPM) in (21), and the four-moment CRRA utility in (22) with $\gamma=100$. The out-of-sample performance is evaluated via an estimation window of five years and a testing window of one month that are rolled over time by one month. The portfolio returns are net of transaction costs using a proportional transaction cost of 30 basis points. Moments are estimated via robust shrinkage estimators. More details on the methodology are provided in Section 3.1. The bold figures indicate the best portfolio(s) for each criterion.

even after adjusting for transaction costs. In addition to a large Sharpe ratio, the OMVE portfolios perform better than the portfolios that have been fully optimized using the three higher moment criteria considered *in all cases except one*. Thus, it is not useful OOS to move away from the MVE frontier. Moreover, on the frontier, it is preferable to choose the OMVE portfolio over the GMV and KWZ portfolios in terms of higher-moment performance in the majority of cases. Compared to the EW portfolio, the OMVE portfolios always achieve better higher-moment criteria, which shows that naive diversification does not work well when the sample size is large.

Concerning turnover, the OMVE portfolio significantly reduces turnover compared to the GOPT portfolio because all higher-moment estimation risk is transferred to a single parameter $\lambda_{\mathcal{C}}$. It also features a higher turnover than the GMV portfolio, which makes the corresponding outperformance even more material given that we adjust the returns for transaction costs.

Another interesting observation is that, even though we use daily returns for which estimation risk is reduced, the GMV portfolio remains a better alternative to the GOPT portfolio $\boldsymbol{w}_{\mathcal{C}}$. Indeed, it has a larger Sharpe ratio in all cases and achieves a better higher-moment objective $\mathcal{C}(\boldsymbol{w})$ compared to the corresponding portfolio $\boldsymbol{w}_{\mathcal{C}}$ in all cases except one. Along the MVE frontier, the investor should however pick the OMVE portfolio rather than the GMV portfolio as evidenced above.

Monthly return datasets. We turn to the *MSCI* and *S&P* monthly return datasets, whose results are reported in Table 4. Most conclusions drawn from the three daily return datasets above continue to hold. First, the OMVE portfolios generally achieve a larger Sharpe ratio than the GMV and KWZ portfolios. Second, compared to the EW portfolio, they have a substantially larger Sharpe ratio for *S&P* but smaller for *MSCI*. However, for *MSCI*, the OMVE portfolios have a much better higher-moment performance than EW in terms of CFVaR, GCLPM, and CRRA, which is what they are designed for. This is also true for the *S&P* dataset. Third, comparing the OMVE and GOPT portfolios for corresponding criteria $\mathcal{C}(\boldsymbol{w})$, OMVE outperforms GOPT *in all cases*. Thus, once again, there is no benefit in moving away from the MVE frontier. The main difference compared to the case with daily returns is that the OMVE portfolios do not outperform the GMV portfolio in terms of the three criteria, although they perform quite closely. This observation is consistent with Martellini and Ziemann (2010) who find that, for monthly returns, higher-moment portfolios do not achieve better higher-moment performance than the GMV portfolio even when using robust moment estimates. However, on the MVE frontier, the OMVE portfolio remains an appealing choice because it has a better higher-moment performance than the GOPT portfolio and a larger mean return and Sharpe ratio than the GMV portfolio for only a small loss in terms of risk.

3.4. Disentangling between skewness and kurtosis

How important is it to consider kurtosis in investors' preferences? Considering the CRRA objective function in (22) for example, what do the results become if we drop out the $m_{4,P}$ term? This is a relevant question because many papers on higher-moment portfolio selection only consider preferences for mean, variance and skewness; see, e.g., Levy (1969), Athayde and Flores (2004), Briec et al. (2007), Adcock (2014), and Boudt et al. (2020).

To examine this question, we compute for each dataset the OMVE and GOPT portfolios, $\boldsymbol{w}(\lambda_{\mathcal{C}})$ and $\boldsymbol{w}_{\mathcal{C}}$, for the CRRA objective function in (22) without the $m_{4,P}$ term. We then report their OOS performance in terms of CRRA computed with and without the $m_{4,P}$ term. We also report results for the GMV portfolio as a benchmark. The results are reported in Table 5 and there are two main observations. First, looking at how the portfolio strategies manage to maximize the CRRA without the kurtosis term, the relative ordering of the GMV, OMVE and GOPT portfolios is similar to that

Table 5 Out-of-sample empirical performance when optimizing the CRRA utility without the kurtosis term

	CRRA with $m_{4,P}$			CRRA without $m_{4,P}$		
	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
10IND	-7.20	-8.04	-10.30	-2.79	-3.16	-3.37
25BTM	-4.24	-3.97	-4.03	-1.88	-1.75	-1.83
25Prof	-5.45	-5.29	-5.44	-2.26	-2.23	-2.29
MSCI	-5.08	-5.21	-5.33	-1.35	-1.39	-1.42
S&P	-0.62	-0.62	-0.61	-0.21	-0.22	-0.22

Notes. The table reports, for the five empirical datasets, the impact of optimizing the CRRA utility in (22), where $\gamma = 100$, without the kurtosis term $m_{4,P}$. We proceed as for the results in Tables 3 and 4 except that the proposed optimal MVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ and global-optimal portfolio $\boldsymbol{w}_{\mathcal{C}}$ are computed by optimizing the CRRA utility without the $m_{4,P}$ term. We then report the out-of-sample performance in terms of CRRA computed with and without the $m_{4,P}$ term. The CRRA is multiplied by 1000 for the datasets with daily returns (*10IND*, *25BTM*, *25Prof*), and by 10 for the datasets with monthly returns (*MSCI*, *S&P*).

with the CRRA including the kurtosis term in Tables 3 and 4. However, there are generally smaller differences between the portfolios than when allowing for optimization of the kurtosis in addition to skewness. Second, we interestingly observe that optimizing the CRRA without the kurtosis term can actually lead to a larger CRRA even when it is evaluated with the kurtosis term. This is for example the case with the OMVE portfolio for all datasets except *10IND*. Thus, investors who care about the first four moments may actually be better off ignoring the fourth moment in the optimization. This can be explained from the large estimation errors in estimates of kurtosis (Martellini and Ziemann 2010). Ignoring both skewness and kurtosis is however not preferable because we observed in Section 3.3 that the OMVE portfolio is overall a better option than the GMV portfolio.

4. Monte Carlo simulations

In this section, we use simulated data to assess more formally the tradeoff between specification and estimation error of the proposed OMVE approach, and to test whether the empirical conclusions hold on a large number of Monte Carlo simulated scenarios. We explain how we measure the tradeoff in Section 4.1, describe the methodology in Section 4.2, illustrate for a simple case of $N = 3$ assets in Section 4.3, and discuss the results in Section 4.4.

4.1. Specification and estimation error

To distinguish between specification and estimation error, we follow the expected-utility-loss approach introduced by Ter Horst et al. (2006) and Kan and Zhou (2007). Given $\boldsymbol{w}_{\mathcal{C}}$ the true optimal portfolio given the criterion $\mathcal{C}(\boldsymbol{w})$, $\boldsymbol{w}$ an arbitrary portfolio, and $\hat{\boldsymbol{w}}$ a sample estimate of $\boldsymbol{w}$, we measure the expected performance loss associated to $\hat{\boldsymbol{w}}$ as

$$\mathcal{L}(\boldsymbol{w}_{\mathcal{C}}, \hat{\boldsymbol{w}}) = \mathbb{E}[\mathcal{C}(\boldsymbol{w}_{\mathcal{C}}) - \mathcal{C}(\hat{\boldsymbol{w}})] = \underbrace{(\mathcal{C}(\boldsymbol{w}_{\mathcal{C}}) - \mathcal{C}(\boldsymbol{w}))}_{\text{Specification error}} + \underbrace{(\mathcal{C}(\boldsymbol{w}) - \mathbb{E}[\mathcal{C}(\hat{\boldsymbol{w}})])}_{\text{Estimation error}}, \tag{24}$$

where the expectation is taken under the true distribution of asset returns. The first term, specification error, measures the performance bias of the portfolio $\boldsymbol{w}$. The second term, estimation error, measures on average the loss in performance due to parameter uncertainty in $\hat{\boldsymbol{w}}$ compared to $\boldsymbol{w}$.

The optimal portfolio $\hat{\boldsymbol{w}}$ minimizes the expected loss $\mathcal{L}(\boldsymbol{w}_{\mathcal{C}}, \hat{\boldsymbol{w}})$, which amounts to maximizing the expected performance $\mathbb{E}[\mathcal{C}(\hat{\boldsymbol{w}})]$. As the results of Section 4.4 exemplify, the OMVE portfolio often achieves the best tradeoff for different values of the number of assets N and sample size T.

4.2. Data, portfolio strategies and performance criterion

We consider a non-Gaussian data-generation process and, for different N and T, we compute the expected performance loss $\mathcal{L}(\boldsymbol{w}_{\mathcal{C}}, \hat{\boldsymbol{w}})$ of four portfolio strategies $\hat{\boldsymbol{w}}$ averaged over 1000 Monte Carlo simulations, which corresponds to the number of simulations used in Simaan (2014).

The four portfolio strategies are: the GMV portfolio $\boldsymbol{w}_{GMV}$, the MSR portfolio $\boldsymbol{w}_{MSR}$, our proposed OMVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ using the CFVaR in (8) with $\alpha = 1\%$ as objective, and the GOPT portfolio $\boldsymbol{w}_{\mathcal{C}}$ in (6). We consider the case with no risk-free asset as in the empirical analysis and short-selling is allowed. For brevity, we restrict to the CFVaR objective function. Because we use daily returns below, we also restrict to sample moment estimates. In the empirical analysis of Section 3, we consider two additional objective functions and robust shrinkage estimates.

We consider a non-Gaussian data-generating process similar to Section 4 in Lassance and Vrins (2021), to which we refer for details. In summary, we simulate $T+\tau$ daily asset returns $\boldsymbol{X}$, where T is the sample size used to estimate the portfolios, and τ is the length of the OOS period. We fix τ to be 20 years of daily returns, whereas T will be varied in our experiments from 2 to 10 years. We consider daily returns because they improve estimation accuracy and allow a fairer comparison with the GOPT portfolio. As commonly done in the literature for Gaussian data-generation processes (MacKinlay and Pastor 2000, DeMiguel et al. 2009a), we simulate the returns from a factor model. Specifically, we simulate our $T + \tau$ asset returns via a non-Gaussian factor model of the form

$$\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{Y} + \boldsymbol{\epsilon}, \tag{25}$$

where $\boldsymbol{\mu}$ is of size $N \times (T+\tau)$ with each row being constant, the loading matrix $\boldsymbol{A}$ is of size $N \times K$, the K factors $\boldsymbol{Y}$ are of size $K \times (T+\tau)$ and are uncorrelated, and the residuals $\boldsymbol{\epsilon}$ are of size $N \times (T+\tau)$ and are independent from the factors $\boldsymbol{Y}$. The factors $\boldsymbol{Y}$ have zero mean and are simulated from a Clayton copula with Student t marginals. The residuals $\boldsymbol{\epsilon}$ have zero mean and are simulated from a Gaussian distribution. They have an annualized standard deviation drawn uniformly between 5% and 15%. The loading matrix $\boldsymbol{A}$ is taken as the eigendecomposition of a randomly simulated covariance matrix with annualized standard deviations drawn uniformly between 10% and 30%, twice larger on average than that of the residuals. Contrary to Lassance and Vrins (2021), we do not take zero-mean asset returns. The mean return of each asset, μ_i, is randomly sampled from a uniform distribution between 0.05/252 and 0.35/252. We fix the number of factors at $K = 5$, and we consider a number of assets $N = 10$, 30 and 50.

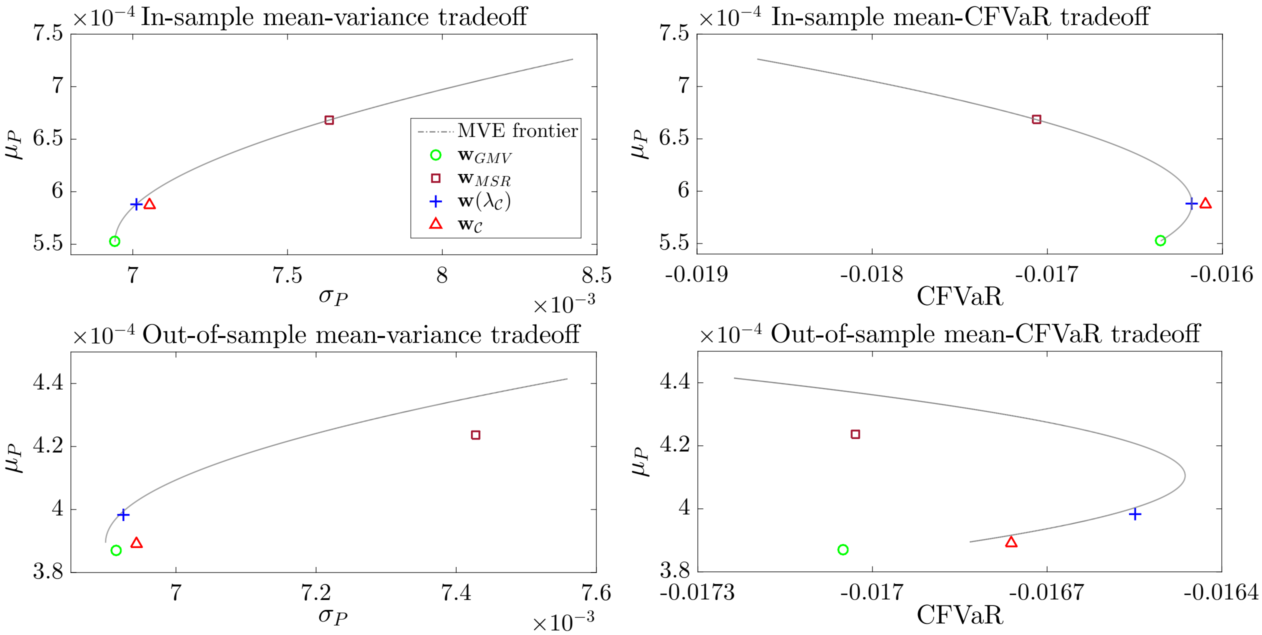
Finally, to measure the expected performance loss $\mathcal{L}(\boldsymbol{w}_{\mathcal{C}}, \hat{\boldsymbol{w}})$ in (24), we need to compute $\mathcal{C}(\boldsymbol{w}_{\mathcal{C}})$ and $\mathbb{E}[\mathcal{C}(\hat{\boldsymbol{w}})]$, where $\mathcal{C}(\boldsymbol{w})$ is the CFVaR in (8). Given that the distribution of asset returns changes randomly for each Monte Carlo draw, we measure $\mathcal{C}(\boldsymbol{w}_{\mathcal{C}})$ by taking the average over all 1000 simulations for the CFVaR-optimal portfolio on the OOS returns. Regarding $\mathbb{E}[\mathcal{C}(\hat{\boldsymbol{w}})]$, it is measured by estimating $\hat{\boldsymbol{w}}$ on the in-sample returns, computing the CFVaR $\mathcal{C}(\hat{\boldsymbol{w}})$ on the OOS returns, and averaging over all 1000 simulations. In addition to the CFVaR, it is also of practical interest to report the expected Sharpe ratio loss of all considered portfolio strategies with respect to the Sharpe ratio of the MSR portfolio computed on the OOS returns.

4.3. Illustration for three assets

Before turning to the main results of the simulation, we illustrate the framework for a particular draw in the simple case with $N = 3$ assets, $K = 2$ factors and a sample size of $T = 2$ years.

In the top two plots of Figure 3, we depict the in-sample daily mean-standard-deviation and mean-CFVaR tradeoff attained by the four portfolios estimated in-sample, which correspond to $\boldsymbol{w}_{GMV} = (0.67, 0.36, -0.03)$, $\boldsymbol{w}_{MSR} = (0.91, 0.47, -0.38)$, $\boldsymbol{w}(\lambda_{\mathcal{C}}) = (0.75, 0.39, -0.14)$, and $\boldsymbol{w}_{\mathcal{C}} = (0.80, 0.37, -0.17)$. We also depict the tradeoff for portfolios on the in-sample MVE frontier. We see in particular that the GOPT portfolio $\boldsymbol{w}_{\mathcal{C}}$ that maximizes the in-sample CFVaR is located outside of the MVE frontier but, as a result, attains a better CFVaR than the best MVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$. Then, in the bottom two plots of Figure 3, we depict the OOS daily mean-standard-deviation and mean-CFVaR tradeoff attained by the four portfolios estimated in-sample. We also depict the OOS MVE frontier. We see in particular that the in-sample portfolios are no longer located on the MVE

Figure 3 In-sample and out-of-sample performance for three simulated asset returns



Notes. The top two plots depict the in-sample mean-variance and mean-Cornish-Fisher-Value-at-Risk (CFVaR) tradeoff of the GMV portfolio $\boldsymbol{w}_{GMV}$ (green circles), the MSR portfolio $\boldsymbol{w}_{MSR}$ (purple squares), our proposed optimal MVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ (blue crosses), and the global-optimal portfolio $\boldsymbol{w}_{\mathcal{C}}$ (red triangles). We use simulated data for $N = 3$ asset returns using a particular draw from the random sampling described in Section 4.2. The portfolios are estimated on a window of $T = 2$ years of daily returns. The bottom two plots depict the out-of-sample mean-variance and mean-CFVaR tradeoff of the in-sample portfolios, which is evaluated on a separate window of $\tau = 20$ years of daily returns drawn from the same distribution as for the in-sample returns.

frontier OOS and it is now the OMVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ that attains the largest CFVaR OOS. This exemplifies that moving away from the MVE frontier may not be optimal OOS, as the empirical analysis and the results of the next section show.

4.4. Results

In Figure 4, we report the expected daily CFVaR and Sharpe ratio loss achieved by the considered portfolios, averaged over the 1000 simulations. Recall that the portfolios are estimated based on the returns from $t = 1$ to $t = T$, and that the expected loss is evaluated on the OOS returns from $t = T + 1$ to $t = T + \tau$. Because the MSR portfolio $\boldsymbol{w}_{MSR}$ always performs much worse than the competitors in terms of expected CFVaR loss, we do not report its performance in the figure.

First, in terms of the CFVaR objective function, our proposed OMVE strategy leads to the smallest expected OOS loss. It is only for a small $N = 10$ that the GOPT portfolio has a smaller loss once the sample size is large enough (7 and 10 years). This is consistent with the empirical results in Table 3 that show that the GOPT portfolio has a better OOS CFVaR than the OMVE portfolio for the *10IND* dataset and a sample size of five years. Otherwise, for $N = 30$ and $N = 50$, the OMVE portfolio has the smallest loss in all considered cases. Referring to (24), this means that the OMVE portfolio achieves a better tradeoff between specification and estimation error than the GMV and GOPT portfolios, which is consistent with the overall empirical conclusions obtained for the daily return datasets in Table 3. The outperformance compared to the GOPT portfolio is particularly visible when the sample size is small, but remains even until $T = 10$ years for $N = 30$ and 50. This is striking because estimation risk is reduced in our Monte Carlo setting as the asset returns are drawn from the same distribution for the in-sample and OOS periods.

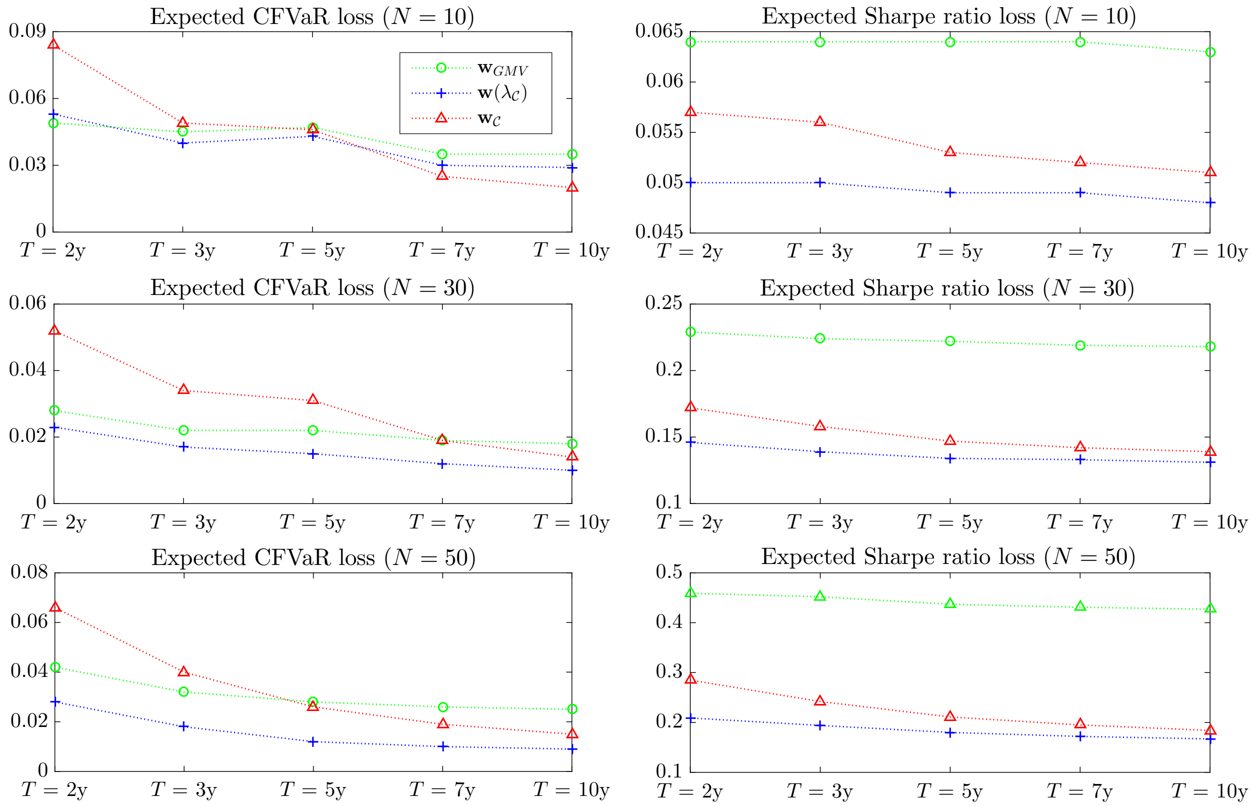
Second, although the OMVE portfolio does not explicitly aim to optimize the Sharpe ratio, it has a smaller expected Sharpe ratio loss OOS than the GMV and GOPT portfolios, which is an important additional benefit of our approach. This can be explained as follows. Contrary to the OMVE and GOPT portfolios, the GMV portfolio ignores information about mean returns, which results in the largest Sharpe ratio loss. Then, among the OMVE and GOPT portfolios, the OMVE portfolio has a smaller Sharpe ratio loss because, by design, it is located on the MVE frontier. This is consistent with the empirical analysis in which we also found that the OMVE portfolio often has a larger Sharpe ratio than the GOPT and GMV portfolios.

In summary, there is no OOS benefit in moving away from the MVE frontier. However, one is better off choosing the optimal portfolio on the frontier rather than the commonly used GMV and MSR portfolios. The results confirm the general conclusions drawn in the empirical analysis, and even more strongly in our favor because the simulated data do not depict the low-risk anomaly observed in real data, which tends to inflate the OOS performance of the GMV portfolio.

5. Conclusion

The main conclusion of the paper is that mean-variance portfolio theory is not incompatible with non-Gaussian returns as long as the right mean-variance-efficient portfolio is chosen. From an out-of-sample perspective, we find no benefit in moving away from the efficient frontier: by adjusting the quadratic-utility risk-tolerance coefficient to improve higher moments, we obtain a portfolio strategy that often achieves a better tradeoff between specification error and estimation error than the global-optimal portfolio, and the GMV portfolio when the sample size is large enough, while retaining a large Sharpe ratio. This result has useful implications for investors who may be willing to rely on mean-variance portfolio theory, but who may fear that the obtained portfolio has undesirable higher-moment performance.

Figure 4 Expected performance loss via Monte Carlo simulations



Notes. The three left plots depict, for the Cornish-Fisher Value-at-Risk (CFVaR) with $\alpha = 1\%$ in (8) as objective function $\mathcal{C}(\boldsymbol{w})$, the expected performance loss of the GMV portfolio $\boldsymbol{w}_{GMV}$ (green circles), our proposed optimal MVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ (blue crosses), and the global-optimal portfolio $\boldsymbol{w}_{\mathcal{C}}$ (red triangles) for the Monte Carlo simulations of Section 4. The expected CFVaR loss is computed from (24) as the difference (multiplied by 100) between the optimal out-of-sample CFVaR and the out-of-sample CFVaR achieved by the portfolio strategies estimated in-sample. Similarly, the three right plots depict the expected Sharpe ratio loss. The way the daily asset returns are randomly sampled is described in Section 4.2. The portfolios are estimated on a window of T returns reported on the x-axis, which goes from 2 to 10 years of daily returns. The out-of-sample performance is evaluated on a different window of $\tau = 20$ years of daily returns. We consider a number of assets $N = 10$, 30 and 50. The expected performance loss is averaged over 1000 Monte Carlo simulations.

The paper has shown that methods proposed by the literature on portfolio selection with parameter uncertainty, which advocates the use of in-sample misspecified portfolios to improve performance, are also useful to improve higher moments. Future research may look at other applications coming from this literature. For example, when a risk-free asset is available, the results in Section 2.4.3 suggest that further combining the OMVE portfolio with the GMV portfolio as in Kan and Zhou (2007) can be beneficial. This can be done in a two-stage process in which one first identifies the OMVE portfolio and then optimally combines it with the GMV portfolio. The present setting can also be extended to a dynamic higher-moment portfolio selection problem for which the same question can be posed: can a misspecified portfolio outperform a correctly specified multiperiod optimal portfolio out-of-sample due to estimation risk? To that end, one can rely on

DeMiguel et al. (2015) who extend the Kan-Zhou methodology to a multiperiod setting.

Acknowledgments

Part of this work was conducted while the author was supported by the Fonds de la Recherche Scientifique (F.R.S.-FNRS) via the ASP grant [grant number FC 17775].

References

Adcock, C. (2014). Mean–variance–skewness efficient surfaces, Stein's lemma and the multivariate extended skew-Student distribution.*European Journal of Operational Research*, 234(2), 392–401.

Alexander, G., & Baptista, A. (2004). A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model. *Management Science*, 50, 1261–1273.

Ang, A., Chen, J., & Yu, H. (2006). Downside risk. *Review of Financial Studies*, 19(4), 1191–1239.

Athayde, G., & Flores, R. (2004). Finding a maximum skewness portfolio: A general solution to three-moments portfolio choice. *Journal of Economic Dynamics and Control*, 28(7), 1335–1352.

Bali, T., Brown, S., & Demirtas, K. (2013). Do hedge funds outperform stocks and bonds? *Management Science*, 59(8), 1887–1903.

Barroso, P., & Saxena, K. (2021). Lest we forget: Learn from out-of-sample forecast errors when optimizing portfolios. *Review of Financial Studies*, forthcoming.

Bauder, D., Bodnar, R., Bodnar, T., & Schmid, W. (2018). Bayesian estimation of the efficient frontier. *Scandinavian Journal of Statistics*, 46(3), 802–830.

Bernard, C., De Vecchi, C., & Vanduffel, S. (2020). When do two- or three-fund separation theorems hold? SSRN3703665.

Boudt, K., Cornilly, D., & Verdonck, T. (2020). A coskewness shrinkage approach for estimating the skewness of linear combinations of random variables. *Journal of Financial Econometrics*, 18(1), 1–23.

Brandtner, M., Kürsten, W., & Rischau, R. (2020). Beyond expected utility: Subjective risk aversion and optimal portfolio choice under convex shortfall risk measures. *European Journal of Operational Research*, 285(3), 1114–1126.

Briec, W., Kerstens, K., & Jokung, O. (2007). Mean-variance-skewness portfolio performance gauging: A general shortage function and dual approach. *Management Science*, 53(1), 135–149.

Briec, W., Kerstens, K., & Van de Woestyne, I. (2013). Portfolio selection with skewness: A comparison of methods and a generalized one fund result. *European Journal of Operational Research*, 230(2), 412–421.

Chen, J., & Yuan, M. (2016). Efficient portfolio selection in a large market. *Journal of Financial Econometrics*, 14(3), 496–524.

Cogneau, P., & Hubner, G. (2014). The (more than) 100 ways to measure portfolio performance. *Journal of Performance Measurement*, 14, 56–69.

Dahlquist, M., Farago, A., & Tédongap, R. (2017). Asymmetries and portfolio choice. *Review of Financial Studies*, 30(2), 667–702.

DeMiguel, V., Garlappi, L., Nogales, F., & Uppal, R. (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, 55(5), 798–812.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009a). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, 22(5), 1915–1953.

DeMiguel, V., Martìn-Utrera, A., & Nogales, F. (2015). Parameter uncertainty in multiperiod portfolio optimization with transaction costs. *Journal of Financial and Quantitative Analysis*, 50(6), 1443–1471.

DeMiguel, V., & Nogales, F. (2009) Portfolio selection with robust estimation. *Operations Research*, 55(5), 798–812.

El Ghaoui, L., Oks, M., & Oustry, F. (2003). Worst-case value-At-risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51(4), 543–556.

Eling, M., & Schuhmacher, F. (2007). Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking and Finance*, 31(9), 2632–2647.

Favre, L., & Galeano, J. (2002). Mean-modified value-at-risk optimization with hedge funds. *Journal of Alternative Investments*, 5(2), 21–25.

Ghalanos, A., Rossi, E., & Urga, G. (2015). Independent factor autoregressive conditional density model. *Econometric Reviews*, 34(5), 594–616.

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

Gormsen, N., & Jensen, C. (2020). Higher-moment risk. SSRN3069617.

Guidolin, M., & Nicodano, G. (2009). Small caps in international equity portfolios: The effects of variance risk. *Annals of Finance*, 5(1), 15–48.

Guidolin, M., Pedio, M., & Petrova, M. (2020). The predictability of real estate excess returns: An out-of-sample economic value analysis. *Journal of Real Estate Financial Economics*.

Guidolin, M., & Timmermann, A. (2008). International asset allocation under regime switching, skew, and kurtosis preferences. *Review of Financial Studies*, 21(2), 889–935.

Harvey, C., Liechty, J., Liechty, M., & Müller, P. (2010). Portfolio selection with higher moments. *Quantitative Finance*, 10(5), 469–485.

Jagannathan, R., & Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance*, 58(4), 1651–1684.

Jeon, J., & Shin, Y. (2019). Finite horizon portfolio selection with a negative wealth constraint. *Journal of Computational and Applied Mathematics*, 356, 329–338.

Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21(3), 279–292.

Kan, R., Wang, X., & Zhou, G. (2021). Optimal portfolio choice with estimation risk: No risk-free asset case. *Management Science*, forthcoming.

Kan, R., & Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, 42(3), 621–656.

Khashanah, K., Simaan, M., & Simaan, Y. (2020). Do we need higher-order comoments to enhance mean-variance portfolios? SSRN3523379.

Kroll, Y., Levy, H., & Markowitz, H. (1984). Mean-variance versus direct utility maximization. *Journal of Finance*, 39(1), 47–61.

Lassance, N. (2020). Information-theoretic approaches to portfolio selection. *PhD Thesis*, Louvain School of Management.

Lassance, N., & Vrins, F. (2021). Portfolio selection with parsimonious higher comoments estimation. *Journal of Banking and Finance*, 126(9), 106–115.

Ledoit, O., & Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2), 365–411.

Ledoit, O., & Wolf, M. (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks. *Review of Financial Studies*, 30(12), 4349–4388.

Leòn, A., & Moreno, M. (2017). One-sided performance measures under Gram-Charlier distributions. *Journal of Banking and Finance*, 74, 38–50.

Levy, H. (1969). A utility function depending on the first three moments. *Journal of Finance*, 24(4), 715–719.

Levy, H., & Levy, M. (2014). The benefits of differential variance-based constraints in portfolio optimization. *European Journal of Operational Research*, 234(2), 372–381.

Levy, H., & Markowitz, H. (1979). Approximating expected utility by a function of mean and variance. *American Economic Review*, 69(3), 308–317.

Lwin, K., Qu, R., & MacCarthy, B. (2017). Mean-VaR portfolio optimization: A nonparametric approach. *European Journal of Operational Research*, 260(2), 751–766.

MacKinlay, A., & Pastor, L. (2000). Asset pricing models: Implications for expected returns and portfolio selection. *Review of Financial Studies*, 13(4), 883–916.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36(4), 417–492.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1), 77–91.

Markowitz, H. (2014). Mean–variance approximations to expected utility. *European Journal of Operational Research*, 234(2), 346–355.

Martellini, L., & Ziemann, V. (2010). Improved estimates of higher-order comoments and implications for portfolio selection. *Review of Financial Studies*, 23(4), 1467–1502.

Price, K., Price, B., & Nantell, T. (1982). Variance and lower partial moment measures of systematic risk: Some analytical and empirical results. *Journal of Finance*, 37(3), 843–855.

Pulley, L. (1979). Mean-variance approximations to expected logarithmic utility. *Operations Research*, 31(4), 685–696.

Scott, R., & Horvath, P. (1980). On the direction of preference for moments of higher order than the variance. *Journal of Finance*, 35(4), 915–919.

Simaan, Y. (1993). Portfolio selection and asset pricing-three-parameter framework. *Management Science*, 39(5), 568–577.

Simaan, Y. (1993a). What is the opportunity cost of mean-variance investment strategies? *Management Science*, 39(5), 578–587.

Simaan, Y. (2014). The opportunity cost of mean-variance choice under estimation risk. *European Journal of Operational Research*, 234(2), 382–391.

Sortino, F., & van der Meer, R. (1991). Downside risk. *Journal of Portfolio Management*, 17(4), 27–31.

Ter Horst, J., De Roon, F., & Werker, B. (2006). Incorporating estimation risk in portfolio choice. In P. Duffhues & L. Renneboog (Eds.), *Advances in Corporate Finance and Asset Pricing*, Emerald Group Publishing.

Supplementary materials to "Reconciling mean-variance portfolio theory with non-Gaussian returns"

Nathan Lassance
UCLouvain (BE), Louvain Finance, Louvain Institute of Data Analysis and Modeling
Chaussée de Binche 151, 7000 Mons, Belgium
nathan.lassance@uclouvain.be

A. Proofs of theoretical results

A.1. Proof of Proposition 1

We need to find the risk-tolerance coefficient $\lambda_{\mathcal{C}}$ such that $\boldsymbol{w}(\lambda_{\mathcal{C}}) = \lambda_{\mathcal{C}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_e$ maximizes $\mathcal{C}(\boldsymbol{w}(\lambda_{\mathcal{C}}))$ in (11). Multiplying $\mathcal{C}(\boldsymbol{w}(\lambda_{\mathcal{C}}))$ by -1 and taking the log, we need to find $\lambda_{\mathcal{C}}$ that minimizes

$$-\lambda_{\mathcal{C}}\boldsymbol{\mu}_e'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_{\boldsymbol{G}} + \frac{\lambda_{\mathcal{C}}^2}{2\lambda}\boldsymbol{\mu}_e'\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{G}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_e + \lambda\ln\left(1-\pi+\pi e^{-\lambda_{\mathcal{C}}\boldsymbol{\mu}_e'\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{\gamma}}/\lambda}\right).$$

Taking the differential with respect to $\lambda_{\mathcal{C}}$ and setting it to zero yields the first-order condition

$$\boldsymbol{\mu}_e'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_{\boldsymbol{G}} - \frac{\lambda_{\mathcal{C}}}{\lambda}\boldsymbol{\mu}_e'\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{G}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_e + \phi(\lambda_{\mathcal{C}})\boldsymbol{\mu}_e'\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{\gamma}} = 0, \tag{A.1}$$

where the function $\phi(x)$ is defined in (12). Then, using the linear approximation of $\phi(x)$ in (13), the condition (A.1) becomes

$$\boldsymbol{\mu}_e'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_{\boldsymbol{G}} - \frac{\lambda_{\mathcal{C}}}{\lambda}\boldsymbol{\mu}_e'\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{G}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_e + \pi\boldsymbol{\mu}_e'\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{\gamma}} - \frac{\lambda_{\mathcal{C}}}{\lambda}\pi(1-\pi)(\boldsymbol{\mu}_e'\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{\gamma}})^2 = 0. \tag{A.2}$$

Isolating $\lambda_{\mathcal{C}}$ in (A.2) and noticing that $\boldsymbol{\mu}_e'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_{\boldsymbol{G}} + \pi\boldsymbol{\mu}_e'\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{\gamma}} = \boldsymbol{\mu}_e'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_e$ yields $\lambda_{\mathcal{C}} = \phi_2\lambda$, where ϕ_2 is defined in (18). Because $\lambda\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_e = \boldsymbol{w}(\lambda)$ in (2), we indeed have that $\boldsymbol{w}(\lambda_{\mathcal{C}}) = \phi_2\boldsymbol{w}(\lambda)$.

A.2. Gram-Charlier expansion of the lower-partial-moment

In this section, we prove the formula for the GCLPM in (21). For a general τ, Leòn and Moreno (2017) show that the four-moment Gram-Charlier expansion of $\mathrm{LPM}(\tau, m=2)$ in (20) is given by

$$\mathrm{LPM}(\tau,2) \approx \mathrm{LPM}_\phi(\tau,2) + \frac{\zeta_P}{\sqrt{6}}\theta_{2,2} + \frac{\kappa_P}{\sqrt{24}}\theta_{3,2}, \tag{A.3}$$

where

$$\begin{aligned}
\mathrm{LPM}_\phi(\tau,2) &= (\tau-\mu_P)\Phi(\tau^\star) + (\tau-\mu_P)\sigma_P\phi(\tau^\star) + \sigma_P^2\Phi(\tau^\star),\\
\theta_{j,2} &= (\tau-\mu_P)^2 A_{0,j} - 2(\tau-\mu_P)\sigma_P A_{1,j} + \sigma_P^2 A_{2,j},\\
A_{k2} &= \frac{1}{\sqrt{6}}(B_{k+3} - 3B_{k+1}),\\
A_{k3} &= \frac{1}{\sqrt{24}}(B_{k+4} - 6B_{k+2} + 3B_k),\\
B_k &= \int_{-\infty}^{\tau^\star} x^k\phi(x)\mathrm{d}x,
\end{aligned} \tag{A.4}$$

with ϕ and Φ the pdf and cdf of the standard Gaussian, and $\tau^\star := (\tau-\mu_P)/\sigma$. For $\tau = \mu_P$, and thus $\tau^\star = 0$, (A.3)–(A.4) reduces to

$$\mathrm{LPM}(\mu_P,2) \approx \sigma_P^2\left(\frac{1}{2} + \frac{\zeta_P}{6}(B_5 - 3B_3) + \frac{\kappa_P}{24}(B_6 - 6B_4 + 3B_2)\right), \quad B_k = \int_{-\infty}^{0} x^k\phi(x)\mathrm{d}x. \tag{A.5}$$

Because, for $\tau^\star = 0$, $B_2 = 1/2$, $B_3 = \sqrt{2/\pi}$, $B_4 = 3/2$, $B_5 = -4\sqrt{2/\pi}$ and $B_6 = 15/2$, we have $B_6 - 6B_4 + 3B_2 = 0$ and $B_5 - 3B_3 = -\sqrt{2/\pi}$, resulting in (21).

B. Additional empirical results

We provide additional tables related to the empirical analysis of Section 3. Tables I and II report the OOS performance without adjusting for transaction costs. Tables III and IV report, adjusting for transaction costs, the OOS empirical performance when using sample estimates of portfolio return moments instead of shrinkage estimates. We do not report the equally weighted portfolio in the last two tables because it is not affected by the choice of moment estimator.

References

Leòn, A., & Moreno, M. (2017). One-sided performance measures under Gram-Charlier distributions. *Journal of Banking and Finance*, 74, 38–50.

Table I Out-of-sample empirical performance for the three daily return datasets without transaction costs

(a) *10IND* dataset				$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$		$\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$		$\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$	
	$\mathbf{1}/N$	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$252 \times \mu_P$	13.09%	13.91%	13.94%	14.06%	14.23%	15.06%	11.36%	14.58%	14.65%
$\sqrt{252} \times \sigma_P$	16.13%	12.29%	13.99%	12.51%	13.80%	12.73%	13.36%	12.55%	12.68%
$\sqrt{252} \times \mu_P/\sigma_P$	0.81	1.13	1.00	1.12	1.03	1.18	0.85	1.16	1.16
CFVaR	-7.35%	-6.55%	-5.99%	-6.41%	-5.75%	-6.24%	-5.95%	-6.34%	-6.32%
$\sqrt{252}$GCLPM	12.40%	9.11%	10.42%	9.29%	10.34%	9.42%	9.60%	9.34%	9.44%
CRRA $\times 1000$	-16.27	-7.14	-9.34	-7.36	-8.75	-7.50	-8.20	-7.35	-7.55
Turnover	0.44%	1.22%	3.08%	1.69%	2.65%	1.69%	2.81%	1.57%	1.38%
(b) *25BTM* dataset				$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$		$\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$		$\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$	
	$\mathbf{1}/N$	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$252 \times \mu_P$	14.30%	19.83%	28.97%	19.89%	21.05%	20.42%	20.48%	20.57%	20.67%
$\sqrt{252} \times \sigma_P$	17.21%	10.61%	13.26%	10.62%	11.50%	10.61%	11.05%	10.61%	10.90%
$\sqrt{252} \times \mu_P/\sigma_P$	0.83	1.87	2.19	1.87	1.83	1.93	1.85	1.94	1.90
CFVaR	-6.66%	-5.41%	-3.55%	-5.41%	-5.34%	-5.41%	-5.21%	-5.25%	-5.12%
$\sqrt{252}$GCLPM	13.33%	7.98%	9.77%	7.99%	8.59%	7.91%	8.24%	7.90%	8.00%
CRRA $\times 1000$	-17.64	-4.01	-4.94	-4.01	-4.91	-3.96	-4.32	-3.86	-4.05
Turnover	0.32%	4.23%	11.62%	4.53%	9.06%	4.43%	5.38%	4.53%	5.87%
(c) *25Prof* dataset				$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$		$\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$		$\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$	
	$\mathbf{1}/N$	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$252 \times \mu_P$	14.03%	16.76%	25.76%	16.87%	16.67%	17.12%	17.62%	17.29%	17.37%
$\sqrt{252} \times \sigma_P$	17.21%	11.36%	14.40%	11.37%	11.74%	11.36%	11.97%	11.37%	12.44%
$\sqrt{252} \times \mu_P/\sigma_P$	0.82	1.47	1.79	1.48	1.42	1.51	1.47	1.52	1.40
CFVaR	-6.46%	-5.92%	-4.27%	-5.82%	-6.32%	-5.89%	-5.55%	-5.79%	-5.47%
$\sqrt{252}$GCLPM	13.29%	8.33%	10.12%	8.33%	8.72%	8.32%	8.68%	8.32%	9.19%
CRRA $\times 1000$	-17.15	-5.23	-7.05	-5.16	-6.06	-5.20	-5.67	-5.13	-6.28
Turnover	0.30%	4.34%	11.35%	4.81%	7.65%	4.43%	6.37%	4.67%	7.03%

Notes. See notes of Table 3. The only difference is that we do not adjust the performance to transaction costs.

Table II Out-of-sample empirical performance for the two monthly return datasets without transaction costs

(a) *MSCI* dataset				$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$		$\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$		$\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$	
	$\mathbf{1}/N$	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$12 \times \mu_P$	5.75%	2.30%	3.02%	3.61%	4.41%	2.34%	0.19%	3.22%	3.89%
$\sqrt{12} \times \sigma_P$	17.01%	13.32%	13.49%	13.53%	14.03%	13.33%	14.21%	13.42%	13.88%
$\sqrt{12} \times \mu_P/\sigma_P$	0.34	0.17	0.22	0.27	0.31	0.18	0.01	0.24	0.28
CFVaR	-14.90%	-10.89%	-11.18%	-10.99%	-12.20%	-10.98%	-11.37%	-10.93%	-12.10%
$\sqrt{12\text{GCLPM}}$	13.24%	10.21%	10.46%	10.42%	10.93%	10.24%	10.83%	10.33%	10.81%
CRRA $\times 10$	-14.97	-5.05	-5.53	-5.38	-7.19	-5.16	-5.87	-5.23	-6.90
Turnover	1.40%	11.34%	18.26%	13.49%	11.38%	11.40%	20.03%	11.95%	7.90%
(b) *S&P* dataset				$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$		$\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$		$\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$	
	$\mathbf{1}/N$	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$12 \times \mu_P$	8.37%	7.52%	8.63%	8.40%	8.28%	7.71%	7.42%	7.99%	8.11%
$\sqrt{12} \times \sigma_P$	8.52%	5.66%	7.62%	6.40%	7.29%	5.71%	5.82%	5.97%	6.44%
$\sqrt{12} \times \mu_P/\sigma_P$	0.98	1.33	1.13	1.31	1.14	1.35	1.28	1.34	1.26
CFVaR	-11.68%	-7.29%	-8.03%	-7.55%	-8.94%	-7.14%	-6.84%	-7.49%	-7.58%
$\sqrt{12\text{GCLPM}}$	7.40%	4.90%	5.68%	5.01%	6.11%	4.93%	4.92%	4.95%	5.39%
CRRA $\times 10$	-2.96	-0.62	-1.14	-0.72	-1.35	-0.62	-0.59	-0.66	-0.83
Turnover	2.69%	3.63%	11.65%	6.49%	4.58%	3.84%	5.92%	4.35%	3.08%

Notes. See notes of Table 4. The only difference is that we do not adjust the performance to transaction costs.

Table III Out-of-sample empirical performance for the three daily return datasets using sample estimates of moments

(a) *10IND* dataset			$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$		$\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$		$\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$	
	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$252 \times \mu_P$	12.60%	10.67%	12.76%	10.67%	12.34%	12.06%	12.49%	12.21%
$\sqrt{252} \times \sigma_P$	12.30%	16.08%	13.40%	13.59%	12.37%	12.50%	12.48%	12.70%
$\sqrt{252} \times \mu_P/\sigma_P$	1.02	0.66	0.95	0.79	1.00	0.96	1.00	0.96
CFVaR	-6.48%	-5.95%	-6.50%	-5.62%	-6.65%	-6.31%	-6.53%	-6.12%
$\sqrt{252\text{GCLPM}}$	9.12%	12.05%	10.08%	10.35%	9.14%	9.22%	9.25%	9.59%
CRRA$\times$1000	-7.15	-13.32	-9.02	-8.49	-7.39	-7.31	-7.47	-7.54
Turnover	1.37%	5.05%	2.70%	3.52%	1.45%	1.80%	1.61%	2.00%

(b) *25BTM* dataset			$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$		$\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$		$\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$	
	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$252\mu_P$	16.67%	20.54%	18.62%	16.35%	17.11%	14.69%	19.01%	18.93%
$\sqrt{252}\sigma_P$	10.65%	18.33%	11.40%	11.57%	10.66%	11.49%	10.89%	11.01%
$\sqrt{252}\mu_P/\sigma_P$	1.56	1.12	1.63	1.41	1.60	1.28	1.75	1.72
CFVaR	-5.38%	-4.17%	-4.98%	-4.86%	-5.29%	-4.99%	-4.63%	-4.79%
$\sqrt{252\text{GCLPM}}$	8.11%	13.42%	8.28%	8.48%	8.05%	8.90%	8.06%	8.06%
CRRA$\times$1000	-4.20	-12.69	-4.54	-4.75	-4.11	-5.04	-3.84	-4.02
Turnover	4.84%	21.78%	7.38%	8.58%	5.03%	6.75%	6.22%	6.78%

(c) *25Prof* dataset			$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$		$\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$		$\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$	
	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$252\mu_P$	13.19%	17.18%	14.39%	12.07%	13.45%	11.91%	14.87%	14.68%
$\sqrt{252}\sigma_P$	11.43%	19.13%	12.32%	12.50%	11.44%	12.00%	11.76%	11.87%
$\sqrt{252}\mu_P/\sigma_P$	1.15	0.90	1.17	0.97	1.18	0.99	1.26	1.24
CFVaR	-5.91%	-4.81%	-5.77%	-5.76%	-5.83%	-5.46%	-5.46%	-5.54%
$\sqrt{252\text{GCLPM}}$	8.39%	13.43%	8.82%	8.34%	8.34%	8.91%	8.46%	8.26%
CRRA$\times$1000	-5.45	-15.96	-6.38	-6.60	-5.38	-5.93	-5.42	-5.55
Turnover	4.96%	20.81%	7.66%	9.26%	5.16%	7.05%	6.28%	6.96%

Notes. See notes of Table 3. The only difference is that we use sample estimates of portfolio return moments instead of shrinkage estimates; see Section 3.1.

Table IV Out-of-sample empirical performance for the two monthly return datasets using sample estimates of moments

(a) *MSCI* dataset			$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$		$\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$		$\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$	
	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$12 \times \mu_P$	-2.43%	-7.55%	-2.82%	-2.84%	-2.44%	-3.01%	-2.54%	-3.61%
$\sqrt{12} \times \sigma_P$	13.40%	15.50%	14.15%	14.02%	13.46%	13.66%	13.45%	13.67%
$\sqrt{12} \times \mu_P/\sigma_P$	-0.18	-0.49	-0.20	-0.20	-0.18	-0.22	-0.19	-0.26
CFVaR	-10.73%	-12.47%	-11.89%	-13.09%	-10.84%	-11.19%	-10.97%	-10.99%
$\sqrt{12\text{GCLPM}}$	10.12%	11.49%	10.78%	10.96%	10.19%	10.38%	10.19%	10.32%
CRRA $\times 10$	-4.58	-7.43	-6.20	-7.60	-4.71	-5.12	-4.84	-4.88
Turnover	123.0%	338.9%	183.8%	139.1%	128.3%	140.6%	134.7%	154.7%
(b) *S&P* dataset			$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$		$\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$		$\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$	
	$\boldsymbol{w}_{GMV}$	$\boldsymbol{w}_{KWZ}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$12 \times \mu_P$	6.26%	12.54%	5.63%	7.18%	5.72%	5.36%	7.01%	7.26%
$\sqrt{12} \times \sigma_P$	5.19%	14.40%	8.02%	5.69%	5.49%	6.27%	5.38%	5.49%
$\sqrt{12} \times \mu_P/\sigma_P$	1.21	0.87	0.70	1.26	1.04	0.86	1.30	1.32
CFVaR	-6.64%	-18.01%	-8.81%	-6.23%	-6.43%	-8.21%	-6.35%	-6.31%
$\sqrt{12\text{GCLPM}}$	4.32%	10.80%	6.47%	4.60%	4.51%	5.26%	4.46%	4.54%
CRRA $\times 10$	-0.41	-15.99	-1.52	-0.45	-0.44	-0.84	-0.42	-0.44
Turnover	6.33%	49.10%	32.85%	11.13%	12.76%	23.48%	8.61%	7.31%

Notes. See notes of Table 4. The only difference is that we use sample estimates of portfolio return moments instead of shrinkage estimates; see Section 3.1.