Reconciling mean-variance portfolio theory with non-Gaussian returns

Nathan Lassance

UCLouvain (BE), Louvain Finance, Louvain Institute of Data Analysis and Modeling

Voie du Roman Pays 34, 1348 Louvain-la-Neuve, Belgium

nathan.lassance@uclouvain.be

This version: August 3, 2020

Abstract

Mean-variance portfolio theory remains frequently used as investment rationale because of its simplicity, its closed-form solution, and the availability of many 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. Via numerical simulations and an empirical analysis, we find that, for three higher-moment objective functions and adjusting for transaction costs, the resulting portfolios outperform the minimum-variance and fully optimized portfolios out of sample both in terms of Sharpe ratio and higher moments, thus striking a favorable tradeoff between specification and estimation error.

Keywords

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

1. Introduction

Mean-variance portfolio theory, introduced by Markowitz (1952) and consolidated by many others ever since, 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 of Sharpe 1964). Second, it has simple and efficient mathematical and numerical solutions as its objective function is convex. Third, a vast literature exists on powerful 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, Tu and Zhou 2011).

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 expected quadratic utility), it is not obvious for the investor which alternative to select as 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, Jondeau and Rockinger 2006, Martellini and Ziemann 2010, Dahlquist et al. 2017), moment constraints (Athayde and Flores 2004, Briec et al. 2007), or downside-risk criteria such as Value-at-Risk, expected shortfall, lower-partial-moments and drawdown measures (Price et al. 1982, Basak and Shapiro 2001, El Ghaoui et al. 2003, Alexander and Baptista 2004, Leòn and Moreno 2017, Lwin et al. 2017, Van Hemert et al. 2020). Second, such alternative objective functions are mathematically and numerically more complex, particularly in high dimension, as they require for example the identification of a mean-variance-skewness(-kurtosis) efficient surface or the solution to a non-convex objective function. Third, including higher moments exacerbates estimation risk and, although a recent literature proposes robust estimation schemes for higher-moment portfolios (Harvey et al. 2010, Martellini and Ziemann 2010, Boudt et al. 2018, Lassance and Vrins 2019, Khashanah et al. 2020), the question remains whether one can really improve upon MVE portfolios out of sample. As Brandt et al. (2009, p.3418) explain, "[...] extending the traditional approach beyond first and second moments, when the investor's utility function is not quadratic, is practically impossible because it requires modeling not only the conditional skewness and kurtosis of each stock but also the numerous high-order cross-moments."

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 one 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 (1983, 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 expected quadratic utility for some choice of risk-tolerance coefficient, they are not expected to all behave similarly in terms of higher moments when asset returns are not Gaussian. We do so 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 appeal of Markowitz mean-variance framework, while improving higher moments and keeping an efficient mean-variance tradeoff.

The paper is structured as follows. In Section 2, we detail the theoretical framework surrounding the proposed approach, and illustrate the method with two examples.

In Section 3, we rely on Monte Carlo simulations and the Value-at-Risk objective function to compare the out-of-sample performance of the global minimum-variance (GMV) portfolio, our portfolio strategy, and the fully optimized portfolio. We find that our strategy always outperforms the fully optimized portfolio, except when the number of assets is small and the sample size is large (10 assets and 10 years of daily returns). It also outperforms the GMV portfolio expect when the sample size is small in which case they perform similarly. Thus, it has lower estimation error than the fully optimized portfolio, and lower specification error (bias) than the GMV portfolio.

In Section 4, we conduct an empirical analysis using daily returns on two benchmark datasets of 25 assets from Kenneth French's library. Looking at out-of-sample performance adjusted for transaction costs as in DeMiguel et al. (2009a), we obtain compelling evidence that our proposed strategy is the best portfolio on the MVE frontier compared to the GMV and maximum-Sharpe-ratio (MSR) portfolios. Moreover, it outperforms the fully optimized portfolio in terms of Sharpe ratio and higher moments, while systematically reducing turnover. These results are in line with Simaan (2014) who, for the log, exponential and power utilities, find that there is an out-of-sample opportunity cost of relying on the maximum-expected-utility portfolio rather than a MVE portfolio. They conclude 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." Contrary to Simaan (2014), the objective functions we consider depend directly on the portfolio-return moments, and thus, the fully optimized portfolios make use of the covariance matrix (and higher-comoment matrices). Therefore, our results indicate that the out-of-sample superiority of mean-variance goes beyond the specific setting of Simaan (2014) based on sample estimates of maximum-expected-utility portfolios.

It is worth noting how our work differs from Simaan (2014). First, whereas Simaan restricts to expected utility as 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 of "locally quadratic", so that Simaan's results are not very surprising. In contrast, we use three distinct higher-moment objectives: the Value-at-Risk, the lower-partial-moment (Price et al. 1982, Sortino and van der Meer 1991, Leòn and Moreno 2017) and the constant-relative-risk-aversion utility (Jondeau and Rockinger 2006, Guidolin and Timmermann 2008, Martellini and Ziemann 2010). Second, Simaan's results are based on monthly returns, and the corresponding low 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 use daily returns to make the comparison more fair, and consider both sample and robust shrinkage estimates. Third, Simaan only exemplifies his method on the Dow Jones index, whereas we consider simulated data in Section 3 and two real datasets in Section 4. Finally, Simaan does not consider the GMV portfolio as benchmark even though it is often the best MVE portfolio out of sample (Jagannathan and Ma 2003). We include it in our results and find indeed that it performs well compared to the fully optimized portfolios. Still, our alternative MVE portfolio strategy does better than GMV both in terms of Sharpe ratio and higher moments even though we adjust for transaction costs.

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

2. Theoretical framework

We place ourselves in 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 $\boldsymbol{\Sigma}$, $\boldsymbol{w} \in \mathbb{R}^N$ is a vector of portfolio weights, and $P := \boldsymbol{w}'\boldsymbol{X}$ is the associated portfolio return. 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 $\kappa_P := m_{4,P}/\sigma_P^4 - 3$. Finally, we denote the set of unconstrained fully invested portfolios by

$$\mathcal{W} := \{\boldsymbol{w} \in \mathbb{R}^N \mid \boldsymbol{w}'\mathbf{1} = 1\}. \tag{1}$$

2.1. Mean-variance-efficient portfolios

Given a mean-variance risk-tolerance coefficient $\lambda \in \mathbb{R}^+$, the unconstrained mean-variance-efficient (MVE) portfolio is defined as the expected-quadratic-utility maximizer:

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

Defining

$$\boldsymbol{Q} := \boldsymbol{\Sigma}^{-1} - \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}'\boldsymbol{\Sigma}^{-1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}},$$

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

$$\boldsymbol{w}(\lambda) = \boldsymbol{w}_{MV} + \lambda\boldsymbol{Q}\boldsymbol{\mu}, \tag{2}$$

where

$$\boldsymbol{w}_{MV} := \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}} = \boldsymbol{w}(0) \tag{3}$$

is the global minimum-variance (GMV) portfolio. Another notable MVE portfolio is the maximum-Sharpe-ratio (MSR) portfolio,

$$\boldsymbol{w}_{MSR} := \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} = \boldsymbol{w}(\lambda_{MSR}), \tag{4}$$

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

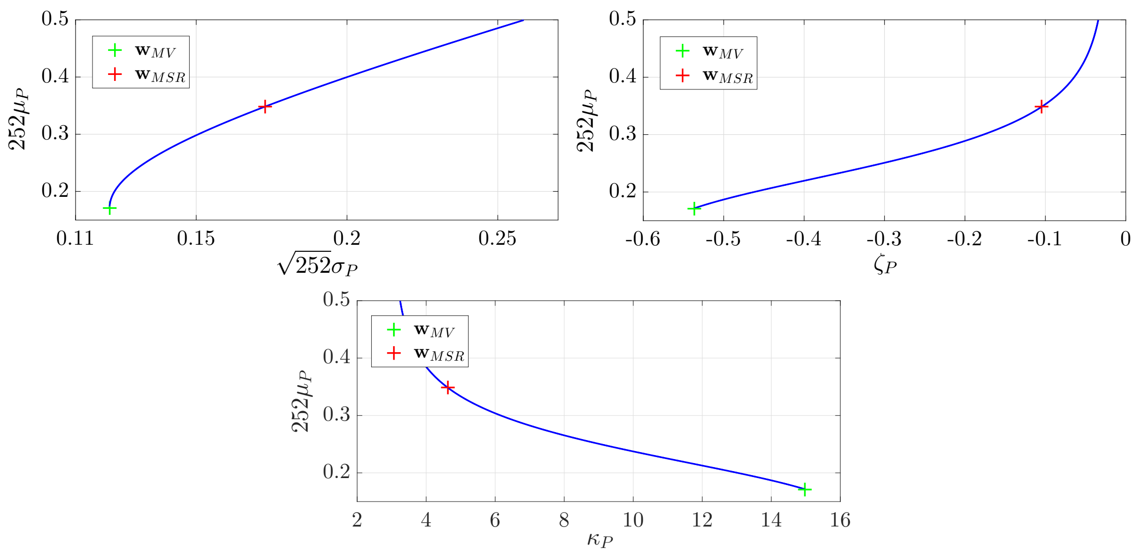
At this stage, it is important to realize that there are an infinite number of MVE portfolios depending on the coefficient λ. For example, $\lambda = 0$ gives the GMV portfolio, $\lambda = \infty$ gives the maximum-mean-return portfolio, and $\lambda = \lambda_{MSR}$ gives the MSR portfolio.

2.2. Optimal MVE portfolio with non-Gaussian returns

While all these portfolios are mean-variance efficient, there is however no reason to believe that they all behave similarly in terms of higher moments. This is illustrated in the next example.

Example 1. We collect value-weighted daily returns on the 25 size-and-book-to-market portfolios dataset from Kenneth French's library, for the time period January 1970 to May 2020. We

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



Notes. The figure is based on value-weighted daily returns for 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 minimum-variance portfolio $\boldsymbol{w}_{MV}$ and the red cross is the maximum-Sharpe-ratio portfolio $\boldsymbol{w}_{MSR}$.

consider all MVE portfolios with annualized mean return between $\mu_{MV} = 0.1715$ and 0.4, corresponding to $\lambda \in [0, 0.1593]$. 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 get very different moment profiles. For example, the minimum-variance portfolio $\boldsymbol{w}_{MV}$, 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 maximum-Sharpe ratio $\boldsymbol{w}_{MSR}$, 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 substantial in-sample higher-moment improvements by going from the GMV to the MSR portfolio, even though they are both MVE.

Because investors care about higher moments, it is theoretically no longer fully optimal to rely on the expected quadratic utility as objective function when asset returns are non-Gaussian. We suppose that the investor has chosen an alternative objective function $\mathcal{C}(\boldsymbol{w})$. 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 portfolio maximizing $\mathcal{C}(\boldsymbol{w})$ as

$$\boldsymbol{w}_{\mathcal{C}} := \underset{\boldsymbol{w} \in \mathcal{W}}{\operatorname{argmax}} \ \mathcal{C}(\boldsymbol{w}). \tag{5}$$

Our proposal is, instead of relying on $\boldsymbol{w}_{\mathcal{C}}$ that may be difficult to find numerically and may be highly prone to estimation risk, to find the MVE portfolio that maximizes $\mathcal{C}(\boldsymbol{w})$. In doing so, one can get 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 portfolios.

Mathematically, we select the portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}}) = \boldsymbol{w}_{MV} + \lambda_{\mathcal{C}}\boldsymbol{Q}\boldsymbol{\mu}$, where

$$\lambda_{\mathcal{C}} := \underset{\lambda \in \mathbb{R}^+}{\operatorname{argmax}} \ \mathcal{C}(\boldsymbol{w}(\lambda)). \tag{6}$$

We name this portfolio optimal MVE (OMVE) portfolio. Contrary to the identification of $\boldsymbol{w}_{\mathcal{C}}$, we are pretty much guaranteed to find the OMVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ as $\lambda_{\mathcal{C}}$ 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 is beneficial out of sample given how challenging it is to estimate higher moments in a robust manner. Sections 3 and 4 will provide evidence that $\boldsymbol{w}(\lambda_{\mathcal{C}})$ achieves a better out-of-sample tradeoff between specification error and estimation error than $\boldsymbol{w}_{MV}$ and $\boldsymbol{w}_{\mathcal{C}}$.

Example 2. Suppose for the sake of mathematical tractability that there are $N = 2$ assets, the investor's objective function is a variance-skewness expansion of the constant-relative-risk aversion (CRRA) expected utility (Martellini and Ziemann 2010 p.1487), and the asset returns are independent. In that case, the MVE portfolio is $\boldsymbol{w}(\lambda) = (w(\lambda), 1 - w(\lambda))'$ with

$$w(\lambda) = \frac{\mu_1 - \mu_2 + \lambda\sigma_2^2}{\lambda(\sigma_1^2 + \sigma_2^2)}, \tag{7}$$

and the objective function is given by

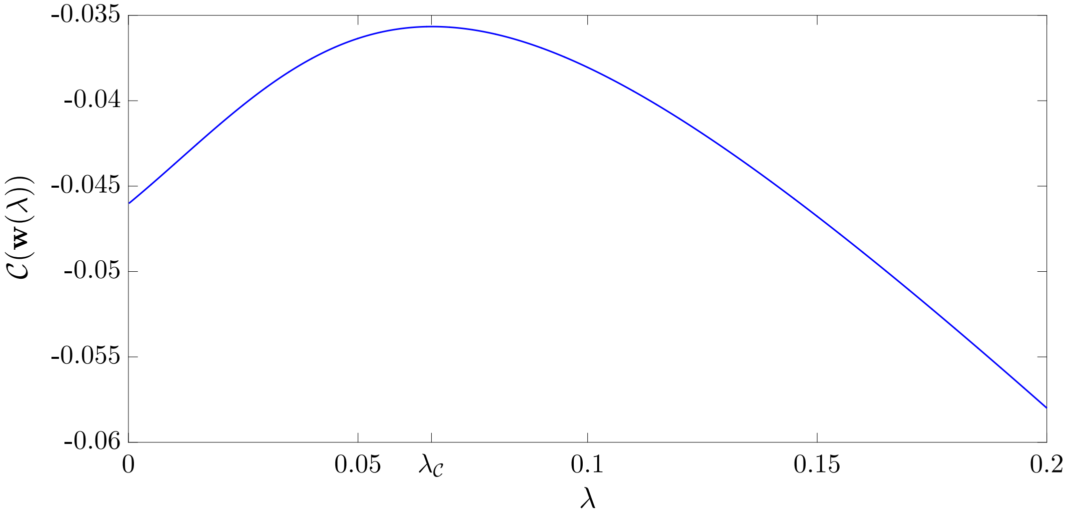
$$\begin{aligned} \mathcal{C}(\boldsymbol{w}(\lambda)) &= -\frac{\gamma}{2}\sigma_P^2 + \frac{\gamma(\gamma+1)}{6}m_{3,P} \\ &= -\frac{\gamma}{2}\left(w(\lambda)^2\sigma_1^2 + (1-w(\lambda))^2\sigma_2^2\right) + \frac{\gamma(\gamma+1)}{6}\left(w(\lambda)^3 m_{3,1} + (1-w(\lambda))^3 m_{3,2}\right). \end{aligned} \tag{8}$$

By differentiating $\mathcal{C}(\boldsymbol{w}(\lambda))$ with respect to λ and with some algebra we find that $\lambda_{\mathcal{C}}$ is the real root with largest $\mathcal{C}(\boldsymbol{w}(\lambda))$ of the polynomial $a\lambda^3 + b\lambda^2 + c\lambda + d$, where

$$\begin{aligned} &a = 2\sigma_1^2\sigma_2^2(\sigma_1^2+\sigma_2^2)(1-\sigma_1^2-\sigma_2^2), \quad b = (\gamma+1)(\mu_1-\mu_2)(\sigma_2^4 m_{3,1} - \sigma_1^4 m_{3,2}), \\ &c = 2(\mu_1-\mu_2)^2\left((\gamma+1)(\sigma_1^2 m_{3,2} + \sigma_2^2 m_{3,1}) - (\sigma_1^2+\sigma_2^2)^2\right), \quad d = (\gamma+1)(\mu_1-\mu_2)^3(m_{3,1} - m_{3,2}). \end{aligned} \tag{9}$$

For example, suppose $(\mu_1, \mu_2) = (0.10, 0.05)$, $(\sigma_1, \sigma_2) = (0.2, 0.3)$ and $(\zeta_1, \zeta_2) = (-2, 2)$. Clearly,

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



Notes. The figure is based on value-weighted daily returns on the 25 size-and-book-to-market portfolios dataset from January 1970 to May 2020. The figure depicts the Cornish-Fisher Value-at-Risk $\mathcal{C}(\boldsymbol{w})$ in (10), with $\alpha = 1\%$, of the MVE portfolio $\boldsymbol{w}(\lambda)$ with a risk-tolerance coefficient $\lambda \in [0, 0.2]$. The value of λ maximizing $\mathcal{C}(\boldsymbol{w}(\lambda))$ is $\lambda_{\mathcal{C}} = 0.066$, corresponding to $\mathcal{C}(\boldsymbol{w}(\lambda_{\mathcal{C}})) = -3.57\%$

from a mean-variance perspective, it is better to invest aggressively in the second asset and, indeed, for a low value of $\gamma = 10$ we have $\lambda_{\mathcal{C}} = 0.59$ and $w(\lambda_{\mathcal{C}}) = 1.35$. However, the first asset has a large negative skewness and, when γ increases to 500, we have $\lambda_{\mathcal{C}} = 6.45$ and $w(\lambda_{\mathcal{C}}) = 0.75$.

Example 1 (Continued)**.** Consider as higher-moment objective function $\mathcal{C}(\boldsymbol{w})$ the Cornish-Fisher Value-at-Risk (Cornish and Fisher 1938, Favre and Galeano 2002), which is a four-moment expansion of the Value-at-Risk given by

$$\mathcal{C}(\boldsymbol{w}) = \text{CFVaR} := \mu_P - \sigma_P \left(\frac{1}{36}(2z_\alpha^3 - 5z_\alpha)\zeta_P^2 - \frac{1}{6}(z_\alpha^2 - 1)\zeta_P - \frac{1}{24}(z_\alpha^3 - 3z_\alpha)\kappa_P - z_\alpha \right), \tag{10}$$

where z_α is the standard Gaussian quantile at confidence level α. For the data of Example 1, Figure 2 shows how $\mathcal{C}(\boldsymbol{w}(\lambda))$ evolves as a function of $\lambda \in [0, 0.2]$ 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.

3. Monte Carlo simulations

To assess the out-of-sample performance of our approach, we rely in this section on Monte Carlo simulations. We describe the methodology in Section 3.1 and we discuss the results in Section 3.2.

3.1. Methodology

We consider a non-Gaussian data-generation process and, for different values of the number of assets N and sample size T, we report the out-of-sample performance of four portfolio strategies averaged over 100 Monte Carlo simulations.

The four portfolio strategies are: the GMV portfolio $\boldsymbol{w}_{MV}$, the MSR portfolio $\boldsymbol{w}_{MSR}$, our proposed OMVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ using the CFVaR with $\alpha = 1\%$ in (10) as objective, and the fully optimized portfolio $\boldsymbol{w}_{\mathcal{C}}$ in (5).[1] The portfolio $\boldsymbol{w}_{\mathcal{C}}$ is the maximum of a non-concave objective function, which rules out the use of local optimizers, particularly for large N. We use Matlab global optimizer *GlobalSearch* to find $\boldsymbol{w}_{\mathcal{C}}$, and similarly for $\lambda_{\mathcal{C}}$. We restrict here to the CFVaR for brevity but, in the empirical analysis of Section 4, we consider two additional higher-moment objective functions. All portfolios are estimated via sample moment estimates only for brevity. In the empirical analysis of Section 4, we also consider robust shrinkage estimates.

We consider a non-Gaussian data-generating process similar to Section 4 in Lassance and Vrins (2020), 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 out-of-sample period. We fix τ to be 20 years of daily returns, whereas T will be varied in our experiments from 1 to 10 years. 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 independent factor model of the form

$$\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{Y} + \boldsymbol{\epsilon}, \tag{11}$$

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 mutually independent, 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. The loading matrix $\boldsymbol{A}$ is taken as the eigendecomposition of a randomly simulated covariance matrix. Contrary to Lassance and Vrins (2020), we do not take zero-mean asset returns as we include the extra term $\boldsymbol{\mu}$. The mean return of each asset, μ_i, $i = 1, \ldots, N$, 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 $N = 10, 30, 50$ in our experiments.

[1] Another common choice is $\alpha = 5\%$. We set $\alpha = 1\%$ because Cavenaile and Lejeune (2012) show that $\alpha > 4.16\%$ is inconsistent with investors' negative preferences for kurtosis.

3.2. Results

In Figure 3, we report the daily CFVaR achieved by the considered portfolios, averaged over the 100 simulations. Recall that the portfolios are estimated based on the returns from $t = 1$ to $t = T$. On the left plots, we report the in-sample CFVaR computed on the estimation window. On the right plots, we report the out-of-sample CFVaR computed on the returns from $t = T + 1$ to $t = T + \tau$. Because the MSR portfolio $\boldsymbol{w}_{MSR}$ always performs worst, we do not report its performance.

We make the following observations. First, in-sample performance is always better than out-of-sample performance: it provides an overoptimistic estimate because it ignores estimation risk. This is why the comparison of portfolio strategies should be conducted out of sample. Second, the in-sample ordering of the portfolios from best to worst is $\boldsymbol{w}_{\mathcal{C}}$, $\boldsymbol{w}(\lambda_{\mathcal{C}})$ and $\boldsymbol{w}_{MV}$ because specification error (bias) increases in that order as well. Third, our proposed OMVE strategy is the best option out of sample. The only instances where it does not perform best is for $N = 10$ and $T = 1$ year where $\boldsymbol{w}_{MV}$ does better because it has lower estimation risk, and for $N = 10$ and $T = 10$ years where $\boldsymbol{w}_{\mathcal{C}}$ does better because it has lower specification error. The outperformance compared to $\boldsymbol{w}_{\mathcal{C}}$ is particularly visible when the sample size is small, but it remains present 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 out-of-sample periods. This contrasts with real data for which the distribution of asset returns evolves over time; we will observe in Section 4 that the outperformance in that case is even more substantial.

In summary, there is no out-of-sample 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.

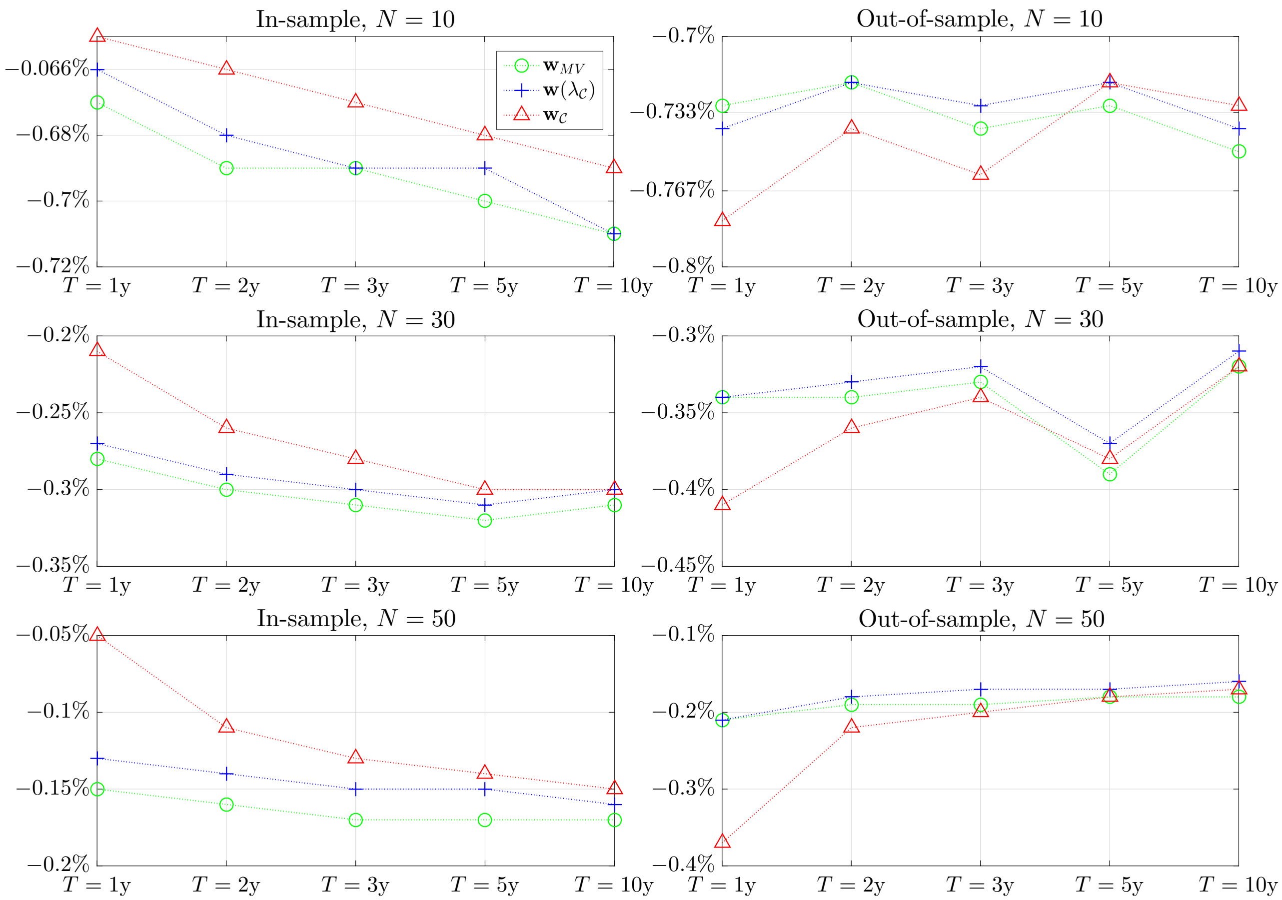
4. Empirical analysis

In this section, we test the out-of-sample performance of the proposed portfolio strategy on real data. We describe the methodology in Section 4.1, and we discuss the results in Section 4.2.

4.1. Methodology

Data. We use value-weighted daily returns of two benchmark equity datasets from Kenneth French's library: 25 size-and-book-to-market portfolios (*25BTM*) and 25 size-and-operating-profitability portfolios (*25Prof*). The sample period is from January 1970 to May 2020. We use daily returns

Figure 3 Out-of-sample performance via Monte Carlo simulations



Notes. The figure depicts, for the Cornish-Fisher Value-at-Risk (CFVaR) with $\alpha = 1\%$ in (10) as objective function $\mathcal{C}(\boldsymbol{w})$, the out-of-sample performance of the GMV portfolio $\boldsymbol{w}_{MV}$ (green circles), our proposed optimal MVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ (blue crosses), and the fully optimized portfolio $\boldsymbol{w}_{\mathcal{C}}$ (red triangles) for the Monte Carlo simulations of Section 3. The way the daily asset returns are randomly sampled is described in Section 3.1. The portfolios are estimated on a window of T returns reported on the x-axis, which goes from 1 to 10 years of daily returns. The y-axis of the left plots depict the in-sample CFVaR computed on the estimation window. The y-axis of the right plots depict the out-of-sample CFVaR computed on a different window of $\tau = 20$ years of daily returns. We consider a number of assets $N = 10, 30, 50$. The performance is averaged over 100 Monte Carlo simulations.

to reduce estimation error and give the optimal portfolios $\boldsymbol{w}_{\mathcal{C}}$ a fair chance. Indeed, Martellini and Ziemann (2010) find that, with monthly returns, higher-moment portfolios cannot outperform the GMV portfolio even when using robust moment estimates.

Portfolio strategies. We consider four portfolio strategies: the GMV portfolio $\boldsymbol{w}_{MV}$, the MSR portfolio $\boldsymbol{w}_{MSR}$, the OMVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$, and the fully optimized portfolio $\boldsymbol{w}_{\mathcal{C}}$. We use three objective functions $\mathcal{C}(\boldsymbol{w})$. $\lambda_{\mathcal{C}}$ and $\boldsymbol{w}_{\mathcal{C}}$ are found using Matlab *GlobalSearch* optimizer.

Objective functions. The first objective is the CFVaR in (10) with $\alpha = 1\%$. The second objective is the lower-partial-moment (LPM); see, for example, Price et al. (1982) and Sortino and van der

Meer (1991). In all generality, the LPM is defined as

$$\mathrm{LPM}(\tau, m) := \int_{-\infty}^{\tau} (\tau - x)^m f_P(x)\mathrm{d}x, \tag{12}$$

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 Appendix A that for our special case $m = 2$ and $\tau = \mu_P$, the expansion reduces to maximizing the objective function[2]

$$\mathcal{C}(\boldsymbol{w}) = -\mathrm{GCLPM} := -\sigma_P^2\big(1/2 - (18\pi)^{-1/2}\zeta_P\big). \tag{13}$$

The third objective is the four-moment Taylor-series expansion of the CRRA expected utility (Jondeau and Rockinger 2006, Guidolin and Timmermann 2008, Martellini and Ziemann 2010):

$$\mathcal{C}(\boldsymbol{w}) = \mathrm{CRRA} := \mu_P - \frac{\gamma}{2}\sigma_P^2 + \frac{\gamma(\gamma+1)}{6}m_{3,P} - \frac{\gamma(\gamma+1)(\gamma+2)}{24}m_{4,P}. \tag{14}$$

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 μ_P, just like Martellini and Ziemann (2010) who remove μ_P from (14).

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

Performance metrics and transaction costs. We report the annualized mean, annualized volatility, annualized Sharpe ratio, daily skewness, daily excess kurtosis, daily CFVaR, square root of annualized GCLPM, daily CRRA, and finally daily turnover. As in DeMiguel et al. (2009a, p.1930), the returns are computed net of transaction costs, using a proportional transaction cost of 30 basis points. Results without transaction costs are available in the supplementary materials.

Moment estimation. All portfolios considered depend on some or all of the first four portfolio-return moments. 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. (2018) for the coskewness and cokurtosis matrices.

[2] Note that the GCLPM in (13) becomes negative if $\zeta_P > \sqrt{9\pi/2} = 3.76$, but this case never occurred.

We use the R package *PerformanceAnalytics* to compute the estimators. For the portfolio mean return, we use the estimator of Jorion (1986). It is consistent with the iid target as 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 the GMV portfolio performs very similarly under both estimators because the shrinkage intensity is close to zero. For the OVME and fully optimized portfolios, we observe that, compared to sample estimators, the portfolios under shrinkage estimators have a lower turnover but most of the time a worse performance even adjusting for transaction costs. This is because, when using daily returns, introducing additional bias may hurt out-of-sample performance (Jagannathan and Ma 2003). Thus, we only report the results for sample estimators; results for shrinkage estimators are available in the supplementary materials.

4.2. Results

The out-of-sample results for the *25BTM* and *25Prof* datasets are reported in Table 1. The results indicate that the OMVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ is the best portfolio on the MVE frontier compared to the GMV and MSR portfolios and, moreover, outperforms the fully optimized 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 for the three objective functions $\mathcal{C}(\boldsymbol{w})$ have the best performance compared to all other portfolios. The improvement is particularly significant compared to the fully optimized portfolios $\boldsymbol{w}_{\mathcal{C}}$. Thus, by staying on the MVE frontier, we enjoy a strong mean-variance tradeoff even after adjusting for transaction costs. In addition to a large Sharpe ratio, the OMVE portfolios perform overall best with respect to the three higher-moment criteria $\mathcal{C}(\boldsymbol{w})$ considered. For $\mathcal{C}(\boldsymbol{w}) = \text{CFVaR}$, $\boldsymbol{w}(\lambda_{\mathcal{C}})$ performs worse than $\boldsymbol{w}_{\mathcal{C}}$ for *25BTM*, and performs equally well for *25Prof*. For $\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$, $\boldsymbol{w}(\lambda_{\mathcal{C}})$ clearly outperforms $\boldsymbol{w}_{\mathcal{C}}$ for both datasets, and does best out of all portfolios. For $\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$, $\boldsymbol{w}(\lambda_{\mathcal{C}})$ outperforms $\boldsymbol{w}_{\mathcal{C}}$ and does best out of all portfolios for *25BTM*, and second best for *25Prof*.

Concerning turnover, the OMVE portfolio $\boldsymbol{w}(\lambda_{\mathcal{C}})$ reduces turnover compared to $\boldsymbol{w}_{\mathcal{C}}$ in all cases because all higher-moment estimation risk is transferred to a single parameter $\lambda_{\mathcal{C}}$. It also has a higher turnover than the GMV portfolio, which makes the outperformance even more material given that we adjust the returns to transaction costs.

Another interesting observation is that, even though we use daily returns for which estimation risk is reduced, the GMV portfolio may already be an appealing alternative to the fully optimized portfolio $\boldsymbol{w}_{\mathcal{C}}$. Indeed, it has a larger Sharpe ratio for $\mathcal{C}(\boldsymbol{w}) = -\text{GCLPM}$ and $\mathcal{C}(\boldsymbol{w}) = \text{CRRA}$, and

Table 1 Out-of-sample empirical performance for the *25BTM* and *25Prof* datasets

(a) *25BTM* dataset			$\mathcal{C}(\boldsymbol{w})$ = CFVaR		$\mathcal{C}(\boldsymbol{w}) = -$GCLPM		$\mathcal{C}(\boldsymbol{w})$ = CRRA	
	$\boldsymbol{w}_{MV}$	$\boldsymbol{w}_{MSR}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$252\mu_P$	16.67%	8.43%	18.62%	16.35%	17.11%	14.69%	**19.01%**	18.93%
$\sqrt{252}\sigma_P$	**10.65%**	27.68%	11.40%	11.57%	10.66%	11.49%	10.89%	11.01%
$\sqrt{252}\mu_P/\sigma_P$	1.56	0.30	1.63	1.41	1.60	1.28	**1.75**	1.72
ζ_P	-0.59	-1.87	**-0.21**	-0.29	-0.53	-0.75	-0.36	-0.27
κ_P	23.48	76.66	19.55	**18.17**	22.94	18.43	18.48	19.31
CFVaR	-5.38%	-35.38%	-4.98%	-4.86%	-5.29%	-4.99%	**-4.63%**	-4.79%
$\sqrt{252\text{GCLPM}}$	8.11%	23.95%	8.28%	8.48%	**8.05%**	8.90%	8.06%	8.06%
CRRA×1000	-4.20	-347.67	-4.54	-4.75	-4.11	-5.04	**-3.84**	-4.02
Turnover	**4.84%**	32.36%	7.38%	8.58%	5.03%	6.75%	6.22%	6.78%

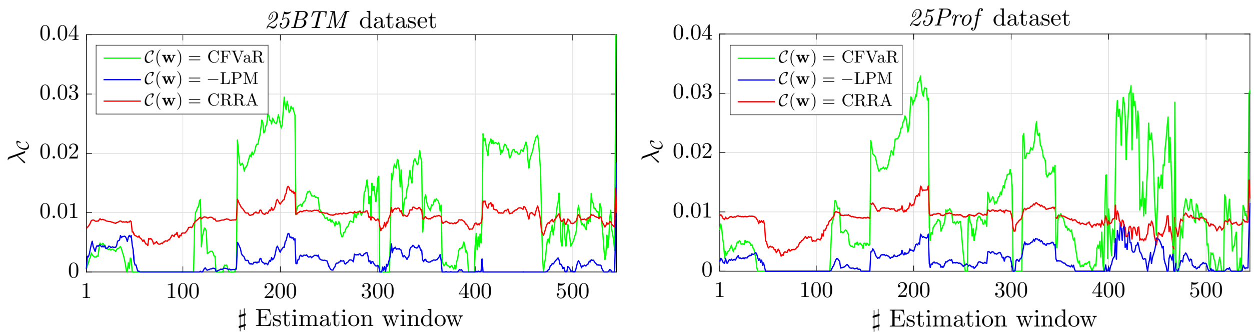
(b) *25Prof* dataset			$\mathcal{C}(\boldsymbol{w})$ = CFVaR		$\mathcal{C}(\boldsymbol{w}) = -$GCLPM		$\mathcal{C}(\boldsymbol{w})$ = CRRA	
	$\boldsymbol{w}_{MV}$	$\boldsymbol{w}_{MSR}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$	$\boldsymbol{w}(\lambda_{\mathcal{C}})$	$\boldsymbol{w}_{\mathcal{C}}$
$252\mu_P$	13.19%	-4.77%	14.39%	12.07%	13.45%	11.91%	**14.87%**	14.68%
$\sqrt{252}\sigma_P$	**11.43%**	26.79%	12.32%	12.50%	11.44%	12.00%	11.76%	11.87%
$\sqrt{252}\mu_P/\sigma_P$	1.15	-0.18	1.17	0.97	1.18	0.99	**1.26**	1.24
ζ_P	-0.29	-3.00	-0.09	**0.42**	-0.24	-0.38	-0.13	0.12
κ_P	24.69	52.67	21.88	23.19	24.32	**20.24**	21.53	22.48
CFVaR	-5.91%	-22.74%	-5.77%	-5.76%	-5.83%	**-5.46%**	**-5.46%**	-5.54%
$\sqrt{252\text{GCLPM}}$	8.39%	25.40%	8.82%	**8.34%**	**8.34%**	8.91%	8.46%	8.26%
CRRA×1000	-5.45	-232.57	-6.38	-6.60	**-5.38**	-5.93	-5.42	-5.55
Turnover	**4.96%**	39.30%	7.66%	9.26%	5.16%	7.05%	6.28%	6.96%

Notes. The table reports, using daily returns on 25 size-and-book-to-market and size-and-operating-profitability portfolios datasets from January 1970 to May 2020, the out-of-sample performance of the GMV portfolio $\boldsymbol{w}_{MV}$, the MSR portfolio $\boldsymbol{w}_{MSR}$, the OMVE portfolios $\boldsymbol{w}(\lambda_{\mathcal{C}})$ and the fully optimized portfolios $\boldsymbol{w}_{\mathcal{C}}$. We use three higher-moment objective functions $\mathcal{C}(\boldsymbol{w})$: the Cornish-Fisher Value-at-Risk (CFVaR) in (10) with $\alpha = 1\%$, the Gram-Charlier lower-partial-moment (GCLPM) in (13), and the four-moment CRRA expected utility in (14) with $\gamma = 100$. The out-of-sample performance is computed 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. More details on the methodology are provided in Section 4.1. The bold figures indicate the best portfolio for each performance metric.

achieves a better higher-moment objective $\mathcal{C}(\boldsymbol{w})$ compared to the corresponding portfolio $\boldsymbol{w}_{\mathcal{C}}$ in three out of six cases. Along the MVE frontier, the investor should however pick the OMVE rather than the GMV portfolio as evidenced above.

In Figure 4, we depict the time evolution of the optimal $\lambda_{\mathcal{C}}$ for the three objective functions $\mathcal{C}(\boldsymbol{w})$. The figure shows that, whereas the magnitude of $\lambda_{\mathcal{C}}$ depends on the objective considered, it remains quite small. In particular, $\lambda_{\mathcal{C}}$ is small compared to λ_{MSR} in (4) that equals 0.1713 and 0.1832 on average for the *25BTM* and *25Prof* datasets, respectively. This means that the OMVE portfolios $\boldsymbol{w}(\lambda_{\mathcal{C}})$ are not very sensitive to estimation risk in asset mean returns $\boldsymbol{\mu}$, and thus, work well out of sample as observed above. The variability in $\lambda_{\mathcal{C}}$ in Figure 4 also explains why the OMVE

Figure 4 Time evolution of optimal risk-tolerance coefficient $\lambda_{\mathcal{C}}$



Notes. The figure draws on the empirical analysis of Section 4 using daily returns on the 25 size-and-book-to-market and size-and-operating-profitability portfolios datasets from January 1970 to May 2020. It depicts the time evolution, across all five-year estimation windows that are rolled over time by one month until April 2020, of the optimal risk-tolerance coefficient $\lambda_{\mathcal{C}}$ in (6). We use three higher-moment objective functions $\mathcal{C}(\boldsymbol{w})$: the Cornish-Fisher Value-at-Risk (CFVaR) in (10) with $\alpha = 1\%$, the Gram-Charlier lower-partial-moment (GCLPM) in (13) and the four-moment CRRA expected utility in (14) with $\gamma = 100$. More details on the methodology are provided in Section 4.1.

portfolios have a larger turnover than 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 MVE portfolio is chosen. From an out-of-sample perspective, we find no benefit in moving away from the MVE frontier: by adjusting the quadratic-utility risk-tolerance coefficient to improve higher moments, we obtain a portfolio strategy that strikes a better tradeoff between specification error and estimation error than the fully optimized portfolio and the GMV and MSR portfolios. This reconciles mean-variance portfolio theory with non-Gaussian returns and is an important implication for investors who may be attracted by the clarity and closed-form solution of mean-variance portfolio theory, but who may fear that the obtained portfolio does not perform well in terms of higher-moment risk.

Acknowledgements

This work was supported by the Fonds de la Recherche Scientifique (F.R.S.-FNRS) under grant ASP [grant number FC 17775].

A. Appendix: Gram-Charlier lower-partial-moment

In this appendix, we prove the formula for the GCLPM in (13). For a general τ, Leòn and Moreno (2017) show that the four-moment Gram-Charlier expansion of $\mathrm{LPM}(\tau, m=2)$ in (12) is given by

$$\mathrm{LPM}(\tau,2) \approx \mathrm{LPM}_\phi(\tau,2) + \frac{\zeta_P}{\sqrt{6}}\theta_{2,2} + \frac{\kappa_P}{\sqrt{24}}\theta_{3,2}, \tag{A.1}$$

where

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

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.1)–(A.2) reduces to

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

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 (13).

References

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.

Basak, S., & Shapiro, A. (2001). Value-at-risk-based risk management: Optimal policies and asset prices. *Review of Financial Studies*, 14(2), 371–405.

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

Boudt, K., Cornilly, D., & Verdonck, T. (2018). A coskewness shrinkage approach for estimating the skewness of linear combinations of random variables. *Journal of Financial Econometrics*, 22, 1–23.

Brandt, M., Santa-Clara, P., & Valkanov, R. (2009). Parametric portfolio policies: Exploiting characteristics in the cross section of equity returns. *Review of Financial Studies*, 22(9), 3411–3447.

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.

Cavenaile, L., & Lejeune, T. (2012). A note on the use of modified value-at-risk. *Journal of Alternative Investments*, 14(4), 79–83.

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.

Cornish, E., & Fisher, R. (1938). Moments and cumulants in the specification of distributions. *Review of the International Statistical Institute*, 5(4), 307–320.

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.

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.

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

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

Gormsen, N., & Jensen, C. (2019). Higher-moment risk. Available on SSRN: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3069617`.

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.

Jondeau, E., & Rockinger, M. (2006). Optimal portfolio allocation under higher moments. *European Financial Management*, 12(1), 29–55.

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

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). A parsimonious approach for higher-order moments in portfolio selection. Available on SSRN: `papers.ssrn.com/sol3/papers.cfm?abstract_id=3523379`.

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, 2—2020.

Lassance, N., & Vrins, F. (2020). Portfolio selection using parsimonious estimation of comoment tensors. In revision. Available on SSRN: `papers.ssrn.com/sol3/papers.cfm?abstract_id=3455400`.

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, 372–381.

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

Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3), 425–442.

Simaan, Y. (1993). 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.

Tu, J., & Zhou, G. (2011). Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics*, 99(1), 204–215.

Van Hemert, O., Ganz, M., Harvey, C., Rattray, S., Martin, E., & Yawitch, D. (2020). Drawdowns. *Journal of Portfolio Management*, 46(7).