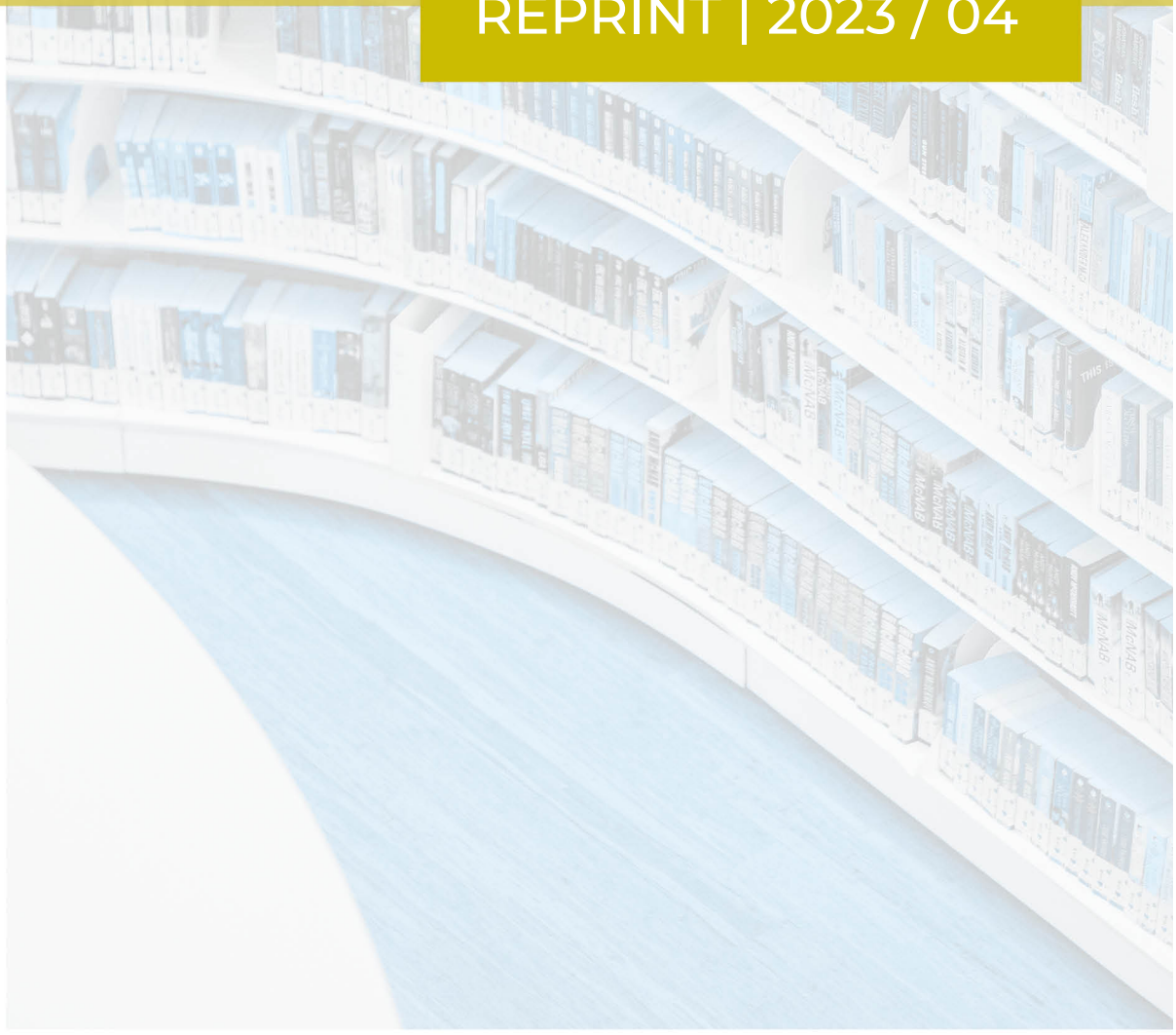
PORTFOLIO SELECTION: A TARGET-DISTRIBUTION APPROACH

Nathan Lassance, Frédéric Vrins

REPRINT | 2023 / 04



UCLouvain

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

LFIN

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

Tel (32 10) 47 43 04

Email: lidam-library@uclouvain.be

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

Portfolio Selection: A Target-Distribution Approach*

Nathan Lassance Frédéric Vrins

Forthcoming in *European Journal of Operational Research* (2023)

Abstract

We introduce a novel framework for the portfolio selection problem in which investors aim to target a return distribution, and the optimal portfolio has a return distribution as close as possible to the targeted one. The proposed framework can be applied to a variety of investment objectives. In this paper, we focus on improving the higher moments of mean-variance-efficient portfolios by designing the target so that its first two moments match those of the chosen efficient portfolio but has more desirable higher moments. We show theoretically that the optimal portfolio is in general different from the mean-variance portfolio, but remains mean-variance efficient when asset returns are Gaussian. Otherwise, it can move away from the efficient frontier to better match the higher moments of the target distribution. An extensive empirical analysis using three characteristic-sorted datasets and a dataset of 100 individual stocks indicates that the proposed framework delivers a satisfying compromise between mean-variance efficiency and improved higher moments.

Keywords: portfolio optimization, higher moments, downside risk, Kullback-Leibler divergence
JEL Classification: G11, G12

*Lassance, LFIN/LIDAM, UCLouvain, corresponding author, email: nathan.lassance@uclouvain.be; Vrins, LFIN/LIDAM, UCLouvain, HEC Montréal department of decision sciences, email: frederic.vrins@uclouvain.be. We gratefully acknowledge Kris Boudt, Victor DeMiguel, Roberto Renò, Guofu Zhou, and conference participants at 9th General AMaMeF Conference, 2019 Paris Financial Management Conference, and 13th Actuarial and Financial Mathematics Conference for comments and discussions. This work is supported by the Fonds de la Recherche Scientifique (F.R.S.-FNRS) under Grant Number T.0221.22 (F. Vrins) and J.0115.22 (N. Lassance) and the Belgian Federal Science Policy Office under Grant Number ARC 18-23/089 (F. Vrins).

1 Introduction

The portfolio selection problem asks how to best allocate wealth to financial assets. It is generally addressed in three steps. First, model the investor's preferences by choosing an appropriate objective function. Second, invest in the assets of a given investment universe to maximize the objective, i.e., find the optimal portfolio weights. This second step requires the investor to estimate the objective function from a sample of data, and several methods exist for that purpose.[1] Finally, the relevance of the optimal portfolio is evaluated in a third step, via an out-of-sample performance analysis.

In this paper, we address the first step of the portfolio selection process. To carry this out, it is important to recognize that, unlike assumed by Markowitz (1952), investors care for other distributional characteristics than just mean and variance; see Ang et al. (2006), Kadan and Liu (2014), and Kelly and Jiang (2014). However, there are dozens of performance/risk criteria in the portfolio literature that go beyond mean and variance, such as higher moments and quantiles; see Cogneau and Hubner (2014) for an extensive review. These criteria reflect different desired aspects of the portfolio-return distribution, and it is not trivial how one should trade off between several criteria of interest. In the case of moment-based approaches for example, which moments should be discarded, and how should the selected ones be aggregated into a single objective function? One standard approach is to maximize a higher-moment expansion of expected utility but, as we review at the end of this section, expected utility is inadequately insensitive to higher moments.

Therefore, in this paper, we propose to directly specify a desired *target-return distribution*. This offers a natural way for investors to balance between the different desired features of the portfolio-return distribution by choosing a target distribution reflecting these features, and then finding the portfolio whose return distribution is as close as possible to the targeted one. The Kullback-Leibler (KL) divergence (Kullback and Leibler, 1951) is the standard way to measure how well the two distributions match one another, and determines the optimal tradeoff between the different desired aspects of the portfolio-return distribution.

The proposed framework can be applied to different investment objectives depending on the choice of target distribution. In this paper, we mainly focus on one application: improving the higher moments of mean-variance-efficient (MVE) portfolios. To achieve this, we choose a parametric family for the target distribution and determine its parameters such that (i) the first two moments coincide with those of a *reference portfolio* chosen on the MVE frontier, but (ii) the higher moments

[1] Some recent approaches in the literature on portfolio selection with parameter uncertainty improve the estimation of eigenvalues (Ledoit and Wolf, 2017, Zhao et al., 2019), combine portfolios (Kan and Zhou, 2007, Kan et al., 2022), constrain portfolio weights (Levy and Levy, 2014), use sparse and robust estimation (DeMiguel and Nogales 2009, Ao et al., 2019), and reduce dimension via factor models (De Nard et al. 2019, Lassance and Vrins, 2021a).

of the target distribution improve upon those of the reference portfolio. The optimal portfolio is the attainable strategy whose return distribution best matches the target distribution, and can be regarded as the MVE reference portfolio *shrunk* toward another portfolio that better reflects the higher moments embedded in the chosen target. We show that our portfolio significantly improves tail risk characteristics relative to MVE portfolios and higher-moment portfolios built on expected utility, which is a salient benefit of the proposed approach.

The proposed methodology can be seen as a compromise between two common approaches to portfolio optimization: expected utility and benchmark tracking. Expected utility is an absolute and distribution-based approach, which delivers the portfolio whose return distribution is most appealing according to several moments, e.g., mean and variance. Benchmark tracking is a relative and path-by-path approach, which evaluates performance via the tracking error to a benchmark asset at a given time horizon, irrespective of the benchmark's absolute characteristics. Our target-distribution methodology sits between the two methods because it is a relative *and* distribution-based approach: performance is evaluated by comparing the portfolio-return distribution to a target distribution. The latter needs not be that of an identified asset, and may even be unattainable.

Our contribution is threefold. First, we introduce a novel framework in which the investor's preferences are captured by a target-return distribution. The optimal portfolio is then obtained by minimizing the extended Kullback-Leibler (EKL) divergence between the portfolio-return and target-return density functions, and is called the minimum-EKL (MEKL) portfolio. The EKL divergence is a novel extension to the standard KL divergence that we introduce to handle target densities with bounded support set. The choice of target distribution is specific to the investor but, to operationalize the approach, we study two classes of target distribution that are parsimonious, analytically tractable, and allow for a wide range of higher moments. On the one hand, the shifted-log-normal (SLN) distribution (Galton, 1879, Johnson et al., 1994), which extends the Gaussian via an additional parameter controlling skewness. On the other hand, the generalized-normal (GN) distribution (Nadarajah, 2005), which extends the Gaussian via an additional parameter controlling kurtosis. We match the mean and variance of the target to an MVE reference portfolio, while the additional parameter controls for higher-moment preferences.

Second, we investigate the theoretical properties of the proposed framework. We begin by deriving closed-form expressions for the objective functions related to the SLN and GN target distributions. We then consider the special case in which the target distribution is Gaussian and show that the objective function decomposes as a sum of three terms. The first two terms measure the fit to the target-return mean and variance. The third term is the entropy of the standardized portfolio return and shrinks the higher moments toward the Gaussian ones, which is desirable in terms of

downside risk. Finally, we study the classical setting in which asset returns are assumed Gaussian and demonstrate that if the target-return mean and variance are located on or above the MVE frontier, then the MEKL portfolio is MVE too, though its first two moments are generally different from those of the target. Otherwise, if asset returns are not Gaussian, the investor might be better off with a portfolio that is not MVE provided that its return distribution is closer to the target.

Third, we evaluate empirically how the MEKL portfolio strategy can help improve the higher moments of MVE portfolios. We consider daily returns from three datasets of characteristic-sorted portfolios and a dataset of 100 individual stocks. We use the global-minimum-variance (GMV) portfolio and the robust mean-variance portfolio of Kan et al. (2022) as reference portfolios, and find that the MEKL portfolio compromises between mean-variance efficiency and the higher moments of the target distribution. Moreover, the MEKL portfolio delivers a statistically smaller downside risk than the chosen reference portfolio according to the modified value-at-risk of Zangari (1996), and a smaller EKL divergence as desired. These improvements offered by our approach result in portfolio-return distributions with significantly thinner left tails.[2] Finally, the tradeoff between mean-variance efficiency and improved higher moments is more satisfying than that in the higher-moment approaches of Athayde and Flores (2004), Briec et al. (2007), Martellini and Ziemann (2010), and Lassance et al. (2022). The results are robust to imposing transaction costs and short-sale constraints, and also to different specifications of the target distribution.

We close with a review of related literature. Chamakh (2021, Chapter IV) extends a previous version of our paper by exploiting different divergence measures between the portfolio-return and target-return distributions. Specifically, the author shows that the Wasserstein distance and Stein discrepancy deliver portfolios robust to misspecification about the distribution of returns.

Our work contributes to the literature on portfolio selection with preferences beyond mean and variance. A first approach is to construct higher-moment efficient surfaces; see Hirschberger et al. (2013) and Adcock (2014). The standard way to then choose a portfolio on the surface is to optimize Taylor-series expansions of expected utility (Guidolin and Timmermann, 2008, Martellini and Ziemann, 2010). However, the third and fourth moments typically have a negligible impact on portfolios optimizing expected utility (Kroll et al., 1984, Markowitz, 2014), which is at odds with prominent behavioral theories of choice under risk (Ebert and Karehnke, 2021). A second approach is to control for higher moments using constraints; see Athayde and Flores (2004) and Briec et al. (2007) for two popular examples. A third approach is to minimize tail risk criteria such

[2]For example, on average across the four datasets, GMV delivers a skewness of -0.71, an excess kurtosis of 23.70, and a daily 1% VaR of 6.30%. In contrast, our MEKL portfolio that targets an SLN distribution whose mean and variance are equal to those of GMV but that has a positive skewness of one delivers, on average, a skewness of -0.54, an excess kurtosis of 16.00, and a daily 1% VaR of 5.42%.

as the value-at-risk (El Ghaoui et al., 2003) and the expected shortfall (Rockafellar and Uryasev, 2000). A final approach is to improve higher moments indirectly, i.e., without including them in the optimization program, via robust optimization (Kim et al., 2014) or factor diversification (Lassance et al., 2022).

Our approach is *comparative*: the preference for one portfolio or an other is assessed relative to a targeted return distribution. Other target-based methods exist that minimize the tracking error relative to a benchmark, possibly while controlling for downside risk or mean return (Gaivoronski et al., 2005, Alexander and Baptista, 2010, Clark et al., 2022).

Our paper is related to other approaches. Bera and Park (2008) minimize the KL divergence between portfolio weights and target weights, subject to moment constraints. Chalabi and Würtz (2012) study the mathematics of a similar problem via a dual representation of the optimization program. The KL divergence is also extensively used in distributionally robust portfolio optimization, which maximizes worst-case performance assuming that the portfolio-return distribution lies in an ambiguity set at some prescribed distance from a nominal distribution; see Calafiore (2007), Glasserman and Xu (2013), Hu and Hong (2013), and Penev et al. (2019). Finally, our method is conceptually related to approaches in derivatives pricing that aim to replicate a payoff distribution, such as the cost-efficient portfolio (Bernard et al., 2014), or a terminal-wealth distribution, such as optimal positioning (Haugh and Lo, 2001).

The paper is organized as follows. Section 2 presents the target-distribution framework. Section 3 derives theoretical results for specific target distributions. Section 4 contains our empirical analysis in which we apply our framework to improve the higher moments of MVE portfolios. Section 5 concludes. The internet appendix contains the proofs of all results and additional material.

2 The target-distribution framework

We now introduce the divergence measure and the MEKL optimal portfolio. In this section, the target-return density is not specified and left free of choice for the investor.

2.1 Notation

We consider a single-period setting and denote by $R \in \mathbb{R}^n$ the column vector of asset returns during that period, which has mean μ and positive-definite covariance matrix $\mathbf{\Sigma}$. Unless otherwise stated, we do not assume any particular distribution for R. The investor searches for an optimal portfolio w, which belongs to a set $\mathcal{W}$ that can include any desired constraint. The return of this portfolio is $P := w'R$, which admits a density f_P with support set Ω_P. The first four portfolio-return moments,

assumed finite, are the mean $\mu_P := w'\mu$, the variance $\sigma_P^2 := w'\Sigma w$, the skewness $\zeta_P := \mathbb{E}\left[P^{\star 3}\right]$, and the excess kurtosis $\kappa_P := \mathbb{E}\left[P^{\star 4}\right] - 3$, where $P^\star$ is the standardized portfolio return:

$$P^\star := (P - \mu_P)/\sigma_P. \tag{1}$$

The central moments are $m_{k,P}$ for $k \geq 3$. The set of attainable portfolio mean and volatility is

$$\mathcal{MV} := \left\{(\mu_0, \sigma_0) \in \mathbb{R} \times \mathbb{R}^+ \mid \exists w \in \mathcal{W} : w'\mu = \mu_0, w'\mathbf{\Sigma} w = \sigma_0^2\right\}. \tag{2}$$

2.2 The divergence measure

We consider a target-return density f_T that captures the investor's preferences. Then, the optimal portfolio has a return density f_P that is as close as possible to f_T. To that end, we first review the KL divergence (Kullback and Leibler, 1951) between f_P and f_T, which requires the support set Ω_P to be included in Ω_T. It is defined as follows.

Definition 1 (Kullback-Leibler divergence)**.** *Let $\Omega_P \subseteq \Omega_T$. Then, the KL divergence between f_P and f_T is defined as*

$$\langle f_P | f_T \rangle_1 := \mathbb{E}\left[\ln \frac{f_P(P)}{f_T(P)}\right] = \int_{\Omega_P} f_P(x) \ln \frac{f_P(x)}{f_T(x)} \mathrm{d}x. \tag{3}$$

The KL divergence is a measure of divergence because $\langle f_P | f_T \rangle_1 \geq 0$ for all f_P, f_T and $\langle f_P | f_T \rangle_1 = 0$ if and only if $f_P = f_T$ almost everywhere.[3] It is a standard measure of divergence and is used in many financial applications; see Jiang et al. (2018) and references therein. We also rely on the KL divergence for three main reasons. First, it takes the form of a log-likelihood ratio in (3), and thus, minimizing the KL divergence has the intuitive interpretation of maximizing the likelihood that P is distributed as T. Second, it accounts for higher moments and can capture investors' preferences beyond mean and variance. Third, it leads to an objective function that is mathematically tractable when the target-return density belongs to the exponential family. Indeed, the KL divergence admits the decomposition

$$-\langle f_P | f_T \rangle_1 = H[P] + \mathbb{E}[\ln f_T(P)], \tag{4}$$

where

$$H[P] := -\mathbb{E}[\ln f_P(P)] \tag{5}$$

is the portfolio-return entropy (Shannon, 1948) and the second term simplifies when f_T has an

[3]The KL divergence is not symmetric: $\langle f_P | f_T \rangle_1 \neq \langle f_T | f_P \rangle_1$, in general. The Gaussian target density we study in Section 3.2 clarifies why the direction $\langle f_P | f_T \rangle_1$ is more appropriate: minimizing the other direction amounts to fit only the first two moments of the target distribution, whereas our chosen direction also fits the higher moments.

exponential expression. The tractability of the KL divergence is one of its key benefits, which has been noted in other contexts. For example, in distributionally robust optimization, Hu and Hong (2013) conclude their paper by writing that "[...] the properties of KL divergence make it a very special and tractable distance measure and a good candidate for modeling ambiguities."

A limitation of the KL divergence is that it is not defined when $\Omega_P \not\subseteq \Omega_T$, a condition that depends on the chosen target density. For example, the SLN target distribution is defined on $\Omega_T = \{x \in \mathbb{R} | x > \delta\}$, where $\delta \in \mathbb{R}$ is a fixed parameter. For this class of target distributions, Ω_P is not contained in Ω_T when $\Omega_P = \mathbb{R}$, for example. To admit such cases, we introduce the *extended KL (EKL) divergence*, which is defined as the KL divergence between the conditional density $f_{P|P\in\Omega_T}$ and the target density f_T.

Definition 2 (Extended KL divergence)**.** *Let $p := \mathbb{P}(P \notin \Omega_T)$. Then, the EKL divergence between f_P and f_T is defined as $\langle f_P | f_T \rangle_2 := +\infty$ if $p = 1$ and, otherwise,*

$$\langle f_P | f_T \rangle_2 := \langle f_{P|P\in\Omega_T} | f_T \rangle_1 = \mathbb{E}\left[\ln \frac{f_P(P)}{f_T(P)} \middle| P \in \Omega_T \right] - \ln(1-p) \geq 0. \tag{6}$$

Minimizing the EKL divergence amounts to fit the two densities on the intersection of their support set, $\Omega_T \cap \Omega_P$, but also to minimize the probability $p = \mathbb{P}(P \notin \Omega_T)$, ensuring that the support sets tend to coincide. The EKL divergence corresponds to the plain KL divergence when $p = 0$ and thus $\Omega_P \subseteq \Omega_T$, and it admits a similar decomposition to that in (4):

$$-\langle f_P | f_T \rangle_2 = H[P | P \in \Omega_T] + \mathbb{E}[\ln f_T(P) | P \in \Omega_T] + \ln(1-p), \tag{7}$$

where $H[P|P \in \Omega_T] := -\mathbb{E}[\ln f_P(P) | P \in \Omega_T]$.

2.3 The optimal portfolio

Given that the EKL divergence generalizes the KL, we define the optimal portfolio such that, given a target-return density f_T, it minimizes the EKL divergence between f_P and f_T.

Definition 3 (Minimum-EKL portfolio)**.** *The MEKL portfolio minimizes the EKL divergence between the portfolio-return density f_P and the target-return density f_T:*[4]

$$w_{f_T} := \underset{w\in\mathcal{W}}{\operatorname{argmax}}\ \mathcal{C}(w; f_T) \quad \textit{with} \quad \mathcal{C}(w; f_T) := -\langle f_{P(w)} | f_T \rangle_2. \tag{8}$$

[4]We consider equalities in different specifications of $\mathcal{C}(w; f_T)$ to hold up to additive and strictly positive multiplicative constants independent of w, because they do not affect the solution w_{f_T}.

Definition 3 can be applied to different problems, which correspond to different choices for the target distribution. For instance, an investor may wish to target the return distribution of a primary portfolio (e.g., built from a large number of assets) by investing in a secondary portfolio (e.g., built from a subset of all assets so as to reduce trading costs).[5] In the empirical analysis of Section 4, we consider an alternative application: improving the higher moments of a chosen MVE portfolio. To achieve this, we choose a parametric family for the target distribution and determine its parameters such that (i) the first two moments coincide with those of the MVE portfolio, but (ii) the higher moments of the target distribution improve upon those of the MVE portfolio.

3 Theoretical results for specific target distributions

To operationalize the proposed approach and analyze its properties, we need to consider particular target distributions. It is important to note that the distributions commonly used in finance to model asset returns, such as the (skew-)Student distribution or the hyperbolic distribution, may not be those that investors would wish to target. Indeed, investors would typically target distributions with positive skewness and thin tails, which is at odds with the characteristics of asset returns.

In this section, we consider two classes of target distribution. First, the SLN distribution, which allows for varying skewness. Second, the GN distribution, which allows for varying kurtosis. For brevity, all results for the GN target are relegated to Section IA.1 of the internet appendix. These two choices are appealing because (i) they are parsimonious (only three parameters), (ii) they allow for a wide range of skewness or kurtosis, (iii) they belong to the exponential family, which makes the EKL divergence analytically tractable, and (iv) we find empirically that they improve the higher moments of portfolio returns in the desired way.

In this section, we only impose a budget constraint for tractability: $\mathcal{W} = \left\{w \in \mathbb{R}^n \mid w'\iota = 1\right\}$, where ι is an n-dimensional vector of ones.

3.1 The shifted-log-normal target distribution

We consider an investor who targets the asymmetric SLN distribution, which dates back from Galton (1879). We refer to Johnson et al. (1994) for a review of its genesis and properties, and Kaplan and Knowles (2004) and Kadan and Liu (2014, Sec. 4.2.1) for more recent financial illustrations.

The SLN distribution is less commonly used than the three-parameter skew-normal distribution of Azzalini (1985), but we opt for the SLN for three reasons. First, the SLN can have any real value of skewness while it is bounded in $[-1, 1]$ (approximately) for the skew-normal. Second, the SLN

[5] We consider such an application in our additional empirical experiments of Section 4.8.

belongs to the exponential family and thus is analytically tractable in our context. Third, we show in Section IA.2 of the internet appendix that the SLN and skew-normal distributions with the same first three moments are nearly indistinguishable and deliver very similar MEKL portfolios.

Definition 4 (Shifted-log-normal target)**.** *The target return follows an SLN distribution,* $T \sim \mathcal{SLN}(\alpha, \beta, \delta)$*, if its density is a log-normal density shifted by* δ*:*

$$f_T(x; \alpha, \beta, \delta) := \frac{1}{(x-\delta)\beta\sqrt{2\pi}} \exp\left(-\frac{1}{2\beta^2}\big(\ln(x-\delta)-\alpha\big)^2\right), \tag{9}$$

where $x > \delta$*,* $\alpha, \delta \in \mathbb{R}$*, and* $\beta > 0$*. The first three moments of* T *are*

$$\mu_T = \delta + \exp\left(\alpha + \beta^2/2\right), \quad \sigma_T^2 = g(\beta)(g(\beta)+1)\exp(2\alpha), \quad \zeta_T = (g(\beta)+3)g(\beta)^{1/2}, \tag{10}$$

where $g(\beta) := \exp(\beta^2) - 1$*.*

Note that the skewness ζ_T is always positive, but can be made negative by considering $-T$ as a target. The investor can calibrate the SLN parameters (α, β, δ) via moment-matching using results by Yuan (1933). Specifically, the parameters (α, β, δ) matching the moments $(\mu_T, \sigma_T^2, \zeta_T)$ are

$$\begin{aligned}
\beta &= \sqrt{\ln(A + A^{-1} - 1)}, \quad A = \left(1 + \frac{\zeta_T^2}{2} + \sqrt{\frac{\zeta_T^4}{4} + \zeta_T^2}\right)^{1/3}, \\
\alpha &= \frac{1}{2}\ln\left(\frac{\sigma_T^2}{g(\beta)(g(\beta)+1)}\right), \\
\delta &= \mu_T - \sigma_T/\sqrt{g(\beta)}.
\end{aligned} \tag{11}$$

As mentioned in Johnson et al. (1994), we recover the Gaussian distribution when the parameter β, and thus the SLN skewness, tends to zero. That is, for a fixed (μ_T, σ_T), we have that

$$\lim_{\zeta_T \to 0^+} f_T(x; \alpha, \beta, \delta) = \lim_{\beta \to 0^+} f_T(x; \alpha, \beta, \delta) = \frac{1}{\sigma_T}\phi\left(\frac{x-\mu_T}{\sigma_T}\right), \tag{12}$$

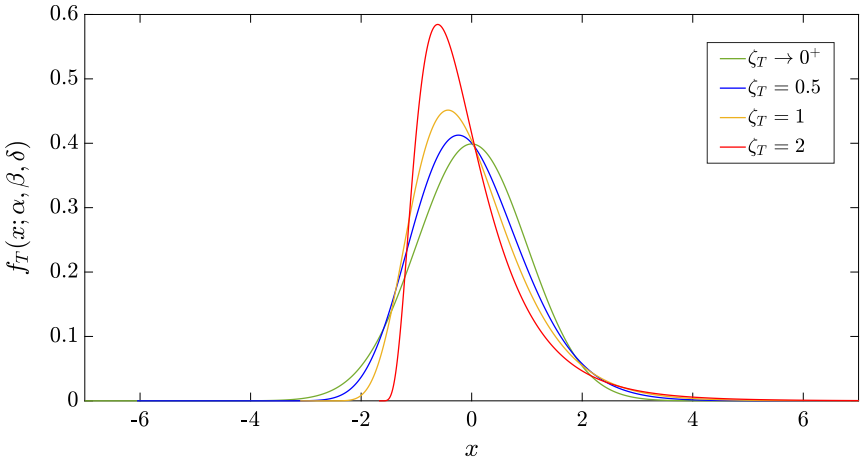
where ϕ is the standard Gaussian density. In Figure 1, we depict the zero-mean unit-variance SLN density with different skewness ζ_T, and observe indeed that it tends to a Gaussian as $\zeta_T \to 0^+$.

In the next proposition, we derive the objective function in (8) associated with the SLN target.

Proposition 1 (MEKL portfolio with SLN target)**.** *Let* $T \sim \mathcal{SLN}(\alpha, \beta, \delta)$ *and* $P_\delta := P - \delta$*. Then, the MEKL portfolio* $w_{f_T} = w_{\alpha,\beta,\delta}$ *is obtained by maximizing*

$$\mathcal{C}(w; \alpha, \beta, \delta) = H[P_\delta | P_\delta \geq 0] - \mathbb{E}[\ln P_\delta | P_\delta \geq 0]$$

Figure 1: Density of the shifted-log-normal distribution



Note. This figure depicts the shifted-log-normal density in (9) with mean zero, variance one, and skewness ζ_T between 0 and 2.

$$- \frac{1}{2\beta^2}\left(\mathbb{V}[\ln P_\delta | P_\delta \geq 0] + (\mathbb{E}[\ln P_\delta | P_\delta \geq 0] - \alpha)^2\right) + \ln(1-p). \tag{13}$$

For tractability purposes, we rely on an approximation of the objective function (13) by i) replacing $\mathbb{E}[\ln P_\delta | P_\delta \geq 0]$ and $\mathbb{V}[\ln P_\delta | P_\delta \geq 0]$ by their second-order Taylor-series expansion, and ii) assuming that $p = \mathbb{P}(P < \delta)$ is small. The resulting approximation can be shown to be[6]

$$\widetilde{\mathcal{C}}(w;\alpha,\beta,\delta) := \; H[P^\star] + \ln \sigma_P - \widetilde{\mathbb{E}}[\ln P_\delta | P_\delta \geq 0] - \frac{1}{2\beta^2}\left(\widetilde{\mathbb{V}}[\ln P_\delta | P_\delta \geq 0] + (\widetilde{\mathbb{E}}[\ln P_\delta | P_\delta \geq 0] - \alpha)^2\right), \tag{14}$$

where

$$\widetilde{\mathbb{E}}[\ln P_\delta | P_\delta \geq 0] := \ln(\mu_P - \delta) - \frac{\sigma_P^2}{2(\mu_P - \delta)^2}, \tag{15}$$

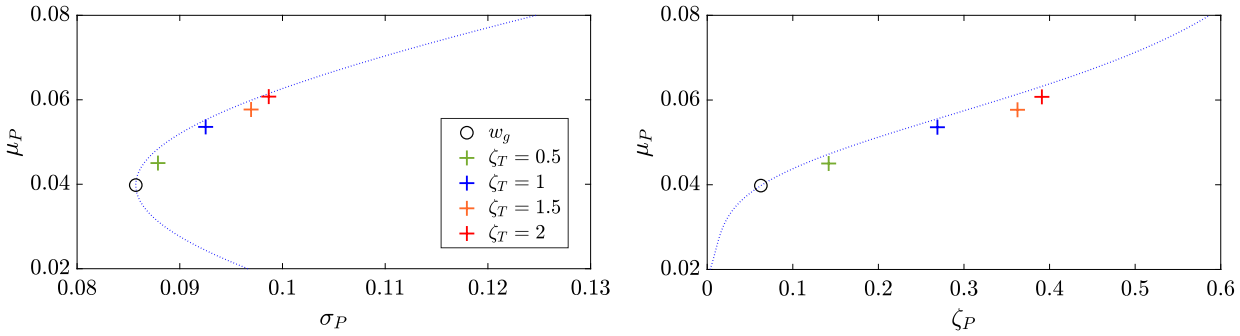
$$\widetilde{\mathbb{V}}[\ln P_\delta | P_\delta \geq 0] := \frac{\sigma_P^2}{(\mu_P - \delta)^2} - \frac{m_{3,P}}{(\mu_P - \delta)^3} + \frac{m_{4,P} - \sigma_P^4}{4(\mu_P - \delta)^4}. \tag{16}$$

The approximate MEKL portfolio that maximizes $\widetilde{\mathcal{C}}(w;\alpha,\beta,\delta)$ is denoted $\widetilde{w}_{\alpha,\beta,\delta}$. We show in Section IA.4 of the internet appendix that this approximation is very accurate, particularly when the skewness ζ_T of the SLN target distribution is not too extreme. Indeed, the smaller the ζ_T, the more δ decreases and the more accurate the approximation, which becomes exact when $\zeta_T \to 0^+$.

In the next proposition, we derive the optimal portfolio-return skewness ζ_P and excess kurtosis κ_P maximizing the objective function $\widetilde{\mathcal{C}}(w;\alpha,\beta,\delta)$.

[6]One could also use a fourth-order expansion in (15) to account for the effect of skewness and kurtosis. However, this complicates the derivation of analytical results because $\widetilde{\mathbb{E}}[\ln P_\delta | P_\delta \geq 0]$ appears squared in (14). In unreported results, we find that the portfolio weights maximizing $\widetilde{\mathcal{C}}(w;\alpha,\beta,\delta)$, and the resulting portfolio-return moments, are very similar with a second or a fourth-order expansion in (15).

Figure 2: Optimal portfolio for the shifted-log-normal target distribution



Note. This figure depicts the results in Example 1, which considers $n = 3$ independent asset returns with mean $(0.02, 0.07, 0.15)$, volatility $(0.10, 0.20, 0.30)$, skewness $(0, 0.5, 1)$, and excess kurtosis $(1, 2, 3)$. We consider a shifted-log-normal target-return distribution, $T \sim \mathcal{SLN}(\alpha, \beta, \delta)$, whose parameters are matched to the return mean and variance of the global-minimum-variance (GMV) portfolio, $w_g = (0.74, 0.18, 0.08)'$, and to a skewness ζ_T between 0.5 and 2. The two plots depict the mean-variance-skewness tradeoff of the minimum-EKL portfolios (colored crosses) and GMV portfolio (black circle) as a function of ζ_T. The blue dotted curves depict the tradeoff for the mean-variance portfolios with mean return $\mu_P \in [0.02, 0.08]$.

Proposition 2 (Optimal skewness and kurtosis with SLN target)**.** *Let* $(\mu_P, \sigma_P) = (\mu_T, \sigma_T)$ *be fixed. Then, the skewness* ζ_P *and excess kurtosis* κ_P *maximizing* $\widetilde{\mathcal{C}}(w; \alpha, \beta, \delta)$ *in* (14) *are*

$$\zeta_P = 3\beta^{-2} g(\beta)^{3/2} \quad \textit{and} \quad \kappa_P = -3\beta^{-2} g(\beta)^2 < 0, \tag{17}$$

provided they are attainable, where ζ_P *is close to* ζ_T*.*[7]

Proposition 2 indicates that targeting an SLN distribution with skewness ζ_T amounts to partly fit ζ_T and is also best achieved for a negative excess kurtosis, which together reduce downside risk. In the next example, we illustrate the MEKL portfolio under an SLN target.

Example 1. Consider $n = 3$ independent asset returns with mean $(0.02, 0.07, 0.15)$, volatility $(0.10, 0.20, 0.30)$, skewness $(0, 0.5, 1)$, and excess kurtosis $(1, 2, 3)$. We match the mean and variance of the SLN target to those of the GMV portfolio, $w_g = (0.74, 0.18, 0.08)'$. The skewness ζ_T varies between 0.5 and 2, which improves upon the skewness of GMV that equals 0.06. We observe that the higher the ζ_T, the more the MEKL portfolio invests in the third asset that has the largest skewness: $w_3 = 0.09$ for $\zeta_T = 0.5$ versus $w_3 = 0.19$ for $\zeta_T = 1$. In Figure 2, we depict the mean-variance-skewness tradeoff of the GMV and MEKL portfolios. The higher the ζ_T, the more the MEKL portfolio departs from the GMV portfolio to improve the fit to ζ_T.

3.2 The Gaussian target distribution

We now study the Gaussian target-return distribution, which is a natural and appealing benchmark because the higher moments of returns are typically negatively skewed and leptokurtic, and thus,

[7] Taking β as a function of ζ_T in (11), the first-order Taylor-series expansion of ζ_P around $\zeta_T = 0$ is $\zeta_P = \zeta_T$.

are less desirable than the Gaussian ones. The Gaussian target, $T \sim \mathcal{N}(\alpha, \beta)$, has a density

$$f_T(x; \alpha, \beta) = \frac{1}{\beta}\phi\left(\frac{x-\alpha}{\beta}\right), \tag{18}$$

where $\beta > 0$. We show in the next proposition that the objective function involves three separate terms that measure the fit to the target-return mean, variance, and higher moments, respectively. Note that the EKL divergence corresponds to the KL divergence in this case because $\Omega_T = \mathbb{R}$.

Proposition 3 (MEKL portfolio with Gaussian target)**.** *Let $T \sim \mathcal{N}(\alpha, \beta)$. Then, the MEKL portfolio $w_{f_T} = w_{\alpha,\beta}$ is obtained by maximizing*

$$\mathcal{C}(w; \alpha, \beta) = -\frac{1}{2}\left(\frac{\mu_P - \alpha}{\beta}\right)^2 + \frac{1}{2}\left(\ln \sigma_P^2 - \frac{\sigma_P^2}{\beta^2}\right) + H[P^\star], \tag{19}$$

where $H[P^\star]$ is the entropy of the standardized portfolio return, which is invariant to μ_P and σ_P.

Three comments are in order. First, the first two terms in (19) measure the fit to the target-return mean and variance, α and β, respectively.

Second, the KL divergence penalizes a misfit to the standard deviation β more strongly than a misfit to the mean α. Indeed, maximizing the sum of the first two terms in (19) amounts to minimize the squared ℓ_2 distance between (μ_P, σ_P) and (α, β) plus an extra penalty $2\beta(\sigma_P - \beta \ln \sigma_P)$ that is minimized for $\sigma_P = \beta$.[8] Being more tolerant about a misfit in mean return than in variance is sensible in practice given the large estimation errors in sample mean returns (Merton, 1980).

Third, concerning the entropy term $H[P^\star]$, it is a well-known result from information theory that the Gaussian distribution has the largest entropy among all other distributions defined on the support set $\mathbb{R}$ (Cover and Thomas, 2006). In other words, $H[P^\star]$ is maximized when P is Gaussian. Therefore, the KL divergence does not only try to fit (μ_P, σ_P) to (μ_T, σ_T) but also to shrink the higher moments of portfolio returns toward those of the Gaussian. This can be seen from the second-order Gram-Charlier expansion of entropy (Hyvärinen et al., 2001),

$$H[P^\star] \approx \frac{1}{2}\ln(2\pi e) - \frac{\zeta_P^2}{12} - \frac{\kappa_P^2}{48}, \tag{20}$$

which is maximized when the skewness and excess kurtosis of P are zero, as expected.

Finally, when the target volatility $\beta \to 0^+$, we obtain a Dirac-delta target distribution, which we study in detail in Section IA.3 of the internet appendix. We show in particular that the resulting MEKL portfolio is a mean-variance portfolio whose mean return is located between that of the

[8] The ℓ_2 distance between (μ_P, σ_P) and (α, β) also corresponds to the squared Wasserstein distance between P and T when both are Gaussian.

target and that of the GMV portfolio.

3.3 Gaussian asset returns and mean-variance efficiency

In this section, we analyze the theoretical properties of the MEKL portfolio targeting an SLN distribution in the classical setting in which asset returns are Gaussian.

For that purpose, we introduce several definitions related to mean-variance portfolios when only a budget constraint is imposed. The mean return and variance of the GMV portfolio $w_g := (\iota'\mathbf{\Sigma}^{-1}\iota)^{-1}\mathbf{\Sigma}^{-1}\iota$ are $\mu_g := \iota'\mathbf{\Sigma}^{-1}\mu(\iota'\mathbf{\Sigma}^{-1}\iota)^{-1}$ and $\sigma_g^2 := (\iota'\mathbf{\Sigma}^{-1}\iota)^{-1}$. The maximum squared Sharpe ratio is $\theta^2 := \mu'\mathbf{\Sigma}^{-1}\mu$ and the difference between θ^2 and the squared Sharpe ratio of the GMV portfolio is denoted $\psi^2 := \theta^2 - \mu_g^2/\sigma_g^2 \geq 0$. Finally, the MVE frontier is

$$\sigma_{EF}(\mu_0) := \sqrt{\sigma_g^2 + \frac{(\mu_0 - \mu_g)^2}{\psi^2}}, \quad \mu_0 \geq \mu_g. \tag{21}$$

We now derive the moments of the Gaussian distribution minimizing the EKL divergence with respect to the SLN distribution.

Proposition 4 (Gaussian distribution with minimum EKL)**.** *The moments of the Gaussian distribution minimizing the EKL divergence with respect to the density of* $T \sim \mathcal{SLN}(\alpha, \beta, \delta)$ *according to the approximation* (14) *satisfy*

$$(\mu_{SLN}^\star, \sigma_{SLN}^\star) := (\mu_T, c_1(\beta)\sigma_T), \tag{22}$$

where

$$c_1(\beta) := \beta/g(\beta)^{1/2} \leq 1, \tag{23}$$

with equality if and only if $\beta \to 0^+$.

We conclude from this result that if there exists a portfolio $w \in \mathcal{W}$ whose first two return moments are equal to (22), then this portfolio will coincide with the MEKL portfolio $\widetilde{w}_{\alpha,\beta,\delta}$. However, this does not tell us what are the properties of the MEKL portfolio when the first two moments of the SLN target are located on or above the MVE frontier. Indeed, in such a case, $\sigma_{SLN}^\star \leq \sigma_T \leq \sigma_{EF}(\mu_T)$ so that $(\mu_T, \sigma_{SLN}^\star) \notin \mathcal{MV}$ and no portfolio $w \in \mathcal{W}$ can deliver those moments. Choosing the target-return mean and variance on the MVE frontier is a natural choice, which we consider in the application of Section 4. As mentioned in Section 2.3, a classical application in which the investor may select the target-return mean and variance above the MVE frontier is when she restricts to investing in only a subset of the investment universe (e.g., to reduce trading

costs), and wishes to target the performance of the optimal portfolio computed from all assets. The next proposition guarantees that the MEKL portfolio $\widetilde{w}_{\alpha,\beta,\delta}$ remains MVE in such a case, for any SLN target distribution.[9]

Proposition 5 (MEKL portfolio with Gaussian asset returns)**.** *Let $R \sim \mathcal{N}(\mu, \mathbf{\Sigma})$ and the target-return mean and variance are located on or above the mean-variance efficient frontier: $\mu_T \geq \mu_g$ and $\sigma_T \leq \sigma_{EF}(\mu_T)$. Then, the MEKL portfolios $\widetilde{w}_{\alpha,\beta,\delta}$ and $w_{\alpha,\beta}$ satisfy the following:*

1. *$\widetilde{w}_{\alpha,\beta,\delta}$ and $w_{\alpha,\beta}$ are mean-variance efficient.*

2. *Focusing on $w_{\alpha,\beta}$ associated with a Gaussian target, the mean return $\mu_P^\star = w_{\alpha,\beta}'\mu$ solves*

$$(\mu_P^\star - \mu_g)(\sigma_P^{\star 2} - \beta^2) = \sigma_P^{\star 2}(\alpha - \mu_P^\star)\psi^2 \quad \textit{with} \quad \sigma_P^\star = \sigma_{EF}(\mu_P^\star). \tag{24}$$

Moreover, $\mu_P^\star = \mu_g$ if $\alpha = \mu_g$ and $\mu_P^\star \leq \mu_{\ell_2} \leq \alpha$ if $\alpha > \mu_g$, where μ_{ℓ_2} is the mean return of the mean-variance portfolio minimizing the ℓ_2 distance to (α, β).

Three comments are in order. First, MEKL portfolios are located on the MVE frontier when asset returns are Gaussian, which is consistent with Markowitz's theory. The core of the proof consists in showing that the EKL divergence has no other stationary point in the mean-variance plane than the global optimum in Proposition 4, which as a by-product is desirable from an optimization perspective. This implies geometrically that the MEKL portfolio is MVE because, from Proposition 4, the global optimum is located above the MVE frontier when so is (μ_T, σ_T).

Second, the results in Proposition 5 hold when asset returns are Gaussian. For non-Gaussian returns, the investor might be better off choosing a portfolio outside of the MVE frontier if it better balances the fit to the target-return mean-variance and higher moments, as in Figure 2.

Third, when targeting a Gaussian distribution, we have more specific bounds on the mean return of the MEKL portfolio. In particular, the EKL divergence induces a larger fit to the target variance than to the target mean relative to the ℓ_2 distance.

4 Application: improving higher moments of MVE portfolios

As a particular application of our framework, we now evaluate how the MEKL portfolio strategy can help improve the higher moments of MVE portfolios. Section 4.1 explains the reference portfolio approach we use to fix the first two moments of the target distribution. Section 4.2 contains our

[9] When targeting a Gaussian distribution, the result in Proposition 5 holds more generally when asset returns are elliptically distributed with finite second moment. This is because the standardized entropy $H[P^\star]$ in (19) is independent of portfolio weights w for elliptically distributed returns as well.

estimation method. Section 4.3 states hypotheses from our theoretical results. Section 4.4 presents the data we use. Section 4.5 details the considered performance criteria and target distributions. Sections 4.6 and 4.7 present the methodology and analyze the results. Finally, Section 4.8 conducts additional empirical experiments.

4.1 The reference portfolio

Unless specified otherwise, we match the target-return mean and variance to those of a reference portfolio located on the MVE frontier, which is justified given the mean-variance dominance of MVE portfolios. Specifically, given a portfolio w_0 satisfying $\mu_0 \geq \mu_g$ and $\sigma_0 = \sigma_{EF}(\mu_0)$, we fix the target-return mean and variance as

$$\left(\mu_T, \sigma_T^2\right) \leftarrow (\mu_0, \sigma_0^2) = \left(w_0'\mu, w_0'\boldsymbol{\Sigma} w_0\right), \tag{25}$$

which means that they are attainable, $(\mu_T, \sigma_T) \in \mathcal{MV}$. As a result, the MEKL portfolio offers a compromise between (i) the optimal mean-variance tradeoff of the reference portfolio and (ii) the desired higher-moment properties embedded in the target distribution.

4.2 Estimation method

We now explain how we estimate the objective functions associated with the SLN and Gaussian target distributions. The estimation method for the GN target distribution is available in Section IA.1 of the internet appendix. We consider a sample of historical asset-return observations $R_{i,t}$, with $i = 1, \ldots, n$ and $t = 1, \ldots, h$. We denote R_t the asset-return vector at time t and, for a given portfolio w, the portfolio return at time t is $P_t = w'R_t$.

We first estimate the MVE reference portfolio w_0 via $\hat{w}_0$ and recover its return mean and variance,

$$\hat{\mu}_0 = \hat{w}_0'\hat{\mu} \quad \text{and} \quad \hat{\sigma}_0^2 = \hat{w}_0'\widehat{\boldsymbol{\Sigma}}\hat{w}_0, \tag{26}$$

where $\hat{\mu} := h^{-1}\sum_{t=1}^{h} R_t$ and $\widehat{\boldsymbol{\Sigma}} := (h-1)^{-1}\sum_{t=1}^{h}(R_t - \hat{\mu})(R_t - \hat{\mu})'$. To estimate the SLN target objective function $\widetilde{\mathcal{C}}(w; \alpha, \beta, \delta)$ in (14), we calibrate the parameters (α, δ) to $(\hat{\mu}_0, \hat{\sigma}_0)$ in (26),

$$\hat{\alpha} = \frac{1}{2}\ln\left(\frac{\hat{\sigma}_0^2}{\exp(\beta^2)g(\beta)}\right) \quad \text{and} \quad \hat{\delta} = \hat{\mu}_0 - \hat{\sigma}_0 g(\beta)^{-1/2}, \tag{27}$$

and β is fixed exogenously by the investor, e.g., to match a desired skewness ζ_T in (11). Using the

property $H[P^\star] = H[P] - \ln \sigma_P$, the estimated objective function depends on β and $\hat{w}_0$ as

$$\begin{aligned}
&\widehat{\mathcal{C}}(w;\beta,\hat{w}_0) = \widehat{H}[P] - \widehat{\mathbb{E}}[\ln P_\delta | P_\delta \geq 0] - \frac{1}{2\beta^2}\left(\widehat{\mathbb{V}}[\ln P_\delta | P_\delta \geq 0] + \left(\widehat{\mathbb{E}}[\ln P_\delta | P_\delta \geq 0] - \hat{\alpha}\right)^2\right),\\
&\widehat{\mathbb{E}}[\ln P_\delta | P_\delta \geq 0] = \ln(\hat{\mu}_P - \hat{\delta}) - \frac{\hat{\sigma}_P^2}{2(\hat{\mu}_P - \hat{\delta})^2},\\
&\widehat{\mathbb{V}}[\ln P_\delta | P_\delta \geq 0] = \frac{\hat{\sigma}_P^2}{(\hat{\mu}_P - \hat{\delta})^2} - \frac{\hat{m}_{3,P}}{(\hat{\mu}_P - \hat{\delta})^3} + \frac{\hat{m}_{4,P} - \hat{\sigma}_P^4}{4(\hat{\mu}_P - \hat{\delta})^4}.
\end{aligned} \tag{28}$$

The estimate (28) involves estimating the portfolio-return first four moments and entropy. Concerning the moments, we rely on the unbiased sample estimates (Kendall and Stuart, 1963):

$$\begin{aligned}
&\hat{\mu}_P = \frac{1}{h}\sum_{t=1}^{h} P_t, \quad \hat{\sigma}_P^2 = \frac{1}{h-1}\sum_{t=1}^{h}(P_t - \hat{\mu}_P)^2, \quad \hat{m}_{3,P} = \frac{h}{(h-1)(h-2)}\sum_{t=1}^{h}(P_t - \hat{\mu}_P)^3,\\
&\hat{m}_{4,P} = \frac{1}{h^2 - 3h + 3}\left(\frac{h^2}{h-1}\sum_{t=1}^{h}(P_t - \hat{\mu}_P)^4 - 3(2h-3)\hat{\sigma}_P^4\right).
\end{aligned} \tag{29}$$

To alleviate estimation risk, we also consider the robust shrinkage estimators of the variance, skewness, and kurtosis in Ledoit and Wolf (2004), Martellini and Ziemann (2010), and Boudt et al. (2020). However, because we use daily return data in this empirical analysis, we find that they perform as well as the sample estimates in (29).

Concerning the entropy $H[P]$, we rely on the m-spacings estimator because it is both robust and accurate; see Vasicek (1976) and Learned-Miller and Fisher (2003).

Definition 5 (Entropy estimator)**.** *The m-spacings estimator of $H[P]$ is*

$$\widehat{H}[P] = \frac{1}{h-m}\sum_{t=1}^{h-m} \ln\left(\frac{h+1}{m}\left(P_{(t+m:h)} - P_{(t:h)}\right)\right) + \ln m - \psi(m), \tag{30}$$

where $P_{(1:h)} \leq P_{(2:h)} \leq \cdots \leq P_{(h:h)}$ are the order statistics, $m \in [1, h-1]$ is an integer, $\psi(m) := \Gamma'(m)/\Gamma(m)$ is the digamma function, and $\ln m - \psi(m)$ is a bias-correction term.

Increasing the parameter m improves robustness to outliers. This estimator is consistent under the two conditions $m, h \to \infty$ and $m/h \to 0$ and therefore m is commonly taken as a power function of h. We fix $m = \left\lfloor h^{2/3} \right\rfloor$ as in Lassance and Vrins (2021b).

Finally, when targeting a Gaussian distribution, the estimated objective function is obtained from Proposition 3 as

$$\widehat{\mathcal{C}}(w;\hat{w}_0) = \widehat{H}[P] - \frac{(\hat{\mu}_P - \hat{\mu}_0)^2 + \hat{\sigma}_P^2}{2\hat{\sigma}_0^2}. \tag{31}$$

4.3 Hypotheses

Given our theoretical results in Section 3, we draw four hypotheses to assess in our application:

1. The MEKL portfolio meets its objective: it delivers a smaller out-of-sample EKL divergence than the reference portfolio. This is the case by construction in sample, but is not necessarily true out of sample if the MEKL portfolio is too sensitive to estimation errors.

2. The mean return and variance of the MEKL portfolio are similar to those of the reference portfolio, but the MEKL portfolio has improved higher moments and thus less downside risk. We expect this because the first two moments of the target match those of the reference portfolio, but the target has more desirable higher moments.

3. The MEKL portfolio fits the target volatility better than the target mean, which we expect from part 2 of Proposition 5.

4. The MEKL portfolio has a smaller volatility than that of the reference portfolio as long as the latter is not too close to the GMV portfolio. We expect this from Proposition 4, except when the reference portfolio is close to the GMV portfolio because then few portfolios have a smaller volatility and thus it is likely more optimal for the MEKL portfolio to accept a larger volatility so as to improve higher moments relative to the reference portfolio.

4.4 Data

We consider three commonly used datasets of characteristic-sorted portfolios from Ken French's library and one dataset of individual stocks from CRSP. The characteristic-sorted portfolios are: (i) 25 portfolios formed on size and book-to-market (25SBTM), (ii) 25 portfolios formed on operating profitability and investment (25OPINV), and (iii) 30 industry portfolios (30IND). The time period spans January 1980 to March 2021. The dataset of 100 individual stocks from CRSP (100CRSP) spans January 1980 to December 2019 and is constructed so as to reduce survivorship bias similarly to Jagannathan and Ma (2003), DeMiguel et al. (2009), and Lassance et al. (2022). For all four datasets, we use daily returns because they improve estimation accuracy and because higher-moment risk is more pronounced at the daily frequency (Martellini and Ziemann, 2010).

4.5 Performance criteria and target distributions

Given a time series of length τ of out-of-sample portfolio returns $P_1, \ldots, P_\tau$, we evaluate portfolio performance in terms of moments, downside risk, and EKL divergence.

First, we report the annualized mean $252\hat{\mu}_P$, annualized volatility $\sqrt{252}\hat{\sigma}_P$, skewness $\hat{\zeta}_P = \hat{m}_{3,P}/\hat{\sigma}_P^{3/2}$, and excess kurtosis $\hat{\kappa}_P = \hat{m}_{4,P}/\hat{\sigma}_P^2 - 3$, computed as in Equation (29).

Second, we evaluate the downside risk of the portfolio via the value-at-risk (VaR). We estimate the VaR with the modified VaR (MVaR) of Zangari (1996) that accounts for the effect of skewness and kurtosis with a second-order Gram-Charlier expansion; see Eling and Schuchmacher (2007), Boudt et al. (2008), Bali et al. (2013), and Lassance et al. (2022) for financial applications. The MVaR at confidence level α is defined as

$$\mathrm{MVaR}_\alpha(P) := -\hat{\mu}_P - \hat{\sigma}_P \left(z_\alpha + \frac{1}{6}(z_\alpha^2 - 1)\hat{\zeta}_P + \frac{1}{24}(z_\alpha^3 - 3z_\alpha)\hat{\kappa}_P - \frac{1}{36}(2z_\alpha^3 - 5z_\alpha)\hat{\zeta}_P^2 \right), \tag{32}$$

where z_α is the standard Gaussian quantile. We fix $\alpha = 1\%$. Moreover, we report p-values for the null hypothesis that the MEKL and reference portfolios have the same MVaR against the hypothesis that the MEKL portfolio has a smaller MVaR. We compute the p-values by bootstrapping with replacement, using 1,000 bootstrap samples as in DeMiguel et al. (2009).

Third, we report the EKL divergence with respect to the three target distributions we consider in this empirical analysis. First, the Gaussian target studied in Section 3.2. Second, the SLN target studied in Section 3.1 with a skewness $\zeta_T = 0.5$. Third, the GN target studied in Section IA.1 of the internet appendix with a parameter $\gamma = 2.78$ corresponding to an excess kurtosis $\kappa_T = -0.5$.[10] These values of skewness and excess kurtosis improve upon the Gaussian higher moments.[11]

4.6 Out-of-sample methodology

We detail the considered portfolio strategies in Sections 4.6.1 and 4.6.2, and how we construct the time series of out-of-sample returns in Section 4.6.3. Consistent with our theory, we assume that short-selling is allowed for constructing the portfolios. In Section IA.5.3 of the internet appendix, we impose short-sale constraints and find the conclusions are robust.

4.6.1 MEKL and reference portfolios

We study the MEKL portfolios associated with the three target-return distributions in Section 4.5. We fix the mean and variance of the target distributions via two estimated reference portfolios $\hat{w}_0$. First, the sample GMV portfolio, which delivers a satisfying mean-variance performance out of sample because it ignores estimates of mean returns (DeMiguel et al., 2009). Second, the combination

[10] Negative values of excess kurtosis are rarely observed in asset returns, and thus in Section IA.5.2 we also consider the GN target with $\gamma = 1$, which corresponds to a Laplace distribution with $\kappa_T = 3$.

[11] The out-of-sample results in Section 4.7 are not very sensitive to the specific parameters chosen; they are similar when using an SLN distribution with skewness $\zeta_T = 1$ and a GN distribution with excess kurtosis $\kappa_T = -1$.

of the sample GMV and sample mean-variance portfolios that maximizes expected out-of-sample utility according to Kan et al. (2022) (KWZ portfolio), using a risk-aversion coefficient $\lambda = 10$.

4.6.2 Higher-moment portfolios from the literature

We compare our MEKL portfolios and the two reference portfolios with three other approaches put forward in the literature to improve higher moments, as discussed in Section 1. First, the approach that optimizes a four-moment Taylor-series expansion of expected utility; see Guidolin and Timmermann (2008) and Martellini and Ziemann (2010). We follow the latter and rely on the constant relative risk aversion (CRRA) utility assuming a constant mean-return vector μ to neutralize its large estimation risk. Thus, we compute the portfolio solving

$$\min_{w\in\mathcal{W}} \ \frac{\lambda}{2}\hat{\sigma}_P^2 - \frac{\lambda(\lambda+1)}{6}\hat{m}_{3,P} + \frac{\lambda(\lambda+1)(\lambda+2)}{24}\hat{m}_{4,P}, \tag{33}$$

where the estimators $\hat{\sigma}_P^2$, $\hat{m}_{3,P}$, and $\hat{m}_{4,P}$ are given by (29). To be consistent with the KWZ reference portfolio, we consider a risk-aversion coefficient $\lambda = 10$. As discussed in Section 1, this portfolio is not expected to bring much higher-moment improvements relative to the sample GMV portfolio.

Second, the approach of Briec et al. (2007), which finds the largest moment improvements relative to a benchmark portfolio w_B. We consider their mean-variance-skewness efficient portfolio and a mean-variance-kurtosis efficient portfolio found similarly. Specifically, given a benchmark portfolio w_B, we consider the two portfolios solving

$$\begin{aligned}
&\max_{\delta\in[0,1],w\in\mathcal{W}} \ \delta \quad \text{subject to} \ \ \hat{\mu}_P \geq \hat{\mu}_B, \ \hat{\sigma}_P^2 \leq \hat{\sigma}_B^2(1-\delta), \ \hat{m}_{3,P} \leq \hat{m}_{3,B}(1+\text{sign}(\hat{m}_{3,P})\delta), \\
&\max_{\delta\in[0,1],w\in\mathcal{W}} \ \delta \quad \text{subject to} \ \ \hat{\mu}_P \geq \hat{\mu}_B, \ \hat{\sigma}_P^2 \leq \hat{\sigma}_B^2(1-\delta), \ \hat{m}_{4,P} \leq \hat{m}_{4,B}(1-\delta).
\end{aligned} \tag{34}$$

We call these two portfolios MVS and MVK, respectively. Briec et al. (2007) demonstrate that the maximum is always global, the portfolio is efficient in terms of moments, and is indirectly equivalent to maximizing a higher-moment utility function. However, a drawback relative to the MEKL portfolio is that the optimal portfolio coincides with the benchmark portfolio w_B if w_B is MVE, and thus, one cannot start from such a natural benchmark. Therefore, as in Boudt et al. (2020), we take the equally weighted portfolio as a benchmark: $w_B = \iota/n$.

Third, the IC-variance-parity (ICVP) portfolio of Lassance et al. (2022) that diversifies the portfolio-return variance across $k \leq n$ independent components. We consider the ICVP portfolio because it is conceptually related to our MEKL portfolio when targeting a Gaussian distribution. Indeed, as explained by Lassance et al. (2022), diversifying variance across independent components

shrinks higher moments toward those of the Gaussian by the central limit theorem. However, contrary to the MEKL portfolio, there is no control over which mean and variance is targeted by the ICVP portfolio. Mathematically, the ICVP portfolio is

$$w_{ICVP} = \frac{\boldsymbol{V}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{R}^{\dagger\prime}\iota_k^{MV}}{\iota'\boldsymbol{V}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{R}^{\dagger\prime}\iota_k^{MV}}, \quad \iota_k^{MV} = \text{sign}(\boldsymbol{R}^{\dagger}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{V}'\iota), \tag{35}$$

where $\boldsymbol{V}$ and $\boldsymbol{\Lambda}$ are the eigenvectors and eigenvalues matrices corresponding to the first k principal components, and $\boldsymbol{R}^{\dagger}$ is the rotation matrix that transforms the principal components into independent components. As in Lassance et al. (2022), we fix k via the method of Velicer (1976).

4.6.3 Out-of-sample returns

For each portfolio strategy, we construct a time series of out-of-sample returns using a rolling-window procedure. At a given time t, we estimate the different portfolio strategies using the past five years of daily data, which corresponds to a sample size $h = 1260$. We keep these portfolio weights constant for the next quarter, thus rebalancing the portfolio every day to account for asset growth, and compute the corresponding out-of-sample daily portfolio returns. After three months, we roll forward the estimation window by three months and repeat this procedure until the end of the sample. We finally use all returns to evaluate performance with the criteria in Section 4.5.[12]. We do not adjust the returns to transaction costs for consistency with the theoretical objective function, but we report the results net of transaction costs in Section IA.5.1 of the internet appendix and find the conclusions are robust.

4.7 Discussion of out-of-sample results

The out-of-sample performance of the considered portfolio strategies is reported in Table 1. First, we observe that in order to better match the target distribution in terms of higher moments, the MEKL portfolio needs to accept a volatility that is close to but generally larger than that of the reference portfolio. This is consistent with hypothesis 4 in Section 4.3 because the KWZ portfolio is close to the GMV portfolio.[13] Concerning the mean return, that of the MEKL portfolio can be larger or smaller than that of the reference portfolio depending on the cases. The difference in mean return relative to the reference portfolio tends to be more pronounced than the difference in volatility, in line with hypothesis 3 in Section 4.3.

[12]Concerning the EKL divergence criteria, we consider the GMV portfolio as a reference portfolio for the four higher-moment portfolios from the literature in Section 4.6.2.

[13]Indeed, we use a rather large risk-aversion coefficient of 10 and the average shrinkage intensity of the sample mean-variance portfolio toward the sample GMV portfolio in the method of Kan et al. (2022) is 83%, 77%, 89%, and 97% for the 25SBTM, 25OPINV, 30IND, and 100CRSP datasets, respectively.

Table 1: Out-of-sample performance of portfolio strategies (continued on next page)

	$252\hat{\mu}_P$	$\sqrt{252}\hat{\sigma}_P$	$\hat{\zeta}_P$	$\hat{\kappa}_P$	EKL (Gauss)	EKL (SLN)	EKL (GN)	MVaR
(a) 25SBTM								
GMV reference portfolio								
GMV	0.186	0.117	-0.855	19.40	0.318	0.392	0.508	5.23%
MEKL (Gaussian)	0.198	0.126	-0.429	12.71	0.251	0.304	0.422	4.31%***
MEKL (SLN)	0.212	0.129	-0.478	12.07	0.241	0.302	0.419	4.31%**
MEKL (GN)	0.200	0.126	-0.523	12.59	0.250	0.310	0.422	4.34%***
KWZ reference portfolio								
KWZ	0.370	0.197	-0.128	4.07	0.196	0.212	0.309	4.02%
MEKL (Gaussian)	0.279	0.209	-0.055	4.22	0.190	0.208	0.323	4.30%
MEKL (SLN)	0.293	0.207	-0.008	3.94	0.181	0.192	0.303	4.13%
MEKL (GN)	0.336	0.201	0.016	4.14	0.182	0.189	0.297	4.03%
Higher-moment portfolios from the literature								
CRRA	0.186	0.117	-0.854	19.12	0.316	0.389	0.504	5.17%
MVS	0.181	0.120	-0.896	17.23	0.320	0.400	0.514	4.98%
MVK	0.187	0.116	-0.878	18.25	0.314	0.386	0.494	4.99%
ICVP	0.208	0.149	-0.647	11.93	0.258	0.383	0.555	5.01%

Second, the MEKL portfolio nearly systematically delivers more desirable higher moments than those of the reference portfolio. The improvement in kurtosis is more substantial than in skewness, which can be a result of the weight of the moments in the EKL divergence, but also because skewness suffers from larger sampling errors than kurtosis (Martellini and Ziemann, 2010). Looking at the first four moments together, the MEKL portfolio provides a better tradeoff from a downside-risk point of view because it delivers a smaller MVaR in nearly all cases, and the improvement is often statistically significant. These results are consistent with hypothesis 2 in Section 4.3.

Third, the objective of minimizing the EKL divergence is better achieved out of sample with the MEKL portfolio: it achieves a smaller EKL divergence than the reference portfolio *in all cases*. This is notable because in-sample efficiency does not automatically translate to better out-of-sample performance, particularly given the large estimation errors in higher moments. This result is consistent with hypothesis 1 in Section 4.3 and suggests that the estimation method proposed in Section 4.2 is robust and delivers good performance even for a larger dataset of 100 stocks.

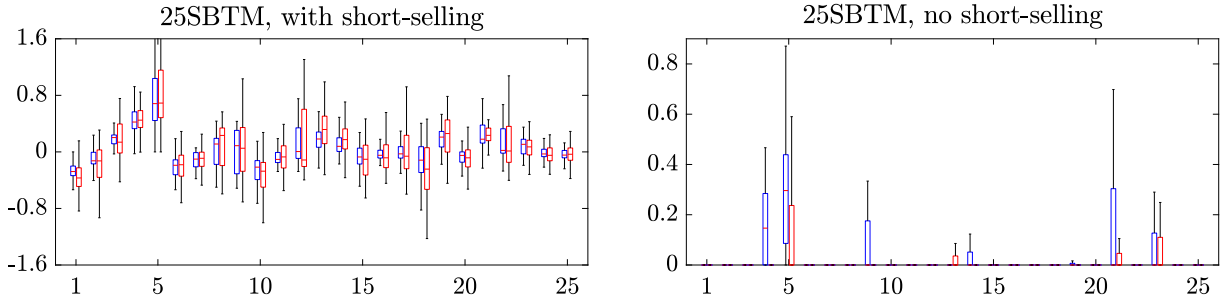
Fourth, we discuss the out-of-sample performance of the benchmark higher-moment portfolios. The CRRA portfolio barely improves higher moments relative to GMV. This is consistent with Ebert and Karehnke (2021) who find that the impact of higher moments in expected utility is limited and at odds with prominent behavioral theories of choice under risk. In contrast, as discussed above, our approach significantly improves higher moments. The MVS and MVK portfolios deliver a modest gain in higher moments relative to the GMV portfolio, for a correspondingly small increase in variance, and thus do not really improve the MVaR. Finally, the shrinkage of higher moments in the ICVP portfolio is achieved with great success because it significantly improves skewness and kurtosis relative to GMV, often more considerably so than with the MEKL portfolio. However, a

Table 1: Out-of-sample performance of portfolio strategies (continued)

	$252\hat{\mu}_P$	$\sqrt{252}\hat{\sigma}_P$	$\hat{\zeta}_P$	$\hat{\kappa}_P$	EKL (Gauss)	EKL (SLN)	EKL (GN)	MVaR
(b) 25OPINV								
GMV reference portfolio								
GMV	0.135	0.154	-0.855	27.08	0.298	0.442	0.586	8.66%
MEKL (Gauss)	0.139	0.167	-0.634	18.84	0.250	0.391	0.533	7.37%***
MEKL (SLN)	0.129	0.172	-0.710	17.64	0.244	0.406	0.534	7.31%***
MEKL (GN)	0.135	0.169	-0.601	18.26	0.253	0.395	0.545	7.31%***
KWZ reference portfolio								
KWZ	0.162	0.168	-0.664	22.12	0.269	0.359	0.474	8.22%
MEKL (Gauss)	0.155	0.176	-0.320	10.37	0.221	0.268	0.382	5.42%
MEKL (SLN)	0.136	0.181	-0.636	14.31	0.218	0.310	0.404	6.75%**
MEKL (GN)	0.147	0.181	-0.730	21.50	0.233	0.358	0.462	8.70%
Higher-moment portfolios from the literature								
CRRA	0.134	0.154	-0.853	26.78	0.294	0.438	0.580	8.60%
MVS	0.138	0.158	-0.713	25.33	0.305	0.446	0.611	8.46%
MVK	0.135	0.154	-0.848	26.83	0.299	0.443	0.588	8.63%
ICVP	0.180	0.206	-0.550	13.56	0.323	0.564	0.634	7.45%
(c) 30IND								
GMV reference portfolio								
GMV	0.103	0.124	-0.989	24.60	0.288	0.421	0.551	6.57%
MEKL (Gauss)	0.087	0.137	-1.016	17.47	0.241	0.406	0.507	5.80%***
MEKL (SLN)	0.092	0.140	-0.960	17.19	0.236	0.407	0.507	5.87%**
MEKL (GN)	0.087	0.139	-1.043	17.34	0.243	0.418	0.521	5.85%*
KWZ reference portfolio								
KWZ	0.124	0.154	-0.804	17.17	0.276	0.399	0.529	6.44%
MEKL (Gauss)	0.085	0.154	-0.793	11.19	0.232	0.333	0.426	5.08%***
MEKL (SLN)	0.075	0.153	-0.914	14.05	0.223	0.343	0.507	5.72%
MEKL (GN)	0.082	0.153	-0.811	11.98	0.226	0.330	0.419	5.25%*
Higher-moment portfolios from the literature								
CRRA	0.102	0.124	-1.010	24.43	0.285	0.418	0.544	6.52%
MVS	0.106	0.132	-0.731	24.03	0.292	0.432	0.603	6.82%
MVK	0.111	0.128	-1.017	23.02	0.276	0.420	0.545	6.46%
ICVP	0.111	0.167	-0.618	10.68	0.262	0.466	0.676	5.34%
(d) 100CRSP								
GMV reference portfolio								
GMV	0.156	0.096	-0.126	23.73	0.442	0.449	0.559	4.73%
MEKL (Gauss)	0.156	0.101	-0.301	17.12	0.359	0.370	0.463	4.07%***
MEKL (SLN)	0.153	0.106	-0.012	17.11	0.346	0.353	0.460	4.19%*
MEKL (GN)	0.158	0.101	-0.284	14.39	0.347	0.356	0.444	3.68%***
KWZ reference portfolio								
KWZ	0.160	0.098	-0.163	22.93	0.648	0.634	0.748	4.73%
MEKL (Gauss)	0.156	0.101	-0.327	15.33	0.569	0.556	0.660	3.84%**
MEKL (SLN)	0.158	0.108	0.000	16.37	0.542	0.526	0.634	4.10%**
MEKL (GN)	0.156	0.103	-0.253	14.91	0.557	0.543	0.647	3.81%***
Higher-moment portfolios from the literature								
CRRA	0.157	0.096	-0.125	23.64	0.440	0.447	0.557	4.72%
MVS	0.063	0.426	-0.918	37.88	3.736	15.28	17.53	30.95%
MVK	0.173	0.430	1.502	56.13	3.478	14.53	16.70	36.51%
ICVP	0.170	0.118	-0.119	15.94	0.286	0.303	0.416	4.48%

Note. This table reports the out-of-sample performance of the portfolio strategies considered in Sections 4.6.1 and 4.6.2 for the four datasets of daily returns of Section 4.4. We consider two reference portfolios: the sample global-minimum-variance (GMV) portfolio and the combination of the sample GMV and sample mean-variance portfolios that maximizes the expected out-of-sample utility according to Kan et al. (2022) (KWZ portfolio), using a risk-aversion coefficient $\lambda = 10$. We then consider three target-return distributions whose mean and variance match those of the two reference portfolios: a Gaussian distribution, a shifted-log-normal (SLN) distribution with a skewness of 0.5, and a generalized-normal (GN) distribution with an excess kurtosis of -0.5. The minimum-EKL (MEKL) portfolios minimize the extended Kullback-Leibler divergence with respect to these target distributions. As performance criteria, we report, in order, the mean, standard deviation, skewness, excess kurtosis, EKL divergence with respect to the target distributions (we take GMV as a reference portfolio for the higher-moment portfolios from the literature), and the 1% modified value-at-risk; see Section 4.5 for more details. The stars $\star$, $\star\star$, and $\star\star\star$ indicate that the MEKL portfolios have a smaller MVaR than the reference portfolios at a confidence level of 10%, 5%, and 1%.

Figure 3: Boxplots of portfolio weights for the MEKL and GMV portfolios



Note. This figure depicts boxplots of portfolio weights for the global-minimum-variance (GMV) portfolio and the minimum-EKL (MEKL) portfolio that targets a Gaussian distribution whose first two moments are matched to those of GMV. The figure is constructed from daily returns on the 25 portfolios of US stocks sorted on size and book-to-market spanning January 1980 to March 2021. The left plot is the case with short-selling, and the right plot without short-selling. In each plot, we depict 25 pairs of boxplots, one per asset. For each pair, the left boxplot (in blue) is for the GMV portfolio, and the right one (in red) is for the MEKL portfolio.

drawback of the ICVP portfolio is that it also delivers a significantly larger volatility than that of GMV, which translates into a larger EKL divergence than that obtained with the MEKL portfolio. Thus, it is more desirable to explicitly target a Gaussian distribution by minimizing the EKL divergence than implicitly targeting it via factor diversification.

Overall, the results suggest that the target-distribution approach is worth considering as an alternative to mean-variance and higher-moment portfolio strategies proposed in the literature.

4.8 Additional experiments

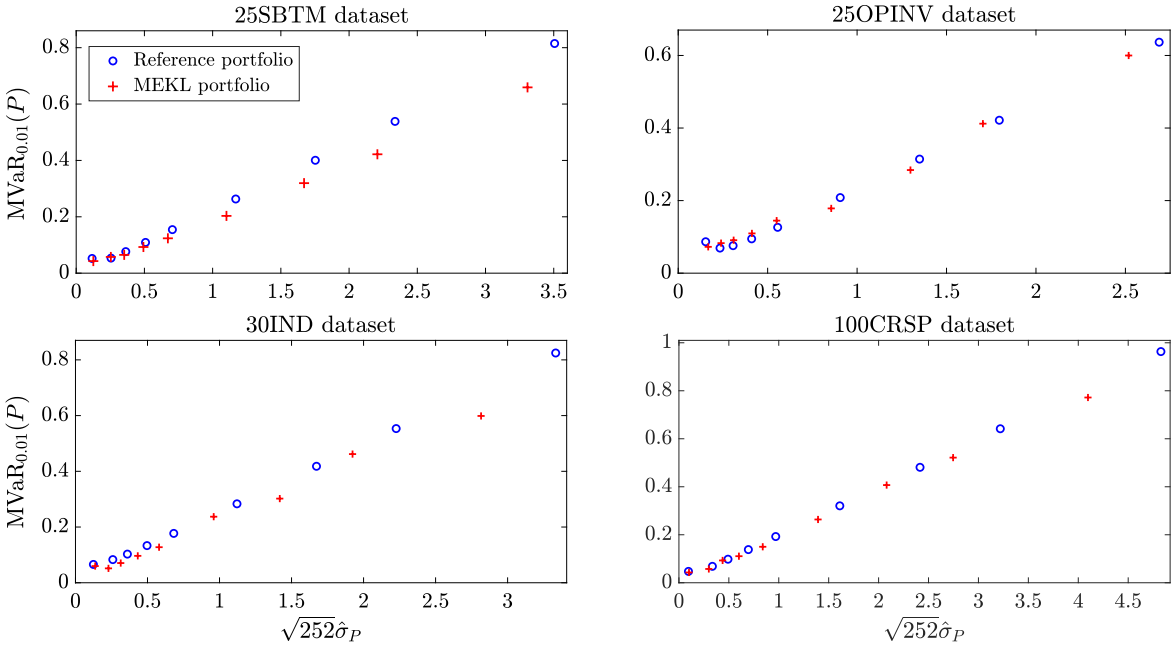
In Figure 3, we depict boxplots of portfolio weights for the MEKL and reference portfolios to visualize how they differ. We focus on the 25SBTM dataset, the GMV reference portfolio, and the Gaussian target distribution.[14] We consider both the case in which short-selling is allowed, and in which it is not allowed as in Section IA.5.3 of the internet appendix. With short-selling, the boxplots are located around the same position but they are wider for the MEKL portfolio, which is due to the entropy term in the objective function (19) that shrink skewness and excess kurtosis toward zero. Without short-selling, the MEKL portfolio tends to be more concentrated: it has more boxplots that stay at zero.[15] This result is consistent with other papers that show that more diversification in a portfolio can actually deteriorate higher moments; see Conine and Tamarkin (1981) and Desmoulins-Lebeault and Kharoubi-Rakotomalala (2012).

In Figure 4, we compare the out-of-sample volatility and MVaR delivered by the MEKL portfolio and several MVE reference portfolios obtained by taking a risk-aversion coefficient $\lambda \in$

[14]The conclusions are robust to considering the other datasets, the KWZ reference portfolio, and the SLN and GN target distributions.

[15]More precisely, the average number of zero weights across all estimation windows as a percentage of the number of assets is 37% in the GMV portfolio and 41% in the MEKL portfolio.

Figure 4: Volatility versus value-at-risk of several MEKL and reference portfolios



Note. This figure depicts the volatility (x-axis) and 1% modified value-at-risk (y-axis) of several sample mean-variance-efficient reference portfolios (blue circles) and minimum-EKL (MEKL) portfolios (red crosses). The reference portfolios are obtained for a risk-aversion coefficient $\lambda \in \{1, 1.5, 2, 3, 5, 7, 10, 15, \infty\}$ and the MEKL portfolios target a Gaussian distribution whose mean and variance match those of the reference portfolio. We consider the four datasets described in Section 4.4.

$\{1, 1.5, 2, 3, 5, 7, 10, 15, \infty\}$. We only consider the Gaussian target distribution for conciseness; the results are similar for the SLN and GN targets. A notable result is that for all reference portfolios except the GMV portfolio ($\lambda = \infty$), the MEKL portfolio delivers a smaller volatility, which is consistent with hypothesis 4 in Section 4.3. Moreover, in all cases except for some values of λ in the 25OPINV dataset, the MEKL portfolio delivers a smaller MVaR because it targets a distribution with zero skewness and excess kurtosis. When the reference portfolio is the GMV portfolio, the MEKL portfolio must necessarily accept a larger volatility as already observed in Table 1, but it nonetheless continues to achieve a smaller MVaR. These results show that the risk benefits of the MEKL portfolio approach appear all along the MVE frontier.

In Table 2, we consider a different way of specifying the SLN target distribution in which the target mean return is fixed instead of varying over time with the reference portfolio. Specifically, we consider an SLN target with an annualized mean of 10%, the minimum variance at that specific mean, and a skewness of either 0.5, 1, or 2. In Table 2, we compare the resulting MEKL portfolio with two other portfolios: 1) the sample mean-variance portfolio with an annualized mean of 10% (called MVE), 2) similar to Athayde and Flores (2004), the portfolio minimizing the variance under the constraint that the annualized mean return is 10% and the skewness is 0.5, 1, or 2 (called CST). For conciseness, we only report the results without transaction costs and with short-selling. The

Table 2: Out-of-sample performance for a fixed target mean return

			25SBTM	25OPINV	30IND	100CRSP
	MVE	$252\hat{\mu}_P$	0.137	0.120	0.100	0.146
		$\sqrt{252}\hat{\sigma}_P$	0.122	0.155	0.125	0.096
		$\hat{\zeta}_P$	-0.873	-0.759	-0.885	-0.070
$\zeta_T = 0.5$	MEKL	$252\hat{\mu}_P$	0.186	0.129	0.091	0.154
		$\sqrt{252}\hat{\sigma}_P$	0.136	0.175	0.141	0.107
		$\hat{\zeta}_P$	-0.472	-0.779	-1.021	-0.054
	CST	$252\hat{\mu}_P$	0.150	0.118	0.133	0.155
		$\sqrt{252}\hat{\sigma}_P$	0.166	0.173	0.137	0.111
		$\hat{\zeta}_P$	-0.809	-0.733	-0.652	-0.548
$\zeta_T = 1$	MEKL	$252\hat{\mu}_P$	0.208	0.121	0.095	0.151
		$\sqrt{252}\hat{\sigma}_P$	0.141	0.182	0.149	0.122
		$\hat{\zeta}_P$	-0.386	-0.570	-1.164	-0.083
	CST	$252\hat{\mu}_P$	0.179	0.101	0.281	0.142
		$\sqrt{252}\hat{\sigma}_P$	0.297	0.286	0.267	0.105
		$\hat{\zeta}_P$	-0.552	-0.218	2.129	0.228
$\zeta_T = 2$	MEKL	$252\hat{\mu}_P$	0.233	0.128	0.096	0.157
		$\sqrt{252}\hat{\sigma}_P$	0.144	0.184	0.150	0.128
		$\hat{\zeta}_P$	-0.368	-0.377	-1.201	-0.235
	CST	$252\hat{\mu}_P$	0.247	0.182	0.135	0.168
		$\sqrt{252}\hat{\sigma}_P$	0.249	0.313	0.439	0.156
		$\hat{\zeta}_P$	0.261	0.326	-3.300	1.508

Note. This table reports the out-of-sample performance of three portfolio strategies for the four datasets of daily returns of Section 4.4; see the analysis of Section 4.8. MVE is the sample mean-variance-efficient portfolio with an annualized mean return of 10%. MEKL is the portfolio that minimizes the EKL divergence with respect to the shifted-log-normal distribution with an annualized mean return of 10%, the minimum variance at that mean return, and a skewness $\zeta_T = 0.5$, 1, or 2. CST is the portfolio with minimum return variance under the constraint that the annualized mean return is equal to 10% and the skewness is equal to ζ_T. As performance criteria, we report, in order, the mean, standard deviation, and skewness.

table shows that except for the 30IND dataset, the MEKL portfolio improves upon the skewness of the MVE portfolio. It achieves this by going up-right in the mean-variance space since the MEKL portfolio has generally a larger mean and variance than MVE. The CST portfolio is much more restrictive because the target mean return and skewness are imposed as binding constraints. As a result, CST is very risky because it delivers a much larger volatility than MEKL and MVE, and it does not achieve a robust performance because its skewness is quite variable across the different cases and not always larger than that of MEKL and MVE, despite the larger volatility it accepts.

Finally, in Table 3 we consider the application we mention in Section 2.3: targeting the return distribution of a primary portfolio by investing in a secondary portfolio. The primary portfolio is a mean-variance portfolio that permits selling short, and we target its return distribution via the MEKL portfolio that prevents short sales. This is done in three steps: 1) compute the GMV and KWZ mean-variance portfolios, 2) take as target an SLN distribution whose parameters are calibrated from the first three moments of GMV and KWZ, 3) compute the MEKL portfolio under short-sale constraints. We compare the out-of-sample moments of GMV and KWZ that can

Table 3: Targeting the distribution of mean-variance portfolios under short-sale constraints

		$252\hat{\mu}_P$	$\sqrt{252}\hat{\sigma}_P$	$\hat{\zeta}_P$	$\hat{\kappa}_P$	EKL
25SBTM	GMV	0.186	0.117	-0.855	19.40	0.401
	MEKLSS	0.130	0.156	-0.714	18.46	0.597
	GMVSS	0.141	0.152	-0.719	19.47	0.602
	KWZ	0.370	0.197	-0.128	4.07	0.198
	MEKLSS	0.152	0.187	-0.888	15.82	0.292
	KWZSS	0.142	0.155	-0.824	18.72	0.432
25OPINV	GMV	0.135	0.154	-0.855	27.08	0.469
	MEKLSS	0.135	0.168	-0.756	22.38	0.519
	GMVSS	0.133	0.163	-0.829	22.58	0.522
	KWZ	0.162	0.168	-0.664	22.12	0.272
	MEKLSS	0.139	0.173	-0.849	22.61	0.311
	KWZSS	0.136	0.164	-0.898	23.37	0.342
30IND	GMV	0.103	0.124	-0.989	24.60	0.389
	MEKLSS	0.117	0.138	-0.900	24.12	0.440
	GMVSS	0.114	0.133	-0.825	27.52	0.452
	KWZ	0.124	0.154	-0.804	17.17	0.315
	MEKLSS	0.123	0.146	-0.712	19.53	0.303
	KWZSS	0.119	0.138	-0.917	27.42	0.364
100CRSP	GMV	0.156	0.096	-0.126	23.73	0.448
	MEKLSS	0.177	0.110	-0.193	10.91	0.354
	GMVSS	0.179	0.102	-0.225	15.89	0.436
	KWZ	0.160	0.098	-0.163	22.93	0.642
	MEKLSS	0.178	0.109	-0.346	10.35	0.542
	KWZSS	0.180	0.102	-0.268	15.88	0.637

Note. This table reports the out-of-sample performance of six portfolio strategies for the four datasets of daily returns of Section 4.4; see the analysis of Section 4.8. GMV and KWZ are the two mean-variance-efficient reference portfolios described in Section 4.6.1 that can short sell. MEKLSS is the short-sale-constrained portfolio minimizing the EKL divergence with respect to a shifted-log-normal (SLN) distribution whose parameters are calibrated from the first three moments of either GMV or KWZ. GMVSS and KWZSS are the short-sale-constrained counterparts to GMV and KWZ. As performance criteria, we report, in order, the mean, standard deviation, skewness, excess kurtosis, and the EKL divergence with respect to the SLN distribution whose parameters are calibrated from the GMV and KWZ portfolios computed with short-selling over the out-of-sample period.

short sell, GMV and KWZ that cannot short sell, and the MEKL portfolio found as above. We also report the out-of-sample EKL divergence with respect to the SLN distribution whose parameters are calibrated from GMV and KWZ computed with short-selling over the out-of-sample period. Table 3 shows that the short-sale-constrained GMV and KWZ and the MEKL portfolios have similar first two moments, because they have a similar objective. However, by targeting an SLN distribution calibrated to the unattainable GMV and KWZ, MEKL also controls for higher moments and generally achieves better skewness and kurtosis than the short-sale-constrained GMV and KWZ.

5 Conclusion

We capture investors' preferences in a novel framework that consists in optimizing the full distribution of returns: given a target-return distribution specified by the investor, the optimal portfolio has

a return density as close as possible to the target-return density. We derive theoretical properties of the proposed framework and illustrate its benefits for a specific application: improving the higher moments of MVE portfolios. Specifically, we consider two classes of target distribution that allow for different skewness and kurtosis, while selecting the mean and variance on the MVE frontier. We show in particular that unless asset returns are Gaussian, the optimal portfolio departs from the efficient frontier to better fit the higher moments embedded in the target distribution. Empirically, the proposed approach trades off between mean-variance efficiency and improved higher moments, has better properties than other higher-moment approaches in the portfolio literature, and is robust to transaction costs, short-sale constraints, and alternative specifications of the target.

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. (2010). Active portfolio management with benchmarking: A frontier based on alpha. *Journal of Banking and Finance*, 34(9), 2185–2197.

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

Ao, M., Yingying, L., & Xinghua, Z. (2019). Approaching mean-variance efficiency for large portfolios. *The Review of Financial Studies*, 32(7), 2890–2919.

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.

Azzalini, A. (1985). A class of distributions which includes the Normal ones. *Scandinavian Journal of Statistics*, 12(2), 171–178.

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

Bera, A., & Park, S. (2008). Optimal portfolio diversification using maximum entropy principle. *Econometric Reviews*, 27(4-6), 484–512.

Bernard, C., Boyle, P., & Vanduffel, S. (2014). Explicit representation of cost-efficient strategies. *Finance*, 35(2), 5–55.

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.

Boudt, K., Peterson, B., & Croux, C. (2008). Estimation and decomposition of downside risk for

portfolios with non-normal returns. *Journal of Risk*, 11(2), 79–103.

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.

Calafiore, G. (2007). Ambiguous risk measures and optimal robust portfolio. *SIAM Journal on Optimization* 18(3), 853–877.

Chalabi, Y., & Würtz, D. (2012). Portfolio optimization based on divergence measures. MPRA working paper 43332.

Chamakh, L. (2021). Quantifying uncertainty in asset management: Kernel methods and statistical fluctuations. PhD Thesis, Institut Polytechnique de Paris.

Clark, B., Edirisinghe, C., & Simaan, M. (2022). Estimation risk and the implicit value of index-tracking. *Quantitative Finance*, 22(2), 303–319.

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

Conine, T., & Tamarkin, M. (1981). On diversification given asymmetry in returns. *The Journal of Finance*, 36(5), 1143–1155.

Cover, T., & Thomas, J. (2006). *Elements of Information Theory* (2nd ed.). John Wiley & Sons, New Jersey.

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., & Nogales, F. (2009). Portfolio selection with robust estimation. *Operations Research*, 57(3), 560–577.

De Nard, G., Ledoit, O., & Wolf, M. (2019). Factor models for portfolio selection in large dimensions: The good, the better and the ugly. *Journal of Financial Econometrics*, 1–22.

Desmoulins-Lebeault, F., & Kharoubi-Rakotomalala, C. (2012). Non-Gaussian diversification: When size matters. *Journal of Banking and Finance*, 36(7), 1987–1996.

Ebert, S., & Karehnke, P. (2021). Skewness preferences in choice under risk. SSRN working paper.

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.

Gaivoronski, A., Krylov, S., & van der Wijst, N. (2005). Optimal portfolio selection and dynamic benchmark tracking. *European Journal of Operational Research*, 163(1), 115–131.

Galton, F. (1879). The geometric mean in vital and social statistics. *Proceedings of the Royal Society*

of London, 29, 365–367.

Glasserman, P., & Xu, X. (2013). Robust risk measurement and model risk. *Quantitative Finance*, 14(1), 29–58.

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

Haugh, M., & Lo, A. (2001). Asset allocation and derivatives. *Quantitative Finance*, 1(1), 45–72.

Hirschberger, M., Steuer, R., Utz, S., Wimmer, M., & Qi, Y. (2013). Computing the nondominated surface in tri-criterion portfolio selection. *Operations Research*, 61(1), 169–183.

Hu, Z., & Hong, L. (2013). Kullback-Leibler divergence constrained distributionally robust optimization. Technical report, The Hong Kong University of Science and Technology.

Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent Component Analysis*. John Wiley & Sons, New York.

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

Jiang, L., Wu, K., & Zhou, G. (2018). Asymmetry in stock comovements: An entropy approach. *Journal of Financial and Quantitative Analysis*, 53(4), 1479–1507.

Johnson, N., Kotz, S., & Balakrishnan, N. (1994). *Continuous Univariate Distributions Volume 1* (2nd ed.). John Wiley & Sons, New York.

Kadan, O., & Liu, F. (2014). Performance evaluation with high moments and disaster risk. *Journal of Financial Economics*, 113(1), 131–155.

Kan, R., Wang, X., & Zhou, G. (2022). Optimal portfolio choice with estimation risk: No risk-free asset case. *Management Science*, 68(3), 2047–2068.

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

Kaplan, P., & Knowles, J. (2004). Kappa: A generalized downside risk-adjusted performance measure. *Journal of Performance Measurement*, 8, 42–54.

Kelly, B., & Jiang, H. (2014). Tail risk and asset prices. *The Review of Financial Studies*, 27(10), 2841–2871.

Kendall, M., & Stuart, A. (1963). *The Advanced Theory of Statistics* (2nd ed.). Charles Griffin, London.

Kim, W., Fabozzi, F., Cheridito, P., & Fox, C. (2014). Controlling portfolio skewness and kurtosis without directly optimizing third and fourth moments. *Economics Letters*, 122, 154–158.

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

Kullback, S., & Leibler, R. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1), 79–86.

Lassance, N., DeMiguel, V., & Vrins, F. (2022). Optimal portfolio diversification via independent component analysis. *Operations Research*, 70(1), 55–72.

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

Lassance, N., & Vrins, F. (2021b). Minimum Rényi entropy portfolios. *Annals of Operations Research*, 299(1), 23–46.

Learned-Miller, E., & Fisher, J. (2003). ICA using spacings estimates of entropy. *Journal of Machine Learning Research*, 4(7-8), 1271–1295.

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. *The Review of Financial Studies*, 30(12), 4349–4388.

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

Markowitz, H. (1952). Portfolio selection. *The 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. *The Review of Financial Studies*, 23(4), 1467–1502.

Merton, R. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8(4), 323–361.

Nadarajah, S. (2005). A generalized normal distribution. *Journal of Applied Statistics*, 32(7), 685–694.

Penev, S., Shevchenko, P., & Wu, W. (2019). The impact of model risk on dynamic portfolio selection under multi-period mean-standard-deviation criterion. *European Journal of Operational Research*, 273(2), 772–784.

Rockafellar, R., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2(3), 21–41.

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379–423 and 623–656.

Vasicek, O. (1976). A test for normality based on sample entropy. *Journal of the Royal Statistical Society Series B*, 38(1), 54–59.

Velicer, W. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41(3), 321–327.

Yuan, P. (1933). On the logarithmic frequency distribution and the semi-logarithmic correlation surface. *The Annals of Mathematical Statistics*, 4(1), 30–74.

Zangari, P. (1996). A VaR methodology for portfolios that include options. *RiskMetrics Monitor First Quarter*, 4–12.

Zhao, L., Chakrabarti, D., & Muthuraman, K. (2019). Portfolio construction by mitigating error amplification: The bounded-noise portfolio. *Operations Research*, 67(4), 965–983.

Internet Appendix to

Portfolio Selection:
A Target-Distribution Approach

This internet appendix to the main paper contains six sections. Section IA.1 contains the theoretical results related to the generalized-normal target distribution. Section IA.2 compares the shifted-log-normal and skew-normal target distributions. Section IA.3 studies the Dirac-delta target distribution. Section IA.4 evaluates the accuracy of the approximation to the EKL divergence with a shifted-log-normal target distribution. Section IA.5 reports robustness tests of our empirical results. Finally, Section IA.6 provides the proofs of all results in the paper and this appendix.

IA.1 The generalized-normal target distribution

In this section, we analyze the generalized-normal (GN) target distribution of Nadarajah (2005), which is a special case of the skewed generalized t distribution of Theodossiou (1998). The GN distribution is tractable under the EKL divergence because it belongs to the exponential family, it is symmetric, and it is defined via three parameters that control for the mean, variance, and fat tails (kurtosis).

IA.1.1 Definition

Definition IA1 (Generalized-normal target)**.** *The target return follows a GN distribution, $T \sim \mathcal{GN}(\alpha, \beta, \gamma)$, if its density function is*

$$f_T(x; \alpha, \beta, \gamma) := \frac{2^{-\frac{\gamma+1}{\gamma}}\gamma}{\beta\Gamma(1/\gamma)} \exp\left(-\frac{1}{2}\left(\frac{|x-\alpha|}{\beta}\right)^{\gamma}\right), \tag{IA1}$$

where $\alpha, x \in \mathbb{R}$ and $\beta, \gamma > 0$. The first four moments of T are

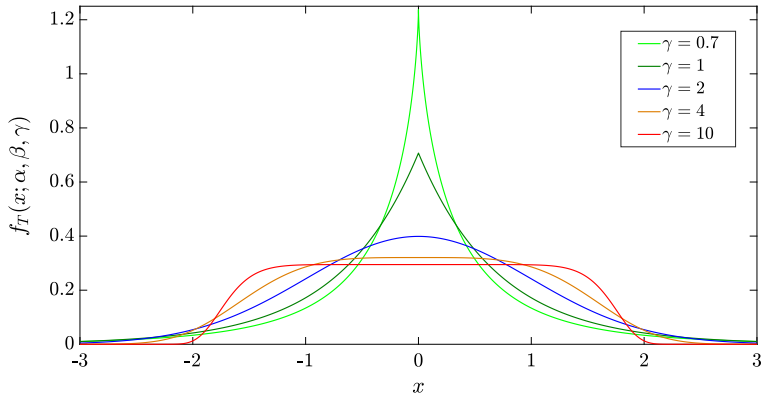
$$\mu_T = \alpha, \quad \sigma_T^2 = \beta^2 4^{1/\gamma}\frac{\Gamma(3/\gamma)}{\Gamma(1/\gamma)}, \quad \zeta_T = 0, \quad \textit{and} \quad \kappa_T = \frac{\Gamma(5/\gamma)\Gamma(1/\gamma)}{\Gamma(3/\gamma)^2} - 3. \tag{IA2}$$

The GN distribution is symmetric and its excess kurtosis is negative for $\gamma > 2$. The minimum excess kurtosis, obtained at the limit $\gamma \to \infty$, is -1.2. The GN distribution retrieves the Laplace distribution for $\gamma = 1$, the Gaussian distribution for $\gamma = 2$, and the Uniform distribution on the interval $[\alpha-\beta, \alpha+\beta]$ for $\gamma \to \infty$. In Figure IA.1, we depict the zero-mean unit-variance GN density for different values of γ, and observe that the smaller the γ the more heavy-tailed the distribution.

IA.1.2 Objective function

In the next proposition, we characterize the objective function in (8) associated with the GN target distribution. Note that $\Omega_T = \mathbb{R}$ and thus $\Omega_P \subseteq \Omega_T$, so that the EKL divergence reduces to the KL divergence.

Figure IA.1: Density of the generalized-normal distribution



Note. This figure depicts the generalized-normal density in (IA1) with mean zero, variance one, and several values of γ that controls the kurtosis via (IA2). We recover the Gaussian density when $\gamma = 2$.

Proposition IA1 (MEKL portfolio with GN target)**.** *Let $T \sim \mathcal{GN}(\alpha, \beta, \gamma)$. Then, the MEKL portfolio $w_{f_T} = w_{\alpha,\beta,\gamma}$ is obtained by maximizing*

$$\mathcal{C}(w; \alpha, \beta, \gamma) = H[P^\star] + \ln \sigma_P - \frac{1}{2\beta^\gamma} \mathbb{E}\left[|P - \alpha|^\gamma\right]. \tag{IA3}$$

In the next example, we illustrate the MEKL portfolio optimizing the objective function (IA3).

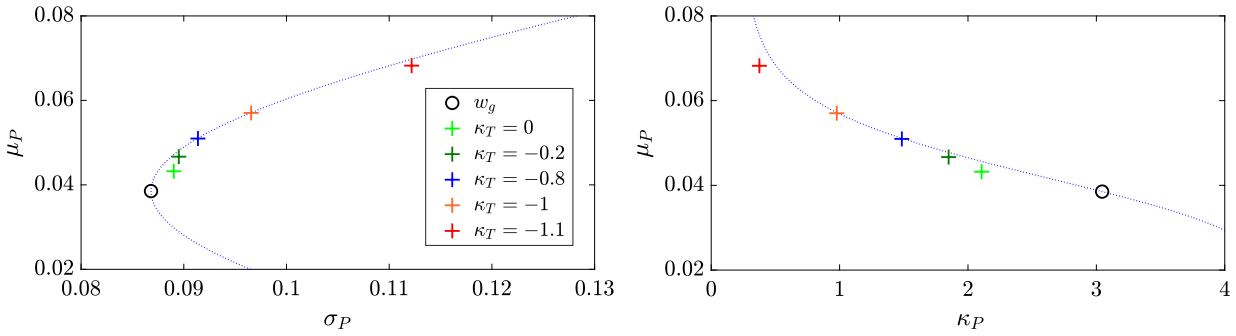
Example IA1. Consider $n = 3$ independent Student t asset returns with the same mean and variance as in Example 1 and excess kurtosis $(8, 4, 0)$. We match the mean and variance of the GN target distribution to those of the GMV portfolio. Moreover, the parameter γ is calibrated to an excess kurtosis κ_T between 0 and -1.1, which improves upon the excess kurtosis of GMV that equals 3.05. We observe that the lower the target kurtosis κ_T, the more the MEKL portfolio invests in the third asset that has the lowest kurtosis: $w_3 = 0.08$ for $\kappa_T = 0$ versus $w_3 = 0.30$ for $\kappa_T = -1.1$. In Figure IA.2, we depict the mean-variance-kurtosis tradeoff for the GMV portfolio and MEKL portfolios. The smaller the κ_T, the more the MEKL portfolio departs from the GMV portfolio so as to improve the kurtosis.

IA.1.3 Gaussian asset returns

We now analyze the properties of the MEKL portfolio when the investor targets a GN distribution and asset returns are Gaussian as in Section 3.3. We first derive the moments of the Gaussian distribution minimizing the EKL divergence with respect to the GN target distribution.

Proposition IA2 (Gaussian distribution with minimum EKL)**.** *The moments of the Gaussian*

Figure IA.2: Optimal portfolio for the generalized-normal target distribution



Note. This figure depicts the results in Example IA1. We consider $n = 3$ independent Student t asset returns with mean $(0.02, 0.07, 0.15)$, volatility $(0.10, 0.20, 0.30)$, and excess kurtosis $(8, 4, 0)$. We consider a generalized-normal target-return distribution whose parameters are matched to the return mean and variance of the GMV portfolio, $w_g = (0.74, 0.18, 0.08)'$, and to an excess kurtosis κ_T between 0 and -1.1. The two plots depict the mean-variance-kurtosis tradeoff of the minimum-EKL portfolios (colored crosses) and GMV portfolio (black circle) as a function of κ_T. The blue dotted curves show the tradeoff for the mean-variance portfolios with mean return $\mu_P \in [0.02, 0.08]$.

distribution minimizing the EKL divergence with respect to the density of $T \sim \mathcal{GN}(\alpha, \beta, \gamma)$ *satisfy*

$$(\mu^\star_{GN}, \sigma^\star_{GN}) = (\mu_T, c_2(\gamma)\sigma_T)\,, \tag{IA4}$$

where

$$c_2(\gamma) := \sqrt{\frac{\Gamma(1/\gamma)}{2\Gamma(3/\gamma)}} \left(\frac{\Gamma\left(\frac{\gamma+1}{2}\right)\gamma}{\pi^{1/2}} \right)^{-1/\gamma} \leq 1, \tag{IA5}$$

with equality if and only if $\gamma = 2$*, corresponding to a Gaussian target.*

Similar to the case with the SLN target distribution, those optimal mean and variance are not attainable when (μ_T, σ_T) are located on or above the MVE frontier. However, we establish in the next proposition that the MEKL portfolio $w_{\alpha,\beta,\gamma}$ remains MVE in such a case, for any GN target distribution.

Proposition IA3 (MEKL portfolio with Gaussian asset returns)**.** *Let* $R \sim \mathcal{N}(\mu, \mathbf{\Sigma})$ *and the target-return mean and variance are located on or above the mean-variance efficient frontier:* $\mu_T \geq \mu_g$ *and* $\sigma_T \leq \sigma_{EF}(\mu_T)$*. Then, the MEKL portfolio* $w_{\alpha,\beta,\gamma}$ *is mean-variance efficient.*

Similar to the case with the SLN target distribution, the proof consists in showing that the EKL divergence has no other stationary point in the mean-variance plane than the global optimum in Proposition IA2.

IA.1.4 Estimation of objective function

To estimate the objective function $\mathcal{C}(w;\alpha,\beta,\gamma)$ in Proposition IA1, we calibrate the parameters (α,β) to $(\hat{\mu}_0,\hat{\sigma}_0)$ in (26),

$$\hat{\alpha}=\hat{\mu}_0 \quad \text{and} \quad \hat{\beta}=\hat{\sigma}_0 2^{-1/\gamma}\sqrt{\frac{\Gamma(1/\gamma)}{\Gamma(3/\gamma)}}, \tag{IA6}$$

where γ is fixed exogenously by the investor, e.g., to match a desired excess kurtosis κ_T in (IA2). Then, the estimated objective function depends on γ and $\hat{w}_0$ as

$$\widehat{\mathcal{C}}(w;\gamma,\hat{w}_0)=\widehat{H}[P]-\frac{1}{2\hat{\beta}^{\gamma}}\widehat{\mathbb{E}}\left[|P-\hat{\alpha}|^{\gamma}\right], \tag{IA7}$$

where $\widehat{H}[P]$ is given by (30) and $\widehat{\mathbb{E}}\left[|P-\hat{\alpha}|^{\gamma}\right]$ is the sample estimate of $\mathbb{E}\left[|P-\hat{\alpha}|^{\gamma}\right]$:

$$\widehat{\mathbb{E}}\left[|P-\hat{\alpha}|^{\gamma}\right]:=\frac{1}{h}\sum_{t=1}^{h}|P_t-\hat{\alpha}|^{\gamma}. \tag{IA8}$$

IA.2 Shifted-log-normal versus skew-normal target distribution

In this section, we compare the SLN target distribution in Section 3.1 with a skewed distribution commonly used in finance: the skew-normal distribution of Azzalini (1985). This target density is given by

$$f_T(x;\alpha,\beta,\gamma)=\frac{2}{\beta}\phi\left(\frac{x-\alpha}{\beta}\right)\Phi\left(\gamma\left(\frac{x-\alpha}{\beta}\right)\right), \tag{IA9}$$

where $x,\alpha,\gamma\in\mathbb{R}$ and $\beta>0$. When $\gamma=0$, we recover the Gaussian distribution, and otherwise the sign of the skewness of f_T is equal to the sign of γ. Relative to the SLN target, the skew-normal distribution has a support set $\Omega_T=\mathbb{R}$, so that the EKL divergence is equal to the plain KL divergence. However, its skewness is limited in the interval $[-0.995,0.995]$ and it is less analytically tractable in our context. Indeed, from Proposition 3, we can show that the MEKL portfolio is obtained by maximizing

$$\mathcal{C}(w;\alpha,\beta,\gamma)=-\frac{1}{2}\left(\frac{\mu_P-\alpha}{\beta}\right)^2+\frac{1}{2}\left(\ln\sigma_P^2-\frac{\sigma_P^2}{\beta^2}\right)+H[P^{\star}]+\mathbb{E}\left[\ln\Phi\left(\gamma\left(\frac{P-\alpha}{\beta}\right)\right)\right]. \tag{IA10}$$

Compared to the Gaussian target objective function in (19), $\mathcal{C}(w;\alpha,\beta,\gamma)$ in (IA10) includes an additional term $\mathbb{E}\left[\ln\Phi\left(\gamma\left(\frac{P-\alpha}{\beta}\right)\right)\right]$ that is mathematically not tractable, even in the sample case in which P is Gaussian.

We now proceed with an experiment showing that relying on the SLN target distribution or the

skew-normal target distribution yields very similar results. We rely on the 25SBTM dataset used in the empirical analysis (see Section 4.4) and we calibrate the parameters of the two target densities so that they have the same first three moments. The first two moments are matched to those of the GMV portfolio, and the skewness is set to $\zeta_T = 0.3$ and 0.7. We make three main observations:

1. The lower bound δ is -8.27% when $\zeta_T = 0.3$ and -3.55% when $\zeta_T = 0.7$, which is 10.03 and 4.36 standard deviations away from the target-return mean, respectively. As a result, the mismatch in support sets between the SLN target and the MEKL portfolio is very small: the probability that the MEKL portfolio return is smaller than δ is 0.011% when $\zeta_T = 0.3$ and 0.42% when $\zeta_T = 0.7$.

2. The left plots of Figure IA.3 show that using the SLN target is actually an excellent proxy for the skew-normal target. Indeed, for both values of ζ_T, the two densities are nearly indistinguishable from one another, despite the SLN density having a bounded support set.

3. The right plots of Figure IA.3 show that the resulting MEKL portfolio weights are consequently very similar, as expected.

This experiment indicates that relying on a target satisfying $\Omega_T \subset \mathbb{R}$ is not a problem *per se* because, in our context, it can be regarded as a tractable version of more complicated distributions with support $\mathbb{R}$. In the SLN and skew-normal cases, the results are essentially the same when matching the first three moments, but the former allows us to derive analytical results, in contrast with the latter. This explains why considering target densities with (semi)-bounded support might be appealing even when the returns are defined on $\mathbb{R}$, and justifies the EKL divergence we introduce in the main body of the paper.

IA.3 The Dirac-delta target distribution

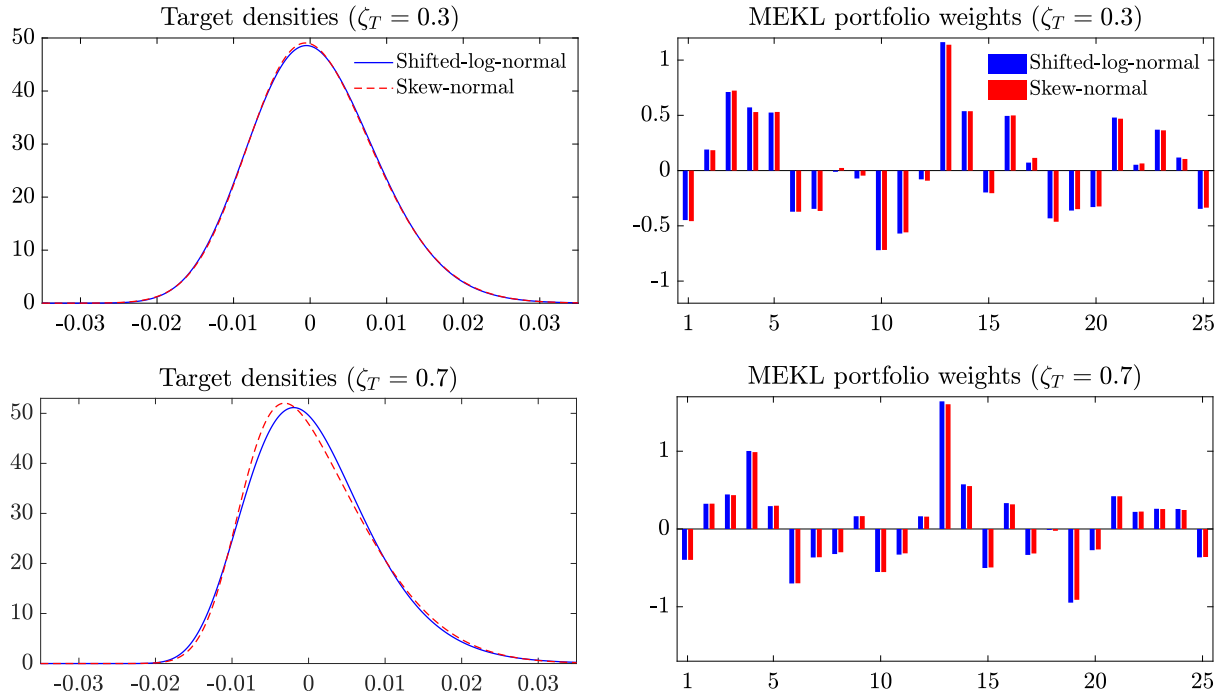
Suppose an investor wishes to target a return α. A first simple way to achieve such an objective is to maximize the probability that the portfolio return P belongs to a small interval around α:

$$\max_{w \in \mathcal{W}} \ \mathbb{P}(\alpha - \epsilon \leq P \leq \alpha + \epsilon), \tag{IA11}$$

where $\epsilon > 0$. We can expect this problem to be related to the mean-variance portfolio that minimizes σ_P^2 under the constraint $\mu_P = \alpha$. The mean-variance portfolio with mean return μ_0 is defined as

$$w_{\mu_0} := w_g + \frac{\mu_g}{\sigma_g^2} \frac{(\mu_0 - \mu_g)}{\psi^2} (w_{msr} - w_g), \tag{IA12}$$

Figure IA.3: Shifted-log-normal versus skew-normal target densities with same first three moments



Note. This figure compares the target densities and minimum-EKL portfolios for the shifted-log-normal and skew-normal target distributions; see Section IA.2. The figure is constructed from daily returns on the 25 portfolios of US stocks sorted on size and book-to-market spanning January 1980 to March 2021. Each target has the same first three moments. The first two moments are matched to the global-minimum-variance portfolio, and the skewness is $\zeta_T = 0.3$ or 0.7. The left plots compare the two target densities, and the right plots compare the resulting 25 minimum-EKL portfolio weights.

where $w_{msr} := (\iota'\boldsymbol{\Sigma}^{-1}\mu)^{-1}\boldsymbol{\Sigma}^{-1}\mu$ is the maximum-Sharpe-ratio portfolio and the other notation is defined in Section 3.3.

Indeed, as we show in the next proposition, the two problems are equivalent when asset returns R are elliptically symmetric. We assume the only constraint in $\mathcal{W}$ is $w'\iota = 1$.

Proposition IA4. *Let the asset returns R be elliptically symmetric, and the mean and covariance matrix of R exist. Then, the portfolio w solving* (IA11) *is the mean-variance portfolio in* (IA12) *with mean return α.*

This result is quite intuitive: when returns are symmetrically distributed, targeting a return with maximum probability is achieved by having the most probability mass in the interval $[\alpha - \epsilon, \alpha + \epsilon]$ and thus having $\mu_P = \alpha$ and σ_P as small as possible.

We now compare this approach with how the objective of targeting a return can be achieved in our framework: by targeting a Dirac-delta distribution located at α. Specifically, the decomposition of the KL divergence in Proposition 3 allows us to study the limit case where the investor targets a distribution with zero variance, $\beta = \sigma_T \to 0^+$, which corresponds to a Dirac-delta density centered

in α, denoted by δ_α. To derive the objective function, notice that, for any $\beta > 0$, maximizing $\mathcal{C}(w;\alpha,\beta)$ in (19) is equivalent to maximizing

$$\breve{\mathcal{C}}(w;\alpha,\beta) := \beta^2\mathcal{C}(w;\alpha,\beta) = -\frac{1}{2}\left((\mu_P-\alpha)^2+\sigma_P^2\right)+\beta^2 H[P]. \tag{IA13}$$

Then, we can deal with the Dirac-delta target by considering the limit of this objective function when $\beta \to 0^+$,

$$\mathcal{C}(w;\alpha) = \mathcal{C}(w;\delta_\alpha) := 2\lim_{\beta\to 0^+}\breve{\mathcal{C}}(w;\alpha,\beta) = -(\mu_P-\alpha)^2-\sigma_P^2, \tag{IA14}$$

leading to the MEKL portfolio $w_{\delta_\alpha} = w_\alpha$. The objective function $\mathcal{C}(w;\alpha)$ corresponds to the squared ℓ_2 distance between (μ_P,σ_P) and $(\mu_T,\sigma_T) = (\alpha,0)$. In the next proposition, we derive the analytical solution for w_α and show that it is located on the MVE frontier if $\alpha \geq \mu_g$.

Proposition IA5 (MEKL portfolio with Dirac-delta target)**.** *The MEKL portfolio w_α maximizing* (IA14) *is a mean-variance portfolio with mean return*

$$w_\alpha'\mu = \mu_g + (\alpha-\mu_g)\frac{\psi^2}{1+\psi^2}. \tag{IA15}$$

Moreover, $w_\alpha = w_g$ if $\alpha = \mu_g$ and is mean-variance efficient if and only if $\alpha \geq \mu_g$.

Proposition IA5 shows that the mean return of the MEKL portfolio w_α is located in between μ_g and α, which is because the objective (IA14) amounts to fit the mean return α but also to minimize the portfolio-return variance σ_P^2. This shows the difference with the maximum-probability formulation: the objective function (IA11) forces to take $\mu_P = \alpha$, so that σ_P^2 will be large if α is large, whereas the objective function (IA14) compromises between fitting α and minimizing σ_P^2.

IA.4 Approximation of EKL divergence for the SLN target

In this section, we evaluate the accuracy of the approximation to the MEKL objective function introduced in Section 3.1. The EKL divergence objective function corresponding to the SLN target is

$$\mathcal{C}(w;\alpha,\beta,\delta) = H[P|P\in\Omega_T] + \mathbb{E}[\ln f_T(P)|P\in\Omega_T] + \ln(1-p), \tag{IA16}$$

where f_T is defined in (9). Throughout this section, we name $\mathbb{E}[\ln f_T(P)|P\in\Omega_T]$ and $H[P|P\in\Omega_T]$ the conditional entropy and the conditional cross-entropy, respectively.

For analytical tractability, we assume that the portfolio return P is Gaussian, $P\sim\mathcal{N}(\mu_P,\sigma_P)$. In this case, $p = \mathbb{P}(P<\delta) = \Phi\left(\frac{\delta-\mu_P}{\sigma_P}\right)$. We begin with the conditional entropy $H[P|P\in\Omega_T]$,

then study the conditional cross-entropy $\mathbb{E}[\ln f_T(P)|P \in \Omega_T]$, explain how the objective function simplifies under the approximation $p = 0$, and finally we compare the different objective functions and corresponding MEKL portfolios via a numerical example. For what follows, it is useful to define the function

$$h(p) := \phi(z_p)/(1-p), \tag{IA17}$$

where ϕ is the standard Gaussian pdf and $z_p := \Phi^{-1}(p)$ is its p-th quantile.

IA.4.1 Conditional entropy

The support set is $\Omega_T = [\delta, \infty)$ and the conditional entropy becomes $H[P|P \in \Omega_T] = H_\delta[P]/(1-p)$ with

$$H_\delta[P] := -\int_\delta^\infty f_P(x) \ln f_P(x) \mathrm{d}x = H[P] + \int_{-\infty}^\delta f_P(x) \ln f_P(x) \mathrm{d}x.$$

Using $f_P(x) = \phi((x-\mu_P)/\sigma_P)/\sigma_P$, the integral simplifies to

$$\begin{aligned} \int_{-\infty}^\delta f_P(x) \ln f_P(x) \mathrm{d}x &= \int_{-\infty}^\delta \frac{1}{\sigma_P} \phi\left(\frac{x-\mu_P}{\sigma_P}\right) \ln\left(\frac{1}{\sigma_P}\phi\left(\frac{x-\mu_P}{\sigma_P}\right)\right) \mathrm{d}x \\ &= \int_{-\infty}^{\frac{\delta-\mu_P}{\sigma_P}} \phi(y) \ln\left(\frac{1}{\sigma_P}\phi(y)\right) \mathrm{d}y \\ &= I(z_p) - p \ln \sigma_P, \end{aligned}$$

where

$$I(a) := \int_{-\infty}^a \phi(x) \ln \phi(x) \mathrm{d}x = \frac{a\phi(a) - \Phi(a)(1+\ln(2\pi))}{2}.$$

Noting that $H[P] = H[P^\star] + \ln \sigma_P$ with $H[P^\star] = \frac{1}{2}\ln 2\pi e$ and $\Phi(z_p) = p$, $H_\delta[P]$ simplifies to

$$H_\delta[P] = H[P] + I(z_p) - p \ln \sigma_P = (1-p)H[P] + \frac{z_p}{2}\phi(z_p).$$

Note that $z_p\phi(z_p) \to 0$ when $p \to 0^+$. Finally, the conditional entropy is equal to

$$H[P|P \geq \delta] = \frac{H_\delta[P]}{1-p} = H[P] + \frac{z_p}{2}h(p). \tag{IA18}$$

IA.4.2 Conditional cross-entropy

We now turn to the conditional cross-entropy term in (IA16), $\mathbb{E}[\ln f_T(P)|P \in \Omega_T]$. We study first the true value and, second, the Taylor-series approximation.

IA.4.2.1 True value

Ignoring the proportionality constant $1/(\beta\sqrt{2\pi})$ in (9), the conditional cross-entropy is equal to

$$\begin{aligned}\mathbb{E}[\ln f_T(P)|P \geq \delta] &= -\mathbb{E}[\ln P_\delta|P_\delta \geq 0] - \frac{1}{2\beta^2}\mathbb{E}[(\ln P_\delta - \alpha)^2|P_\delta \geq 0] \\ &= -\mathbb{E}[\ln P_\delta|P_\delta \geq 0] - \frac{1}{2\beta^2}\left(\mathbb{E}[(\ln P_\delta)^2|P_\delta \geq 0] - 2\alpha\mathbb{E}[\ln P_\delta|P_\delta \geq 0] + \alpha^2\right) \\ &= -\frac{1}{2\beta^2}\left(\mathbb{E}[(\ln P_\delta)^2|P_\delta \geq 0] + 2(\beta^2 - \alpha)\mathbb{E}[\ln P_\delta|P_\delta \geq 0] + \alpha^2\right), \end{aligned} \tag{IA19}$$

where

$$\mathbb{E}[(\ln P_\delta)^k|P_\delta \geq 0] = \frac{1}{1-p}\int_\delta^\infty f_P(x)\,(\ln(x-\delta))^k \mathrm{d}x = \frac{1}{1-p}\int_0^\infty f_P(x+\delta)\,(\ln x)^k \mathrm{d}x. \tag{IA20}$$

This integral does not simplify further for $f_P(x) = \phi((x-\mu_P)/\sigma_P)/\sigma_P$, but can be evaluated numerically.

IA.4.2.2 Taylor-series approximation

Let us write the conditional cross-entropy in (IA19) as

$$\mathbb{E}[\ln f_T(P)|P \geq \delta] = -\mathbb{E}[\ln P_\delta|P_\delta \geq 0] - \frac{1}{2\beta^2}\left(\mathbb{V}[\ln P_\delta|P_\delta \geq 0] + (\mathbb{E}[\ln P_\delta|P_\delta \geq 0] - \alpha)^2\right). \tag{IA21}$$

Define $m(X) := \mathbb{E}[X|X \geq 0]$, $m_k(X) := \mathbb{E}\big[(X - m(X))^k|X \geq 0\big]$, and $\psi_k(X) = m_k(X)/m(X)^k$. Then, a second-order Taylor-series expansion of $\ln x$ yields the approximations

$$\begin{aligned}&\mathbb{E}[\ln P_\delta|P_\delta \geq 0] \approx \ln(m(P_\delta)) - \psi_2(P_\delta)/2, \\ &\mathbb{V}[\ln P_\delta|P_\delta \geq 0] \approx \psi_2(P_\delta) - \psi_3(P_\delta) + \psi_4(P_\delta)/4 - \frac{\psi_2(P_\delta)}{4m(P_\delta)^2}. \end{aligned} \tag{IA22}$$

Moreover, when $P \sim \mathcal{N}(\mu_P, \sigma_P)$, we have from Pender (2015) that

$$\begin{aligned} m(P_\delta) &= (\mu_P - \delta) + \sigma_P h(p), \\ \frac{m_2(P_\delta)}{\sigma_P^2} &= 1 + z_p h(p) - h^2(p), \\ \frac{m_3(P_\delta)}{\sigma_P^3} &= (z_P^2 - 1)h(p) - 3z_p h^2(p) + 2h^3(p), \\ \frac{m_4(P_\delta)}{\sigma_P^4} &= 3 + 12 z_p h^3(p) - 4(z_p^2 - 1)h^2(p) - 3z_p h^2(p) - 6h^4(p) + (z_p^3 - 3z_p)h(p). \end{aligned} \tag{IA23}$$

IA.4.3 Approximation for $p = 0$

The lower bound δ of the target-return support set Ω_T is typically many standard deviations below the target-return mean. Specifically, the quantity $(\mu_T - \delta)/\sigma_T = g(\beta)^{-1/2}$, which decreases with the skewness ζ_T, is large: it varies between 30 and 3.1 for ζ_T between 0.1 and 1. Therefore, as long as (μ_P, σ_P) is not too different from (μ_T, σ_T), the probability $p = \mathbb{P}(P \leq \delta)$ is negligible.

If we plug in $p = 0$ in the objective function (IA16) with the conditional entropy given by (IA18) and the Taylor-series approximation to the conditional cross-entropy given by (IA21)–(IA23), we obtain the approximation to the objective function $\widetilde{\mathcal{C}}(w; \alpha, \beta, \delta)$ in (14), which for P being Gaussian is equal to

$$\widetilde{\mathcal{C}}(w; \alpha, \beta, \delta) = H[P] - \widetilde{\mathbb{E}}[\ln P_\delta | P_\delta \geq 0] - \frac{1}{2\beta^2}\left(\widetilde{\mathbb{V}}[\ln P_\delta | P_\delta \geq 0] + (\widetilde{\mathbb{E}}[\ln P_\delta | P_\delta \geq 0] - \alpha)^2\right), \quad \text{(IA24)}$$

where

$$\widetilde{\mathbb{E}}[\ln P_\delta | P_\delta \geq 0] = \ln(\mu_P - \delta) - \frac{\sigma_P^2}{2(\mu_P - \delta)^2}, \quad \text{(IA25)}$$

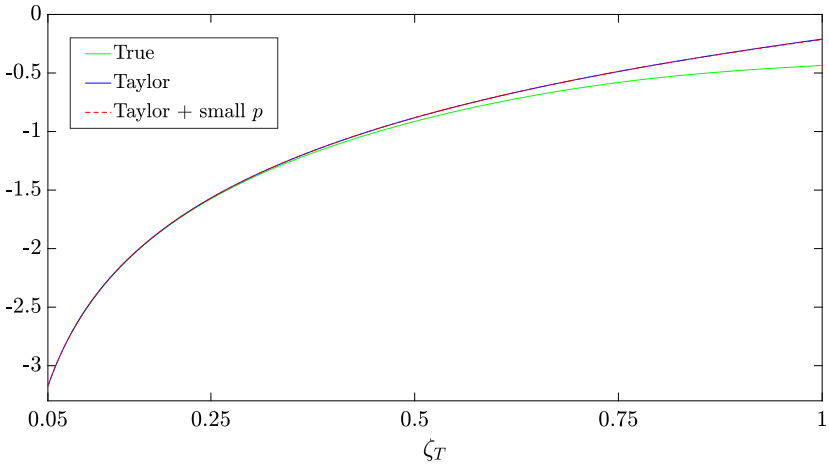
$$\widetilde{\mathbb{V}}[\ln P_\delta | P_\delta \geq 0] = \frac{\sigma_P^2}{(\mu_P - \delta)^2} + \frac{\sigma_P^4}{2(\mu_P - \delta)^4}. \quad \text{(IA26)}$$

IA.4.4 Numerical evaluation of approximation accuracy

In Figure IA.4, we set $(\mu_P, \sigma_P) = (\mu_T, \sigma_T) = (0.1, 0.2)$ and depict, as a function of the target-return skewness ζ_T between 0.05 and 1, the true objective function, the one approximated via a Taylor-series expansion, and the one approximated via a Taylor-series expansion plus assuming $p = 0$, for P being Gaussian as considered above. We find that the Taylor-series approximation is accurate as long as ζ_T is not too large, and that the small-p approximation introduces a very negligible error. Note that when $\zeta_T \to 0^+$, which corresponds to targeting a Gaussian distribution, the three objective functions correspond to one another.

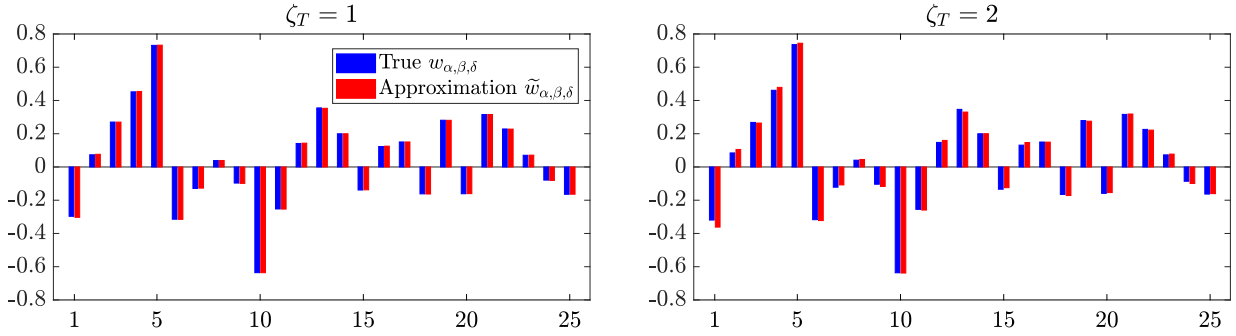
In Figure IA.5, we use the closed-form expressions introduced previously to compare the MEKL portfolio maximizing the true objective function $\mathcal{C}(w; \alpha, \beta, \delta)$ in (13) to the MEKL portfolio maximizing the approximation $\widetilde{\mathcal{C}}(w; \alpha, \beta, \delta)$ in (14). To do so, we calibrate the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ to the 25SBTM dataset, we match the target-return mean and variance to the GMV portfolio, and we take a large target-return skewness $\zeta_T = 1$ and 2. The figure shows that although the approximation of the objective function is less accurate when ζ_T is large as shown above, the resulting approximation to the MEKL portfolio is very accurate.

Figure IA.4: Accuracy of approximation to the objective function for the shifted-log-normal target



Note. This figure depicts the true objective function corresponding to the shifted-log-normal target-return distribution (solid green), its approximation via a Taylor-series expansion (solid blue), and its approximation via a Taylor-series expansion plus assuming that the probability $p = \mathbb{P}(P \leq \delta)$ is small (dotted red). We consider the special case in which the portfolio return P has a Gaussian distribution, for which we derive closed-form expressions for the objective functions in Section IA.4. We fix $(\mu_P, \sigma_P) = (\mu_T, \sigma_T) = (0.1, 0.2)$, and the target-return skewness ζ_T varies between 0.05 and 1.

Figure IA.5: Accuracy of approximation to the MEKL portfolio for the shifted-log-normal target



Note. This figure depicts the portfolio weights maximizing the true objective function for the shifted-log-normal target-return distribution (in blue), corresponding to $\mathcal{C}(w; \alpha, \beta, \delta)$ in (13), compared to those maximizing the approximation of the objective function via a Taylor-series expansion and assuming that the probability $p = \mathbb{P}(P \leq \delta)$ is small (in red), corresponding to $\widetilde{\mathcal{C}}(w; \alpha, \beta, \delta)$ in (14). We consider the special case in which the portfolio return P has a Gaussian distribution, for which we derived closed-form expressions for the two objective functions in Section IA.4. The mean and covariance matrix of asset returns are calibrated to a dataset of daily returns on the 25 portfolios of US stocks sorted on size and book-to-market spanning January 1980 to March 2021. The return mean and variance of the target are matched to those of the global-minimum-variance portfolio, and the target-return skewness equals $\zeta_T = 1$ (left plot) and $\zeta_T = 2$ (right plot).

IA.5 Robustness tests of empirical results

In Section IA.5.1, we report results net of transaction costs. In Section IA.5.2, we target a Laplace distribution. In Section IA.5.3, we impose short-sale constraints.

IA.5.1 Out-of-sample performance net of transaction costs

In Table IA.1, we replicate the results in Table 1 with proportional transaction costs of 10 basis points. This is the cost level used by Ao et al. (2019) and is in line with the average market impact of the trades in the 1998–2011 sample considered by Frazzini et al. (2018). We observe that introducing transaction costs reduces mean returns, but otherwise leaves the other performance criteria nearly intact. Therefore, the conclusions drawn in the paper for the case without transaction costs continue to hold here. We also find that the mean returns of MEKL portfolios are more impacted by transaction costs than those of the reference portfolios because they have a larger turnover due to the optimization of higher moments.

IA.5.2 Targeting a Laplace distribution

In the main results of Table 1, we consider two values of the parameter γ in the GN target, $\gamma = 2.78$ and $\gamma = 2$, which correspond to an excess kurtosis of -0.5 and 0, respectively. However, such values of excess kurtosis can be difficult to attain, and the MEKL portfolio may be better-behaved when targeting a positive excess kurtosis. To evaluate this, we report in Tables IA.2 and IA.3 the out-of-sample performance of the MEKL portfolio when $\gamma = 1$, which corresponds to a Laplace distribution with excess kurtosis $\kappa_T = 3$. Table IA.2 contains the results without transaction costs, and Table IA.3 the results net of transaction costs.

The conclusions are intuitive. First, the Laplace distribution is more closely attainable because the EKL divergence is smaller than with the GN distribution with $\gamma = 2$ or 2.78. Second, because the Laplace distribution has a larger kurtosis than the two other GN targets, the resulting MEKL portfolio has nearly systematically a larger kurtosis. Third, the counterpart of this less intensive penalty on kurtosis is that with the Laplace target the MEKL portfolio attains a volatility closer to that of the reference portfolio, relative to the MEKL portfolios targeting a GN distribution with $\gamma = 2$ or 2.78. Overall, the results are also reassuring because targeting a GN target with $\gamma = 1$, 2, or 2.78 delivers MEKL portfolios that differ in intuitive ways but also that are not excessively different from one another, which points to the robustness of our approach.

IA.5.3 Short-sale constraints

We now replicate the results in Table 1 but imposing short-sale constraints: $w_i \geq 0$ for $i = 1, \ldots, n$. The results without transaction costs are in Table IA.4, and the results net of transaction costs are in Table IA.5. For the MEKL portfolios, we adjust the GMV and KWZ reference portfolios so that they are positively invested as well and thus that their first two moments are attainable.

The conclusions drawn in the case with short-selling are robust to imposing short-sale constraints. In particular, relative to the reference portfolio, the MEKL portfolio has overall a larger mean return and a larger variance, more skewness and less excess kurtosis, and the tradeoff between the four moments translates into a generally smaller MVaR. Moreover, consistent with the underlying objective function, the MEKL portfolio delivers a smaller EKL divergence than the reference portfolio in all cases except one.

These results show that our target-distribution framework continues to be applicable and advantageous in the practically relevant case in which portfolio weights are positive.

Table IA.1: Out-of-sample performance of portfolio strategies net of transaction costs (continued on next page)

	$252\hat{\mu}_P$	$\sqrt{252}\hat{\sigma}_P$	$\hat{\zeta}_P$	$\hat{\kappa}_P$	EKL (Gauss)	EKL (SLN)	EKL (GN)	MVaR
(a) 25SBTM								
GMV reference portfolio								
GMV	0.171	0.117	-0.889	19.43	0.318	0.394	0.508	5.24%
MEKL (Gaussian)	0.174	0.126	-0.469	12.67	0.251	0.307	0.423	4.33%**
MEKL (SLN)	0.187	0.129	-0.535	12.23	0.242	0.307	0.422	4.37%**
MEKL (GN)	0.176	0.126	-0.565	12.56	0.250	0.313	0.423	4.35%***
KWZ reference portfolio								
KWZ	0.332	0.197	-0.144	4.07	0.196	0.213	0.309	4.05%
MEKL (Gaussian)	0.220	0.209	-0.073	4.18	0.191	0.211	0.324	4.34%
MEKL (SLN)	0.234	0.208	-0.039	3.92	0.182	0.197	0.306	4.19%
MEKL (GN)	0.283	0.202	-0.008	4.14	0.183	0.192	0.300	4.08%
Higher-moment portfolios from the literature								
CRRA	0.171	0.117	-0.888	19.14	0.316	0.391	0.504	5.18%
MVS	0.166	0.120	-0.934	17.23	0.320	0.401	0.513	4.99%
MVK	0.172	0.116	-0.912	18.27	0.314	0.387	0.494	5.00%
ICVP	0.162	0.146	-0.575	10.54	0.231	0.335	0.555	4.61%
(b) 25OPINV								
GMV reference portfolio								
GMV	0.125	0.154	-0.871	27.08	0.298	0.443	0.585	8.66%
MEKL (Gauss)	0.121	0.167	-0.647	18.82	0.250	0.394	0.534	7.37%***
MEKL (SLN)	0.109	0.173	-0.723	17.61	0.244	0.409	0.535	7.32%***
MEKL (GN)	0.116	0.169	-0.615	18.24	0.253	0.398	0.546	7.31%***
KWZ reference portfolio								
KWZ	0.146	0.168	-0.680	22.09	0.269	0.360	0.474	8.22%
MEKL (Gauss)	0.131	0.176	-0.335	10.35	0.221	0.270	0.382	5.44%
MEKL (SLN)	0.112	0.181	-0.651	14.28	0.218	0.312	0.405	6.76%**
MEKL (GN)	0.123	0.181	-0.744	21.46	0.234	0.360	0.463	8.71%
Higher-moment portfolios from the literature								
CRRA	0.124	0.154	-0.869	26.78	0.294	0.439	0.579	8.61%
MVS	0.127	0.158	-0.734	25.29	0.304	0.448	0.609	8.46%
MVK	0.124	0.154	-0.864	26.83	0.298	0.444	0.587	8.63%
ICVP	0.142	0.209	-0.476	10.72	0.337	0.549	0.635	6.64%

Table IA.1: Out-of-sample performance of portfolio strategies net of transaction costs (continued)

	$252\hat{\mu}_P$	$\sqrt{252}\hat{\sigma}_P$	$\hat{\zeta}_P$	$\hat{\kappa}_P$	EKL (Gauss)	EKL (SLN)	EKL (GN)	MVaR
(c) 30IND								
GMV reference portfolio								
GMV	0.095	0.124	-1.005	24.62	0.287	0.422	0.550	6.57%
MEKL (Gauss)	0.073	0.137	-1.028	17.54	0.241	0.408	0.508	5.82%***
MEKL (SLN)	0.077	0.140	-0.972	17.25	0.237	0.410	0.508	5.89%**
MEKL (GN)	0.073	0.139	-1.056	17.41	0.243	0.420	0.521	5.87%*
KWZ reference portfolio								
KWZ	0.112	0.154	-0.823	17.19	0.275	0.400	0.528	6.44%
MEKL (Gauss)	0.066	0.154	-0.804	11.22	0.232	0.335	0.426	5.10%**
MEKL (SLN)	0.056	0.153	-0.924	14.07	0.223	0.345	0.508	5.74%**
MEKL (GN)	0.062	0.153	-0.822	12.02	0.227	0.332	0.419	5.26%***
Higher-moment portfolios from the literature								
CRRA	0.095	0.124	-1.026	24.47	0.285	0.419	0.543	6.53%
MVS	0.098	0.131	-0.749	23.96	0.292	0.433	0.602	6.80%
MVK	0.103	0.128	-1.033	23.04	0.275	0.421	0.545	6.46%
ICVP	0.072	0.171	-0.886	13.63	0.299	0.612	0.679	6.30%
(d) 100CRSP								
GMV reference portfolio								
GMV	0.147	0.096	-0.124	23.73	0.441	0.448	0.559	4.73%
MEKL (Gauss)	0.144	0.101	-0.298	17.12	0.359	0.369	0.463	4.07%**
MEKL (SLN)	0.140	0.106	-0.010	17.10	0.346	0.352	0.460	4.19%*
MEKL (GN)	0.145	0.101	-0.282	14.38	0.347	0.355	0.444	3.69%***
KWZ reference portfolio								
KWZ	0.151	0.098	-0.160	22.92	0.648	0.633	0.748	4.73%
MEKL (Gauss)	0.143	0.101	-0.324	15.31	0.569	0.555	0.660	3.84%**
MEKL (SLN)	0.145	0.108	-0.001	16.38	0.542	0.526	0.634	4.11%**
MEKL (GN)	0.143	0.103	-0.250	14.90	0.557	0.543	0.647	3.81%**
Higher-moment portfolios from the literature								
CRRA	0.148	0.096	-0.123	23.64	0.439	0.446	0.557	4.73%
MVS	0.018	0.426	-0.987	37.69	3.731	15.525	17.475	30.84%
MVK	0.116	0.431	1.461	55.95	3.478	14.821	16.973	36.64%
ICVP	0.160	0.120	0.098	21.78	0.311	0.330	0.416	5.50%

Note. This table reports the out-of-sample performance of the portfolio strategies considered in Sections 4.6.1 and 4.6.2 for the four datasets of daily returns of Section 4.4. We consider two reference portfolios: the sample global-minimum-variance (GMV) portfolio and the combination of the sample GMV and sample mean-variance portfolios that maximizes the expected out-of-sample utility according to Kan et al. (2022) (KWZ portfolio), using a risk-aversion coefficient $\lambda = 10$. We then consider three target-return distributions whose mean and variance match those of the two reference portfolios: a Gaussian distribution, a shifted-log-normal (SLN) distribution with a skewness of 0.5, and a generalized-normal (GN) distribution with an excess kurtosis of -0.5. The minimum-EKL (MEKL) portfolios minimize the extended Kullback-Leibler divergence with respect to these target distributions. The out-of-sample returns are adjusted with proportional transaction costs of 10 basis points. As performance criteria, we report, in order, the mean, standard deviation, skewness, excess kurtosis, EKL divergence with respect to the three target distributions (we take GMV as a reference portfolio for the higher-moment portfolios from the literature), and the 1% modified value-at-risk; see Section 4.5 for more details. The stars $\star$, $\star\star$, and $\star\star\star$ indicate that the MEKL portfolios have a smaller MVaR than the reference portfolios at a confidence level of 10%, 5%, and 1%.

Table IA.2: Out-of-sample performance of portfolio strategies under the Laplace target distribution (no transaction costs)

	$252\hat{\mu}_P$	$\sqrt{252}\hat{\sigma}_P$	$\hat{\zeta}_P$	$\hat{\kappa}_P$	EKL (Gauss)	EKL (SLN)	EKL (Laplace)	MVaR
(a) 25SBTM								
GMV reference portfolio								
GMV	0.186	0.117	-0.855	19.40	0.318	0.392	0.508	5.23%
MEKL (GN)	0.200	0.126	-0.523	12.59	0.250	0.310	0.169	4.34%***
MEKL (Gaussian)	0.198	0.126	-0.429	12.71	0.251	0.304	0.169	4.31%***
MEKL (Laplace)	0.186	0.122	-0.538	18.43	0.299	0.366	0.181	5.25%
KWZ reference portfolio								
KWZ	0.370	0.197	-0.128	4.07	0.196	0.212	0.309	4.02%
MEKL (GN)	0.336	0.201	0.016	4.14	0.182	0.189	0.155	4.03%
MEKL (Gaussian)	0.279	0.209	-0.055	4.22	0.190	0.208	0.157	4.30%
MEKL (Laplace)	0.308	0.194	-0.108	4.49	0.205	0.220	0.153	4.11%
(b) 25OPINV								
GMV reference portfolio								
GMV	0.135	0.154	-0.855	27.08	0.298	0.442	0.174	8.66%
MEKL (GN)	0.135	0.169	-0.601	18.26	0.253	0.395	0.167	7.31%***
MEKL (Gaussian)	0.139	0.167	-0.634	18.84	0.250	0.391	0.165	7.37%***
MEKL (Laplace)	0.123	0.162	-0.894	24.64	0.292	0.464	0.170	8.57%**
KWZ reference portfolio								
KWZ	0.162	0.168	-0.664	22.12	0.269	0.359	0.170	8.22%
MEKL (GN)	0.147	0.181	-0.730	21.50	0.233	0.358	0.163	8.70%
MEKL (Gaussian)	0.155	0.176	-0.320	10.37	0.221	0.268	0.159	5.42%
MEKL (Laplace)	0.133	0.171	-0.942	20.00	0.283	0.394	0.174	7.89%**
(c) 30IND								
GMV reference portfolio								
GMV	0.103	0.124	-0.989	24.60	0.288	0.421	0.175	6.57%
MEKL (GN)	0.087	0.139	-1.043	17.34	0.243	0.418	0.167	5.85%*
MEKL (Gaussian)	0.087	0.137	-1.016	17.47	0.241	0.406	0.169	5.80%***
MEKL (Laplace)	0.105	0.129	-0.836	19.49	0.254	0.378	0.167	5.85%**
KWZ reference portfolio								
KWZ	0.124	0.154	-0.804	17.17	0.276	0.399	0.168	6.44%
MEKL (GN)	0.082	0.153	-0.811	11.98	0.226	0.330	0.163	5.25%*
MEKL (Gaussian)	0.085	0.154	-0.793	11.19	0.232	0.333	0.165	5.08%***
MEKL (Laplace)	0.124	0.147	-0.419	13.46	0.264	0.328	0.168	5.25%**
(d) 100CRSP								
GMV reference portfolio								
GMV	0.156	0.096	-0.126	23.73	0.442	0.449	0.276	4.73%
MEKL (GN)	0.158	0.101	-0.284	14.39	0.347	0.356	0.230	3.68%***
MEKL (Gaussian)	0.156	0.101	-0.301	17.12	0.359	0.370	0.375	4.07%***
MEKL (Laplace)	0.155	0.101	-0.142	21.55	0.393	0.403	0.249	4.70%**
KWZ reference portfolio								
KWZ	0.160	0.098	-0.163	22.93	0.648	0.634	0.423	4.73%
MEKL (GN)	0.156	0.103	-0.253	14.91	0.557	0.543	0.367	3.81%***
MEKL (Gaussian)	0.156	0.101	-0.327	15.33	0.569	0.556	0.375	3.84%**
MEKL (Laplace)	0.160	0.102	-0.137	21.59	0.593	0.580	0.388	4.74%

Note. This table reports the out-of-sample performance of the portfolio strategies considered in Sections 4.6.1 and 4.6.2 for the four datasets of daily returns of Section 4.4. We consider two reference portfolios: the sample global-minimum-variance (GMV) portfolio and the combination of the sample GMV and sample mean-variance portfolios that maximizes the expected out-of-sample utility according to Kan et al. (2022) (KWZ portfolio), using a risk-aversion coefficient $\lambda = 10$. We then consider three generalized-normal target-return distributions whose mean and variance match those of the two reference portfolios and having an excess kurtosis of -0.5, 0 (Gaussian distribution), and 3 (Laplace distribution). The minimum-EKL (MEKL) portfolios minimize the extended Kullback-Leibler divergence with respect to these target distributions. As performance criteria, we report, in order, the mean, standard deviation, skewness, excess kurtosis, EKL divergence with respect to the three target distributions (we take GMV as a reference portfolio for the higher-moment portfolios from the literature), and the 1% modified value-at-risk; see Section 4.5 for more details. The stars $\star$, $\star\star$, and $\star\star\star$ indicate that the MEKL portfolios have a smaller MVaR than the reference portfolios at a confidence level of 10%, 5%, and 1%.

Table IA.3: Out-of-sample performance of portfolio strategies under the Laplace target distribution (net of transaction costs)

	$252\hat{\mu}_P$	$\sqrt{252}\hat{\sigma}_P$	$\hat{\zeta}_P$	$\hat{\kappa}_P$	EKL (Gauss)	EKL (SLN)	EKL (Laplace)	MVaR
(a) 25SBTM								
GMV reference portfolio								
GMV	0.171	0.117	-0.889	19.43	0.318	0.394	0.508	5.24%
MEKL (GN)	0.176	0.126	-0.565	12.56	0.250	0.313	0.169	4.35%***
MEKL (Gaussian)	0.174	0.126	-0.469	12.67	0.251	0.307	0.169	4.33%**
MEKL (Laplace)	0.164	0.122	-0.573	18.39	0.298	0.368	0.181	5.26%
KWZ reference portfolio								
KWZ	0.332	0.197	-0.144	4.07	0.196	0.213	0.309	4.05%
MEKL (GN)	0.283	0.202	-0.008	4.14	0.183	0.192	0.155	4.08%
MEKL (Gaussian)	0.220	0.209	-0.073	4.18	0.191	0.211	0.158	4.34%
MEKL (Laplace)	0.260	0.195	-0.125	4.44	0.205	0.221	0.153	4.14%
(b) 25OPINV								
GMV reference portfolio								
GMV	0.125	0.154	-0.871	27.08	0.298	0.443	0.174	8.66%
MEKL (GN)	0.116	0.169	-0.615	18.24	0.253	0.398	0.167	7.31%***
MEKL (Gaussian)	0.121	0.167	-0.647	18.82	0.250	0.394	0.165	7.37%***
MEKL (Laplace)	0.106	0.162	-0.907	24.64	0.292	0.466	0.169	8.58%**
KWZ reference portfolio								
KWZ	0.146	0.168	-0.680	22.09	0.269	0.360	0.170	8.22%
MEKL (GN)	0.123	0.181	-0.744	21.46	0.234	0.360	0.163	8.71%
MEKL (Gaussian)	0.131	0.176	-0.335	10.35	0.221	0.270	0.159	5.44%
MEKL (Laplace)	0.113	0.172	-0.952	19.97	0.283	0.395	0.173	7.90%**
(c) 30IND								
GMV reference portfolio								
GMV	0.095	0.124	-1.005	24.62	0.287	0.422	0.175	6.57%
MEKL (GN)	0.073	0.139	-1.056	17.41	0.243	0.420	0.167	5.87%*
MEKL (Gaussian)	0.073	0.137	-1.028	17.54	0.241	0.408	0.169	5.82%***
MEKL (Laplace)	0.093	0.129	-0.848	19.54	0.253	0.379	0.167	5.86%**
KWZ reference portfolio								
KWZ	0.112	0.154	-0.823	17.19	0.275	0.400	0.168	6.44%
MEKL (GN)	0.062	0.153	-0.822	12.02	0.227	0.332	0.163	5.26%***
MEKL (Gaussian)	0.066	0.154	-0.804	11.22	0.232	0.335	0.165	5.10%***
MEKL (Laplace)	0.108	0.147	-0.437	13.44	0.263	0.329	0.168	5.26%**
(d) 100CRSP								
GMV reference portfolio								
GMV	0.147	0.096	-0.124	23.73	0.441	0.448	0.276	4.73%
MEKL (GN)	0.145	0.101	-0.282	14.38	0.347	0.355	0.230	3.69%***
MEKL (Gaussian)	0.144	0.101	-0.298	17.12	0.359	0.369	0.237	4.07%**
MEKL (Laplace)	0.143	0.101	-0.139	21.51	0.393	0.402	0.249	4.70%**
KWZ reference portfolio								
KWZ	0.151	0.098	-0.160	22.92	0.648	0.633	0.423	4.73%
MEKL (GN)	0.143	0.103	-0.250	14.90	0.557	0.543	0.368	3.81%**
MEKL (Gaussian)	0.143	0.101	-0.324	15.31	0.569	0.555	0.375	3.84%**
MEKL (Laplace)	0.148	0.102	-0.135	21.56	0.593	0.579	0.388	4.74%

Note. This table reports the out-of-sample performance of the portfolio strategies considered in Sections 4.6.1 and 4.6.2 for the four datasets of daily returns of Section 4.4. We consider two reference portfolios: the sample global-minimum-variance (GMV) portfolio and the combination of the sample GMV and sample mean-variance portfolios that maximizes the expected out-of-sample utility according to Kan et al. (2022) (KWZ portfolio), using a risk-aversion coefficient $\lambda = 10$. We then consider three generalized-normal target-return distributions whose mean and variance match those of the two reference portfolios and having an excess kurtosis of -0.5, 0 (Gaussian distribution), and 3 (Laplace distribution). The minimum-EKL (MEKL) portfolios minimize the extended Kullback-Leibler divergence with respect to these target distributions. The out-of-sample returns are adjusted with proportional transaction costs of 10 basis points. As performance criteria, we report, in order, the mean, standard deviation, skewness, excess kurtosis, EKL divergence with respect to the three target distributions (we take GMV as a reference portfolio for the higher-moment portfolios from the literature), and the 1% modified value-at-risk; see Section 4.5 for more details. The stars $\star$, $\star\star$, and $\star\star\star$ indicate that the MEKL portfolios have a smaller MVaR than the reference portfolios at a confidence level of 10%, 5%, and 1%.

Table IA.4: Out-of-sample performance of portfolio strategies under short-sale constraints without transaction costs (continued on next page)

	$252\hat{\mu}_P$	$\sqrt{252}\hat{\sigma}_P$	$\hat{\zeta}_P$	$\hat{\kappa}_P$	EKL (Gauss)	EKL (SLN)	EKL (GN)	MVaR
(a) 25SBTM								
GMV reference portfolio								
GMV	0.141	0.152	-0.719	19.47	0.378	0.462	0.593	6.87%
MEKL (Gaussian)	0.148	0.174	-0.663	13.58	0.324	0.435	0.538	6.32%*
MEKL (SLN)	0.138	0.170	-0.672	14.09	0.319	0.423	0.526	6.30%*
MEKL (GN)	0.144	0.169	-0.621	14.96	0.331	0.432	0.546	6.48%*
KWZ reference portfolio								
KWZ	0.142	0.155	-0.824	18.72	0.373	0.466	0.588	6.82%
MEKL (Gaussian)	0.145	0.176	-0.667	13.27	0.320	0.432	0.534	6.30%*
MEKL (SLN)	0.136	0.171	-0.721	13.88	0.321	0.430	0.530	6.31%
MEKL (GN)	0.143	0.170	-0.693	14.34	0.329	0.434	0.541	6.39%
Higher-moment portfolios from the literature								
CRRA	0.141	0.152	-0.691	19.63	0.376	0.457	0.589	6.86%
MVS	0.181	0.120	-0.896	17.23	0.320	0.400	0.566	4.98%
MVK	0.187	0.116	-0.878	18.25	0.314	0.386	0.569	4.99%
ICVP	0.141	0.152	-0.711	19.48	0.376	0.460	0.591	6.86%
(b) 25OPINV								
GMV reference portfolio								
GMV	0.133	0.163	-0.829	22.58	0.328	0.467	0.574	8.14%
MEKL (Gauss)	0.141	0.172	-0.737	19.34	0.315	0.462	0.565	7.74%*
MEKL (SLN)	0.129	0.170	-0.853	23.14	0.299	0.467	0.552	8.63%
MEKL (GN)	0.135	0.168	-0.812	21.97	0.309	0.458	0.557	8.19%
KWZ reference portfolio								
KWZ	0.136	0.164	-0.898	23.37	0.327	0.462	0.574	8.38%
MEKL (Gauss)	0.139	0.173	-0.765	19.46	0.310	0.447	0.563	7.80%*
MEKL (SLN)	0.127	0.171	-0.852	23.42	0.301	0.455	0.560	8.74%
MEKL (GN)	0.131	0.168	-0.824	22.27	0.307	0.444	0.555	8.27%
Higher-moment portfolios from the literature								
CRRA	0.133	0.164	-0.825	22.59	0.327	0.466	0.572	8.14%
MVS	0.138	0.158	-0.713	25.33	0.305	0.446	0.537	8.46%
MVK	0.135	0.154	-0.848	26.83	0.299	0.443	0.526	8.63%
ICVP	0.134	0.163	-0.815	22.19	0.325	0.461	0.566	8.02%

Table IA.4: Out-of-sample performance of portfolio strategies under short-sale constraints without transaction costs (continued)

	$252\hat{\mu}_P$	$\sqrt{252}\hat{\sigma}_P$	$\hat{\zeta}_P$	$\hat{\kappa}_P$	EKL (Gauss)	EKL (SLN)	EKL (GN)	MVaR
(c) 30IND								
GMV reference portfolio								
GMV	0.114	0.133	-0.825	27.52	0.331	0.447	0.569	7.60%
MEKL (Gauss)	0.130	0.147	-0.640	18.76	0.276	0.395	0.501	6.45%**
MEKL (SLN)	0.118	0.145	-0.969	25.23	0.286	0.448	0.529	7.77%
MEKL (GN)	0.127	0.144	-0.656	19.94	0.285	0.401	0.512	6.59%**
KWZ reference portfolio								
KWZ	0.119	0.138	-0.917	27.42	0.326	0.448	0.569	7.83%
MEKL (Gauss)	0.132	0.150	-0.637	18.43	0.275	0.387	0.504	6.49%**
MEKL (SLN)	0.116	0.147	-0.995	25.18	0.283	0.439	0.528	7.90%
MEKL (GN)	0.129	0.146	-0.626	19.00	0.279	0.383	0.503	6.48%**
Higher-moment portfolios from the literature								
CRRA	0.114	0.132	-0.854	25.53	0.333	0.429	0.552	7.15%
MVS	0.106	0.132	-0.731	24.03	0.314	0.389	0.513	6.81%
MVK	0.111	0.128	-1.017	23.02	0.305	0.384	0.483	6.45%
ICVP	0.111	0.138	-0.719	24.32	0.316	0.431	0.552	7.21%
(d) 100CRSP								
GMV reference portfolio								
GMV	0.179	0.102	-0.225	15.89	0.429	0.438	0.547	3.89%
MEKL (Gauss)	0.180	0.110	-0.417	11.87	0.348	0.365	0.460	3.62%
MEKL (SLN)	0.181	0.113	-0.212	8.22	0.318	0.326	0.417	3.05%***
MEKL (GN)	0.178	0.110	-0.409	10.78	0.341	0.357	0.449	3.46%**
KWZ reference portfolio								
KWZ	0.180	0.102	-0.268	15.88	0.501	0.500	0.606	3.92%
MEKL (Gauss)	0.179	0.110	-0.395	12.10	0.418	0.422	0.517	3.66%
MEKL (SLN)	0.182	0.114	-0.229	8.42	0.381	0.380	0.470	3.11%***
MEKL (GN)	0.179	0.111	-0.387	10.63	0.406	0.409	0.501	3.44%**
Higher-moment portfolios from the literature								
CRRA	0.176	0.101	-0.196	16.09	0.435	0.443	0.553	3.89%
MVS	0.078	0.427	-0.922	37.69	3.725	15.22	17.52	30.86%
MVK	0.183	0.431	1.494	55.97	3.479	14.50	17.00	36.47%
ICVP	0.185	0.103	-0.242	14.94	0.406	0.416	0.522	3.80%

Note. This table reports the out-of-sample performance of the portfolio strategies considered in Sections 4.6.1 and 4.6.2 for the four datasets of daily returns of Section 4.4. We consider two reference portfolios: the sample global-minimum-variance (GMV) portfolio and the combination of the sample GMV and sample mean-variance portfolios that maximizes the expected out-of-sample utility according to Kan et al. (2022) (KWZ portfolio), using a risk-aversion coefficient $\lambda = 10$. We then consider three target-return distributions whose mean and variance match those of the two reference portfolios: a Gaussian distribution, a shifted-log-normal (SLN) distribution with a skewness of 0.5, and a generalized-normal (GN) distribution with an excess kurtosis of -0.5. The minimum-EKL (MEKL) portfolios minimize the extended Kullback-Leibler divergence with respect to these target distributions. All portfolio strategies are computed under short-sale constraints. As performance criteria, we report, in order, the mean, standard deviation, skewness, excess kurtosis, EKL divergence with respect to the three target distributions (we take GMV as a reference portfolio for the higher-moment portfolios from the literature), and the 1% modified value-at-risk; see Section 4.5 for more details. The stars $\star$, $\star\star$, and $\star\star\star$ indicate that the MEKL portfolios have a smaller MVaR than the reference portfolios at a confidence level of 10%, 5%, and 1%.

Table IA.5: Out-of-sample performance of portfolio strategies under short-sale constraints net of transaction costs (continued on next page)

	$252\hat{\mu}_P$	$\sqrt{252}\hat{\sigma}_P$	$\hat{\zeta}_P$	$\hat{\kappa}_P$	EKL (Gauss)	EKL (SLN)	EKL (GN)	MVaR
(a) 25SBTM								
GMV reference portfolio								
GMV	0.139	0.152	-0.725	19.46	0.378	0.463	0.593	6.87%
MEKL (Gaussian)	0.144	0.174	-0.671	13.54	0.323	0.435	0.537	6.31%*
MEKL (SLN)	0.135	0.170	-0.681	14.05	0.318	0.423	0.525	6.29%*
MEKL (GN)	0.140	0.169	-0.631	14.92	0.331	0.433	0.545	6.47% *
KWZ reference portfolio								
KWZ	0.140	0.155	-0.829	18.71	0.373	0.466	0.588	6.82%
MEKL (Gaussian)	0.141	0.176	-0.675	13.24	0.320	0.432	0.533	6.30%*
MEKL (SLN)	0.132	0.171	-0.730	13.83	0.321	0.430	0.530	6.30%
MEKL (GN)	0.140	0.170	-0.702	14.29	0.328	0.435	0.540	6.38%*
Higher-moment portfolios from the literature								
CRRA	0.139	0.151	-0.697	19.61	0.376	0.457	0.589	6.85%
MVS	0.166	0.120	-0.934	17.23	0.320	0.401	0.566	4.99%
MVK	0.172	0.116	-0.912	18.27	0.314	0.387	0.569	5.00%
ICVP	0.139	0.152	-0.716	19.46	0.376	0.460	0.591	6.86%
(b) 25OPINV								
GMV reference portfolio								
GMV	0.130	0.163	-0.835	22.58	0.328	0.468	0.573	8.14%
MEKL (Gauss)	0.138	0.172	-0.744	19.32	0.315	0.462	0.565	7.74%*
MEKL (SLN)	0.125	0.170	-0.858	23.14	0.299	0.467	0.552	8.63%
MEKL (GN)	0.131	0.168	-0.818	21.96	0.309	0.458	0.556	8.19%
KWZ reference portfolio								
KWZ	0.133	0.164	-0.904	23.37	0.326	0.462	0.574	8.38%
MEKL (Gauss)	0.135	0.173	-0.771	19.44	0.310	0.447	0.562	7.80%*
MEKL (SLN)	0.124	0.171	-0.858	23.41	0.301	0.455	0.559	8.73%
MEKL (GN)	0.128	0.168	-0.830	22.26	0.307	0.444	0.555	8.27%
Higher-moment portfolios from the literature								
CRRA	0.130	0.163	-0.830	22.58	0.327	0.466	0.571	8.14%
MVS	0.127	0.158	-0.734	25.29	0.304	0.448	0.536	8.46%
MVK	0.124	0.154	-0.864	26.83	0.298	0.444	0.525	8.63%
ICVP	0.131	0.163	-0.819	22.19	0.325	0.461	0.566	8.02%

Table IA.5: Out-of-sample performance of portfolio strategies under short-sale constraints net of transaction costs (continued)

	$252\hat{\mu}_P$	$\sqrt{252}\hat{\sigma}_P$	$\hat{\zeta}_P$	$\hat{\kappa}_P$	EKL (Gauss)	EKL (SLN)	EKL (GN)	MVaR
(c) 30IND								
GMV reference portfolio								
GMV	0.112	0.133	-0.832	27.50	0.331	0.447	0.569	7.60%
MEKL (Gauss)	0.127	0.147	-0.649	18.71	0.276	0.395	0.500	6.44%**
MEKL (SLN)	0.115	0.144	-0.976	25.24	0.286	0.448	0.529	7.77%
MEKL (GN)	0.124	0.144	-0.665	19.91	0.284	0.401	0.511	6.58%**
KWZ reference portfolio								
KWZ	0.117	0.138	-0.923	27.41	0.326	0.448	0.569	7.83%
MEKL (Gauss)	0.129	0.149	-0.646	18.38	0.275	0.387	0.503	6.48%**
MEKL (SLN)	0.113	0.147	-1.002	25.18	0.282	0.439	0.527	7.90%
MEKL (GN)	0.126	0.146	-0.635	18.96	0.279	0.383	0.503	6.47%**
Higher-moment portfolios from the literature								
CRRA	0.112	0.132	-0.861	25.51	0.327	0.436	0.552	7.15%
MVS	0.098	0.131	-0.749	23.96	0.292	0.433	0.512	6.81%
MVK	0.103	0.128	-1.033	23.04	0.275	0.421	0.483	6.46%
ICVP	0.108	0.138	-0.726	24.28	0.316	0.432	0.552	7.20%
(d) 100CRSP								
GMV reference portfolio								
GMV	0.173	0.102	-0.223	15.90	0.429	0.438	0.547	3.89%
MEKL (Gauss)	0.173	0.110	-0.415	11.88	0.348	0.365	0.460	3.62%
MEKL (SLN)	0.176	0.114	-0.227	8.43	0.381	0.380	0.470	3.11%***
MEKL (GN)	0.172	0.110	-0.407	10.79	0.341	0.357	0.449	3.46%*
KWZ reference portfolio								
KWZ	0.174	0.102	-0.265	15.89	0.501	0.500	0.606	3.93%
MEKL (Gauss)	0.172	0.110	-0.393	12.11	0.419	0.422	0.518	3.66%
MEKL (SLN)	0.176	0.114	-0.227	8.43	0.381	0.380	0.470	3.11%***
MEKL (GN)	0.173	0.110	-0.385	10.64	0.406	0.409	0.501	3.44%**
Higher-moment portfolios from the literature								
CRRA	0.170	0.101	-0.193	16.10	0.435	0.443	0.553	3.89%
MVS	0.035	0.427	-0.991	37.49	3.719	15.47	17.47	30.75%
MVK	0.126	0.431	1.453	55.79	3.479	14.79	16.98	36.59%
ICVP	0.179	0.103	-0.240	14.95	0.406	0.416	0.522	3.81%

Note. This table reports the out-of-sample performance of the portfolio strategies considered in Sections 4.6.1 and 4.6.2 for the four datasets of daily returns of Section 4.4. We consider two reference portfolios: the sample global-minimum-variance (GMV) portfolio and the combination of the sample GMV and sample mean-variance portfolios that maximizes the expected out-of-sample utility according to Kan et al. (2022) (KWZ portfolio), using a risk-aversion coefficient $\lambda = 10$. We then consider three target-return distributions whose mean and variance match those of the two reference portfolios: a Gaussian distribution, a shifted-log-normal (SLN) distribution with a skewness of 0.5, and a generalized-normal (GN) distribution with an excess kurtosis of -0.5. The minimum-EKL (MEKL) portfolios minimize the extended Kullback-Leibler divergence with respect to these target distributions. All portfolio strategies are computed under short-sale constraints. The out-of-sample returns are adjusted with proportional transaction costs of 10 basis points. As performance criteria, we report, in order, the mean, standard deviation, skewness, excess kurtosis, EKL divergence with respect to the three target distributions (we take GMV as a reference portfolio for the higher-moment portfolios from the literature), and the 1% modified value-at-risk; see Section 4.5 for more details. The stars $\star$, $\star\star$, and $\star\star\star$ indicate that the MEKL portfolios have a smaller MVaR than the reference portfolios at a confidence level of 10%, 5%, and 1%.

IA.6 Proofs of all results

Proof of Proposition 1

The support set of the SLN density is $\Omega_T = (\delta, \infty]$. Applying the decomposition of the EKL divergence in (7) to the SLN target density $f_T(x; \alpha, \beta, \delta)$ in (9) yields

$$\mathcal{C}(w; \alpha, \beta, \delta) = \mathbb{E}[\ln f_T(P)|P_\delta > 0] + H[P|P_\delta > 0] + \ln(1-p),$$

where the conditional cross-entropy, ignoring the normalization constant, simplifies to

$$\begin{aligned}\mathbb{E}[\ln f_T(P)|P \geq \delta] &= -\mathbb{E}[\ln P_\delta|P_\delta \geq 0] - \frac{1}{2\beta^2}\mathbb{E}[(\ln P_\delta - \alpha)^2|P_\delta \geq 0] \\ &= -\mathbb{E}[\ln P_\delta|P_\delta \geq 0] - \frac{1}{2\beta^2}\left(\mathbb{V}[\ln P_\delta|P_\delta \geq 0] + (\mathbb{E}[\ln P_\delta|P_\delta \geq 0] - \alpha)^2\right).\end{aligned}$$

Proof of Proposition 2

From the Gram-Charlier expansion of the entropy in (20), we have that $H[P^\star]$ depends on the skewness via the term $\zeta_P^2/12$. Then, because ζ_P only appears in $H[P^\star]$ and $\widetilde{\mathbb{V}}[\ln P_\delta|P_\delta \geq 0]$, $\widetilde{\mathcal{C}}(w; \alpha, \beta, \delta)$ depends on ζ_P as

$$\widetilde{\mathcal{C}}(w; \alpha, \beta, \delta) = \frac{\zeta_P \sigma_P^3}{2\beta^2(\mu_P - \delta)^3} - \frac{\zeta_P^2}{12} + \text{constant}, \tag{IA27}$$

where the constant is independent from ζ_P. Moreover, because we assume $(\mu_P, \sigma_P) = (\mu_T, \sigma_T)$, we have from (11) that $\mu_P - \delta = \sigma_P g(\beta)^{1/2}$. Thus, (IA27) becomes

$$\widetilde{\mathcal{C}}(w; \alpha, \beta, \delta) = \frac{\zeta_P g(\beta)^{3/2}}{2\beta^2} - \frac{\zeta_P^2}{12} + \text{constant}, \tag{IA28}$$

The optimal skewness ζ_P maximizing (IA28) is easily found to be given by (17), provided it is attainable. The result for the kurtosis is found in a similar way from the fact that $H[P^\star]$ depends on the excess kurtosis via the term $\kappa_P^2/48$ in (20). Therefore, when $(\mu_P, \sigma_P) = (\mu_T, \sigma_T)$, $\widetilde{\mathcal{C}}(w; \alpha, \beta, \delta)$ depends on κ_P as

$$\widetilde{\mathcal{C}}(w; \alpha, \beta, \delta) = -\frac{(\kappa_P + 2)g(\beta)^2}{8\beta^2} - \frac{\kappa_P^2}{48} + \text{constant}, \tag{IA29}$$

The optimal excess kurtosis κ_P maximizing (IA29) is easily found to be given by (17), provided it is attainable.

Proof of Proposition 3

The Gaussian distribution is a special case of the GN distribution in (IA1) for $\gamma = 2$. Therefore, plugging $\gamma = 2$ in the GN-target objective function in (IA3) gives

$$\mathcal{C}(w;\alpha,\beta,2) = H[P^\star] + \ln\sigma_P - \frac{1}{2\beta^2}\mathbb{E}\left[(P-\alpha)^2\right] = H[P^\star] + \ln\sigma_P - \frac{1}{2}\left(\frac{\mu_P-\alpha}{\beta}\right)^2 - \frac{\sigma_P^2}{2\beta^2},$$

which corresponds indeed to (19). The result that $H[P^\star]$ is invariant to μ_P and σ_P is well known in information theory; see Cover and Thomas (2006).

Proof of Proposition 4

As a preliminary useful result, note that $H[P] = H[P^\star] + \ln\sigma_P$ and, when $P^\star \sim \mathcal{N}(0,1)$, $H[P^\star] = \frac{1}{2}\ln 2\pi e$ is a constant that can be removed from the objective function.

Noting that $\alpha = \ln(\mu_T - \delta) - \beta^2/2$, the objective function in (14) is given by the following function of (μ_P, σ_P) when $(\zeta_P, \kappa_P) = (0,0)$:

$$\begin{aligned}\mathcal{C}_{SLN}(\mu_P,\sigma_P) := \ln\sigma_P - \ln(\mu_P-\delta) + \frac{\sigma_P^2}{2(\mu_P-\delta)^2} \\ - \frac{1}{2\beta^2}\left[\frac{\sigma_P^2}{(\mu_P-\delta)^2} + \frac{\sigma_P^4}{2(\mu_P-\delta)^4} + \left(\ln\left(\frac{\mu_P-\delta}{\mu_T-\delta}\right) - \frac{\sigma_P^2}{2(\mu_P-\delta)^2} + \frac{\beta^2}{2}\right)^2\right].\end{aligned} \tag{IA30}$$

To prove the result, we show that

$$\left.\frac{\partial\mathcal{C}_{SLN}(\mu_P,\sigma_P)}{\partial\sigma_P^2}\right|_{\mu_P=\mu_T} = 0 \text{ for } \sigma_P = c_1(\beta)\sigma_T \tag{IA31}$$

and

$$\left.\frac{\partial\mathcal{C}_{SLN}(\mu_P,\sigma_P)}{\partial\mu_P}\right|_{\sigma_P=c_1(\beta)\sigma_T} = 0 \text{ for } \mu_P = \mu_T. \tag{IA32}$$

In the proof of part 1 of Proposition 5, it is shown that $\mathcal{C}_{SLN}(\mu_P,\sigma_P)$ has no other local optima and, thus, that this optimum is global.

First, we prove Equation (IA31). Noting that $\mu_T - \delta = \sigma_T g(\beta)^{-1/2}$, when $\mu_P = \mu_T$ the objective function (IA30) becomes, after removing terms independent from σ_P,

$$\mathcal{C}_{SLN}(\mu_P,\sigma_P) = \frac{\sigma_P^4}{\sigma_T^4}\left(-\frac{3g(\beta)^2}{8\beta^2}\right) + \frac{\sigma_P^2}{\sigma_T^2}\left(g(\beta)\left(\frac{3}{4} - \frac{1}{2\beta^2}\right)\right) + \frac{1}{2}\ln\sigma_P^2.$$

Setting the derivative with respect to σ_P^2 equal to zero, the optimal σ_P^2 must solve

$$\frac{\sigma_P^4}{\sigma_T^4}\left(\frac{3g(\beta)^2}{4\beta^2}\right)+\frac{\sigma_P^2}{\sigma_T^2}\left(g(\beta)\left(\frac{1}{2\beta^2}-\frac{3}{4}\right)\right)-\frac{1}{2}=0.$$

Of the two solutions to this second-degree polynomial, there is always only one positive solution, which is equal after simplifications to $\sigma_P = c_1(\beta)\sigma_T$.

Second, we prove Equation (IA32). Setting the derivative of $\mathcal{C}_{SLN}(\mu_P, \sigma_P)$ in (IA30) with respect to μ_P equal to zero, we need to prove that

$$\begin{aligned}&\frac{1}{\mu_P-\delta}+\frac{\sigma_P^2}{(\mu_P-\delta)^3}+\frac{1}{\beta^2}\left(\ln(\mu_P-\delta)-\alpha-\frac{\sigma_P^2}{2(\mu_P-\delta)^2}\right)\left(\frac{1}{\mu_P-\delta}+\frac{\sigma_P^2}{(\mu_P-\delta)^3}\right)\\&-\frac{1}{\beta^2}\left(\frac{\sigma_P^2}{(\mu_P-\delta)^3}+\frac{\sigma_P^4}{(\mu_P-\delta)^5}\right)=0\end{aligned}$$

for $\mu_P = \mu_T$ if $\sigma_P = c_1(\beta)\sigma_T$. Setting $\mu_P = \mu_T$ and further simplifications result in the equality

$$\frac{\sigma_P^4}{\sigma_T^4}\left(\frac{3g(\beta)^2}{2\beta^2}\right)+\frac{\sigma_P^2}{\sigma_T^2}\left(g(\beta)\left(\frac{3}{2\beta^2}-\frac{3}{2}\right)\right)-\frac{3}{2}=0,$$

which must hold for $\sigma_P = c_1(\beta)\sigma_T$. Solving for σ_P^2 and keeping the only positive root of the polynomial results in

$$\sigma_P=\sigma_T\sqrt{\frac{\beta^2-1+\sqrt{\beta^4+2\beta^2+1}}{2g(\beta)}}=c_1(\beta)\sigma_T=\sigma_{SLN}^{\star}.$$

Finally, $c_1(\beta) = 1$ if and only if $\beta \to 0^+$ because $\lim_{\beta\to 0^+} c_1(\beta) = 1$ and because, for $\beta > 0$, $c_1(\beta) < 1$ since $\exp(\beta^2) - 1 > \beta^2$ for all $\beta > 0$.

Proof of Proposition 5

Part 1. The strategy of the proof of part 1 of the proposition is to show that the partial derivatives of the MEKL objective function with respect to μ_P and σ_P vanish simultaneously at a single point in the plane $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}^+$, which, of course, coincides with the maximum $(\mu_T, \sigma_{SLN}^{\star})$ for the SLN target distribution in Proposition 4. This implies that there is no other stationary point than the global optimum, and thus, that starting from any point of the (μ, σ) plane, there always exists a geodesic connecting the point to the global optimum along which the MEKL objective function increases. This, in turn, shows that if (μ_T, σ_T) is chosen on or above the MVE frontier (implying that so is the global optimum because $\sigma_{SLN}^{\star} \leq \sigma_T$), then the maximum of the MEKL objective function in the restriction of the (μ, σ) plane associated with the set of attainable portfolio-return

mean and volatility, $\mathcal{MV}$, is reached on the MVE frontier. We focus in our proof on the MEKL portfolio $w_{\alpha,\beta,\delta}$ associated with the SLN target distribution; the proof for the MEKL portfolio $w_{\alpha,\beta}$ associated with the Gaussian target distribution follows as a special case by letting $\beta \to 0^+$.

Because the optimal $\mu_P = \mu_T$ and $\mu_T > \delta$, we have $\mu_P > \delta$, which is useful below. To simplify the formulas, we introduce the notation

$$\nu_P := \frac{\sigma_P}{(\mu_P - \delta)} > 0.$$

First, we derive the unique zero of $\frac{\partial \mathcal{C}_{SLN}(\mu_P,\sigma_P)}{\partial \sigma_P^2}$ for a fixed μ_P, where $\mathcal{C}_{SLN}(\mu_P, \sigma_P)$ is defined in (IA30). After some algebra, this amounts to find the roots of a second-degree polynomial in ν_P^2:

$$\nu_P^4 \left[-\frac{3}{4\beta^2} \right] + \nu_P^2 \left[\frac{1}{2\beta^2} \left(\frac{3\beta^2}{2} - 1 + \ln\left(\frac{\mu_P - \delta}{\mu_T - \delta} \right) \right) \right] + \frac{1}{2} = 0.$$

There is a unique positive root to this polynomial, which is

$$\nu_P^2 = \frac{1}{3} \left[\frac{3\beta^2}{2} - 1 + \ln\left(\frac{\mu_P - \delta}{\mu_T - \delta} \right) + \sqrt{6\beta^2 + \left(\frac{3\beta^2}{2} - 1 + \ln\left(\frac{\mu_P - \delta}{\mu_T - \delta} \right) \right)^2} \right]. \tag{IA33}$$

Second, we turn to the derivative with respect to μ_P for a fixed σ_P. Because $\mu_P > \delta$, the derivative of $\mathcal{C}_{SLN}(\mu_P, \sigma_P)$ with respect to $(\mu_P - \delta)^2$ has the same sign as that with respect to μ_P, and we find after some algebra that

$$-2\beta^2 (\mu_P - \delta)^2 \frac{\partial \mathcal{C}_{SLN}(\mu_P, \sigma_P)}{\partial (\mu_P - \delta)^2} = (1 + \nu_P^2) \left(\frac{3}{2}(\beta^2 - \nu_P^2) + \ln\left(\frac{\mu_P - \delta}{\mu_T - \delta} \right) \right). \tag{IA34}$$

This means that the derivative has a unique zero. Indeed, $\frac{\partial \mathcal{C}_{SLN}(\mu_P,\sigma_P)}{\partial (\mu_P - \delta)^2} = 0$ when

$$\frac{3}{2}(\beta^2 - \nu^2) + \ln\left(\frac{\mu_P - \delta}{\mu_T - \delta} \right) = 0 \Leftrightarrow -\frac{3\sigma_P^2}{(\mu_P - \delta)^2} + 2\ln(\mu_P - \delta) = 2\ln(\mu_T - \delta) - 3\beta^2.$$

The left-hand side of this last equality is continuous, spans $\mathbb{R}$ as a function of $\mu_P > \delta$, and is strictly increasing with μ_P. Therefore, it equals the right-hand for a single value of μ_P and the zero is unique.

Third, we show that the two partial derivatives of $\mathcal{C}_{SLN}(\mu_P, \sigma_P)$ vanish simultaneously at a single point, corresponding to the global optimum $(\mu_P, \sigma_P) = (\mu_T, \sigma^\star_{SLN})$. As explained above, this then proves part 1 of the proposition. From (IA34), the optimal μ_P for a given σ_P is such that $\ln\left(\frac{\mu_P - \delta}{\mu_T - \delta} \right) = \frac{3}{2}(\nu_P^2 - \beta^2)$. Inserting this condition into the optimal σ_P^2 in (IA33) means that the two

partial derivatives vanish simultaneously when

$$\left(\frac{3\nu_P^2}{2}+1\right)^2 = 6\beta^2 + \left(\frac{3\nu_P^2}{2}-1\right)^2 \;\Leftrightarrow\; 6\nu_P^2 = 6\beta^2 \;\Leftrightarrow\; \nu_P = \beta,$$

because both ν_P and β are positive. This unique solution coincides with the global optimum $(\mu_P, \sigma_P) = (\mu_T, \sigma^\star_{SLN})$ because $\nu_P = \beta$ in that case.

Part 2. When the asset returns are Gaussian, as noted in the proof of Proposition 4, the standardized entropy $H[P^\star]$ is a constant that can be removed from the objective function. Thus, the Lagrangian function associated with the objective function in (19) is

$$\mathcal{L} = -\frac{1}{2\beta^2}\big((w'\mu-\alpha)^2 + w'\boldsymbol{\Sigma}w\big) + \frac{1}{2}\ln w'\boldsymbol{\Sigma}w + \tau(w'\iota - 1),$$

where τ is the Lagrange multiplier associated with the budget constraint $w'\iota = 1$. Setting the derivative of $\mathcal{L}$ with respect to w equal to zero yields, after some developments,

$$w = \frac{w'\boldsymbol{\Sigma}w}{\beta^2 - w'\boldsymbol{\Sigma}w}\left(\frac{\mu_g(w'\mu-\alpha)}{\sigma_g^2}w_{msr} - \frac{\tau\beta^2}{\sigma_g^2}w_g\right). \tag{IA35}$$

Replacing (IA35) in the condition $w'\iota = 1$ gives the following value for the Lagrange multiplier τ:

$$\tau = \frac{\sigma_g^2}{\beta^2}\left(\frac{\mu_g(w'\mu-\alpha)}{\sigma_g^2} + 1 - \frac{\beta^2}{w'\boldsymbol{\Sigma}w}\right). \tag{IA36}$$

Combining (IA35) and (IA36), and denoting $(\mu_P^\star, \sigma_P^{\star 2})$ the optimal MEKL portfolio-return mean and variance, results in the following solution for the MEKL portfolio $w_{\alpha,\beta}$:

$$w_{\alpha,\beta} = w_g + \frac{\sigma_P^{\star 2}\mu_g(\mu_P^\star - \alpha)}{\sigma_g^2(\beta^2 - \sigma_P^{\star 2})}(w_{msr} - w_g), \tag{IA37}$$

which, as a special case of part 1 of the proposition, is an MVE portfolio when $\mu_T \geq \mu_g$ and $\sigma_T \leq \sigma_{EF}(\mu_T)$, i.e., $\sigma_P^\star = \sigma_{EF}(\mu_P^\star)$. To find the optimal mean return $\mu_P^\star$, one must solve $w'_{\alpha,\beta}\mu = \mu_P^\star$, which from (IA37) reduces to solving (24). Note that the division by $\beta^2 - \sigma_P^{\star 2}$ in (IA37) is not problematic. When $(\alpha,\beta) \notin \mathcal{MV}$, i.e., (α,β) is above the MVE frontier, then $\sigma_P^\star \neq \beta$ because of the presence of the first term in (19) that favors having a larger mean return than $\sigma_{EF}^{-1}(\beta)$. When $(\alpha,\beta) \in \mathcal{MV}$, i.e., (α,β) is on the MVE frontier, then the limit of $(\mu_P^\star - \alpha)/(\beta^2 - \sigma_P^{\star 2})$ in (IA37) as $\mu_P \to \alpha$ is necessarily such that $w_{\alpha,\beta}$ in an MVE portfolio with mean return $\mu_P = \alpha$.

Part 2. We first prove that $\mu_P^\star = \mu_g$ if $\alpha = \mu_g$. This is obvious when $\beta = \sigma_{EF}(\alpha) = \sigma_g$ because $(\alpha,\beta) \in \mathcal{MV}$ in such a case. Consider now the case $\beta < \sigma_g$ and denote $\beta^2 = \sigma_g^2/c$ with $c > 1$, in

which case $(\alpha, \beta) \notin \mathcal{MV}$. Denote the MEKL objective function as a function of (μ_P, σ_P) as

$$\mathcal{C}_N(\mu_P, \sigma_P) := -\frac{1}{2}\left(\frac{\mu_P - \alpha}{\beta}\right)^2 + \frac{1}{2}\left(\ln \sigma_P^2 - \frac{\sigma_P^2}{\beta^2}\right). \tag{IA38}$$

When $(\mu_P, \sigma_P) = (\mu_g, \sigma_g)$, we easily find that

$$\mathcal{C}_N(\mu_P, \sigma_P) = \frac{1}{2}(\ln \sigma_g^2 - c). \tag{IA39}$$

Let us now show that choosing an MVE portfolio with $\mu_P > \mu_g$ will necessarily result in a smaller value for the objective function. Specifically, consider $\mu_P = \mu_g + \epsilon$, $\epsilon > 0$, and thus $\sigma_P^2 = \sigma_{EF}^2(\mu_P) = \sigma_g^2 + \epsilon^2/\psi^2$. In that case, the objective function becomes

$$\begin{aligned}\mathcal{C}_N(\mu_P, \sigma_P) &= \frac{1}{2}\left(-\frac{c\epsilon^2}{\sigma_g^2}\left(1 + \frac{1}{\psi^2}\right) + \ln\left(\sigma_g^2 + \frac{\epsilon^2}{\psi^2}\right) - c\right) \\ &\leq \frac{1}{2}\left(-\frac{c\epsilon^2}{\sigma_g^2}\left(1 + \frac{1}{\psi^2}\right) + \ln \sigma_g^2 + \frac{\epsilon^2}{\sigma_g^2\psi^2} - c\right),\end{aligned} \tag{IA40}$$

from the inequality $\ln(x + a) \leq \ln x + a/x$. Comparing (IA39) and (IA40), we have that $\mu_P = \mu_g$ is optimal if

$$-\frac{c\epsilon^2}{\sigma_g^2}\left(1 + \frac{1}{\psi^2}\right) + \frac{\epsilon^2}{\sigma_g^2\psi^2} < 0 \quad \text{for} \quad \epsilon^2 > 0,$$

which holds because the inequality reduces to $-c + (1 - c)/\psi^2 < 0$ and the left-hand side is indeed negative because $c > 1$.

Finally, we prove that $\mu_P^\star \leq \mu_{\ell_2} \leq \alpha$ if $\alpha > \mu_g$. It is clear that the statement holds with equalities when $\beta = \sigma_{EF}(\alpha)$ because $(\alpha, \beta) \in \mathcal{MV}$ in such a case. Therefore, we focus on proving that the statement holds with strict inequalities when $\beta < \sigma_{EF}(\alpha)$, in which case $(\alpha, \beta) \notin \mathcal{MV}$. The mean return of the MVE portfolio minimizing the ℓ_2 distance to (α, β) solves

$$\mu_{\ell_2} = \min_{\mu_P} (\mu_P - \alpha)^2 + (\sigma_{EF}(\mu_P) - \beta)^2. \tag{IA41}$$

Observe from the geometry of the problem that μ_{ℓ_2} is unique. Setting the derivative of the ℓ_2 distance in the right-hand side of (IA41) equal to zero, μ_{ℓ_2} is thus the unique value of μ satisfying

$$f(\mu) := \frac{\alpha - \mu}{\mu - \mu_g}\psi^2\sigma_{EF}(\mu) = \sigma_{EF}(\mu) - \beta =: h(\mu).$$

Moreover, $\mu_{\ell_2} > \mu_g$ (because μ_{ℓ_2} sits on the MVE frontier when $\alpha > \mu_g$) and, from the concavity of the MVE frontier, we have $\mu_{\ell_2} < \alpha$ if $\beta \geq \sigma_g$. From (24), we find after some algebra that the

mean return $\mu_P^\star$ of the MEKL portfolio is the value of μ solving

$$f(\mu) = h(\mu)\frac{\sigma_{EF}(\mu)+\beta}{\sigma_{EF}(\mu)} =: l(\mu).$$

Because $\beta > 0$, we necessarily have $\mu_P^\star < \mu_{\ell_2}$. To see this, recall that $f(\mu) - h(\mu)$ has a unique root $\mu_{\ell_2} < \alpha$ and that $h(\alpha) > 0 = f(\alpha)$, which shows that $h(\mu) > f(\mu)$ for all $\mu > \mu_{\ell_2}$. Let us now prove that l dominates h on (μ_{ℓ_2}, ∞), which amounts to show that $h(\mu) > 0$ for all $\mu > \mu_{\ell_2}$. Observe that $h(\mu) > 0$ for all μ if $\beta < \sigma_g$. If $\beta \geq \sigma_g$, $\sigma_{EF}^{-1}(\beta)$ exists. Using $\mu_{\ell_2} > \sigma_{EF}^{-1}(\beta)$ and the increasing property of h,

$$h(\mu_{\ell_2}) > h(\sigma_{EF}^{-1}(\beta)) = 0.$$

Because h is increasing, we find that $h(\mu) > 0$ for all $\mu > \mu_{\ell_2}$, leading to $l(\mu) > h(\mu)$ on the right of μ_{ℓ_2}. We have thus shown that on (μ_{ℓ_2}, ∞), l dominates h and h dominates f, and conclude that l dominates f on the right of μ_{ℓ_2}. This proves that the root of $f(\mu) - l(\mu)$ must satisfy $\mu_P^\star < \mu_{\ell_2}$.

Proof of Proposition IA1

Because $\Omega_T = \mathbb{R}$ for $T \sim \mathcal{GN}(\alpha, \beta, \gamma)$, we have $\Omega_P \subseteq \Omega_T$ and the EKL divergence reduces to the KL divergence. Applying the decomposition of the KL divergence in (4) to the GN target density $f_T(x; \alpha, \beta, \gamma)$ in (IA1) yields

$$\mathcal{C}(w; \alpha, \beta, \gamma) = H[P] + \ln \frac{2^{-\frac{\gamma+1}{\gamma}}\gamma}{\beta\Gamma(1/\gamma)} - \frac{1}{2\beta^\gamma}\mathbb{E}\left[|P-\alpha|^\gamma\right],$$

where $H[P] = H[P^\star] + \ln \sigma_P$. Removing the constant middle term completes the proof.

Proof of Proposition IA2

Let $Z \sim \mathcal{N}(\mu_Z, \sigma_Z)$ be a Gaussian random variable and $k > -1$ a real number. Then, the kth absolute moment of Z is (Winkelbauer, 2014)

$$\mathbb{E}\left[|Z|^k\right] = \frac{\sigma_Z^k 2^{k/2}\Gamma\left(\frac{k+1}{2}\right)}{\pi^{1/2}} {}_1F_1\left(-\frac{1}{2}\left(\frac{\mu_Z}{\sigma_Z}\right)^2; -\frac{k}{2}, \frac{1}{2}\right), \tag{IA42}$$

where the function ${}_1F_1$ is the confluent hypergeometric function. Using this property with $k = 1$, the objective function in (IA3) is given by the following function of μ_P and σ_P when $P \sim \mathcal{N}(\mu_P, \sigma_P)$:

$$\mathcal{C}_{GN}(\mu_P, \sigma_P) := \ln \sigma_P - \frac{2^{\frac{\gamma-2}{2}}\Gamma\left(\frac{\gamma+1}{2}\right)}{\pi^{1/2}}\left(\frac{\sigma_P}{\beta}\right)^\gamma {}_1F_1\left(-\frac{1}{2}\left(\frac{\mu_P-\alpha}{\sigma_P}\right)^2; -\frac{\gamma}{2}, \frac{1}{2}\right). \tag{IA43}$$

To prove the result, we show that $\frac{\partial \mathcal{C}_{GN}(\mu_P,\sigma_P)}{\partial \mu_P} = 0$ if and only if $\mu_P = \alpha$, and then that $\frac{\partial \mathcal{C}_{GN}(\mu_P,\sigma_P)}{\partial \sigma_P}\big|_{\mu_P=\alpha} = 0$ if and only if $\sigma_P = c_2(\gamma)\sigma_T$. The first result is verified because

$$\frac{\partial_1 F_1\left(-\frac{1}{2}\left(\frac{\mu_P-\alpha}{\sigma_P}\right)^2; -\frac{\gamma}{2}, \frac{1}{2}\right)}{\partial \mu_P} = \gamma_1 F_1\left(-\frac{1}{2}\left(\frac{\mu_P-\alpha}{\sigma_P}\right)^2; 1-\frac{\gamma}{2}, \frac{3}{2}\right)\left(\frac{\mu_P-\alpha}{\sigma_P^2}\right) = 0$$

if and only if $\mu_P = \alpha$. This is the unique zero because ${}_1F_1(-x^2; 1-\gamma/2, 3/2) > 0$ for all $x \in \mathbb{R}$ and $\gamma > 0$. Then, because ${}_1F_1\left(0; -\frac{\gamma}{2}, \frac{1}{2}\right) = 1$ for all γ, the function $\mathcal{C}_{GN}(\mu_P, \sigma_P)$ in (IA43) becomes, for $\mu_P = \alpha$,

$$\mathcal{C}_{GN}(\mu_P, \sigma_P) = \ln \sigma_P - \frac{2^{\frac{\gamma-2}{2}}\Gamma\left(\frac{\gamma+1}{2}\right)}{\pi^{1/2}}\left(\frac{\sigma_P}{\beta}\right)^{\gamma}. \tag{IA44}$$

From (IA44), it is straightforward to check that the unique solution to $\frac{\partial \mathcal{C}_{GN}(\mu_P,\sigma_P)}{\partial \sigma_P}\big|_{\mu_P=\alpha} = 0$ is $\sigma_P = c_2(\gamma)\sigma_T = \sigma^\star_{GN}$. Given the unicity of the zeros to the partial derivatives, the optimum $(\mu_P, \sigma_P) = (\mu_T, \sigma^\star_{GN})$ is global and there is no other local optimum. Finally, $c_2(\gamma) \leq 1$ with equality if and only if $\gamma = 2$ because $\lim_{\gamma\to 0^+} c_2(\gamma) = \lim_{\gamma\to+\infty} c_2(\gamma) = 0$, and $c_2'(\gamma) > 0$ for $\gamma < 2$, $c_2'(\gamma) < 0$ for $\gamma > 2$.

Proof of Proposition IA3

Concerning the MEKL portfolio $w_{\alpha,\beta,\gamma}$ associated with the GN target distribution, it is demonstrated in the proof of Proposition IA2 that $(\mu_T, \sigma^\star_{GN})$ is the unique stationary point of the objective function $\mathcal{C}_{GN}(\mu_P, \sigma_P)$ in (IA43). As explained in the proof of part 1 of Proposition 5, this suffices to prove the proposition.

Proof of Proposition IA4

Because R is elliptically symmetric, the standardized portfolio return $(P - \mu_P)/\sigma_P$ has the same distribution for all w, and we denote its cdf by F, which has the property $F(x) = 1 - F(-x)$ because the density F' is even. Then, we have that

$$\mathbb{P}(\alpha - \epsilon \leq P \leq \alpha + \epsilon) = F\left(\frac{\alpha+\epsilon-\mu_P}{\sigma_P}\right) - F\left(\frac{\alpha-\epsilon-\mu_P}{\sigma_P}\right). \tag{IA45}$$

For any fixed σ_P, this probability is decreasing in $|\mu_P - \alpha|$, and thus it is optimal to have $\mu_P = \alpha$, which is attainable because any value of μ_P is attainable when the only constraint on w is $w'\iota = 1$. In this case, the probability becomes $F(\epsilon/\sigma_P) - F(-\epsilon/\sigma_P) = 2F(\epsilon/\sigma_P) - 1$. Because F in an increasing function and $\epsilon > 0$, $F(\epsilon/\sigma_P)$ is decreasing in σ_P, and thus given $\mu_P = \alpha$ it is optimal to take σ_P as small as possible. The portfolio w minimizing σ_P under the constraint $\mu_P = \alpha$ is the

mean-variance portfolio with mean return α.

Proof of Proposition IA5

The Lagrangian associated with the objective function $\mathcal{C}(w;\alpha)$ is

$$\mathcal{L} = -(w'\mu - \alpha)^2 - w'\boldsymbol{\Sigma}w + \tau(w'\iota - 1),$$

where τ is the Lagrange multiplier associated with the budget constraint $w'\iota = 1$. Define $\boldsymbol{A} := (\boldsymbol{\Sigma} + \mu\mu')^{-1}$. Then, setting the derivative of $\mathcal{L}$ with respect to w equal to zero yields, after some developments,

$$w = \boldsymbol{A}^{-1}\left(\alpha\mu - \frac{\tau}{2}\iota\right). \tag{IA46}$$

After replacing (IA46) in the constraint $w'\iota = 1$, the Lagrange multiplier τ is

$$\tau = \frac{2\big(\alpha\iota'\boldsymbol{A}^{-1}\mu - 1\big)}{\iota'\boldsymbol{A}^{-1}\iota}. \tag{IA47}$$

Combining (IA46) and (IA47) results in the following solution for w_α:

$$w_\alpha = \boldsymbol{A}\left(\alpha\mu - \frac{(\alpha\iota'\boldsymbol{A}\mu - 1)\,\iota}{\iota'\boldsymbol{A}\iota}\right). \tag{IA48}$$

From the updating formula for the inverse matrix (Greene, 1993, p.25), we have that

$$\boldsymbol{A} = \boldsymbol{\Sigma}^{-1} - \frac{1}{1+\theta^2}\boldsymbol{\Sigma}^{-1}\mu\mu'\boldsymbol{\Sigma}^{-1}.$$

As a result, we find that

$$\boldsymbol{A}\mu = \frac{\mu_g}{\sigma_g^2(\theta^2+1)}w_{msr}, \quad \text{and} \quad \boldsymbol{A}\iota = \frac{w_g}{\sigma_g^2} - \frac{w_{msr}}{\theta^2+1}\left(\frac{\mu_g}{\sigma_g^2}\right)^2,$$

and plugging these two equalities in (IA48) results in the following formula for w_α after simplifications:

$$w_\alpha = w_g\left[\frac{(\theta^2+1)\sigma_g^2 - \alpha\mu_g}{(\theta^2+1)\sigma_g^2 - \mu_g^2}\right] + w_{msr}\left[\frac{\mu_g(\mu_g - \alpha)}{\mu_g^2 - (\theta^2+1)\sigma_g^2}\right]. \tag{IA49}$$

Being a combination of w_g and w_{msr}, w_α is a mean-variance portfolio, whose mean return $w_\alpha'\mu$ in (IA15) is obtained by noting that $w_{msr}'\mu = \theta^2\sigma_g^2/\mu_g$.

Finally, $w_\alpha = w_g$ if $\alpha = \mu_g$ because, from (IA15), $w_\alpha'\mu = \mu_g$ if $\alpha = \mu_g$. Moreover, w_α is MVE if and only if $w_\alpha'\mu \geq \mu_g$, which, from (IA15), happens when $\alpha \geq \mu_g$.

References in internet appendix

Ao, M., Yingying, L., & Xinghua, Z. (2019). Approaching mean-variance efficiency for large portfolios. *The Review of Financial Studies*, 32(7), 2890–2919.

Azzalini, A. (1985). A class of distributions which includes the Normal ones. *Scandinavian Journal of Statistics*, 12(2), 171–178.

Cover, T., & Thomas, J. (2006). *Elements of Information Theory* (2nd ed.). John Wiley & Sons, New Jersey.

Frazzini, A., Israel, R., & Moskowitz, T. (2018). Trading costs. Available at SSRN 3229719.

Greene, W. (1993). *Econometric Analysis*. Macmillan, New York.

Nadarajah, S. (2005). A generalized normal distribution. *Journal of Applied Statistics*, 32(7), 685–694.

Pender, J. (2015). The truncated normal distribution: Applications to queues with impatient customers. *Operations Research Letters*, 43(1), 40–45.

Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science*, 44(12), 1650–1661.

Winkelbauer, A. (2014). Moments and absolute moments of the normal distribution. Available at arXiv 1209.4340v2.