Quantitative Finance, 2021
Vol. 21, No. 2, 199–219, https://doi.org/10.1080/14697688.2020.1762913

Portfolio optimization under the generalized hyperbolic distribution: optimal allocation, performance and tail behavior

JOHN R. BIRGE† and L. CHAVEZ-BEDOYA *‡

†University of Chicago Booth School of Business, 5807 S. Woodlawn Ave., Chicago, IL 60637, USA
‡ESAN Graduate School of Business, Alonso de Molina 1652, Surco, Lima, Peru

(*Received 23 September 2019; accepted 24 April 2020; published online 2 September 2020*)

In this paper, we analyze the asset allocation problem under the generalized hyperbolic (GH) distribution of returns and exponential utility. We provide closed-form expressions to compute the optimal portfolio weights; and we introduce two new measures, associated with a more general mean-risk trade-off, that allow us to express the optimal solution as an affine combination of two efficient portfolios: one minimizing risk and the other maximizing mean given a particular level of risk. Also, we prove that optimal portfolio performance is not monotonic in tail behavior since it increases when tails become lighter or heavier with respect to a particular threshold; however, distributions with heavier tails produce more conservative allocations in terms of the weight given to the minimum-risk portfolio increments. Finally, the practical relevance of our paper show that tail behavior greatly affects portfolio construction and performance, and that including non-normality features of short-term asset returns, through a GH distribution, has the potential to significantly improve the investor's certainty equivalent excess return.

Keywords: Portfolio optimization; Generalized hyperbolic distribution; Mean-variance; Minimum-risk portfolio; Tail density

JEL Classification: G11,

1. Introduction

The Markowitz (1952) mean-variance (MV) framework for portfolio selection is consistent with expected utility maximization when asset returns follow an elliptical distribution or the investor has a quadratic utility function. Nevertheless, the presence of short-term† investment opportunities whose returns are not normally distributed suggests the inclusion of higher moments (especially skewness and kurtosis) in the portfolio selection process (Madan and Yen 2008).

The fact that asset returns are not considered elliptically distributed is one of the most important criticisms to the MV approach.‡ Therefore, researchers and practitioners have incorporated different types of non-elliptical distributions to model asset returns more realistically. A prominent example, the generalized hyperbolic (GH) distribution, introduced by Bandorff-Nielsen (1977, 1978) and further explained in Bandorff-Nielsen and Blæsild (1981), has received attention in the financial literature for its ability to model 'stylized' facts of equity returns, especially different classes of tail behavior and asymmetry (Prause 1999).§ For example, Eberlein and Keller (1995) showed that GH distribution provides a convincing fit to financial data, Aas and Hobæk Haff (2006) presented evidence that a particular case of the GH distribution, the Skew Student's t-distribution (Skew t-distribution), was superior when modeling tail behavior and skew in financial

*Corresponding author. Email: lchavezbedoya@esan.edu.pe

† The short-term investment horizon is also suitable for investors which are able to rebalance positions with a greater frequency.

‡ One of the first published criticisms of the normality assumption appeared in Mandelbrot (1963) and Fama (1965). In terms of portfolio optimization Pulley (1981), Kroll *et al.* (1984), Reid and Tew (1986) and Simaan (1993) basically agreed that the MV formulation provides a very good local approximation to the expected utility maximization problem. However, the use of the Sharpe (1966) ratio (performance measure in the MV framework) can potentially lead to wrong conclusions and paradoxes such as the ones in Hodges (1998) and Bernardo and Ledoit (2000).

§ The GH distribution is also consistent with continuous-time models where log-prices follow univariate or multivariate Lévy processes (see Eberlein and Keller (1995) for more details).

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

data and, McNeil *et al.* (2005) used it heavily in quantitative risk management applications.

Additionally, the GH distribution, and its particular cases like the Normal Inverse Gaussian (NIG), Variance Gamma (VG) and Skew *t*-distribution, has been used in portfolio optimization because of its analytical tractability and the fact that it is closed under linear transformations (the so called 'portfolio property').† Lillestøl (1998) discussed problems of risk analysis and portfolio choice in a GH context using the NIG distribution. Hu and Karcheval (2010) showed that a significant amount of available return (for a fixed level of risk) can be captured using a GH Skew *t*-distribution in contrast to classical distributions employed in the MV approach. In Madan and Yen (2008), asset returns are assumed to follow a VG distribution and they found closed-form expressions for the optimal portfolio weights using independent component analysis. Hellmich and Kassberger (2011) and Yu *et al.* (2009) applied GH distributions to portfolio optimization using Conditional Value at Risk (CVaR) as the risk measure, and the mean-variance-skewness frontier for portfolio selection when returns follow a general GH distribution was described in Mencía and Sentana (2009). Recently, Birge and Chavez-Bedoya (2016) analyzed a portfolio problem under a GH Skew *t*-distribution and assessed the quality of different Taylor series expansions of expected utility, Kwak and Pirvu (2018) solved a cumulative prospect theory problem under the same distribution, and Vanduffel and Yao (2017) presented an extension of Stein lemma useful to solve a portfolio problem under a GH distribution of returns.

The GH distribution has also motivated alternative portfolio performance measures which take into account higher moments. These new measures mitigate deficiencies of the Sharpe ratio when applied to non-elliptical distributions. In this line of research, Zakamouline and Koekabakker (2009) developed a more general performance measure (the Adjusted for Skewness and Kurtosis Sharpe Ratio) based on a univariate NIG distribution and the concept of Generalized Sharpe Ratio (Hodges 1998). Similarly, the Higher Moment Ranking (HRM) statistic in Madan and McPhail (2000) is based on the assumption of a VG distribution of returns.

Considering the importance and flexibility of the GH distribution, we analyze an asset allocation problem in which the investor applies exponential utility and asset returns follow the aforementioned distribution. This problem presents many interesting features and properties which will help us understand some relevant aspects of asset allocation under non-elliptical distributions. The following are the most important theoretical contributions to the literature contained in this paper:

(i) We find closed-form expressions for the optimal portfolio weights, generalizing the results of Lillestøl (1998), Madan and Yen (2008) and Birge and Chavez-Bedoya (2016) to the GH class. Additionally, the optimal solution motivates a performance measure that generalizes the ones of Zakamouline and Koekabakker (2009) and Madan and McPhail (2000).

(ii) We introduce two new measures (associated with mean and risk) that allow us to express the optimal solution as the one with the best trade-off between them. The mean component is related to the location of the GH distribution while risk is based on its dispersion adjusted by the skewness parameter. Also, the weight given to each of these components in the investor's objective function depends on the relationship between the parameters that control the tail behavior of a univariate GH distribution. Consequently, we prove that in this particular case it is not necessary to rely on the three-dimensional mean-variance-skewness frontier of Mencía and Sentana (2009) to characterize the optimal allocation set.

(iii) The performance decomposition in (2) makes it possible to express the optimal portfolio as an affine combination of two portfolios: one minimizing risk and the other maximizing mean given a particular level of risk. This affine combination generates an efficient frontier in this new two-dimensional space. These theoretical facts help us to analyze the behavior of the optimal portfolio and its performance as tail density changes. For example, we demonstrate that optimal performance is not monotonic since it increases as tails become lighter or heavier with respect to a particular threshold; but, heavier tails always produce more conservative allocations by increasing the optimal weight given to the minimum-risk portfolio.

(iv) Interestingly, in this framework the concept of a minimum-risk portfolio changes from the risk-free asset to one that assures that the right tail of the excess return distribution generated by the optimal portfolio is always the heaviest or, equivalently, the distribution is skewed to the right. Finally, we present evidence that the power component of the semi-heavy tails has more influence on the optimal asset allocation than the exponential component.

As mentioned earlier, it is a well-known fact that stock returns, especially when measured over short time horizons, i.e. daily or weekly, are characterized by a behavior deviating greatly from normality.‡ Consequently, the practical relevance of our paper is in showing that, when modeling asset returns, replacing a Gaussian distribution with an asymmetric GH distribution has the potential to increase the certainty equivalent excess return (CE) that can be achieved by the investor. For example, in a sample of 241 stocks of the Russell 2000 index, the optimal portfolio under a particular GH distribution generated an ex-ante CE approximately 0.035% per day (approximately 10% annual) higher than that of the

† The GH distribution also provides analytical expressions for the central moments. All these features motivate the incorporation of higher order moments in portfolio selection through a Taylor's series approximation of expected utility, as given initially in Samuelson (1970) and Tsiang (1972). In this context, Hlawitschka (1994) and Birge and Chavez-Bedoya (2016) evaluated the effect of the order on the quality of the expected utility approximation, concluding that the selection of the appropriate order expansion is merely an empirical issue.

‡ The empirical distribution of such returns is more peaked and has heavier tails than the normal distribution, which implies that extreme returns occur with a higher rate than under a Gaussian assumption. In addition, it is often skewed, having one heavy, and one semi-heavy tail (Aas and Hobæk Haff 2006).

optimal MV portfolio. Consequently, we were able to extract performance from the non-normality features of the daily returns in a better fashion than when investing in the optimal MV portfolio.

Our results can also help investors understand the importance of achieving an asset allocation with positive skewness and how this relates to the behavior of the heavy and light tail of the resulting portfolio distribution. Then, other moments of the distribution and tail behavior have a very important effect on portfolio performance, and the typical MV trade-off is not enough to predict or maximize portfolio performance. Additionally, we provide a measure of performance that is independent of the level of risk aversion or initial endowment, and this measure will characterize the potential of a particular GH distribution to generate CE. Furthermore, we completely describe the mean-risk trade-off involved when a GH distribution of returns is adopted.

The outline of the paper is as follows. In Section 2, we introduce the GH distribution as a mean-variance mixture. Assuming a GH distribution of returns, and an investor with exponential utility, the general utility maximization problem is stated and solved in Section 3. The proposed measure used to evaluate portfolio performance is also stated and studied in this section. Sensitivity analysis of the optimal portfolio and its performance with respect to the parameters controlling tail behavior is treated in Section 4. Numerical experiments illustrating some of our theoretical results and showing the in-sample and out-of-sample performance of selected portfolios are presented in Section 5. Finally, Section 6 summarizes our results and provides future extensions of this research.

2. Generalized hyperbolic distribution

In this section, we define the GH distribution by using mean-variance mixtures as presented in McNeil *et al.* (2005). The random vector $\mathbf{X}$ has a multivariate normal mean-variance mixture (in d dimensions) if

$$\mathbf{X} = \boldsymbol{\mu} + Y\boldsymbol{\gamma} + \sqrt{Y}\mathbf{A}\mathbf{Z} \tag{1}$$

where $\mathbf{Z} \sim \mathrm{N}_k(\mathbf{0}, \mathbf{I}_k)$ (multivariate normal in k dimensions with zero mean and covariance matrix $\mathbf{I}_k$); $Y \geq 0$ is a non-negative scalar-valued random variable independent of $\mathbf{Z}$; $\mathbf{A} \in \mathbb{R}^{d\times k}$, $\boldsymbol{\mu}$ (location parameter) and $\boldsymbol{\gamma}$ (skewness parameter) belong to $\mathbb{R}^d$. Therefore, $\mathbf{X}|Y = y \sim \mathrm{N}_d(\boldsymbol{\mu} + y\boldsymbol{\gamma}, y\boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^\top$ (dispersion parameter). We assume that $\operatorname{rank}(\mathbf{A}) = d \leq k$, and then $\boldsymbol{\Sigma}$ is a full-rank and positive-definite matrix. If the first two moments of the mixing variable Y are finite, then the mean vector and covariance matrix of $\mathbf{X}$ are given by

$$\begin{aligned} \mathbb{E}[\mathbf{X}] &= \boldsymbol{\mu} + \mathbb{E}[Y]\boldsymbol{\gamma}, \quad \text{and} \\ \operatorname{Cov}(\mathbf{X}) &= \mathbb{E}[Y]\boldsymbol{\Sigma} + \operatorname{Var}(Y)\boldsymbol{\gamma}\boldsymbol{\gamma}^\top. \end{aligned} \tag{2}$$

There are many possible choices for the random variable Y, but the most studied case in the literature is when $Y \sim \mathrm{GIG}(\lambda, \chi, \psi)$, i.e. Y follows a generalized inverse Gaussian (GIG) distribution with parameters λ, χ, ψ.† In such a case, $\mathbf{X}$ is said to have a GH distribution, briefly $\mathbf{X} \sim \mathrm{GH}_d(\lambda, \chi, \psi, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$ and, in general, $\mathbf{X}$ does not have an elliptical distribution and the marginal distributions are asymmetrical.

The moment-generating function (MGF) of $\mathbf{X} \sim \mathrm{GH}_d(\lambda, \chi, \psi, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$, $M_{\mathbf{X}}(\mathbf{s})$, with $\mathbf{s} \in \mathbb{R}^d$, is given by

$$\begin{aligned} M_{\mathbf{X}}(\mathbf{s}) = e^{\mathbf{s}^\top\boldsymbol{\mu}} &\left(\frac{\psi}{\psi - (\mathbf{s}^\top\boldsymbol{\Sigma}\mathbf{s} + 2\mathbf{s}^\top\boldsymbol{\gamma})}\right)^{\lambda/2} \\ &\times \frac{K_\lambda(\sqrt{\chi(\psi - (\mathbf{s}^\top\boldsymbol{\Sigma}\mathbf{s} + 2\mathbf{s}^\top\boldsymbol{\gamma}))})}{K_\lambda(\sqrt{\chi\psi})} \end{aligned} \tag{3}$$

for $\mathbf{s} \in \mathbb{R}^d$ such that $2\mathbf{s}^\top\boldsymbol{\gamma} + \mathbf{s}^\top\boldsymbol{\Sigma}\mathbf{s} < \psi$ (and where the equality case should be considered as a limit and only in the case when $\lambda < 0$) and K_λ is the modified Bessel function of the second kind with index λ.‡ In the next proposition, we provide a specification of the GH distribution, which converges to a multivariate normal distribution under certain limiting assumptions.

PROPOSITION 1 *Assume* $\mathbf{X} \sim \mathrm{GH}_d(-\frac{1}{2}, \psi, \psi, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$. *If* $\psi \to \infty$, *then* $\mathbf{X}$ *converges in distribution to a multivariate normal distribution with mean* $\boldsymbol{\mu} + \boldsymbol{\gamma}$ *and covariance matrix* $\boldsymbol{\Sigma}$.

Proof See Appendix 3.§ ■

A key property of the GH class is that it is closed under linear operations (the so called 'portfolio property'). More specifically, if $\mathbf{X} \sim \mathrm{GH}_d(\lambda, \chi, \psi, \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma})$ and $\mathbf{Y} = \mathbf{B}\mathbf{X} + \mathbf{b}$, where $\mathbf{B} \in \mathbb{R}^{k\times d}$ and $\mathbf{b} \in \mathbb{R}^k$, then

$$\mathbf{Y} \sim \mathrm{GH}_k(\lambda, \chi, \psi, \mathbf{B}\boldsymbol{\mu} + \mathbf{b}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^\top, \mathbf{B}\boldsymbol{\gamma}). \tag{4}$$

The proof of this claim is given in McNeil *et al.* (2005) and this property together with the moment-generating function in (3) will be crucial for the solution and tractability of our portfolio optimization problem.

Moroever, Aas and Hobæk Haff (2006) stated that in the tails the density of a univariate $X \sim \mathrm{GH}(\lambda, \chi, \psi, \mu, \sigma^2, \gamma)$ behaves as

$$f(x) \sim \text{const}|x|^{\lambda-1}\exp\{-\varphi|x| + \beta x\} \quad \text{as } x \to \pm\infty, \tag{5}$$

with

$$\varphi = \sqrt{\frac{\psi}{\sigma^2} + \beta^2} \quad \text{and} \quad \beta = \frac{\gamma}{\sigma^2}. \tag{6}$$

† See details of the GIG distribution in Appendix 1.

‡ For definition and useful properties of function $K_\lambda(\cdot)$, the reader is referred to Appendix 2.

§ When $\lambda = -\frac{1}{2}$ the GH distribution is often called Normal Inverse Gaussian (NIG) distribution and therefore convergence occurs for a particular specification of a NIG distribution. Another case of convergence also arises for the Variance Gamma (VG) distribution. When $\lambda > 0$ and $\chi = 0$ the GH distribution is identified as VG. Moreover, it can be proved that if $\mathbf{X} \sim \mathrm{GH}_d(\lambda, 0, 2\lambda, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$ and $\lambda \to \infty$, then $\mathbf{X}$ converges in distribution to the same multivariate normal distribution.

From (5), we have that as long as $|\beta| \neq \varphi$ both tails are semi-heavy but they behave differently.† Since β in (6) is equal to γ/σ^2, we notice that the sign of the skewness parameter γ determines which of the tails is the heaviest or the lightest one (if $\gamma > 0$, then the right tail will be the heaviest). Finally, we can observe in (5) that as λ increases or ψ decreases the density of the tails increases, i.e. they become heavier.

3. Portfolio optimization under GH distribution of returns and exponential utility

In this section, we introduce and solve our portfolio optimization problem. Then, we present an equivalent measure to assess the performance of feasible portfolios. We also decompose this performance measure and describe an efficient frontier in this framework.

3.1. Portfolio problem and performance decomposition

Assume that an investor has exponential utility of wealth, i.e. $U(W) = -e^{-aW}$, $a > 0$, and the investment opportunity set consists of N risky assets and one risk-free asset. Also, the vector of log-returns $\mathbf{R} = (R_1, \ldots, R_N)^\top$ of the risky assets follows a N-dimensional generalized hyperbolic distribution, i.e. $\mathbf{R} \sim \mathrm{GH}_N(\lambda, \chi, \psi, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$. The vector of arithmetic returns will be denoted by $\mathbf{R^a}$ and the risk-free asset return by $r_f > 0$ in the period.

Given an initial endowment $W_0 > 0$, the investor must determine portfolio weights $\mathbf{x}$ on the risky assets such that the expected utility of the next period wealth, $\mathbb{E}[U(W)]$, is maximized. Clearly $W = W_0(1 + (1 - \mathbf{x}^\top \mathbf{1})r_f + \mathbf{x}^\top \mathbf{R^a})$ and the investor's problem **(P)** can be written as

$$\begin{array}{ll} \text{Maximize} & \mathbb{E}[U(W_0(1 + (1 - \mathbf{x}^\top \mathbf{1})r_f + \mathbf{x}^\top \mathbf{R^a}))] \\ \text{s.t.} & \mathbf{x} \in \mathbb{S}_\mathbf{a}, \end{array}$$

where $\mathbb{S}_\mathbf{a} \subseteq \mathbb{R}^N$ is such that the objective is finite. To have some degree of analytical tractability for problem **(P)**, we approximate $\mathbf{R^a}$ by $\mathbf{R}$, which is equivalent to working with the linearized return of the portfolio. Under this assumption, the investor's new optimization problem **(P1)** is given by

$$\begin{array}{ll} \text{Maximize} & \mathbb{E}[U(W(\mathbf{x}))] = \mathbb{E}[-\exp\{-aW_0(1 + (1 - \mathbf{x}^\top \mathbf{1}) \\ & \qquad r_f + \mathbf{x}^\top \mathbf{R})\}] \\ \text{s.t.} & \mathbf{x} \in \mathbb{S}, \end{array}$$

where $\mathbf{R} \sim \mathrm{GH}_N(\lambda, \chi, \psi, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$, and $\mathbb{S} \subseteq \mathbb{R}^N$ is such that $E[U(W(\mathbf{x}))]$ is finite for all $\mathbf{x} \in \mathbb{S}$. The portfolio property of the GH distribution implies that the excess returns of a feasible portfolio $\mathbf{x}$ of **(P1)**, $R_{EX}(\mathbf{x}) = \mathbf{x}^\top(\mathbf{R} - r_f \mathbf{1})$, satisfies

$$R_{EX}(\mathbf{x}) \sim \mathrm{GH}(\lambda, \chi, \psi, \mathbf{x}^\top(\boldsymbol{\mu} - r_f \mathbf{1}), \mathbf{x}^\top \boldsymbol{\Sigma} \mathbf{x}, \mathbf{x}^\top \boldsymbol{\gamma}). \quad (7)$$

Using (7) and the MGF in (3), we provide the optimal solution to **(P1)** in the next theorem.

† If $\lambda = -\upsilon/2$ with $\upsilon > 0$ and $\varphi \to |\beta|$ we obtain the GH Skew Student's t-distribution which is the only case in which the GH distribution will have one heavy and one semi-heavy tail.

THEOREM 1 *Defining*

$$A = \boldsymbol{\gamma}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma}, \quad (8)$$

$$C = (\boldsymbol{\mu} - r_f \mathbf{1})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - r_f \mathbf{1}), \quad (9)$$

and

$$\theta = \sqrt{\frac{\psi + A}{C}}, \quad (10)$$

the optimal portfolio $\mathbf{x}^*$ *for the investing problem* ***(P1)*** *is given by one of the following cases*:

Case 1: *If* $\boldsymbol{\mu} - r_f \mathbf{1} = \mathbf{0}$, *then*

$$\mathbf{x}^* = \frac{1}{aW_0} \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma}. \quad (11)$$

Case 2: *If* $\boldsymbol{\mu} - r_f \mathbf{1} \neq \mathbf{0}$ *and* $\lambda \geq -1$, *the optimal portfolio* $\mathbf{x}^*$ *solves the following systems of equations*

$$\mathbf{x}^* = \frac{1}{aW_0} (\mathbf{L}(\mathbf{x}^*) \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - r_f \mathbf{1}) + \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma}), \quad (12)$$

where

$$\mathbf{L}(\mathbf{x}^*) = \frac{\sqrt{\chi(\psi + aW_0 \mathbf{x}^{*\top}(2\boldsymbol{\gamma} - aW_0 \boldsymbol{\Sigma} \mathbf{x}^*))}}{\chi} \times \frac{K_\lambda(\sqrt{\chi(\psi + aW_0 \mathbf{x}^{*\top}(2\boldsymbol{\gamma} - aW_0 \boldsymbol{\Sigma} \mathbf{x}^*))})}{K_{\lambda+1}(\sqrt{\chi(\psi + aW_0 \mathbf{x}^{*\top}(2\boldsymbol{\gamma} - aW_0 \boldsymbol{\Sigma} \mathbf{x}^*))})}. \quad (13)$$

Case 3: *If* $\boldsymbol{\mu} - r_f \mathbf{1} \neq \mathbf{0}$, $\lambda < -1$ *and* $(\chi/2(-\lambda - 1))\theta > 1$, *the optimal portfolio* $\mathbf{x}^*$ *also solves the system of equations given by (12).*

Case 4: *If* $\boldsymbol{\mu} - r_f \mathbf{1} \neq \mathbf{0}$, $\lambda < -1$ *and* $(\chi/2(-\lambda - 1))\theta \leq 1$, *then*

$$\mathbf{x}^* = \frac{1}{aW_0} (\theta \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - r_f \mathbf{1}) + \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma}). \quad (14)$$

Proof See Appendix 4.‡ ■

The optimal portfolio, $\mathbf{x}^*$, combines the risk-free asset and two risky portfolios: $\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f \mathbf{1})$ and $\boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma}$. As $\mathbf{L}(\mathbf{x}^*) > 0$ and $\theta > 0$, we always have an long position in $\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f \mathbf{1})$. Moreover, portfolio $\mathbf{x}^*$ provides a generalization of the results of Birge and Chavez-Bedoya (2016) for the skewed-t distribution, and it can also be derived using the results of Vanduffel and Yao (2017). In the latter reference, it was also stated that $\mathbf{x}^*$ will be a linear combination of the aforementioned portfolios, but our theoretical derivation motivates a new decomposition of $\mathbf{x}^*$ associated with different definitions of mean and risk arising in this particular setting. This fact will provide a better understanding of the asset allocation problem and its optimal solution and performance.

‡ The optimal solution in Case 4 must be interpreted as a limit allocation following a feasible approximation path.

3.2. Performance decomposition

We propose a performance measure to evaluate portfolios under the assumptions of exponential utility and a GH distribution of returns. The proposed measure will be called Expected Utility Transformation under GH distribution (EUT). It will help us disaggregate the performance of a portfolio into 'mean' and 'risk' components. This decomposition will allow us to perform sensitivity analysis on the optimal asset allocation with respect to the parameters of the mixing variables.

Using the MGF of $\mathbf{R}$ given in (3), we can restate problem **(P1)** as

$$\begin{aligned} \text{Maximize} \quad & \tfrac{1}{2}\text{EUT}(\mathbf{x}) = Q(\mathbf{x}) - \ln\left(\left(\frac{\psi}{\psi - KE(\mathbf{x})}\right)^{\lambda/2} \right. \\ & \left. \times \frac{K_\lambda(\sqrt{\chi(\psi - KE(\mathbf{x}))})}{K_\lambda(\sqrt{\chi\psi})}\right) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbb{S}, \end{aligned}$$

where

$$Q(\mathbf{x}) = aW_0\mathbf{x}^\top(\boldsymbol{\mu} - r_f\mathbf{1}), \tag{15}$$

$$KE(\mathbf{x}) = (aW_0)^2\mathbf{x}^\top\boldsymbol{\Sigma}\mathbf{x} - 2aW_0\mathbf{x}^\top\boldsymbol{\gamma}, \tag{16}$$

and $\mathbb{S} = \{\mathbf{x} \in \mathbb{R}^N \mid \psi - KE(\mathbf{x}) > 0\}$ when $\lambda \geq 0$ and $\mathbb{S} = \{\mathbf{x} \in \mathbb{R}^N \mid \psi - KE(\mathbf{x}) \geq 0\}$ when $\lambda < 0$. It is easy to verify that if $\mathbf{x}^1$ and $\mathbf{x}^2$ are feasible portfolios of **(P1)**, then $\mathbb{E}[U(W(\mathbf{x}^1))] \geq \mathbb{E}[U(W(\mathbf{x}^2))]$ if and only if $\text{EUT}(\mathbf{x}^1) \geq \text{EUT}(\mathbf{x}^2)$. This holds because the performance measure EUT is a monotonic transformation of expected utility.

The value of Q in (15) is a multiple of the location parameter $\mathbf{x}^\top(\boldsymbol{\mu} - r_f\mathbf{1})$ and is closely related to the expected excess return of a portfolio without considering the portion corresponding to the skewness parameter. On the other hand, *KE* is equal to a multiple of the dispersion parameter $\mathbf{x}^\top\boldsymbol{\Sigma}\mathbf{x}$ minus twice the skewness parameter $\mathbf{x}^\top\boldsymbol{\gamma}$ generated by a particular portfolio. Notice that both Q and *KE* are independent of λ, χ and ψ (the parameters of the mixing variable Y), and the latter ones control the weight given to the *KE* component in the objective. But, most important, for every feasible portfolio $\mathbf{x}$ of **(P1)** we have

$$\begin{aligned} &\frac{\partial \text{EUT}}{\partial Q} = 2 > 0, \quad \text{and} \\ &\frac{\partial \text{EUT}}{\partial KE} = -\frac{\chi}{\sqrt{\chi(\psi - KE(\mathbf{x}))}} \frac{K_{\lambda+1}(\sqrt{\chi(\psi - KE(\mathbf{x}))})}{K_\lambda(\sqrt{\chi(\psi - KE(\mathbf{x}))})} < 0. \end{aligned} \tag{17}$$

Consequently, if we have two feasible portfolios $\mathbf{x}^1$ and $\mathbf{x}^2$ such that $Q(\mathbf{x}^1) = Q(\mathbf{x}^2)$ and $KE(\mathbf{x}^1) < KE(\mathbf{x}^2)$, then portfolio $\mathbf{x}^1$ is preferred to portfolio $\mathbf{x}^2$ since $\mathbb{E}[U(W(\mathbf{x}^1))] > \mathbb{E}[U(W(\mathbf{x}^2))]$. Similarly, if $KE(\mathbf{x}^1) = KE(\mathbf{x}^2)$, the chosen portfolio will be the one with the highest value of Q.

In Proposition 1, we provided a set of conditions for a d-dimensional GH distribution to converge (in distribution) to a multivariate normal distribution. Now, we want to study the consistency of $\text{EUT}(\mathbf{x})$, $\mathbf{x}^*$ and $\text{EUT}(\mathbf{x}^*)$ with the MV approach under the same limiting assumptions.

PROPOSITION 2 *If* $\mathbf{x}$ *is a feasible portfolio for* **(P1)**, $\lambda = -\frac{1}{2}$ *and* $\chi = \psi$, *then*

$$\begin{aligned} \lim_{\psi\to\infty} \frac{1}{2}\text{EUT}(\mathbf{x}) &= aW_0\mathbf{x}^\top\mathbb{E}[\mathbf{R}_{mv} - r_f\mathbf{1}] \\ &\quad - \frac{1}{2}(aW_0)^2\,\text{Var}(\mathbf{x}^\top\mathbf{R}_{mv}), \end{aligned} \tag{18}$$

$$\lim_{\psi\to\infty} \mathbf{x}^* = \frac{1}{aW_0}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} + \boldsymbol{\gamma} - r_f\mathbf{1}) = \mathbf{x}^*_{mv}, \tag{19}$$

$$\begin{aligned} \lim_{\psi\to\infty} \text{EUT}(\mathbf{x}^*) &= \left(\frac{\mathbb{E}[\mathbf{x}^{*\top}_{mv}(\mathbf{R}_{mv} - r_f\mathbf{1})]}{\sqrt{\text{Var}(\mathbf{x}^{*\top}_{mv}\mathbf{R}_{mv})}}\right)^2 \\ &= (\boldsymbol{\mu} + \boldsymbol{\gamma} - r_f\mathbf{1})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} + \boldsymbol{\gamma} - r_f\mathbf{1}). \end{aligned} \tag{20}$$

where $\mathbf{R}_{mv}$ *follows a multivariate normal distribution with mean* $\boldsymbol{\mu} + \boldsymbol{\gamma}$ *and covariance matrix given by* $\boldsymbol{\Sigma}$.

Proof See Appendix 5. ■

The last proposition implies that our results are consistent (in the limit) with the MV approach. Expression (18) concludes that the trade-off between Q and *KE* reduces to the typical mean-variance trade-off for elliptical distributions of returns. Expression (19) implies that in the limit $\mathbf{x}^*$ coincides with the optimal MV portfolio with exponential utility and normal returns given in Ingersoll (1987), and (20) shows that the optimal EUT converges to the typical optimal Sharpe ratio squared as the distribution of returns converges to a multivariate normal distribution. Moreover, $\sqrt{\text{EUT}(\mathbf{x}^*)}$ constitutes another example of the Generalized Sharpe Ratio of Hodges (1998), it is related to the Higher Moment Rank Statistic (HMRS) of Madan and McPhail (2000) and generalizes the Adjusted for Skewness and Kurtosis Sharpe Ratio (ASKSR) of Zakamouline and Koekabakker (2009) to multiple dimensions.

It is important to recognize that EUT is closely related to the certainty equivalent return. First, the monotonic transformation given by $r(\cdot)$ in (A25) directly implies that $\mathbb{E}[U(W(\mathbf{x}))] = -\exp(-aW_0(1 + r_f + \text{CE}(\mathbf{x})))$ where $\text{CE}(\mathbf{x})$ is the certainty equivalent excess return generated by a feasible portfolio $\mathbf{x}$ of **(P1)**. Second, $\text{CE}(\mathbf{x})$ is just a multiple of $\text{EUT}(\mathbf{x})$ because the following relation holds:

$$\text{CE}(\mathbf{x}) = \frac{1}{aW_0}\frac{\text{EUT}(\mathbf{x})}{2}. \tag{21}$$

At the optimal portfolio $\mathbf{x}^*$ of **(P1)**, $\text{EUT}(\mathbf{x}^*)$ becomes independent of both the initial wealth W_0 and the risk aversion coefficient a; however, by equation (21) we observe that $\text{CE}(\mathbf{x}^*)$ depends on the aforementioned parameters. Therefore, $\text{EUT}(\mathbf{x}^*)$ measures the potential of the GH distribution $\mathbf{R}$ to generate certainty equivalent excess returns, and this is independent of the value of aW_0. Moreover, $\text{EUT}(\mathbf{x}^*)$ implies an optimal trade-off between 'mean' and 'risk' because $\frac{1}{2}\text{EUT}(\mathbf{x}^*) = Q(\mathbf{x}^*) - \text{Risk}(\mathbf{x}^*)$ where for $\mathbf{x} \in \mathbb{S}$ we define

$$\text{Risk}(\mathbf{x}) = \ln\left(\left(\frac{\psi}{\psi - KE(\mathbf{x})}\right)^{\lambda/2} \frac{K_\lambda(\sqrt{\chi(\psi - KE(\mathbf{x}))})}{K_\lambda(\sqrt{\chi\psi})}\right). \tag{22}$$

Equation (22) can be easily deduced from the optimization problem of the beginning of Section 3.2. Finally, it is straightforward to notice that CE($\mathbf{x}$) can also be expressed as a trade-off between $Q(\mathbf{x})$ and Risk($\mathbf{x}$) but adjusted by aW_0.

In the next theorem, we prove that any optimal portfolio $\mathbf{x}^*$ of **(P1)** can be expressed as an affine combination of two feasible portfolios closely related to Q and KE and that do not involve parameters λ, χ and ψ.

THEOREM 2 *If $\boldsymbol{\mu} - r_f\mathbf{1} \neq \mathbf{0}$ and $\boldsymbol{\gamma} \neq \mathbf{0}$, any optimal portfolio $\mathbf{x}^*$ of **(P1)** can be expressed as an affine combination of the following two feasible portfolios*:

$$\mathbf{x}^Q = \frac{1}{aW_0}(\nu\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}) \tag{23}$$

and

$$\mathbf{x}^{KE} = \frac{1}{aW_0}\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}, \tag{24}$$

where

$$\nu = \sqrt{\frac{A}{C}}, \tag{25}$$

and A, C were defined in (8) and (9), respectively.

Proof See Appendix 6. ■

From the proof of Theorem 2, portfolio $\mathbf{x}^{KE}$ minimizes $KE(\mathbf{x})$ and can be interpreted as a 'minimum-risk' portfolio. It also satisfies $Q(\mathbf{x}^{KE}) = B$ and $KE(\mathbf{x}^{KE}) = -A$ where the constant B is

$$B = (\boldsymbol{\mu} - r_f\mathbf{1})^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}. \tag{26}$$

On the other hand, $\mathbf{x}^Q$ can be regarded as an 'efficient' portfolio in our mean-risk objective: it maximizes $Q(\mathbf{x})$ subject to the constraint $KE(\mathbf{x}) = 0$. It is easy to show that $KE(\mathbf{x}^Q) = 0$ and $Q(\mathbf{x}^Q) = \sqrt{AC} + B$ and $Q(\mathbf{x}^Q) \geq 0$ by the Cauchy-Schwarz inequality.

Given a, W_0, $A > 0$, B, and $C > 0$, it was shown in the proof of Theorem 2 that

$$\begin{aligned}\mathbf{x}^* &= \alpha^*\mathbf{x}^Q + (1-\alpha^*)\mathbf{x}^{KE}\\ &= \frac{1}{aW_0}(\alpha^*\nu\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}),\end{aligned} \tag{27}$$

where α^* is the solution to the following equation

$$\begin{aligned}\alpha^* = \sqrt{\frac{C}{A}}\frac{\sqrt{\chi(\psi + (1-\alpha^{*2})A)}}{\chi}\\ \times \frac{K_\lambda(\sqrt{\chi(\psi + (1-\alpha^{*2})A)})}{K_{\lambda+1}(\sqrt{\chi(\psi + (1-\alpha^{*2})A)})}\end{aligned} \tag{28}$$

when $\lambda \geq -1$ or when $\lambda < -1$ and $(\chi/2(-\lambda - 1))\theta > 1$ with θ given by (10). Moreover, $\alpha^* = \sqrt{1 + \psi/A}$ when $\lambda < -1$ and $(\chi/2(-\lambda - 1))\theta \leq 1$.

Finally, another interesting portfolio to analyze is:

$$\mathbf{x}^B = \frac{1}{aW_0}(\theta\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}), \tag{29}$$

where θ is given by (10). Portfolio $\mathbf{x}^B$ is located at the boundary of the feasible region and is infeasible when $\lambda \geq 0$, but it is feasible (as a limit) when $\lambda < 0$. It produces the supremum for both $Q(\mathbf{x}^*)$ and $KE(\mathbf{x}^*)$. Those values are given by $Q(\mathbf{x}^B) = \sqrt{C(\psi + A)} + B$ and $KE(\mathbf{x}^B) = \psi$.

3.3. The $Q - KE$ frontier

In this section, we study the portfolios that minimize KE for a certain level of Q. Those portfolios will generate the $Q - KE$ frontier, which resembles in some feature the MV frontier. Also, we show that the optimal portfolio of our utility maximization problem lies in the 'efficient' part of this frontier.

THEOREM 3 *Assuming $\boldsymbol{\mu} - r_f\mathbf{1} \neq \mathbf{0}$, $\boldsymbol{\gamma} \neq \mathbf{0}$ and defining*

$$\mathbf{x}^*(Q) = \operatorname{argmin}\{KE(\mathbf{x}) | Q(\mathbf{x}) = Q\}, \tag{30}$$

the function $Q \to KE(\mathbf{x}^*(Q))$ is a parabola given by

$$KE(\mathbf{x}^*(Q)) = KE(Q) = \frac{(Q-B)^2}{C} - A. \tag{31}$$

We call this parabola the $Q - KE$ frontier.

Proof See Appendix A.1. ■

Notice that portfolio $\mathbf{x}^*(Q)$ given by (30) is not necessarily feasible for problem **(P1)** since not all levels of KE are attainable for that problem. In the next theorem, we show that any portfolio $\mathbf{x}$ in the $Q - KE$ frontier (the corresponding values of $Q(\mathbf{x})$ and $KE(\mathbf{x})$ satisfy (31)) can be expressed as an affine combination of portfolios $\mathbf{x}^Q$ and $\mathbf{x}^{KE}$. This result is similar to that of Theorem 2.

THEOREM 4 *Assuming $\boldsymbol{\mu} - r_f\mathbf{1} \neq \mathbf{0}$, $\boldsymbol{\gamma} \neq \mathbf{0}$, if $\mathbf{x}$ is a portfolio located in the $Q - KE$ frontier, then $\mathbf{x} = \alpha\mathbf{x}^Q + (1-\alpha)\mathbf{x}^{KE}$ for some $\alpha \in \mathbb{R}$ with portfolios $\mathbf{x}^Q$ and $\mathbf{x}^{KE}$ given by (23) and (24), respectively.*

Proof See Appendix A.2. ■

From Theorem 4, the feasible portfolios that belong to the $Q - KE$ frontier have $\alpha \in [-\sqrt{1 + \psi/A}, \sqrt{1 + \psi/A}]$ when $\lambda < 0$, and $\alpha \in (-\sqrt{1 + \psi/A}, \sqrt{1 + \psi/A})$ when $\lambda \geq 0$. This fact can be verified if we find the boundary values for α with the conditions $KE(\mathbf{x}) \leq \psi$ or $KE(\mathbf{x}) < \psi$ when $\mathbf{x} = \alpha\mathbf{x}^Q + (1-\alpha)\mathbf{x}^{KE}$.

The optimal solution $\mathbf{x}^*$ to **(P1)** is going to be located somewhere in the $Q - KE$ frontier because the portfolios in such frontier have the best possible Q vs. KE trade-off. Next, we provide the definition of the $Q - KE$ *efficient* frontier and the proof that $\mathbf{x}^*$ will be located in that region.

DEFINITION 1 *Assuming $\boldsymbol{\mu} - r_f\mathbf{1} \neq \mathbf{0}$, $\boldsymbol{\gamma} \neq \mathbf{0}$, the $Q - KE$ efficient frontier is the truncated parabola given by*

$$KE(Q) = \frac{(Q-B)^2}{C} - A, \tag{32}$$

with

$$B \leq Q < \sqrt{C(\psi + A)} + B, \quad \text{if } \lambda \geq 0, \tag{33}$$

$$B \leq Q \leq \sqrt{C(\psi + A)} + B, \quad \text{if } \lambda < 0. \tag{34}$$

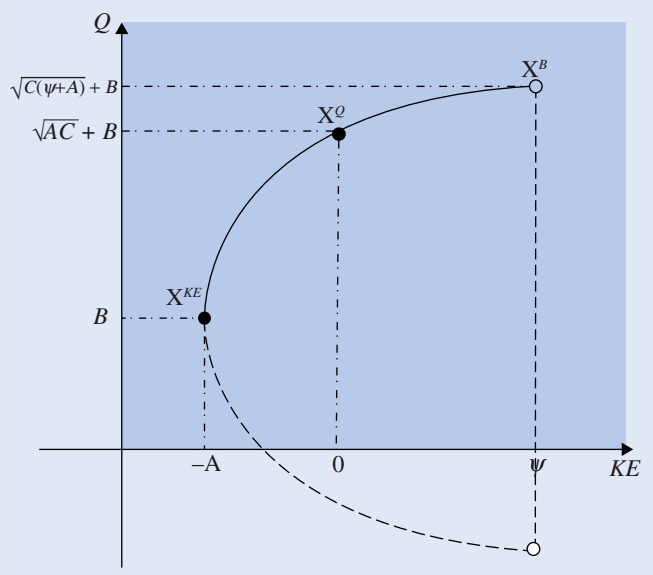

Figure 1. $Q-KE$ Efficient frontier and feasible portfolios when $\lambda \geq 0$.

THEOREM 5 *Assuming* $\boldsymbol{\mu} - r_f\mathbf{1} \neq \mathbf{0}$, $\boldsymbol{\gamma} \neq \mathbf{0}$ *and that* $\mathbf{x}^*$ *is the optimal solution to* ***(P1)***, $(Q(\mathbf{x}^*), KE(\mathbf{x}^*))$ *is located in the* $Q-KE$ *efficient frontier.*

Proof See Appendix A.3. ■

Figure 1 shows a possible $Q-KE$ efficient frontier when $\lambda \geq 0$.† In the plot, the dotted line in the parabola represents the 'on-efficient' part of the $Q-KE$ frontier. All the feasible portfolios for problem **(P1)** are in the area inside the parabola and the line $KE = \psi$ (the line is not included in the set of feasible portfolios when $\lambda \geq 0$). We also locate portfolios $\mathbf{x}^Q$, $\mathbf{x}^{KE}$ and $\mathbf{x}^B$.

Using Theorem 4, we can observe that all portfolios in the $Q-KE$ efficient frontier will be of the form $\mathbf{x} = \alpha\mathbf{x}^Q + (1-\alpha)\mathbf{x}^{KE}$ with $\alpha \in [0, \sqrt{1+\psi/A}]$ when $\lambda < 0$, and with $\alpha \in [0, \sqrt{1+\psi/A})$ when $\lambda \geq 0$. In the next proposition, we study the conditions under which portfolio $\rho\mathbf{x}$ belongs to the $Q-KE$ frontier when $\mathbf{x}$ is on such frontier and $\rho \in \mathbb{R}$.

PROPOSITION 3 *Let* $\boldsymbol{\mu} - r_f\mathbf{1} \neq \mathbf{0}$, $\boldsymbol{\gamma} \neq \mathbf{0}$, *and* $\mathbf{x}$ *be a portfolio in the* $Q-KE$ *frontier. If* $AC = B^2$, *then portfolio* $\rho\mathbf{x}$ *belongs to the* $Q-KE$ *frontier for all* $\rho \in \mathbb{R}$.

Proof See Appendix A.4. ■

There are some important observations that can be derived from the previous proposition and the condition $AC = B^2$. Under the last condition, the $Q-KE$ efficient frontier is given by the following equation,

$$KE(Q) = \frac{Q^2 - 2QB}{C}. \tag{35}$$

Consequently, portfolio $\mathbf{x} = \mathbf{0}$ (investing only in the risk-free asset) is located in the $Q-KE$ frontier since $Q(\mathbf{0}) = 0$ and $KE(\mathbf{0}) = 0$. Additionally, if $\mathbf{x}$ is a portfolio satisfying (35) with $Q(\mathbf{x}) \neq 0$, then all the portfolios in the $Q-KE$ frontier can be expressed as multiples of $\mathbf{x}$, i.e. $\rho\mathbf{x}$ spans the $Q-KE$ frontier. In the case of $B < 0$, portfolio $\mathbf{x}^Q$ given by (23) satisfies $Q(\mathbf{x}^Q) = 0$, and therefore has the same $Q-KE$ trade-off as the portfolio that is fully invested in the risk-free asset.

4. Sensitivity analysis

The excess return of the optimal portfolio, $R_{EX}(\mathbf{x}^*)$, will follow a univariate GH distribution given by $R_{EX}(\mathbf{x}^*) \sim \text{GH}(\lambda, \chi, \psi, \mu^*, \sigma^{*2}, \gamma^*)$ where

$$\mu^* = \mathbf{x}^{*\top}(\boldsymbol{\mu} - r_f\mathbf{1}), \quad \sigma^{*2} = \mathbf{x}^{*\top}\boldsymbol{\Sigma}\mathbf{x}^*, \quad \text{and} \quad \gamma^* = \mathbf{x}^{*\top}\boldsymbol{\gamma}. \tag{36}$$

Moreover, using (5) and φ^* in (38), we can express the tail density of $R_{EX}(\mathbf{x}^*)$ as

$$f(x) \sim \text{const}|x|^{\lambda-1}\exp\{-\varphi^*|x| + \beta^* x\} \quad \text{as } x \to \pm\infty, \tag{37}$$

with

$$\varphi^* = \sqrt{\frac{\psi}{\sigma^{*2}} + \beta^{*2}} \quad \text{and} \quad \beta^* = \frac{\gamma^*}{\sigma^{*2}}. \tag{38}$$

Having $\beta^* > 0$ at the optimal allocation could be desirable because it will assure that the distribution is skewed to the right and the right tail is the heaviest. The case of $\beta^* < 0$ is also interesting since the left tail is the heaviest but the portfolio will be aggressive in the sense that it will tend to invest more in $\mathbf{x}^Q$ than in $\mathbf{x}^{KE}$.‡

By Theorem 2, $\mathbf{x}^*$ is also an affine combination of portfolios $\mathbf{x}^Q$ and $\mathbf{x}^{KE}$ given by (23) and (24), respectively. Now, if we express $\mathbf{x}^*$ as a function of the parameters of the mixing variable Y, we have:

$$\mathbf{x}^*(\lambda, \chi, \psi) = \alpha^*(\lambda, \chi, \psi)\mathbf{x}^Q + (1 - \alpha^*(\lambda, \chi, \psi))\mathbf{x}^{KE}. \tag{39}$$

From tail density (37), λ and ψ play an important role in the tail behavior of $R_{EX}(\mathbf{x}^*)$ because as λ increases and ψ decreases the tails of the distribution become heavier.§ Consequently, under these circumstances, it will be reasonable to optimally invest more in 'minimum-risk' portfolio $\mathbf{x}^{KE}$ and less in efficient portfolio $\mathbf{x}^Q$. Proposition 4 will confirm our intuition as it studies changes in the optimal weights given to $\mathbf{x}^Q$ and $\mathbf{x}^{KE}$ due to changes in ψ and λ.

PROPOSITION 4 *Let* $\boldsymbol{\mu} - r_f\mathbf{1} \neq \mathbf{0}$, $\boldsymbol{\gamma} \neq \mathbf{0}$, $\chi > 0$ *and* $\psi > 0$. *For a fixed* $\lambda \geq -1$, *the function* $\psi \to \alpha^*(\lambda, \chi, \psi)$ *is strictly increasing in* ψ. *Also, if* $\lambda > -1 - \chi\theta/2$ *with* $\theta = \sqrt{(\psi + A)/C}$, *the function* $\lambda \to \alpha^*(\lambda, \chi, \psi)$ *is strictly decreasing in* λ.

† If $\lambda < 0$, then portfolio $\mathbf{x}^B$ (as a limit) will be feasible in Figure 1.

‡ We have that $\gamma^* = \nu(\sqrt{AC} + \alpha^* B)$ and when $B < 0$ the condition $\alpha^* \leq 1$ implies $\gamma^* > 0$.

§ In tail density (37), parameter λ controls the power component and parameter ψ the exponential one. However, coefficient φ^* in (38) not only is affected by ψ but also by the choice of $\mathbf{x}^*$ through parameters σ^{2*} and γ^* (which in turn depend on ψ and λ). This interaction and the behavior of φ^* will be addressed later in the section.

Proof See Appendix 8.† ■

Under the assumptions of Proposition 4, we have $\lim_{\lambda\to\infty}\alpha^*(\lambda,\chi,\psi)=0$; thus in the limit we tend to invest only in $\mathbf{x}^{KE}$. However, this fact does not always happen when changing ψ because $\lim_{\psi\to 0^+}\alpha^*(\lambda,\chi,\psi)=\bar{\alpha}$ where $\bar{\alpha}\in(0,1)$ and in the limit we will have a long position of $\mathbf{x}^{KE}$ in $\mathbf{x}^*$. In the following proposition, we will study the effects on optimal performance due to changes in parameters ψ and λ.

PROPOSITION 5 *Let* $\boldsymbol{\mu}-r_f\mathbf{1}\neq\mathbf{0}$, $\boldsymbol{\gamma}\neq\mathbf{0}$, $\chi>0$, $\mathrm{EUT}^*(\lambda,\chi,\psi)=\mathrm{EUT}(\mathbf{x}^*(\lambda,\chi,\psi))$ *and* θ *in (10). If* $\lambda\geq -1$ *is fixed, then there exists* $\widehat{\psi}>0$ *such that* $\alpha^*(\lambda,\chi,\widehat{\psi})=1$ *and the function* $\psi\to\mathrm{EUT}^*(\lambda,\chi,\psi)$ *is strictly decreasing for* $0<\psi<\widehat{\psi}$ *and strictly increasing for* $\psi>\widehat{\psi}$. *Moreover, for a fixed* $\psi>0$ *and* $\lambda>-1-\chi\theta/2$, *there exists* $\widehat{\lambda}\in\mathbb{R}$ *such that* $\alpha^*(\widehat{\lambda},\chi,\psi)=1$ *and the function* $\lambda\to\mathrm{EUT}^*(\lambda,\chi,\psi)$ *is strictly decreasing for* $\lambda<\widehat{\lambda}$ *and strictly increasing for* $\lambda>\widehat{\lambda}$.

Proof See Appendix 9. ■

Proposition 5 states that the optimal performance, in general, is not a monotonic function of ψ or λ, and there exist thresholds $\widehat{\psi}$ and $\widehat{\lambda}$ (ones that make $\mathbf{x}^*=\mathbf{x}^Q$) determining when $\mathrm{EUT}^*(\lambda,\chi,\psi)$ will be a decreasing or increasing function of ψ or λ. Surprisingly, we can infer that optimal performance improves as tail density tend to become heavier or lighter with respect to a particular threshold.

In the case of $\lambda,\widehat{\lambda}$ of Proposition 5 makes $\mathbf{x}^*=\mathbf{x}^Q$ and then $KE(\mathbf{x}^*)=0$. If λ increases above threshold $\widehat{\lambda}$, then $KE(\mathbf{x}^*)$ reduces to a negative value, $\mathbf{x}^*$ starts having a long position in $\mathbf{x}^{KE}$, and the gain in optimal performance due to the reduction in KE turns out to be more significant than the loss in performance due to the corresponding reduction in Q. This behavior intensifies as we consider even greater values of λ. Then, we are rewarded the most for choosing conservative portfolios in the presence of distributions with a particular tail behavior. Moreover, the most conservative portfolio is $\mathbf{x}^{KE}$ (occurs when $\lambda\to\infty$) and it is such that $\gamma^*=\boldsymbol{\gamma}^\top\mathbf{x}^{KE}=(1/aW_0)A>0$, so we invest in an optimal allocation with a positive skewness parameter and assure that the heavier tail is always the one of the right. The case of ψ (for a fixed $\lambda\geq -1$) is analogous but as $\psi\to 0^+$ there is no guarantee to converge to portfolio $\mathbf{x}^{KE}$ but to one with positive γ^*.

The exponential component of tail density (37) is determined by parameters φ^* and β^* in (38), and they depend also on ψ and λ. Next we study changes on the decay coefficients of the exponential component of the tail density of $R_{EX}(\mathbf{x}^*)$ due to changes in the aforementioned parameters. As usual, we introduce the notation $\varphi^*(\lambda,\chi,\psi)$ and $\beta^*(\lambda,\chi,\psi)$ to make the dependence explicit.

PROPOSITION 6 *Let* $\boldsymbol{\mu}-r_f\mathbf{1}\neq\mathbf{0}$, $\boldsymbol{\gamma}\neq\mathbf{0}$, $\chi>0$, $\psi>0$, $\lambda\in\mathbb{R}$, B *given by (26) and* θ *in (10). If* $B>0$, *then the heaviest tail of the distribution of* $R_{EX}(\mathbf{x}^*)$ *decays as*

$$f^*(x)\sim \mathrm{const}|x|^{\lambda-1}\exp\{\vartheta_H^*(\lambda,\chi,\psi)x\}\quad \text{as } x\to+\infty, \tag{40}$$

and the lightest as

$$f^*(x)\sim \mathrm{const}|x|^{\lambda-1}\exp\{\vartheta_L^*(\lambda,\chi,\psi)|x|\}\quad \text{as } x\to+\infty, \tag{41}$$

where

$$\vartheta_H^*(\lambda,\chi,\psi)=-\varphi^*(\lambda,\chi,\psi)+\beta^*(\lambda,\chi,\psi), \tag{42}$$

$$\vartheta_L^*(\lambda,\chi,\psi)=-\varphi^*(\lambda,\chi,\psi)-\beta^*(\lambda,\chi,\psi), \tag{43}$$

with $\varphi^*(\lambda,\chi,\psi)$ *and* $\beta^*(\lambda,\chi,\psi)>0$ *given by (38). Moreover, functions* $(\psi,\lambda)\to\vartheta_i^*(\lambda,\chi,\psi)$ *with* $i\in\{\mathrm{H},\mathrm{L}\}$ *are strictly decreasing in* ψ *for a fixed* $\lambda\geq -1$; *and also they are strictly decreasing in* λ *for* $\lambda>-1-\chi\theta/2$ *and a fixed* $\psi>0$.

Proof See Appendix 10. ■

We can notice from the proof of Proposition 6 that $B>0$ implies $\gamma^*>0$, i.e. the optimal allocation always has positive skewness. Additionally, since $\vartheta_i^*(\lambda,\chi,\psi)$ is a decreasing function of ψ (with $\lambda\geq -1$ fixed) both tails become lighter, generating less aggressive allocations. This behavior is also observed when λ increases (with $\psi>0$ fixed); but, the power component of tail density also increases and is the one effectively driving the optimal allocation to be more conservative. In the case of $B<0$ we do not observe the same monotonicity properties. For example, β^* goes from negative to positive as λ increases and from positive to negative as ψ increases. In the first case, the heaviest tail switches from the left to the right, and in the second case it switches from the right to the left. Moreover, there are no monotonicity properties for coefficients ϑ_H^* and ϑ_L^*.

5. Numerical examples

5.1. Data description and general considerations

We consider the set of stocks of the Russell 2000 Index‡ with a market capitalization between 20 and 2.5 billion USD as of 01/08/2020 and complete daily price data from 05/02/2014 to 08/01/2020. The number of stocks satisfying these conditions is 241. In order to calibrate the appropriate multivariate distributions, we use daily logarithmic excess returns from 05/02/2014 to 05/07/2019. The number of observations in this calibration period is 1262 which corresponds roughly to five years of trading data. We will fit a multivariate normal distribution (using MLE) to the data, as well as multivariate NIG, VG, and GH distributions.§ We

† We can extend the first result when $\lambda>-1-\chi\theta/2$ as the solution to **(P1)** is still in the interior of the feasible region; but, we wanted to avoid conditions that introduce dependence among parameters λ and ψ. We also avert feasibility issues arising when $\lambda\leq -1-\chi\theta/2$ and $\alpha^*=\sqrt{1+\psi/A}$ although the sensitivity analysis in this case is straightforward.

‡ This market-cap weighted index measures the performance of approximately 2000 smallest-cap US companies in the Russell 3000 Index, which consists of 3000 of the largest US stocks.

§ In the case of NIG we have $\lambda=-0.5$ and assume for calibration that $\chi=\psi$. For VG we have $\chi\to 0$ and $\lambda>0$, and assume $\psi=2\lambda$. Finally, for GH we consider the case in which $\lambda<0$, $\chi>0$ and $\psi\to 0$.

Table 1. Statistics for daily excess returns of ARES, BCO, UNF, SSD, UFPI, and DAN from 05/02/2014 to 05/07/2019.

Statistic	ARES	BCO	UNF	SSD	UFPI	DAN
Mean (%)	0.05	0.11	0.07	0.08	− 0.02	0.06
Standard deviation (%)	1.39	1.86	1.46	1.98	2.28	1.95
Skewness	0.13	0.88	2.23	1.12	− 0.45	− 0.14
Kurtosis	11.17	16.10	36.53	14.66	6.51	7.41
Minimum (%)	− 9.45	− 14.02	− 9.73	− 9.29	− 14.19	− 10.28
Maximum (%)	9.69	15.02	19.56	17.89	9.89	10.60

Table 2. Fitting statistics of the six assets for multivariate normal, NIG, VG, and GH distributions.

Distribution	Log-likelihood (LL)	Number of Param.	AIC
Mulivariate Normal	− 14 465.96	27	28 985.91
Asymmetric NIG	− 13 449.66	34	26 967.32
Symmetric NIG	− 13 459.26	28	26 974.52
Asymmetric VG	− 13 518.15	34	27 104.30
Symmetric VG	− 13 527.53	28	27 111.05
Asymmetric GH	− 13 413.25	35	26 896.50
Symmetric GH	− 13 423.23	29	26 904.46

also consider the elliptical (symmetrical) and non-elliptical (asymmetrical) cases for NIG, VG, and GH. Recall that the symmetrical cases assume $\boldsymbol{\gamma} = \mathbf{0}$. The calibration of the corresponding generalized hyperbolic distributions was performed using the Expectation-Maximization (EM) algorithm described in McNeil *et al.* (2005). The out-of-sample performance of selected portfolios will be assessed in the period 05/08/2019–08/01/2020 (170 daily observations).

Our numerical experiments will be divided in two parts. In the first one (Section 5.2), we consider an investment problem with 6 risky assets (selected from the 241) and a risk-free asset. The six stocks were selected such that $B > 0$ for all the calibrated asymmetric GH, NIG and VG distributions. Then, we will find the optimal solution to **(P1)** and verify the results of sensitivity analysis given by Propositions 4, 5 and 6 of Section 4. In the second part (Section 5.3), we consider the complete sample of 241 stocks and study the in-sample and out-of-sample performance of three portfolios: (i) the optimal solution to **(P1)**, (ii) the MV portfolio, (iii) and the optimal portfolio assuming a symmetric GH distribution. Finally, we include a short numerical analysis of Proposition 6 (tail behavior) when $B < 0$.

5.2. Six risky assets: sensitivity analysis

The six risky assets considered in this section are: Ares Management (ARES), Brink's (BCO), Unifirst (UNF), Simpson Manufacturing Company (SSD), Universal Forest Products Inc. (UFPI), and Dana Inc. (DAN). In Table 1, we present some basic statistics of the aforementioned assets: mean excess return, standard deviation, normalized skewness and kurtosis, minimum and maximum daily returns. As we can observe, the excess returns appear to be non-gaussian presenting cases with positive and negative skewness, and kurtosis largely greater than three (with minimum value of 6.51 and as high as 36.53).

Table 2 shows the fitting results: Log-likelihood (LL), number of fitted parameters, and Akaike Information Criteria (AIC). We can observe that the proposed multivariate distributions perform better than the multivariate normal distribution in both LL and AIC. As expected, the GH case provides the best fitting; however, the difference between the symmetrical and asymmetrical cases is almost negligible in terms of AIC. The NIG distribution fits the data better than the corresponding VG distribution.

Next, we solve our portfolio optimization problem with six risky assets assuming a NIG distribution of returns, i.e. fixing $\lambda = -\frac{1}{2}$, and we assume this distribution as reference. This choice simplifies the presentation of results and the computation of optimal portfolios,† and also the fitting difference with respect to the GH distribution is very small. From the calibration we have: $A = 0.0102$, $C = 0.0044$, $B = 0.0037$, where the fitted values of χ and ψ were both equal to 1.019. We also assume $aW_0 = 1$ and for the optimal portfolio under NIG we have: $\text{EUT}(\mathbf{x}^*) = 0.0220$, $Q(\mathbf{x}^*) = 0.0081$, $KE(\mathbf{x}^*) = -0.0058$, $\mathbf{L}(\mathbf{x}^*) = 1.0028$ and $\alpha^* = 0.6592$. With regard to the exponential tail behavior of Proposition 6, at the optimal allocation we have $\beta^* = 0.6351$, $\vartheta_H = -6.2015$ and $\vartheta_L = -7.4645$.

Figure 2 shows the values of α^*, $\text{EUT}(\mathbf{x}^*)$, $Q(\mathbf{x}^*)$, $\text{Risk}(\mathbf{x}^*) = \frac{1}{2}\text{EUT}(\mathbf{x}^*) - Q(\mathbf{x}^*)$, density functions of $R_{EX}(\mathbf{x}^*)$, β^* and ϑ_i^* for different values of λ and keeping the other parameters of the fitted NIG distribution fixed at their calibrated values.‡ Also, the results for $\lambda = -\frac{1}{2}$ (original calibrated NIG) are shown as a black dot in the plots. Panels (a) and (b) corroborate the theoretical results of Propositions 4 and 5 for changes in λ: α^* decreases towards zero and $\text{EUT}(\mathbf{x}^*)$ decreases until $\widehat{\lambda} = -1.1$ (in this point we achieve the minimum EUT which is equal to 0.02076) and

† Under NIG distribution a closed-form expression for the optimal portfolio weights is available. If returns follow $\mathbf{R} \sim \text{NIG}_N(\chi, \psi, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$, the optimal portfolio for **(P1)** is given by:

$$\mathbf{x}^* = \frac{1}{aW_0}\left(\sqrt{\frac{A+\psi}{C+\chi}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}\right).$$

‡ Recall that α^*, $\text{EUT}(\mathbf{x}^*)$, $Q(\mathbf{x}^*)$, $\text{Risk}(\mathbf{x}^*) = \frac{1}{2}\text{EUT}(\mathbf{x}^*) - Q(\mathbf{x}^*)$ are independent of the choice of aW_0; but, this is not the case for $R_{EX}(\mathbf{x}^*)$, β^* and ϑ_i^*.

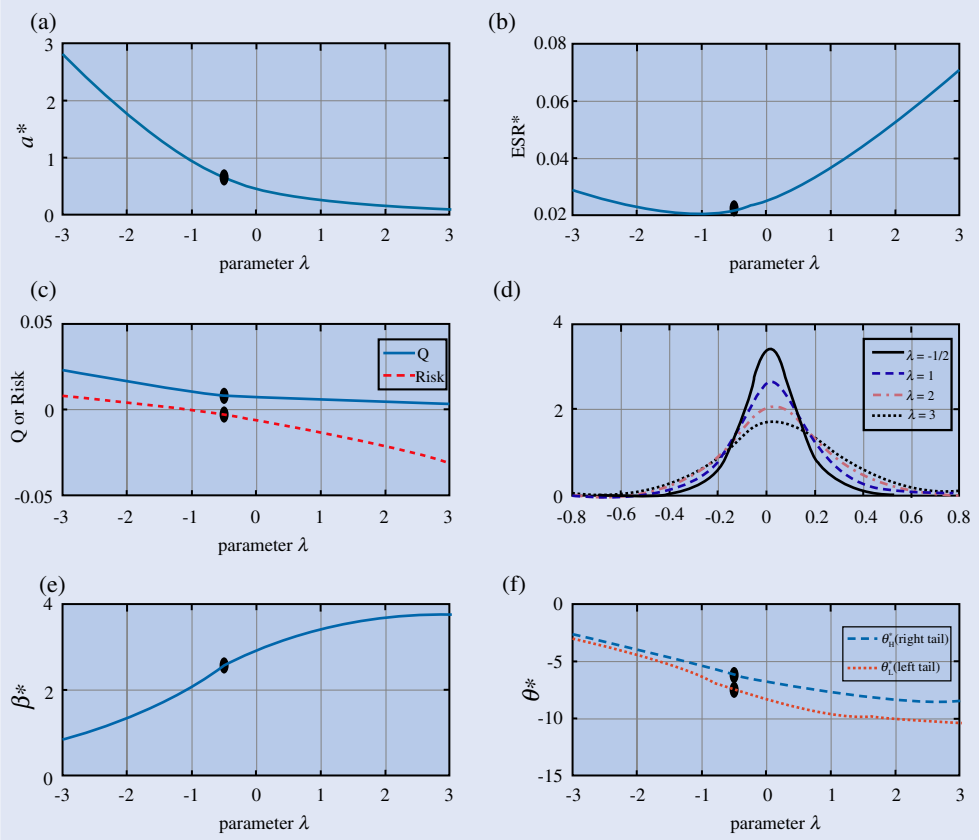

Figure 2. For six risky assets and different values of parameter λ: panel (a) shows α^* (the optimal position in $\mathbf{x}^Q$), panel (b) displays the optimal performance EUT($\mathbf{x}^*$), panel (c) contains the values of $Q(\mathbf{x}^*)$ and Risk($\mathbf{x}^*$) $= \frac{1}{2}$EUT($\mathbf{x}^*$) $- Q(\mathbf{x}^*)$, panel (d) presents some density functions of the NIG distribution of the excess returns (in %) generated by $\mathbf{x}^*$, panel (e) exhibits the values of $\beta^* = \frac{\gamma^*}{\sigma^{2*}}$ and panel (d) shows the behavior of coefficients ϑ_H^* (heavy/right tail) and ϑ_L^* (light/left tail). The black dot in the plots represents the results for the calibrated NIG distribution ($\lambda = -\frac{1}{2}$). We have assumed $aW_0 = 1$.

then it increases indefinitely. Notice that the calibrated NIG ($\lambda = -\frac{1}{2}$) generates one of the lowest optimal performances observed in the figure. Panel (c) shows the optimal trade-off between $Q(\mathbf{x}^*)$ and Risk($\mathbf{x}^*$), and for $\widehat{\lambda} = -1.1$ they should have the same slope. Panel (d) shows some density functions of the GH distribution of the excess returns (in %) generated by $\mathbf{x}^*$, and, as λ increases the distribution becomes skewed to the right and its tails are heavier. Panel (e) displays β^* which is strictly increasing as expected from the proof of Proposition 6. Finally, panel (f) shows coefficients ϑ_H^* (heavy/right tail) and ϑ_L^* (light/left tail) and, as expected from Proposition 6, they are strictly decreasing in λ and $\vartheta_H^* > \vartheta_L^*$ holds.

Similarly, Figure 3 presents the values of α^*, EUT($\mathbf{x}^*$), $Q(\mathbf{x}^*)$, Risk($\mathbf{x}^*$) $= \frac{1}{2}$EUT($\mathbf{x}^*$) $- Q(\mathbf{x}^*)$, density functions of $R_{EX}(\mathbf{x}^*)$, β^* and ϑ_i^* for different values of ψ keeping the other parameters of the fitted NIG distribution fixed at their calibrated values. Also, the results for $\psi = 1.019$ (original calibrated NIG) are shown as a black dot. Panels (a) and (b) corroborate the theoretical results of Propositions 4 and 5 for changes in ψ: α^* increases and EUT($\mathbf{x}^*$) decreases until $\widehat{\psi} = 2.36$ and then it increases (for $\psi = \widehat{\psi}$ we have EUT $= 0.02076$ and for $\psi = 3$ we have EUT $= 0.02085$). Panel (c) displays the optimal trade-off between $Q(\mathbf{x}^*)$ and Risk($\mathbf{x}^*$) and both are increasing. Panel (d) shows some density functions of the GH distribution of the excess returns (in %) generated by $\mathbf{x}^*$; and, as ψ increases the tails of the distribution become lighter and the skewness of the distribution slightly diminishes. Panel (e) shows β^* and it is strictly decreasing (as expected from the proof of Proposition 6). Finally, panel (f) shows exponents ϑ_H^* (heavy/right tail) and ϑ_L^* (light/left tail) and, from Proposition 6 they are strictly decreasing in ψ, $\vartheta_H^* > \vartheta_L^*$ holds, and when ψ approaches zero we have $\vartheta_H^* = 0$.

5.3. *Complete sample: optimal solution, performance and tail behavior*

Table 3 shows statistics of the daily excess returns of the 241 stocks during the period 05/02/2014 – 05/07/2019. We consider different percentiles to summarize the information for each statistic. The first six lines of Table 3 were constructed using 241 observations each; but, the line 'Correlation Coefficient' was constructed using the sample correlation matrix (28,920 observed correlations). It is interesting to notice that the mean excess return is negative only in the 5% percentile and, more precisely, only 23 out of 241 stocks exhibited negative mean. The skewness starts to be positive in the 75% percentile and only 80 stocks have a positive value. Moreover, the kurtosis is consistently greater than 4 with a median of 9.75. The correlation coefficient fluctuates between 0.1 and 0.46 in 90% of the cases. Clearly, the daily data deviates

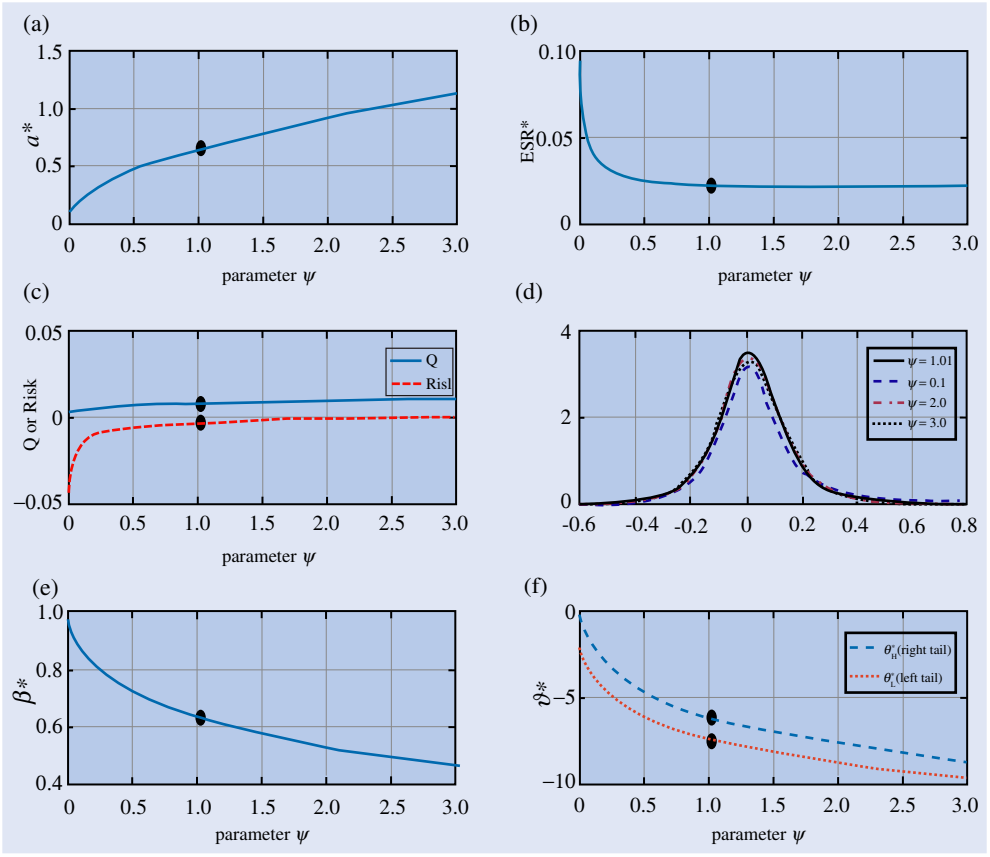

Figure 3. For six risky assets and different values of parameter ψ: panel (a) shows α^* (the optimal position in $\mathbf{x}^Q$), panel (b) displays the optimal performance EUT($\mathbf{x}^*$), panel (c) contains the values of $Q(\mathbf{x}^*)$ and Risk($\mathbf{x}^*$) $= \frac{1}{2}$EUT($\mathbf{x}^*$) $- Q(\mathbf{x}^*)$, panel (d) presents some density functions of the NIG distribution of the excess returns (in %) generated by $\mathbf{x}^*$, panel (e) exhibits the values of $\beta^* = \frac{\gamma^*}{\sigma^{2*}}$ and panel (d) shows the behavior of coefficients ϑ_H^* (heavy/right tail) and ϑ_L^* (light/left tail). The black dot in the plots represents the results for the calibrated NIG distribution ($\psi = 0.67$). We have assumed $aW_0 = 1$.

Table 3. Statistics for daily excess returns of the 241 risky assets from 05/02/2014 to 05/07/2019.

Statistic/Percentile	5%	10%	25%	50%	75%	90%	95%
Mean (%)	− 0.02	0.00	0.02	0.04	0.07	0.10	0.12
Standard deviation. (%)	1.17	1.29	1.54	1.87	2.33	3.01	3.69
Skewness	− 2.34	− 1.54	− 0.59	− 0.26	0.23	0.80	1.16
Kurtosis	4.39	4.62	6.06	9.75	16.16	31.31	51.26
Minimum (%)	− 38.49	− 27.61	− 18.92	− 12.87	− 8.42	− 6.55	− 6.01
Maximum (%)	4.52	5.24	7.36	12.47	17.91	25.51	32.21
Correlation coefficient	0.10	0.13	0.18	0.24	0.31	0.40	0.46

Note: The line 'Correlation Coefficient' was constructed using the 28 920 sample correlations.

greatly from normality justifying the use of GH multivariate distributions.

Table 4 shows the fitting information when the multivariate normal and the symmetric and asymmetric NIG, VG and GH distributions are fitted to the calibration data. First, we observe an important improvement when the multivariate normal is replaced by the other distributions. Second, the differences between the fitting statistics of the symmetric and asymmetric versions of NIG, VG and GH are small, but the non-symmetric versions tend to represent the data better. Consequently, and for the same reasons given in Section 5.2, we will assume an asymmetric NIG to model excess returns. From the calibration to NIG we have: $A = 0.9535$, $C = 0.9415$, $B = -0.8901$, where the fitted values of χ and ψ were both equal to 3.2127. This calibrated distribution of excess returns will be denoted as $\mathbf{R}_{\mathrm{E}}$.

Next, we study the behavior and performance of three portfolios: $\mathbf{x}^*$, the optimal solution to **(P1)** assuming the calibrated asymmetric NIG distribution $\mathbf{R}_{\mathrm{E}}$; $\mathbf{x}_{\mathrm{mv}}$, the optimal MV portfolio; and, $\mathbf{x}_{\mathrm{S}}$, the optimal solution to **(P1)** assuming a symmetric NIG distribution. In the case of $\mathbf{x}_{\mathrm{mv}}$, we calibrate a multivariate normal distribution to the excess return data and $\mathbf{x}_{\mathrm{mv}} = (aW_0)^{-1}\mathbf{V}^{-1}\mathbf{u}$ where $\mathbf{u}$ and $\mathbf{V}$ are the estimated mean vector and covariance matrix, respectively. Portfolio

Table 4. Fitting statistics of the 241 assets for multivariate normal, NIG, VG, and GH distributions.

Distribution	Log-likelihood (LL)	Number of Param.	AIC
Mulivariate normal	− 53 0124.50	29 402	1 119 053
Asymmetric NIG	− 51 4932.00	29 645	1 089 154
Symmetric NIG	− 51 5290.00	29 403	1 089 386
Asymmetric VG	− 51 5020.00	29 645	1 089 330
Symmetric VG	− 51 5376.00	29 403	1 089 558
Asymmetric GH	− 51 4900.00	29 646	1 089 092
Symmetric GH	− 51 5260.50	29 404	1 089 329

Table 5. Composition information, statistics and performance measures of portfolios $\mathbf{x}^*$, $\mathbf{x}_{\text{mv}}$, $\mathbf{x}_{\text{S}}$, $\mathbf{x}^{KE}$ and $\mathbf{x}^{Q}$.

Portfolio	$\mathbf{x}^*$	$\mathbf{x}_{\text{mv}}$	$\mathbf{x}_{\text{S}}$	$\mathbf{x}^{KE}$	$\mathbf{x}^{Q}$
Average long position (%)	0.13	0.14	0.13	0.37	0.13
Maximum long position (%)	0.81	1.00	0.95	2.59	0.81
Total long position (%)	15.61	15.87	15.92	44.78	15.65
Average short position (%)	− 0.12	− 0.12	− 0.13	− 0.37	− 0.12
Minimum short position (%)	− 0.60	− 0.60	− 0.96	− 2.21	− 0.59
Total short position (%)	− 14.78	− 15.06	− 15.00	− 45.35	− 14.82
Mean (%)	0.92	0.94	0.80	0.51	0.92
Standard deviation (%)	2.73	2.88	2.84	8.90	2.73
Skewness	0.17	0.11	− 0.32	0.80	0.16
Kurtosis	3.97	3.95	4.07	4.79	3.97
Sharpe ratio	0.337	0.327	0.283	0.057	0.337
Q	0.053	0.075	0.221	− 0.890	0.057
KE	− 0.009	0.043	0.363	− 0.954	0.000
EUT	0.115	0.106	0.068	− 0.889	0.115
Certainty Equivalent excess return (CE) (%)	0.459	0.425	0.273	− 3.554	0.459

Note: All moments and performance measures where computed assuming the calibrated asymmetric NIG distribution for daily excess returns.

$\mathbf{x}_{\text{S}}$ can be found using Theorem 1 when excess returns are assumed to follow a symmetric multivariate NIG.† We assume $aW_0 = 0.125$, and this choice guarantees that for the three portfolios the maximum position in any asset (either long or short) is limited to 1%.

Table 5 presents composition information, statistics and performance measures of portfolios $\mathbf{x}^*$ (including components $\mathbf{x}^{Q}$ and $\mathbf{x}^{KE}$), $\mathbf{x}_{\text{mv}}$ and $\mathbf{x}_{\text{S}}$. The statistics (mean, standard deviation, skewness and kurtosis) and the performance measures (Sharpe ratio, Q, KE, EUT and Certainty Equivalent) were calculated using the calibrated asymmetric NIG distribution $\mathbf{R}_{\text{E}}$. For the optimal portfolio $\mathbf{x}^*$ we have: $\text{EUT}(\mathbf{x}^*) = 0.1148$, $Q(\mathbf{x}^*) = 0.0527$, $KE(\mathbf{x}^*) = -0.0093$, $\alpha^* = 0.9951$ and $\text{CE}(\mathbf{x}^*) = 0.459\%$. The value of α^* is very close to one and therefore the optimal solution will be dominated by $\mathbf{x}^{Q}$. From the table, we observe that for all portfolios (with the exception of $\mathbf{x}^{KE}$) the difference between the total long and short positions is around 1% and their compositions look similar. However, $\mathbf{x}^{KE}$ is the most aggressive portfolio and its positions are roughly three time the ones of the other portfolios. Portfolio $\mathbf{x}^*$ has lower mean and higher kurtosis than $\mathbf{x}_{\text{mv}}$; but, at the same time it has higher skewness and lower standard deviation. Moreover, portfolio $\mathbf{x}_{\text{S}}$ has lower mean, higher standard deviation, lower skewness (negative) and higher kurtosis than $\mathbf{x}^*$. Therefore, portfolio $\mathbf{x}^*$ is superior to $\mathbf{x}_{\text{S}}$ in each of the distributional moments. As expected, $\mathbf{x}^*$ shows the best performance in terms of EUT and CE. It

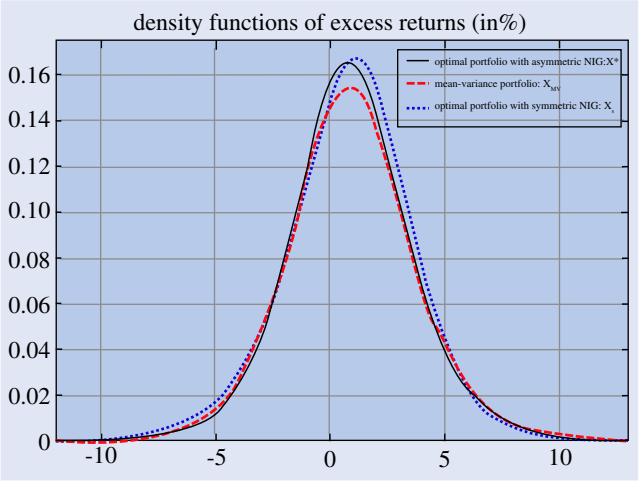

Figure 4. Density of the excess returns generated by portfolios $\mathbf{x}^*$ (optimal portfolio assuming asymmetric NIG), $\mathbf{x}_{\text{mv}}$ (optimal MV portfolio) and $\mathbf{x}_{\text{S}}$ (optimal portfolio assuming symmetric NIG). Densities were computed using the asymmetric NIG distribution of excess returns calibrated from daily data.

† We find $\mathbf{x}_{\text{S}}$ assuming that excess returns follow $\mathbf{R}_{\text{E,S}} \sim \text{GH}_N(-0.5, \chi_{\text{S}}, \psi_{\text{S}}, \boldsymbol{\mu}_{\text{S}}, \boldsymbol{\Sigma}_{\text{S}}, \mathbf{0})$. From the calibration we obtain $\chi_{\text{S}} = \psi_{\text{S}} = 3.1999$ and $C_{\text{S}} = \boldsymbol{\mu}_{\text{S}}^{\top} \boldsymbol{\Sigma}_{\text{S}}^{-1} \boldsymbol{\mu}_{\text{S}} = 0.13$. Also, $\mathbf{x}_{\text{S}}$ can be found using the following formula:

$$\mathbf{x}_{\text{S}} = \frac{1}{aW_0} \sqrt{\frac{\psi_{\text{S}}}{C_{\text{S}} + \chi_{\text{S}}}} \boldsymbol{\Sigma}_{\text{S}}^{-1} \boldsymbol{\mu}_{\text{S}}.$$

Table 6. Out-of-sample statistics and performance measures of portfolios $\mathbf{x}^*$, $\mathbf{x}_{\text{mv}}$, $\mathbf{x}_{\text{S}}$, $\mathbf{x}^{KE}$ and $\mathbf{x}^{Q}$.

Portfolio	$\mathbf{x}^*$	$\mathbf{x}_{\text{mv}}$	$\mathbf{x}_{\text{S}}$	$\mathbf{x}^{KE}$	$\mathbf{x}^{Q}$
Mean (%)	1.15	1.06	0.99	0.64	1.15
Standard deviation (%)	2.93	2.94	2.82	9.35	2.93
Skewness	0.47	0.07	0.19	1.38	0.46
Kurtosis	4.40	3.26	3.75	6.25	4.40
Average turnover (%/day)	0.368	0.376	0.787	1.107	0.368
Sharpe ratio	0.393	0.362	0.351	0.068	0.393
Certainty Equivalent excess return (CE) (%)	0.639	0.530	0.504	− 3.130	0.641

Note: The out-of-sample period is 05/08/2019–08/01/2020 (170 days).

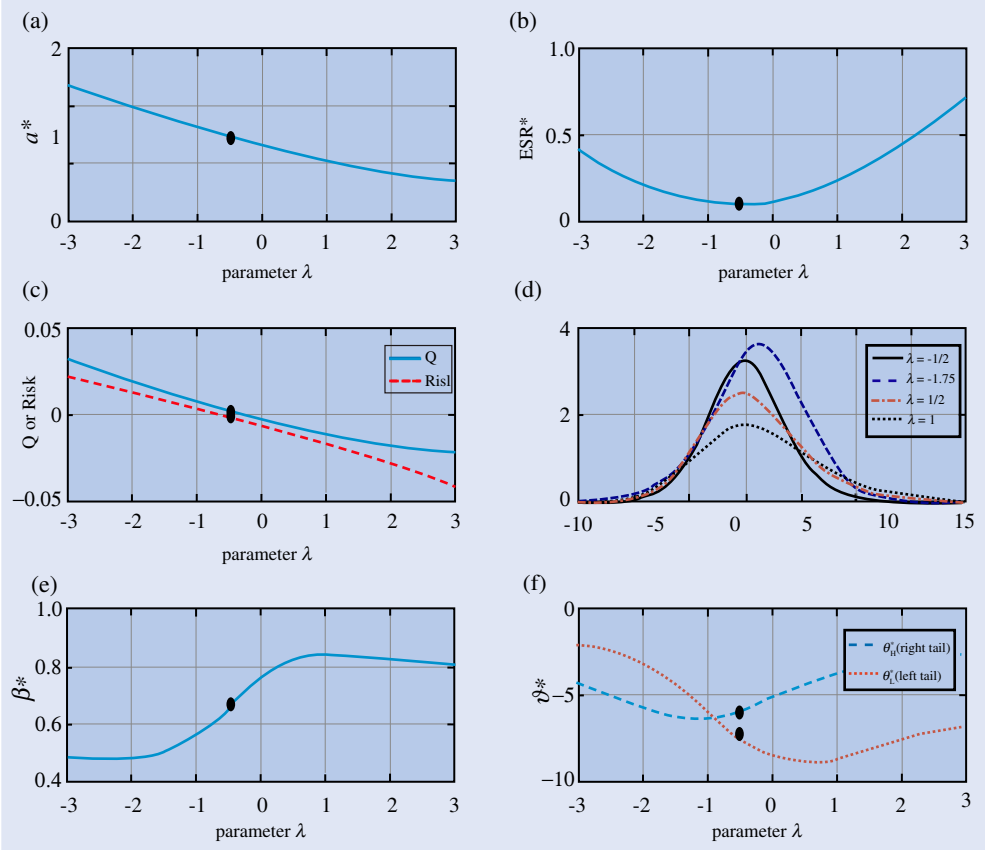

Figure 5. For the 241 risky assets and different values of parameter λ: panel (a) shows α^* (the optimal position in $\mathbf{x}^Q$), panel (b) displays the optimal performance EUT($\mathbf{x}^*$), panel (c) contains the values of $Q(\mathbf{x}^*)$ and $\text{Risk}(\mathbf{x}^*) = \frac{1}{2}\text{EUT}(\mathbf{x}^*) - Q(\mathbf{x}^*)$, panel (d) presents some density functions of the NIG distribution of the excess returns (in %) generated by $\mathbf{x}^*$, panel (e) exhibits the values of $\beta^* = \gamma^*/\sigma^{2*}$ and panel (d) shows the behavior of coefficients ϑ^* for the right and left tails. The black dot in the plots represents the results for the calibrated NIG distribution ($\lambda = -\frac{1}{2}$). We have assumed $aW_0 = 0.125$ in the portfolio optimization problem.

also has a Sharpe ratio greater than the one of $\mathbf{x}_{\text{mv}}$.† The performance of $\mathbf{x}_{\text{mv}}$ is close to the one of $\mathbf{x}^*$ even though their values of Q and KE are significantly different. The EUT of $\mathbf{x}_{\text{S}}$ is significantly lower compared to the ones of $\mathbf{x}^*$ and $\mathbf{x}_{\text{mv}}$.

Figure 4 shows the density plots of the asymmetric NIG distributions generated by portfolios $\mathbf{x}^*$, $\mathbf{x}_{\text{mv}}$ and $\mathbf{x}_{\text{S}}$ under $\mathbf{R}_{\text{E}}$. For example, the dotted line represents the density of $R_{\text{EX,S}} = \mathbf{x}_{\text{S}}{}^{\top}\mathbf{R}_{\text{E}}$ and the solid line the one of $R^*_{\text{EX}} = \mathbf{x}^{*\top}\mathbf{R}_{\text{E}}$. The values of the ratio β for portfolios $\mathbf{x}^*$, $\mathbf{x}_{\text{mv}}$ and $\mathbf{x}_{\text{S}}$ are $\beta^* = 0.0676$, $\beta_{\text{mv}} = 0.0416$ and $\beta_{\text{S}} = -0.1244$, respectively. This implies that the right tail is the heaviest for $\mathbf{x}^*$ and $\mathbf{x}_{\text{mv}}$; but, the left tail is the heaviest for $\mathbf{x}_{\text{S}}$. On the one hand, the right tail of $\mathbf{x}_{\text{mv}}$ decays slower than the one of $\mathbf{x}^*$ because $\vartheta^*_H = -0.5967$ and $\vartheta^{mv}_H = -0.5836$. On the other hand, the left tail of $\mathbf{x}^*$ decays faster than the one of $\mathbf{x}_{\text{mv}}$ because $\vartheta^*_L = -0.7319$ and $\vartheta^{mv}_L = -0.6667$. These facts can be observed from the tails of the distributions presented in the figure and can explain the differences in EUT among portfolios.

Table 6 shows the out-of-sample results of our portfolios in the period 05/08/2019–08/01/2020 which corresponds to 170 trading days. We report the out-of-sample mean, standard deviation, skewness and kurtosis, as well as, average daily turnover, Sharpe ratio and certainty equivalent excess return. To implement the out-of-sample analysis, we assume that portfolios were adjusted daily to match the ones of

† Although it is expected for $\mathbf{x}_{\text{mv}}$ to have the highest Sharpe ratio, this is not necessarily true because $\mathbf{x}_{\text{mv}}$ was not constructed using the mean vector and covariance matrix generated by $\mathbf{R}_{\text{E}}$.

the calibration period, i.e. $\mathbf{x}^*$, $\mathbf{x}_S$, $\mathbf{x}_{mv}$ remain constant. We observe that portfolio $\mathbf{x}^*$ had the best performance in terms of Sharpe ratio and CE, and also it had the lowest turnover. Additionally, portfolio $\mathbf{x}_S$ improved its performance with respect to the one of the calibration period; but, its turnover is almost twice as high as the ones of $\mathbf{x}^*$ and $\mathbf{x}_{mv}$. As expected, the performance of $\mathbf{x}^*$ is similar to the one of $\mathbf{x}^Q$; and, $\mathbf{x}^{KE}$ is the most aggressive portfolio in terms of turnover and exhibits the highest out-of-sample skewness and kurtosis. Portfolio $\mathbf{x}^*$ generated a CE approximately 0.11% higher than that of $\mathbf{x}_{mv}$ and this difference is greater than the 0.04% obtained in the calibrating period. Consequently, $\mathbf{x}^*$ continued to extract performance from the non-normality features of the daily returns in a better fashion than the optimal MV portfolio does.

Finally, we will study tail behavior when $B < 0$. For this purpose, we present Figure 5 which is similar to Figure 2, but considers the optimal portfolio with 241 risky assets. Notice that the values of EUT($\mathbf{x}^*$) and $Q(\mathbf{x}^*)$ have increased significantly compared to their values in the case of 6 risky assets. Recall that the aforementioned values are independent of aW_0. Since $B < 0$, the results of Proposition 6 for tail behavior may not hold. That is the case because in panels (e) and (f) of Figure 5 we observe: (i) a switch in the sign of β^*, (ii) β^* is not strictly increasing in λ, (iii) the values of ϑ^* for the right and left tails are not strictly increasing in λ; and, (iv) the heaviest tail switches from the left to the right at the point in which $\beta^* = 0$. The values of β^* can be negative but this happens under negative values of λ; and, this is associated with more aggressive allocations (in $\mathbf{x}^Q$ compared to $\mathbf{x}^{KE}$) since the corresponding α^* is greater than one.

6. Conclusions and future extensions

In this paper, we provide closed-form expressions to compute optimal portfolio weights when excess returns are modeled with a GH distribution, the investor has an exponential utility function and a risk-free asset is available. A new portfolio performance measure, the Expected Utility Transformation under the GH distribution (EUT), offers a decomposition leading to the definition of an efficient frontier (the $Q - KE$ frontier) that is constructed by an affine combination of two efficient portfolios. In this decomposition, Q can be associated with return (location) while KE is related to risk (dispersion adjusted by skewness), respectively. By means of the $Q - KE$ frontier, we determine interesting properties related to the optimal allocation and performance, and we are able to conduct sensitivity analysis with respect to the parameters of the mixing variable of the underlying GH returns.

Our theoretical results imply that as the tail density of a GH distribution becomes heavier, we invest more in a portfolio that minimizes risk and enforces positive skewness; and, simultaneously, we invest less in another more aggressive efficient portfolio that maximizes mean (location) subject to a fixed level of risk. We also provide theoretical evidence that the optimal performance is not a monotonic function of the parameters controlling tail behavior, and we find that optimal performance improves either as tails tend to become lighter or heavier (with respect to some threshold). In the numerical experiments, we show that, when modeling asset returns, replacing a gaussian distribution with an asymmetric GH distribution has the potential to increase the certainty equivalent excess return achieved by the investor. Future research can extend our theoretical results to the case when the feasible set is a linear combination of predetermined portfolios and another risk-averse utility function is considered. Finally, our results can generate new performance measures to assess the behavior and suitability of different portfolio construction methods.

Acknowledgments

John R. Birge gratefully acknowledges the support of the University of Chicago Booth School of Business.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

L. Chavez-Bedoya http://orcid.org/0000-0002-0992-9495

References

Aas, K. and Hobæk Haff, I.H., The generalized hyperbolic skew student's t-distribution. *J. Financ. Econ.*, 2006, **4**, 275–309.

Abramowitz, M. and Stegun, I.M., *Handbook of Mathematical Functions*, 1968 (Dover: New York).

Bandorff-Nielsen, O., Exponentially decreasing distributions for the logarithm of the particle size. *Proc. R. Soc. A: Math. Phys. Sci.*, 1977, **350**, 401–419.

Bandorff-Nielsen, O., Hyperbolic distributions and distributions on hyperbolae. *Scand. J. Stat.*, 1978, **5**, 151–157.

Bandorff-Nielsen, O. and Blæsild, P., Hyperbolic distributions and ramifications: Contributions to theory and application. In *Statistical Distributions in Scientific Work*, edited by C. Taillie, G. Patil, and B. Baldessari, pp. 19–44, 1981 (Reidel: Dordrecht).

Bernardo, A.E. and Ledoit, O., Gain, loss, and asset pricing. *J. Political Econ.*, 2000, **108**, 144–172.

Birge, J.R. and Chavez-Bedoya, L., Portfolio optimization under a generalized hyperbolic skewed t distribution and exponential utility. *Quant. Finance*, 2016, **16**, 1019–1036.

Eberlein, E. and Keller, U., Hyperbolic distributions in finance. *Bernoulli*, 1995, **1**, 281–299.

Fama, E.F., The behavior of stock-market prices. *J. Bus.*, 1965, **38**, 34–105.

Hellmich, M. and Kassberger, S., Efficient and robust portfolio optimization in the multivariate generalized hyperbolic framework. *Quant. Finance*, 2011, **11**, 1503–1516.

Hlawitschka, W., The empirical nature of taylor-series approximations to expected utility. *Am. Econ. Rev.*, 1994, **84**, 713–719.

Hodges, S., A Generalization of the Sharpe Ratio and Its Applications to Valuation Bounds and Risk Measures. Working Paper, Financial Options Research Centre, University of Warwick, 1998.

Hu, W. and Karcheval, A., Portfolio optimization for student *t* and skewed *t* returns. *Quant. Finance*, 2010, **10**, 91–105.

Ingersoll, J.E., *Theory of Financial Decision Making*, 1987 (Rowman and Littlefield: Totowa, NJ).

Ismail, M. and Muldoon, M., Monotonicity of the zeros of a cross-product of bessel functions. *SIAM J. Math. Anal.*, 1978, **9**, 759–767.
Kroll, Y., Levy, H. and Markowitz, H., Mean-variance versus direct utility maximization. *J. Finance*, 1984, **39**, 47–61.
Kwak, M. and Pirvu, T.A., Cumulative prospect theory with generalized hyperbolic skewed *t* distribution. *SIAM J. Financ. Math.*, 2018, **9**, 54–89.
Lillestøl, J., Fat and skew: Can NIG cure? On the Prospects of Using the Normal Inverse Gaussian Distribution in Finance. Working Paper, Norwegian School of Economics and Business Administration, 1998.
Madan, D. and McPhail, G., Investing in skews. *J. Risk Finance*, 2000, **2**, 10–18.
Madan, D. and Yen, J., Asset allocation with multivariate non-gaussian returns. Chapter 23 in *Handbooks in OR&MS*, edited by J.R. Birge and V. Linetsky, 2008 (Elsevier: North Holland).
Mandelbrot, B., The variation of certain speculative prices. *J. Bus.*, 1963, **36**, 394–419.
Markowitz, H., Portfolio selection. *J. Finance*, 1952, **7**, 77–91.
McNeil, A., Frey, R. and Embrechts, P., *Quantitative Risk Management: Concepts, Techniques, and Tools*, 2005 (Princeton University Press: Princeton).
Mencía, J. and Sentana, E., Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation. *J. Econom.*, 2009, **153**, 105–121.
Prause, K., The generalized hyperbolic model: Estimation, financial derivatives, and risk Measures. *Dissertation zur Erlangung des Doktorgrades der Mathematischen Fakultat der Albert-Ludwigs-Universitat Freiburg i. Br.*, 1999.
Pulley, L., A general mean-variance approximation to expected utility for short holding periods. *J. Financ. Quant. Anal.*, 1981, **16**, 361–373.
Reid, D.W. and Tew, B.V., Mean-variance versus direct utility maximization: A comment. *J. Finance*, 1986, **41**, 1177–1179.
Samuelson, P.A., The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *Rev. Econ. Stud.*, 1970, **37**, 537–542.
Sharpe, W.F., Mutual fund performance. *J. Bus.*, 1966, **39**, 119–138.
Simaan, Y., What is the opportunity cost of mean-variance investment strategies? *Manage. Sci.*, 1993, **39**, 578–587.
Tsiang, S.C., The rationale of the mean-standard deviation analysis, skewness preference, and the demand for money. *Am. Econ. Rev.*, 1972, **62**, 354–371.
Vanduffel, S. and Yao, J., A stein type lemma for the multivariate generalized hyperbolic distribution. *Eur. J. Oper. Res.*, 2017, **261**, 606–612.
Yu, J., Yang, X. and Li, S., Portfolio optimization with CVaR under VG process. *Res. Int. Bus. Finance*, 2009, **23**, 107–116.
Zakamouline, V. and Koekabakker, S., Portfolio performance evaluation with generalized Sharpe ratios: Beyond the mean and variance. *J. Bank. Finance*, 2009, **33**, 1242–1254.

Appendix 1. Generalized inverse gaussian distribution

The random variable Y has a generalized inverse Gaussian (GIG) distribution, $Y \sim \text{GIG}(\lambda, \chi, \psi)$, if its density is given by

$$f(y) = \frac{\chi^{-\lambda}(\sqrt{\chi\psi})^{\lambda}}{2K_{\lambda}(\sqrt{\chi\psi})} y^{\lambda-1} e^{-(1/2)(\chi y^{-1}+\psi y)}, \quad y > 0, \tag{A1}$$

where K_λ is the modified Bessel function of the second kind with index λ, and the parameters satisfy $\chi > 0$, $\psi \geq 0$ if $\lambda < 0$; $\chi > 0$, $\psi > 0$ if $\lambda = 0$; and $\chi \geq 0$, $\psi > 0$ if $\lambda > 0$. In the limiting cases $\chi = 0$ or $\psi = 0$, the density function (A1) must be interpreted as a limit that can be evaluated using properties (A10), (A11) or (A12) of the modified Bessel function of the second kind. The moment-generation of Y, $M_Y(s)$, is given by

$$M_Y(s) = \left(\frac{\psi}{\psi - 2s}\right)^{\lambda/2} \frac{K_\lambda(\sqrt{\chi(\psi - 2s)})}{K_\lambda(\sqrt{\chi\psi})} \tag{A2}$$

for $s \in \mathbb{R}$ such that $s < \psi/2$. We are also interested in computing $\lim_{s\to(\psi/2)^-} M_Y(s)$. Assuming $\psi > 0$ and $\chi > 0$, the previous limit depends on the value of λ. In the case in which $\lambda \geq 0$, the corresponding limit goes to infinity. In the case of $\lambda < 0$, we have the following result using property (A12):

$$\lim_{s\to(\psi/2)^-} M_Y(s) = \frac{\Gamma(-\lambda)}{2^{\lambda+1}} \frac{\sqrt{\chi\psi}^{\lambda}}{K_\lambda(\sqrt{\chi\psi})}. \tag{A3}$$

For the non-limiting case when $\psi > 0$ and $\chi > 0$, we can find the raw moments of Y by differentiating the moment-generating function and making $s = 0$. Then, after some algebraic calculations, we have the following:

$$\mathbb{E}[Y] = \left(\frac{\chi}{\psi}\right)^{1/2} \frac{K_{\lambda+1}(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi\psi})} \tag{A4}$$

and

$$\text{Var}(Y) = \frac{\chi}{\psi}\left[\frac{K_{\lambda+2}(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi\psi})} - \left(\frac{K_{\lambda+1}(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi\psi})}\right)^2\right]. \tag{A5}$$

Appendix 2. Specific properties and functions involving the modified Bessel function

The modified Bessel function of the second kind with index η has the following integral representation as given by Bandorff-Nielsen and Blæsild (1981):

$$K_\eta(x) = \frac{1}{2}\int_0^\infty y^{\eta-1} \exp\left\{-\frac{x}{2}(y + y^{-1})\right\} \text{dy}, \quad x > 0. \tag{A6}$$

Some important properties of K also given in Bandorff-Nielsen and Blæsild (1981) are:

$$K_\eta(x) = K_{-\eta}(x), \tag{A7}$$

$$K_{\eta+1}(x) = \frac{2\eta}{x} K_\eta(x) + K_{\eta-1}(x), \tag{A8}$$

and

$$K_{1/2}(x) = \sqrt{\frac{\pi}{2}}\frac{e^{-x}}{\sqrt{x}}. \tag{A9}$$

Additionally, K_η satisfies the following asymptotic relations:

$$K_\eta(x) \sim \Gamma(\eta)2^{\eta-1}x^{-\eta}, \quad \text{as } x \to 0^+ \text{ for } \eta > 0, \tag{A10}$$

$$K_0(x) \sim -\ln(x), \quad \text{as } x \to 0^+, \tag{A11}$$

$$K_\eta(x) \sim \Gamma(-\eta)2^{-\eta-1}x^{\eta}, \quad \text{as } x \to 0^+ \text{ for } \eta < 0, \tag{A12}$$

and

$$K_\eta(x) \sim \sqrt{\frac{\pi}{2x}} e^{-x}, \quad \text{as } x \to \infty. \tag{A13}$$

where $\Gamma(\cdot)$ is the gamma function. The derivative of the modified Bessel function of the second kind, as given in Abramowitz and Stegun (1968), satisfies

$$K'_\eta(x) = -\tfrac{1}{2}(K_{\eta+1}(x) + K_{\eta-1}(x)), \tag{A14}$$

and

$$K'_\eta(x) = -\frac{\eta}{x}K_\eta(x) - K_{\eta-1}(x). \tag{A15}$$

The following propositions are useful in proving some of our analytical results.

PROPOSITION A.1 *If $\lambda \in \mathbb{R}$, $\chi > 0$, $\psi > 0$ and $\psi > z$ with $z \neq 0$, the function*

$$f(\psi) = -\ln\left(\left(\frac{\psi}{\psi - z}\right)^{\lambda/2} \frac{K_\lambda(\sqrt{\chi(\psi - z)})}{K_\lambda(\sqrt{\chi\psi})}\right) \tag{A16}$$

is strictly decreasing when $z < 0$ and strictly increasing when $z > 0$.

Proof We start by computing the derivative of the function. Using properties (A8) and (A14) of the modified Bessel function, we obtain the following expression for $f'(\psi)$ after some algebraic manipulations:

$$f'(\psi) = \frac{\lambda z}{\psi(\psi - z)} + \frac{\chi}{2} \times \left(\frac{K_{\lambda-1}(\sqrt{\chi(\psi - z)})}{\sqrt{\chi(\psi - z)}K_\lambda(\sqrt{\chi(\psi - z)})} - \frac{K_{\lambda-1}(\sqrt{\chi\psi})}{\sqrt{\chi\psi}K_\lambda(\sqrt{\chi\psi})}\right). \tag{A17}$$

By Lemma 2.6 of Ismail and Muldoon (1978), the function $K_{\lambda-1}(x)/xK_\lambda(x)$ decreases with x, when $0 < x < \infty$ and for any $\lambda \in \mathbb{R}$. Therefore, $f'(\psi) < 0$ when $z < 0$, and $f'(\psi) > 0$ when $z > 0$. ■

PROPOSITION A.2 *Given $a > 0$, $b > 0$ and function $g(\nu) = K_\nu(a)/K_\nu(b)$. If $b > a$, then $g(\nu)$ is strictly increasing for $\nu \geq 0$ and strictly decreasing for $\nu \leq 0$. If $a > b$, then $g(\nu)$ is strictly increasing for $\nu \leq 0$ and strictly decreasing for $\nu \geq 0$.*

Proof First assume $\nu \geq 0$. We have to prove that for all $\epsilon > 0$, $g(\nu + \epsilon) > g(\nu)$ when $b > a$. Expanding $g(\nu + \epsilon) - g(\nu)$, we obtain

$$\frac{K_{\nu+\epsilon}(a)}{K_{\nu+\epsilon}(b)} - \frac{K_\nu(a)}{K_\nu(b)} = \frac{K_\nu(a)}{K_\nu(b)}\left(\frac{K_{\nu+\epsilon}(a)/K_\nu(a)}{K_{\nu+\epsilon}(b)/K_\nu(b)} - 1\right). \tag{A18}$$

By Lemma 2.4 of Ismail and Muldoon (1978), the function $K_{\nu+\epsilon}(x)/K_\nu(x)$ is decreasing in $x > 0$ for all $\epsilon > 0$ such that $\nu > -\epsilon/2$. Since we are working with an arbitrary ϵ, it implies that for $\nu \geq 0$, and the right-hand side of (A18) is greater than zero when $b > a$. The proof when $\nu \geq 0$ and $a > b$ follows similarly. Finally, (A7) implies directly the results when $\nu \leq 0$. ■

PROPOSITION A.3 *If $\chi > 0$, $\psi > 0$ and $\psi > z$ with $z \neq 0$, the function*

$$f(\lambda) = -\ln\left(\left(\frac{\psi}{\psi - z}\right)^{\lambda/2} \frac{K_\lambda(\sqrt{\chi(\psi - z)})}{K_\lambda(\sqrt{\chi\psi})}\right) \tag{A19}$$

is strictly increasing when $z < 0$ and strictly decreasing when $z > 0$.

Proof First consider $\lambda \geq 0$. The function $f(\lambda)$ is equivalent to

$$f(\lambda) = \frac{\lambda}{2}\ln\left(\frac{\psi - z}{\psi}\right) + \ln(r(\lambda)), \tag{A20}$$

where

$$r(\lambda) = \frac{K_\lambda(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi(\psi - z)})}. \tag{A21}$$

Notice that $r(\lambda) > 0$ for all λ. Taking the derivative of $f(\lambda)$, we obtain

$$f'(\lambda) = \frac{1}{2}\ln\left(\frac{\psi - z}{\psi}\right) + \frac{r'(\lambda)}{r(\lambda)}. \tag{A22}$$

If $z < 0$, $(\psi - z)/\psi > 1$, $r'(\lambda) > 0$ (by Proposition A.2), and $f'(\lambda) > 0$. When $z > 0$, $(\psi - z)/\psi < 1$, $r'(\lambda) < 0$ (again by Proposition A.2), and $f'(\lambda) < 0$. Finally, the case when $\lambda \leq 0$ follows in a similar fashion. ■

Appendix 3. Proof of Proposition 1

Using properties (A7) and (A9) of the modified Bessel function, it is true that $K_{1/2}(z) = K_{-1/2}(z) = \sqrt{\pi/2}(e^{-z}/\sqrt{z})$ when $z > 0$. In the case of $\lambda = -\frac{1}{2}$ and $\chi = \psi$, expression (3) reduces to:

$$M_{\mathbf{X}}(\mathbf{s}) = \exp\{\mathbf{s}^\top\boldsymbol{\mu} + \psi - \sqrt{\psi(\psi - (\mathbf{s}^\top\boldsymbol{\Sigma}\mathbf{s} + 2\mathbf{s}^\top\boldsymbol{\gamma}))}\}. \tag{A23}$$

Finally, $\lim_{\psi\to\infty} M_{\mathbf{X}}(\mathbf{s}) = \exp\{\mathbf{s}^\top(\boldsymbol{\mu} + \boldsymbol{\gamma}) + \frac{1}{2}\mathbf{s}^\top\boldsymbol{\Sigma}\mathbf{s}\}$ when $M_{\mathbf{X}}(\mathbf{s})$ is given by (A23).

Appendix 4. Proof of Theorem 1

Using the moment-generating function given in (3), we can restate problem **(P1)** as

$$\begin{aligned} \text{Maximize} \quad & -e^{-aW_0(1+(1-\mathbf{x}^\top\mathbf{1})r_f+\mathbf{x}^\top\boldsymbol{\mu})}\left(\frac{\psi}{\psi+aW_0\mathbf{x}^\top(2\boldsymbol{\gamma}-aW_0\boldsymbol{\Sigma}\mathbf{x})}\right)^{\lambda/2} \\ & \frac{K_\lambda(\sqrt{\chi(\psi+aW_0\mathbf{x}^\top(2\boldsymbol{\gamma}-aW_0\boldsymbol{\Sigma}\mathbf{x}))})}{K_\lambda(\sqrt{\chi\psi})} \\ \text{s.t.} \quad & \mathbf{x} \in \mathbb{S}. \end{aligned}$$

If we define

$$\mathbf{m}(\mathbf{x}) = aW_0\mathbf{x}^\top(2\boldsymbol{\gamma} - aW_0\boldsymbol{\Sigma}\mathbf{x}), \tag{A24}$$

we will have $\mathbb{S} = \{\mathbf{x} \in \mathbb{R}^N \mid \psi + \mathbf{m}(\mathbf{x}) > 0\}$ when $\lambda \geq 0$ and $\mathbb{S} = \{\mathbf{x} \in \mathbb{R}^N \mid \psi + \mathbf{m}(\mathbf{x}) \geq 0\}$ when $\lambda < 0$. Notice that $\mathbb{S}$ is always a convex set because $\mathbf{m}(\mathbf{x})$ is concave ($\boldsymbol{\Sigma}$ is positive-definite). From the moment-generation function of the mixing variable Y given by (A2), if a portfolio $\mathbf{x}$ does not belong to $\mathbb{S}$, then $\mathbb{E}[U(W(\mathbf{x}))] = -\infty$. The differences in the feasible set $\mathbb{S}$ for $\lambda \geq 0$ and $\lambda < 0$ are caused by the existence of the limit (A3) in the case of $\lambda < 0$.

Then, we define

$$r(y) = -\frac{1}{aW_0}(\ln(-y) + aW_0(1 + r_f)), \quad y < 0. \tag{A25}$$

The function $r(y)$ is strictly increasing in y and since the exponential utility is ordinal, we can formulate an equivalent optimization problem to **(P1)**:

$$\begin{aligned} \text{Maximize} \quad & r(\mathbf{f}(\mathbf{x})) = \mathbf{q}(\mathbf{x}) = \mathbf{x}^\top(\boldsymbol{\mu} - r_f\mathbf{1}) - \frac{\mathbf{g}(\mathbf{x})}{aW_0} \\ \text{s.t.} \quad & \mathbf{x} \in \mathbb{S}, \end{aligned}$$

where

$$\mathbf{g}(\mathbf{x}) = \ln\left(\frac{1}{K_\lambda(\sqrt{\chi\psi})}\left(\frac{\psi}{\psi + \mathbf{m}(\mathbf{x})}\right)^{\lambda/2} K_\lambda(\sqrt{\chi(\psi + \mathbf{m}(\mathbf{x}))})\right), \tag{A26}$$

$\mathbf{f}(\mathbf{x})$ is the objective function of **(P1)**, and $\mathbf{m}(\mathbf{x})$ was defined in (A24). Notice that $-\mathbf{g}(\mathbf{x})$ is a concave function in the interior of $\mathbb{S}$, and this alternative problem is a standard concave optimization problem over a convex set.

After some re-arrangement, the gradient of the objective function is given by

$$\nabla\mathbf{f}(\mathbf{x}) = (\boldsymbol{\mu} - r_f\mathbf{1}) - (aW_0\boldsymbol{\Sigma}\mathbf{x} - \boldsymbol{\gamma})\frac{1}{\mathbf{L}(\mathbf{x})}, \tag{A27}$$

where

$$\mathbf{L}(\mathbf{x}) = \frac{\sqrt{\chi(\psi + \mathbf{m}(\mathbf{x}))}}{\chi}\frac{K_\lambda(\sqrt{\chi(\psi + \mathbf{m}(\mathbf{x}))})}{K_{\lambda+1}(\sqrt{\chi(\psi + \mathbf{m}(\mathbf{x}))})}. \tag{A28}$$

Then, the FOC of the optimization problem lead to the following system of equations

$$\mathbf{x}^* = \frac{1}{aW_0}(\mathbf{L}(\mathbf{x}^*)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}). \tag{A29}$$

We can see immediately that if $\boldsymbol{\mu} - r_f\mathbf{1} = \mathbf{0}$, then $\mathbf{x}^* = (1/aW_0)\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}$, and $\mathbf{x}^* \in \mathbb{S}$. Therefore, Case 1 has been proved.

Now, we assume $\boldsymbol{\mu} - r_f \mathbf{1} \neq \mathbf{0}$. Since $\boldsymbol{\Sigma}$ is positive definite, we have $C > 0$ and $A \geq 0$ (it is zero when $\boldsymbol{\gamma} = \mathbf{0}$). If system (A29) has a solution in $\mathbb{S}$, then it is obvious that $\mathbf{x}^*$ will be the solution to the expected utility maximization problem. Defining ζ^* such that $\zeta^* = \mathbf{L}(\mathbf{x}^*)$ and, since $\mathbf{x}^* = \mathbf{t}(\zeta^*)$, it is possible to find ζ^* by solving $\zeta^* = \mathbf{L}(\mathbf{t}(\zeta^*))$. After some algebraic manipulations, we have that $\mathbf{m}(\mathbf{x}^*) = aW_0\mathbf{x}^{*\top}(\boldsymbol{\gamma} - aW_0\boldsymbol{\Sigma}\mathbf{x}^*) = A - C\zeta^{*2}$. Consequently, ζ^* is the solution of the following equation:

$$\zeta = v(\zeta) = \frac{\sqrt{\chi(\psi + A - C\zeta^2)}}{\chi} \frac{K_\lambda(\sqrt{\chi(\psi + A - C\zeta^2)})}{K_{\lambda+1}(\sqrt{\chi(\psi + A - C\zeta^2)})}. \tag{A30}$$

The function $v(\zeta)$ can be expressed as a compound function, i.e.$v(\zeta) = w(u(\zeta))$, where

$$w(y) = \frac{y}{\chi} \frac{K_\lambda(y)}{K_{\lambda+1}(y)}, \quad y > 0 \tag{A31}$$

and

$$u(\zeta) = \sqrt{\chi(\psi + A - C\zeta^2)}, \quad -\theta < \zeta < \theta. \tag{A32}$$

In the next proposition, we investigate the first derivative of $w(y)$ and its limit when y approaches zero. Both results are important to determine the existence of ζ^*.

PROPOSITION A.4 *The following properties are satisfied by the function $w(y)$ defined in (A31)*:

(i) *The first derivative of $w(y)$ is always positive, and it is given by*

$$w'(y) = \frac{y}{\chi} \frac{K_{\lambda+2}(y)K_\lambda(y) - K_{\lambda+1}(y)^2}{K_{\lambda+1}(y)^2}. \tag{A33}$$

(ii)

$$\lim_{y\to 0^+} w(y) = \begin{cases} 0 & \text{if } \lambda \geq -1, \\ \dfrac{2(-\lambda - 1)}{\chi} & \text{if } \lambda < -1. \end{cases} \tag{A34}$$

Proof We begin proving the first statement. Using expressions (A14) and (A15), we have

$$w'(y) = \frac{y}{\chi}\left(\frac{K_\lambda(y)^2 - K_{\lambda+1}(y)^2 + (\frac{2(\lambda+1)}{y}K_{\lambda+1}(y))K_\lambda(y)}{K_{\lambda+1}(y)^2}\right). \tag{A35}$$

By property (A8) of the modified Bessel function, we can replace $(2(\lambda+1)/y)K_{\lambda+1}(y)$ in the expression above with $K_{\lambda+2}(y) - K_\lambda(y)$ to obtain expression (A33). Finally, having $w'(y) > 0$ for all y is equivalent to having $K_{\lambda+2}(y)K_\lambda(y) - K_{\lambda+1}(y)^2 > 0$. The last inequality is always true as stated in Ismail and Muldoon (1978).

To prove (2), we need to work with five cases: $\lambda > 0, \lambda = 0, -1 < \lambda < 0, \lambda - 1$ and $\lambda < -1$. For $\lambda > 0$, we have

$$\begin{aligned} \lim_{y\to 0^+} w(y) &= \lim_{y\to 0^+} \frac{y}{\chi} \lim_{y\to 0^+} \frac{K_\lambda(y)}{K_{\lambda+1}(y)} \\ &= \lim_{y\to 0^+} \frac{y}{\chi} \frac{\Gamma(\lambda)2^{\lambda-1}y^{-\lambda}}{\Gamma(\lambda+1)2^{\lambda+1-1}y^{-\lambda-1}} \\ &= \frac{1}{2\chi}\frac{\Gamma(\lambda)}{\Gamma(\lambda+1)} \lim_{y\to 0^+} y^2 = 0. \end{aligned} \tag{A36}$$

In the intermediate step above, we have used property (A7).

In the case of $\lambda = 0$, we have

$$\begin{aligned} \lim_{y\to 0^+} w(y) &= \lim_{y\to 0^+} \frac{y}{\chi} \lim_{y\to 0^+} \frac{K_0(y)}{K_1(y)} \\ &= \lim_{y\to 0^+} \frac{y}{\chi} \frac{-\ln(y)}{y^{-1}} \\ &= -\frac{1}{\chi} \lim_{y\to 0^+} \ln(y)y^2 = 0. \end{aligned} \tag{A37}$$

In the intermediate step above, we have used both (A7) and (A11).

In the case of $-1 < \lambda < 0$, we have

$$\begin{aligned} \lim_{y\to 0^+} w(y) &= \lim_{y\to 0^+} \frac{y}{\chi} \lim_{y\to 0^+} \frac{K_\lambda(y)}{K_{\lambda+1}(y)} \\ &= \lim_{y\to 0^+} \frac{y}{\chi} \frac{\Gamma(-\lambda)2^{-\lambda-1}y^{\lambda}}{\Gamma(\lambda+1)2^{\lambda+1-1}y^{-\lambda-1}} \\ &= \frac{1}{\chi}\frac{\Gamma(-\lambda)}{\Gamma(\lambda+1)}\frac{1}{2^{2\lambda+1}} \lim_{y\to 0^+} y^{2(\lambda+1)} = 0. \end{aligned} \tag{A38}$$

In the intermediate step above, we have used both (A7) and (A10).

In the case of $\lambda = 1$, we have

$$\begin{aligned} \lim_{y\to 0^+} w(y) &= \lim_{y\to 0^+} \frac{y}{\chi} \lim_{y\to 0^+} \frac{K_{-1}(y)}{K_0(y)} \\ &= \lim_{y\to 0^+} \frac{y}{\chi} \frac{y^{-1}}{-\ln(y)} \\ &= -\frac{1}{\chi} \lim_{y\to 0^+} \frac{1}{\ln(y)} = 0. \end{aligned} \tag{A39}$$

In the intermediate step above, we have used both (A10) and (A11).

In the case of $\lambda < -1$, we have

$$\begin{aligned} \lim_{y\to 0^+} w(y) &= \lim_{y\to 0^+} \frac{y}{\chi} \lim_{y\to 0^+} \frac{K_\lambda(y)}{K_{\lambda+1}(y)} \\ &= \lim_{y\to 0^+} \frac{y}{\chi} \frac{\Gamma(-\lambda)2^{-\lambda-1}y^{\lambda}}{\Gamma(-\lambda-1)2^{-\lambda-1-1}y^{\lambda+1}} \\ &= \frac{2}{\chi}\frac{\Gamma(-\lambda)}{\Gamma(-\lambda-1)} = \frac{2}{\chi}(-\lambda - 1). \end{aligned} \tag{A40}$$

In the intermediate step above, we have used (A10), and then the proof is complete. ■

Using the derivative of $w(y)$ given in Proposition A.4, it is easy to verify that the derivative of $v(\zeta)$ with respect to ζ is given by:

$$v'(\zeta) = -\zeta C \frac{K_{\lambda+2}(\sqrt{\chi(\psi + A - C\zeta^2)})K_\lambda(\sqrt{\chi(\psi + A - C\zeta^2)}) - K_{\lambda+1}(\sqrt{\chi(\psi + A - C\zeta^2)})^2}{K_{\lambda+1}(\sqrt{\chi(\psi + A - C\zeta^2)})^2}. \tag{A41}$$

From the last expression we can observe that $v'(0) = 0$ and $v'(\zeta) < 0$ when $\zeta > 0$, and $v'(\zeta) > 0$ when $\zeta < 0$. Notice that v is defined for $-\theta < \zeta < \theta$, and by Proposition C.1 we have: $\lim_{\zeta\to\theta^-} v(\zeta) = \lim_{\zeta\to\theta^+} v(\zeta) = 0$ when $\lambda \geq -1$, and $\lim_{\zeta\to\theta^-} v(\zeta) = \lim_{\zeta\to\theta^+} v(\zeta) = (2/\chi)(-\lambda - 1)$ when $\lambda < -1$. It can also be shown that $v(\zeta)$ is strictly concave but this fact is not necessary in the proof.

In Figure A1, we show the plot of a typical function v and the location of ζ^* for $\lambda \geq -1$. Notice that ζ^* is located at the intersection of $v(\zeta)$ and the 45-degree line starting at the origin. Because of Proposition A.4, it is clear that ζ^* will always exist when $\lambda \geq -1$, proving Case 2 of the theorem. In Figure A2, we show the function v assuming $\lambda < -1$ and $\theta > 2(-\lambda - 1)/\chi$. In this particular case, we can observe that the system $v(\zeta) = \zeta$ will always have a solution, and this proves Case 3. Figure 3 shows the plot of $v(\zeta)$ when $\lambda < -1$ and $\theta \leq 2(-\lambda - 1)/\chi$. We can observe that the system $v(\zeta) = \zeta$ does not have a solution since the 45 degree line does not intersect $v(\zeta)$. Consequently, $\mathbf{x}^*$ is not in the interior of $\mathbb{S}$ and is located on the boundary. The previous conditions correspond to the ones of Case 4. Next, we find the optimal portfolio under these assumptions.

Consider a portfolio $\mathbf{x}^B$ such that $\psi + \mathbf{m}(\mathbf{x}^b) = 0$, i.e.$\mathbf{x}^B$ is in the boundary of the feasible region $\mathbb{S}$ (we are also assuming $\lambda < -1$).

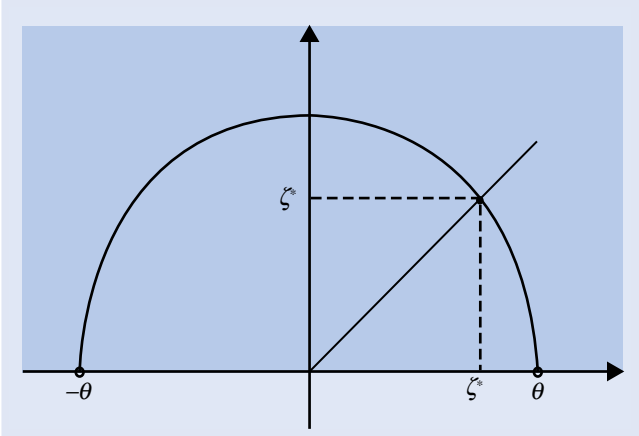

Figure A1. Plot of $v(\zeta)$ for $\lambda \geq -1$ with the location of ζ^*.

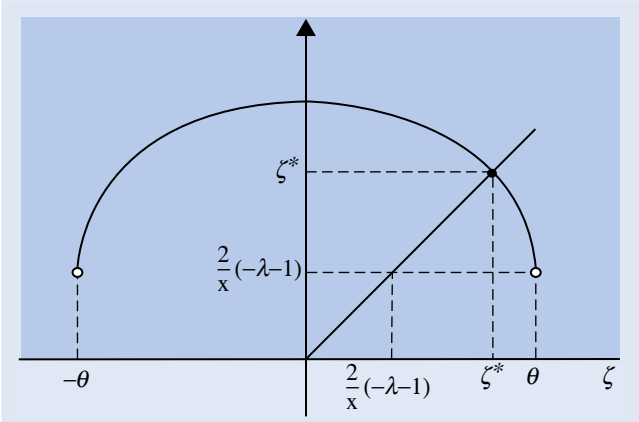

Figure A2. Plot of $v(\zeta)$ for $\lambda < -1$ and $\theta > 2(-\lambda - 1)/\chi$ with the location of ζ^*.

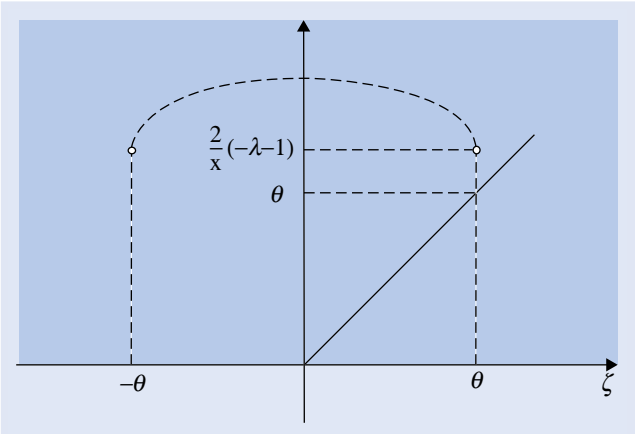

Figure A3. Plot of $v(\zeta)$ for $\lambda < -1$ and $\theta \leq 2(-\lambda - 1)/\chi$.

It can be shown that when $\mathbf{x}$ tends to $\mathbf{x}^B$, the function $\mathbf{g}(\mathbf{x})$ tends to a constant. Consequently, the utility maximization problem at the boundary is given by:

$$\begin{array}{ll} \text{Maximize} & \mathbf{f}(\mathbf{x}) = \mathbf{x}^\top(\boldsymbol{\mu} - r_f\mathbf{1}) \\ \text{s.t.} & 2\mathbf{x}^\top\boldsymbol{\gamma} - aW_0\mathbf{x}^\top\boldsymbol{\Sigma}\mathbf{x} = 0. \end{array}$$

We can write the FOC of the boundary optimization problem as

$$(\boldsymbol{\mu} - r_f\mathbf{1}) + \frac{1}{\theta}(\boldsymbol{\gamma} - aW_0\boldsymbol{\Sigma}\mathbf{x}^*) = \mathbf{0}, \tag{A42}$$

with $\theta \neq 0$. The conditions lead to

$$\mathbf{x}^* = \frac{1}{aW_0}(\theta\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}). \tag{A43}$$

Now, we replace $\mathbf{x}^*$ in the boundary condition to find the value of θ. After solving the equation for θ, we have

$$\theta = \pm\sqrt{\frac{\psi + A}{C}}. \tag{A44}$$

Finally, using the second order conditions of optimality, we keep the positive solution. This completes the proof for Case 4.

Appendix 5. Proof of Proposition 2

To prove (18), we begin by computing the following limit:

$$\begin{aligned} \lim_{\psi\to\infty}\frac{1}{2}\,\mathrm{EUT}(\mathbf{x}) &= Q(\mathbf{x}) - \lim_{\psi\to\infty}\psi - \sqrt{\psi(\psi - KE(\mathbf{x}))} \\ &= aW_0\mathbf{x}^\top(\boldsymbol{\mu} + \boldsymbol{\gamma} - r_f\mathbf{1}) - \frac{1}{2}(aW_0)^2\mathbf{x}^\top\boldsymbol{\Sigma}\mathbf{x}. \end{aligned} \tag{A45}$$

Additionally, if $\mathbf{R}_{mv} \sim N(\boldsymbol{\mu} + \boldsymbol{\gamma}, \boldsymbol{\Sigma})$, then $\mathbb{E}[\mathbf{x}^\top(\mathbf{R}_{mv} - r_f\mathbf{1})] = \mathbf{x}^\top(\boldsymbol{\mu} + \boldsymbol{\gamma} - r_f\mathbf{1})$ and $\mathrm{Var}(\mathbf{x}^\top\mathbf{R}_{mv}) = \mathbf{x}^\top\boldsymbol{\Sigma}\mathbf{x}$. Replacing those values in expression (A45), we obtain (18). To prove (19) and (20), first we need to find $\mathbf{x}^*$ when $\lambda = -\frac{1}{2}$ and $\chi = \psi$. Under the previous assumptions, the optimal portfolio satisfies $aW_0\mathbf{x}^* = \zeta^*\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}$, and ζ^* is the solution to

$$\zeta = \sqrt{\frac{\psi + A - C\zeta^2}{\psi}}. \tag{A46}$$

Notice that the right-hand side of the equation above is equal to $v(\zeta)$ given by (A30) when $\lambda = -\frac{1}{2}$ and $\chi = \psi$. Solving the previous system, we obtain $\zeta^* = \sqrt{(\psi + A)/(\psi + C)}$, and we can verify that $\lim_{\psi\to\infty}\zeta^* = 1$, which implies (19). Additionally, the expression for EUT given in (**P1**) reduces to $\frac{1}{2}\,\mathrm{EUT}(\mathbf{x}) = Q(\mathbf{x}) - (\psi - \sqrt{\psi(\psi - KE(\mathbf{x}))})$ when $\lambda = -\frac{1}{2}$ and $\chi = \psi$. After finding $Q(\mathbf{x}^*)$ and $KE(\mathbf{x}^*)$, it can be shown that $\frac{1}{2}\,\mathrm{EUT}(\mathbf{x}^*) = C\zeta^* + B - (\psi - \sqrt{\psi(\psi - \zeta^{*2}C + A)})$. To complete the proof, we have to evaluate $p = \lim_{\psi\to\infty}\psi - \sqrt{\psi(\psi - \zeta^{*2}C + A)}$. After the appropriate computations, we have $p = (C - A)/2$. This implies that $\lim_{\psi\to\infty}\frac{1}{2}\,\mathrm{EUT}(\mathbf{x}^*) = C + B + (A - C)/2$, and $\lim_{\psi\to\infty}\mathrm{EUT}(\mathbf{x}^*) = (\boldsymbol{\mu} + \boldsymbol{\gamma} - r_f\mathbf{1})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} + \boldsymbol{\gamma} - r_f\mathbf{1})$. Notice that the last expression corresponds to the Sharpe ratio squared of portfolio $\mathbf{x}_E^*$ which is optimal in the case of a multivariate normal distribution of returns.

Appendix 6. Proof of Theorem 2

We start by showing that both $\mathbf{x}^Q$ and $\mathbf{x}^{KE}$ solve two simple optimization problems involving Q and KE. After that, we establish that any optimal portfolio $\mathbf{x}^*$ can be expressed as an affine combination of the two aforementioned portfolios.

First, we find the portfolio with the lowest possible value of KE, and call this portfolio $\mathbf{x}^{KE}$. Consequently,

$$\mathbf{x}^{KE} = \mathrm{argmin}_{\mathbf{x}\in\mathbb{S}}\{KE(\mathbf{x})\}. \tag{A47}$$

Taking the Lagrangian of $KE(\mathbf{x})$, the FOC for the above optimization problem are given by the following system of equations

$$aW_0\boldsymbol{\Sigma}\mathbf{x}^{KE} - \boldsymbol{\gamma} = \mathbf{0}, \tag{A48}$$

implying that $\mathbf{x}^{KE} = (1/aW_0)\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}$. The portfolio $\mathbf{x}^{KE}$ is feasible since $KE(\mathbf{x}^{KE}) = -\boldsymbol{\gamma}^\top\boldsymbol{\Sigma}^{-1}\gamma = -A$ and $\psi - KE(\mathbf{x}^{KE}) = \psi + A >$

0 since $\boldsymbol{\gamma} \neq \mathbf{0}$ and $\boldsymbol{\Sigma}$ is positive-definite. The SOC are satisfied since *KE* is a strictly convex function.

Second, we find the portfolio that maximizes Q with the condition to produce a value of *KE* equal to zero. We call this portfolio $\mathbf{x}^Q$, and it solves the following optimization problem:

$$\begin{array}{ll} \text{Maximize} & Q(\mathbf{x}) = aW_0\mathbf{x}^\top(\boldsymbol{\mu} - r_f\mathbf{1}) \\ \text{s.t.} & KE(\mathbf{x}) = (aW_0)^2\mathbf{x}^\top\boldsymbol{\Sigma}\mathbf{x} - 2aW_0\mathbf{x}^\top\boldsymbol{\gamma} = 0, \\ & \mathbf{x} \in \mathbb{S}. \end{array}$$

Since $aW_0 > 0$, we can write the Lagrangian of the problem above as

$$\mathbf{La}(\mathbf{x}) = \mathbf{x}^\top(\boldsymbol{\mu} - r_f\mathbf{1}) + \frac{\epsilon}{2}(2\mathbf{x}^\top\boldsymbol{\gamma} - aW_0\mathbf{x}^\top\boldsymbol{\Sigma}\mathbf{x}), \tag{A49}$$

and after some simplifications in $\nabla\mathbf{La}(\mathbf{x})$, the FOC are given by the following system of equations:

$$\frac{1}{\epsilon}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\gamma} = aW_0\boldsymbol{\Sigma}\mathbf{x}^Q, \tag{A50}$$

which imply that $\mathbf{x}^Q = (1/aW_0)(\nu\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\Sigma}^{-1})$ with $\nu = 1/\epsilon$. Next, we find the value of ν using the constraint $KE(\mathbf{x}^Q) = 0$. After solving a quadratic equation and choosing the appropriate sign of the solution, the value of ν is given by $\nu = \sqrt{A/C} > 0$, where $C > 0$ and $A > 0$ by the assumptions made in the theorem. Moreover, portfolio $\mathbf{x}^Q$ is always feasible since $\psi - KE(\mathbf{x}^Q) = \psi > 0$.

It is important to remember that both $\mathbf{x}^Q$ and $\mathbf{x}^{KE}$ are independent of the parameters λ, χ, and ψ, and are always feasible. To complete the proof, we need to find α^* such that $\mathbf{x}^* = \alpha^*\mathbf{x}^Q + (1 - \alpha^*)\mathbf{x}^{KE}$. Here, $\mathbf{x}^*$ is the optimal solution of **(P1)** given by (12) or (14). Using the expressions for $\mathbf{x}^Q$ and $\mathbf{x}^{KE}$,

$$\mathbf{x}^* = \frac{1}{aW_0}(\nu\alpha^*\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}). \tag{A51}$$

Using Theorem 1, it is clear when either $\lambda \geq -1$ (Case 2) or $\lambda < -1$ and $(\chi/2(-\lambda - 1))\theta > 1$ (Case 3), we can find α^* by solving the following equation:

$$\alpha^* = \sqrt{\frac{C}{A}}\frac{\sqrt{\chi(\psi + (1 - \alpha^{*2})A)}}{\chi}\frac{K_\lambda(\sqrt{\chi(\psi + (1 - \alpha^{*2})A)})}{K_{\lambda+1}(\sqrt{\chi(\psi + (1 - \alpha^{*2})A)})}. \tag{A52}$$

Finally, when $\lambda < -1$ and $\frac{\chi}{2(-\lambda-1)}\theta \leq 1$ (Case 4 of Theorem 3.1.1), we have

$$\alpha^* = \sqrt{\frac{\psi}{A} + 1}. \tag{A53}$$

Appendix 7. Proof of Theorems and Propositions for the $Q - KE$ frontier

A.1. Proof of Theorem 3

In order to find $\mathbf{x}^*(Q)$, we need to solve the following optimization problem:

$$\begin{array}{ll} \text{Minimize} & \frac{1}{2}(aW_0)^2\mathbf{x}^\top\boldsymbol{\Sigma}\mathbf{x} - aW_0\mathbf{x}^\top\boldsymbol{\gamma} \\ \text{s.t.} & aW_0\mathbf{x}^\top(\boldsymbol{\mu} - r_f\mathbf{1}) = Q, \end{array}$$

Since $aW_0 > 0$, we can write the Lagrangian of the problem above as

$$\mathbf{La}(\mathbf{x}) = \tfrac{1}{2}(aW_0)^2\mathbf{x}^\top\boldsymbol{\Sigma}\mathbf{x} - aW_0\mathbf{x}^\top\boldsymbol{\gamma} + \epsilon(Q - aW_0\mathbf{x}^\top(\boldsymbol{\mu} - r_f\mathbf{1})); \tag{A54}$$

after some simplifications in $\nabla\mathbf{La}(\mathbf{x})$, the FOC are given by the following system of equations:

$$aW_0\boldsymbol{\Sigma}\mathbf{x}^*(Q) = \epsilon(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\gamma}, \tag{A55}$$

which imply that $\mathbf{x}^*(Q) = (1/aW_0)\boldsymbol{\Sigma}^{-1}(\epsilon(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\gamma})$. Now, we replace $\mathbf{x}^*(Q)$ in the constraint of the optimization problem to find ϵ. Then, $\epsilon = (Q - B)/C$ and

$$\mathbf{x}^*(Q) = \frac{1}{aW_0}\boldsymbol{\Sigma}^{-1}\left(\frac{(Q - B)}{C}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\gamma}\right). \tag{A56}$$

Finally, using the expression for $\mathbf{x}^*(Q)$ above, we can verify that

$$KE(\mathbf{x}^*(Q)) = \frac{(Q - B)^2}{C} - A. \tag{A57}$$

A.2. Proof of Theorem 4

Let $Q = Q(\mathbf{x}) = aW_0\mathbf{x}^\top(\boldsymbol{\mu} - r_f\mathbf{1})$ and because $\mathbf{x}$ belongs to the $Q - KE$ frontier, we have:

$$\begin{aligned} \mathbf{x} &= \frac{1}{aW_0}\boldsymbol{\Sigma}^{-1}\left(\left(\frac{Q - B}{C}\right)(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\gamma}\right) \\ &= \sqrt{\frac{C}{A}}\left(\frac{Q - B}{C}\right)\frac{1}{aW_0}\boldsymbol{\Sigma}^{-1}\left(\sqrt{\frac{A}{C}}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\gamma}\right) \\ &\quad + \left(1 - \sqrt{\frac{C}{A}}\left(\frac{Q - B}{C}\right)\right)\frac{1}{aW_0}\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma} \\ &= \left(\frac{Q - B}{\sqrt{AC}}\right)\mathbf{x}^Q + \left(1 - \frac{Q - B}{\sqrt{AC}}\right)\mathbf{x}^{KE}. \end{aligned}$$

Finally, we complete the proof by taking

$$\alpha = \frac{Q - B}{\sqrt{AC}}. \tag{A58}$$

A.3. Proof of Theorem 5

From Theorem 2, we have

$$\mathbf{x}^* = \frac{1}{aW_0}(\alpha^*\nu\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}), \tag{A59}$$

where $\nu = \sqrt{A/C}$. Consequently,

$$\begin{aligned} Q(\mathbf{x}^*) &= aW_0\mathbf{x}^{*\top}(\boldsymbol{\mu} - r_f\mathbf{1}) \\ &= \nu\alpha^*C + B, \end{aligned} \tag{A60}$$

and

$$\begin{aligned} KE(\mathbf{x}^*) &= (aW_0)^2\mathbf{x}^{*\top}\boldsymbol{\Sigma}\mathbf{x}^* - 2aW_0\mathbf{x}^{*\top}\boldsymbol{\gamma} \\ &= (\nu\alpha^*(\boldsymbol{\mu} - r_f\mathbf{1})^\top + \boldsymbol{\gamma}^\top)(\nu\alpha^*\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}) \\ &\quad - 2\boldsymbol{\gamma}^\top(\nu\alpha^*\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}) + \boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}) \\ &= (\nu\alpha^*)^2C + 2\nu\alpha^*B + A - 2\nu\alpha^*B - 2A \\ &= (\nu\alpha^*)^2C - A. \end{aligned} \tag{A61}$$

From the expression of $Q(\mathbf{x}^*)$, we have $\nu\alpha^* = (Q(\mathbf{x})^* - B)/C$. Then,

$$KE(Q(\mathbf{x}^*)) = \left(\frac{Q(\mathbf{x}^*) - B}{C}\right)^2 C - A = \frac{(Q(\mathbf{x}^*) - B)^2}{C} - A, \tag{A62}$$

which is the equation of a parabola in the $Q - KE$ space. However, α^* imposes a constraint in the possible values of $Q(\mathbf{x}^*)$ since $KE(\mathbf{x}^*) < \psi$ if $\lambda \geq 0$ or $KE(\mathbf{x}^*) \leq \psi$ if $\lambda < 0$. This constraint is given by either:

$$Q_{\min} \leq Q(\mathbf{x}^*) < Q_{\max}, \quad \text{if } \lambda \geq 0,$$

or

$$Q_{\min} \leq Q(\mathbf{x}^*) \leq Q_{\max}, \quad \text{if } \lambda < 0,$$

with $Q_{\min} = B$ and $Q_{\max} = \sqrt{C(\psi + A)} + B$.

A.4. Proof of Proposition 3

Since $\mathbf{x}$ belongs to the $Q-KE$ frontier, by Theorem 4 there exists $\alpha \in \mathbb{R}$ such that $\mathbf{x} = \alpha\mathbf{x}^Q + (1-\alpha)\mathbf{x}^{KE}$. If we define $\mathbf{x}(\rho) = \rho\mathbf{x} = \rho(\alpha\mathbf{x}^Q + (1-\alpha)\mathbf{x}^{KE})$, then $Q(\mathbf{x}(\rho)) = \rho(\nu\alpha C + B)$ and

$$KE(\mathbf{x}(\rho)) = \rho^2((\nu\alpha)^2 C + 2\nu\alpha B + A) - 2\rho(\nu\alpha B + A) \tag{A63}$$

with $\nu = \sqrt{A/C}$. In order for $\mathbf{x}(\rho)$ to be in the $Q-KE$ frontier the equation,

$$KE(\mathbf{x}(\rho)) = \frac{(Q(\mathbf{x}(\rho)) - B)^2}{C} - A, \tag{A64}$$

needs to hold. The equation above reduces to

$$\frac{(AC - B^2)(\rho - 1)^2}{C} = 0. \tag{A65}$$

Consequently, if $AC - B^2 = 0$ the portfolio $\rho\mathbf{x}$ will belong to the $Q-KE$ frontier for all $\rho \in \mathbb{R}$.

Appendix 8. Proof of Proposition 4

We start proving that $\partial\alpha^*(\lambda,\chi,\psi)/\partial\psi > 0$. From Cases 2 and 3 of Theorems 1 and 2, it is clear that $\mathbf{L}(\mathbf{x}^*) = \nu\alpha^*(\lambda,\chi,\psi)$ where $\nu = \sqrt{A/C}$ and $\mathbf{L}(\mathbf{x}^*)$ is given by (13). Moreover, we can define $\mathbf{L}(\mathbf{x}^*) = \zeta^*(\lambda,\chi,\psi)$ and $\zeta^*(\lambda,\chi,\psi)$ is the solution to

$$\zeta = \nu(\zeta) = \frac{\sqrt{\chi(\psi + A - C\zeta^2)}}{\chi} \frac{K_\lambda(\sqrt{\chi(\psi + A - C\zeta^2)})}{K_{\lambda+1}(\sqrt{\chi(\psi + A - C\zeta^2)})}. \tag{A66}$$

Then, it is enough to prove that

$$\frac{\partial\zeta^*(\lambda,\chi,\psi)}{\partial\psi} > 0, \tag{A67}$$

where $\zeta^*(\lambda,\chi,\psi)$ is the solution to (A66). We define $f(\psi) = r(g(\psi))$ where

$$r(y) = \frac{y}{\psi}\frac{K_\lambda(y)}{K_{\lambda+1}(y)}, \quad y > 0, \tag{A68}$$

and

$$g(\psi) = \sqrt{\chi(\psi + A - C\zeta^2)}. \tag{A69}$$

Hence, we need to show that $f'(\psi) > 0$. Using properties (A8) and (A14) of the modified Bessel function and some algebraic manipulation, we obtain

$$f'(\psi) = \frac{1}{2}\frac{(K_{\lambda+2}(g(\psi))K_\lambda(g(\psi)) - K_{\lambda+1}(g(\psi)))^2}{K_{\lambda+1}(g(\psi))^2}. \tag{A70}$$

Since

$$(K_{\lambda+2}(g(\psi))K_\lambda(g(\psi)) - K_{\lambda+1}(g(\psi)))^2 > 0 \tag{A71}$$

for all λ, we have $f'(\psi) > 0$.

To proove that $\partial\alpha^*(\lambda,\chi,\psi)/\partial\lambda < 0$, it is enough to show that

$$\frac{\partial\zeta^*(\lambda,\chi,\psi)}{\partial\lambda} < 0, \tag{A72}$$

where $\zeta^*(\lambda,\chi,\psi)$ is the solution to (A66). If we define

$$f(\lambda) = \frac{\sqrt{\kappa\chi}K_\lambda(\sqrt{\kappa\chi})}{\chi K_{\lambda+1}(\sqrt{\kappa\chi})}, \tag{A73}$$

with $\kappa = \psi + A - \zeta^2 C$ and $\chi > 0$, we need to show that $f'(\lambda) < 0$ for all ζ such that $\kappa > 0$. By Lemma 2.2 of Ismail and Muldoon (1978), the function $K_{\lambda+1}(y)/K_\lambda(y)$ with $y > 0$ is strictly increasing in λ. Therefore, it is clear that $f'(\lambda) < 0$.

Appendix 9. Proof of Proposition 5

First, consider the case in which $\widehat{\psi} > \psi > \psi^\diamond > 0$. For notational convenience define $Q = Q(\mathbf{x}^*(\lambda,\chi,\psi))$, $Q^\diamond = Q(\mathbf{x}^*(\lambda,\chi,\psi^\diamond))$, $KE = KE(\mathbf{x}^*(\lambda,\chi,\psi))$, and $KE^\diamond = KE(\mathbf{x}^*(\lambda,\chi,\psi^\diamond))$. Recall that $Q(\mathbf{x})$ and $KE(\mathbf{x})$ were defined in (15) and (16) for a feasible portfolio $\mathbf{x}$. From Proposition 4, we have $KE^\diamond < 0$ and $KE < 0$; then,

$$\begin{aligned}
&\frac{1}{2}\mathrm{EUT}^*(\lambda,\chi,\psi^\diamond) \\
&\quad \geq Q - \ln\left(\left(\frac{\psi^\diamond}{\psi^\diamond - KE}\right)^{\lambda/2} \frac{K_\lambda(\sqrt{\chi(\psi^\diamond - KE)})}{K_\lambda(\sqrt{\chi^\diamond\psi})}\right) \\
&\quad > Q - \ln\left(\left(\frac{\psi}{\psi - KE}\right)^{\lambda/2} \frac{K_\lambda(\sqrt{\chi(\psi - KE)})}{K_\lambda(\sqrt{\chi\psi})}\right) \\
&\quad = \frac{1}{2}\mathrm{EUT}^*(\lambda,\chi,\psi),
\end{aligned}$$

The first inequality holds since $\mathrm{EUT}(\mathbf{x}^*(\lambda,\chi,\psi^\diamond)) \geq \mathrm{EUT}(\mathbf{x})$ for any feasible portfolio $\mathbf{x}$. In particular, it holds for $\mathbf{x}^*(\lambda,\chi,\psi)$ with $\psi > \psi^\diamond$. The second inequality is justified by Proposition A.1. Therefore,

$$\mathrm{EUT}^*(\lambda,\chi,\psi^\diamond) > \mathrm{ES}^*(\lambda,\chi,\psi) \tag{A74}$$

for $\psi > \psi^\diamond$ and $\widehat{\psi} > \psi$. Now, we consider the case in which $\widehat{\psi} < \psi^\diamond < \psi$. From Proposition 4, we have $0 < KE^\diamond < KE$, which implies $0 < \psi - KE < \psi - KE^\diamond$. Using the definition of EUT in (**P1**) and the optimality of $\mathbf{x}^*(\lambda,\chi,\psi)$, we have

$$\begin{aligned}
&\frac{1}{2}\mathrm{EUT}^*(\lambda,\chi,\psi) \\
&\quad \geq Q^\diamond - \ln\left(\left(\frac{\psi}{\psi - KE^\diamond}\right)^{\lambda/2} \frac{K_\lambda(\sqrt{\chi(\psi - KE^\diamond)})}{K_\lambda(\sqrt{\chi\psi})}\right) \\
&\quad > Q^\diamond - \ln\left(\left(\frac{\psi^\diamond}{\psi^\diamond - KE^\diamond}\right)^{\lambda/2} \frac{K_\lambda(\sqrt{\chi(\psi^\diamond - KE^\diamond)})}{K_\lambda(\sqrt{\chi\psi^\diamond})}\right) \\
&\quad = \frac{1}{2}\mathrm{EUT}^*(\lambda,\chi,\psi^\diamond),
\end{aligned}$$

where the second inequality is again justified by Proposition A.1. Consequently, we have verified that

$$\mathrm{EUT}^*(\lambda,\chi,\psi) > \mathrm{EUT}^*(\lambda,\chi,\psi^\diamond) \tag{A75}$$

for $\psi > \psi^\diamond$ and $\psi^\diamond < \widehat{\psi}$. It is important to mention that it is always possible to find $\widehat{\psi}$ such that $\alpha^*(\lambda,\chi,\widehat{\psi}) = 1$ when $\lambda \geq -1$. Defining

$$h(\psi) = \frac{\sqrt{\chi\psi}}{\chi}\frac{K_\lambda(\sqrt{\chi\psi})}{K_{\lambda+1}(\sqrt{\chi\psi})}, \tag{A76}$$

we obtain $\lim_{\psi\to 0} h(\psi) = 0$, and $\lim_{\psi\to\infty} h(\psi) = \infty$. Then, there always exists $\widehat{\psi}$ such that $h(\widehat{\psi}) = \nu$. In the case of λ, the proof is similar to the one of Proposition 5. However, it uses Proposition A.3 instead of Proposition A.1.

Appendix 10. Proof of Proposition 6

Without loss of generality we assume $aW_0 = 1$. Also, we simplify the notation indicating only the parameter of interest, if no parameter is explicitly mentioned, the results are valid for either ψ and λ. Then, using Proposition 4 and ν in (25) it is straightforward to verify that

$$\gamma^* = \nu B\alpha^* + A > 0, \tag{A77}$$

$$\partial_\psi\gamma^*(\psi) = \nu B \times \partial_\psi\alpha^*(\psi) > 0, \tag{A78}$$

$$\partial_\lambda\gamma^*(\lambda) = \nu B \times \partial_\lambda\alpha^*(\lambda) < 0. \tag{A79}$$

Expression (A77) implies that the right tail is always the heaviest when $B > 0$, i.e. $\beta^* > 0$. Additionally, we have

$$\sigma^{2*} = \nu^2 C\alpha^{*2} + 2\nu B\alpha^* + A > 0, \tag{A80}$$

$$\partial_\psi \sigma^{2*}(\psi) = (2\nu^2 C\alpha^*(\psi) + 2\nu B) \times \partial_\psi \alpha^*(\psi) > 0, \tag{A81}$$

$$\partial_\lambda \sigma^{2*}(\lambda) = (2\nu^2 C\alpha^*(\lambda) + 2\nu B) \times \partial_\lambda \alpha^*(\lambda) < 0. \tag{A82}$$

Considering the following functions

$$g_1(\lambda, \psi) = \frac{\psi}{\sigma^{2*}(\lambda, \psi)}, \quad \text{and} \quad \beta^*(\lambda, \psi) = \frac{\gamma^*(\lambda, \psi)}{\sigma^{2*}(\lambda, \psi)}, \tag{A83}$$

taking the corresponding derivatives with respect to λ and using (A79) and (A82), we have

$$\partial_\lambda g_1(\lambda) = -\frac{\psi \times \partial_\lambda \sigma^{2*}(\lambda)}{(\sigma^{2*}(\lambda))^2} > 0, \tag{A84}$$

and

$$\partial_\lambda \beta^*(\lambda) = -\partial_\lambda \alpha^*(\lambda) \times \frac{\nu \times (\nu^2 CB\alpha^*(\lambda) + 2\nu AC\alpha^*(\lambda) + AB)}{(\sigma^{2*}(\lambda))^2} > 0. \tag{A85}$$

Expressions (A84) and (A85) implies that $\partial_\lambda \vartheta_H(\lambda) < 0$ and $\partial_\lambda \vartheta_L(\lambda) < 0$. By (A78) and (A81), it is immediate to show that

$$\partial_\psi \beta^*(\psi) = \partial_\psi \alpha^*(\psi) \times \frac{\nu \times (-\nu^2 CB\alpha^*(\lambda) - 2\nu AC\alpha^*(\lambda) - AB)}{(\sigma^{2*}(\lambda))^2} < 0. \tag{A86}$$

Next, we have

$$\varphi(\psi) = \sqrt{g_1(\psi) + \beta^*(\psi)^2} = \frac{\sqrt{\nu^2\alpha^{*2}(\psi C + B^2) + 2\nu B\alpha^*(A + \psi) + A(A + \psi)}}{\nu^2 C\alpha^{*2} + 2\nu B\alpha^* + A}. \tag{A87}$$

Expressions (A81) and Proposition 4 imply $\partial_\psi \varphi(\psi) > 0$, and therefore $\partial_\psi \vartheta_H(\psi) < 0$ and $\partial_\psi \vartheta_L(\psi) < 0$.