Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Justifying Mean-Variance Portfolio Selection when Asset Returns Are Skewed

Frank Schuhmacher, Hendrik Kohrs, Benjamin R. Auer

informs
http://pubsonline.informs.org/journal/mnsc
MANAGEMENT SCIENCE
Vol. 67, No. 12, December 2021, pp. 7812–7824
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Justifying Mean-Variance Portfolio Selection when Asset Returns Are Skewed

Frank Schuhmacher,[a] Hendrik Kohrs,[a,b] Benjamin R. Auer[a,c,d,*]

[a] Department of Finance, University of Leipzig, 04109 Leipzig, Germany; [b] Department of Risk Management and Quantitative Analysis, VNG Handel and Vertrieb GmbH, 04347 Leipzig, Germany; [c] Chair of Finance, Brandenburg University of Technology Cottbus-Senftenberg, 03046 Cottbus, Germany; [d] Research Network Area Macro, Money and International Finance, CESifo Munich, 80539 Munich, Germany

*Corresponding author

Contact: schuhm@wifa.uni-leipzig.de (FS); hendrik.kohrs@vng-handel.de (HK); auer@b-tu.de, https://orcid.org/0000-0002-4795-2821 (BRA)

Received: August 5, 2019
Revised: March 17, 2020; July 1, 2020
Accepted: July 30, 2020
Published Online in Articles in Advance: January 18, 2021

https://doi.org/10.1287/mnsc.2020.3846

Abstract. We show that, in the presence of a risk-free asset, the return distribution of every portfolio is determined by its mean and variance if and only if asset returns follow a specific skew-elliptical distribution. Thus, contrary to common belief among academics and practitioners, skewed returns do not allow a rejection of mean-variance analysis. Our work differs from Chamberlain's [Chamberlain G (1983) A characterization of the distributions that imply mean-variance utility functions. *J. Econom. Theory* 29(1):185–201.] by focusing on the returns of portfolios, where the weights over the risk-free asset and the risky assets sum to unity. Furthermore, it extends Meyer's [Meyer J, Rasche RH (1992) Sufficient conditions for expected utility to imply mean-standard deviation rankings: Empirical evidence concerning the location and scale condition. *Econom. J. (London)* 102(410):91–106.] by introducing elliptical noise into their generalized location-scale framework. To emphasize the relevance of our skew-elliptical model, we additionally provide empirical evidence that it cannot be rejected for the returns of typical portfolios of common stocks or popular alternative investments.

History: Accepted by Kay Giesecke, finance.
Supplemental Material: Data and the online appendix are available at https://doi.org/10.1287/mnsc.2020.3846.

Keywords: portfolio constraints • location-scale condition • skew-elliptical distributions

1. Introduction

In the 1950s, Markowitz (1952, 1959) laid the foundations for efficient portfolio optimization frameworks by linking expected utility maximization to mean-variance analysis. His ideas have been extended in various ways and are at the core of many approaches in modern finance. The capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966) is probably one of the most important follow-up results. Despite the increasing complexity and uncertainty in global financial markets, the framework introduced by Markowitz continues to find wide application (see Markowitz 2000, Levy and Levy 2004, Bernard and Vanduffel 2014, Kolm et al. 2014, Ray and Jenamani 2016, Gao and Nardari 2018). However, the theoretical and empirical *justification of mean-variance analysis remains controversial*. There are even *widespread misunderstandings concerning the necessary and sufficient conditions for its application* (see Markowitz 2014). For example, if asset returns follow a multivariate normal distribution, or if the utility function is quadratic, it is sufficient to consider mean and variance of returns in portfolio selection because these conditions ensure consistency between the mean-variance and the expected utility model (see Baron 1977). However, many researchers argue that these conditions either have theoretical defects or lack empirical support. Quadratic utility implies increasing absolute risk aversion and that marginal utility eventually becomes negative for risk-averse agents (see Pratt 1964, Brockett and Golden 1987). Furthermore, normality of asset returns is typically rejected based on empirical evidence of skewness or heavy tails (see Mandelbrot 1963, Peiró 1999). Although these observations are correct, researchers are often unaware that there are more general conditions that imply mean-variance utility functions such that we cannot reject Markowitz's principles based on such simplistic arguments.

Two strands of literature provide more general justifications for mean-variance analysis. Empirical studies examine whether functions of mean and variance at least serve as good approximations for expected utility in practice. Markowitz (2014) reviews this issue and stresses that many authors have ignored the favorable results of this body of research. Using different methodologies, preference settings, and asset return data, Markowitz (1959), Levy and Markowitz (1979), Dexter et al. (1980), Pulley (1983), Kroll et al. (1984), Hlawitschka (1994) and Simaan (1993b, 2014) show that mean-variance-efficient portfolios can provide good approximations for expected-utility-maximizing portfolios. Hence, for the practical use of mean-variance analysis, it is necessary and sufficient to ensure that a portfolio choice derived from the mean-variance-efficient frontier approximately maximizes expected utility for a wide variety of utility functions. Although this pragmatic justification motivates intensive empirical research using different data sets, it does not explain which general circumstances justify mean-variance analysis. Identifying theoretical models that imply mean-variance portfolio utility makes it much easier to justify Markowitz's approach. Moreover, as we observe good approximations, there has to be a (at least approximate) theoretical foundation for this phenomenon.

This brings us to the theoretical literature. General consistency conditions should hold for any risk-averse agent. If we allow the utility function to be an arbitrary increasing function, we can equivalently search for conditions under which the distributions of portfolio *returns* are determined by mean and variance. This refers to the *standard* (see Black 1972, Markowitz 2000) or *canonical* (see Ingersoll 1987) portfolio optimization problem, which is characterized by a *full investment condition*, that is, the constraint that the portfolio weights of all assets (including the risk-free asset, if it exists) must sum to unity. In an important contribution, Chamberlain (1983) instead considers the distributions of all *scaled* portfolios. He argues that expected utility is a function of mean and variance for every increasing utility function if and only if every portfolio with equal mean and variance has the same distribution. Then he uses this relationship to provide some insights into which distributional assumption is *necessary and sufficient* for portfolio distributions to be determined by mean and variance. Specifically, he shows that, in portfolio selection with a risk-free asset, portfolio distributions are determined by mean and variance if and only if asset returns are jointly elliptically distributed (see Chamberlain 1983, theorem 1). Without a risk-free asset, elliptical symmetry is still important; however, there is one additional degree of freedom with respect to the asset return distributions (see Chamberlain 1983, theorem 2). His result for the portfolio problem with a risk-free asset has been cited quite frequently because it is usually the focus of academic interest – as it is in our study – and is often used by practitioners to reject the Sharpe ratio as a measure of investment performance. Another important contribution has been made by Meyer and Rasche (1992). They point out that Chamberlain's findings build on the premise of an unlimited portfolio universe, which even includes portfolios with negative wealth, and that restricting this universe extends the class of distributions implying mean-variance utility functions. Specifically, they illustrate that a generalized location-scale (GLS) distribution, which allows asymmetries in asset returns, is *sufficient* to imply mean-variance utility for the primary assets in the choice set (in other words, for very simple portfolios with only one risky asset) but not necessarily for portfolios thereof.

In contrast to Chamberlain (1983) and Meyer and Rasche (1992), who look at rather unusual portfolio universes, we consider asset combinations, which are the basis of a wide variety of portfolio applications. Our research question is as follows: Which distributions imply mean-variance utility functions in the context of standard portfolio optimization with *full investment constraint*? With this in mind, we automatically search for the distributions relevant for portfolio *returns* as opposed to those relevant for *scaled* portfolio values, which do not necessarily represent returns.[1] Our answer can be divided between a preliminary result (sufficient condition) and a main result (necessary condition, full characterization). To derive the former, we modify the GLS distribution used in Meyer and Rasche (1992) by introducing *elliptical noise*. We show that, if the returns of risky assets follow what we call skew-elliptical GLS distribution, all investments are ranked equivalently under expected utility theory and the mean-variance approach. This holds regardless whether there is a risk-free asset in the choice set or not. In comparison with Meyer and Rasche (1992), we require a larger set of portfolios to be ranked equivalently, namely, all portfolio returns. In this respect, our result is stronger but requires the specification of elliptical noise, which is not needed in the original framework of Meyer and Rasche (1992). Our preliminary result is not in conflict with the condition of elliptical symmetry contained in Chamberlain (1983, theorem 1) because the full investment restriction eliminates, for example, portfolios with negative wealth. The sufficient condition of elliptical symmetry is still sufficient. However, now there is a more general sufficient condition that directly calls for an analysis of the necessary condition for mean-variance utility. Our main result highlights that the skew-elliptical GLS distribution is both a sufficient and a necessary condition. We arrive at this core finding by two important insights. First, we show that switching the perspective from returns to *excess returns*, which is possible under the full investment constraint, allows us to connect the situation with risk-free asset and full investment constraint to the situation without risk-free asset and no full investment constraint. Consequently, when what we call skew-elliptical mean-standard deviation (MS) distribution used in Chamberlain (1983, theorem 2) holds for excess returns, it implies mean-variance utility for portfolio returns. Second, we show that this distribution is equivalent to a version of our skew-elliptical GLS distribution.

This set of results highlights that introducing a full investment constraint into Chamberlain's framework with risk-free asset – in other words, focusing on portfolio returns – extends the class of distributions implying mean-variance utility by, in particular, asymmetric distributions. It allows us to emphasize that, *in the presence of a risk-free asset*, conducting *mean-variance analysis is justified even when asset and portfolio returns are skewed*. Many seem unaware of this issue because they tend to dismiss mean-variance analysis based on the lack of empirical evidence for elliptical return symmetry (or even normality) in decision problems where all wealth is invested.[2] Our skew-elliptical GLS representation captures the relevant family of distributions by a simple two-factor linear equation and allows risk to be broken down into two components: an elliptical variance component and a nonelliptical component that covers skewness (and higher odd cumulants) of asset returns. This simplicity (in contrast to the rather abstract definition of Chamberlain's skew-elliptical MS distribution) can help to make the conditions for mean-variance utility more accessible to academics and practitioners and thus prevent misunderstandings in future research. Furthermore, our GLS model is a valuable tool for the derivation of additional results, including more intuitive ways of obtaining results of the early mean-variance literature and the generalization of results provided in more recent research. We give some examples of such applications after the derivation and discussion of our main findings.

To round off our theoretical results, we finish our study by presenting some evidence on the empirical relevance of our skew-elliptical GLS model. To this end, we modify the testing approach proposed by Meyer and Rasche (1992) to be applicable to our specific GLS case. Roughly speaking, our skew-elliptical GLS model can be tested by checking normalized out-of-sample CAPM residuals for sphericity (a special case of ellipticality) via the uniform techniques of Liang et al. (2008). Using a sample of common stocks, well-known factor portfolios (capturing size, value, momentum, reversal and industry effects), and alternative investment vehicles (i.e., indices reflecting the performance of commodity futures and hedge funds), we find that the skew-elliptical GLS distribution cannot be rejected in the majority of considered settings.

2. Preliminary Result

Let $X = (X_1, \ldots, X_n)'$ be a $n \times 1$ vector of stochastic (primary) asset returns with finite variances and $r > 0$ the nonstochastic return on the risk-free asset, if there is one.[3] Suppose that we have an initial budget $B > 0$ to construct a portfolio consisting of these assets. In this context, we have to decide on the amount of capital $w_i B$ to be assigned to each asset i $(i = 0, 1, \ldots, n)$, where $w_i \in \mathbb{R}$ is a weighting factor for asset i. The term w_0 refers to the risk-free asset, whereas the terms for the risky assets are collected in the vector $w = (w_1, \ldots, w_n)'$. We allow the weights to take negative values, that is, short sales are permitted. To preserve the budget constraint, however, the weights have to sum to unity. That is, we introduce the full investment condition $w_0 + w'e = 1$, where e is a vector of ones, such that the sum of invested capital equals B. If the full investment condition is imposed, the linear combination $P := w_0 r + w'X$ can be interpreted as the return of a portfolio of assets weighted by $(w_0, w) \in \mathbb{R}^{n+1}$. In standard portfolio optimization, we choose portfolio weights by maximizing the expected utility of the portfolio return (see Ingersoll 1987, section 3.1). That is, we solve the optimization problem

$$\max_{(w_0,w)\in\mathbb{R}^{n+1}} \mathbb{E}[u(w_0 r + w'X)]$$
$$\text{subject to } w_0 + w'e = 1, \tag{P}$$

where the utility function u represents our preferences. In the portfolio selection process, the full investment constraint ensures a well-posed optimization problem, which delivers optimal asset ratios, that is, relative amounts of investment capital, for each asset. Consequently, all investors can utilize the results of the optimization regardless of the sum they wish to invest because they can proportionally adapt an obtained optimal portfolio to their needs.

With this focus, it is clear that theoretical justification of mean-variance analysis has to build upon portfolio *returns*. This is important and often overlooked in articles using Chamberlain (1983) to reject mean-variance analysis. By pushing the limits of an initially introduced budget constraint to infinity in the "early" optimization steps, he describes the distributions of an unrestricted universe of scaled portfolios $B(w_0 r + w'X)$ for an arbitrary budget $B \in \mathbb{R}$. From a scientific perspective, there is nothing wrong with this approach. However, readers tend to misinterpret the results. Chamberlain's findings cannot be used to reject mean-variance analysis based on skewness in empirical returns because he is not looking at portfolio returns.[4]

By imposing the full investment restriction, we limit the portfolio universe to practically relevant constellations and simultaneously deal with this interpretation problem because the restriction enforces a portfolio return focus in our distributional considerations. As emphasized by Meyer and Rasche (1992), when additional constraints are imposed, larger families of distributions can be sufficient to justify mean-variance analysis. In the following, we show that our constraint leads to a significant and highly relevant expansion.

Within our standard portfolio optimization setting, the following distribution will be essential for our results:

Definition (Skew-elliptical GLS distribution)**.** A random vector $X \in \mathbb{R}^n$ is said to have a skew-elliptical generalized location-scale distribution with constant $r \in \mathbb{R}$, if its components X_i $(i = 1, \dots, n)$ can be written as

$$X_i = r + \beta_i \cdot Y + \gamma_i \cdot Z_i, \tag{1}$$

where, conditional on Y, the vector $Z = (Z_1, \dots, Z_n)'$ is spherically distributed and Y is a real-valued random variable with $\mathbb{E}[Y] \neq 0$ and $\mathrm{Var}[Y] = 1$. The coefficients β_i, γ_i are real numbers with $\beta_i \neq 0$ for at least one $i = 1, \dots, n$.

Meyer and Rasche (1992) introduced the GLS representation in a more general form, that is, they did not restrict the specific distributional form of Y and Z, and showed that their GLS model leads to mean-variance utility functions for the primary assets in the choice set but not necessarily for portfolios thereof. In order to analyze not only primary assets but also nontrivial portfolios of the primary assets, we differ from their specification by defining Z to be spherical while leaving Y untouched. Note that spherical distributions are invariant under orthogonal transformations and, per definition, all Z_i $(i = 1, \dots, n)$ are identically distributed with zero mean and unit variance. We do not require $\gamma_i \geq 0$ because the distribution of the Z_i is symmetric around the origin such that Z_i and $-Z_i$ have identical distributions.[5] Our denomination "skew-elliptical" arises from two facts. First, the noise vector $\xi := (\gamma_1 Z_1, \dots, \gamma_n Z_n)'$ is elliptically distributed about the origin because elliptical random variables are generated by affine transformations of spherically distributed ones.[6] Second, although the spherical component Z is symmetric, the distribution of the nonspherical component Y is not restricted. This means that a skewed Y can introduce skewness into X (and higher odd moments are influenced analogously). Fat tails of X can originate from excess kurtosis in Z or Y.

Formulation (1) has the appealing property that it can be interpreted as a market model (see Elton et al. 2007, chapter 7), where Y and Z_i are systematic and unsystematic components driving the returns X_i of asset i. Thus, from this perspective, our distributional specification of Y and Z implies that nonzero skewness in asset returns comes primarily from market-wide shocks. A risk-free asset can be easily introduced, if $r > 0$ is interpreted as the risk-free rate of return. That is, if $\beta_i = 0$ and $\gamma_i = 0$ for one $i = 1, \dots, n$, we have the risk-free return $X_i = r$. This is important because, at first glance, it may appear that some of the following theoretical results only hold for portfolios of risky assets, when, in fact, they also hold for portfolios including the risk-free asset.

In general, any linear combination of an elliptically distributed random vector is also elliptical (see McNeil et al. 2005, section 3.3.3). With our modification, the GLS distribution receives a similar *stability* property. That is, with elliptical noise, a linear combination of GLS random variables inherits the GLS distribution from its components. Because this important feature drives our preliminary result, we state it in a lemma.

Lemma 1. *Let $X = (X_1, \dots, X_n)'$ be a random vector satisfying the skew-elliptical GLS property* (1)*. Then, every linear combination $P = w'X$ with $w \in \mathbb{R}^n$ subject to $w'e = 1$ belongs to the same skew-elliptical GLS family* (1)*.*

Proof. See appendix. □

We are now in a position to formulate our preliminary result (sufficient condition). If we interpret the X_i as asset returns and if they satisfy (1), the portfolio returns P have a representation of the form (1) as well. In other words, portfolio building does not enlarge the family of return distributions. The stability (or transmission) property has the consequence that not only are the distributions of the primary asset returns determined by mean and variance, so are those of portfolio returns. The following proposition captures this result.

Proposition 1. *If a vector X of asset returns satisfies the skew-elliptical GLS property* (1)*, then the distribution of a portfolio return $P = w'X$ is determined by its mean and variance for every $w \in \mathbb{R}^n$ with $w'e = 1$.*

Proof. See appendix. □

Researchers and practitioners often use the ellipticality statement of Chamberlain (1983, theorem 1) and the observation that empirical returns tend to be asymmetric to argue that mean-variance analysis should be abandoned. In addition to our previous discussion, Proposition 1 shows that Chamberlain (1983, theorem 1) cannot be used in such a way. It highlights that, in portfolio selection with a risk-free asset, mean-variance analysis is justified even when asset returns and portfolio returns are skewed. We can nicely see that the set of relevant distributions significantly extends from distributions with no skewness to distributions with skewness, when the full investment condition is imposed, that is, when we are interested in a theory for portfolio returns. Intuitively, the reason for this is that fewer portfolios have to be ranked such that the conditions for equal rankings under expected utility and mean-variance framework relax.

3. Main Result

Chamberlain (1983, theorem 1) provides a full characterization of the distributions that imply mean-variance utility functions for all scaled portfolios, including those with negative wealth. Because our preliminary result shows that the full investment condition crucially influences his result, that is, enlarges the set of distributions that imply mean-variance utility functions, we now aim at the question of whether we can provide a full characterization of the distributions that imply mean-variance utility for portfolio returns.

Interestingly, it turns out that the mathematical tools for deriving an answer can be found in Chamberlain (1983, theorem 2) because the setting of this theorem and our portfolio selection problem exhibit, in a specific way, the same degrees of freedom of distributional assumptions for portfolio returns to be determined by mean and variance. The family of distribution relevant in this theorem is defined as follows.

Definition (Skew-Elliptical MS Distribution)**.** A random vector $X \in \mathbb{R}^n$ is said to have a skew-elliptical mean-standard deviation distribution if it is a linear transformation of a random vector in which the last $n-1$ components are spherically distributed conditional on the first component, which has an arbitrary distribution. Formally, we require

$$TX = \begin{pmatrix} m \\ S \end{pmatrix}, \tag{2}$$

where T is a nonsingular matrix, m is a real-valued random variable with $\mathbb{E}[m] \neq 0$ and $S = (S_2, \dots, S_n)'$ is a random vector with $S|m$ being spherically distributed.

According to Chamberlain (1983, theorem 2), in the absence of a risk-free asset, the distribution of every scaled portfolio is determined by its mean and variance if and only if the primary assets have a skew-elliptical MS distribution (2). Proposition 1 shows that the skew-elliptical GLS distribution (1) applied to asset returns is sufficient for mean-variance-determined portfolio returns. Our first main result is that these two distributions can be considered identical.[7]

Theorem 1. *A random vector $X \in \mathbb{R}^n$ follows the skew-elliptical MS distribution (2) if and only if it follows the skew-elliptical GLS distribution (1) with no constant.*

Proof. See appendix. □

Because the *skew-elliptical MS distribution and the skew-elliptical GLS distribution are different representations of the same family of distributions*, our GLS framework provides a concise summary of skewed distributions implying mean-variance utility functions.

With this insight, we can now turn to our core objective: the full characterization of which distributional assumptions are necessary and sufficient for portfolio returns to be determined by mean and variance. To our surprise, its derivation is quite simple. The introduction of the full investment condition in the situation with a risk-free asset and n risky assets has the consequence that the portfolio excess return is equal to the weighted sum of asset excess returns:[8]

$$P - r = w_1(X_1 - r) + \cdots + w_n(X_n - r). \tag{3}$$

Otherwise, this equality would not hold. If we denote the risky asset excess returns with $\bar{X}_i := X_i - r$ and the portfolio excess return with $\bar{P} := P - r$, we can rewrite the portfolio excess return as $\bar{P} = w_1\bar{X}_1 + \cdots + w_n\bar{X}_n$. Because, formally, the risk-free asset and the full investment condition "disappear," this is exactly the situation analyzed in Chamberlain (1983, theorem 2). Hence, the distribution in Chamberlain (1983, theorem 2) applied to excess returns characterizes the distributions that imply mean-variance utility functions for portfolio returns.[9] With the help of Theorem 1, we are now in a position to formulate our second main result.

Theorem 2. *Assume there exists at least one $i = 1, \dots, n$ such that $\mathbb{E}[X_i] \neq r$, where X_i is the ith element of the risky asset vector X. In the presence of a risk-free asset, the distribution of portfolio returns $P = w_0 r + w'X$ is determined by its mean and variance for every $(w_0, w) \in \mathbb{R}^{n+1}$ with $w_0 + w'e = 1$ if and only if the asset returns X have a skew-elliptical GLS distribution (1) with r being the risk-free rate.*

Proof. See appendix. □

Theorem 2 contains the core message of our study: *elliptical symmetry is not necessary for mean-variance utility of portfolio returns*. In standard portfolio optimization, mean-variance analysis is justified even when returns are skewed.

To put an end to the longstanding debate on the distributions that imply mean-variance utility for portfolio returns, we require one more theoretical step. Theorem 2 presents a necessary and sufficient condition for the situation with a risk-free asset. For the situation *without a risk-free asset*, Proposition 1 shows that the same condition is sufficient. Recall that Chamberlain (1983, theorem 2) combined with our Theorem 1 states that, in the latter situation, the skew-elliptical GLS distribution (1) is a necessary condition if we consider all scaled portfolios. Thus, the question arises whether introducing the full investment constraint also relaxes Chamberlain (1983, theorem 2) such that a larger family of distributions justifies mean-variance analysis. As illustrated in the following, it does indeed. We assume that there is a risky asset, say $X_0 = \mathcal{R}$, such that $X_i - \mathcal{R}$ and $\mathcal{R}$ are stochastically independent for each $i = 1, \ldots, n$. If $\mathcal{R}$ is just a constant, this condition holds vacuously true and $\mathcal{R}$ can again be interpreted as the risk-free rate of return. The premise of this setup, which might look artificial at first glance, is that with a variable $\mathcal{R}$, which can be treated just as the constant risk-free rate, our proof can be based on the same chain of reasoning that we used before.[10]

Corollary 1. *Let $X_0 = \mathcal{R}$ be a risky asset such that $X_i - \mathcal{R}$ and $\mathcal{R}$ are stochastically independent for each $i = 1, \ldots, n$ in the risky asset vector X. Further, assume $\mathbb{E}[X_i] \neq \mathbb{E}[\mathcal{R}]$ for at least one $i = 1, \ldots, n$. Then, in the absence of a risk-free asset, the distribution of a portfolio return $P = w_0 X_0 + w'X$ is determined by its mean and variance for every $(w_0, w) \in \mathbb{R}^{n+1}$ with $w_0 + w'e = 1$ if and only if the asset returns X_i have a skew-elliptical GLS distribution of the form*

$$X_i = \mathcal{R} + \beta_i \cdot Y + \gamma_i \cdot Z_i, \tag{4}$$

where, conditional on Y, the vector $Z = (Z_1, \ldots, Z_n)'$ is spherically distributed, Y is a real-valued random variable with $\mathbb{E}[Y] \neq 0$ and $\mathrm{Var}[Y] = 1$, and $\mathcal{R}$ is independent of Y and Z.

Proof. See appendix. □

Instead of having a constant risk-free rate, the family of distributions (4) offers space for another risky return component $\mathcal{R}$. However, because there is no additional coefficient to choose, this enlargement is rather unimportant. We did not expect it to be significant because, if we wish to remain in a mean-variance world, introducing a real third degree of freedom is not possible.[11] The variance measures symmetric uncertainty such that one component of the distribution also has to be symmetric. The other component, which corresponds to the mean, allows an arbitrary distribution. Thus, the skew-elliptical GLS distribution (1) already offers the maximum degrees of freedom.

4. Additional Results

Our analysis has produced many additional theoretical results because, based on our framework, findings of the portfolio literature can be derived and/or generalized quite easily. So as not to distract the reader from our core message, this section presents a summary of these extras. The full discussion and technical details are located in Sections A to D of our online appendix.

First, we can alternatively derive the sufficiency part of Chamberlain (1983, theorem 1) by realizing that our GLS model contains an elliptical version of the Sinn (1983) and Meyer (1987) location-scale (LS) model. If this special case holds, the distributions of all scaled portfolios and portfolio returns are determined by mean and variance. This result also helps to illustrate the distributional expansion we observe with the introduction of the full investment constraint because we are jumping from a LS model to a GLS model.

Second, we generalize the findings of Schuhmacher and Auer (2014). Most importantly, we show that under our skew-elliptical GLS model and some mild axiomatic requirements (with respect to the quantification of risk), the investment performance rankings of multiasset portfolios are identical regardless of the choice of performance measure. This explains the so far puzzling empirical observation that the Sharpe ratio and many other popular measures tend to produce the same rank ordering of investment funds with highly non-normal return distributions (see Eling and Schuhmacher 2007). It also highlights that switching from the Sharpe ratio to conceptionally different (typically asymmetric) alternatives − an approach practitioners often promote for skewed return environments (see Bacon 2008) − does not necessarily improve the investment process. In other words, using a symmetric risk measure in an asymmetric (skew-elliptical GLS) world still leads to the optimal portfolio decision because the efficient sets of various mean-risk models are identical.

Third, by linking our GLS model to the framework of Simaan (1993a), we can adopt his efficient set calculations to show that the famous two-fund separation property of Tobin (1958) holds in our setting and to derive the corresponding tangency portfolio formula.

Finally, we illustrate that the parameters determining the distribution of portfolio returns can be identified in a very elegant fashion and derive a *generally* applicable two-step guide for use in future research. According to this guide, in the first step, we have to show that a given family of distributions is stable with respect to portfolio formation. If this is true, we can, in the second step, identify the parameters that determine the distribution of a primary asset belonging to this family by following the implicit function approach used in our proofs and in Meyer and Rasche (1992). The resulting parameters consequently determine the return distribution of the primary asset and the return distribution of portfolios.

5. Empirical Perspective

Even though the focus of our study is theoretical in nature, we shed some light on whether our skew-elliptical GLS distribution has empirical support. Meyer and Rasche (1992) propose an intuitive testing procedure for their original GLS model, which uses

Management Science, 2021, vol. 67, no. 12, pp. 7812–7824, © 2021 INFORMS

the CAPM to fill it with life, that is, to establish a link to empirically observable variables. We adopt their general idea but, at a certain point, have to deviate from it because our GLS model is more specific.

According to the CAPM, the random return of an asset or portfolio i is given by $X_i = r + \beta_i(X_m - r) + \xi_i$, where X_m is the random return of the market portfolio, β_i governs systematic risk and ξ_i represents the random idiosyncratic component of the return. In this model, X_m and ξ_i are independent and the ξ_i are independent of one another. If the ξ_i follow the form $\xi_i = d + \gamma_i Z_i$, then the return of i becomes $X_i = (r + d) + \beta_i(X_m - r) + \gamma_i Z_i$, which is of the original GLS form. Consequently, to obtain empirical evidence in favor of the original GLS condition in a sample of N $(i = 1, \ldots, N)$ assets with T $(t = 1, \ldots, T)$ returns each, we should focus on the properties of $Z_{it} = [X_{it} - (r + d) - \beta_i(X_{mt} - r)]/\gamma_i$.

To obtain empirical observations of Z_{it}, Meyer and Rasche (1992) use a recent 30-year sample of stocks with continuous NASDAQ trading history. In our study, such a selection yields monthly total return data from January 1989 to December 2018 for 503 stocks. The market portfolio is proxied by the Center for Research in Security Prices (CRSP) value-weighted index, and the risk-free rate is captured by the one-month Treasury bill rate (from Ibbotson Associates).[12] The stock returns are obtained from Thomson Reuters Datastream, whereas the two remaining variables are taken from Kenneth R. French's data library.[13] The data are split into two subsamples, A and B, of the same number of return observations for each stock. Subsample A serves to construct values for the parameters determining the Z_{it}. That is, the excess returns for each stock are regressed against the contemporary excess return of the market. No intercept is included in the regression equation. The regression slope coefficient for each stock is used as its β_i. Next, the β_i are used to obtain the $\xi_{it} = (X_{it} - r) - \beta_i(X_{mt} - r)$. The average of ξ_{it} over all i and t gives d, whereas the standard deviation of ξ_{it} for each i delivers γ_i. In subsample B, these parameter values then allow the calculation of the Z_{it}.

Because this procedure only looks at individual stocks, Meyer and Rasche (1992) employ a vector sampling procedure to construct portfolios. In contrast to pure random sampling, this method selects blocks of stocks from a beta-sorted stock vector such that the resulting portfolios do not overlap (allowing independent idiosyncratic risk terms) and provide diversity with respect to β_i. Within each portfolio, the stocks are weighted equally (as often in practice; see DeMiguel et al. 2009) and the Z_{it} of the portfolio are determined similar to individual stocks. To illustrate typical empirical test results, we use $K = 3, 5, 7$ portfolios of $S = 2, 4, 6, 8$ stocks in each, which are common sizes for individual investors (see Statman 1987, Kumar and Lim 2008, Barasinska et al. 2012).[14]

To test whether there is empirical support for their GLS specification, Meyer and Rasche (1992) analyze whether empirical realizations of Z_{it} are likely to have resulted as independent samples of the same population and whether their mean is significantly different from zero or not. Specifically, they use Kolmogorov-Smirnov (KS) multisample tests with simulated probability distribution (see Gardner et al. 1980) and classic t-tests for this purpose.[15] Implicitly testing a *joint hypothesis* that the GLS condition holds across portfolios and that its variables and parameters are adequately chosen, they cannot reject their model in almost all settings they consider. Because we closely follow Meyer and Rasche (1992), our testing procedure will have a similar joint nature. In other words, we build a *conservative* setup, which is not tilted in favor of the GLS model because the additional components of the joint hypothesis (such as our potentially suboptimal market portfolio choice for the common factor) make our tests more likely to reject the GLS distribution.[16]

Because our GLS requirements are more restrictive with respect to the model constant (risk-free rate) and the noise component (elliptical), we drop d and cannot apply the KS test. At first glance, we might jump to established tests for elliptical symmetry, like the ones of Manzotti et al. (2002), Schott (2002), or Huffer and Park (2007), and apply them to the ξ_{it} of subsample B.[17] However, testing for spherical symmetry of the Z_{it} is more suitable because, apart from several technical advantages (related to scaling parameter measurement), it takes into account that the mean of the noise has to be zero. In general, there are many statistics for this purpose. However, most of them either converge slowly to their limiting distributions or cannot be easily evaluated in numerical computation. We use some of the tests recently outlined and extended in Liang et al. (2008, 2019) because they do not suffer from such shortcomings.

According to Liang et al. (2008), the null hypothesis that the cumulative distribution function underlying an independently identically distributed K-dimensional sample $\{z_{tj}\}_{t=1,\ldots,T, j=1,\ldots,K}$ is spherical can be tested via univariate or multivariate uniform statistics calculated with observations $\{v_{tj}\}_{t=1,\ldots,T, j=1,\ldots,K-1}$ originating from well-defined transformations of the sample data.[18] This is because a test for spherical symmetry can be substituted by a test of the null hypothesis H_0^1: *the* v_{tj} *are uniformly distributed in* $(0, 1)$ or a test of the null hypothesis H_0^2: *the* $v_t = (v_{t1}, \ldots, v_{t,K-1})'$ *are uniformly distributed in the hypercube* $[0, 1]^{K-1}$. Based on extensive simulations, Quesenberry and Miller Jr. (1977) and Miller Jr. and Quesenberry (1979) recommend the (modified) Watson (1961, 1962) statistic or the Neyman (1937) smooth statistic with fourth-degree polynomials for testing univariate uniformity. Critical values for these tests have been provided by Stephens (1970) and Miller Jr. and Quesenberry (1979), respectively. As far as multivariate uniformity is concerned, Liang et al. (2001) propose two statistics: one asymptotically normal and one asymptotically chi-square with two degrees of freedom. Each of them can be computed using three different measures of discrepancy: symmetric, centered, and star (see Hickernell 1998). For all tests, high statistics (in absolute terms) indicate evidence of nonuniformity and thus nonsphericality. In the following, we implement the (modified) Watson (1961, 1962) test and the normal form of the Liang et al. (2001) test because, in comparison with the two alternatives, they are characterized by better size and power under many known spherical distributions.

Table 1 presents our test results. We find that the null hypothesis of sphericity – and thus our GLS model – cannot be rejected in the majority of cases. Given that our approach is conservative and yet the number of rejections is low, this is strong evidence that our distribution model is relevant. As expected, the few observable rejections mainly occur for less diversified portfolios with $S = 2$.[19] There are also some rejections for more diversified portfolios with $S = 8$. Similar to the empirical findings of Meyer and Rasche (1992), there is no clear-cut relationship between rejection and portfolio size. That is, test statistics do not monotonically fall with a rising number of assets within the portfolios. However, they tend to be somewhat higher when more portfolios enter the test procedure. All this also holds in other sampling exercises.[20]

To analyze larger portfolios (with different weighting schemes) and to go beyond our initial (limited) stock selection, we also test whether well-known portfolios of French's library, which are typically used as benchmarks in asset pricing studies and have high relevance in institutional trading (see Ang 2014, chapter 10), are in line with our GLS model. That is, we perform additional tests for 6 size/book-to-market, 6 size/momentum, 6 size/reversal, as well as 5 and 10 industry portfolios (defined and used in, for example, Fama and French 2012, Abhakorn et al. 2013). Furthermore, we look at two additional asset classes by considering the 5 futures-based subindices of the Goldman Sachs Commodity Index (GSCI) and the 13 strategy subindices of the Credit Suisse Hedge Fund Index (CSHFI).[21] Although the stock portfolios and the commodity indices have the same sample size as our previous analysis, the availability of the hedge fund data are limited to the period from January 1994 to December 2018.

Table 2 presents our additional test results. Based on the Watson statistic, our stock, commodity, and hedge fund data are not in conflict with a skew-elliptical GLS world. The test statistics are insignificant for all data sets. Turning to the different specifications of the normal Liang test, we see a slightly different picture. Although there are still many instances of nonrejection, the hedge fund results may appear less fortunate. However, this is only a slight

Table 1. NASDAQ Stock Returns

		Liang		
	Watson	Symmetric	Centered	Star
3 Portfolios				
2 stocks	**0.35**	0.03	**4.30**	2.17
4 stocks	0.08	0.16	0.42	0.67
6 stocks	0.03	0.38	0.47	−0.07
8 stocks	0.05	−1.05	0.31	0.27
5 Portfolios				
2 stocks	0.07	1.50	0.62	1.12
4 stocks	0.05	−0.91	0.68	0.25
6 stocks	0.04	0.34	−0.60	0.29
8 stocks	**0.31**	**−2.65**	2.37	**−3.41**
7 Portfolios				
2 stocks	**0.74**	1.83	0.89	**5.86**
4 stocks	0.10	1.65	0.96	**2.91**
6 stocks	0.13	1.34	−0.05	−0.89
8 stocks	0.19	0.15	1.73	−1.53

Notes. Following the procedure outlined in the main text, this table tests the null hypothesis of sphericity against the alternative of nonsphericity within a universe of NASDAQ stocks. For several sets of portfolios with specific sizes, we obtain CAPM-filtered normalized noise, transform the noise as suggested by Liang et al. (2008), and report the uniform statistics of Watson (1961, 1962) and Liang et al. (2001), where the latter is calculated in its normal version. A significant statistic implies that the normalized noise may be considered nonspherical and the return data are in conflict with our GLS model specification. Significance at the 1% level is highlighted in bold print.

Table 2. Fama-French Portfolios and Additional Asset Classes

		Liang		
	Watson	Symmetric	Centered	Star
Additional stock portfolios				
FF 6 size/book-to-market	0.21	**5.55**	**2.98**	−0.18
FF 6 size/momentum	0.11	**4.53**	1.72	0.70
FF 6 size/reversal	0.21	1.39	1.72	−1.99
FF 5 industries	0.04	1.39	1.22	−1.32
FF 10 industries	0.05	1.11	**2.60**	−1.60
Additional asset classes				
GS 5 commodity indices	0.03	1.40	1.90	0.11
CS 13 hedge fund indices	0.23	**8.78**	**8.66**	**4.16**

Notes. This table repeats the tests of Table 1 for additional data sets. Specifically, we use the well-known Fama-French (FF) size, book-to-market, momentum, reversal, and industry portfolios. Furthermore, we look at the subindices of the Goldman Sachs (GS) commodity futures index and the Credit Suisse (CS) hedge fund index. Again, significance at the 1% level is marked bold and implies rejection of our GLS specification.

bump in the otherwise supportive nature of our empirical results because investors typically trade individual hedge funds that are more GLS compatible (see Schuhmacher 2012). Furthermore, we must keep in mind that we are testing a joint hypothesis and that rejection may simply stem from the fact that, instead of the market portfolio proxy, a more suitable variable capturing common variation of returns is needed.

Because our conclusions may be sensitive to some of the settings in our research design, we have conducted several robustness checks with respect to a potential survivorship bias, the choice of sample split in the filter procedure, and the selected common factor in the case of commodity and hedge fund data. The results, presented in Section E of our online appendix, provide additional support for the GLS model and the practical relevance of our work.

6. Conclusion

Almost 70 years after Markowitz's groundbreaking work on mean-variance portfolio theory and decades of ambiguity over whether its application is justified when returns are skewed, we build on the important contributions of Chamberlain (1983) and Meyer and Rasche (1992) to offer a resolution: if there is a risk-free asset and we introduce a full investment condition, that is, focus on the distributions of portfolio returns, then mean-variance analysis is justified even when asset and portfolio returns are skewed. This cuts across the widespread belief that mean-variance analysis should be abandoned in the case of asymmetric return distributions. As we also show, using the symmetric risk measure variance in an asymmetric world even leads to the same optimal portfolio decisions as popular asymmetric risk measures.

The key to this result is the discovery that, if the skew-elliptical MS distribution used in Chamberlain (1983, theorem 2) holds for excess asset returns, it completely characterizes the distributions that imply mean-variance utility for portfolio returns. With the additional insight that this class of distributions is identical to an extension of the GLS family of Meyer and Rasche (1992), which we obtain by specifying elliptical noise, we can provide a concise representation of the distributions that justify mean-variance analysis. The resulting factor-model-like framework is not only easy to interpret but also allows the findings of previous studies to be reproduced and generalized in a quick and elegant fashion. In some empirical tests, we additionally show that its validity cannot be rejected in several data sets containing, for example, portfolios typically held by private investors or portfolios popular in asset pricing tests.

We have imposed a wealth restriction. Another common assumption in portfolio selection bans short sales (see Board and Sutcliffe 1994, Jagannathan and Ma 2003, Agarwal and Naik 2004, Grundy et al. 2012), that is, requires all portfolio weights to be nonnegative. Meyer and Rasche (1992) emphasize that the necessity of elliptical symmetry in Chamberlain (1983) is related to the fact that, for each portfolio, the mirror counterpart likewise has to be appropriately evaluated. Hence, future theoretical research might examine whether banning short sales also enlarges the family of distributions that are sufficient for portfolio return distributions to be determined by mean and variance. To derive specific results, one might think of modifying the assumption that the noise in our GLS model is elliptically distributed. A closer look at the stability discussion in Cass and Stiglitz (1970) and Owen and Rabinovitch (1983) may be a suitable starting point for such an endeavor.[22] A separate analysis of self-financing portfolios, where the sum of the investment weights is zero (see Korkie and Turtle 2002), is similarly interesting. As far as additional empirical work is concerned, Schuhmacher and Auer (2014) provide some thoughts on how we might find or construct a common GLS factor that is not a market portfolio. Consequently, our empirical setup should be considered as the starting point of a full-scale empirical analysis of the skew-elliptical GLS model. Given the considerable flexibility of the model, in that we can set any distribution for the common factor, Monte Carlo studies may also, in a wide variety of distributional settings, compare the size and power characteristics of the many alternative methods (identifiable in the references of Liang et al. 2008) we have to test our GLS model.[23]

Acknowledgments

The authors thank Kay Giesecke (the department editor), an anonymous associate editor, and three anonymous reviewers for valuable comments and suggestions, which have significantly improved the quality of their manuscript. The authors also thank Fred Huffer and Jiajuan Liang for personal correspondence on ellipticality and sphericity testing. Finally, the authors are indebted to Horst Rottmann, Julia Mehlitz, and Kerstin Lamert for statistical and mathematical comments.

Appendix. Proofs

Proof of Lemma 1. Assume that the X_i $(i = 1, \dots, n)$ satisfy (1) where Y is a real-valued random variable and, using Endnote 6, $Z|Y \sim E_n(0, \Sigma, \psi)$ is n-dimensionally elliptically distributed about the origin with covariance matrix Σ and characteristic generator ψ. Note that $\sum_{i=1}^n w_i r = r \sum_{i=1}^n w_i = r$, if $\sum_{i=1}^n w_i = 1$. For $w \in \mathbb{R}^n$ with $w'e = 1$, we hence have $P = w'X = r + Y \sum_{i=1}^n w_i\beta_i + \sum_{i=1}^n w_i\gamma_i Z_i$. Set $\tilde{\beta} := \sum_{i=1}^n w_i\beta_i$ and $\tilde{Z} := \sum_{i=1}^n w_i\gamma_i Z_i$. Because linear combinations of elliptical random vectors remain elliptical with the same characteristic generator (see McNeil et al. 2005, section 3.3.3), it follows that $\tilde{Z}|Y \sim E_1(0, \sigma^2, \psi)$ is elliptically distributed with variance $\sigma^2 = c'\Sigma c$, where $c := (w_1\gamma_1, \dots, w_n\gamma_n)'$. □

Proof of Proposition 1. Assume that X satisfies (1). From Lemma 1, it follows that portfolio returns $P = w'X$ belong to the same skew-elliptical GLS family for every $w \in \mathbb{R}^n$ with $w'e = 1$. To prove the proposition, it is hence sufficient that the distribution of every primary asset X_i $(i = 1, \dots, n)$ is determined by its mean and variance. The two equations $\mathbb{E}[X_i] = r + \beta_i\mathbb{E}[Y] + \gamma_i\mathbb{E}[Z_i] = r + \beta_i\mathbb{E}[Y]$ and $\mathrm{Var}[X_i] = \beta_i^2\mathrm{Var}[Y] + \gamma_i^2\mathrm{Var}[Z_i] = \beta_i^2 + \gamma_i^2$ can be solved for β_i and $|\gamma_i|$, that is, $\beta_i = (\mathbb{E}[X_i] - r)/\mathbb{E}[Y]$ and $|\gamma_i| = (\mathrm{Var}[X_i] - ((\mathbb{E}[X_i] - r)/\mathbb{E}[Y])^2)^{\frac{1}{2}}$. Note that $\mathbb{E}[Y] \neq 0$ such that these terms are well defined. Recall that Z_i and $-Z_i$ have the same distribution such that β_i and $|\gamma_i|$ are in fact enough to determine the distribution of X_i. □

Proof of Theorem 1. We have to prove that the random vector X satisfies (2) if and only if X satisfies (1) with no constant.

"⇐": Assume that X satisfies condition (1) with no constant, that is, $X_i = \beta_i Y + \gamma_i Z_i$ for $i = 1, \dots, n$. Additionally, we have by definition $\beta_i \neq 0$ for one $i = 1, \dots, n$. Without loss of generality, let $\beta_1 \neq 0$.

First step: Find a nonsingular matrix T such that $\mathbb{E}[S] = 0$ and $\mathrm{Cov}[S] = I_{n-1}$, where $(m, S')' = TX$.[24] First, choose

$$T_1 := \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\beta_2/\beta_1 & 1 & 0 & \cdots & 0 \\ -\beta_3/\beta_1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\beta_n/\beta_1 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

and set $\tilde{X} := (\tilde{X}_1, \dots, \tilde{X}_n)' := T_1 X$. Note that we have $\tilde{X}_i = \beta_i Y + \gamma_i Z_i - \beta_i/\beta_1 \cdot (\beta_1 Y + \gamma_1 Z_1) = \gamma_i Z_i - \gamma_1\beta_i/\beta_1 Z_1$ and hence $\mathbb{E}[\tilde{X}_i] = 0$ for $i = 2, \dots, n$. Second, set $\tilde{\Sigma} := \mathrm{Cov}[(\tilde{X}_2, \dots, \tilde{X}_n)]$ and let $\tilde{\Sigma} = Q\Lambda Q'$ be its spectral decomposition, that is, Q is an orthogonal matrix and $\Lambda = \mathrm{diag}(\lambda_2, \dots, \lambda_n)$ is a diagonal matrix with $\lambda_i > 0$ for $i = 2, \dots, n$ (see Strang 2009, section 6.4). Define the block matrix

$$T_2 := \begin{pmatrix} 1 & 0 \\ 0 & \tilde{\Sigma}^{-\frac{1}{2}} \end{pmatrix},$$

where $\tilde{\Sigma}^{-\frac{1}{2}} := Q\Lambda^{-\frac{1}{2}}Q'$ is the inverse of the positive definite square root of $\tilde{\Sigma}$ and we have $\Lambda^{-\frac{1}{2}} := diag(\lambda_2^{-\frac{1}{2}}, \dots, \lambda_n^{-\frac{1}{2}})$. Let $(m, S')' := TX$, where $T := T_2 T_1$. Note that T is nonsingular because both T_1 and T_2 are nonsingular. Further, we have $\mathbb{E}[S] = 0$ because of T_1 and $\mathrm{Cov}[S] = I_{n-1}$ because of T_2. This is because $\mathrm{Cov}[TX] = \mathrm{Cov}[T_2\tilde{X}] = T_2\mathrm{Cov}[\tilde{X}]T_2'$ and hence $\mathrm{Cov}[S] = \tilde{\Sigma}^{-\frac{1}{2}}\tilde{\Sigma}(\tilde{\Sigma}^{-\frac{1}{2}})' = Q\Lambda^{-\frac{1}{2}}Q'Q\Lambda Q'Q\Lambda^{-\frac{1}{2}}Q' = Q\Lambda^{-\frac{1}{2}}\Lambda\Lambda^{-\frac{1}{2}}Q' = QI_{n-1}Q' = I_{n-1}$.

Second step: Show that S given m is spherically distributed. Similar to the proof of Lemma 1, we can easily verify that, for $r = 0$, all linear transformations of random vectors satisfying condition (1) inherit the GLS property, that is, we have the representation $TX = (\tilde{\beta}_i Y + \tilde{\gamma}_i\tilde{Z}_i)_{i=1,\dots,n}$ with some suitable $\tilde{\beta}_i$, $\tilde{\gamma}_i$ and elliptically distributed $\tilde{Z} = (\tilde{Z}_1, \dots, \tilde{Z}_n)'$, conditional on Y. Because $\mathbb{E}[S] = (0, \dots, 0)'$, $\mathbb{E}[Y] \neq 0$, and $\mathbb{E}[\tilde{Z}_2, \dots, \tilde{Z}_n] = (0, \dots, 0)'$, we have $\tilde{\beta}_1 \neq 0$ and $\tilde{\beta}_2 = \dots = \tilde{\beta}_n = 0$, that is, $S = (\tilde{\gamma}_2\tilde{Z}_2, \dots, \tilde{\gamma}_n\tilde{Z}_n)'$. We have $\mathbb{E}[m] \neq 0$ as $\tilde{\beta}_1 \neq 0$ and $\mathbb{E}[Y] \neq 0$. Further, as $(\tilde{Z}_1, S')'$ is elliptically distributed (conditional on Y), the conditional distribution of S given $\tilde{Z}_1$ (and Y) is also elliptical, although generally with a different characteristic generator (see McNeil et al. 2005, section 3.3.3). For $\tilde{\gamma}_1 = 0$, we have $m = \tilde{\beta}_1 Y$ such that $S|m$ is elliptically distributed. In the case of $\tilde{\gamma}_1 \neq 0$, $S|m$ is also elliptically distributed because S is elliptically distributed conditional on $\{Y = y, \tilde{Z}_1 = z\}$ for every $y, z \in \mathbb{R}$ and $\{m = k\} = \{\tilde{\beta}_1 Y + \tilde{\gamma}_1\tilde{Z}_1 = k\} = \cup_{l\in\mathbb{R}}\{Y = l/\tilde{\beta}_1, \tilde{Z}_1 = (k - l)/\tilde{\gamma}_1\}$. Finally, considering $\mathrm{Cov}[S] = I_{n-1}$, it follows that S given m is spherically distributed.

"⇒": Assume that there is a nonsingular matrix T such that $TX = (m, S')'$. If we denote $T^{-1} = (\bar{t}_{i,j})_{i,j=1,\dots,n}$, we have $X = T^{-1}(m,\ S')'$ and hence $X_i = \bar{t}_{i,1}m + \sum_{j=2}^n \bar{t}_{i,j}S_j$ for $i = 1, \dots, n$. Setting $\beta_i := \bar{t}_{i,1}$, $Y := m$, and $\tilde{Z}_i := \sum_{j=2}^n \bar{t}_{i,j}S_j$, we have the representation $X_i = \beta_i Y + \tilde{Z}_i$. As elliptical distributions are obtained by linear transformations of spherical distributions (see McNeil et al. 2005, section 3.3.2), $\tilde{Z} = (\tilde{Z}_1, \dots, \tilde{Z}_n)$ is elliptically distributed, conditional on Y. □

Proof of Theorem 2. As we have already proven the sufficient condition in Proposition 1, we now turn to the necessary condition. Assume that $\mathbb{E}[X_i] \neq r$ for at least one $i = 1, \dots, n$ and the distribution of $P = w_0 r + w'X$ is determined by its mean and variance for every $(w_0, w) \in \mathbb{R}^{n+1}$ with $w_0 + w'e = 1$. Note that $\mathbb{E}[P - r] = \mathbb{E}[P] - r$, $\mathrm{Var}[P - r] = \mathrm{Var}[P]$, and $\{(w_0, w) : (w_0, w) \in \mathbb{R}^{n+1} \text{ and } w_0 + w'e = 1\} = \{(1 - w'e, w) : w \in \mathbb{R}^n\}$. Hence, in light of the fact that, for the excess returns $\bar{P}$ and $\bar{X}$, we have $\bar{P} = w'\bar{X}$ under the full investment constraint, our assumption states that the distribution of $w'\bar{X}$ is determined by its mean and variance for every $w \in \mathbb{R}^n$. Note that $\mathbb{E}[\bar{X}_i] \neq 0$ for at least one $i = 1, \dots, n$. Following from Chamberlain (1983, theorem 2), $\bar{X}$ must satisfy (2). By Theorem 1 the family of distributions (1) with no constant and (2) are equivalent. □

Proof of Corollary 1. "⇐": Assume the $X_1, \dots, X_n$ in X satisfy (4) and, furthermore, w_0 and the $w_1, \dots, w_n$ in w are real numbers with $w_0 + w'e = 1$. Then, $P = w_0 X_0 + w'X = (w_0 + w'e)\mathcal{R} + w'\beta Y + w'\xi$ with $\beta := (\beta_1, \dots, \beta_n)'$ and $\xi := (\gamma_1 Z_1, \dots, \gamma_n Z_n)'$. Note that $(w_0 + w'e)\mathcal{R} = \mathcal{R}$, $w'\beta \in \mathbb{R}$, and $w'\xi$ is elliptically distributed about the origin (with the same arguments as before). Thus, in analogy to the proof of Proposition 1, we have mean-variance determination.

"⇒": To prove the necessary condition, we can proceed similar to the proof of Theorem 2. Assume that $P = w_0 X_0 + w'X$ is determined by its mean and variance for every $(w_0, w) \in \mathbb{R}^{n+1}$ with $w_0 + w'e = 1$. Let $\tilde{X}_i := X_i - \mathcal{R}$ and note that $\tilde{P} := P - \mathcal{R} = w'X + (w_0 - 1)\mathcal{R} = (w_0 - 1)\mathcal{R} + w'\tilde{X} + w'e\mathcal{R} = (w'e + w_0 - 1)\mathcal{R} + w'\tilde{X} = w'\tilde{X}$ similar to before. Further, we have $\mathbb{E}[\tilde{P}] = \mathbb{E}[P - \mathcal{R}] = \mathbb{E}[P] - \mathbb{E}[\mathcal{R}]$ and $\mathrm{Var}[\tilde{P}] = \mathrm{Var}[w'\tilde{X}] = \mathrm{Var}[w'(X - e\mathcal{R})] + \mathrm{Var}[\mathcal{R}] - \mathrm{Var}[\mathcal{R}] = \mathrm{Var}[w'(X - e\mathcal{R}) + \mathcal{R}] - \mathrm{Var}[\mathcal{R}] = \mathrm{Var}[w'(X - e\mathcal{R}) + (w_0 + w'e)\mathcal{R}] - \mathrm{Var}[\mathcal{R}] = \mathrm{Var}[w_0\mathcal{R} + w'X] - \mathrm{Var}[\mathcal{R}] = \mathrm{Var}[P] - \mathrm{Var}[\mathcal{R}]$. Here we used that $X_i - \mathcal{R}$ and $\mathcal{R}$ are independent for each $i = 1, \dots, n$, such that $w'(X - e\mathcal{R})$ is independent of $\mathcal{R}$. Hence, our assumption states that the distribution of $w'\tilde{X}$ is determined by its mean and variance for every $w \in \mathbb{R}^n$. Note that $\mathbb{E}[\tilde{X}_i] \neq 0$ for at least one $i = 1, \dots, n$, as $\mathbb{E}[X_i] \neq \mathbb{E}[\mathcal{R}]$ for at least one $i = 1, \dots, n$. Following from Chamberlain (1983, theorem 2), $\tilde{X}$ must satisfy (2). In analogy to our Theorem 1, X must follow (4). □

Endnotes

[1] In Section 2, we discuss in more detail why this difference is important and why ignoring it has led to crucial misunderstandings in the literature.

[2] Grootveld and Hallerbach (1999), Kapsos et al. (2014), and Lwin et al. (2017) are examples motivating their otherwise high-quality work in such a way. Studies generally discarding mean-variance analysis in the light of non-normality are, for example, Zakamouline and Koekebakker (2009), Homm and Pigorsch (2012), and Levy and Kaplanski (2015).

[3] In the following, all random variables are assumed to have finite variances.

[4] In other words, when critics argue based on returns, they are assuming that the full investment condition holds. However, in this case, they cannot use Chamberlain (1983) to back up their arguments.

[5] Furthermore, although Y and Z_i might be stochastically dependent, we still have $\mathrm{Cov}[Z_i, Y] = 0$ because $\mathbb{E}[Z_i|Y] = 0$ (see Hunter 1972). In other words, Y and Z_i might affect each other, but there is no linear relationship between them.

[6] In our definition, we could alternatively assume $Z = (Z_1, \ldots, Z_n)'$ to be elliptically distributed about the origin and appropriately scale the coefficients γ_i to obtain the same conclusions in our following discussions. In subsequent proofs, we often use this perspective (with "dropped" scaling argument) when verifying that a random vector follows a skew-elliptical GLS distribution.

[7] Showing the equivalence of the distributional representations is not trivial. An elliptically distributed random vector $Z \in \mathbb{R}^n$ is, in general, generated by an affine transformation of a spherically distributed random vector of the same dimension n. The vector S, however, is of dimension $n-1$. We can still choose the arbitrary random variable m appropriately such that it covers the missing last degree of freedom and the random variable Y.

[8] For $w_0 + w_1 + \ldots + w_n = 1$, we have $P - r = (w_0 r + w_1 X_1 + .. + w_n X_n) - r = (w_0 r + w_1 X_1 + \ldots + w_n X_n) - (w_0 + w_1 + \ldots + w_n) r = w_1(X_1 - r) + \ldots + w_n(X_n - r)$.

[9] It is not important to distinguish whether portfolio returns or portfolio excess returns are determined by mean and variance because they contain the same information: $\mathbb{E}[\bar{P}] = \mathbb{E}[P - r] = \mathbb{E}[P] - r$ and $\mathrm{Var}[\bar{P}] = \mathrm{Var}[P - r] = \mathrm{Var}[P]$.

[10] Note that, because the following Corollary 1 requires $\mathcal{R}$ to satisfy specific conditions, the necessary condition for the very general case of n risky assets and no risk-free asset remains open.

[11] In this context, also see our discussion in Section C of our online appendix.

[12] Following Eling and Schuhmacher (2007), we have used several alternative ways of modeling the risk-free rate (e.g., initial fixes or diverse averages). None of them affected our main conclusion.

[13] See https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[14] For full transparency, the results for other K and S are available from the authors upon personal request. Also note that individual investors can, of course, also form portfolios based on characteristics other than betas (see Green et al. 2017). However, considering them all is beyond the scope of our study.

[15] For alternative testing procedures (like the Cramer-von Mises test), which can also be used in the context of the original GLS model, see Vassalos et al. (2012).

[16] Also note that, by relying on the CAPM framework, we assume that Y and Z are independent, which is not required in our GLS world. Besides simplifying the testing procedure (because conditional and unconditional distributions become identical), this introduces an additional hurdle for empirical data to pass our test.

[17] More recently proposed ellipticality tests have not found wide acceptance because their implementation is computationally too expensive for large numbers of dimensions and data points (see Paolella 2019, section C.2.4).

[18] For technical details on these transformations and the formal reasons for the reduction from K to $K-1$, see Liang et al. (2008, pp. 684–685).

[19] However, in line with earlier evidence (see McNeil et al. 2005, section 3.3.5), very small portfolios and even single stocks are not generally in conflict with spherical symmetry.

[20] In addition to vector sampling, we implemented random sampling similar to Paolella (2019, section C.2.4). That is, we randomly drew $S \cdot K$ of our NASDAQ stocks, formed equal-weighted portfolios, calculated our test statistics, and repeated this 1,000 times. With few exceptions, density plots of the statistics nicely matched the null distributions of our tests. Performing this procedure for industry subsets (given by the Worldscope general industry classification) delivered similar results.

[21] The GSCI data are available in Thomson Reuters Datastream. The CSHFI data can be obtained via https://lab.credit-suisse.com.

[22] One may also study Genton and Loperfido (2005) and Shushi (2016, 2018).

[23] Test extensions in the spirit of Bai (2003) and Li and Tkacz (2011) may also have their merits.

[24] Chamberlain (1983) does not prove this statement because arguments are drawn from basic linear algebra only. Nonetheless, we provide a proof for the sake of completeness.

References

Abhakorn P, Smith P, Wickens M (2013) What do the Fama-French factors add to C-CAPM? *J. Empirical Finance* 22:113–127.

Agarwal V, Naik N (2004) Risks and portfolio decisions involving hedge funds. *Rev. Financial Stud.* 17(1):63–98.

Ang A (2014) *Asset Management: A Systematic Approach to Factor Investing* (Oxford University Press, Oxford, UK).

Bacon C (2008) *Practical Portfolio Performance Measurement and Attribution*, 2nd ed. (John Wiley & Sons, Hoboken, NJ).

Bai J (2003) Testing parametric conditional distributions of dynamic models. *Rev. Econom. Statist.* 85(3):531–549.

Barasinska N, Schäfer D, Stephan A (2012) Individual risk attitudes and the composition of financial portfolios: Evidence from German household portfolios. *Quart. Rev. Econom. Finance* 52(1):1–14.

Baron D (1977) On the utility theoretic foundations of mean-variance analysis. *J. Finance* 32(5):1683–1697.

Bernard C, Vanduffel S (2014) Mean-variance optimal portfolios in the presence of a benchmark with applications to fraud detection. *Eur. J. Oper. Res.* 234(2):469–480.

Black F (1972) Capital market equilibrium with restricted borrowing. *J. Bus.* 45(3):444–455.

Board J, Sutcliffe C (1994) Estimation methods in portfolio selection and the effectiveness of short sales restrictions: UK evidence. *Management Sci.* 40(4):516–534.

Brockett P, Golden L (1987) A class of utility functions containing all the common utility functions. *Management Sci.* 33(8):955–964.

Cass D, Stiglitz JE (1970) The structure of investor preferences and asset returns, and separability in portfolio allocation: A contribution to the pure theory of mutual funds. *J. Econom. Theory* 2(2):122–160.

Chamberlain G (1983) A characterization of the distributions that imply mean-variance utility functions. *J. Econom. Theory* 29(1):185–201.

DeMiguel V, Garlappi L, Uppal R (2009) Optimal vs. naive diversification: How inefficient is the 1/N portfolio strategy? *Rev. Financial Stud.* 22(5):1915–1953.

Dexter A, Yu J, Ziemba W (1980) Portfolio selection in a lognormal market when the investor has a power utility function: Computational results. Dempster M, ed. *Stochastic Programming* (Academic Press, London), 507–523.

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

Elton E, Gruber M, Brown S, Goetzmann W (2007) *Modern Portfolio Theory and Investment Analysis*, 7th ed. (John Wiley & Sons, Hoboken, NJ).

Fama E, French K (2012) Size, value, and momentum in international stock returns. *J. Financial Econom.* 105(3):457–472.

Gao X, Nardari F (2018) Do commodities add economic value in asset allocation? New evidence from time-varying moments. *J. Financial Quant. Anal.* 53(1):365–393.

Gardner R, Pinder J, Wood R (1980) Monte Carlo estimation of percentiles for the multi-Smirnov test. *J. Statist. Comput. Simulation* 10(3-4):243–249.

Genton M, Loperfido N (2005) Generalized skew-elliptical distributions and their quadratic forms. *Ann. Inst. Statist. Math.* 57(2):389–401.

Green J, Hand J, Zhang X (2017) The characteristics that provide independent information about average U.S. monthly stock returns. *Rev. Financial Stud.* 30(12):4389–4436.

Grootveld H, Hallerbach W (1999) Variance vs downside risk: Is there really that much difference? *Eur. J. Oper. Res.* 114(2):304–319.

Grundy B, Lim B, Verwijmeren P (2012) Do option markets undo restrictions on short sales? Evidence from the 2008 short-sale ban. *J. Financial Econom.* 106(2):331–348.

Hickernell F (1998) A generalized discrepancy and quadrature error bound. *Math. Comput.* 67(221):299–322.

Hlawitschka W (1994) The empirical nature of Taylor-series approximations to expected utility. *Amer. Econom. Rev.* 84(3):713–719.

Homm U, Pigorsch C (2012) Beyond the Sharpe ratio: An application of the Aumann-Serrano index to performance measurement. *J. Banking Finance* 36(8):2274–2284.

Huffer F, Park C (2007) A test for elliptical symmetry. *J. Multivariate Anal.* 98(2):256–281.

Hunter J (1972) Independence, conditional expectation, and zero covariance. *Amer. Statist.* 26(5):22–24.

Ingersoll JE (1987) *Theory of Financial Decision Making*, vol. 3 (Rowman & Littlefield, Lanham, MD).

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

Kapsos M, Christofides N, Rustem B (2014) Worst-case robust Omega ratio. *Eur. J. Oper. Res.* 234(2):499–507.

Kolm PN, Tütüncü R, Fabozzi FJ (2014) 60 years of portfolio optimization: Practical challenges and current trends. *Eur. J. Oper. Res.* 234(2):356–371.

Korkie B, Turtle H (2002) A mean-variance analysis of self-financing portfolios. *Management Sci.* 48(3):427–443.

Kroll Y, Levy H, Markowitz HM (1984) Mean-variance vs. direct utility maximization. *J. Finance* 39(1):47–61.

Kumar A, Lim S (2008) How do decision frames influence the stock investment choices of individual investors? *Management Sci.* 54(6):1052–1064.

Levy H, Levy M (2004) Prospect theory and mean-variance analysis. *Rev. Financial Stud.* 17(4):1015–1041.

Levy M, Kaplanski G (2015) Portfolio selection in a two-regime world. *Eur. J. Oper. Res.* 242(2):514–524.

Levy H, Markowitz HM (1979) Approximating expected utility by a function of mean and variance. *Amer. Econom. Rev.* 69(3):308–317.

Li F, Tkacz G (2011) A consistent test for multivariate conditional distributions. *Econometric Rev.* 30(3):251–273.

Liang J, Fang K, Hickernell F (2008) Some necessary uniform tests for spherical symmetry. *Ann. Inst. Statist. Math.* 60(3):679–696.

Liang J, Ng K, Tian G (2019) A class of uniform tests for goodness-of-fit of the multivariate l_p-norm spherical distributions and the l_p-norm symmetric distributions. *Ann. Inst. Statist. Math.* 71(1):137–162.

Liang J, Fang K, Hickernell FJ, Li R (2001) Testing multivariate uniformity and its applications. *Math. Comput.* 70(233):337–355.

Lintner J (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Rev. Econom. Statist.* 47(1):13–37.

Lwin KT, Qu R, MacCarthy BL (2017) Mean-VaR portfolio optimization: A nonparametric approach. *Eur. J. Oper. Res.* 260(2):751–766.

Mandelbrot B (1963) The variation of certain speculative prices. *J. Bus.* 36(4):394–419.

Manzotti A, Pérez F, Quiroz A (2002) A statistic for testing the null hypothesis of elliptical symmetry. *J. Multivariate Anal.* 81(2): 274–285.

Markowitz H (1952) Portfolio selection. *J. Finance* 7(1):77–91.

Markowitz H (1959) *Portfolio Selection: Efficient Diversification of Investments* (John Wiley & Sons, New York).

Markowitz H (2000) *Mean-Variance Analysis in Portfolio Choice and Capital Markets* (Fabozzi Associates, New Hope, PA).

Markowitz H (2014) Mean-variance approximations to expected utility. *Eur. J. Oper. Res.* 234(2):346–355.

McNeil AJ, Frey R, Embrechts P (2005) *Quantitative Risk Management: Concepts, Techniques and Tools* (Princeton University Press, Princeton, NJ).

Meyer J (1987) Two moment decision models and expected utility maximization. *Amer. Econom. Rev.* 77:421–430.

Meyer J, Rasche RH (1992) Sufficient conditions for expected utility to imply mean-standard deviation rankings: Empirical evidence concerning the location and scale condition. *Econom. J. (London)* 102(410):91–106.

Miller F Jr, Quesenberry C (1979) Power studies of tests for uniformity, ii. *Comm. Statist. Simulation Comput.* 8(3):271–290.

Mossin J (1966) Equilibrium in a capital asset market. *Econometrica* 34(4):768–783.

Neyman J (1937) Smooth test for goodness of fit. *Scandinavian Actuarial J.* (3-4):149–199.

Owen J, Rabinovitch R (1983) On the class of elliptical distributions and their applications to the theory of portfolio choice. *J. Finance* 38(3):745–752.

Paolella M (2019) *Linear Models and Time-Series Analysis: Regression, ANOVA, ARMA and GARCH* (John Wiley & Sons, Hoboken, NJ).

Peiró A (1999) Skewness in financial returns. *J. Banking Finance* 23(6):847–862.

Pratt J (1964) Risk aversion in the small and in the large. *Econometrica* 32(1/2):122–136.

Pulley LB (1983) Mean-variance approximations to expected logarithmic utility. *Oper. Res.* 31(4):685–696.

Quesenberry C, Miller F Jr (1977) Power studies of some tests for uniformity. *J. Statist. Comput. Simulation* 5(3):169–191.

Ray P, Jenamani M (2016) Mean-variance analysis of sourcing decision under disruption risk. *Eur. J. Oper. Res.* 250(2):679–689.

Schott J (2002) Testing for elliptical symmetry in covariance-matrix-based analyses. *Statist. Probab. Lett.* 60(4):395–404.

Schuhmacher F (2012) The Sharpe ratio is better than you may think. *J. Bus. Econom.* 82(6):685–705.

Schuhmacher F, Auer BR (2014) Sufficient conditions under which SSD- and MR-efficient sets are identical. *Eur. J. Oper. Res.* 239(3):756–763.

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

Shushi T (2016) A proof for the conjecture of characteristic function of the generalized skew-elliptical distributions. *Statist. Probab. Lett.* 119:301–304.

Shushi T (2018) Generalized skew-elliptical distributions are closed under affine transformations. *Statist. Probab. Lett.* 134:1–4.

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

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

Simaan Y (2014) The opportunity cost of mean-variance choice under estimation risk. *Eur. J. Oper. Res.* 234(2):382–391.

Sinn H-W (1983) *Economic Decisions Under Uncertainty* (North-Holland Publishing Company, New York).

Statman M (1987) How many stocks make a diversified portfolio? *J. Financial Quant. Anal.* 22(3):353–363.

Stephens M (1970) Use of the Kolmogorov Smirnov, Cramér-von Mises and related statistics without extensive tables. *J. Roy. Statist. Soc. B.* 32(1):115–122.

Strang G (2009) *Introduction to Linear Algebra*, 4th ed. (Wellesley-Cambridge Press, Wellesley, MA).

Tobin J (1958) Liquidity preference as behavior toward risk. *Rev. Econom. Stud.* 25(2):65–86.

Vassalos M, Dillon C, Childs P (2012) Empirically testing for the location-scale condition: A review of the economic literature. *J. Risk Model Validation* 6(3):51–66.

Watson G (1961) Goodness-of-fit tests on a circle. *Biometrika* 49(1-2):109–114.

Watson G (1962) Goodness-of-fit tests on a circle. ii. *Biometrika* 49(1-2):57–63.

Zakamouline V, Koekebakker S (2009) Portfolio performance evaluation with generalized Sharpe ratios: Beyond the mean and variance. *J. Banking Finance* 33(7):1242–1254.