

Assessing skewness in financial markets

Giovanni Campisi¹  | Luca La Rocca²  | Silvia Muzzioli¹ 

¹Marco Biagi Department of Economics, University of Modena and Reggio Emilia, Modena, Italy

²Department of Physics, Informatics and Mathematics, University of Modena and Reggio Emilia, Modena, Italy

Correspondence

Giovanni Campisi, Marco Biagi
Department of Economics, University of
Modena and Reggio Emilia, Modena
41121, Italy.
Email: giovannicampisi6@gmail.com

Funding information

Università degli Studi di Modena e Reggio
Emilia, Grant/Award Numbers: FAR2017,
FAR2019

It is a matter of common observation that investors value substantial gains but are averse to heavy losses. Obvious as it may sound, this translates into an interesting preference for right-skewed return distributions, whose right tails are heavier than their left tails. Skewness is thus not only a way to describe the shape of a distribution, but also a tool for risk measurement. We review the statistical literature on skewness and provide a comprehensive framework for its assessment. Then, we present a new measure of skewness, based on the decomposition of variance in its upward and downward components. We argue that this measure fills a gap in the literature and show in a simulation study that it strikes a good balance between robustness and sensitivity.

KEYWORDS

convex ordering, distribution asymmetry, risk asymmetry, unimodality, volatility

1 | INTRODUCTION

The financial literature offers a wealth of papers dealing with the construction and implementation of risk measures to allow investors to make informed trading decisions. In light of the recent financial crises, these measures have become increasingly important in order to prevent tail risk events.

One of the most important risk measures is the VIX index adopted by the Chicago Board Options Exchange (CBOE). This is a forward-looking measure of volatility that investors expect in the coming month (Whaley, 2009). Volatility indexes are deemed by market operators to capture *market fear*: high index values are associated with high levels of uncertainty in the underlying market, whereas low index values with stable conditions; see Whaley (2000) and

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Statistica Neerlandica* published by John Wiley & Sons Ltd on behalf of Netherlands Society for Statistics and Operations Research.

Muzzioli (2013b). Traditionally, financial returns are assumed to follow a normal distribution. In this connection, volatility is a good measure of risk based on the idea that investors are averse to uncertainty. However, a significant number of studies highlight the fact that financial returns are nonnormally distributed; see for example, Fama (1965), Peiró (1999), Chen, So, and Gerlach (2005), Lempérière et al. (2017), and Elyasiani, Gambarelli, and Muzzioli (2018). Specifically, financial returns are found to present an empirical distribution with heavy tails and a negative skew. In other words, the occurrence of extreme and negative events is more probable than in the normal distribution. This has consequences in terms of the need to include higher-order moments as indicators of market risk.

This may be illustrated by the CBOE SKEW index. The CBOE SKEW index has been listed on the CBOE since February 2011 to measure the tail risk not fully captured by the VIX index. While VIX measures the overall risk in the 30-day S&P500 log-returns without disentangling the probabilities attached to positive and negative returns, the skewness index (CBOE SKEW) is intended to measure the perceived tail risk, that is, the probability that investors attach to extreme negative returns. The CBOE SKEW index relies on Pearson's (third order) moment coefficient of skewness. It is well known that Pearson's moment coefficient of skewness is not a robust measure of skewness. In the statistical literature there are several cases in which Pearson's moment coefficient of skewness leads to controversial conclusions. From a financial point of view, this could have serious problems. Indeed, the role of the CBOE SKEW index as an indicator of market fear has been questioned since it frequently moves in the same direction as returns (Elyasiani, Gambarelli, & Muzzioli, 2020).

There is an extensive statistical literature on skewness, which we review in this paper in order to outline a comprehensive framework for its assessment. We find that only the classical measures of skewness rely on higher-order moments, while a number of more recent measures do not. We also propose a new measure of skewness: the risk asymmetry index (RAX) based on the same measure introduced by Elyasiani et al. (2018). The authors derive the RAX index by estimating upside, downside and total volatility of returns using option prices in a model-free setting (implied volatility). On the other hand, in this paper we derive the RAX index by computing upside, downside, and total volatility of returns from the physical distribution. By proposing the RAX index and reviewing existing measures of skewness, we seek to contribute to the investigation of skewness measures from a statistical perspective, arguing that the notion of skewness can play an important role in risk measurement.

The choice of the RAX index is based on its ability to detect not only risk, i.e. the volatility of returns, but also asymmetry, that is, the different volatility of positive and negative returns, as empirically found in Elyasiani et al. (2018). In particular, we highlight two main points. First, it has been found that the RAX index subsumes all the information of the skewness index (ITSKEW) and volatility index (ITVIX) for the Italian market. Indeed, when these three indexes are included in the same model, both the ITSKEW and the ITVIX fail to have a significant explanatory power for future returns. In this connection, the contribution of the RAX is useful for investors who can exploit its information in order to make profitable trades. Moreover, Elyasiani et al. (2018) found that the RAX is able to indicate future fear or greed since extremely high or low level of the index are related to positive or negative future returns. While the volatility index ITVIX provides useful information on future returns only in the high volatility period, the RAX index provides useful information about future returns during the entire sample period. The RAX index gives a clear and unambiguous signal to investors as extremely low (high) values of the risk-asymmetry index signal a buy (sell) opportunity.

The paper proceeds as follows. In Section 2 we outline our framework for the study of univariate skewness, with a review of the tools available for both qualitative and quantitative assessments, identifying the statistical properties that a valid measure of skewness should satisfy. In Section 3 we propose the new measure of skewness, focussing on its statistical properties and interpretation. Section 4 contains an analysis of the robustness to outliers and sensitivity to changes in the shape of the distribution of the RAX and the most relevant measures of skewness. Section 5 concludes the paper with a brief discussion.

2 | SYMMETRY AND ASYMMETRY

Skewness is defined as a relaxation of symmetry to allow for asymmetry in a specific direction. We therefore start our investigation by introducing the notion of symmetry, which is uncontroversial, at least in the univariate case. A univariate random variable X is *symmetric* with respect to the real value m when $X - m$ and $m - X$ have the same distribution; see for instance Doksum (1975). This property of the distribution of X can be written as $\mathbb{P}\{X \leq m - t\} = \mathbb{P}\{X \geq m + t\}$ for all $t > 0$, which amounts to saying that all corresponding left and right tails of X with respect to m have the same weight. Letting $t \rightarrow 0$, it is straightforward to see that m has to be the *median* of X , uniquely defined as the midpoint of the interval formed by all v such that $\mathbb{P}\{X \leq v\} \geq 1/2$ and $\mathbb{P}\{X \geq v\} \geq 1/2$. It is also evident that symmetry is invariant with respect to affine transformations: if X is symmetric and $Y = \alpha + \beta X$, then Y is also symmetric. Further considerations will be eased by the introduction of a suitable distributional framework.

Let F be the distribution function of X . We define the *support interval* of F as the open interval $]a, b[$ with left endpoint $a = \inf\{x \in \mathbb{R} | F(x) > 0\}$ and right endpoint $b = \sup\{x \in \mathbb{R} | F(x) < 1\}$. Note that it is possible to have $a = -\infty$ or $b = +\infty$. We assume that F is continuous on the real line and strictly increasing on its support interval. More specifically, we assume that F is obtained from a probability density function f that is continuous on $]a, b[$ and such that $f(x) > 0$ for all $x \in]a, b[$, while $f(x) = 0$ if $x \leq a$ or $x \geq b$. Let \mathcal{F}_0 be the class of all such distributions. Piecewise continuous density functions could be allowed to enlarge \mathcal{F}_0 , but here we favor simplicity over generality. Nonetheless, we point out that such an enlargement would provide scope for probability density histograms (representing an important class of data based distributions). Interesting subclasses of \mathcal{F}_0 are obtained by assuming that X has finite moments up to some order; let $\mathcal{F}_k = \{F \in \mathcal{F}_0 | \mathbb{E}|X|^k < \infty\}$, $k = 1, 2, \dots$, be such classes. If $F \in \mathcal{F}_1$, we denote by $\mu = \mathbb{E}(X)$ the *mean* of X . If $F \in \mathcal{F}_2$, we denote by $\sigma^2 = \mathbb{E}(X - \mu)^2$ the *variance* of X (with σ denoting the *standard deviation* of X); note that $\sigma^2 > 0$ because F is continuous.

Assuming $F \in \mathcal{F}_0$, the symmetry condition can be written as

$$F(m - t) + F(m + t) = 1 \quad \text{for all } t > 0, \quad (1)$$

where $m = F^{-1}(1/2)$ with F^{-1} uniquely defined on the open interval $]0, 1[$ as the inverse function of F . Note that F^{-1} , called the *quantile function* of X , is also continuous and strictly increasing. Since f is the derivative of F , condition (1) can be rewritten as $f(m - t) = f(m + t)$ for all $t > 0$, so that $\mu = m$, for $F \in \mathcal{F}_1$, and μ can replace m in (1).

A special case of interest is given by unimodal distributions. Following Dharmadhikari and Joag-Dev (1988), p. 2, we say that X is *unimodal* at $x_\star \in]a, b[$ if F is convex on $]a, x_\star[$ and concave on $]x_\star, b[$, which corresponds to f increasing on $]a, x_\star[$ and decreasing on $]x_\star, b[$. We say that X is unimodal (*tout court*) if it is unimodal at some x_\star . In this case, the *mode* of X (denoted by M) can

be uniquely defined as the midpoint of the interval formed by all x_* such that X is unimodal at x_* . For all $k = 0, 1, 2, \dots$, we define $\mathcal{F}_k^* = \{F \in \mathcal{F}_k | X \text{ is unimodal}\}$. If X is symmetric, it follows from $f(m-t) = f(m+t)$ for all $t > 0$ that $M = m$, for $F \in \mathcal{F}_0^*$, and M can replace m in (1). Hence, for all symmetric X with $F \in \mathcal{F}_1^*$, we have $M = m = \mu$ and the three classical measures of central tendency coincide.

If X is not symmetric, we say that X is *asymmetric*. While all symmetric distributions are alike in symmetry, each asymmetric distribution is asymmetric in its own way. Skewness relaxes symmetry to allow for a specific type of asymmetry: a random variable is left-skewed when its left tails are heavier than its right tails, whereas it is right-skewed when its right tails are heavier than its left tails. A strictly skewed variable is a skewed variable that is not symmetric and, as such, it represents a way of being asymmetric. We formalize these notions in the following: Section 2.1 provides an assessment of when a random variable is manifestly skewed; Section 2.2 provides an assessment of how much of a skew a given random variable exhibits, even though such a skew may not be manifest.

2.1 | Qualitative assessment of skewness

We start with a simplifying remark: since the left tails of a random variable X are the right tails of the opposite random variable $-X$, and vice versa, it will be sufficient to assess when X is right-skewed; X will be left-skewed when $-X$ is right-skewed. We proceed with a discussion of when a random variable Y is more right-skewed than another random variable X , which is both an interesting problem in itself and one whose solution will enable us to consider X as right-skewed when it is more right-skewed than $-X$, as recommended by MacGillivray (1986).

Following van Zwet (1964), we compare X with Y by means of the function $R(x) = G^{-1}(F(x))$, $a < x < b$, where G is the distribution function of Y . We call $R = G^{-1} \circ F$ the *quantile-quantile function* of Y against X , because it is the function whose graph is represented in the Q-Q plot with X on the horizontal axis and Y on the vertical axis. We say that X is *less right-skewed* than Y , or Y is *more right-skewed* than X , and write $F \lesssim G$, or $G \gtrsim F$, if R is convex, which indicates that the left tails of X are progressively heavier than the left tails of Y and the right tails of Y are progressively heavier than the right tails of X . Note that R is strictly increasing (as well as continuous) on $]a, b[$. The quantile-quantile function of X against Y is given by $R^{-1}(y) = F^{-1}(G(y))$, $c < y < d$, where $]c, d[$ is the support interval of G . Since $R^{-1} = F^{-1} \circ G$ is convex if and only if R is concave, we have $F \gtrsim G$ if and only if R is concave.

An interesting characterization of the above described comparison is that $F \lesssim G$ if and only if Y is equal in distribution to a strictly increasing convex transformation of X : on the one hand, the variable $R(X)$ has the same distribution as Y ; on the other hand, if $Y = \phi(X)$ with ϕ strictly increasing and convex, then $R = \phi$. If both $F \lesssim G$ and $F \gtrsim G$, we write $F \sim G$. Clearly, this is the case when R is a positive affine function, that is, $Y = \alpha + \beta X$ with $\beta > 0$. This makes relative skewness a property of location-scale models rather than individual distributions. The distribution function of $-X$ is $\bar{F}(x) = 1 - F(-x)$, $x \in \mathbb{R}$. Since $F \lesssim G$ if and only if $\bar{F} \gtrsim \bar{G}$, we find that X is less right-skewed than Y if and only if $-X$ is more right-skewed than $-Y$; this means that we can safely interpret $F \lesssim G$ as X being *more left-skewed* than Y , or Y being *less left-skewed* than X .

It is a simple matter to check that \lesssim is *reflexive* and *transitive*: $F \lesssim F$ and $F \lesssim G, G \lesssim H$ implies $F \lesssim H$, where H denotes the distribution function of a third variable Z . This means that the relationship \lesssim is a *preorder* on \mathcal{F}_0 , which justifies its common name of *convex ordering* of distributions and qualifies \sim as the equivalence relationship defined by \lesssim . Several other orderings

of distributions have been proposed in the skewness literature (Arnold & Groeneveld, 1993; MacGillivray, 1986; Oja, 1981) and it turns out that the convex ordering is the strongest one. This is because it only considers the convexity of the quantile-quantile function, without reference to any measure of central tendency, and thus it only signals the most manifest cases of relative skewness.

As anticipated, we say that X is *right-skewed* when $-X \lesssim X$ (and *left-skewed* when $X \lesssim -X$); note that here, for the sake of expressiveness, we apply \lesssim to $-X$ and X rather than to \bar{F} and F . Since the convex ordering of distributions actually compares location-scale models, we can compare $m - X$ with $X - m$ instead of $-X$ with X . This leads us to focus on the quantile-quantile function \bar{R} of $X - m$ against $m - X$ and its slope \bar{r} , which are given by

$$\bar{R}(t) = F^{-1}(\bar{F}(t - m)) - m \quad \text{and} \quad \bar{r}(t) = \frac{f(m - t)}{f(F^{-1}(\bar{F}(t - m)))}, \quad -(b - m) < t < m - a. \quad (2)$$

Since $\bar{R}(0) = 0$ and $\bar{r}(0) = 1$, it follows from the convexity of \bar{R} that $x \leq \bar{R}(x)$ for all x in (2), and therefore

$$F(m - t) + F(m + t) \leq 1 \quad \text{for all } t > 0, \quad (3)$$

which is known as *van Zwet's condition* (Abadir, 2005) after van Zwet (1979) introduced it to prove the celebrated mode-median-mean inequality (usually considered a sign of right skewness).

When (3) holds, the right tails of X with respect to m are uniformly heavier than the corresponding left tails. This notion of *skew to the right* (Doksum, 1975) is easy to interpret, but it is relative to a specific measure of central tendency. If the mean or mode of X are also available, different tails can be compared and, in general, they will lead to different assessments of skewness: see Sato (1997) for an illustration of this phenomenon.

We conclude our discussion of qualitative skewness by noting that a random variable X is symmetric if and only if it is both right-skewed and left-skewed ($-X \sim X$). A notion of *strict skewness* can then be obtained by excluding symmetric distributions ($-X < X$ or $X < -X$). In this way, we are able to partition \mathcal{F}_0 into four groups of distributions: symmetric distributions, strictly left-skewed distributions, strictly right-skewed distributions, and other asymmetric distributions. As demonstrated in the next subsection, the distribution in the last group can be adjudicated as cases of negative (left) or positive (right) skewness through the choice of a suitable measure of skewness.

2.2 | Quantitative assessment of skewness

Here we consider the more ambitious problem of measuring the extent to which a given distribution is skewed. This means associating to every distribution of interest a real number, whose sign captures the direction of skewness and whose absolute value is greater when skewness is more pronounced. Formally, within our distributional framework, for distributions with finite moments up to order $k \in \{0, 1, 2, \dots\}$, we aim to specify a functional $\text{Sk} : \mathcal{F}_k \rightarrow \mathbb{R}$, which will be called a *measure of skewness* of order k and will be required to satisfy the following properties:

- (P1) $\text{Sk}(\bar{F}) = -\text{Sk}(F)$ for all $F \in \mathcal{F}_k$;
- (P2) $\text{Sk}(F) \leq \text{Sk}(G)$ whenever $F \lesssim G$.

If we are only interested in unimodal distributions, we can replace \mathcal{F}_k with \mathcal{F}_k^* and specify a *unimodal measure of skewness* of the same order. We will say that Sk is *valid* to stress that (P1) and (P2) hold.

The meaning of (P1) and (P2) is that we intend our quantitative assessment to respect our qualitative assessment. Indeed, it follows from (P1) and (P2) that $\text{Sk}(F) \geq 0$ if F is right-skewed, while $\text{Sk}(F) \leq 0$ if F is left-skewed, so that $\text{Sk}(F) = 0$ if F is symmetric. Furthermore, it follows from (P2) that $\text{Sk}(F) = \text{Sk}(G)$ if $F \sim G$, that is, we require any valid measure of skewness to be location-scale invariant. Hence, to all intents and purposes, we are making the same assumptions as Groeneveld and Meeden (1984), rooted in the foundational work of Oja (1981) and consistently used in later work (Arnold & Groeneveld, 1995; Groeneveld, 1991a, 1991b; Groeneveld & Meeden, 2009; Tajuddin, 1996, 1999). If we define *positive skewness* as $\text{Sk}(F) \geq 0$ and *negative skewness* as $\text{Sk}(F) \leq 0$, we are able to partition \mathcal{F}_k (or \mathcal{F}_k^*) in three groups of distributions: distributions with strictly positive skew, distributions with strictly negative skew, and distributions with null skew. Properties (P1) and (P2) ensure that such a partition does not contradict the partition obtained in the previous subsection in terms of left and right skewness. We will call two measures of skewness *equivalent* when they give rise to the same partition.

In the following, we review the literature on measuring skewness with (P1) and (P2) in mind. We also take into account the values taken by Sk . Let $\underline{s} = \inf_F \text{Sk}(F)$ and $\bar{s} = \sup_F \text{Sk}(F)$ be its extrema. Note that $\underline{s} = -\bar{s}$ by (P1). If Sk is a valid measure of skewness and ϕ is an odd and (strictly) increasing real function defined on the image of Sk , then $\phi \circ \text{Sk}$ is another (equivalent) valid measure of skewness. This means that any measure of skewness can be transformed so as to have -1 and $+1$ as extrema, if \bar{s} is known. However, this is not necessarily the case and, even if it is, there is an interest in understanding whether the extrema can be attained or not and which distributions come close to them. Finally, we pay attention to the estimation of $\text{Sk}(F)$ when a random sample X_1, \dots, X_n from F is available (beyond the general but noisy strategy of computing Sk on a density estimate belonging to its domain of definition).

2.2.1 | Higher-order measures

The study of skewness was pioneered by Pearson (1895, 1901, 1916). In fact, the most classical measure of skewness goes under the name of *Pearson's moment coefficient of skewness*:

$$\gamma_j = \frac{\mathbb{E}(X - \mu)^{2j+1}}{\sigma^{2j+1}}, \quad (4)$$

where typically $j = 1$, but possibly $j = 2, 3, \dots$; since γ_j is well-defined when the distribution of X has finite moments up to order $2j + 1$, the choice $j = 1$ is the least demanding in terms of distributional assumptions. The rationale behind (4) is to use higher-order moments to gauge the extent to which the right tails of X are heavier than its left tails. This strategy obtains a valid $(2j + 1)$ th order measure of skewness: it is apparent that (4) satisfies (P1) and it was shown by van Zwet (1964) that (4) satisfies (P2); see also Oja (1981) and MacGillivray (1986).

The third-order measure of skewness γ_1 (standardized third central moment of X) is also called the *Fisher–Pearson coefficient of skewness* (Doane & Seward, 2011); see Arnold and Groeneveld (1995) for its historical attribution and Holgersson (2010) for the link to Fisher (1929). As an aside, note that Karl Pearson was not interested in the sign of skewness and used $\beta_1 = \gamma_1^2$ in place of γ_1 . If X follows a Pareto distribution with unit scale and large enough shape, that is,

$f(x) = \theta/x^{\theta+1}$ for $x > 1$ and $f(x) = 0$ for $x \leq 1$, with $\theta > 3$, then γ_1 is arbitrarily large for θ close enough to 3 and we conclude that $\bar{s} = +\infty$ (no finite upper bound on γ_1); see Groeneveld and Meeden (1984) for details.

A natural estimator of γ_1 is the sample moment coefficient of skewness $\hat{\gamma}_1$, defined as the ratio of the sample centered third moment $\langle (X - \hat{\mu})^3 \rangle = n^{-1} \sum_{i=1}^n (X_i - \hat{\mu})^3$, where $\hat{\mu} = \langle X \rangle = n^{-1} \sum_{i=1}^n X_i$ is the sample mean, to $\hat{\sigma}^3$, where $\hat{\sigma} = \langle (X - \hat{\mu})^2 \rangle^{1/2}$ is the sample sd. The value $\hat{\gamma}_1$ obtained in this way (sample moment coefficient of skewness) can be adjusted for sample size, but we are not interested in such an adjustment here; see Doane and Seward (2011) for information and references on this topic. Egon Sharpe Pearson, together with H. O. Hartley, provided tables to use $\hat{\gamma}_1$ as a test for departure from normality (Doane & Seward, 2011); see also Holgersson (2010) on testing asymmetry. Finally, the sharp algebraic bound $|\hat{\gamma}_1| \leq (n - 2)/(n - 1)^{1/2}$ holds for all samples of size n (Cox, 2010; Kirby, 1974; Wilkins, 1944) even though we have seen that γ_1 can take arbitrarily large values.

As illustrated by Li and Morris (1991), in some cases γ_1 may not express asymmetry well. Furthermore, as it is based on the third-order moment, γ_1 is strongly influenced by outliers; see for instance Groeneveld (1991a). This lack of robustness, together with an appetite for broadening the domain of definition, motivates the investigation of alternative measures of skewness.

2.2.2 | Unimodal measures

A second measure of skewness that dates back to the pioneering work of Pearson (1895) is called *Pearson's mode coefficient of skewness* or *Pearson's first coefficient of skewness*:

$$S'_K = \frac{\mu - M}{\sigma}, \tag{5}$$

well-defined for $F \in \mathcal{F}_2^*$. For instance, if $f(x) = x^{\alpha-1}e^{-x}/\Gamma(\alpha)$ for $x > 0$ and $f(x) = 0$ for $x \leq 0$, with $\alpha > 1$, we find $\mu = \alpha$, $M = \alpha - 1$ and $\sigma = \sqrt{\alpha}$, so that $S'_K = 1/\sqrt{\alpha}$. Remarkably, in this case (gamma distribution with unit scale), the equality $S'_K = \gamma_1/2$ holds (Arnold & Groeneveld, 1995). The quantity $\alpha_1 = \gamma_1/2$ is called the *coefficient of momental skewness* Zwillinger and Kokoska (1999, p. 18) and is clearly equivalent to γ_1 . In general, of course, α_1 and γ_1 are not equivalent to S'_K . The rationale behind (5) is that, as discussed in Section 2.1, the mode-median-mean inequality is a sign of right skewness. If $-X \leq X$, then $S'_K > 0$ and (5) gauges the width of inequality. It is clear that S'_K satisfies (P1). However, as illustrated by Arnold and Groeneveld (1995), the coefficient S'_K does not satisfy (P2): compatibility with right skewness does not extend to full compatibility with the convex ordering of distributions. Hence, we cannot regard S'_K as a valid measure of skewness.

Arnold and Groeneveld (1995) proposed replacing (5) with

$$\gamma_{AG} = \mathbb{P}\{X \geq M\} - \mathbb{P}\{X \leq M\} = 1 - 2F(M), \tag{6}$$

which we call the *Arnold–Groeneveld coefficient of skewness*. The coefficient γ_{AG} is well-defined for all $F \in \mathcal{F}_0^*$, because it does not involve any moment of X , which is an improvement in itself. The rationale behind (6) lies in an implicit comparison between M and m (in place of μ): if $M \leq m$ then $F(M) \leq 1/2$ and $\gamma_{AG} \geq 0$. In this way, as in the previous case, right skewness implies $\gamma_{AG} \geq 0$ through the mode-median-mean inequality. In addition, unlike the previous case, property (P2) is satisfied. This was shown by Arnold and Groeneveld (1995) assuming differentiable probability

density functions, but it holds for all $F, G \in \mathcal{F}_0^*$ with modes M_X and M_Y , respectively, that $F \leq G$ implies $F(M_X) \geq G(M_Y)$; this follows from $F = G \circ R$, where $R = G^{-1} \circ F$, and the definition of unimodality. Since (P1) follows from the equality $\bar{F}(-M) = 1 - F(M)$, we find that γ_{AG} is a valid unimodal measure of skewness of order zero (best possible order).

The coefficient γ_{AG} takes values in $[-1, 1]$ and the equality $\gamma_{AG} = 1$ is attained when $M = a$, which requires $a > -\infty$ in the support interval of F , while $\gamma_{AG} = -1$ when $M = b$, which requires $b < +\infty$. It follows that all decreasing densities exhibit maximal positive skewness, while all increasing densities exhibit maximal negative skewness. This is clearly a limitation, because γ_{AG} cannot discriminate between monotone densities of the same type. A sample version $\hat{\gamma}_{AG}$ of (6) can be obtained from an estimator \hat{M} of the mode and an estimator \hat{F} of the distribution function; the latter can be the empirical distribution function, for simplicity, while the former can be one of the estimators of the mode implemented in package `modeest` (Poncet, 2019) for R (R Core Team, 2019); see the references therein.

2.2.3 | First-order measures

A third classical measure of skewness is called *Pearson's median coefficient of skewness* or *Pearson's second coefficient of skewness*:

$$S'_K = 3 \frac{\mu - m}{\sigma}, \quad (7)$$

where the leading (arbitrary) multiplicative constant stems from an approximation of (5); see Yule (1911, p. 150). Equation (7) is well-defined for $F \in \mathcal{F}_2$ and is based, like (5), on the mode-median-mean inequality. It is clear that S'_K satisfies (P1), but like S'_K , as shown by van Zwet (1964), S'_K does not satisfy (P2). As a result we cannot consider S'_K a valid measure of skewness. However, a valid replacement for (7) is provided by Groeneveld and Meeden (1984):

$$\gamma_{GM} = \frac{\mu - m}{\mathbb{E}|X - m|}, \quad (8)$$

which is well-defined for all $F \in \mathcal{F}_1$ and we call the *Groeneveld–Meeden coefficient of skewness*. The broader domain of definition is an advantage in itself, property (P1) is clearly preserved and, moreover, γ_{GM} satisfies (P2), as shown by Groeneveld and Meeden (1984). The coefficient γ_{GM} is thus a valid measure of skewness of order one (best possible order using the mean). The mean absolute error turns out to be the right denominator for the difference between the mean and the median, if this is to be used as a measure of skewness.

We know from Jensen's inequality that $|\mathbb{E}(X - m)| \leq \mathbb{E}|X - m|$ with equality if and only if $\mathbb{P}\{X \geq m\} = 1$ or $\mathbb{P}\{X \leq m\} = 1$. It follows that $-1 < \gamma_{GM} < 1$ and the extrema of γ_{GM} are unattainable by continuous distributions; see Groeneveld (1991b) for the case of discrete distributions. A sample version $\hat{\gamma}_{GM}$ of (8) will be obtained by replacing m with the sample median \hat{m} and $\mathbb{E}|X - m|$ with its sample counterpart $\langle |X - \hat{m}| \rangle$, as well as μ with $\hat{\mu}$. Finally, we point out that (8) admits an interpretation in terms of a player betting that an observation X will exceed its median m , which is especially interesting from a financial viewpoint; see Groeneveld and Meeden (1984) for details.

A simple alternative first-order measure of skewness was suggested by Tajuddin (1999) in parallel to γ_{AG} :

$$\gamma_T = \mathbb{P}\{X \leq \mu\} - \mathbb{P}\{X \geq \mu\} = 2F(\mu) - 1. \quad (9)$$

We call γ_T in (9) the *Tajuddin coefficient of skewness*, noting that Tajuddin (1996) had previously proposed the equivalent measure $\log(F(\mu)/\{1 - F(\mu)\}) = \log(1 + \gamma_T)/(1 - \gamma_T)$. Equation (9) is clearly well-defined for all $F \in \mathcal{F}_1$, it satisfies (P1), because $\bar{F}(-\mu) = 1 - F(\mu)$, and it satisfies (P2), because Jensen's inequality gives $\mathbb{E}(Y) = \mathbb{E}(G^{-1}(F(X))) \geq G^{-1}(F(\mu))$ if $X \leq Y$ ($F \leq G$); see also Tajuddin (1996). It follows that γ_T is a valid alternative to γ_{GM} .

The rationale behind (9) is again the mode-median-mean inequality for right-skewed distributions: if a return is right-skewed, then it is probably below average. It may sound counterintuitive that investors like such returns, but a different wording is possible: if a return is right-skewed, then on average it is in the right tail of its distribution. This may sound more palatable, but neither formulation has any impact on the validity of γ_T . As for the values that γ_T can take, it is immediate to see that $-1 < \gamma_T < 1$. The extrema cannot be attained, because F is continuous, but in Section 3 we present an example where $\gamma_T = 1 - 2\lambda \rightarrow 1$ as $\lambda \downarrow 0$. Finally, a sample version $\hat{\gamma}_T$ of γ_T can be obtained from (9) by estimating F with the empirical distribution function and μ with $\hat{\mu} = \langle X \rangle$.

2.2.4 | Zeroth-order measures

None of the measures of skewness presented until now is well-defined for all $F \in \mathcal{F}_0$. A possibility in this sense is offered by the *quantile coefficient of skewness*

$$B_\alpha = \frac{\{F^{-1}(1 - \alpha) - m\} - \{m - F^{-1}(\alpha)\}}{F^{-1}(1 - \alpha) - F^{-1}(\alpha)} = \frac{\{F^{-1}(1 - \alpha) + F^{-1}(\alpha)\}/2 - m}{\{F^{-1}(1 - \alpha) - F^{-1}(\alpha)\}/2}, \quad (10)$$

where $\alpha \in]0, 1/2[$ and a typical choice is $\alpha = 1/4$. The quartile coefficient of skewness $B_{1/4}$ dates back to Bowley (1920, p. 116) and is known as the *Bowley–Yule coefficient of skewness*, because the coefficient $2B_{1/4}$ can be traced back to Yule (1911, p. 150). Groeneveld and Meeden (1984) introduced B_α , inspired by Hinkley (1975), and also let $\alpha \downarrow 0$ to obtain the coefficient $B_0 = (a + b - 2m)/(b - a)$ for distributions with a bounded support interval, that is, with $a > -\infty$ and $b < +\infty$. Remarkably, if both the numerator and denominator in (10) are integrated with respect to α from 0 to 1/2, before taking their ratio, the coefficient γ_{GM} in (8) emerges (assuming $F \in \mathcal{F}_1$). It was shown by Groeneveld and Meeden (1984) that B_α satisfies (P1) and (P2) for all $\alpha \in [0, 1/2[$. Hence, we have a family of valid zeroth-order measures of skewness (making no assumptions on the moments of the distribution).

Groeneveld and Meeden (2009) suggest a variant of (10) that is appropriate when the direction of skewness is known *a priori*, but we do not deal with this case here. Brys, Hubert, and Struyf (2003) argue that the octile coefficient $B_{1/8}$ is more appropriate to detect asymmetry than the quartile coefficient $B_{1/4}$, because it uses more information from the tails of the distribution, but they also note that $B_{1/4}$ is less sensitive to outliers (more robust) than $B_{1/8}$. In the end, this tension between sensitivity and robustness is at the heart of the choice of α in (10) and, more generally, of a measure of skewness or any other distributional summary.

It is easy to see that $-1 < B_\alpha < 1$ for all $\alpha \in [0, 1/2[$; the extreme values are unattainable by continuous distributions, because $B_\alpha = -1$ would require $F^{-1}(1 - \alpha) = m$ and $B_\alpha = 1$ would require $F^{-1}(\alpha) = m$, but see Groeneveld (1991b) for discrete distributions. A sample version of (10) will be obtained by replacing all quantiles of F by their sample counterparts (quantiles of the empirical distribution function); in particular, of course, \hat{m} will replace m . The coefficients B_0 and $B_{1/4}$ admit an interpretation analogous to that of γ_{GM} ; see Groeneveld and Meeden (1984). The coefficient B_α features, together with the median m and the α th interquantile range

$S_\alpha = F^{-1}(1 - \alpha) - F^{-1}(\alpha)$, in an interesting decomposition of $F^{-1}(\alpha)$, where α varies in $\alpha \in [0, 1/2[$; see Benjamini and Krieger (1996).

3 | THE RISK ASYMMETRY INDEX

Let X be a random variable with mean μ and standard deviation σ . The centered variable $X - \mu$ can be written as the sum of its positive part $(X - \mu)_+ = \max(0, X - \mu)$ and its negative part $(X - \mu)_- = \max(0, \mu - X)$. Accordingly, the variance of X can be written as $\sigma^2 = \sigma_U^2 + \sigma_D^2$, where $\sigma_U^2 = \mathbb{E}(X - \mu)_+^2$ is called the *upside variance* of X and $\sigma_D^2 = \mathbb{E}(X - \mu)_-^2$ is called the *downside variance* of X ; note, however, that σ_U^2 and σ_D^2 are not ordinary variances, and in particular not the variances of $(X - \mu)_+$ and $(X - \mu)_-$, but rather their second order moments. The quantities σ_U and σ_D are called the *upside sd* and *downside sd* of X , respectively. From a financial viewpoint, σ_U represents “good” volatility and σ_D represents “bad” volatility, while σ represents “total” volatility. The *risk asymmetry index* (Elyasiani et al., 2018) is defined as $\text{RAX} = (\sigma_U - \sigma_D)/\sigma$ and represents the relative excess of “good” volatility (with respect to “bad” volatility) in the distribution of returns modeled by X . The rationale behind this definition is to compare above average returns with below average returns in terms of their root mean squared residuals. We show in the following that such a comparison results in a valid measure of skewness.

The RAX can be rewritten as:

$$\text{RAX} = \frac{\sigma_U - \sigma_D}{\sigma} = \frac{\sigma_U}{\sigma} - \frac{\sigma_D}{\sigma} = \sqrt{\frac{\sigma_U^2}{\sigma^2}} - \sqrt{1 - \frac{\sigma_U^2}{\sigma^2}} = \sqrt{1 - \frac{\sigma_D^2}{\sigma^2}} - \sqrt{\frac{\sigma_D^2}{\sigma^2}}, \quad (11)$$

that is, as a strictly increasing function of the *relative upside variance* σ_U^2/σ^2 or, alternatively, as the opposite function of the *relative downside variance* $\sigma_D^2/\sigma^2 = 1 - \sigma_U^2/\sigma^2$. This rewriting is useful to show that RAX is a valid measure of skewness. In fact, property (P1) follows directly from the fact that the upside variance of $-X$ is the downside variance of X . As for property (P2), we first note that (11) is location-scale invariant. This allows us to focus on the standard case $\mu = 0$ and $\sigma = 1$. In this case, we have $\sigma_U^2 = \mathbb{E}X_+^2$ and it follows from theorem 5.3 in Oja (1981) that $X \leq Y$ implies $\mathbb{E}X_+^2 \leq \mathbb{E}Y_+^2$. Then, by (11), the same inequality holds for RAX and (P2) holds. We conclude that RAX is a valid second-order measure of skewness and, as such, it fills a gap in the literature reviewed in Section 2.2.

It is clear from the decomposition of variance in its upside and downside components that $0 < \sigma_U^2/\sigma^2 < 1$. If X is symmetric, then $\sigma_U^2 = \sigma_D^2$ and $\sigma_U^2/\sigma^2 = 1/2$, so that $\text{RAX} = 0$. The following example shows that the relative upside variance can come arbitrarily close to 1 for a suitable choice of F in \mathcal{F}_2 . Let X be a random variable with probability density function defined by $f(x) = (1 - \lambda)f_-(x) + \lambda f_+(x)$, $x \in \mathbb{R}$, with

$$f_-(x) = \frac{2\lambda^2}{(1 - \lambda)^2} f_u(x) + \frac{1 - 2\lambda - \lambda^2}{(1 - \lambda)^2} f_t(x), \quad x < 0 \quad \text{and} \quad f_+(x) = \lambda e^{-\lambda x}, \quad x \geq 0, \quad (12)$$

$$f_t(x) = (1 - \lambda)\{1 - |1 + (1 - \lambda)x|\} \quad \text{and} \quad f_u(x) = \frac{1 - \lambda}{2}, \quad -\frac{2}{(1 - \lambda)} < x < 0, \quad (13)$$

while $f_-(x) = 0$ for $x \geq 0$, $f_+(x) = 0$ for $x < 0$, $f_t(x) = f_u(x) = 0$ for $x \leq -2/(1 - \lambda)$, and $\lambda \in]0, \sqrt{2} - 1[$; note that f is continuous at 0 and therefore on its support interval $]a, b[=] - 2/(1 - \lambda), +\infty[$. Since $\mathbb{E}(X) = 0$, $\mathbb{E}(X_+^2) = 2/\lambda$ and $\mathbb{E}(X_-^2) \leq \{2/(1 - \lambda)\}^2(1 - \lambda) = 4/(1 - \lambda)$, so that $\sigma_D^2/\sigma_U^2 \leq 2\lambda/(1 - \lambda)$, we find $\sigma_U^2/\sigma^2 = 1/(1 + \sigma_D^2/\sigma_U^2) \rightarrow 1$, as $\lambda \downarrow 0$. Similarly, since $\sigma_D^2/\sigma^2 \rightarrow 0$, the relative

upside variance of $-X$ can get arbitrarily close to 0. It follows that $RAX \rightarrow 1$ for X and $RAX \rightarrow -1$ for $-X$. Note that $\mathbb{P}\{X \leq 0\} = 1 - \lambda$, so that $\gamma_T = 1 - 2\lambda$ as anticipated in Section 2.2.

A sample version of RAX can be obtained from (11) by replacing σ^2 with the sample variance $\hat{\sigma}^2$ and σ_U^2 with the *sample upside variance* $\hat{\sigma}_U^2 = \langle (X - \hat{\mu})_+^2 \rangle$ or σ_D^2 with the *sample downside variance* $\hat{\sigma}_D^2 = \langle (X - \hat{\mu})_-^2 \rangle$, where of course $\hat{\mu} = \langle X \rangle$ is the sample mean. An advantage of RAX is that risk-neutral versions of these quantities are easy to obtain from option data in a model-free set-up (Bakshi, Kapadia, & Madan, 2003; Muzzioli, 2013a, 2013b); this was indeed the original setting of Elyasiani et al. (2018). In this setting σ_U and σ_D represent the upside corridor implied volatility and the downside corridor implied volatility, respectively (Carr & Madan, 1998; Muzzioli, 2013a, 2013b). From an economic point of view the upside corridor is associated with “good” volatility, as it refers to the potential for substantial gains. On the other hand, the downside corridor is associated with “bad” volatility due to the risk of heavy losses for investors. Note that, in principle, the entire risk-neutral distribution function of returns can be recovered from option data, because it is the discounted first derivative of the European put price, but in practice it can be tricky to go beyond the first moments; see Birru and Figlewski (2012) for an example of work in this direction.

4 | EMPIRICAL FINDINGS

In this section we follow the approach of Brys et al. (2003), Tajuddin (1996, 1999) to examine the robustness to outliers and the sensitivity to changes in the shape of the distribution of the RAX introduced in Section 3, together with other valid measures of skewness. More specifically, we consider simulated data from four common distributions (Gamma, Weibull, Lognormal, and Pareto) for different values of their shape parameter and we compare four measures of skewness: the Groeneveld–Meeden (γ_{GM}), the Bowley–Yule ($B_{1/4}$), the RAX and the Tajuddin (γ_T) coefficient of skewness (all taking values between -1 and 1).

Recall that a Gamma distribution has probability density function given by

$$f_{\alpha,\beta}(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), \quad 0 < x < +\infty, \tag{14}$$

where $\alpha > 0$ is a shape parameter, while β is a scale parameter and can therefore be ignored (set to 1) as far as skewness is concerned. Arnold and Groeneveld (1995) highlight that the parameterization of Gamma distributions in terms of α respects the convex ordering of distributions. Indeed, the Fisher–Pearson coefficient of skewness is given by $\gamma_1 = 2/\sqrt{\alpha}$ and it decreases as α increases. A Weibull distribution has probability density function given by

$$f_{\alpha,\lambda}(x) = \frac{\alpha}{\lambda} \left(\frac{x}{\lambda}\right)^{\alpha-1} \exp\left(-\left(\frac{x}{\lambda}\right)^\alpha\right), \quad 0 < x < +\infty, \tag{15}$$

where $\alpha > 0$ is a shape parameter and λ is a scale parameter (set to 1 without loss of generality). The Fisher–Pearson coefficient of skewness for a Weibull distribution is known to be positive for small α , decreasing with α and negative for large α (Tajuddin, 1999). A Log-normal distribution has the following probability density function:

$$f_{\alpha,\mu}(x) = \frac{\alpha}{x\sqrt{2\pi}} \exp\left(-\frac{\alpha^2(\ln x - \mu)^2}{2}\right), \quad 0 < x < +\infty, \tag{16}$$

where $\alpha > 0$ is a shape parameter ($\alpha = 1/\sigma$) and $\mu = 0$, without loss of generality, because $\exp(\mu)$ is a scale parameter. The Fisher–Pearson coefficient of skewness for a Log-normal distribution is $\gamma_3 = (\exp(-\alpha^2) + 2)\sqrt{\exp(-\alpha^2) - 1}$, which decreases with α ; see for example Tajuddin (1999). Finally, the probability density function of a Pareto distribution is

$$f_{\alpha, x_0}(x) = \frac{\alpha x_0^\alpha}{x^{\alpha+1}}, \quad x_0 < x < +\infty, \quad (17)$$

where $\alpha > 0$ is a shape parameter, while x_0 is set to 1 (shape parameter). Note that, for a Pareto distribution, the expected value and variance are finite if $\alpha > 1$ and $\alpha > 2$, respectively, while the Fisher–Pearson coefficient of skewness requires $\alpha > 3$, and it is decreasing with α .

In the Appendix, we report the average estimated skewness by varying α in a distribution specific set of values, and letting $n = 30, 100, 1,000$. Since the standard errors are small, and the four measures of skewness behave similarly for all sample sizes, in Figure 1 we focus on the

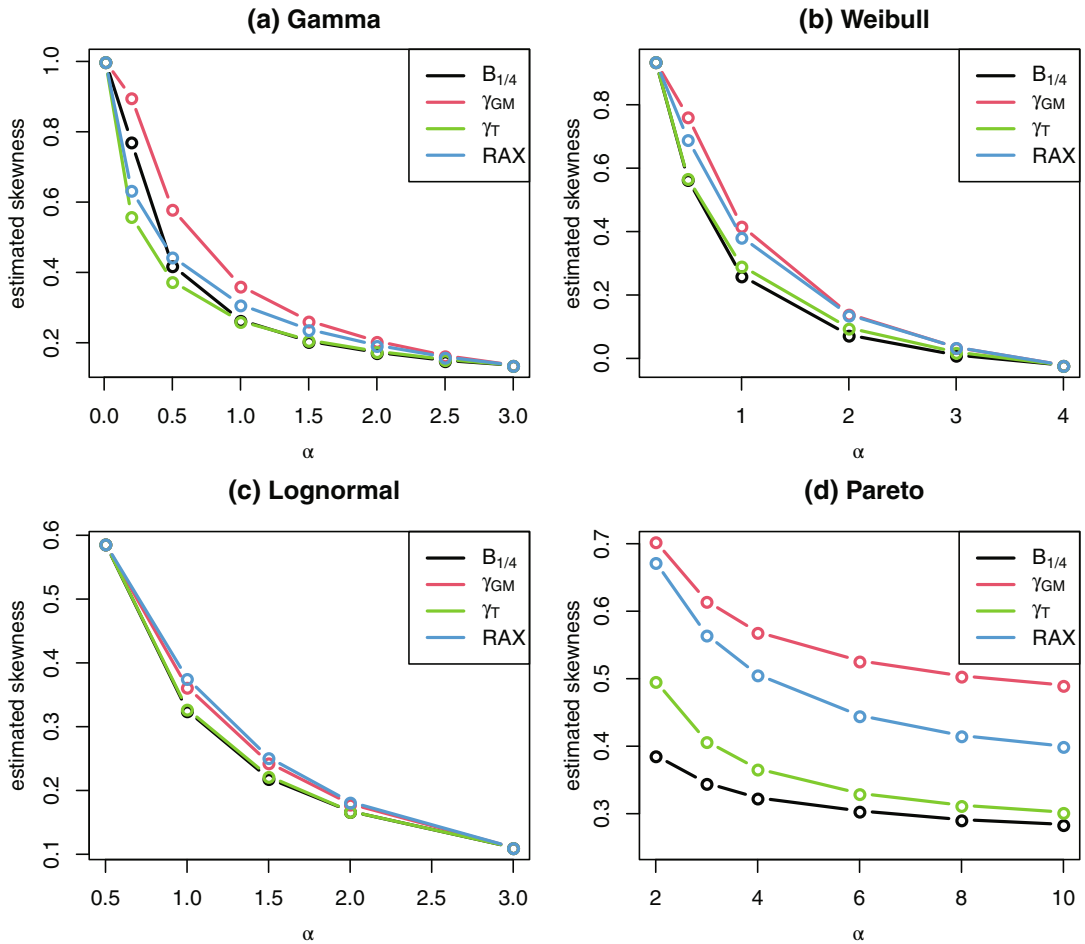


FIGURE 1 Average of skewness estimates over 1,000 samples of size $n = 100$ for several values of the shape parameter α of the Gamma (panel a), Weibull (panel b), Lognormal (panel c) and Pareto (panel d) distributions

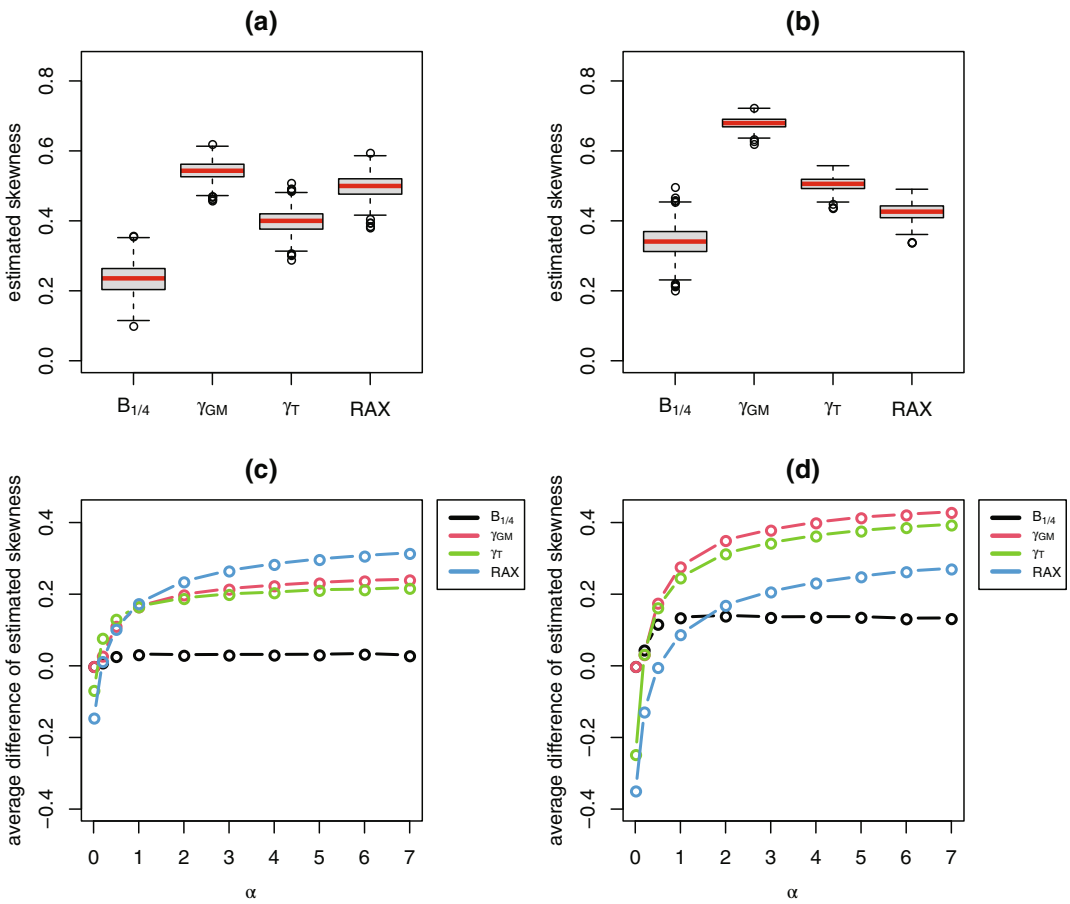


FIGURE 2 Boxplots of skewness estimates on 1000 random samples of $n = 1,000$ observations from a Gamma distribution with shape parameter $\alpha = 1.5$ without contamination (a) and with 15% contamination (b). Difference between average skewness estimates at contaminated and at uncontaminated data for different values of α of the Gamma distribution and 5% (c) and 15% (d) contamination, respectively

average estimated skewness, as a function of α , for samples of size $n = 100$. In the four cases considered, we expect all four measures to decrease monotonically in α and Figure 1 confirms this expectation. Each measure starts close to 1 when α assumes its minimum value (indicating high skewness) and falls toward 0 as α grows (indicating low skewness). For the Gamma distribution, in Figure 1a, we see that $B_{1/4}$ and γ_T attain the smallest values, while RAX maintains between γ_{GM} (upper limit) and $B_{1/4}$ or γ_T (lower limit) for every α . Similar results hold for the Weibull distribution, in Figure 1b, and for the Pareto distribution, in Figure 1d, whereas for the Log-normal distribution, in Figure 1c, the RAX remains slightly above the other measures of skewness, which are here very close to each other. This means that, across a variety of probability distributions, RAX displays an intermediate sensitivity to changes in the shape of each distribution.

As a further step, we analyze the robustness of the four measures of skewness with respect to the influence of a number of outliers. For this purpose, in panels a of Figures 2–5 we

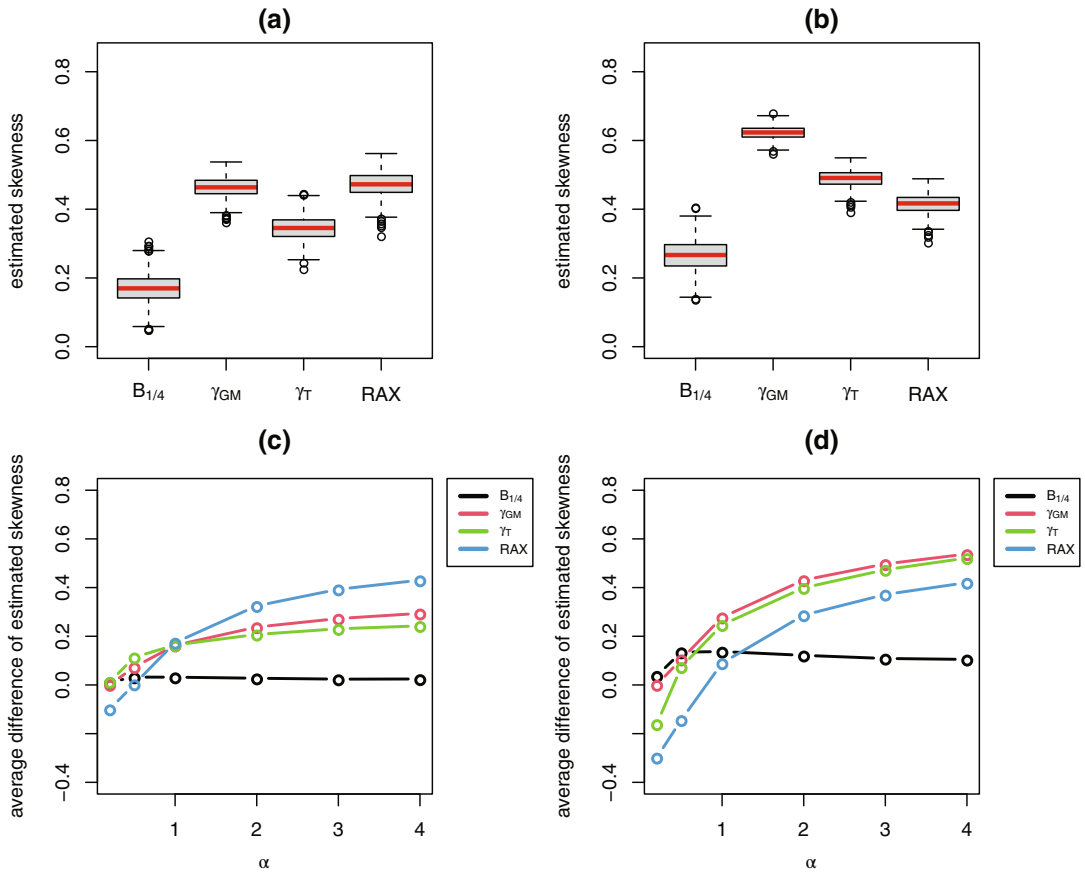


FIGURE 3 Boxplots of skewness estimates on 1000 random samples of $n = 1,000$ observations from a Weibull distribution with shape parameter $\alpha = 1.5$ without contamination (a) and with 15% contamination (b). Difference between average skewness estimates at contaminated and at uncontaminated data for different values of α of the Weibull distribution and 5% (c) and 15% (d) contamination, respectively

propose the boxplots of the skewness estimates on 1,000 random samples of $n = 1,000$ observations, while panels b of Figures 2–5 depict the same boxplots where we replaced 15% of the data with outliers spaced 8 sds to the right of the mean. The value of α for each distribution is fixed, and it is equal to $\alpha = 1.5$ for the Gamma and the Weibull distributions, and $\alpha = 0.8$ and $\alpha = 3$ for the Log-normal and Pareto distributions. In the four numerical simulations conducted, it may be seen that the median values increase for all measures of skewness, bringing the boxes with them, whereas on the other hand the RAX shows a decrease. We will see that γ_T can also decrease upon contamination, for smaller values of α , while this will not happen to γ_{GM} and $B_{1/4}$.

We deeper explore robustness in panels c and d of Figures 2–5, where further simulations are depicted. Specifically, panels c and d of Figures 2–5 show, for each measure of skewness and several values of α , the difference between the average estimated value at the contaminated and at the original datasets. We replaced a percentage of the original data with outliers under two different contamination levels: we contaminated our data at 5% in Figures 2–5c and

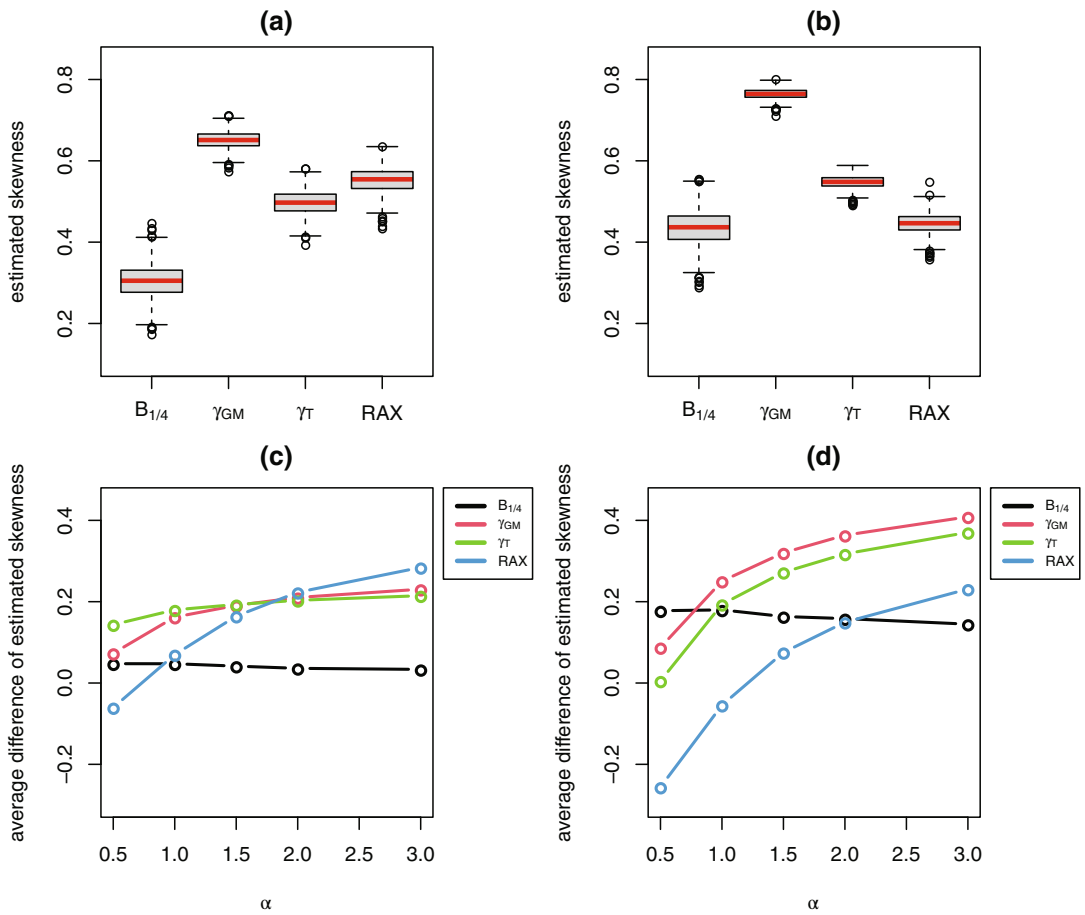


FIGURE 4 Boxplots of skewness estimates on 1,000 random samples of $n = 1,000$ observations from a Log-normal distribution with shape parameter $\alpha = 0.8$ without contamination (a) and with 15% contamination (b). Difference between average skewness estimates at contaminated and at uncontaminated data for different values of α of the Log-normal distribution and 5% (c) and 15% (d) contamination, respectively

at 15% in Figure 2–5d. Focusing on the absolute skewness change upon contamination, we note that $B_{1/4}$ stands out as rather insensitive to the presence of outliers, while the performance of RAX is competitive with that of γ_T and γ_{GM} . As for the sign of change, as anticipated, we observe in panels c and d of Figures 2–5 a possible decrease of γ_T and RAX, but not of $B_{1/4}$, nor of γ_{GM} .

As an aside, following Tajuddin (1999), we analyze the skewness sign pattern according to the relationships between the mean, median, and mode of the Weibull distribution. Table 1 presents the signs of the four measures of skewness depending on the value of α . It may be observed that the sign of $B_{1/4}$ is opposite to the signs of all other measures when $3.2589 < \alpha < 3.4395$. In this case, $B_{1/4}$ fails to preserve the sign of $(\mu - m)$, while RAX, γ_T , and γ_{GM} respect the relationship between μ and m . Based on our empirical findings, we can conclude that RAX strikes a good balance between robustness to outliers and sensitivity to changes in the shape of the distribution.

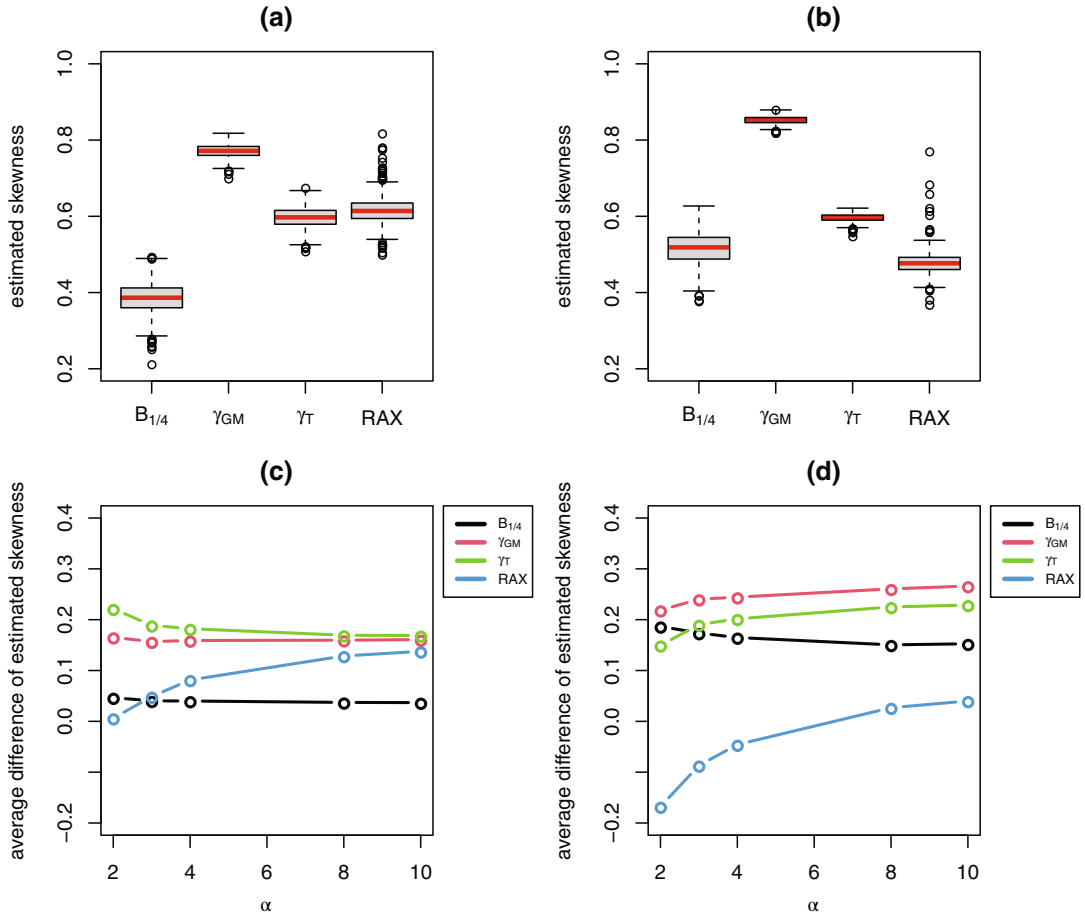


FIGURE 5 Boxplots of skewness estimates on 1000 random samples of $n = 1,000$ observations from a Pareto distribution with shape parameter $\alpha = 3$ without contamination (a) and with 15% contamination (b). Difference between average skewness estimates at contaminated and at uncontaminated data for different values of α of the Pareto distribution and 5% (c) and 15% (d) contamination, respectively

TABLE 1 Sign patterns of skewness values given by different measures in the Weibull distribution case

α	Relationship of M, m, μ	γ_{GM}	$B_{1/4}$	RAX	γ_T
$0 < \alpha \leq 1$	$m < \mu$	+	+	+	+
$1 < \alpha < 3.25889$	$M < m < \mu$	+	+	+	+
$3.2589 < \alpha < 3.3125$	$m < M < \mu$	+	-	+	+
$3.3125 < \alpha < 3.4395$	$m < \mu < M$	+	-	+	+
$3.4395 < \alpha$	$\mu < m < M$	-	-	-	-

5 | DISCUSSION

In this paper we presented a comprehensive framework for the assessment of univariate skewness. We reviewed existing measures and proposed a new one, called RAX, based on Elyasiani et al. (2018). We showed that RAX is a valid measure of skewness and can be safely used by scholars. By using simulated data, we found that RAX strikes a good balance between robustness to outliers and sensitivity to changes in the shape of the distribution.

RAX is the relative difference between upside and downside volatility. We used volatility, following Elyasiani et al. (2018), due to its high standing in finance. In principle, we could also compare above average returns with below average returns in terms of their mean absolute residuals, rather than root mean squared residuals, but we would still need volatility in the denominator to satisfy (P2).

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from the FAR2017 and FAR2019 projects “The role of Asymmetry and Kolmogorov equations in financial Risk Modelling (ARM)” and “Risk assessment in the EU: new indices based on machine learning methods (REU)”. We also wish to thank William Bromwich for his painstaking attention to the copy-editing of this paper. Open Access Funding provided by Università degli Studi di Modena e Reggio Emilia within the CRUI-CARE Agreement.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Giovanni Campisi  <https://orcid.org/0000-0002-6725-8490>

Luca La Rocca  <https://orcid.org/0000-0003-0495-1308>

Silvia Muzzioli  <https://orcid.org/0000-0003-0738-6690>

REFERENCES

- Abadir, K. M. (2005). The mean-median-mode inequality: Counterexamples. *Econometric Theory*, 21, 477–482.
- Arnold, B. C., & Groeneveld, R. A. (1993). *Skewness and kurtosis orderings: An introduction*. In M. Shaked & Y. L. Tong (Eds.), *Stochastic inequalities Lecture Notes-Monograph Series* (Vol. 22, pp. 17–24). Hayward: Institute of Mathematical Statistics.
- Arnold, B. C., & Groeneveld, R. A. (1995). Measuring skewness with respect to the mode. *The American Statistician*, 49, 34–38.
- Bakshi, G., Kapadia, N., & Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies*, 16, 101–143.
- Benjamini, Y., & Krieger, A. M. (1996). Concepts and measures for skewness with data-analytic implications. *Canadian Journal of Statistics*, 24, 131–140.
- Birru, J., & Figlewski, S. (2012). Anatomy of a meltdown: The risk neutral density for the S&P 500 in the fall of 2008. *Journal of Financial Markets*, 15, 151–180.
- Bowley, A. L. (1920). *Elements of statistics* (4th ed.). New York: Scribner.
- Brys, G., Hubert, M., & Struyf, A. (2003). *A comparison of some new measures of skewness*. In R. Dutter, P. Filzmoser, U. Gather, & P. J. Rousseeuw (Eds.), *Developments in robust statistics* (pp. 98–113). Heidelberg: Physica.

- Carr, P., & Madan, D. (1998). *Towards a theory of volatility trading*. In R. A. Jarrow (Ed.), *Volatility: New estimation techniques for pricing derivatives* (pp. 417–427). London: Risk Books.
- Chen, C. W., So, M. K., & Gerlach, R. H. (2005). Asymmetric response and interaction of us and local news in financial markets. *Applied Stochastic Models in Business and Industry*, 21, 273–288.
- Cox, N. J. (2010). Speaking Stata: The limits of sample skewness and kurtosis. *The Stata Journal*, 10, 482–495.
- Dharmadhikari, S., & Joag-Dev, K. (1988). *Unimodality, convexity, and applications*. San Diego: Academic Press.
- Doane, D. P., & Seward, L. E. (2011). Measuring skewness: A forgotten statistic? *Journal of Statistics Education*, 19, 1–18.
- Doksum, K. A. (1975). Measures of location and asymmetry. *Scandinavian Journal of Statistics*, 2, 11–22.
- Elyasiani, E., Gambarelli, L., & Muzzioli, S. (2018). The risk-asymmetry index as a new measure of risk. *Multinational Finance Journal*, 22, 173–210.
- Elyasiani, E., Gambarelli, L., & Muzzioli, S. (2020). Moment risk premia and the cross-section of stock returns in the european stock market. *Journal of Banking & Finance*, 111, 105732.
- Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38, 34–105.
- Fisher, R. A. (1929) Moments and product moments of sampling distributions. *Proceedings of the London Mathematical Society: Series 2*, 30, 199–238. Corrigenda in 30(1), 552–552.
- Groeneveld, R. A. (1991a). An influence function approach to describing the skewness of a distribution. *The American Statistician*, 45, 97–102.
- Groeneveld, R. A. (1991b). Sharp inequalities for skewness measures. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 40, 387–392.
- Groeneveld, R. A., & Meeden, G. (1984). Measuring skewness and kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 33, 391–399.
- Groeneveld, R. A., & Meeden, G. (2009). An improved skewness measure. *METRON*, 67, 325–337.
- Hinkley, D. V. (1975). On power transformations to symmetry. *Biometrika*, 62, 101–111.
- Holgerson, T. H. E. (2010). A modified skewness measure for testing asymmetry. *Communications in Statistics: Simulation and Computation*, 39, 335–346.
- Kirby, W. (1974). Algebraic boundedness of sample statistics. *Water Resources Research*, 10, 220–222.
- Lempérière, Y., Deremble, C., Nguyen, T.-T., Seager, P., Potters, M., & Bouchaud, J.-P. (2017). Risk premia: Asymmetric tail risks and excess returns. *Quantitative Finance*, 17, 1–14.
- Li, X., & Morris, J. M. (1991). On measuring asymmetry and the reliability of the skewness measure. *Statistics & Probability Letters*, 12, 267–271.
- MacGillivray, H. L. (1986). Skewness and asymmetry: Measures and orderings. *The Annals of Statistics*, 14, 994–1011 Correction in *The Annals of Statistics* 1987, Vol. 15(2), 884.
- Muzzioli, S. (2013a). The forecasting performance of corridor implied volatility in the Italian market. *Computational Economics*, 41, 359–386.
- Muzzioli, S. (2013b). The information content of option-based forecasts of volatility: Evidence from the Italian stock market. *The Quarterly Journal of Finance*, 3, 1350005.
- Oja, H. (1981). On location, scale, skewness and kurtosis of univariate distributions. *Scandinavian Journal of Statistics*, 8, 154–168.
- Pearson, K. (1895). Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London: Series A*, 186, 343–414.
- Pearson, K. (1901). Mathematical contributions to the theory of evolution. X. Supplement to a memoir on skew variation. *Philosophical Transactions of the Royal Society of London: Series A*, 197, 443–459.
- Pearson, K. (1916). Mathematical contributions to the theory of evolution. XIX. Second supplement to a memoir on skew variation. *Philosophical Transactions of the Royal Society of London: Series A*, 216, 429–457.
- Peiró, A. (1999). Skewness in financial returns. *Journal of Banking & Finance*, 23, 847–862. <http://www.sciencedirect.com/science/article/pii/S0378426698001198>
- Poncet, P. (2019) modeest: Mode estimation. R package version 2.4.0. <http://CRAN.R-project.org/package=modeest>
- R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Sato, M. (1997). Some remarks on the mean, median, mode and skewness. *Australian Journal of Statistics*, 39, 219–224.

- Tajuddin, I. H. (1996). A simple measure of skewness. *Statistica Neerlandica*, 50, 362–366.
- Tajuddin, I. H. (1999). A comparison between two simple measures of skewness. *Journal of Applied Statistics*, 26, 767–774.
- van Zwet, W. R. (1964). Convex transformations: A new approach to skewness and kurtosis. *Statistica Neerlandica*, 18, 433–441.
- van Zwet, W. R. (1979). Mean, median, mode II. *Statistica Neerlandica*, 33, 1–5.
- Whaley, R. E. (2000). The investor fear gauge. *The Journal of Portfolio Management*, 26, 12–17.
- Whaley, R. E. (2009). Understanding the VIX. *The Journal of Portfolio Management*, 35, 98–105.
- Wilkins, J. E. (1944). A note on skewness and kurtosis. *The Annals of Mathematical Statistics*, 15, 333–335.
- Yule, G. U. (1911). *An introduction to the theory of statistics*. London: Griffin.
- Zwillinger, D., & Kokoska, S. (1999). *CRC standard probability and statistics tables and formulae*. Boca Raton: CRC Press.

How to cite this article: Campisi, G., La Rocca, L., & Muzzioli, S. (2022). Assessing skewness in financial markets. *Statistica Neerlandica*, 1–23. <https://doi.org/10.1111/stan.12273>

APPENDIX A. SENSITIVITY ANALYSIS

In this Appendix we conduct a sensitivity analysis of the four measures of skewness by varying α in a distribution specific set of values in a sample of n observations, with $n = 30, 100, 1,000$. Table A1 is for the Gamma distribution, Table A2 is for the Weibull distribution, Table A3 is for the Log-normal distribution, Table A4 is for the Pareto distribution. Each table reports results based on 1,000 samples. Note that, in all the tables, the standard errors are small, and the four measures of skewness behave similarly for all sample sizes. Therefore, we can conclude, that the results of Section 4 on the average estimated skewness, as a function of α , are valid for different sample sizes.

TABLE A1 Average estimated skewness and standard error of the Gamma distribution for 1,000 samples of $n = 30, 100, 1,000$ observations and for α ranging from 0.01 to 3

	n	γ_{GM}	$B_{1/4}$	RAX	γ_T
$\alpha = 0.01$	30	1.0000	1.0000	0.7663	0.8756
		(0.0000)	(0.0000)	(0.0014)	(0.0019)
	100	1.0000	1.0000	0.8413	0.9109
		(0.0000)	(0.0000)	(0.0011)	(0.0007)
	1,000	1.0000	1.0000	0.8881	0.9197
		(0.0000)	(0.0000)	(0.0005)	(0.0002)
$\alpha = 0.2$	30	0.8780	0.7185	0.5210	0.5105
		(0.0030)	(0.0058)	(0.0030)	(0.0035)
	100	0.9027	0.7558	0.5704	0.5218
		(0.0013)	(0.0031)	(0.0019)	(0.0019)
	1,000	0.9116	0.7730	0.5928	0.5282
		(0.0004)	(0.0009)	(0.0007)	(0.0006)
$\alpha = 0.5$	30	0.6121	0.4004	0.3925	0.3478
		(0.0047)	(0.0074)	(0.0030)	(0.0040)
	100	0.6262	0.4119	0.4281	0.3623
		(0.0026)	(0.0041)	(0.0019)	(0.0021)
	1,000	0.6343	0.4198	0.4473	0.3651
		(0.0009)	(0.0013)	(0.0007)	(0.0007)
$\alpha = 1$	30	0.4202	0.2430	0.3029	0.2530
		(0.0051)	(0.0075)	(0.0030)	(0.0041)
	100	0.4365	0.2562	0.3308	0.2634
		(0.0028)	(0.0041)	(0.0019)	(0.0023)
	1000	0.4413	0.2592	0.3423	0.2633
		(0.0009)	(0.0013)	(0.0006)	(0.0007)
$\alpha = 3$	30	0.2382	0.1315	0.1860	0.1436
		(0.0051)	(0.0074)	(0.0028)	(0.0044)
	100	0.2444	0.1341	0.2037	0.1512
		(0.0029)	(0.0040)	(0.0016)	(0.0024)
	1,000	0.2473	0.1372	0.2094	0.1545
		(0.0009)	(0.0014)	(0.0005)	(0.0008)

TABLE A2 Average estimated skewness and standard error of the Weibull distribution for 1,000 samples of $n = 30, 100, 1,000$ observations and for α ranging from 0.5 to 4

	n	γ_{GM}	$B_{1/4}$	RAX	γ_T
$\alpha = 0.5$	30	0.7478	0.5086	0.3291	0.4777
		(0.0039)	(0.0071)	(0.0120)	(0.0044)
	100	0.7939	0.5539	0.5203	0.5001
		(0.0020)	(0.0037)	(0.0059)	(0.0024)
	1,000	0.8134	0.5682	0.6152	0.5125
		(0.0006)	(0.0012)	(0.0019)	(0.0008)
$\alpha = 1$	30	0.4068	0.2470	0.2457	0.2531
		(0.0048)	(0.0072)	(0.0077)	(0.0041)
	100	0.4267	0.2515	0.3057	0.2373
		(0.0029)	(0.0043)	(0.0043)	(0.0023)
	1,000	0.4423	0.2602	0.3412	0.2646
		(0.0009)	(0.0013)	(0.0013)	(0.0008)
$\alpha = 2$	30	0.1241	0.0608	0.0884	0.0825
		(0.0050)	(0.0073)	(0.0070)	(0.0042)
	100	0.1433	0.0775	0.1156	0.0898
		(0.0029)	(0.0042)	(0.0035)	(0.0022)
	1,000	0.1436	0.0754	0.1218	0.0880
		(0.0009)	(0.0013)	(0.0011)	(0.0007)
$\alpha = 3$	30	0.0317	0.0154	0.0223	0.0194
		(0.0049)	(0.0075)	(0.0070)	(0.0044)
	100	0.0305	0.0133	0.0274	0.0187
		(0.0029)	(0.0042)	(0.0036)	(0.0023)
	1,000	0.0303	0.0102	0.0310	0.0184
		(0.0010)	(0.0014)	(0.0011)	(0.0007)
$\alpha = 4$	30	-0.0202	-0.0130	-0.0173	-0.0129
		(0.0050)	(0.0077)	(0.0065)	(0.0042)
	100	-0.0313	-0.0227	-0.0184	-0.0182
		(0.0029)	(0.0042)	(0.0037)	(0.0024)
	1,000	-0.0283	-0.0193	-0.0193	-0.0177
		(0.0009)	(0.0013)	(0.0012)	(0.0008)

TABLE A3 Average estimated skewness and standard error of the Log-normal distribution for 1,000 samples of $n = 30, 100, 1,000$ observations and for α ranging from 0.5 to 3

	n	γ_{GM}	$B_{1/4}$	RAX	γ_T
$\alpha = 0.5$	30	0.8265	0.5321	0.1613	0.5870
		(0.0034)	(0.0017)	(0.0006)	(0.0064)
	100	0.8790	0.5684	0.4863	0.6467
		(0.0037)	(0.0011)	(0.0173)	(0.0111)
	1,000	0.9012	0.5851	0.7827	0.6746
		(0.0003)	(0.0054)	(0.0034)	(0.0013)
$\alpha = 1$	30	0.5311	0.3065	0.3005	0.3623
		(0.0046)	(0.0026)	(0.009)	(0.0074)
	100	0.5635	0.3210	0.4406	0.3763
		(0.0041)	(0.0014)	(0.0104)	(0.0058)
	1,000	0.5725	0.3218	0.5181	0.3814
		(0.0021)	(0.0050)	(0.0028)	(0.0008)
$\alpha = 1.5$	30	0.3571	0.1896	0.2352	0.2414
		(0.0041)	(0.0024)	(0.0007)	(0.0069)
	100	0.3895	0.2136	0.3190	0.2534
		(0.0040)	(0.0013)	(0.0133)	(0.0083)
	1,000	0.4018	0.2204	0.3630	0.2603
		(0.0027)	(0.0052)	(0.0034)	(0.0011)
$\alpha = 2$	30	0.2706	0.1407	0.1876	0.1817
		(0.0053)	(0.0028)	(0.0009)	(0.0076)
	100	0.2977	0.1618	0.2488	0.1942
		(0.0042)	(0.0013)	(0.0081)	(0.0044)
	1000	0.3065	0.1664	0.2758	0.1970
		(0.0014)	(0.0046)	(0.0025)	(0.0008)
$\alpha = 3$	30	0.1977	0.1160	0.1364	0.1292
		(0.0053)	(0.0028)	(0.0009)	(0.0076)
	100	0.1987	0.1042	0.1674	0.1293
		(0.0042)	(0.0013)	(0.0081)	(0.0044)
	1000	0.2049	0.1098	0.1849	0.1324
		(0.0014)	(0.0046)	(0.0025)	(0.0008)

TABLE A4 Average estimated skewness and se of the Pareto distribution for 1,000 samples of $n = 30, 100, 1,000$ observations and for α ranging from 2.1 to 10

	n	γ_{GM}	$B_{1/4}$	RAX	γ_T
$\alpha = 2.1$	30	0.7030	0.3844	0.6691	0.4967
		(0.0010)	(0.0013)	(0.0035)	(0.0014)
	100	0.7032	0.3842	0.6747	0.4975
		(0.0011)	(0.0013)	(0.0036)	(0.0015)
	1,000	0.7008	0.3833	0.6740	0.4948
		(0.0009)	(0.0012)	(0.0032)	(0.0013)
$\alpha = 3$	30	0.6136	0.3440	0.5666	0.4060
		(0.0009)	(0.0013)	(0.0003)	(0.0009)
	100	0.6134	0.3437	0.5690	0.4062
		(0.0009)	(0.0013)	(0.0003)	(0.0009)
	1,000	0.6135	0.3423	0.5665	0.4072
		(0.0008)	(0.0012)	(0.0002)	(0.0008)
$\alpha = 4$	30	0.5694	0.3219	0.5048	0.3672
		(0.0009)	(0.0013)	(0.0024)	(0.0009)
	100	0.5699	0.3243	0.5044	0.3672
		(0.0009)	(0.0013)	(0.0022)	(0.0009)
	1,000	0.5691	0.3641	0.5012	0.3664
		(0.0008)	(0.0012)	(0.0022)	(0.0008)
$\alpha = 8$	30	0.5052	0.2929	0.4175	0.3127
		(0.0009)	(0.0013)	(0.0016)	(0.0008)
	100	0.5024	0.2898	0.4160	0.3127
		(0.0009)	(0.0014)	(0.0017)	(0.0008)
	1,000	0.5034	0.2916	0.4139	0.3118
		(0.0009)	(0.0014)	(0.0016)	(0.0008)
$\alpha = 10$	30	0.4933	0.2880	0.4004	0.3027
		(0.0009)	(0.0013)	(0.0016)	(0.0008)
	100	0.4908	0.2846	0.4010	0.3026
		(0.0009)	(0.0014)	(0.0015)	(0.0007)
	1,000	0.4917	0.2855	0.3993	0.3027
		(0.0009)	(0.0014)	(0.0016)	(0.0008)