DEMS WORKING PAPER SERIES

Robust measures of skewness and kurtosis for macroeconomic and financial time series

Andrea Bastianin

No. 408 – May 2019

Robust measures of skewness and kurtosis for macroeconomic and financial time series

Andrea Bastianin

University of Milan–Bicocca

May 6, 2019

Abstract: The sample skewness and kurtosis of macroeconomic and financial time series are routinely scrutinized in the early stages of model-building and are often the central topic of studies in economics and finance. Notwithstanding the availability of several robust estimators, most scholars in economics rely on method-of-moments estimation that is known to be very sensitive to outliers. We carry out an extensive Monte Carlo analysis to compare the bias and root mean squared error of twelve different estimators of skewness and kurtosis. We consider nine statistical distributions that approximate the range of data generating processes of many macroeconomic and financial time series. Both in independently and identically distributed samples and in data generating processes featuring serial correlation L-moments and trimmed L-moments estimators are particularly resistant to outliers and deliver the lowest root mean squared error. The application to 128 macroeconomic and financial time series sourced from a large, monthly frequency, database (i.e. the FRED-MD of McCracken and Ng, 2016) confirms the findings of the simulation study.

Corresponding author: Andrea Bastianin, Department of Economics, Management, and Statistics, University of Milan – Bicocca, Via Bicocca degli Arcimboldi, 8, Building U7, 20126 - Milano – Italy. Email: `andrea.bastianin@unimib.it`.

1   Introduction

The assumption of normality is central in many theoretical models and empirical applications in macroeconomics and finance (see e.g. Cecchetti et al., 1990; DeLong and Summers, 1986; Scott and Horvath, 1980). Normality tests, sample skewness and kurtosis of time series are routinely used as diagnostic tools for econometric model-building and to investigate substantive questions in economics and finance. Bontemps and Meddahi (2005) and Kilian and Demiroglu (2000) provide several examples where normality testing is of interest to investigate key issues in economics and finance. While robust estimators and tests do exist, most scholars in economics rely on the method-of-moments to estimate the skewness and kurtosis of data and on the Jarque-Bera test as a diagnostic check. It is well known that measures of skewness and kurtosis based on sample moments are highly sensitive to outliers and even a single large observation can tear them apart (see e.g. Bonato, 2011; Kim and White, 2004; Brys et al., 2004; Thomas, 2009).

We rely on a Monte Carlo exercise to assess the bias and root mean squared error (RMSE) of twelve estimators of skewness and kurtosis. These include method-of-moments estimators, as well as more robust measures based on quantiles, L-moments and trimmed L-moments (TL-moments). While the statistical literature on robust estimation is widespread, L-moments are relatively new in time series econometrics.[1] Much like conventional moments, L-moments can be used to characterize the location, scale and shape of a probability distribution (Hosking, 1990). We show that, both in independently and identically distributed (iid) samples and in data generating processes featuring serial correlation, L-moments and trimmed L-moments (TL-moments) estimators are particularly resistant to outliers, feature a very small bias and deliver the lowest RMSE. The empirical application, that involves 128 macroeconomic and financial time series sourced from a large, monthly frequency, database (i.e. the FRED-MD; see McCracken and Ng, 2016, for details), confirms that L- and TL-moments estimators are particularly resistant to outliers.

Five studies are closely related to this paper. Kim and White (2004) and Bonato (2011)

---

[1]An exception is Darolles et al. (2009) who constructed measures of fund performance based on L-moments to overcome the drawbacks of more traditional performance metrics such as the Sharpe ratio.

deal with the estimation of skewness and kurtosis in the presence of outlying observations. Although these authors do consider robust measures of skewness and kurtosis, they do not analyze L-, nor TL-moments. Karvanen (2006) proposed estimating quantile mixtures via L-moments and presented an empirical analysis involving stock market returns, however the study did not explicitly focus on the comparative performance of skewness and kurtosis estimators. Thomas (2009) assessed the small sample behavior of measures symmetry based on quantiles, L-moments, and TL-moments with a Monte Carlo analysis, but did not consider measures of kurtosis. Harri and Coble (2011) introduced Normality tests based on L-moments, but focused exclusively on iid samples.

This paper makes three main contributions to this strand of the literature. First, by considering Monte Carlo designs with serially dependent observations, we introduce L- and TL-moments in the time series econometrics literature. Second, we present an extensive set of results based Monte Carlo designs with sample size and outlying observations aimed at mimicking the dataset typically analyzed in macroeconomics and finance. Third, we present a detailed empirical application involving a prominent large, monthly, dataset of macroeconomic and financial time series.

The rest of the paper is organized as follows. Section 2 reviews different estimators of skewness and kurtosis. Sections 3 and 4 present the Monte Carlo and empirical analysis, respectively. Section 5 concludes. An Appendix with additional methodological details and results completes the paper.

2   Measures of skewness & kurtosis

2.1   Notation

Let $\{X_t\}_{t=1}^{T}$ be a series with mean μ and r-th central moment $\mu_r = E(X - \mu)^r$. We define the variance of X_t as $\mu_2 = \sigma^2$ and its standard deviation as σ. Moreover, $Q_\alpha = F^{-1}(\alpha)$ is the quantile function, while, $X_{1:T} \leq X_{2:T} \leq \ldots \leq X_{n:T}$ are the order statistics[2] of a random sample of size T drawn from the distribution of X_t.

---

[2]The k-th order statistics of a sample of size T is equal to its k-th smallest value. For the sample $\{5, 3, 9, 1\}$, we have: $X_{1:4} = 1, X_{2:4} = 3, X_{3:4} = 5, X_{4:4} = 9$.

Measures of skewness and kurtosis are denoted as SK_j and KR'_j. We focus on excess kurtosis (KR_j), defined as the difference between the j-th measure of kurtosis $\left(KR'_j\right)$ and the corresponding reference value for the Normal distribution. These reference values are shown in the first column of Table 1(b). Details on the estimation procedures, as well as further methodological issues are discussed in Section A and B of the Appendix.

2.2 Conventional measures of skewness & kurtosis

The two most widely used measures of skewness and kurtosis are based on standardized moments. The moment-based skewness is defined as:

$$SK_1 = \frac{\mu_3}{\sigma^3} \tag{1}$$

for the Normal distribution $SK_1 = 0$. Symmetric distributions have zero skewness, while for asymmetric distributions with longer left (right) tail skewness SK_1 is negative (positive).[3]

The kurtosis of a distribution measures its tail behavior, that is its propensity to generate observations far away from the center of its support (Westfall, 2014). Sample kurtosis is conventionally based on method-of-moments estimation of:

$$KR'_1 = \frac{\mu_4}{\sigma^4} \tag{2}$$

for the Normal distribution $KR'_1 = 3$, therefore the excess kurtosis of a distribution is defined as $KR_1 = KR'_1 - 3$. A distribution is thus said to have "fat-tails" if $KR_1 > 3$.

It is well known that both SK_1 and KR_1 have a very low breakdown value, in fact a single outlier can severely affect them, leading to extremely large values (Kim and White, 2004; Hubert and Debruyne, 2009). Another pitfall of these measures is that they cannot be defined for distributions that do not posses finite moments up to the fourth order. A case in

---

[3]Zero skewness however is only necessary, but not sufficient for the symmetry of a distribution (von Hippel, 2010). A case in point is the discrete Student's t distribution (Ord, 1968). In the case of the Normal mixture distribution (Meijer, 2000) showed that a mixture of two Normal distributions with $\mu_1 = -2, \sigma_1 = 1, \mu_2 = 1, \sigma_2 = \sqrt{2}$ and probability of drawing from the first distribution equal to 1/3 has zero central moment but is asymmetric, that is: $E(X_t - \mu) \neq -E(X_t - \mu)$. A third example of an asymmetric distribution with zero skewness is reported by Bai and Ng (2001, p. 230).

point is Stable distribution family, often used to describe financial time series (Bonato, 2011; Mittnik and Rachev, 1993; Loretan and Phillips, 1994). Method-of-moments estimation of SK_1 and KR_1 is discussed in Section B.1 of the Appendix.

2.3 Robust measures of skewness

Given the drawbacks of SK_1 and KR_1, both Kim and White (2004) and Bonato (2011) have highlighted the advantages of alternative measures of skewness and kurtosis in financial applications.[4]

Several robust estimators of skewness rely on quantiles and are encompassed by the measure proposed by Hinkley (1975):

$$SK(\alpha) = \frac{Q_{1-\alpha} + Q_\alpha - 2Q_{0.5}}{Q_{1-\alpha} - Q_\alpha} \quad \text{for } 0 \leq \alpha \leq 0.5 \tag{3}$$

The quartile skewness due to Bowley (1920) is obtained setting $\alpha = 0.25$:

$$\begin{aligned} SK_2 &= \frac{Q_{0.75} + Q_{0.25} - 2Q_{0.50}}{Q_{0.75} - Q_{0.25}} \\ &= \frac{(Q_{0.75} - Q_{0.50}) - (Q_{0.50} - Q_{0.25})}{Q_{0.75} - Q_{0.25}} \end{aligned} \tag{4}$$

where $-1 \leq SK_2 \leq 1$. The denominator in Equation (4) is the interquartile range and it re-scales the numerator so that $SK_2 = 1$ ($SK_2 = -1$) in the case of extreme right (left) skewness. For symmetric distributions $(Q_{0.75} - Q_{0.50}) = (Q_{0.50} - Q_{0.25})$ and hence $SK_2 = 0$.

One drawback of the Hinkley's measure of skewness is its dependence on the level of α. To overcome this limitation Groeneveld and Meeden (1984) proposed to integrate out α:

$$SK_3 = \frac{\int_0^{0.5} [Q_{1-\alpha} + Q_\alpha - 2Q_{0.5}] \mathrm{d}\alpha}{\int_0^{0.5} [Q_{1-\alpha} - Q_\alpha] \mathrm{d}\alpha} = \frac{\mu - Q_{0.5}}{E\left(|X_t - Q_{0.5}|\right)} \tag{5}$$

notice that also in this case $-1 \leq SK_3 \leq 1$ and $SK_3 = 0$ for symmetric distributions.

The denominator of Equation (5) is the mean absolute deviation (MAD); substituting

---

[4]There is a voluminous literature on robust statistics and on tests of symmetry. See e.g. Brys et al. (2003, 2004) and references therein.

the MAD with the standard deviation, we get the Pearson coefficient of skewness:

$$SK_4 = \frac{\mu - Q_{0.5}}{\sigma} \tag{6}$$

Notice that, unlike SK_2 and SK_3, SK_4 is not bounded in the unit interval.

Of the four measures of skewness introduced so far only SK_2 can be applied to any distribution, independently of whether its moments exist. In fact, SK_1, SK_3 and SK_4 require the existence of moments up to order three, one and two, respectively.

2.4 Robust measures of kurtosis

Moors (1988) proposed to interpret kurtosis of X_t as dispersion around $\mu \pm \sigma$:

$$KR_2' = \frac{(Q_{3/8} - Q_{1/8}) + (Q_{7/8} - Q_{5/8})}{Q_{6/8} - Q_{2/8}} \tag{7}$$

KR_2' relies on "octiles", that is $Q_{i/8} = F^{-1}(i/8)$ for $i = 1, \ldots, 7$. The two terms in the numerator are large (small) if relatively little (much) probability mass is concentrated around $Q_{2/8}$ and $Q_{6/8}$, corresponding with large (small) dispersion in the neighborhoods of $\mu \pm \sigma$. Table 1(b) shows that for the standard Normal distribution $KR_2' = 1.233$, therefore the centered measure is: $KR_2 = KR_2' - 1.233$.

Hogg (1972, 1974) introduced a measure of kurtosis based on averages of lower and upper quantiles, defined as $L_\alpha = (1/\alpha) \int_0^\alpha Q(u) \mathrm{d}u$ and $U_\alpha = (1/\alpha) \int_{1-\alpha}^1 Q(u) \mathrm{d}u$, respectively. In practice, we rely on:

$$KR_3' = \frac{U_{0.05} - L_{0.05}}{U_{0.50} - L_{0.50}} \tag{8}$$

for the standard Normal distribution $KR_2' = 2.585$, therefore: $KR_3 = KR_3' - 2.585$.

A third measure of kurtosis is due to Crow and Siddiqui (1967):

$$KR_4' = \frac{Q_{0.975} + Q_{0.025}}{Q_{0.75} - Q_{0.25}} \tag{9}$$

as shown in Table 1(b) for the standard Normal distribution $KR_4' = 2.906$ and hence $KR_4 = KR_4' - 2.906$.

Since KR'_2 and KR'_4 depend only on quantiles, they are defined for any distribution, independently on whether they have finite moments or not.

2.5 Robust measures based on L- and trimmed L-moments

Hosking (1990) proposed to describe the shape of distributions relying on linear functions of expectations of order statistics called "L-moments", where the "L" emphasizes the linearity of these functions. Being linear functions of the data L-moments are more robust to outliers than conventional moments that raise the difference from the mean to the third or fourth power. Moreover, since L-moments exist for any random variable with finite mean, they uniquely characterize a class of distributions that is wider than the set of distributions for which conventional moments can be applied. In fact, a distribution with finite mean is uniquely characterized by its L-moments, even when conventional moments do not exist.

The first four L-moments of a random variable are:[5]

$$\lambda_1 = E\left(X_{1:1}\right) = \int_0^1 Q(u)\mathrm{d}u \tag{10}$$

$$\lambda_2 = \frac{1}{2}E\left(X_{2:2} - X_{1:2}\right) = \int_0^1 Q(u)(2u-1)\mathrm{d}u \tag{11}$$

$$\lambda_3 = \frac{1}{3}E\left(X_{3:3} - 2X_{2:3} + X_{1:3}\right) = \int_0^1 Q(u)(6u^2 - 6u + 1)\mathrm{d}u \tag{12}$$

$$\lambda_4 = \frac{1}{4}E\left(X_{4:4} - 3X_{3:4} + 3X_{2:4} + X_{1:4}\right) = \int_0^1 Q(u)(20u^3 - 30u^2 + 12u - 1)\mathrm{d}u \tag{13}$$

while λ_1 and λ_2 can be regarded as measures location and scale, population L-skewness and L-kurtosis are defined as ratios of L-moments:

$$SK_5 = \frac{\lambda_3}{\lambda_2} \tag{14}$$

$$KR'_5 = \frac{\lambda_4}{\lambda_2} \tag{15}$$

for a standard Normal variate $KR'_5 = 0.123$, hence the excess L-kurtosis is $KR_5 = KR'_5 -$

---

[5]Notice that the notation relies on the so-called "conceptual sample", that is different from the sample used for inference (see e.g. Elamir and Seheult, 2003, p. 300). For instance, the population mean is defined in terms of a conceptual sample of size one, therefore $\mu \equiv E(X) = E(X_{1:1})$. Similarly, the population variance is defined in terms of conceptual sample of size two. More generally, the r-th moment is defined in terms of a conceptual sample of size r.

0.123. Hosking (1992) showed that L-skewness and L-kurtosis identify deviations from Normality better than conventional moment-based measures. L-moment ratios are bounded: $|\lambda_r/\lambda_2| < 1$ for $r \geq 3$. The boundedness of SK_5 and KR'_5 makes their interpretation easier than conventional skewness and kurtosis that can take arbitrarily large values.

Elamir and Seheult (2003) generalized L-moments to distributions that do not posses finite mean: a case in point are Cauchy random variables. Since this approach relies on trimmed samples, they are more robust to outliers than L-moments. Trimmed L-moments (TL-moments), denoted as $\lambda_r^{(s,k)}$, are linear functions of expectations of order statistics that exclude the s smallest and/or the k largest order statistics of the sample. In the empirical analysis we set $s = k = 1$, therefore we eliminate the largest and smallest observations in the sample. The TL-skewness and TL-kurtosis are:

$$SK_6 = \frac{\lambda_3^{(1,1)}}{\lambda_2^{(1,1)}} \tag{16}$$

$$KR'_6 = \frac{\lambda_4^{(1,1)}}{\lambda_2^{(1,1)}} \tag{17}$$

one drawback of trimming is that, contrarily to SK_5 and KR'_5, SK_6 and KR'_6 are not bounded in the $(-1, 1)$ interval. Using results in Hosking (2007) we can show that $|SK_6| \leq 1.11$ and $|KR'_6| \leq 1.25$. For a standard Normal variate $KR'_6 = 0.062$, hence the excess TL-kurtosis is $KR_6 = KR'_6 - 0.062$. Thomas (2009) studied tests of symmetry based on SK_2, SK_3, SK_5 and SK_6 and showed that TL-skewness yields the most reliable confidence intervals and tests. Further details, as well as a discussion of the estimation of L- and TL-moments are presented in Section C of the Appendix.

3   Simulations

In this section we present simulation results for six alternative distributions both for iid samples and for serially correlated data generating processes. Next, we discuss four additional Monte Carlo designs used to assess the robustness different measures of skewness and kurtosis to outliers. While the design based on iid samples might look simplistic, it used mimic the situation in which regression residuals are analyzed. Nevertheless, we stress that the focus of

the paper is on unconditional skewness and kurtosis, therefore we do not directly analyze the effects of filtering the series with a model. See Bai and Ng (2005) and Kilian and Demiroglu (2000) for details on this issue. In Section 3.3 we assess the impact of serial correlation on the set of estimators considered in this paper.

3.1 Monte Carlo design

We design our Monte Carlo experiments following Kim and White (2004) and Bonato (2011). We generate 1000 random samples of size $T = 50, 250, 500, 1000, 2500, 5000$ for each design of the simulation analysis. Samples of these dimensions are chosen to mimic the size of time series used in macroeconometrics or financial econometrics. For instance, a sample of size $T = 50$ is representative of applications that rely on yearly macroeconomic data, while $T = 5000$ is typical when dealing with daily or weekly financial returns. In Section 3.2 we focus on three symmetric and three asymmetric distributions. The three symmetric distributions are: the standard Normal and the Student's t distribution with 10 and 5 degrees of freedom. The asymmetric ones are: the Log-Normal distribution with parameters $\mu = 1$ and $\sigma = 0.4$, the unit Exponential distribution and the χ^2 distribution with 3 degrees of freedom.

3.1.1 Outliers

The sample distributions of macroeconomic and financial time series are often asymmetric and heavy-tailed. We thus consider three additional simulation designs aimed at capturing the impact of outliers on the properties of skewness and excess kurtosis estimators.

In a first set of experiments we contaminate draws from a standard Normal distribution with a single outlier. Specifically, we use two designs in which we contaminate the Normal distribution with a single, extremely low, value. Following Kim and White (2004), these outliers are calibrated on the empirical distribution of daily log-returns on the S&P500 index to mimic the stock market crashes of 1987 and 2008. Details in Appendix A.1.

Table 1: Values of skewness and excess kurtosis for various distributions

(a) Skewness measures (SK_i)

	N(0,1)	t_{10}	t_5	Log N	Exp(1)	$\chi^2_{(3)}$	N-mix
SK_1	0.000	0.000	0.000	1.322	2.000	1.633	-2.271
SK_2	0.000	0.000	0.000	0.134	0.262	0.203	0.000
SK_3	0.000	0.000	0.000	0.247	0.443	0.356	-0.009
SK_4	0.000	0.000	0.000	0.185	0.307	0.259	-0.007
SK_5	0.000	0.000	0.000	0.194	0.333	0.272	-0.013
SK_6	0.000	0.000	0.000	0.121	0.222	0.176	-0.001

(b) Excess kurtosis measures (KR_i)

	N(0,1)		t_{10}	t_5	Log N	Exp(1)	$\chi^2_{(3)}$	N-mix
	KR'_j	KR_j						
KR_1	3.000	0.000	1.000	6.000	3.260	6.000	4.000	52.575
KR_2	1.233	0.000	0.044	0.094	0.044	0.073	0.042	0.001
KR_3	2.585	0.000	0.200	0.460	0.187	0.279	0.168	0.088
KR_4	2.906	0.000	0.278	0.632	0.268	0.429	0.248	0.009
KR_5	0.123	0.000	0.032	0.071	0.030	0.044	0.027	0.014
KR_6	0.062	0.000	0.013	0.029	0.013	0.021	0.012	0.000

Notes: distributions are denoted as follows: "N(0,1)" = standard Normal, "t_ν" = Student's t with ν degrees of freedom (dof), "Log N" = Log Normal(1,0.4), "Exp(1)" = Exponential, $\chi^2_{(\nu)}$ = χ^2 with 3 dof, "N-mix" = Normal mixture. In panel (b) the column headed KR'_j reports the value of the kurtosis for the standard Normal distribution, while the remaining columns show the kurtosis measure for a given distribution minus the corresponding theoretical value for the standard Normal distribution.

Lastly, we rely on a Normal mixture distribution with parameters calibrated using the empirical distribution of daily log-returns on the S&P500 index. This approach also follows that of Kim and White (2004, p. 68) and leads to a mixture of the following two distributions: $N(0,1)$ and $N(-7,10)$ with the first occurring with probability $p = .9988$. Notice that the second regime captures stock market crashes. As shown in Table 1 the theoretical values of the conventional skewness and excess kurtosis of this distribution are -2.27 and 52.58.

3.1.2 Evaluation of estimators

We evaluate estimators of skewness and excess kurtosis in terms of bias and root mean squared error. The bias is defined as $B_j = E\left(\theta_j - \theta_j^{(0)}\right)$ where $\theta_j^{(0)}$ represents the true value of the j-th measure of skewness or excess kurtosis as reported in Table 1. The Monte Carlo estimates of the bias are: $\hat{B}_j = \bar{\hat{\theta}}_j - \theta_j^{(0)}$ where $\bar{\hat{\theta}}_j = M^{-1}\sum_{m=1}^M \hat{\theta}_{j,m}$ and $\hat{\theta}_{j,m}$ is the sample estimate obtained in the m-th simulation. We stress that the magnitude of the bias of different estimators of skewness and kurtosis cannot be directly compared, because some of them are constrained to lie in a finite interval (e.g. L-skewness and kurtosis), while others

can take on any real value (e.g. standard measures of skewness and kurtosis). Nevertheless, providing an estimate of the bias and assessing whether and how quickly vanishes as the sample size grows provides useful information to evaluate different estimators.

Given the previous warning, we also assess the root mean squared error (RMSE), defined as: $RMSE_j = \sqrt{E\left(\theta_j - \theta_0\right)^2}$. To ease the comparison, we rely on RMSE ratios. These are defined as: $rRMSE_j = RMSE_j/RMSE_1$, where $RMSE_1$ identifies the RMSE of the conventional moment-based skewness or excess kurtosis. Therefore, a RMSE ratio lower than one signals that the j-th estimator is more tightly clustered around its population value than the conventional measures (SK_1 or KR_1).

3.2 Results for iid samples

In this section we evaluate the performance of measures of skewness and excess kurtosis for iid random samples generated from the the standard Normal, Student's t, Log-Normal, Exponential and χ^2.

Table 2 shows that independently of the distribution or sample size, the RMSE of the conventional measure of skewness, SK_1, is always the highest. The bias of SK_1 increases as we move away from the Normal distribution, especially is small samples and for asymmetric distributions. This fact is clearly visible in Figure 1 where, for each data generating process and sample size, we summarize the sampling distribution SK_1 with a boxplot.[6]

For the Log-Normal, Exponential and χ^2 distributions there is a substantial downward bias, that is still visible even in samples of 250 observations. Numbers in Figure 1 represent observations that are greater than $Q_{0.75} + IQR$ or lower than $Q_{0.25} + IQR$, where IQR is the interquartile range. These numbers show that the dispersion of SK_1 increases when we move from the Normal to the Student's t distribution and also when considering asymmetric distributions.

As for the robust measures, Table 2 conveys two main messages. First the bias of SK_2-SK_6 is in most cases negligible; second, the TL-moment skewness, SK_6, is always associated with the lowest RMSE.

---

[6]Boxplots with the sampling distributions of all estimators are shown in Appendix 5.

Table 2: Bias and RMSE ratio of skewness estimators

| T	50	250	500	1000	2500	5000
			(a) N(0,1)			
SK_1	0.003 (1.000)	0.007 (1.000)	0.002 (1.000)	0.001 (1.000)	0.001 (1.000)	0.000 (1.000)
SK_2	-0.009 (0.575)	0.000 (0.505)	0.001 (0.583)	-0.001 (0.518)	-0.001 (0.564)	0.001 (0.569)
SK_3	-0.002 (0.403)	-0.000 (0.366)	0.002 (0.402)	-0.000 (0.364)	-0.000 (0.394)	0.001 (0.392)
SK_4	-0.001 (0.321)	<u>-0.000</u> (0.292)	0.001 (0.321)	-0.000 (0.290)	-0.000 (0.314)	0.001 (0.313)
SK_5	<u>0.001</u> (0.197)	0.001 (0.179)	0.001 (0.181)	<u>0.000</u> (0.174)	0.000 (0.176)	<u>0.000</u> (0.174)
SK_6	-0.001 (<u>0.179</u>)	-0.000 (<u>0.157</u>)	<u>0.000</u> (<u>0.165</u>)	-0.000 (<u>0.150</u>)	0.000 (<u>0.157</u>)	0.000 (<u>0.157</u>)
			(b) t_{10}			
SK_1	-0.007 (1.000)	-0.007 (1.000)	-0.001 (1.000)	-0.008 (1.000)	0.001 (1.000)	-0.000 (1.000)
SK_2	0.000 (0.394)	0.003 (0.315)	-0.003 (0.283)	-0.001 (0.270)	-0.001 (0.285)	-0.001 (0.282)
SK_3	-0.001 (0.271)	0.001 (0.219)	-0.003 (0.198)	-0.002 (0.196)	-0.000 (0.201)	-0.001 (0.196)
SK_4	-0.001 (0.211)	0.001 (0.169)	-0.002 (0.154)	-0.001 (0.151)	-0.000 (0.156)	-0.000 (0.152)
SK_5	<u>-0.000</u> (0.153)	-0.001 (0.116)	<u>-0.000</u> (0.106)	-0.001 (0.108)	<u>-0.000</u> (0.108)	<u>-0.000</u> (0.103)
SK_6	0.000 (<u>0.116</u>)	<u>0.000</u> (<u>0.095</u>)	-0.001 (<u>0.083</u>)	<u>-0.001</u> (<u>0.086</u>)	-0.000 (<u>0.088</u>)	-0.000 (<u>0.083</u>)
			(c) t_5			
SK_1	0.012 (1.000)	0.004 (1.000)	<u>0.000</u> (1.000)	-0.012 (1.000)	-0.013 (1.000)	-0.023 (1.000)
SK_2	0.009 (0.235)	-0.005 (0.097)	-0.003 (0.090)	0.002 (0.086)	0.001 (0.075)	<u>0.000</u> (0.039)
SK_3	0.005 (0.171)	-0.004 (0.072)	-0.002 (0.065)	-0.000 (0.064)	-0.000 (0.054)	0.000 (0.029)
SK_4	0.003 (0.128)	-0.003 (0.053)	-0.002 (0.047)	-0.000 (0.047)	<u>-0.000</u> (0.039)	0.000 (0.021)
SK_5	<u>0.001</u> (0.115)	<u>-0.001</u> (0.051)	-0.001 (0.044)	-0.001 (0.043)	-0.001 (0.037)	0.000 (0.019)
SK_6	0.001 (<u>0.075</u>)	-0.001 (<u>0.031</u>)	-0.001 (<u>0.029</u>)	0.000 (<u>0.028</u>)	-0.000 (<u>0.024</u>)	0.000 (<u>0.013</u>)
			(d) Log-N(1,0.4)			
SK_1	-0.235 (1.000)	-0.072 (1.000)	-0.046 (1.000)	-0.025 (1.000)	-0.009 (1.000)	-0.006 (1.000)
SK_2	-0.012 (0.310)	-0.003 (0.238)	-0.004 (0.221)	-0.002 (0.196)	<u>0.000</u> (0.185)	-0.001 (0.179)
SK_3	-0.007 (0.214)	-0.004 (0.160)	-0.003 (0.150)	-0.002 (0.137)	-0.000 (0.124)	-0.000 (0.122)
SK_4	-0.006 (0.156)	-0.003 (0.116)	-0.002 (0.108)	-0.001 (0.098)	-0.000 (0.089)	-0.000 (0.088)
SK_5	<u>-0.001</u> (0.118)	<u>-0.001</u> (0.084)	<u>-0.001</u> (0.079)	<u>-0.001</u> (0.073)	-0.000 (0.064)	0.000 (0.066)
SK_6	-0.019 (<u>0.102</u>)	-0.003 (<u>0.072</u>)	-0.001 (<u>0.067</u>)	-0.001 (<u>0.060</u>)	-0.000 (<u>0.054</u>)	<u>-0.000</u> (<u>0.055</u>)
			(e) Exp(1)			
SK_1	-0.369 (1.000)	-0.079 (1.000)	-0.063 (1.000)	-0.026 (1.000)	-0.002 (1.000)	-0.007 (1.000)
SK_2	-0.011 (0.256)	-0.004 (0.191)	-0.003 (0.188)	0.001 (0.165)	-0.000 (0.155)	-0.000 (0.161)
SK_3	-0.013 (0.176)	-0.004 (0.126)	-0.003 (0.127)	0.001 (0.111)	-0.001 (0.106)	-0.000 (0.111)
SK_4	<u>-0.005</u> (0.117)	-0.002 (0.082)	-0.001 (0.081)	0.001 (0.072)	-0.000 (0.067)	<u>-0.000</u> (0.072)
SK_5	-0.005 (0.098)	<u>-0.000</u> (0.068)	<u>-0.001</u> (0.067)	0.000 (0.058)	<u>-0.000</u> (0.057)	-0.000 (0.061)
SK_6	-0.031 (<u>0.094</u>)	-0.003 (<u>0.060</u>)	-0.002 (<u>0.060</u>)	<u>-0.000</u> (<u>0.051</u>)	-0.000 (<u>0.049</u>)	-0.000 (<u>0.052</u>)
			(f) $\chi^2_{(3)}$			
SK_1	-0.291 (1.000)	-0.060 (1.000)	-0.040 (1.000)	-0.023 (1.000)	-0.001 (1.000)	-0.007 (1.000)
SK_2	-0.019 (0.299)	-0.006 (0.260)	-0.003 (0.222)	0.001 (0.228)	0.001 (0.209)	-0.001 (0.200)
SK_3	-0.018 (0.207)	-0.004 (0.171)	-0.003 (0.150)	0.000 (0.155)	0.001 (0.141)	-0.001 (0.137)
SK_4	-0.012 (0.145)	-0.003 (0.118)	-0.002 (0.102)	0.000 (0.106)	0.000 (0.095)	-0.000 (0.094)
SK_5	<u>-0.007</u> (0.111)	<u>-0.001</u> (0.087)	<u>-0.002</u> (0.077)	<u>-0.000</u> (0.079)	0.000 (0.071)	-0.000 (0.069)
SK_6	-0.026 (<u>0.104</u>)	-0.004 (<u>0.081</u>)	-0.002 (<u>0.067</u>)	-0.000 (<u>0.068</u>)	<u>0.000</u> (<u>0.062</u>)	<u>-0.000</u> (<u>0.061</u>)

Notes: for each sample size T and distribution the table shows the bias (B_j) and the Root Mean Squared Error (RMSE) ratio of the estimator (in brackets). RMSE ratios are defined as: $rRMSE_j = RMSE_j/RMSE_1$, where $RMSE_1$ identifies the RMSE of the method-of-moments estimator. Therefore, a $rRMSE_j$ lower than one signals that the j-th estimator is more tightly clustered around its population value than the conventional measure. Estimates of B_j and $RMSE_j$ are based on 1000 Monte Carlo simulations. See Section 3 for details.

Table 3: Bias and RMSE ratio of excess kurtosis estimators

| T	50	250	500	1000	2500	5000
			(a) N(0,1)			
KR_1	-0.119 (1.000)	-0.032 (1.000)	-0.019 (1.000)	-0.002 (1.000)	-0.000 (1.000)	-0.002 (1.000)
KR_2	0.019 (0.425)	0.006 (0.382)	0.001 (0.388)	0.002 (0.353)	0.000 (0.377)	0.000 (0.368)
KR_3	-0.074 (0.407)	-0.017 (0.368)	-0.009 (0.365)	-0.003 (0.343)	-0.007 (0.363)	-0.009 (0.378)
KR_4	0.035 (0.990)	-0.014 (0.825)	0.002 (0.830)	0.013 (0.786)	0.001 (0.813)	0.003 (0.795)
KR_5	0.000 (0.075)	-0.001 (0.064)	-0.000 (0.063)	0.000 (0.059)	0.000 (0.061)	0.000 (0.061)
KR_6	-0.148 (0.251)	-0.010 (0.062)	-0.003 (0.056)	-0.001 (0.048)	-0.000 (0.052)	0.000 (0.050)
			(b) t_{10}			
KR_1	-0.534 (1.000)	-0.101 (1.000)	-0.057 (1.000)	-0.027 (1.000)	-0.025 (1.000)	0.007 (1.000)
KR_2	0.036 (0.227)	0.005 (0.105)	0.003 (0.090)	0.002 (0.100)	0.000 (0.095)	0.001 (0.082)
KR_3	-0.102 (0.249)	-0.017 (0.119)	-0.010 (0.102)	-0.010 (0.114)	-0.016 (0.120)	-0.022 (0.110)
KR_4	0.110 (0.617)	-0.013 (0.258)	0.006 (0.220)	0.005 (0.245)	0.003 (0.241)	0.002 (0.199)
KR_5	-0.001 (0.042)	-0.000 (0.019)	-0.000 (0.016)	0.000 (0.018)	0.000 (0.018)	0.000 (0.015)
KR_6	-0.155 (0.126)	-0.012 (0.018)	-0.003 (0.013)	-0.001 (0.014)	-0.000 (0.013)	0.000 (0.011)
			(c) t_5			
KR_1	-4.436 (1.000)	-2.249 (1.000)	-1.944 (1.000)	-1.892 (1.000)	-1.536 (1.000)	-0.614 (1.000)
KR_2	0.034 (0.058)	0.009 (0.015)	-0.000 (0.013)	0.003 (0.011)	0.000 (0.008)	-0.001 (0.002)
KR_3	-0.133 (0.077)	-0.019 (0.023)	-0.011 (0.020)	-0.018 (0.016)	-0.040 (0.015)	-0.056 (0.005)
KR_4	0.247 (0.213)	-0.003 (0.044)	0.009 (0.039)	0.005 (0.032)	-0.000 (0.024)	0.001 (0.006)
KR_5	-0.003 (0.012)	0.000 (0.003)	0.000 (0.003)	0.000 (0.002)	-0.000 (0.002)	-0.000 (0.000)
KR_6	-0.164 (0.033)	-0.013 (0.003)	-0.004 (0.002)	-0.001 (0.002)	-0.000 (0.001)	-0.000 (0.000)
			(d) Log-N(1,0.4)			
KR_1	-1.614 (1.000)	-0.583 (1.000)	-0.388 (1.000)	-0.202 (1.000)	-0.057 (1.000)	-0.062 (1.000)
KR_2	0.031 (0.093)	0.013 (0.053)	0.004 (0.042)	0.002 (0.033)	-0.001 (0.027)	0.000 (0.029)
KR_3	-0.088 (0.110)	-0.013 (0.064)	-0.011 (0.054)	-0.012 (0.041)	-0.016 (0.038)	-0.020 (0.040)
KR_4	-0.241 (0.212)	-0.067 (0.131)	-0.027 (0.108)	-0.007 (0.087)	-0.008 (0.073)	-0.001 (0.073)
KR_5	0.001 (0.019)	0.000 (0.010)	-0.000 (0.008)	-0.000 (0.007)	-0.000 (0.006)	-0.000 (0.006)
KR_6	-0.155 (0.051)	-0.011 (0.008)	-0.003 (0.006)	-0.001 (0.005)	-0.000 (0.004)	-0.000 (0.004)
			(e) Exp(1)			
KR_1	-2.889 (1.000)	-0.756 (1.000)	-0.603 (1.000)	-0.264 (1.000)	-0.012 (1.000)	-0.067 (1.000)
KR_2	0.024 (0.069)	0.008 (0.037)	0.005 (0.032)	0.000 (0.027)	-0.000 (0.025)	0.001 (0.026)
KR_3	-0.114 (0.094)	-0.012 (0.050)	-0.013 (0.043)	-0.010 (0.036)	-0.015 (0.035)	-0.021 (0.040)
KR_4	-0.418 (0.184)	-0.098 (0.105)	-0.041 (0.094)	-0.008 (0.075)	-0.006 (0.075)	0.003 (0.078)
KR_5	-0.003 (0.015)	0.000 (0.008)	-0.000 (0.007)	0.000 (0.005)	0.000 (0.005)	0.000 (0.005)
KR_6	-0.160 (0.036)	-0.012 (0.006)	-0.003 (0.005)	-0.001 (0.004)	-0.000 (0.004)	0.000 (0.004)
			(f) $\chi^2_{(3)}$			
KR_1	-1.917 (1.000)	-0.531 (1.000)	-0.299 (1.000)	-0.213 (1.000)	-0.003 (1.000)	-0.048 (1.000)
KR_2	0.044 (0.091)	0.001 (0.057)	-0.002 (0.047)	0.000 (0.045)	0.002 (0.036)	0.001 (0.030)
KR_3	-0.108 (0.109)	-0.008 (0.070)	-0.015 (0.058)	-0.010 (0.057)	-0.012 (0.046)	-0.018 (0.043)
KR_4	-0.328 (0.207)	-0.093 (0.147)	-0.058 (0.119)	-0.008 (0.119)	-0.008 (0.096)	-0.002 (0.081)
KR_5	-0.002 (0.018)	0.001 (0.011)	-0.001 (0.009)	-0.000 (0.009)	0.000 (0.007)	-0.000 (0.006)
KR_6	-0.153 (0.046)	-0.011 (0.009)	-0.004 (0.007)	-0.001 (0.006)	-0.000 (0.005)	-0.000 (0.004)

Notes: see notes to 2.

Figure 1: Sample distribution of SK_1
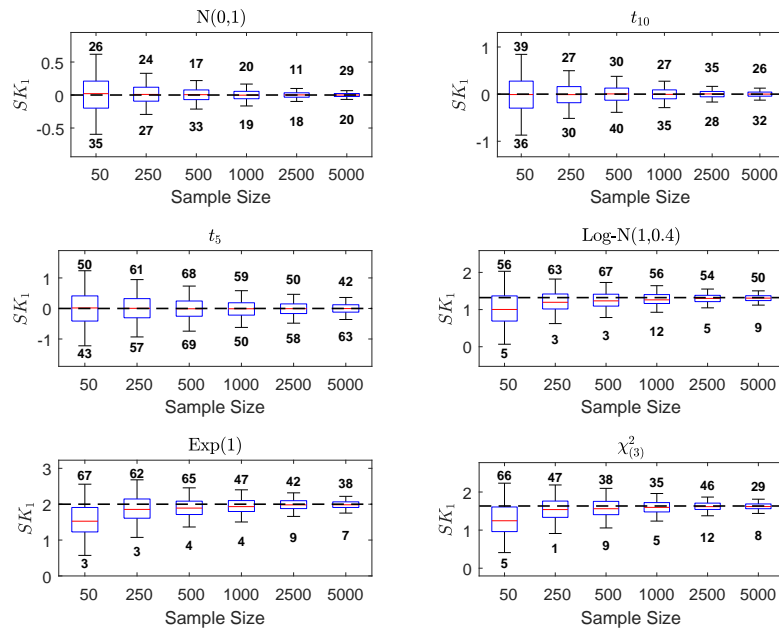
Notes: for each distribution and sample size (reported on the x-axis) the figure represents the sampling distribution of the estimator with a boxplot. The line in the middle of the box is the median, the size of the box is proportional to the interquartile range (IQR), namely the distance between the 75-th ($Q_{0.75}$) and 25-th percentile $Q_{0.25}$. The bottom (top) external line, known as whisker, is drawn in correspondence of $Q_{0.25} - IQR$ ($Q_{0.75} + IQR$). Numbers outside the two whiskers represent the number of observations lower (greater) than the bottom (top) whisker.

Figure 2: Sample distribution of KR_1
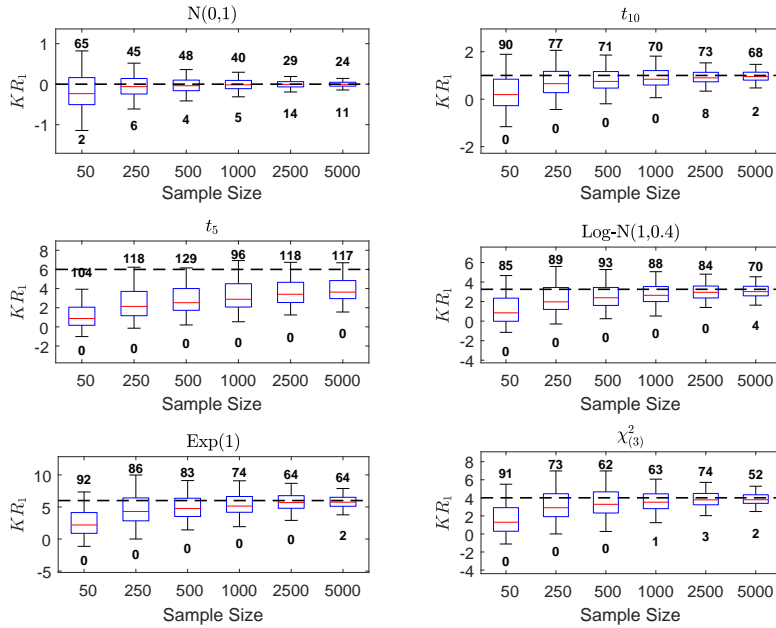
Notes: see notes to Figure 1.

However, while trimming reduces the RMSE, it also leads to a bias in the case of samples of size 50 drawn from asymmetric distributions. As it can been seen, this does not happen for L-skewness, SK_5. To sum up, measures based on L- or TL-moments represent the best choice both in terms of bias and RMSE ratio.

Table 3 and Figure 2 highlight that the conventional measure of excess kurtosis KR_1 is severely biased, especially in small samples. For the Student's t_5 distribution the downward bias can be visually assessed in Figure 2, even in a sample of 5000 observations. Moreover, as in the case of the method of moments skewness measure, KR_1 is always associated with the highest RMSE. In samples of size up to 250 the best options in terms of bias are KR_2 and KR_5: in both cases the bias vanishes as we move from 50 to 250 observations. On the contrary, when the underlying distribution is asymmetric, the remaining measures remain downward biased even in samples of 250 observations. However, when considering also the RMSE, the kurtosis estimator based on L-moments, KR_5, is the best option. Much like in the case of skewness, trimming leads to a deterioration of the performance of the estimator based on L-moments; in fact, irrespective of the distribution, KR_6 displays a severe bias in samples of size 50.

3.3 Serial correlation

Up to now we have analysed iid samples, but macroeconomic and financial time series are typically serially correlated and often highly persistent. For this reason, we now briefly assess the effects of serial correlation on the measures of skewness and excess kurtosis under investigation. We do so by focusing on two Autoregressive (AR) models of order one with different error distributions and degrees of persistence.[7] For the sake of brevity, we present only three distributions: the standard Normal, Student's t_5 and the Log-Normal distribution. In Tables 4 and 5 we consider an AR(1) model with autoregressive parameter equal to 0.9. Results for an AR(1) model with autoregressive parameter 0.5 appears in Appendix D.

Moving from the iid simulation design to the AR(1) data generating process, we see that the bias of all skewness and excess kurtosis measures increases in absolute terms. As for skewness, TL-moments outperform other estimators for what concerns the RMSE ratio. This result carries over, at least in part to excess kurtosis. As the sample size increase, for non-Normal distributions, the bias of KR_6 tends to vanish more quickly than that of its competitors.

All in all, also in the presence of serial correlation the bias of skewness and kurtosis measures based L- or TL-moments vanishes very quickly as the sample size grow. Moreover, these estimators are often associated with the lowest RMSE errors. Table D1 and D2 of the Appendix present results for AR(1) models with autoregressive parameter equal to 0.5. These Tables confirm the findings of this Section, therefore the main results are not affected by the degree of persistence of the AR(1) model. The degree of persistence is positively associated with the size of the bias, but does not alter the ranking of estimators.

---

[7]We simulate two zero mean AR(1) processes with autoregressive parameter $\rho = 0.5$ and $\rho = 0.9$. If we measure persistence in terms of "speed of mean reversion", we can rely on the half-life to quantify the number of periods needed for the effects of a unit shock to halve. For an AR(1) model, the half-life is: $HL = \log(0.5)/\log(|\rho|)$. Then, we get $HL = 1$ when $\rho = 0.5$ and $HL = 6.6$ when $\rho = 0.9$.

Table 4: Bias and RMSE ratio of skewness estimators - AR(1) model ($\rho = 0.9$)

T	50	250	500	1000	2500	5000
			(a) N(0,1)			
SK_1	0.017 (1.000)	-0.008 (1.000)	-0.009 (1.000)	-0.002 (1.000)	0.002 (1.000)	0.002 (1.000)
SK_2	0.005 (0.522)	0.001 (0.363)	0.002 (0.336)	<u>-0.000</u> (0.310)	0.001 (0.294)	0.001 (0.311)
SK_3	0.006 (0.451)	0.000 (0.335)	0.000 (0.303)	-0.001 (0.283)	0.001 (0.278)	0.001 (0.280)
SK_4	0.005 (0.362)	<u>0.000</u> (0.267)	<u>0.000</u> (0.242)	-0.001 (0.226)	0.001 (0.222)	0.001 (0.223)
SK_5	<u>0.005</u> (0.236)	-0.001 (0.194)	-0.001 (0.183)	-0.001 (0.177)	0.000 (0.173)	0.001 (0.170)
SK_6	0.005 (<u>0.225</u>)	-0.001 (<u>0.159</u>)	0.001 (<u>0.145</u>)	-0.001 (<u>0.137</u>)	<u>0.000</u> (<u>0.135</u>)	<u>0.000</u> (<u>0.132</u>)
			(c) t_5			
SK_1	0.002 (1.000)	-0.006 (1.000)	<u>0.000</u> (1.000)	0.006 (1.000)	-0.003 (1.000)	0.003 (1.000)
SK_2	-0.011 (0.468)	0.002 (0.310)	-0.001 (0.270)	-0.002 (0.220)	0.002 (0.188)	0.001 (0.182)
SK_3	-0.003 (0.406)	0.001 (0.280)	-0.001 (0.248)	-0.000 (0.203)	0.001 (0.176)	0.001 (0.173)
SK_4	-0.003 (0.325)	0.001 (0.220)	-0.000 (0.195)	<u>-0.000</u> (0.159)	0.001 (0.138)	0.001 (0.136)
SK_5	<u>-0.000</u> (0.220)	-0.001 (0.171)	-0.000 (0.157)	0.000 (0.132)	<u>-0.000</u> (0.116)	0.000 (0.117)
SK_6	-0.002 (<u>0.194</u>)	<u>-0.000</u> (<u>0.134</u>)	-0.000 (<u>0.122</u>)	-0.000 (<u>0.099</u>)	0.000 (<u>0.086</u>)	<u>0.000</u> (<u>0.084</u>)
			(d) Log-N(1,0.4)			
SK_1	-1.112 (1.000)	-0.976 (1.000)	-0.940 (1.000)	-0.940 (1.000)	-0.925 (1.000)	-0.921 (1.000)
SK_2	-0.104 (0.209)	-0.095 (0.146)	-0.086 (0.121)	-0.090 (0.110)	-0.090 (0.104)	-0.091 (0.102)
SK_3	-0.189 (0.226)	-0.170 (0.194)	-0.162 (0.183)	-0.166 (0.181)	-0.166 (0.181)	-0.166 (0.181)
SK_4	-0.139 (0.174)	-0.124 (0.145)	-0.117 (0.134)	-0.120 (0.132)	-0.120 (0.132)	-0.121 (0.132)
SK_5	-0.147 (0.150)	-0.132 (0.141)	-0.128 (0.139)	-0.131 (0.140)	-0.130 (0.141)	-0.130 (0.141)
SK_6	<u>-0.089</u> (<u>0.110</u>)	<u>-0.082</u> (<u>0.094</u>)	<u>-0.079</u> (<u>0.089</u>)	<u>-0.082</u> (<u>0.089</u>)	<u>-0.081</u> (<u>0.089</u>)	<u>-0.081</u> (<u>0.089</u>)

Notes: see notes to 2.

Table 5: Bias and RMSE ratio of excess kurtosis estimators - AR(1) model ($\rho = 0.9$)

T	50	250	500	1000	2500	5000
			(a) N(0,1)			
KR_1	-0.461 (1.000)	-0.167 (1.000)	-0.117 (1.000)	-0.054 (1.000)	-0.018 (1.000)	-0.007 (1.000)
KR_2	<u>0.003</u> (0.456)	<u>-0.006</u> (0.311)	0.004 (0.274)	0.001 (0.261)	0.000 (0.214)	0.002 (0.245)
KR_3	-0.258 (0.513)	-0.072 (0.412)	-0.046 (0.389)	-0.020 (0.376)	-0.012 (0.336)	-0.010 (0.351)
KR_4	-0.252 (0.945)	-0.049 (0.804)	-0.015 (0.756)	0.002 (0.726)	-0.001 (0.614)	-0.001 (0.660)
KR_5	-0.033 (<u>0.093</u>)	-0.010 (0.071)	-0.006 (0.066)	-0.003 (0.063)	-0.001 (0.055)	-0.000 (0.058)
KR_6	-0.153 (0.212)	-0.012 (<u>0.055</u>)	<u>-0.004</u> (<u>0.046</u>)	<u>-0.001</u> (<u>0.044</u>)	<u>-0.000</u> (<u>0.037</u>)	<u>0.000</u> (<u>0.041</u>)
			(c) t_5			
KR_1	-6.335 (1.000)	-5.923 (1.000)	-5.744 (1.000)	-5.620 (1.000)	-5.528 (1.000)	-5.496 (1.000)
KR_2	<u>-0.072</u> (0.060)	-0.074 (0.032)	-0.073 (0.024)	-0.071 (0.020)	-0.074 (0.016)	-0.076 (0.015)
KR_3	-0.688 (0.119)	-0.466 (0.087)	-0.417 (0.078)	-0.395 (0.073)	-0.389 (0.071)	-0.383 (0.070)
KR_4	-0.810 (0.176)	-0.562 (0.126)	-0.523 (0.110)	-0.515 (0.101)	-0.520 (0.097)	-0.513 (0.095)
KR_5	-0.099 (<u>0.019</u>)	-0.070 (0.013)	-0.063 (0.012)	-0.060 (0.011)	-0.058 (0.011)	-0.057 (0.010)
KR_6	-0.180 (0.029)	<u>-0.036</u> (<u>0.008</u>)	<u>-0.027</u> (<u>0.006</u>)	<u>-0.024</u> (<u>0.005</u>)	<u>-0.024</u> (<u>0.005</u>)	<u>-0.024</u> (<u>0.004</u>)
			(d) Log-N(1,0.4)			
KR_1	-3.670 (1.000)	-3.234 (1.000)	-3.085 (1.000)	-3.034 (1.000)	-2.958 (1.000)	-2.939 (1.000)
KR_2	<u>-0.042</u> (0.086)	-0.039 (0.048)	-0.030 (0.037)	-0.034 (0.028)	-0.037 (0.020)	-0.038 (0.018)
KR_3	-0.432 (0.138)	-0.238 (0.096)	-0.195 (0.079)	-0.180 (0.070)	-0.168 (0.062)	-0.169 (0.060)
KR_4	-0.555 (0.222)	-0.302 (0.158)	-0.240 (0.123)	-0.233 (0.105)	-0.228 (0.090)	-0.230 (0.086)
KR_5	-0.060 (<u>0.023</u>)	-0.036 (0.015)	-0.030 (0.013)	-0.028 (0.011)	-0.026 (0.010)	-0.025 (0.009)
KR_6	-0.164 (0.045)	<u>-0.023</u> (<u>0.010</u>)	<u>-0.014</u> (<u>0.007</u>)	<u>-0.012</u> (<u>0.006</u>)	<u>-0.011</u> (<u>0.005</u>)	<u>-0.011</u> (<u>0.004</u>)

Notes: see notes to 2.

3.4 Robustness to outliers

Joint inspection of Figures 3 and 5 highlights that the effects of outlying observations on the sampling distributions of conventional skewness and kurtosis measures depends heavily on the data generating process used in the simulations.

3.4.1 Normal mixture distribution

Figure 3 shows that the dispersion of SK_1 increases with the size of the sample. This happens because the mixture distribution assigns a very low probability to the "market crash regime", that therefore materializes more often in larger samples. Independently of the sample size, Table 6 shows that SK_1 is upward biased with the size of the bias monotonically decreasing as T increases. On the contrary, the robust measures of skewness are largely unaffected by the presence of very large negative observations generated in the second regime of the mixture. TL-skewness is always associated with the lowest RMSE ratio. See Figure 4.

Figure 5 shows that KR_1 is severely downward biased even in samples of size 5000. The Normal mixture data generating process proves challenging also for KR_3 that always displays a negative bias. On the contrary, KR_2, KR_4 and KR_5 are centered around their true values. TL-kurtosis, KR_6, is severely biased in samples of size 50, but then quickly converges to its true value. See Figure 6. Table 7 illustrates that either KR_5 or KR_6 yield the lowest RMSE ratio. Their bias also quickly vanishes as the sample size grows.

3.4.2 Contaminated Normal distribution

Table 6(c,d) shows that for sample of 50 observations contaminating the standard Normal draws with a single outlier has noticeable effects on all measures, except SK_2 and SK_6. While SK_1 remains downward biased even in samples of size $T = 5000$, SK_3, SK_4 and SK_5 converge to their true values as T grows. In this case, TL-skewness represents the best option in terms RMSE.

The effects of a single outlier on the conventional measure of kurtosis KR_1 are noteworthy: KR_1 becomes completely unreliable, even for $T = 5000$. See Figure 5. As expected, a comparison of Table 7(c) and 7(d) illustrates that this detrimental impact is positively
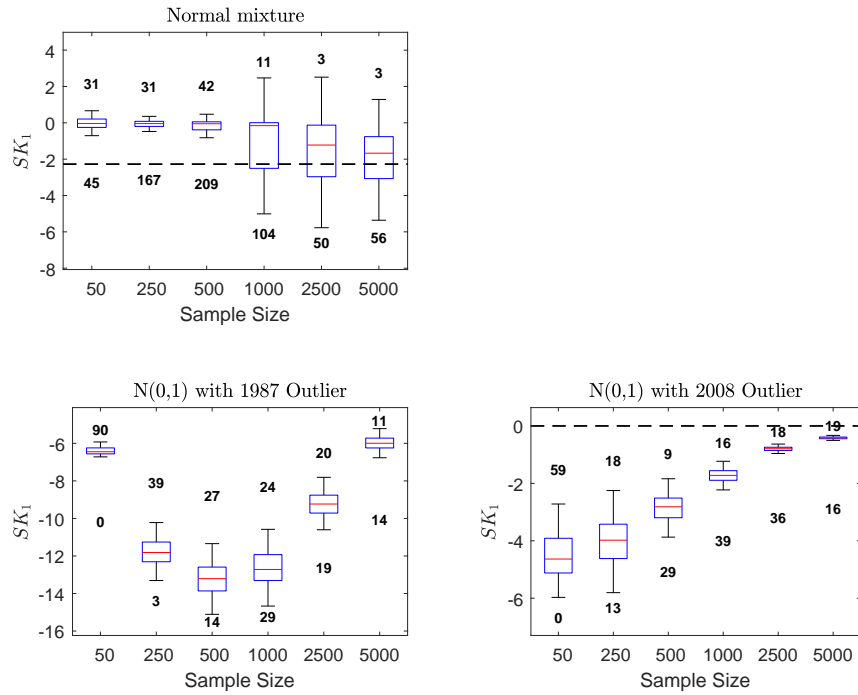
Table 6: Bias and RMSE ratio of skewness estimators – outliers

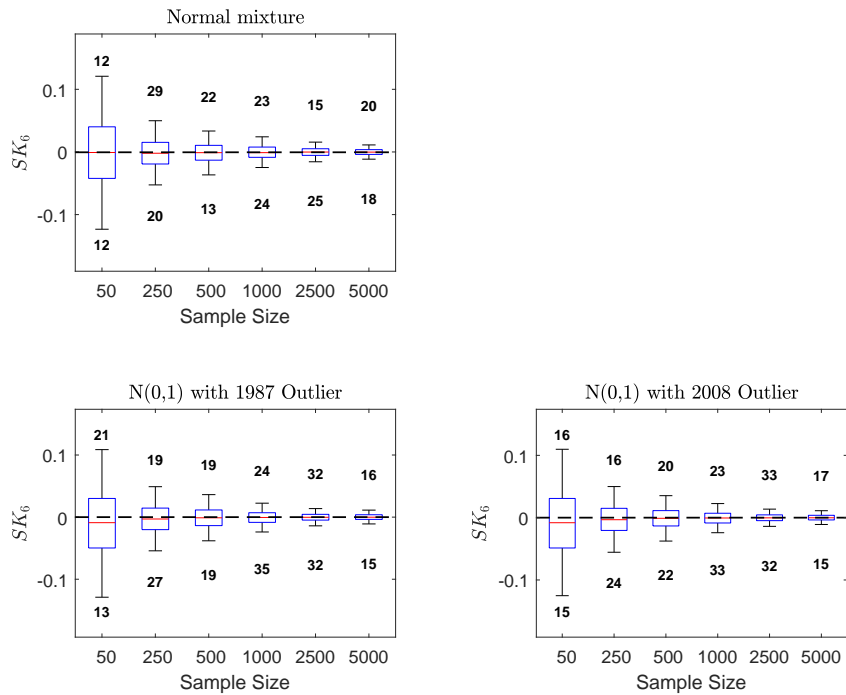| T	50	250	500	1000	2500	5000
			(a) Normal mixture			
SK_1	2.147 (1.00)	1.532 (1.00)	1.402 (1.00)	0.816 (1.00)	0.447 (1.00)	0.199 (1.00)
SK_2	-0.002 (0.077)	-0.001 (0.032)	-0.003 (0.021)	-0.003 (0.016)	0.001 (0.012)	0.001 (0.011)
SK_3	0.004 (0.057)	-0.001 (0.025)	0.000 (0.016)	-0.001 (0.013)	0.001 (0.010)	0.001 (0.008)
SK_4	0.005 (0.043)	0.001 (0.019)	0.001 (0.012)	-0.000 (0.009)	0.001 (0.007)	0.001 (0.006)
SK_5	0.003 (0.036)	-0.001 (0.018)	0.002 (0.011)	0.000 (0.009)	0.000 (0.007)	0.000 (0.006)
SK_6	-0.001 (0.025)	-0.001 (0.010)	-0.000 (0.006)	0.000 (0.005)	0.000 (0.004)	0.000 (0.003)
			(b) N(0,1) with 1987 Outlier			
SK_1	-6.319 (1.00)	-11.722 (1.00)	-13.200 (1.00)	-12.654 (1.00)	-9.246 (1.00)	-5.984 (1.00)
SK_2	-0.018 (0.029)	-0.002 (0.007)	-0.000 (0.005)	-0.001 (0.003)	-0.001 (0.003)	0.001 (0.003)
SK_3	-0.449 (0.073)	-0.143 (0.013)	-0.076 (0.007)	-0.040 (0.004)	-0.017 (0.003)	-0.008 (0.003)
SK_4	-0.135 (0.022)	-0.056 (0.005)	-0.035 (0.003)	-0.023 (0.002)	-0.011 (0.002)	-0.005 (0.002)
SK_5	-0.522 (0.083)	-0.188 (0.016)	-0.104 (0.008)	-0.055 (0.004)	-0.023 (0.003)	-0.011 (0.002)
SK_6	-0.009 (0.009)	-0.003 (0.002)	-0.001 (0.001)	-0.001 (0.001)	-0.000 (0.001)	0.000 (0.001)
			(c) N(0,1) with 2008 Outlier			
SK_1	-4.422 (1.00)	-4.008 (1.00)	-2.862 (1.00)	-1.733 (1.00)	-0.798 (1.00)	-0.415 (1.00)
SK_2	-0.019 (0.041)	-0.002 (0.020)	-0.000 (0.022)	-0.001 (0.024)	-0.001 (0.033)	0.001 (0.046)
SK_3	-0.227 (0.056)	-0.056 (0.019)	-0.027 (0.018)	-0.015 (0.019)	-0.006 (0.024)	-0.002 (0.032)
SK_4	-0.112 (0.027)	-0.037 (0.013)	-0.019 (0.012)	-0.011 (0.014)	-0.005 (0.019)	-0.002 (0.025)
SK_5	-0.277 (0.063)	-0.074 (0.019)	-0.038 (0.015)	-0.020 (0.014)	-0.008 (0.014)	-0.004 (0.017)
SK_6	-0.009 (0.013)	-0.003 (0.006)	-0.001 (0.006)	-0.001 (0.007)	-0.000 (0.009)	0.000 (0.013)

Notes: see notes to Table 2.

associated with the size of the outlier. With the exception of KR_2, all robust measures are biased when $T = 50$. TL-kurtosis is once again the best option both in terms of speed with which the bias vanishes asymptotically and RMSE. See also Figure 6.

Figure 3: Sample distribution of SK_1
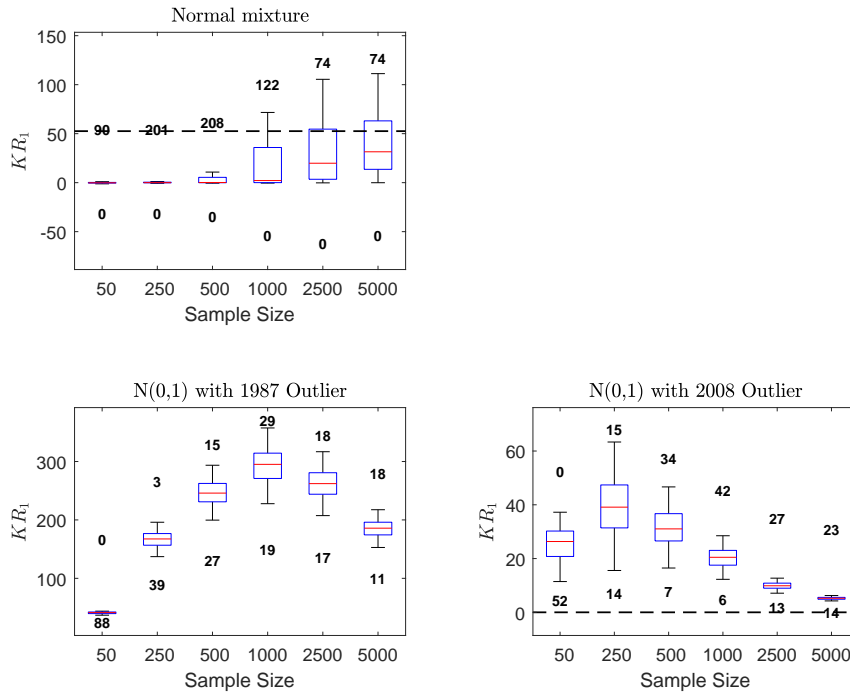
Figure 4: Sample distribution of SK_6

Table 7: Bias and RMSE ratio of excess kurtosis estimators – outliers

| T	50	250	500	1000	2500	5000
			(a) Normal mixture			
KR_1	-51.774 (1.00)	-43.619 (1.00)	-37.998 (1.00)	-26.370 (1.00)	-15.624 (1.00)	-7.918 (1.00)
KR_2	0.026 (0.005)	0.007 (0.002)	<u>0.000</u> (0.002)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
KR_3	-0.088 (0.008)	-0.013 (0.005)	-0.016 (0.003)	-0.036 (0.002)	-0.087 (0.002)	-0.128 (0.003)
KR_4	0.326 (0.039)	-0.008 (0.005)	0.014 (0.003)	-0.002 (0.002)	0.003 (0.002)	0.003 (0.001)
KR_5	<u>-0.003</u> (<u>0.001</u>)	<u>-0.001</u> (0.001)	-0.001 (0.000)	<u>-0.000</u> (0.000)	0.000 (0.000)	0.000 (0.000)
KR_6	-0.148 (0.003)	-0.010 (<u>0.000</u>)	-0.003 (<u>0.000</u>)	-0.001 (<u>0.000</u>)	<u>0.000</u> (<u>0.000</u>)	<u>0.000</u> (<u>0.000</u>)
			(b) N(0,1) with 1987 Outlier			
KR_1	40.32 (1.00)	165.59 (1.00)	245.97 (1.00)	293.92 (1.00)	262.98 (1.00)	185.50 (1.00)
KR_2	<u>0.042</u> (0.007)	0.010 (0.001)	0.003 (0.000)	0.003 (0.000)	0.001 (0.000)	0.000 (0.000)
KR_3	3.186 (0.080)	1.057 (0.006)	0.496 (0.002)	0.134 (0.000)	-0.100 (0.000)	-0.146 (0.001)
KR_4	23.36 (0.597)	0.028 (0.002)	0.023 (0.001)	0.023 (0.000)	0.005 (0.000)	0.005 (0.000)
KR_5	0.453 (0.011)	0.163 (0.001)	0.091 (0.000)	0.048 (0.000)	0.020 (0.000)	0.010 (0.000)
KR_6	-0.145 (<u>0.004</u>)	<u>-0.009</u> (<u>0.000</u>)	<u>-0.002</u> (<u>0.000</u>)	<u>-0.000</u> (<u>0.000</u>)	<u>0.000</u> (<u>0.000</u>)	<u>0.000</u> (<u>0.000</u>)
			(c) N(0,1) with 2008 Outlier			
KR_1	24.977 (1.00)	39.550 (1.00)	31.900 (1.00)	20.611 (1.00)	9.964 (1.00)	5.231 (1.00)
KR_2	<u>0.038</u> (0.010)	0.009 (0.003)	0.003 (0.002)	0.002 (0.003)	0.000 (0.004)	0.000 (0.005)
KR_3	1.492 (0.059)	0.381 (0.010)	0.175 (0.006)	0.051 (0.003)	-0.033 (0.005)	-0.051 (0.011)
KR_4	7.772 (0.313)	0.028 (0.006)	0.023 (0.005)	0.022 (0.006)	0.004 (0.008)	0.005 (0.010)
KR_5	0.236 (0.009)	0.062 (0.002)	0.033 (0.001)	0.017 (0.001)	0.007 (0.001)	0.004 (0.001)
KR_6	-0.146 (<u>0.006</u>)	<u>-0.009</u> (<u>0.000</u>)	<u>-0.002</u> (<u>0.000</u>)	<u>-0.000</u> (<u>0.000</u>)	<u>0.000</u> (<u>0.000</u>)	<u>0.000</u> (<u>0.001</u>)

Notes: see notes to Table 2.

Figure 5: Sample distribution of KR_1

Notes: See notes to Figure 1.

Figure 6: Sample distribution of KR_6

Notes: See notes to Figure 1.

22
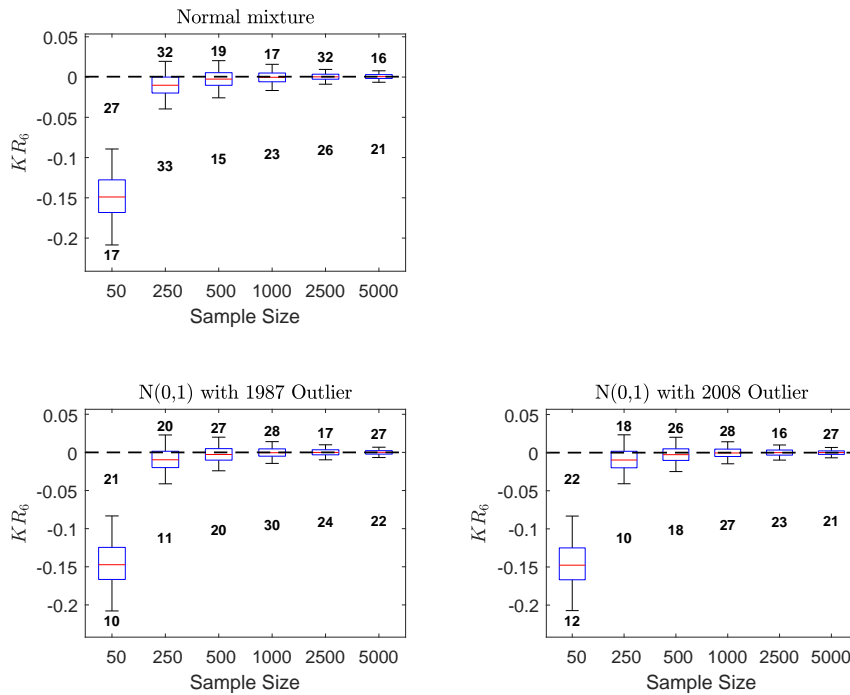
4 Empirical application: the FRED-MD

In this section we estimate the skewness and excess kurtosis of 128 US macroeconomic and financial time series in the FRED-MD (see McCracken and Ng, 2016, for details). This is a large, monthly frequency, dataset that focuses on the US economy and is updated in real-time through the FRED database. We rely on data for the period April 1959 - December 2018. Because of differences in data availability, the sample size ranges from 320 to 717 and is on average equal to 709. Transformations aimed at making the series stationary are applied when necessary.[8]

To gauge the maximum moment that exists for each series in the FRED-MD, we have estimated their maximal moment exponents (MME) relying (Hill, 1975). The MME is defined as $\alpha_\delta = \sup_q E|x_t|^q < \infty$ and can be estimated as follows (Hall, 1982):

$$\hat{\alpha}_\delta = \left[\left(\frac{1}{m-1} \sum_{i=1}^{m-1} \log X_{i:T} \right) - \log X_{m:T} \right]^{-1} \qquad \text{with } \delta = \frac{m}{T} \quad m > 0 \qquad (18)$$

we set $\delta = 0.1$, so that we focus on the upper 10% of the empirical distribution of the series in the FRED-MD.[9]

Earlier empirical analyses showed that estimates of the MME for exchange rates and stock returns often lie between 2 and 4 (Chen et al., 2000; De Lima, 1997; Mittnik and Rachev, 1993; Loretan and Phillips, 1994); therefore, while the second moment seems to be finite, the third and fourth moments often are not. Our results are largely in line with these findings. In fact, Table 8 highlight that some of the series in the FRED-MD do not possess all moments required to compute conventional or even robust measures of skewness and kurtosis. The conventional skewness measure, that requires existence of moments up to order three, can be computed for 104 series out of 128. For the conventional kurtosis this number drops to 46, meaning that only about 36% of the series possess moments up to the fourth or higher. Therefore, these estimates suggest that there are 24 series (19% of the

---

[8] We apply the transformations suggested by McCracken and Ng (2016), but contrarily to them we do not remove outliers. In fact, we want to asses their impact of skewness and kurtosis. See the online Appendix to McCracken and Ng (2016) for details.

[9] It is well known that the Hill estimator is biased in small samples and is very sensitive to the choice of δ (see e.g. Kearns and Pagan, 1997; Huisman et al., 2001). For these reasons, we present some robustness checks in Section D.4 of the Appendix.

total) that do not have finite third moment. Lastly, the estimated MME of the series in the FRED-MD exceeds 8 only in 15 out of 128 cases (11.7% of the total).

All in all, there is evidence that at least some of the monthly US macroeconomic time series in the FRED–MD database might not posses all moments that are necessary to conduct normality tests. For instance the Jarque-Bera procedure requires existence of the eight moment, while a test of symmetry based on SK_1 requires existence of the sixth moment (Bai and Ng, 2001). Problems associated with moment condition failure also arise when testing for nonlinearity (see e.g. Chen et al., 2000; De Lima, 1997).

For each estimator, Figure 7 presents a scatterplot that compares the skewness estimated on the whole sample and after deleting the largest observation in absolute value. Each point in Figure 7 refers to the estimate of the skewness for a different series in the FRED-MD. We can see that SK_2 and SK_6 display the lowest dispersion and a high degree of correlation, thus suggesting that they are the most robust measures. The same exercise for excess kurtosis is presented in Figure 8. The correlation between the estimated kurtosis for the whole sample and the "leave-one-out" estimate is highest for KR_6, that also displays the lowest dispersion. Moreover, Figures 7 and 8 highlight that the conventional measures of skewness and excess kurtosis display the highest dispersion, thus confirming their low resistance to outlying observations.

4.1   Selected series

Following McCracken and Ng (2016) the 128 series in the FRED-MD can be assigned to eight groups: "Output & Income" (16 series), "Labor Market" (31 series), "Housing" (10 series), "Consumption, Orders & Inventories" (10 series), "Money & Credit" (14 series), "Interest & Exchange Rates" (22 series), "Prices" (20 series), "Stock Market" (5 series). We select one representative series in each group and in Table 8 we report the estimate of its MME, SK_1, SK_5, KR_1 and KR_5. Column 3 shows that, with exception of Housing Starts, none of the series has finite fourth moment. Notice that the lowest MME, 1.58, is associated with the Fed Fund rate that also displays the largest excess kurtosis as measured both by KR_1 and KR_5. Comparing columns 6 and 7, we can see that KR_1 is particularly sensitive to
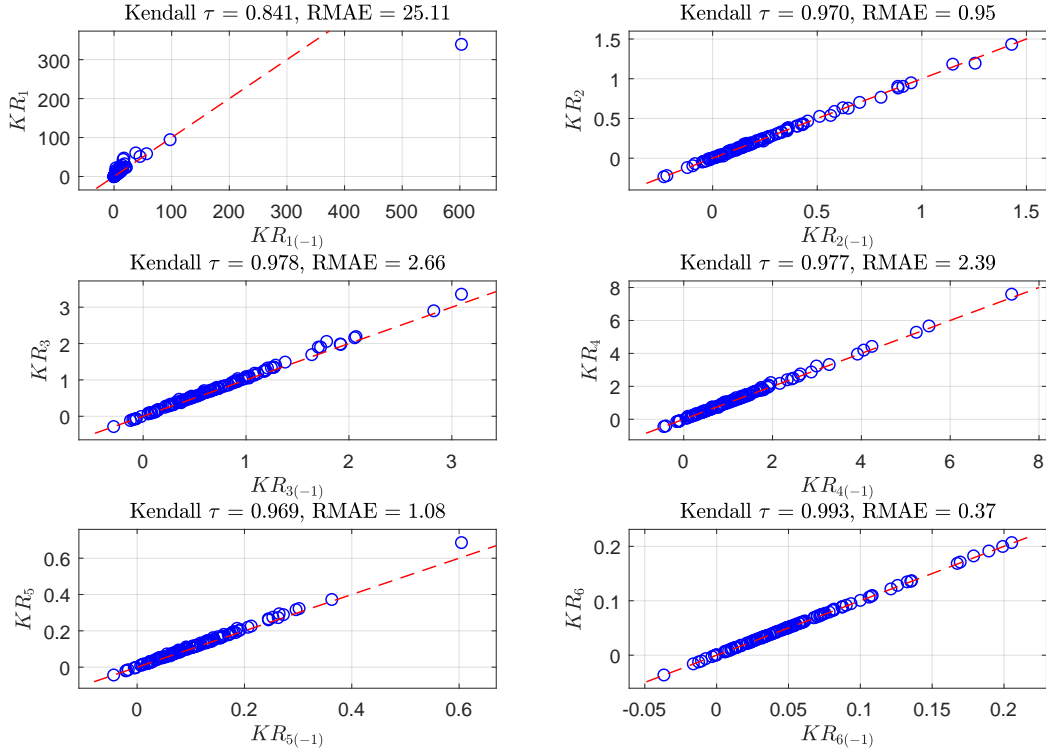
Figure 7: Robustness of skewness measures for the FRED-MD
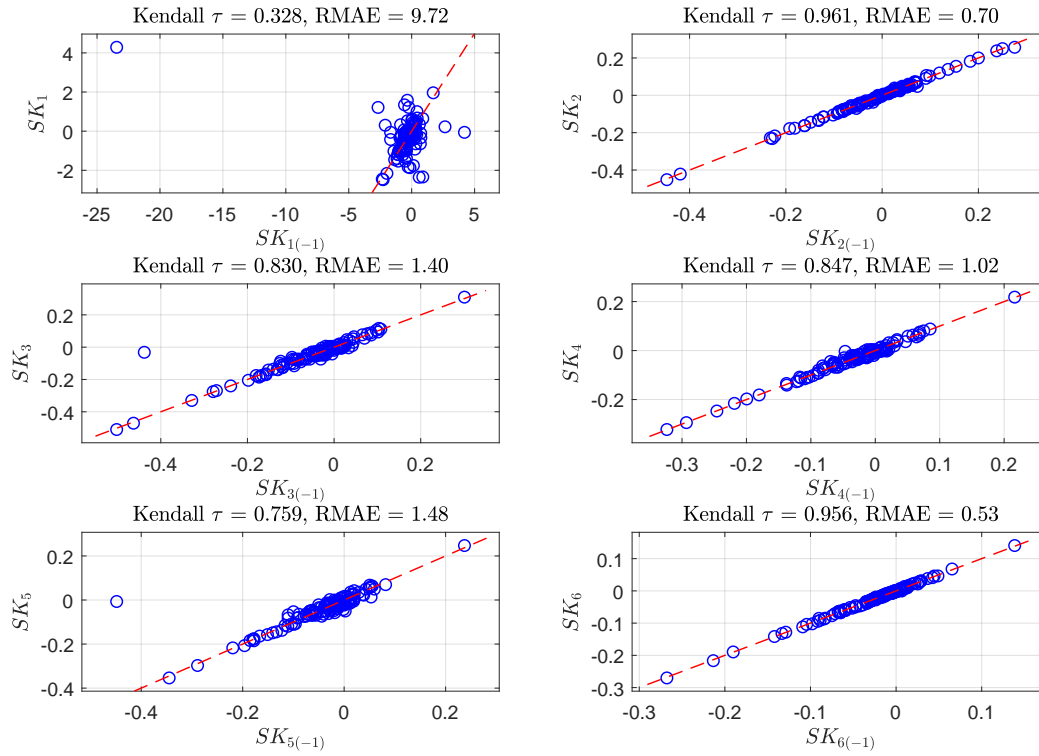
Notes: the figure shows the estimated skewness for each series in the FRED-MD (circles) with (x-axis) and without (y-axis) the largest observation (in absolute value). The dashed line is the 45° line. Each figure also shows the Kemdall's τ correlation coefficient and the root mean absolute error (RMAE).

outliers. For instance, KR_1 for log-returns on the S&P500 index is 4 if computed on the whole sample, while it halves when omitting the largest observation is absolute value. This sensitivity to outliers is also a feature of SK_1. On the contrary, SK_5 and KR_5 remain largely unaffected when one outlying observation is deleted.

Figure 8: Robustness of excess kurtosis measures for the FRED-MD

Notes: the figure shows the estimated excess kurtosis for each series in the FRED-MD (circles) with (x-axis) and without (y-axis) the largest observation (in absolute value). The dashed line is the 45° line. See notes to 7.

Table 8: Robustness of skewness and kurtosis for selected macroeconomic series in the FRED-MD.

Series	Trans.	$\hat{\alpha}_{10\%}$	SK_1	$SK_{1(-1)}$	KR_1	$KR_{1(-1)}$	SK_5	$SK_{5(-1)}$	KR_5	$KR_{5(-1)}$
IP	$\Delta \log x_t$	2.898	-0.344	-0.940	7.319	4.555	-0.051	-0.068	0.102	0.091
Un. Rate	Δx_t	2.153	0.512	0.373	1.786	1.177	0.064	0.055	0.070	0.064
House	$\log x_t$	94.533	-1.012	-1.021	1.185	1.192	-0.176	-0.178	0.063	0.062
Real PCE	$\Delta \log x_t$	3.075	-0.205	-0.006	2.875	2.078	-0.011	-0.001	0.088	0.081
M2	$\Delta^2 \log x_t$	2.353	-0.864	-0.424	8.805	6.446	-0.043	-0.028	0.145	0.135
Fed Funds	Δx_t	1.588	-2.340	0.884	47.247	17.065	-0.043	-0.001	0.323	0.302
CPI	$\Delta^2 \log x_t$	2.165	0.015	-0.402	4.843	2.814	-0.014	-0.028	0.104	0.095
S&P 500	$\Delta \log x_t$	3.263	-1.031	-0.713	4.184	2.095	-0.108	-0.096	0.081	0.072

Notes: column 1 reads as follows: IP = Industrial Production; Un. Rate = Unemployment Rate; House = Housing Starts; Real PCE = Real personal consumption expenditures; M2 = M2 Money Stock; Fed Funds = Effective Federal Funds Rate; CPI = CPI All Items; S&P500 = S&P 500 Index. Column 2 shows the transformation applied to make the series stationary (see McCracken and Ng, 2016); column 3 reports point estimates of maximal moment exponent; columns 4-11 show skewness and excess kurtosis measures computed on the whole sample (SK_i, KR_i) or leaving out the largest observation (in absolute value) of the series ($SK_{i,(-1)}$, $KR_{i,(-1)}$).

5 Conclusions

In this paper we have assessed the relative merits of different estimators of skewness and excess kurtosis. Our Monte Carlo analysis highlights that standard estimators based on the method-of-moments are outperformed by more robust and easily implementable alternatives. This result is not confined to data generating processes contaminated with outlying observations, but carries over to non-Gaussian samples and even to normally distributed random variables featuring serial correlation and persistence.

In the empirical application to the FRED-MD we provide further evidence that method-of-moments estimators of skewness and kurtosis are unreliable for macroeconomic time series. Moreover, we also show that several time series do not possess enough finite moments to perform standard test of Normality, symmetry or nonlinearity (Ramsey and Rothman, 1996). Since these tests are often used not only for model-building purposes, but also to investigate substantive questions in economics and finance (Bontemps and Meddahi, 2005; Kilian and Demiroglu, 2000) our results have far reaching implications for empirical analyses of macroeconomic and financial time series. These are not limited to estimation of skewness and kurtosis, but possibly extend to testing normality and symmetry (Bai and Ng, 2001).

On a more positive note, we show that robust estimators are available and perform much better – both in term of bias reduction and precision – than standard measures of skewness and kurtosis. In particular, our main contribution is to introduce L- and TL-moments in the time series econometrics literature illustrating that these estimators outperform those based on the method-of-moments, as well as alternative robust estimators based on quantiles. Moreover, L- and TL-moments can be used to characterize a very large class of statistical distributions because they are less demanding in terms the moments. This opens the doors for their use in macroeconomic and financial applications, where there is often an issue of moment failure (Chen et al., 2000; De Lima, 1997; Mittnik and Rachev, 1993; Loretan and Phillips, 1994). Two potential applications of L-moments, that we leave for future investigation, are tests of normality and symmetry in the time series domain and the estimation of realized volatility.

References

Bai, J. and Ng, S. (2001). A consistent test for conditional symmetry in time series models. *Journal of Econometrics*, 103(1-2):225–258.

Bai, J. and Ng, S. (2005). Tests for skewness, kurtosis, and normality for time series data. *Journal of Business & Economic Statistics*, 23(1):49–60.

Bonato, M. (2011). Robust estimation of skewness and kurtosis in distributions with infinite higher moments. *Finance Research Letters*, 8(2):77–87.

Bontemps, C. and Meddahi, N. (2005). Testing normality: a GMM approach. *Journal of Econometrics*, 124(1):149–186.

Bowley, A. L. (1920). *Elements of Statistics*. P.S. King & Sons.

Brys, G., Hubert, M., and Struyf, A. (2003). A comparison of some new measures of skewness. In Dutter, R., Filzmoser, P., Gather, U., and Rousseeuw, P. J., editors, *Developments in Robust Statistics*, pages 98–113. Heidelberg: Springer-Verlag.

Brys, G., Hubert, M., and Struyf, A. (2004). A robust measure of skewness. *Journal of Computational and Graphical Statistics*, 13(4):996–1017.

Cecchetti, S. G., Lam, P.-S., and Mark, N. C. (1990). Mean reversion in equilibrium asset prices. *American Economic Review*, 80(3):398–418.

Chen, Y.-T., Chou, R. Y., and Kuan, C.-M. (2000). Testing time reversibility without moment restrictions. *Journal of Econometrics*, 95(1):199–218.

Crow, E. L. and Siddiqui, M. (1967). Robust estimation of location. *Journal of the American Statistical Association*, 62(318):353–389.

Darolles, S., Gouriéroux, C., and Jasiak, J. (2009). L-performance with an application to hedge funds. *Journal of Empirical Finance*, 16(4):671–685.

David, H. A. and Nagaraja, H. N. (2017). *Order Statistics*. Wiley.

De Lima, P. J. (1997). On the robustness of nonlinearity tests to moment condition failure. *Journal of Econometrics*, 76(1-2):251–280.

DeLong, J. B. and Summers, L. H. (1986). Are business cycles symmetrical. In Gordon, R. J., editor, *The American Business Cycle: Continuity and Change*, pages 166–178. University of Chicago Press.

Elamir, E. A. and Seheult, A. H. (2003). Trimmed L-moments. *Computational Statistics & Data Analysis*, 43(3):299–314.

Groeneveld, R. A. and Meeden, G. (1984). Measuring skewness and kurtosis. *Journal of the Royal Statistical Society. Series D*, 33(4):391–399.

Hall, P. (1982). On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society. Series B*, 44(1):37–42.

Harri, A. and Coble, K. H. (2011). Normality testing: two new tests using L-moments. *Journal of Applied Statistics*, 38(7):1369–1379.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174.

Hinkley, D. V. (1975). On power transformations to symmetry. *Biometrika*, 62(1):101–111.

Hogg, R. V. (1972). More light on the kurtosis and related statistics. *Journal of the American Statistical Association*, 67(338):422–424.

Hogg, R. V. (1974). Adaptive robust procedures: A partial review and some suggestions for future applications and theory. *Journal of the American Statistical Association*, 69(348):909–923.

Hosking, J. (1992). Moments or L moments? an example comparing two measures of distributional shape. *The American Statistician*, 46(3):186–189.

Hosking, J. (2007). Some theory and practical uses of trimmed L-moments. *Journal of Statistical Planning and Inference*, 137(9):3024–3039.

Hosking, J. R. (1990). L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B*, pages 105–124.

Hubert, M. and Debruyne, M. (2009). Breakdown value. *Wiley International Reviews – Computational Statistics*, 1:296–302.

Huisman, R., Koedijk, K. G., Kool, C. J. M., and Palm, F. (2001). Tail-index estimates in small samples. *Journal of Business & Economic Statistics*, 19(2):208–216.

Karvanen, J. (2006). Estimation of quantile mixtures via l-moments and trimmed l-moments. *Computational Statistics & Data Analysis*, 51(2):947–959.

Kearns, P. and Pagan, A. (1997). Estimating the density tail index for financial time series. *Review of Economics and Statistics*, 79(2):171–175.

Kilian, L. and Demiroglu, U. (2000). Residual-based tests for normality in autoregressions: asymptotic theory and simulation evidence. *Journal of Business & Economic Statistics*, 18(1):40–50.

Kim, T.-H. and White, H. (2004). On more robust estimation of skewness and kurtosis. *Finance Research Letters*, 1(1):56–73.

Loretan, M. and Phillips, P. C. B. (1994). Testing the covariance stationarity of heavy-tailed time series: an overview of the theory with applications to several financial datasets. *Journal of Empirical Finance*, 1(2):211–248.

McCracken, M. W. and Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4):574–589.

McCracken, M. W. and Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4):574–589.

Meijer, E. (2000). An asymmetric distribution with zero skewness. Technical report, Available at SSRN: http://dx.doi.org/10.2139/ssrn.2531847.

Mittnik, S. and Rachev, S. T. (1993). Modeling asset returns with alternative stable distributions. *Econometric Reviews*, 12(3):261–330.

Moors, J. J. A. (1988). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society. Series D*, 37(1):25–32.

Ord, J. K. (1968). The discrete Student's t distribution. *The Annals of Mathematical Statistics*, 39(5):1513–1516.

Ramsey, J. B. and Rothman, P. (1996). Time irreversibility and business cycle asymmetry. *Journal of Money, Credit and Banking*, 28(1):1–21.

Scott, R. C. and Horvath, P. A. (1980). On the direction of preference for moments of higher order than the variance. *Journal of Finance*, 35(4):915–919.

Thomas, G. (2009). Assessing symmetry using quantiles and L-moments. *Communications in Statistics- Simulation and Computation*, 38(2):335–354.

von Hippel, P. (2010). Skewness. In Lovric, M., editor, *International Encyclopedia of Statistical Science*. Springer.

Westfall, P. H. (2014). Kurtosis as peakedness, 1905–2014. RIP. *The American Statistician*, 68(3):191–195.

Appendix to "Robust measures of skewness and kurtosis for financial and macroeconomic time series"

A Monte Carlo design: further details

A.1 Standard Normal distributions contaminated with a single outlier

To produce the two Monte Carlo designs in which we contaminate the standard Normal distribution with a single outlier, we follow Kim and White (2004) and rely on the sample distribution of daily log-returns on the S&P500 index. We downloaded adjusted close prices from Yahoo! Finance. We consider two sample periods: 04/01/1982-06/05/1999 and 29/06/2001-30/11/2018. In both cases the size of the sample is 4384. For each period, we identify the minimum return that corresponds to -22.90% in 19/10/1987 (first sample) and to -9.47% in 15/10/2008 (second sample). Outliers are injected in the simulated distributions at time τ_i for $i = 1, 2$, that corresponds to the location of the outlier in the sample divided by the total number of observations. This approach leads to $\tau_1 = 0.30$ and $\tau_2 = 0.48$. The size of outliers, m_i for $i = 1, 2$ is determined as: $m_i = x_{\lfloor \tau_i T \rfloor}/Q_{0.25}$. This yields: $m_1 = -22.90/-0.45 = 50.89$ and $m_2 = -9.47/-0.49 = 19.33$. The last step is as follows: in each of the two experiment, we simulate of a sample of size T from the standard Normal distribution, we calculate $Q_{0.25}$ and then we replace $x_{\lfloor \tau_i T \rfloor}$ with the outlier $m_i \times Q_{0.25}$.

B Estimation of sample skewness & kurtosis

B.1 Conventional skewness and kurtosis

Given a sample of size T, $\{x\}_{t=1}^{T}$, SK_1 and KR_1 are estimated as follows:

$$\widehat{SK}_1 = \frac{1}{T} \sum_{t=1}^{T} \left(\frac{x_t - \hat{\mu}}{\hat{\sigma}} \right)^3$$

$$\widehat{KR}_1 = \frac{1}{T} \sum_{t=1}^{T} \left(\frac{x_t - \hat{\mu}}{\hat{\sigma}} \right)^4 - 3 \tag{1}$$

where $\hat{\mu} = T^{-1} \sum_{t=1}^{T} x_t$ and $\hat{\sigma} = \sqrt{T^{-1} \sum_{t=1}^{T} (x_t - \hat{\mu})^2}$.

C L-moments & TL moments: further details

C.1 Estimation of L-moment

L-moments can be generally defined as:

$$\lambda_r = \frac{1}{r} \sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j} E\left(X_{r-j:r} \right) \quad \text{for } r = 1, 2, \dots \tag{2}$$

The expectation of an order statistics, $E\left(X_{j:r} \right)$, can be written as (David and Nagaraja, 2017):

$$E\left(X_{j:r} \right) = \frac{r!}{(j-1)!(r-j)!} \int_{-inf}^{inf} x F^{j-1}(x) \left[1 - F(x)\right]^{r-j} f(x)\mathrm{d}x$$

$$= \frac{r!}{(j-1)!(r-j)!} \int_0^1 u^{j-1} \left[1 - u\right]^{r-j} Q(u)\mathrm{d}u \tag{3}$$

Substituting (3) in (2) and rearranging, we get:

$$\lambda_r = \int_0^1 P_{r-1}^*(u) Q(u)\mathrm{d}u \tag{4}$$

where $P_r^*(u) = \sum_{j=0}^{r} p_{r,j}^* u^j$ and $p_{r,j}^* = (-1)^{r-j} \binom{r}{j}\binom{r+j}{j}$. Notice that $P_r^*(u)$ is the r-th shifted Legendre polynomial, related to the Legendre polynomial by $P_r^*(u) = P_r(2u - 1)$.

L-moments can be estimated by sample L-moments. These are defined as:

$$\ell_r = \sum_{j=0}^{r-1} p^*_{r-1,j} b_j \tag{5}$$

where:

$$b_j = \frac{1}{n} \sum_{i=1}^{n} \frac{(i-1)(i-2)\ldots(i-j)}{(n-1)(n-2)\ldots(n-j)} x_{i:n} \tag{6}$$

C.2 TL-moments: existence, bounds and estimation

The general definition of TL-moment is as follows:

$$\lambda_r^{(s,k)} = \frac{1}{r} \sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j} E\left(X_{r+s-j:r+s+k}\right) \tag{7}$$

Hosking (2007) provided a sufficient condition for the existence of TL-moments that are showed to exist whenever $E\left\{\left[\max\left(-X_t, 0\right)^1 / (s+1)\right]\right\} < \infty$ and $E\left\{\left[\max\left(X_t, 0\right)^1 / (k+1)\right]\right\} < \infty$.

In the empirical analysis we set $s = k = 1$, therefore the first four TL-moments can be written as:[10]

$$\lambda_1^{(1,1)} = E\left(X2:3\right) = 6 \int_0^1 Q(u)u\left(1-u\right)\mathrm{d}u \tag{8}$$

$$\lambda_2^{(1,1)} = \frac{1}{2}E\left(X_{3:4} - X_{2:4}\right) = 6 \int_0^1 Q(u)u\left(1-u\right)\left(2u-1\right)\mathrm{d}u \tag{9}$$

$$\lambda_3^{(1,1)} = \frac{1}{3}E\left(X_{4:5} - 2X_{3:5} + X_{2:5}\right) = \frac{20}{3} \int_0^1 Q(u)u\left(1-u\right)\left(5u^2 - 5u + 1\right)\mathrm{d}u \tag{10}$$

$$\lambda_4^{(1,1)} = \frac{1}{4}E\left(X_{5:6} - 3X_{4:6} + 3X_{3:6} + X_{2:6}\right) = \ldots$$

$$= \frac{15}{2} \int_0^1 Q(u)u\left(1-u\right)\left(14u^3 - 21u^2 + 9u - 1\right)\mathrm{d}u \tag{11}$$

$\lambda_1^{(1,1)}$ and $\lambda_2^{(1,1)}$ are thus robust measures of location and scale. With $s = k = 1$ the condition for the existence of TL-moments is $E\left(|X_t|^{0.5}\right) < \infty$. For example, since the mean of the Cauchy distribution the cannot be defined, its L-moments cannot be defined either. However,

---

[10]In this case, with $k = s = 1$, the r-th TL-moment is defined in terms of a conceptual sample of size equal to $r + s + k = r + 2$.

since $E(|X_t|^{0.5}) = 1.41$ its TL-moment, $\lambda_r^{(1,1)}$, does exist.[11]

Bounds on TL moment ratios. Hosking (2007) showed that for the generic TL-moment ratio $\tau_r^{(s,k)} \equiv \lambda_r^{(r,s)}/\lambda_2^{(r,s)}$ $\forall r \geq 3$ a bound can be defined as follows:

$$\left|\tau_r^{(s,k)}\right| \leq \frac{2(m+1)!(r+s+k)!}{r(m+r-1)!(2+s+k)!} \quad \text{where } m = min(s,k) \tag{12}$$

for $s = k = 0$ we get $|\tau_r| \leq r!/r(r-1)!$ that simplifies to $|\tau_r| < 1$ $\forall r \geq 3$. For $s = k = 1$ the bound on TL-moment ratios is: $\left|\tau_r^{(1,1)}\right| \leq (1/6)(r+2)!/(r \times r!)$. Therefore, we get: $\left|\tau_3^{(1,1)}\right| = |SK_6| \leq 1.11$ and $\left|\tau_4^{(1,1)}\right| = |KR_6'| \leq 1.25$.

Sample TL moments. TL-moments can be consistently and unbiasedly estimated as:

$$\ell_r^{(s,k)} = \frac{1}{r\binom{n}{t+s+k}} \sum_{j=s+1}^{n-k} \sum_{i=0}^{r-1} (-1)^i \binom{r-1}{i} \binom{j-1}{r+s-i-1} \binom{n-j}{k+i} x_j \tag{13}$$

---

[11]We used Matlab's numerical integration routine "`integral`" to solve $E(|X_t|^{0.5}) = \int_{-\infty}^{\infty} |X_t|^{0.5} f(X_t) \mathrm{d}X_t$ assuming that X_t has a standard Cauchy distribution with pdf $f(X_t) = 1/\left[\pi\left(1+X_t^2\right)\right]$.

D   Additional tables & figures

D.1   Additional figures – iid samples

Figure A1: Sample distribution of SK_2
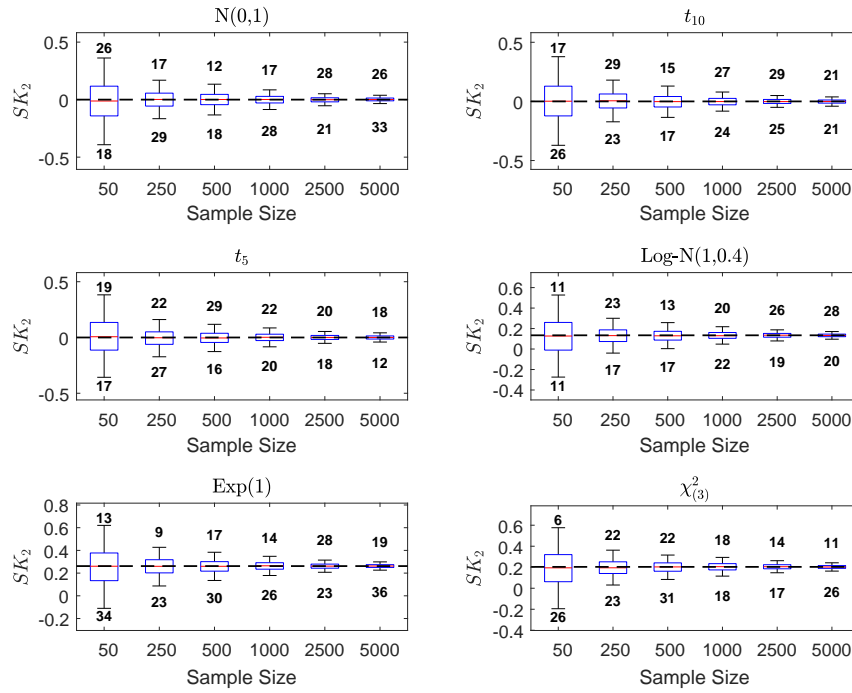
Notes: See notes to Figure 1.

Figure A2: Sample distribution of SK_3

Notes: See notes to Figure 1.

6
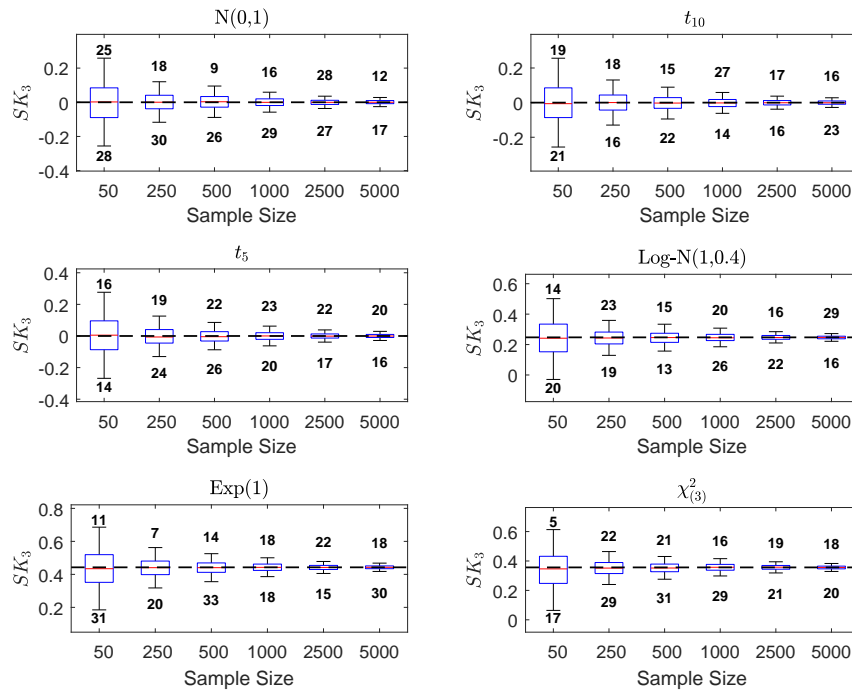
Figure A3: Sample distribution of SK_4
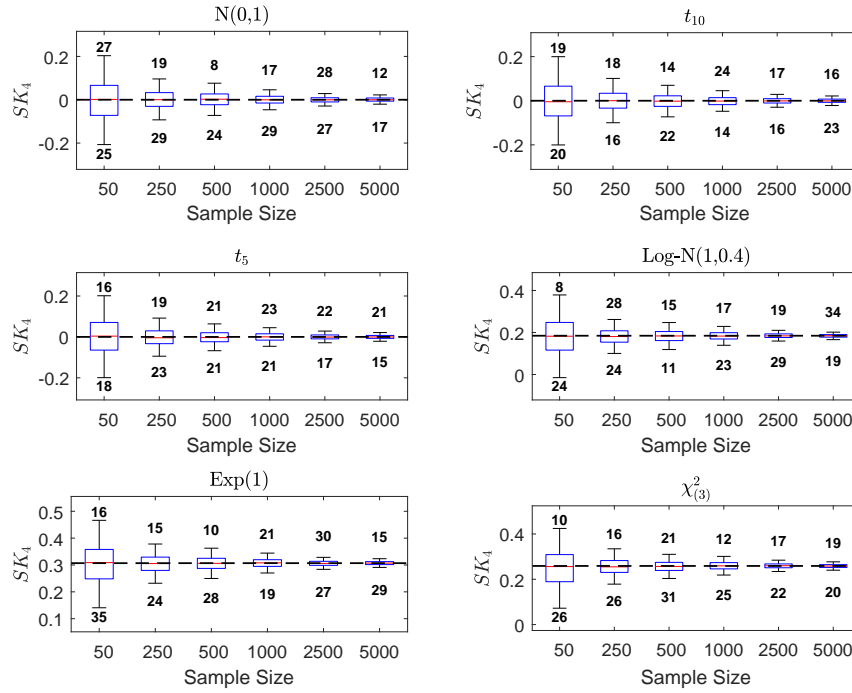
Notes: See notes to Figure 1.

Figure A4: Sample distribution of SK_5

Notes: See notes to Figure 1.
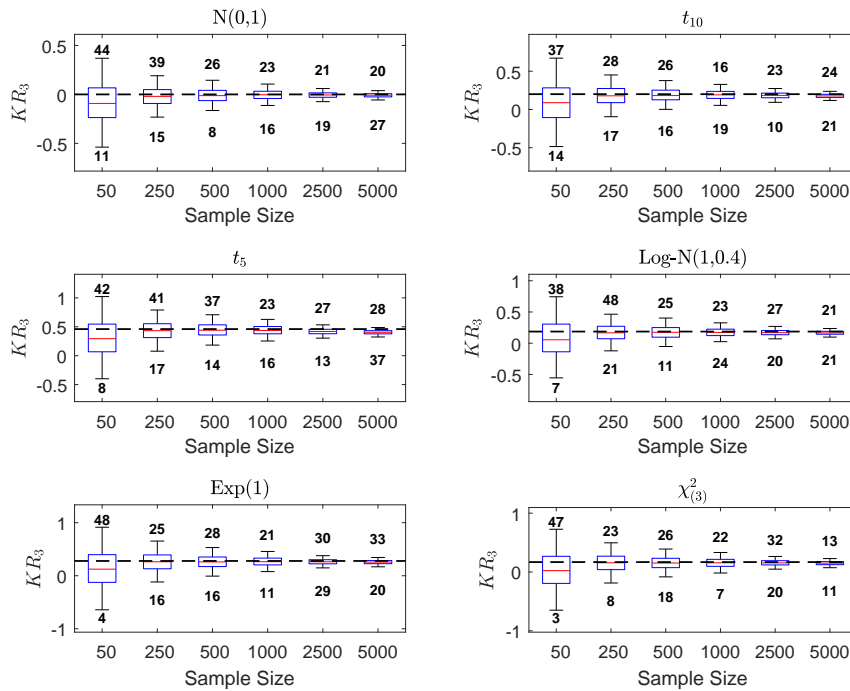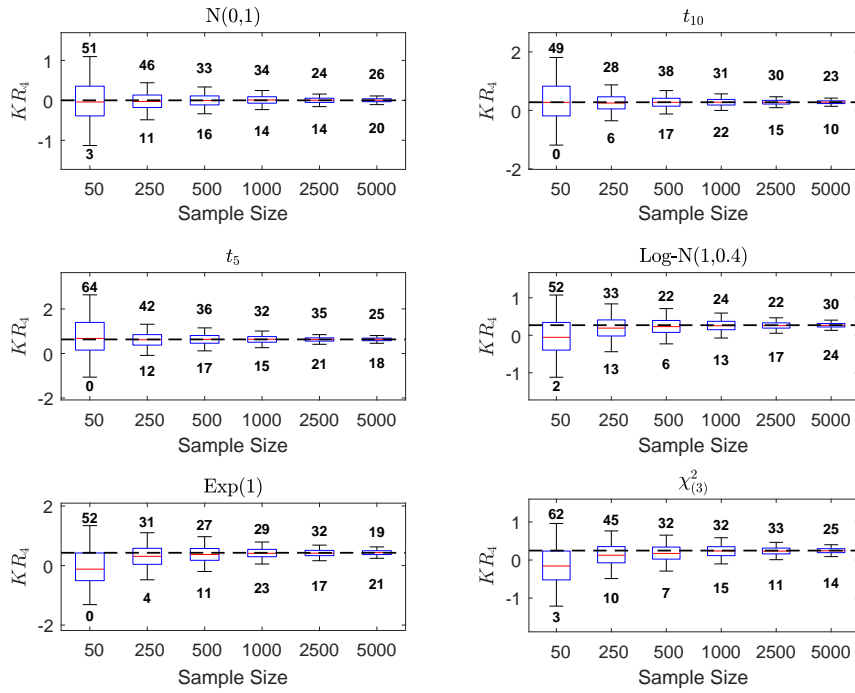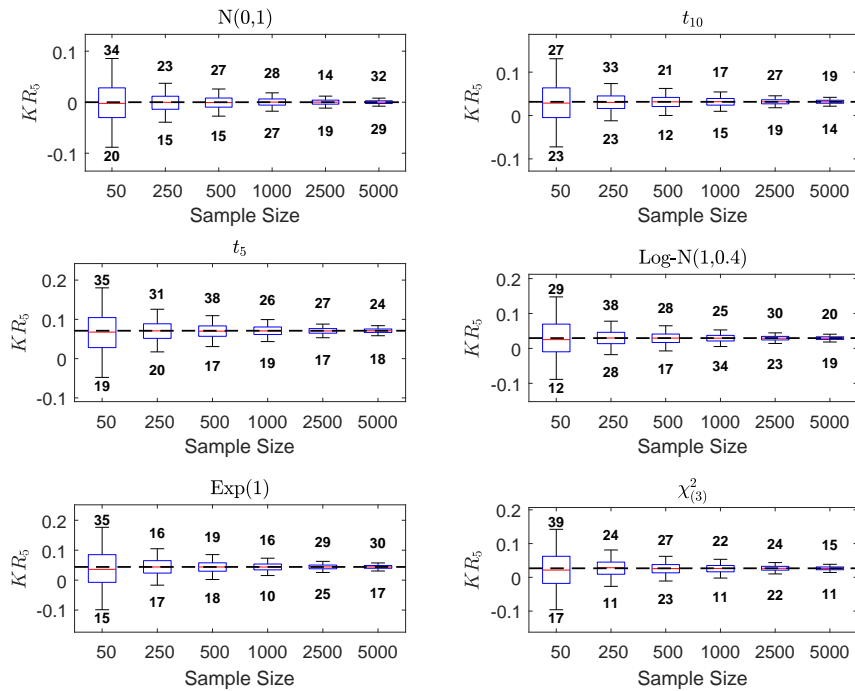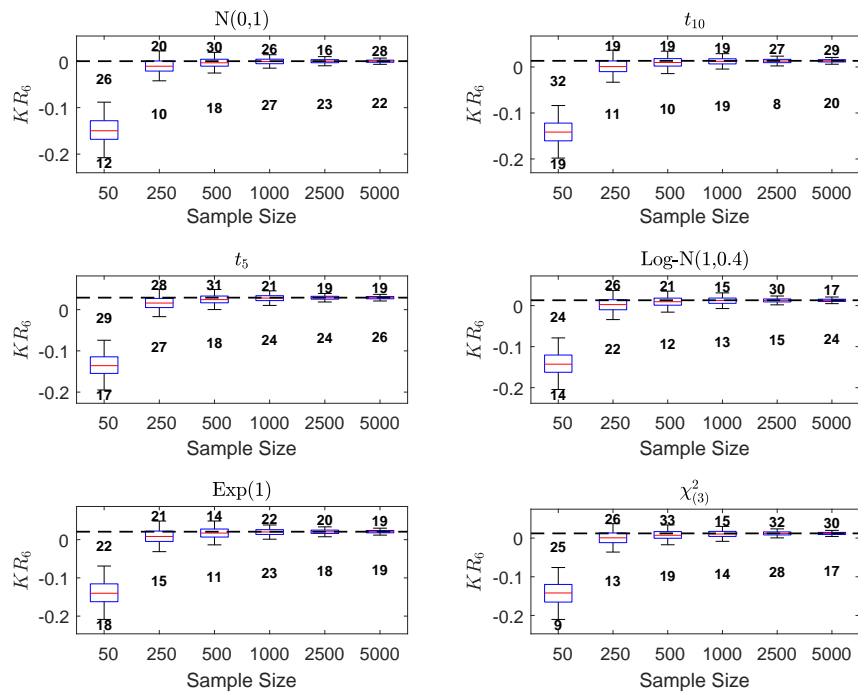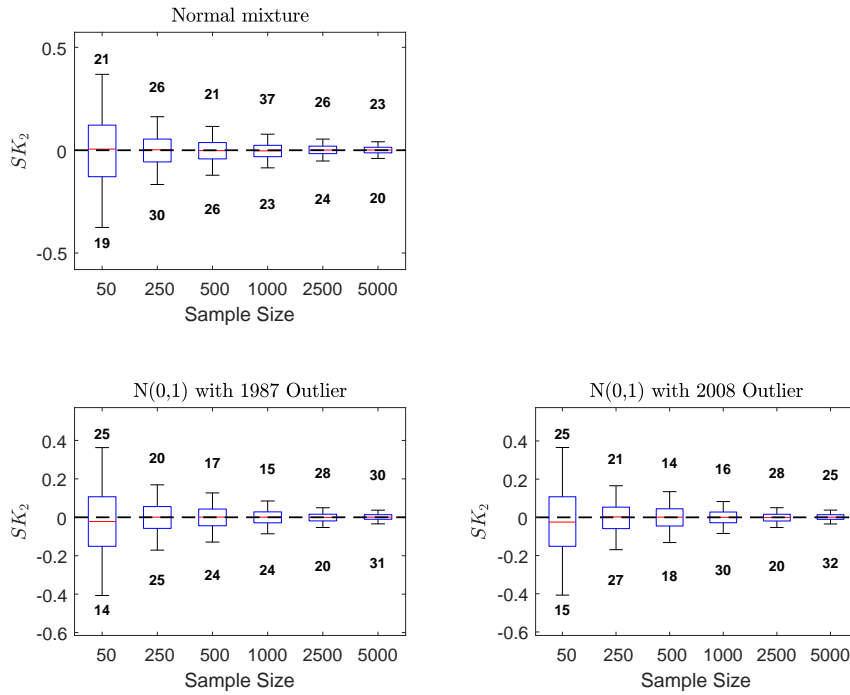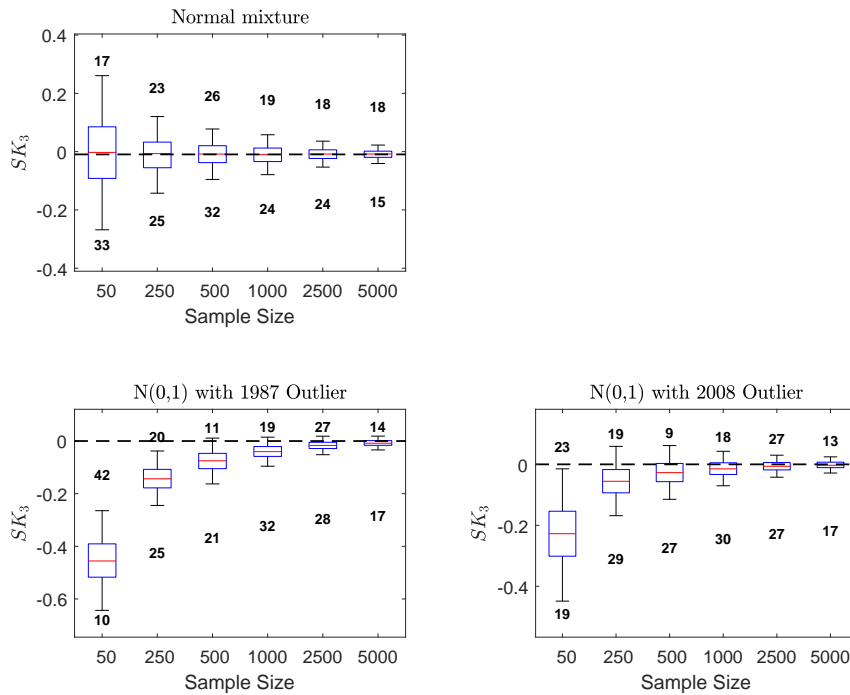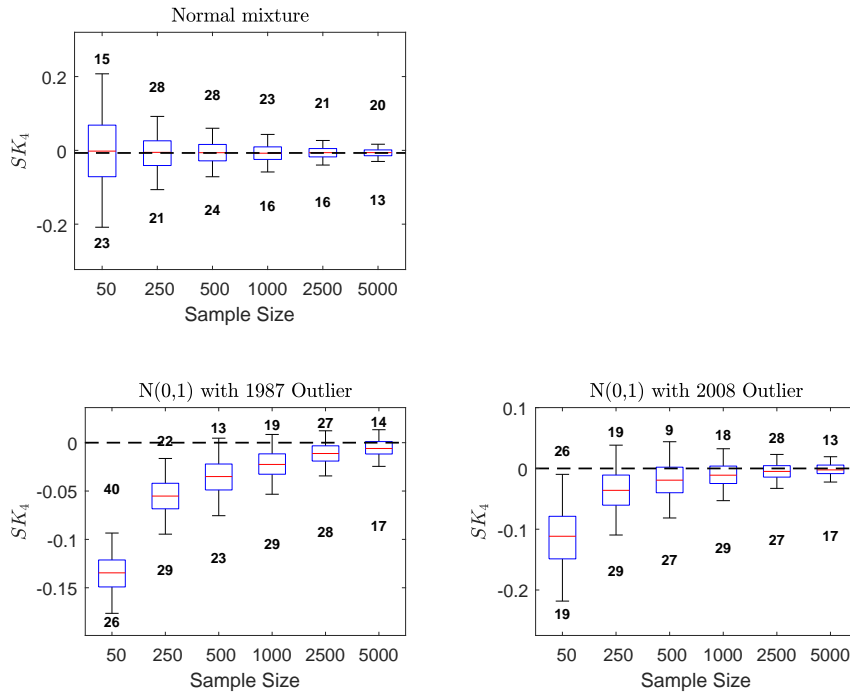
7
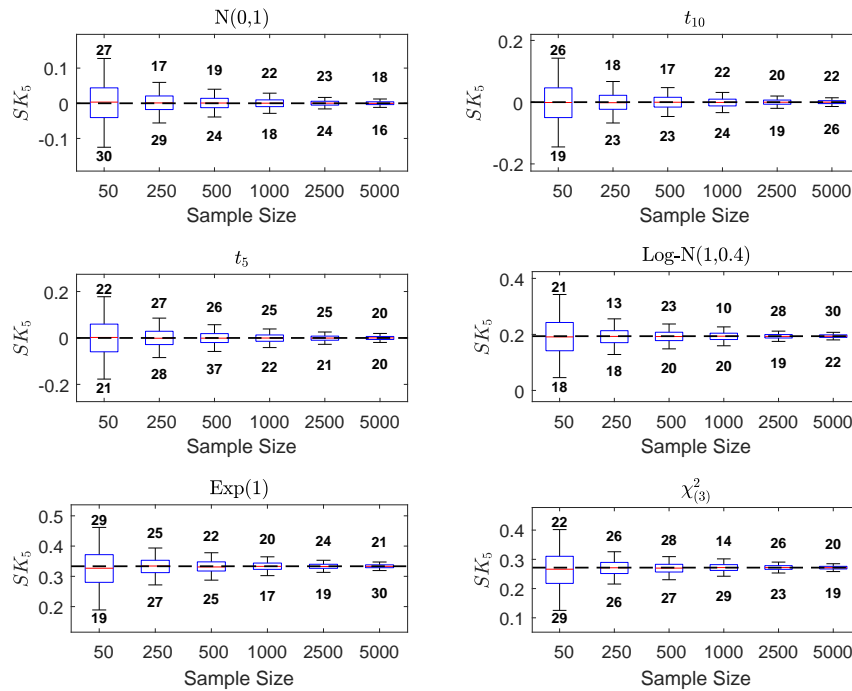
Figure A5: Sample distribution of SK_6
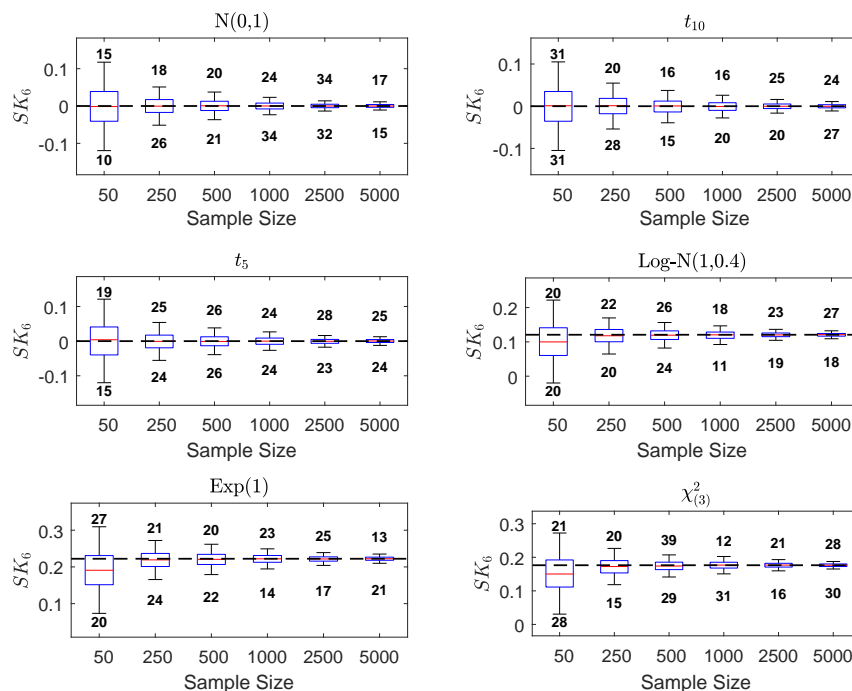
Notes: See notes to Figure 1.

Figure A6: Sample distribution of KR_2
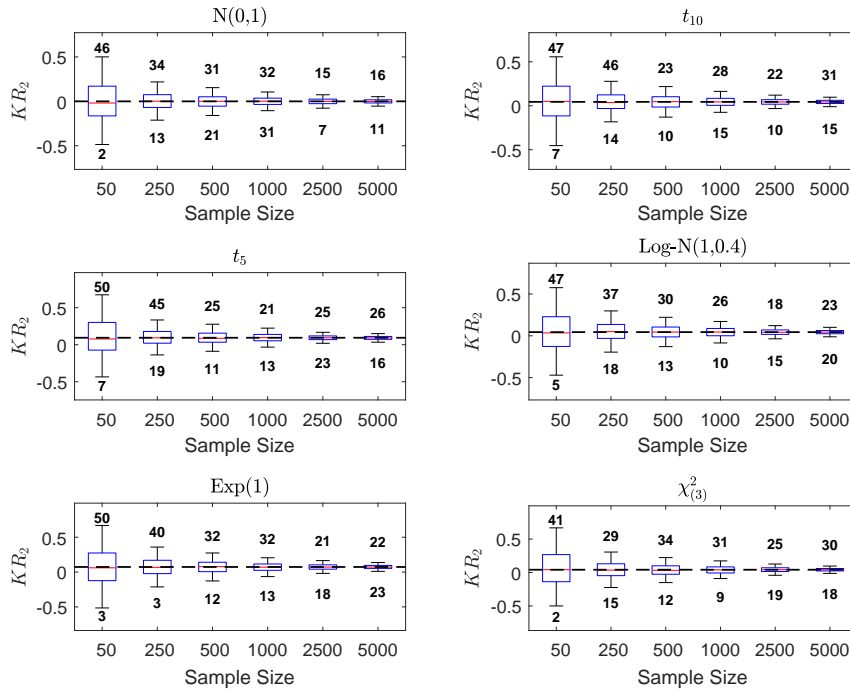
Figure A7: Sample distribution of KR_3

9

Figure A8: Sample distribution of KR_4

Notes: See notes to Figure 1.

Figure A9: Sample distribution of KR_5

Notes: See notes to Figure 1.

Figure A10: Sample distribution of KR_6

Notes: See notes to Figure 1.

D.2 Additional figures – outliers

Figure A11: Sample distribution of SK_2

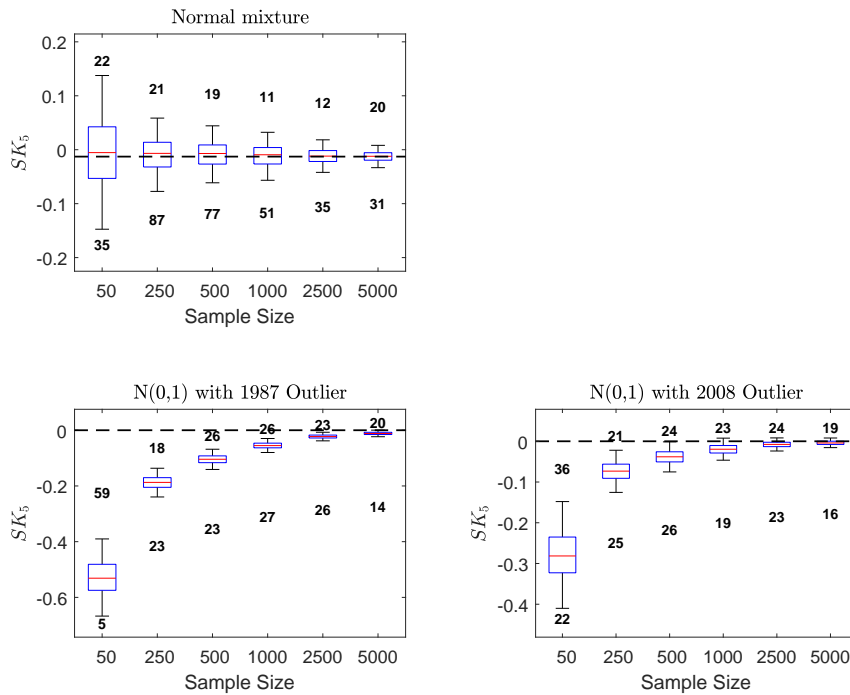*Notes*: See notes to Figure 1.

Figure A12: Sample distribution of SK_3

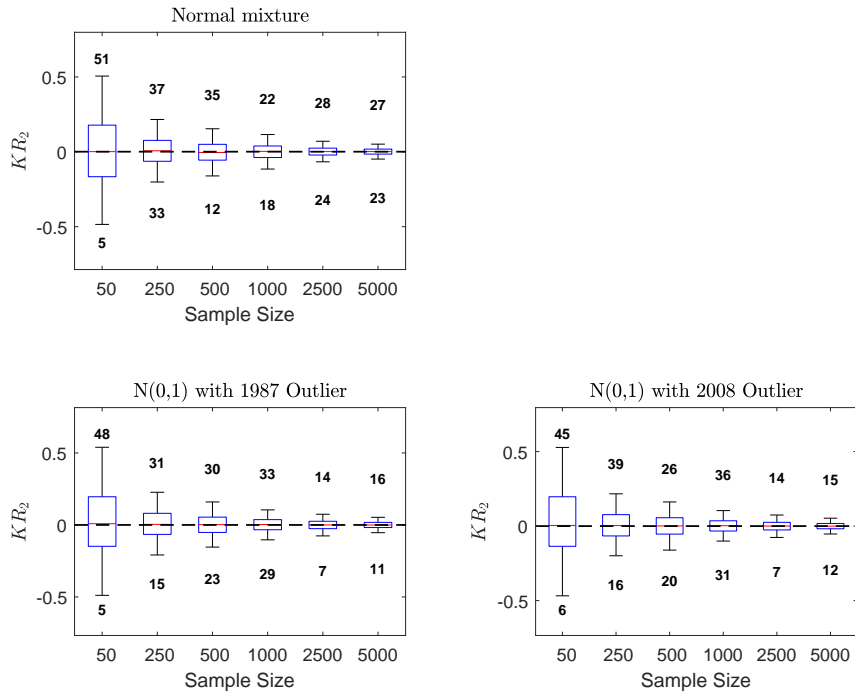*Notes*: See notes to Figure 1.

Figure A13: Sample distribution of SK_4

Notes: See notes to Figure 1.
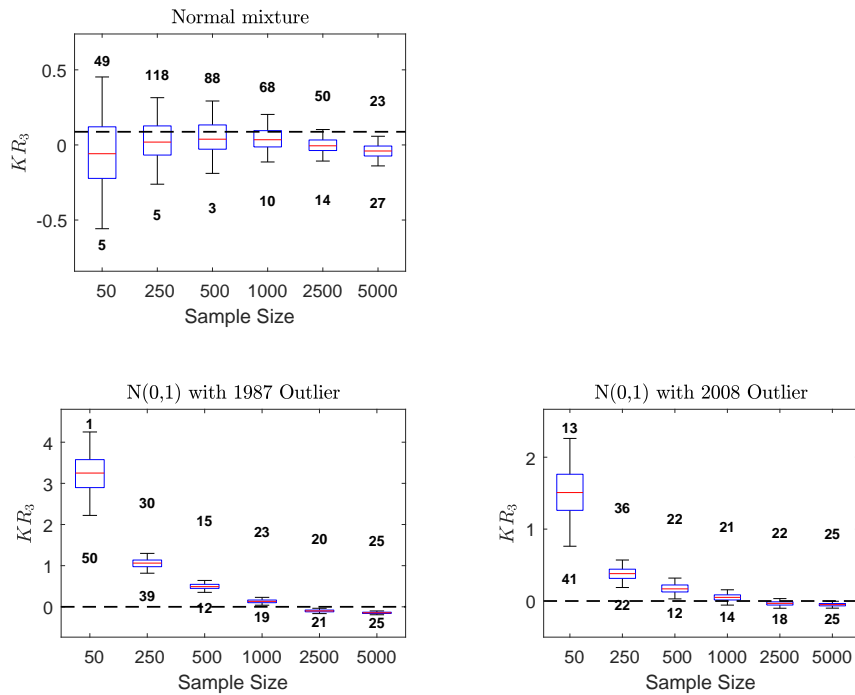
Figure A14: Sample distribution of SK_5

Notes: See notes to Figure 1.

14

Figure A15: Sample distribution of KR_2

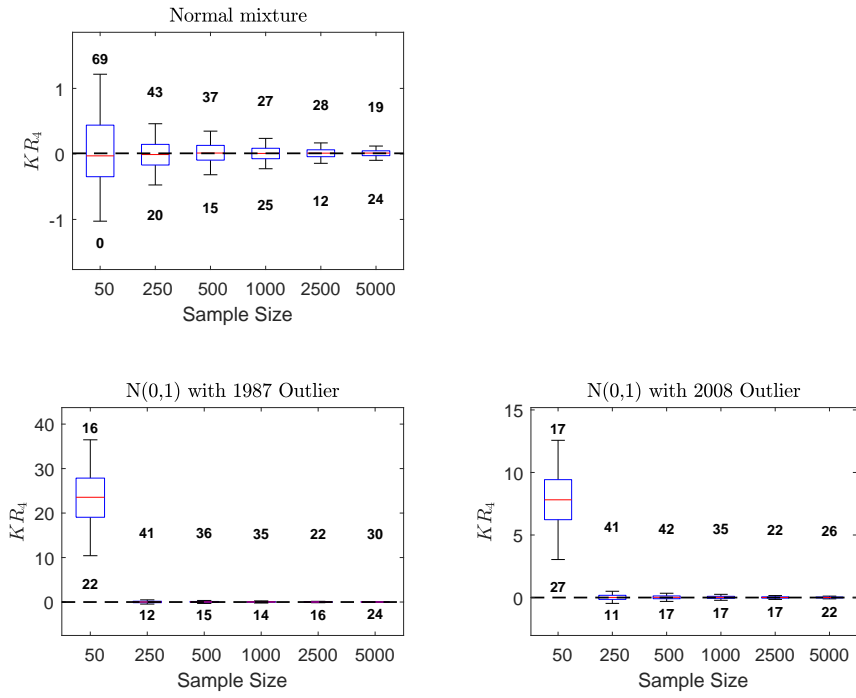*Notes*: See notes to Figure 1.

Figure A16: Sample distribution of KR_3

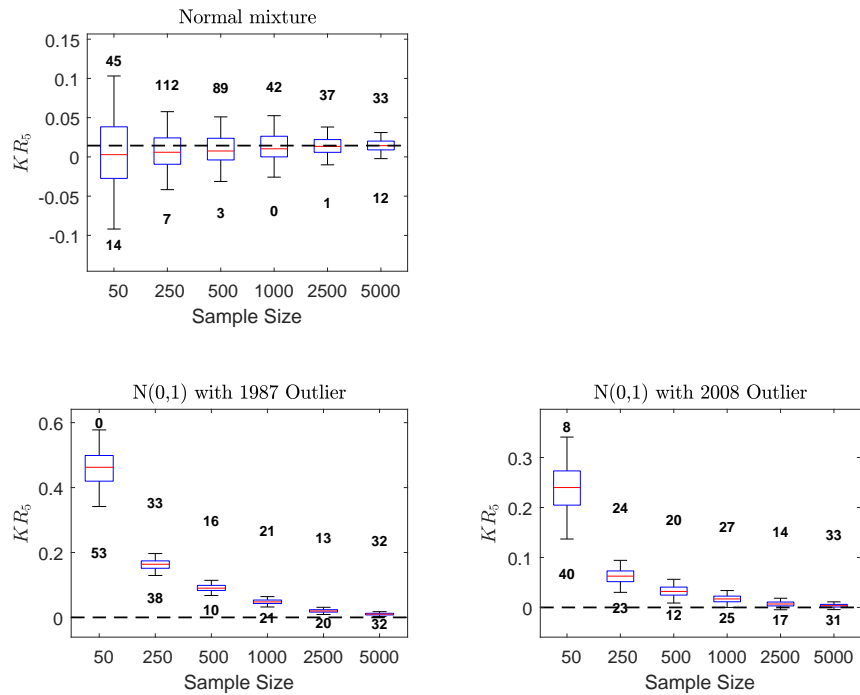

Notes: See notes to Figure 1.

15

Figure A17: Sample distribution of KR_4

Notes: See notes to Figure 1.

Figure A18: Sample distribution of KR_5

Notes: See notes to Figure 1.

D.3 Additional tables – Serial Correlation

Table D1: Bias and RMSE ratio of skewness estimators - AR(1) model ($\rho = 0.5$)

T	50	250	500	1000	2500	5000
			(a) N(0,1)			
SK_1	0.010 (1.000)	0.001 (1.000)	0.004 (1.000)	0.003 (1.000)	0.001 (1.000)	-0.000 (1.000)
SK_2	0.003 (0.580)	0.002 (0.505)	-0.001 (0.501)	-0.002 (0.490)	-0.000 (0.513)	0.000 (0.539)
SK_3	0.004 (0.420)	0.000 (0.374)	-0.000 (0.372)	-0.001 (0.348)	0.000 (0.369)	-0.000 (0.381)
SK_4	0.003 (0.336)	0.000 (0.299)	-0.000 (0.297)	-0.001 (0.278)	0.000 (0.295)	-0.000 (0.304)
SK_5	0.002 (0.204)	0.000 (0.183)	0.000 (0.179)	0.000 (0.172)	0.000 (0.175)	-0.000 (0.175)
SK_6	0.002 (0.191)	0.000 (0.159)	-0.000 (0.155)	-0.000 (0.147)	0.000 (0.152)	0.000 (0.155)
			(c) t_5			
SK_1	0.007 (1.000)	0.002 (1.000)	-0.007 (1.000)	0.019 (1.000)	-0.009 (1.000)	-0.000 (1.000)
SK_2	-0.017 (0.290)	0.005 (0.155)	-0.000 (0.129)	0.001 (0.094)	0.001 (0.074)	-0.001 (0.081)
SK_3	-0.011 (0.224)	0.003 (0.121)	-0.002 (0.105)	0.002 (0.077)	-0.000 (0.059)	-0.000 (0.065)
SK_4	-0.009 (0.170)	0.002 (0.092)	-0.001 (0.079)	0.001 (0.058)	-0.000 (0.045)	-0.000 (0.049)
SK_5	-0.001 (0.139)	0.001 (0.080)	-0.001 (0.072)	0.001 (0.051)	-0.001 (0.041)	0.000 (0.044)
SK_6	-0.003 (0.094)	0.001 (0.055)	-0.001 (0.046)	0.001 (0.033)	-0.000 (0.026)	-0.000 (0.029)
			(d) Log-N(1,0.4)			
SK_1	-0.568 (1.000)	-0.385 (1.000)	-0.371 (1.000)	-0.357 (1.000)	-0.349 (1.000)	-0.343 (1.000)
SK_2	-0.041 (0.257)	-0.034 (0.182)	-0.039 (0.159)	-0.037 (0.141)	-0.037 (0.124)	-0.036 (0.116)
SK_3	-0.073 (0.211)	-0.063 (0.181)	-0.066 (0.175)	-0.065 (0.182)	-0.065 (0.184)	-0.064 (0.186)
SK_4	-0.052 (0.158)	-0.044 (0.130)	-0.046 (0.124)	-0.045 (0.128)	-0.045 (0.128)	-0.045 (0.129)
SK_5	-0.056 (0.128)	-0.049 (0.121)	-0.050 (0.125)	-0.050 (0.133)	-0.050 (0.137)	-0.050 (0.142)
SK_6	-0.045 (0.106)	-0.033 (0.087)	-0.033 (0.085)	-0.033 (0.089)	-0.032 (0.090)	-0.032 (0.092)

Notes: see notes to Table 2.

Table D2: Bias and RMSE ratio of excess kurtosis estimators - AR(1) model ($\rho = 0.5$)

T	50	250	500	1000	2500	5000
			(*a*) N(0,1)			
KR_1	-0.194 (1.000)	-0.030 (1.000)	-0.020 (1.000)	-0.010 (1.000)	-0.004 (1.000)	-0.002 (1.000)
KR_2	0.013 (0.438)	0.006 (0.378)	-0.004 (0.359)	-0.003 (0.337)	0.001 (0.349)	0.001 (0.348)
KR_3	-0.097 (0.431)	-0.015 (0.381)	-0.009 (0.359)	-0.006 (0.357)	-0.007 (0.350)	-0.009 (0.370)
KR_4	-0.010 (0.967)	-0.007 (0.848)	0.002 (0.783)	0.000 (0.759)	0.002 (0.738)	0.001 (0.754)
KR_5	<u>-0.004</u> (<u>0.078</u>)	<u>0.000</u> (0.066)	<u>-0.001</u> (0.062)	<u>-0.000</u> (0.060)	0.000 (0.058)	<u>0.000</u> (0.058)
KR_6	-0.149 (0.259)	-0.010 (<u>0.062</u>)	-0.003 (<u>0.053</u>)	-0.001 (<u>0.048</u>)	<u>0.000</u> (<u>0.047</u>)	0.000 (<u>0.047</u>)
			(*c*) t_5			
KR_1	-5.189 (1.000)	-4.155 (1.000)	-3.756 (1.000)	-3.300 (1.000)	-3.064 (1.000)	-3.075 (1.000)
KR_2	<u>-0.017</u> (0.051)	-0.031 (0.025)	-0.030 (0.020)	-0.033 (0.012)	-0.036 (0.009)	-0.035 (0.011)
KR_3	-0.317 (0.084)	-0.177 (0.048)	-0.160 (0.045)	-0.167 (0.033)	-0.180 (0.033)	-0.192 (0.048)
KR_4	-0.180 (0.165)	-0.230 (0.081)	-0.210 (0.070)	-0.228 (0.051)	-0.230 (0.044)	-0.230 (0.060)
KR_5	-0.034 (<u>0.012</u>)	-0.025 (0.007)	-0.023 (0.007)	-0.023 (0.005)	-0.023 (0.004)	-0.022 (0.006)
KR_6	-0.171 (0.031)	<u>-0.023</u> (<u>0.006</u>)	<u>-0.014</u> (<u>0.004</u>)	<u>-0.011</u> (<u>0.003</u>)	<u>-0.011</u> (<u>0.002</u>)	<u>-0.010</u> (<u>0.003</u>)
			(*d*) Log-N(1,0.4)			
KR_1	-2.559 (1.000)	-1.641 (1.000)	-1.525 (1.000)	-1.435 (1.000)	-1.370 (1.000)	-1.317 (1.000)
KR_2	<u>-0.001</u> (0.090)	-0.012 (0.050)	-0.017 (0.039)	-0.012 (0.033)	-0.014 (0.024)	-0.013 (0.020)
KR_3	-0.181 (0.113)	-0.070 (0.069)	-0.064 (0.058)	-0.064 (0.054)	-0.067 (0.051)	-0.070 (0.053)
KR_4	-0.308 (0.207)	-0.150 (0.137)	-0.110 (0.111)	-0.089 (0.095)	-0.088 (0.080)	-0.081 (0.071)
KR_5	-0.015 (<u>0.018</u>)	<u>-0.009</u> (0.011)	-0.009 (0.009)	-0.009 (0.008)	-0.009 (0.007)	-0.008 (0.007)
KR_6	-0.158 (0.054)	-0.015 (<u>0.010</u>)	<u>-0.008</u> (<u>0.006</u>)	<u>-0.005</u> (<u>0.005</u>)	<u>-0.004</u> (<u>0.004</u>)	<u>-0.004</u> (<u>0.004</u>)

*Notes*: see notes to Table 2.

D.4 Additional tables & figures – FRED-MD

i

Table D3: Point estimates of maximal moment exponents

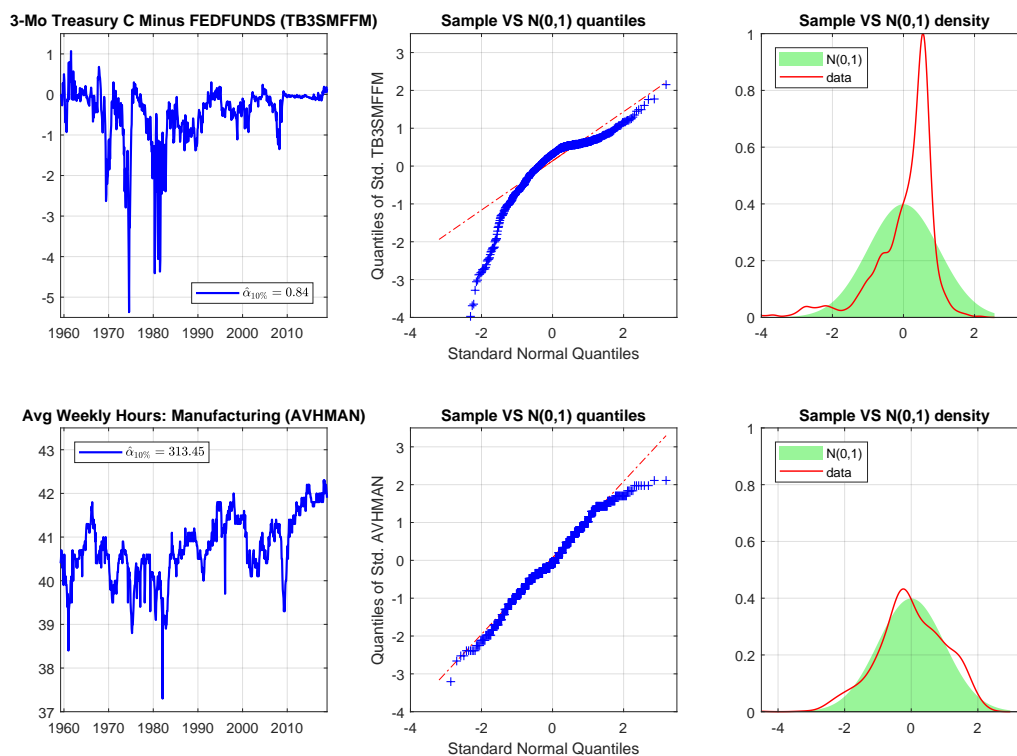| δ	$\hat{\alpha}_\delta$		
	min	median	max
5.0%	1.17	3.24	302.07
7.5%	1.26	2.91	233.98
10.0%	0.84	2.73	313.45

Notes: the table shows the minimum, maximum and median value of the estimates of the maximal moment exponents ($\hat{\alpha}_\delta$) for the series in the FRED-MD. The value of δ displayed in the first column of the table indicate the percentage of observations used to estimate $\hat{\alpha}_\delta$. Therefore $\hat{\alpha}_{10\%}$ indicates that we rely on the upper 10% of the empirical distribution of the series.

Figure A19: Robustness of point estimates of maximal moment exponents
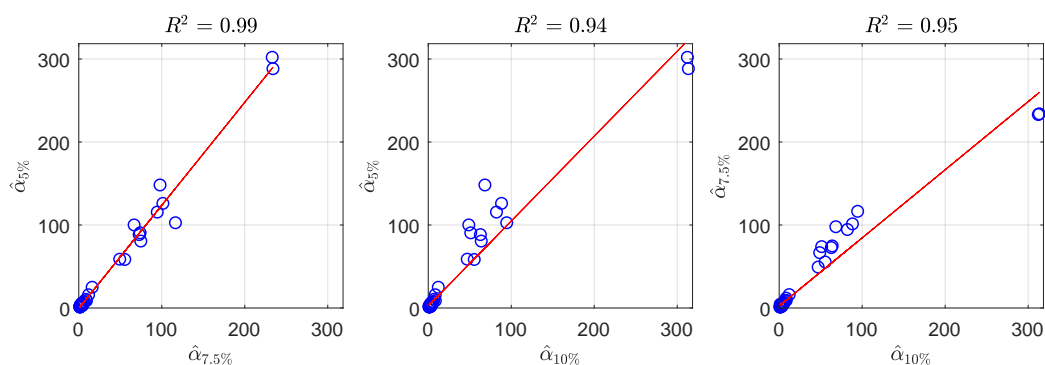
Notes: the figures report estimates of maximal moment exponents for series in the FRED-MD with different δ.

Figure A20: Series in the FRED-MD with highest and lowest maximal moment exponent

Notes: