How Fat-Tailed is US Output Growth?

*Reiner Franke*\(^a,\)\(^*\)

September 2013

\(^a\)University of Kiel, Germany

---

**Abstract**

Several studies have recently rejected the common hypothesis that aggregate output is normally distributed. The present paper reconsiders this issue for US output growth. To this end it focusses on the shape parameter \(b\) of the exponential power distribution (EPD), the two polar values of which constitute the normal distribution and the Laplace distribution with its fatter tails, respectively. The paper first warns against premature conclusions that neglect a structural break in output volatility. Distinguishing the two periods of the Great Inflation and the Great Moderation, it is then found that for quarterly industrial production (IP) the Laplacian cannot be rejected in both periods. By contrast, for monthly IP and quarterly GDP or firm output, the evidence is mixed and even contradictory. The question of whether non-normality can be considered a new stylized fact for macroeconomic modelling has therefore no unambiguous answer.

*JEL classification:* C 22, E 32.

*Keywords:* Exponential power distribution, spurious non-normality, Monte Carlo experiments, asymmetric confidence intervals.

---

1. **Introduction**

Macroeconomic variables or the shock processes that drive them are commonly considered to be largely compatible with normal distributions. This supposition is not just a semantic issue, for example in assessing whether or not our economic system is essentially well-behaved, so that the risk from the random shocks to it can be satisfactorily controlled (occasional crises notwithstanding). The issue is also important in the academic field, where the flourishing business of the estimation of DSGE models with its present predominance of likelihood techniques heavily rests on the assumption of normally distributed innovations.

*Email address:* franke@uni-bremen.de.
There have nevertheless been always some doubts about the prevalent view that takes normality for granted, in the first instance with respect to output and employment variables (as far as the empirical side is concerned, the following will refer to US data). Independently of discussions among heterodox theorists (like Blatt, 1983), concerns about asymmetry were already raised some thirty years ago. The possibly best-known contribution at that time was DeLong and Summers (1985) who, however, did not get much evidence for this. More recently, the tests by Bai and Ng (2005) failed to establish significant support of skewness or excess kurtosis. In contrast to these negative results, Christiano (2007) obtained significant excess kurtosis in the residuals of an unconstrained estimated vector autoregression. Interestingly, on the other hand, this did not prove sufficient to distort their Bayesian analyses that use the normal likelihood.

Other work finds indications of non-normality, too, and tries to take them into account in their theoretical models. De Grauwe (2012) concentrates on the kurtosis of output and emphasizes that, despite the normal shock processes, the nonlinear mechanisms in his model are able to reproduce this feature with great success.¹ While De Grauwe’s discussion remains more informal, Ruge-Murcia (2012) rigorously estimates the third-order approximation of a DSGE model by the simulated method of moments, an approach that need not rely on normality assumptions. Within this framework he can reject the null that productivity innovations are normally distributed in favour of the alternative that they follow an asymmetric Skew normal distribution.²

Besides studying skewness and kurtosis, there is another approach that assesses deviations of the empirical distribution from normality in their entirety. To this end, the analysis refers to the class of the exponential power (or Subbotin) distributions. Their general shape is governed by a parameter \( b \), where \( b = 2 \) yields the normal distribution and lower values give rise to progressively fatter tails. A second benchmark is here \( b = 1 \), at which the Laplace distribution prevails.³ Several papers by Fagiolo et al. (2007, 2008, 2009) claim it a universal phenomenon for output growth rates that estimates of \( b \) are so low that normality has to be rejected. Ascari et al. (2012) even raise this into the category of a stylized fact, i.e., a standard that macroeconomic models should seek to meet. Checking this with calibrated versions of a Real Business Cycle and a New-Keynesian model, they find out that the former can replicate this type of fat tails exogenously but not endogenously, and the latter neither endogenously nor exogenously. Thus, this work levels a serious criticism at the current practice of DSGE modelling.

¹ The strong nonlinearities originate with the agents’ switching between different types of boundedly rational expectations, which he puts forward as a pronounced alternative to DSGE modelling.
² The moments themselves on which the model is estimated include the third-order moments of hours worked and of the growth rates of consumption and investment; see Ruge-Murcia (2012, p. 931, Table 10).
³ It plays a prominent role in cross-sectional distributions of firm characteristics such as their rates of growth and profit; see, e.g., Alfarano et al. (2012).
As a number of other non-normality tests did not prove to be very helpful in identifying non-normal behaviour (see Franke, 2013, where the Jarque-Bera test, two of its generalizations and the Anderson-Darling test are considered), the present paper follows the parametric approach of the exponential power distribution (EPD) and concentrates on the estimates of its shape parameter \( b \) as a rough-and-ready measure of a normal or non-normal fatness of the tails when the amount of data is limited. We start out from the observation for US output data that the aforementioned studies are typically based on samples covering 40 years and more. Whereas such a sample size is certainly desirable from an econometric point of view, so many things have historically changed over this span of time that also the economy can be suspected of having shifted from one regime to another—which might introduce a spurious fatness in the distribution of the output growth rates. Somewhat strangely, this possible problem has not been systematically investigated, although the distinction between the two periods of the Great Inflation (GI) and the Great Moderation (GM) is a well-known topic in macroeconomics. In addition and more specifically, McConnel and Perez-Quiros (2000) provide firm statistical evidence of a structural break in the volatility of output growth around 1984.

The significance of the estimates of the EPD shape parameter can be conveniently investigated by elementary Monte Carlo experiments with autocorrelated random draws from EPDs with hypothetical values of \( b \). We first employ them for a long sample period to check the suspicion just mentioned of a deceptive non-normality, and second when searching for a possible non-normality over one or both of the two subsamples before and after the structural break. These experiments may also help evaluate the precision of the estimated parameters. Regarding the data, four different US growth series are studied in this way: GDP and firm output, which are quarterly data, and industrial production both as a monthly and a quarterly series.

The remainder of the paper is organized as follows. The next section introduces the EPD and how its shape parameter \( b \) can be conveniently and accurately estimated (without having to maximize a likelihood function). It also points out the necessity of being clear about what ‘normality’ precisely is to refer to. Section 3 checks whether the low estimates of \( b \) over a 47-years sample might be simply brought about by two different normal distributions in two subperiods. Section 4 is concerned with estimations over GI and GM, and with tests of whether or to what extent the normal distribution or the Laplace distribution could be possibly rejected. Section 5 inquires into a certain biasedness of the estimations and how one may assess their precision. These issues are of particular relevance for models builders who may have the ambition to reproduce a possible fatness in the tails of the data. Section 6 concludes. An appendix elaborates on several technical details.
2. The exponential power distribution and its estimation

The exponential power distribution (EPD) is a convenient parametric approach to a quantitative assessment of the degree of normality or non-normality of data, which is easy to interpret and has recently also been employed in a number of empirical studies on this subject. With respect to a real variable \( x \in \mathbb{R} \), it is specified by a density

\[
f(x; b, a, m) = \frac{1}{2a b^{1/2} \Gamma(1+1/b)} \exp \left\{ -\frac{1}{b} \left| \frac{x - m}{a} \right|^b \right\}
\]

(1)

where \( \Gamma(\cdot) \) is the Gamma function. The three parameters identifying the EP distributions are the location parameter \( m \), the scale parameter \( a \), and most importantly the shape parameter \( b \). An attractive feature of this family of distributions is that as a special case it nests the normal distribution, which is recovered with \( b = 2 \) (in this case the parameter \( a \) is equal to its standard deviation usually denoted \( \sigma \)). Higher values of \( b \) flatten the entire distribution, such that for \( b \to \infty \) the distribution tends towards the uniform distribution with support \([-a, a]\). Conversely, as \( b \) decreases from the Gaussian benchmark, the shoulders of the distribution become slimmer and the tails become fatter. The benchmark of practical concern in this direction is the Laplace distribution at \( b = 1 \), the density of which yields the famous tent shape in a semi-log diagram (to be illustrated in Figure 1 further below).

Since \( b \) is a parameter characterizing the global shape of the distribution, it can be expected that it will be a more robust measure of the fatness of tails than the kurtosis. Theoretically, the kurtosis implied by an EPD is given by \( K = K(b) = \Gamma(1/b) \Gamma(5/b) / [\Gamma(3/b)]^2 \) (Chiodi, 1995, Section 2). If one prefers the kurtosis as a more familiar measure of fatness, this relationship may be used as a check of the direct empirical calculation of \( K \). To get a first impression of its order of magnitude, vis-à-vis \( K = K(2) = 3 \) for the normal distribution, the kurtosis of the Laplace distribution rises to \( K = K(1) = 6 \).

There are several likelihood methods to estimate the parameters of an EPD. More convenient for us is a moment matching procedure proposed by Mineo (1994, 2003). Based on a generalized index of kurtosis as it can be derived for EPDs, it permits an estimation of \( b \) that is independent of the other two parameters. For sample sizes of 100 or 200 observations it also appears to give the most accurate results (Mineo, 2003).

4 To check possible alternatives to fit fat- and medium-tailed distributions of output growth data, Fagiolo et al. (2008) experimented with the Cauchy, the Student-\( t \) and the Lévy-Stable distribution. They found evidence of fat tails for all of them but conclude that the EP density seems to outperform the other three density families.

5 More general versions of (1) can also account for asymmetries; see, e.g., Botazzi and Secchi (2008), or Zhu and Zinde-Walsh (2009). We neglect this extension, mainly since we feel that our samples are too small for any reliable results—if skewness is present at all.

6 These estimations are not always without problems as for sample sizes less than 100 it may happen that the likelihood function has no minimum within a reasonable range (Agró, 1995, p.527).
Referring to a sample \( \{x_t\}_{t=1}^T \), the method does not require the optimization of an objective function but only the solution of an implicit equation in \( b \), which is indicated by the exclamation mark in the first equation:

\[
\sqrt{\frac{\Gamma(3/b) \Gamma(1/b)}{\Gamma(2/b)}} = \hat{d}_2, \quad \hat{d}_k = \frac{1}{T} \sum_{t=1}^T |x_t - \bar{x}|^k, \quad k = 1, 2
\]

(\( \bar{x} \) being the mean value of the series). The expression on the left-hand side of the first equation is the aforementioned alternative index of kurtosis. It is strictly decreasing in \( b \) over a sufficiently wide range, hence the equation has a unique root \( \hat{b} \). On the other hand, the right-hand side of the equation shows that in the determination of \( \hat{b} \) only the second (and not the fourth) moment is involved, which confirms the expectation articulated above of more robustness.

Once the shape parameter is estimated, the parameters \( m \) and \( a \) characterizing the location and scale of the EPD can be successively obtained from two extra equations (the details are given in the Appendix). With our focus on the general shape of the distributions, they will only be needed for a standardization of the empirical frequency distributions to compare different periods or output data later in Figure 1.

Before turning to an application of the class of EPDs it may be clarified what precisely ‘non-normality’ is meant to be and what precisely ‘fat tails’ refer to. In this respect it is important to note that several studies are not concerned with the raw data, i.e. growth rates, but they first filter them and estimate the shape parameter \( b \) of the resulting residuals. The ‘depuration’ is typically done by a VAR in a multivariate context (see Christiano, 2007) or by an ARMA approach to a univariate series (see Fagiolo et al., 2008). The motivation for filtering is to avoid a possible bias and to get to the core of the economy.

It should, however, be pointed out that underlying such procedures is a certain view of the world, namely, that the economy is driven by possibly non-normal shocks that propagate in essentially linear ways. It may be nice to know that the economy is essentially non-normal in this sense. On the other hand, even if a model builder shares the methodological view but has the ambition that his or her model displays some non-normality, low empirical estimates \( \hat{b} \) of a specific series of residuals will only be of limited help to him or her when the model has a more detailed and perhaps not even comparable block of structural random innovations.

Moreover, fat tails in the residuals from a linear stochastic process are not very informative for researchers entertaining the alternative vision that (an approximation to) the real-world data generation process is strongly nonlinear and that this is the main

---

\(^7\) Estimation of \( b \) via the implicit equation in (2) is also mentioned in Botazzi (1994, p. 4), the manual on the software SUBBOTOOLS, which is online freely available at http://cafim.sssup.it/~giulio/software/subbotools. It has, however, to be noted that there are two misprints in his formula (the software source code in file subbofit.c is correct).
cause for any non-normality phenomena. When examining whether their models are able to reproduce some of them, they will therefore have to resort to the unfiltered data.\(^8\) This and the previous observation are the main reason why we will here abstain from ‘depurating’ the growth rates.

In evaluating the significance of the estimations of the shape parameter, we will nevertheless wish to generate artificial data on which subsequently arbitrarily many re-estimations of \(b\) can be carried out. Without committing ourselves to a specific dynamic process, we will assume that the data have been randomly drawn from an EPD with a hypothesized value of \(b\) and then check whether this hypothesis can be maintained or has to be discarded. Apart from requiring that the elements in such a random sample are autocorrelated with the same coefficient as in the corresponding empirical period, no further assumption will be involved. Formally, this amounts to simulating an AR(1) process with innovations from an EPD.\(^9\)

These tests are (parametric) bootstraps, or Monte Carlo (MC) experiments. Accordingly, ‘normality’ means that an empirically estimated value \(\hat{b}\) has a sufficiently high probability of arising from EPD random draws with \(b = 2\); otherwise ‘non-normality’ is said to prevail. Likewise, the polar case of Laplacian behaviour is (not) rejected if the hypothesis \(b = 1\) is (not) rejected as being compatible with \(\hat{b}\) in such an experiment.

Regarding the data itself, our investigations are concerned with the quarterly growth rates of GDP and firm output (essentially non-financial corporate business). We refer to them as GDP and YF, respectively. The latter series is of interest since it is closer to the theoretical frameworks in most of the macroeconomic models. An even narrower output concept is that of industrial production. It has the additional advantage that it is available as a monthly series, denoted IPm. For a better comparison with the former data, we also specify it as quarterly growth rates, IPq. Interestingly, the results for quarterly and monthly industrial production can be somewhat different, where the monthly frequency does not necessarily produce fatter tails than the quarterly frequency. (See the Appendix for the data sources.)

3. A warning against premature conclusions

The total sample period underlying our study covers a time span of not quite 50 years. Disregarding the 1950s it begins in 1960:q1 and we let it end in 2007:q2 before the first signs of the financial crisis in the real sector. This amounts to a total of \(T = 190\) quarters, or \(T = 570\) months.

\(^8\) Filtering their simulated data in the same linear fashion as in the empirical investigations would in principle be possible but does not appear very meaningful when pronounced nonlinear mechanisms are present in such a model.

\(^9\) The details of how to generate random variates of an EPD are given in the Appendix. Since \(b\) can be estimated independently of the other parameters, the mean and variance of these random draws play no role. The only exception are the pooled random samples in the next section.
The estimates \( \hat{b} \) of the EPD shape parameter over this period for our four output growth rates are reported in the first row of Table 1 (in the column ‘pooled’, an expression to be explained in a moment). The four statistics are more or less close to 1 and in any case considerably less than 2. As a side result from the experiments in the next section it can be said that, when taken as such, normality would be rejected for all of them.

This conclusion should, however, be given a second thought. Apart from a possible bias by serial correlation in the data, one may be aware of the underlying assumption that the data generation process has not essentially changed over time, which is not obvious. Already on the basis of fairly general arguments, a greater part of macroeconomic research divides a period of 40 or 50 years after 1960 into two subsamples, which according to the common terms ‘Great Inflation’ and ‘Great Moderation’ constitute two different regimes. While the names primarily refer to inflation rates and the conduct of monetary policy, there is also rigorous econometric work showing a significant decline in the volatility of output growth. In fact, with quarterly GDP growth rates from 1953:q2 to 1992:q2, McConnell and Perez-Quiros (2000) reveal strong evidence for one—and only one—structural break, where the most suitable point estimate for the break date is 1984:q1. We follow the upshot of their analysis and, maintaining the expressions Great Inflation (GI) and Great Moderation (GM), distinguish two subperiods of nearly equal length, GI: 1960:q1 – 1983:q4 \((T = 96\) observations\) and GM: 1984:q1 – 2007:q2 \((T = 94)\); correspondingly for the monthly data we have \(T = 288\) and \(T = 282\) observations, respectively.

The middle part of Table 1 documents greater differences between GI and GM for all four output variables. In the first instance, from GI to GM the standard deviations of the growth rates were reduced by up to 50 per cent. Here no formal statistical tools are needed to classify these changes as significant. We also observe a fall in the general level of growth, but this change is much more moderate. A similar statement holds true for the first-order autocorrelation \(\rho\), though there is some difference between industrial production and the other two growth rates. With respect to output we can therefore summarize that GI and GM, quite apart from inflation, constitute two different growth regimes of high and low volatility, respectively.

Now, if we have a time series with strong noise in the first half and weak noise in the second, it may be supposed that this gives rise to a higher kurtosis since the higher values (in modulus) in the first half become a rarer, although not exceptional, event when considering the full sample. The argument might be valid even if the random forces are normally distributed in each of the two subsamples.

To check this conjecture, namely, that a structural change in the volatility alone may contribute to fat tails in the distribution over the entire sample, an elementary Monte Carlo (MC) experiment of the type described in the previous section can be designed. Letting \(T_1\) \((T_2)\) be the sample size of a given growth series in GI (GM), we first capture GI by \(T_1\) random draws from a normal distribution with the empirical mean, variance
Table 1: Statistics for GI, GM, and GI + GM.

<table>
<thead>
<tr>
<th></th>
<th>GDP</th>
<th>YF</th>
<th>IPq</th>
<th>IPm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GI GM pooled</td>
<td>GI GM pooled</td>
<td>GI GM pooled</td>
<td>GI GM pooled</td>
</tr>
<tr>
<td>$\hat{b}$</td>
<td>— —</td>
<td>1.18</td>
<td>— —</td>
<td>1.43</td>
</tr>
<tr>
<td>std. dev.</td>
<td>4.45</td>
<td>2.12</td>
<td>— —</td>
<td>4.12</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.26</td>
<td>0.23</td>
<td>— —</td>
<td>0.30</td>
</tr>
<tr>
<td>median $b$</td>
<td>— —</td>
<td>1.37</td>
<td>— —</td>
<td>1.69</td>
</tr>
<tr>
<td>crit. $b$</td>
<td>— —</td>
<td>1.10</td>
<td>— —</td>
<td>1.32</td>
</tr>
<tr>
<td>$p$-value</td>
<td>— —</td>
<td>12.64</td>
<td>— —</td>
<td>13.10</td>
</tr>
</tbody>
</table>

Note: The mean, the standard deviation (std. dev.) and the autocorrelation $\rho$ are the statistics of the empirical series over GI and GM, while $\hat{b}$ is the estimate over the entire sample GI + GM. Median $b$ is the median of the two-regime MC experiment described in the text and ‘crit. $b$’ its 5%-quantile, where normality is rejected if $\hat{b} < \text{crit. } b$. The corresponding $p$-values are given in per cent, and the bold face figures emphasize rejection of normality.

and autocorrelation $\rho_1$ over this period. Secondly, we generate $T_2$ draws from the normal distribution with the GM statistics and add this series to the first one. Subsequently the shape parameter $b$ of this compound series is estimated.\(^{10}\)

Repeating this $C = 5000$ times, a collection $\{b^c\}_{c=1}^C$ of estimates from the two pooled normal distributions is obtained. Table 1 confirms the above supposition that this scenario introduces a certain fatness in the tails; that is, for all four output concepts in the experiment the median is not around 2 but falls definitely short of it. Roughly, the higher the ratio of the two standard deviations in GI and GM, the lower the median of $\{b^c\}$ and the fatter these tails.

In order to rule out at a 5% significance level that the empirical shape $\hat{b}$ could have been generated from the two pooled normal distributions, a one-sided test requires $\hat{b}$ to be less than the 5%-quantile of the $\{b^c\}$ distribution. According to Table 1, this is clearly the case for the monthly as well as quarterly growth rates of industrial production, but not for GDP and firm output. Despite the low estimates of $b$, it would be premature to attribute the shape estimates of the latter two growth rates to a non-normal fatness in their distribution.

\(^{10}\)This experiment is similar in spirit to the regime-switching models that Fagiolo et al. (2008, p. 666) propose to employ for future research.
The last row in the table shows the corresponding p-values from \{b^c\}, which are determined as the percentage of cases in the MC distribution where \(b^c < \hat{b}\). Their interpretation is that if instead of the 5%-quantile the empirical estimate \(\hat{b}\) were employed as a benchmark for rejecting the null hypothesis of ‘pooled normality’, then this \(p\) would be the probability of thus committing an error (a type I error, that is).

In sum, we find full confirmation of the non-normality of industrial production, but statements in the literature about fat-tailed output growth for other series are to be taken with care. While admittedly our experimental design with the sudden change at the break date is somewhat crude, the \(p\)-values in Table 1 for GDP and firm output do not necessarily give us much reason to expect that a smoother transition from one regime to the other would provide a broader scope for fat tails.

It should nevertheless be mentioned that our inability to reject the pooled normality for GDP is in some contrast to Fagiolo et al. (2008, p. 654) who reject normality for the residuals from an ARMA estimation, although over a much longer sample period from 1947 to 2005. One may, however, wonder whether as intended this filter really removes the structural change and the thus resulting downward bias from the data.\(^{11}\) At any rate, the opposite conclusion by Fagiolo et al. demonstrates that for comparative discussions it can become very important to have a clear idea about the specific notion of the null of ‘normality’.

4. The null hypotheses of normality and the Laplacian

In the remainder the paper will be concerned with the growth rate distributions over each of the two subsamples, in order to see whether these results are similar to the full sample or whether the fatness may perhaps have changed from GI to GM.\(^{12}\) The first row in Table 2 further below reports the point estimates \(\hat{b}\) of the shape parameter in GI and GM for our four series. Before a statistical evaluation of their significance let us, however, consider Figure 1, which gives a first geometric account of the goodness-of-fit as well as an impression of how much the distributions differ from the two benchmarks of the normal and the Laplace distribution. It is in this respect informative to plot the logs of the corresponding densities. For a direct comparison of the different series the original growth rates \(x\) are moreover standardized, which means that in addition to \(b\) also the other two parameters \(m\) and \(a\) are estimated (see the Appendix) and the values of \(x\) are transformed into \(z = (x - \hat{m})/\hat{a}\).

\(^{11}\) While according to the ARCH-based tests employed by Fagiolo et al. (2008, p. 652) the residuals show no more signs of heteroscedasticity, it is not clear whether these tests are also fully appropriate to recognize a permanent regime shift. Apart from this, there may be some discussion on employing a full-fledged ARMA filter, which is perhaps overly sophisticated.

\(^{12}\) This is another issue that Fagiolo et al. (2008, p. 666) suggest for future research.
The top-left panel in Figure 1 shows the distribution of the GDP growth rates in the GI period (incidentally, firm output looks fairly similar and is therefore not included in the figure). The log of the exponential power density constituted by the estimated shape parameter $\hat{b} = 1.91$ (as indicated in the header) is drawn as the bold (red) line. The dots distributed around it are the density values of the $T=96$ observations of this period. That is, we use a standard nonparametric procedure to compute a kernel density estimator $\hat{f}(z)$ of the empirical values of $z$ and for each observed $z_t$ plot the point $(z_t, \ln[\hat{f}(z_t)])$. Over a wide range and especially in the middle part of the growth rates, these points really nestle into the smooth curve of the theoretical density function. The estimation thus inspires confidence, its outcome being more than just the result of a somewhat abstract and technical concept.

![Figure 1: Estimated densities (nonparametric: dots; EPDs: solid lines).](image)

The density of the normal distribution (where $b = 2$) looks very similar to the EP density with $b = 1.91$, which is the reason why it has not been plotted in this diagram. The

---

$^{13}$ We employ the Epanechnikov kernel for this purpose; see Davidson and MacKinnon (2004, pp. 678–683) for the computational details.
other polar case of the Laplacian density, where $b=1$, is the tent-shaped thin solid (green) line. The semi-log diagram brings out very nicely the higher degree of fatness in its tails. It might be argued that perhaps the Laplacian provides a better fit for the moderate and more extreme negative values of $z_t$, which could even suggest an asymmetric estimation approach with different values of $b$ for positive and negative values of $z$. However, we better abstain from this idea since given the relatively small sample size, this would also require an econometric discussion of the risk of overfitting.

The top-right panel in Figure 1 presents the density for the GM period (again, YF does not look very different). Note first that there are more extreme points, beyond $z=\pm 2$ say, than in GI with still a non-negligible probability. This is one reason why the shape is estimated at $b=1.29$ and the EPD spreads wider away from the normal distribution (the curve with $b=2$ in the diagram) than in GI. It may appear that here even the Laplacian density would not accomplish too bad a fit, an issue that we will return to shortly.

The lower four panels show the results for the quarterly (IPq) and monthly (IPm) growth rates of industrial production (the two middle and two lower panels, respectively). Again, EPD seems to be a decent parametric approach to describe the empirical densities, at least as long as the symmetry postulate is maintained. Comparing the quarterly to the monthly series, it will not come unexpectedly that we observe more extreme events for the latter, even after the standardization. It may now, however, be somewhat surprising that this does not necessarily mean fatter tails. On the contrary, the monthly frequency tends to exhibit a lower degree of fatness than the quarterly frequency; a mildly lower one in GI ($\hat{b}=1.35$ versus $b=1.18$) and a stronger effect in this direction in GM ($\hat{b}=1.65$ versus $\hat{b}=1.11$).\(^{14}\) Hence also the shape of the density in its centre appears to play a non-negligible role for estimation, although the naked eye is not very good at recognizing this.\(^{15}\)

Given that some estimates are fairly close to $b=2$ or $b=1$, it is only natural to ask if the empirical distributions are, or are not, significantly different from the normal or the Laplace distribution, respectively. To this end, we can make use of the same kind of Monte Carlo experiment as in the preceding section, this time with the correspondingly shorter samples. Since there are no more regime shifts (and the estimation of $b$ is independent of scale), the random draws for GI and GM now only need to take the different autocorrelations into account (as they are reported in Table 1).

Table 2 shows the results for the null hypothesis of normality, i.e., the (autocorrelated) random numbers in the $C=5000$ MC samples of the empirical size $T$ are drawn from the normal distribution with $b=2$. Computing the shape parameter for each of them, the 5%-quantile ‘crit. $b’ of the distribution $\{b^c\}_{c=1}^C$ can serve as a criterion to reject normality

\(^{14}\)For a long sample 1947 – 2005 of IP, a qualitatively similar result holds true; see Fagiolo et al. (2008, p. 664, Figure 7).

\(^{15}\)It may be noted that the empirical density of IPm does not display the relatively flat slope for $z \in (1,3)$ that we find for the quarterly growth rates of IP, GDP as well as (not shown) YF.
Table 2: Estimation and evaluation of $b$ under the null of normality.

Note: The critical value ‘crit. $b$’ is the 5%-quantile of the MC experiment, where normality is rejected if $\hat{b} < \text{crit. } b$. The corresponding $p$-values are given in per cent, and the bold face figures emphasize rejection of normality.

<table>
<thead>
<tr>
<th></th>
<th>GI</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>GM</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GDP</td>
<td>YF</td>
<td>IPq</td>
<td>IPm</td>
<td>GDP</td>
<td>YF</td>
<td>IPq</td>
<td>IPm</td>
<td>GDP</td>
<td>YF</td>
<td>IPq</td>
</tr>
<tr>
<td>$\hat{b}$</td>
<td>1.91</td>
<td>1.84</td>
<td>1.18</td>
<td>1.35</td>
<td>1.29</td>
<td>1.40</td>
<td>1.11</td>
<td>1.65</td>
<td>1.46</td>
<td>1.46</td>
<td>1.46</td>
</tr>
<tr>
<td>crit. $b$</td>
<td>1.46</td>
<td>1.46</td>
<td>1.46</td>
<td>1.63</td>
<td>1.46</td>
<td>1.46</td>
<td>1.45</td>
<td>1.63</td>
<td>1.46</td>
<td>1.46</td>
<td>1.46</td>
</tr>
<tr>
<td>$p$-value</td>
<td>37.29</td>
<td>31.42</td>
<td>0.37</td>
<td>0.04</td>
<td>1.02</td>
<td>3.22</td>
<td>0.08</td>
<td>6.05</td>
<td>37.29</td>
<td>31.42</td>
<td>0.37</td>
</tr>
</tbody>
</table>

in a one-sided test: the recipe being ‘reject’ if the empirical $\hat{b}$ in the first row of the table falls short of ‘crit. $b$’. We note in passing that, of course, the critical values increase with the sample size, but they are practically independent of the underlying autocorrelation.

The table confirms that the two high estimates $\hat{b} = 1.91$ and $\hat{b} = 1.84$ in GI for GDP and firm output YF cannot be told apart from a normal distribution. By contrast, normality is definitely rejected for the same series in GM. The corresponding $p$-values are given in the last row of the table, where their interpretation is analogous to the one in the previous section: they are the probability of being wrong when rejecting normality after observing $\hat{b}$.

We can learn from these opposite results for GI and GM that a moderation in the quarterly output volatility is not necessarily tantamount to a lower fatness in the tails of the growth rate distribution—if reference is made to the standardized growth rates, that is. Remarkably, the conclusion for monthly industrial production is almost the other way around; for these growth rates normality is clearly rejected in GI but not quite in GM. Only quarterly IP with a very strong rejection of normality shows the same behaviour in GI and GM.

While a rejection of normality in Table 2 does not necessarily imply that these distributions are close to the Laplacian, the lowest values of $\hat{b}$ suggest that here a null hypothesis $b = 1$ cannot be rejected. The question is easy to decide with $C = 5000$ random samples of size $T$ from the Laplace distribution and comparisons of $\hat{b}$ to the 95%-quantile of the resulting MC distribution $\{b^c\}_{c=1}^C$. The latter quantiles are the values ‘crit. $b$’ in Table 3, and the Laplacian is rejected if $\hat{b} > \text{crit. } b$. Unlike their counterparts in Table 2, the values crit. $b$ are not practically identical for samples of the same size $T$. The not inconsiderable differences across the same periods are explained by the underlying autocorrelations $\rho$, where a distribution $\{b^c\}$ is the wider the higher the autocorrelation (cf. the values of $\rho$
Table 3: Evaluation of $\hat{b}$ under the null of the Laplace distribution.

Note: The critical value ‘crit. $b$’ is the 95%-quantile of the MC experiment, where the Laplacian is rejected if $\hat{b} > \text{crit. } b$. The corresponding $p$-values are given in per cent, and the bold face figures emphasize non-rejection of the Laplace distribution.

<table>
<thead>
<tr>
<th></th>
<th>GI</th>
<th>GM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{b}$</td>
<td>1.91 1.84 1.18 1.35</td>
<td>1.29 1.40 1.11 1.65</td>
</tr>
<tr>
<td>crit. $b$</td>
<td>1.67 1.73 2.03 1.52</td>
<td>1.65 1.69 1.76 1.25</td>
</tr>
<tr>
<td>p-value</td>
<td>1.48 2.84 73.87 21.81</td>
<td>28.46 20.50 64.78 0.00</td>
</tr>
</tbody>
</table>

in Table 1). For IPq in GI with $\rho = 0.47$, the 95%-quantile is even as high as $b = 2.03$.

The results for this MC experiment in Table 3 speak for themselves. Consistently, the Laplacian is rejected for the three series for which normality was not rejected (GDP and YF in GI and IPm in GM). On the other hand, all five definitely non-normal growth rate samples are well compatible with the Laplace distribution. This is emphasized by the bold face figures in Table 2 and 3. Nevertheless, this nice support for a pronounced non-Gaussian behaviour only holds for a subset of our eight samples.

5. Bias and precision of the estimations

Model builders with an ambition for possible non-normality properties would like to have some guidance about a reasonable tolerance with which their models may miss the empirically estimated shape $\hat{b}$. In other words, they will be interested in a recipe for a confidence interval within which a model-implied shape parameter should preferably be contained, or in checking different proposals for this concept. This is what the present section is concerned with.

Let us begin with the obvious idea of employing another set of MC experiments for this purpose, the only difference being that instead of an hypothesized value $b = 1$ or $b = 2$ they are now based on the empirical estimate $\hat{b}$ itself. A 95% confidence interval for a given output series is then simply determined by the 2.5% and 97.5% quantiles of the resulting distribution $\{b^c\}$, which in order to make its dependence on $\hat{b}$ explicit may also be denoted MC($\hat{b}$).\(^{16}\) For such an interval to make sense, it should be centred around $\hat{b}$. However, as it is readily checked by referring to the median of MC($\hat{b}$), this condition

\(^{16}\)It is understood that the length $T$ of the single MC samples and their autocorrelation $\rho$ derive from the underlying empirical sample period.
cannot be taken for granted. As a matter of fact, the second row in Table 4 shows that the median comes out consistently higher than \( \hat{b} \).

It can thus be said that the re-estimations in the experiments exhibit an upward bias—a bias, to be precise, not with respect to a true data generation process but with respect to our autocorrelated random samples, the device on which we have decided to generate artificial data. The impression from a first quick glance at the third row in the table is that the lower the estimate \( \hat{b} \), the stronger the bias, i.e. the difference between the median of the simulated MC(\( \hat{b} \)) and \( \hat{b} \). At a second look a counterexample may be noticed. It is given by the monthly series IPm in GI with \( \hat{b} = 1.348 \) versus GDP in GM with the lower parameter \( \hat{b} = 1.295 \), where the latter, despite the shorter sample size of the quarterly frequency, implies a weaker bias. The reason for this is the much higher autocorrelation of IPm, \( \rho = 0.40 \) versus \( \rho = 0.23 \) (see Table 1). A battery of additional simulations indeed verify that the bias increases with lower estimates \( \hat{b} \) as well as with higher autocorrelations (at least if the latter are higher than 0.25, say).\(^{17}\)

<table>
<thead>
<tr>
<th></th>
<th>GI</th>
<th>GM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GDP</td>
<td>YF</td>
</tr>
<tr>
<td>( \hat{b} )</td>
<td>1.912</td>
<td>1.843</td>
</tr>
<tr>
<td>Median MC(( \hat{b} ))</td>
<td>1.992</td>
<td>1.936</td>
</tr>
<tr>
<td>Bias</td>
<td>0.081</td>
<td>0.093</td>
</tr>
<tr>
<td>( \hat{b}_{corr} )</td>
<td>1.831</td>
<td>1.751</td>
</tr>
<tr>
<td>Med. MC(( \hat{b}_{corr} ))</td>
<td>1.919</td>
<td>1.861</td>
</tr>
<tr>
<td>CI(_{0.10})</td>
<td>1.352</td>
<td>1.314</td>
</tr>
<tr>
<td>as. SE(( \hat{b} ))</td>
<td>3.000(^b)</td>
<td>2.975</td>
</tr>
<tr>
<td>s-s. SE(( \hat{b}_{corr} ))</td>
<td>0.429</td>
<td>0.409</td>
</tr>
</tbody>
</table>

**Table 4:** Bias and precision of the estimates of \( b \).

*Note:* a) These values of \( \hat{b}_{corr} \) are chosen such that Median MC(\( \hat{b}_{corr} \)) \( \approx \hat{b} \). b) Re-estimates of \( b \) are truncated at \( \hat{b} = 3.00 \). Confidence intervals are given by (3).

\(^{17}\) Fagiolo et al. (2008, p. 643) note the possibility of an upward bias but do not find it relevant for their estimations. This can be explained by their ARMA filtering of the data, which removes the serial correlation.
Because of the systematic overestimation, the envisaged confidence intervals cannot be derived from \( MC(\hat{b}) \). A straightforward device is a correction of \( \hat{b} \) for the bias and to repeat the MC experiments with EPDs based on \( \hat{b}_{corr} := (\hat{b} - \text{bias}) \) instead of \( \hat{b} \), the values of which are noted down in the fourth row of Table 4. The fifth row reports the median of \( MC(\hat{b}_{corr}) \) and confirms the expectation that the original \( \hat{b} \) is in the centre of this MC distribution. Here it has, however, to be added that for IPq in GI and GM with their low estimates of \( b \), medians of \( MC(\hat{b} - \text{bias}) \) are obtained that still exceed \( \hat{b} \) considerably. So even lower values of \( b \) were tried until the median of the resulting MC distribution was sufficiently close to \( \hat{b} \). In this way we arrive at \( \hat{b}_{corr} = 0.300 \) and \( \hat{b}_{corr} = 0.500 \) for GI and GM, respectively.

After consistency of the re-estimations has thus been established, we can use the standard percentile intervals as our confidence intervals. Formally, if \( Q_\alpha(b) \) denotes the \( \alpha \)-quantile of an MC distribution \( \{b_c\}_{c=1}^{5000} \) of re-estimates of simulated random samples based on \( b \) (and the suitable values of \( T \) and \( \rho \), of course), the corresponding confidence interval at a significance level \( \alpha \) is given by

\[
\text{CI}_\alpha = [Q_{\alpha/2}(\hat{b}_{corr}), Q_{1-\alpha/2}(\hat{b}_{corr})]
\] (3)

Unfortunately, these intervals are typically fairly wide. In documenting an interval in Table 4 we therefore choose a level \( \alpha = 0.10 \), but even then several of the intervals are wider than 1.00, the difference between the shape parameters of the normal and Laplacian density.

It will also be observed that the confidence intervals are asymmetric, that is, the difference between the upper bound and \( \hat{b} \) (or the median of \( MC(\hat{b}_{corr}) \), for that matter) is larger than between the lower bound and \( \hat{b} \). This can be explained by the following feature, which refers to the standardized EPDs \( f = f(z; b) \) and is easily verified. Let \( z_o \) be relatively large, \( z_o \geq 3 \) say. Furthermore let \( z_1, z_2 \) be symmetric around this \( z_o; z_1 < z_o < z_2, z_o - z_1 = z_2 - z_o \), and let \( b_o, b_1, b_2 \) bring about \( f(z_1; b_1) = f(z_2; b_2) = f(z_o; b_o) \). Then we do not only have \( b_1 > b_o > b_2 \), but \( b_2 \) as the indicator of the fatter tail is also closer to \( b_o \) than the higher value of \( b_1 \) indicating the thinner tail, \( b_o - b_2 < b_1 - b_o \).

If \( z_o \) is viewed as representative of the more extreme empirical data, and \( z_1 \) and \( z_2 \) as equally wide and equally probable deviations from \( z_o \) in two MC samples, then for the estimates \( b_o, b_1, b_2 \) of the corresponding shape parameters this induces a tendency that \( (b_1 - b_o) \) is larger than \( (b_o - b_2) \), implying that a confidence interval for \( b \) will be wider to the right than to the left.

The MC experiments demonstrate that the common symmetric Cramér-Rao intervals \( \hat{b} \pm 2\hat{\sigma} \) (where \( \hat{\sigma} \) is an estimated standard error) can be easily misleading,\(^{18}\) so that there would be no need to compute a standard error at all. The concept may, however,
become useful when one wants to estimate a model by the method of simulated moments (MSM). This approach refers to a given set of summary statistics, or ‘moments’, and seeks to minimize the distance between the model-generated and empirical moments. In a simple setting this means that the single moment deviations are normalized, squared, and then added up, and it is this type of a loss function that is to be minimized. Usually, normalization is done by dividing the deviations by the standard errors of the corresponding moments; in other words, the loss function is here just the sum of the squared $t$-statistics of the moments.

Certainly, the shape parameter of the output growth rates could be one of these moments. Since it is readily available one can here make use of the asymptotic standard error, which reads,

$$\text{as. } \text{SE}(\hat{b}) = \sqrt{\text{Var}(\hat{b}) / T}, \quad \text{where } \text{Var}(\hat{b}) = \frac{\hat{b}^3}{(1+1/\hat{b}) \Psi'(1+1/\hat{b}) - 1}$$

($\Psi'(\cdot)$ is the trigamma function, i.e. the second derivative of the logarithm of the Gamma function; see Agró, 1995, pp. 524f; Bottazzi and Secchi, 2008, p. 5). As it should be, $\text{Var}(\hat{b})$ is independent of the location and scale of the distribution. On the other hand, the variance changes with the level of the estimate. While the denominator in (4) is rising with $\hat{b}$, the increase in the numerator is stronger. Hence the more normal the distribution, so to speak, the higher the variance. These variations are sizeable. For example, $\hat{b}=1$ and $\hat{b}=2$ give rise to a variance of 3.45 and 19.89, respectively, meaning that the standard error more than doubles.

The second-last row of Table 4 reports the different values of the asymptotic standard errors. Even the lowest estimates of $b$ for IPq in GI and GM give rise to Cramér-Rao intervals that are wider than $2 \times 2 \times 0.2 = 0.8$, and the interval for the monthly IP series over GI is still as wide as $4 \times 0.158 = 0.632$.

Not knowing the small-sample properties of the asymptotic concept, it is interesting to compare this statistic to a standard error derived from the MC distributions. However, adopting the standard deviation of MC($\hat{b}_{\text{corr}}$) for this purpose is possibly not fully appropriate since the high re-estimates of $b$ in these distributions could unduly distort it. To avoid this tendency, a narrower range of MC($\hat{b}_{\text{corr}}$) may be considered, for which we recall that the interval of the mean ± one standard deviation of a normal distribution is given by the quantiles of 15.865% and 84.135%. Our idea is that this is the range that should be relevant for evaluating the dispersion of MC($\hat{b}_{\text{corr}}$). Thus, as an alternative to as. SE($\hat{b}$), we put forward the following expression as a small-sample (s-s.) standard error (with the notation from above):

$$\text{s-s. } \text{SE}(\hat{b}_{\text{corr}}) = \frac{Q_{0.84135}(\hat{b}_{\text{corr}}) - Q_{0.15865}(\hat{b}_{\text{corr}})}{2}$$

The last row in Table 4 shows that this standard error is not dramatically different from the asymptotic one, and that one is not not systematically higher than the other. 

\textit{A priori}
one of the two standard errors may therefore be as good as the other. Furthermore, the Cramér-Rao confidence intervals of both of them are not too dissimilar in width from $ CI_{\alpha} $ in (4) with $ \alpha = 5\% $. It is so far an open question whether (small) dynamic models of the macro economy would exhibit a comparable sample variability. Actually, it is not even clear whether such models would be able at all to reproduce values of $ b $ within the wide range that has just been indicated.\textsuperscript{19}

Lastly, let us return to the loss function for MSM and the deviations of the model-generated shapes, designated $ b^m $, from the empirical shape $ \hat{b} $. After learning of the asymmetry of the MC confidence intervals, one may not wish to apply the same penalty expression to the positive and negative deviations alike. Instead of the abovementioned term $ (b^m - \hat{b})^2/(\text{std. error})^2 $ in the loss function, two different normalizations may be chosen for $ b^m < \hat{b} $ and $ b^m > \hat{b} $. Here we can take up the discussion on eq. (5) and adopt $ [Q_{0.84135}(\hat{b}_{\text{corr}}) - Q_{0.50}(\hat{b}_{\text{corr}})] $ as a standard error for positive deviations, and $ [Q_{0.50}(\hat{b}_{\text{corr}}) - Q_{0.15865}(\hat{b}_{\text{corr}})] $ as a standard error for negative deviations. Thus, based on the above MC experiments, we propose to respecify the ordinary ‘distance’ between $ b^m $ and $ \hat{b} $ as follows:

\[
\text{dist}(b^m, \hat{b}) = \begin{cases} 
\frac{(b^m - \hat{b})^2}{SE_1} & \text{if } b^m < \hat{b}; \quad SE_1 = Q_{0.50}(\hat{b}_{\text{corr}}) - Q_{0.15865}(\hat{b}_{\text{corr}}) \\
\frac{(b^m - \hat{b})^2}{SE_2} & \text{if } b^m > \hat{b}; \quad SE_2 = Q_{0.84135}(\hat{b}_{\text{corr}}) - Q_{0.50}(\hat{b}_{\text{corr}})
\end{cases}
\]

(6)

Normally, a term like this will be combined with many other moment terms (for example, with the lagged auto-covariances mentioned in the previous footnote). If it is feared that the fat-tail moment term (6) is dominated by them, the model builder can express his or her priorities by multiplying (6) with another weighting coefficient greater than one.

6. Conclusion

The point of departure of this paper were claims in the literature that “fat tails are an extremely robust stylized fact characterizing the time series of aggregate-output growth in most industrialized economies” (Fagiolo et al., 2008, p. 664). The topic is not just a struggle for bold and pithy words demanding attention, it also provides hints “in favor of an increasingly ‘non-Gaussian’ economics and econometrics” (\textit{ibid.}, p. 664). Thus, in the latter field, non-normality might undermine the common likelihood techniques that are presently dominating the estimations of macroeconomic models. In economic theory, on

\textsuperscript{19}Franke (2012) designed an ‘Old-Keynesian’ model of the new macroeconomic consensus that seeks to match the lagged auto- and cross-covariances of output, inflation and the interest rate. As a novel feature, his subsequent MSM estimation also includes the different raggedness of these series in the set of moments. This feature might somehow be related to fat tails, but a direct measure of fat tails has so far not been considered in this type of literature.
the other hand, non-normality might become an established property that an ambitious
model builder, among other things, would like to reproduce.

A most suitable approach to measure normality and deviations from it are estimations
of the shape parameter $b$ of the exponential power distribution, where the special case
$b = 2$ yields the normal distribution and lower values produce progressively fatter tails.
Somewhat curiously, even though all of the studies using this method for US output
data have long sample periods of more than 40 years in common, none of them mentions
the problem that the structural change known to be present in this data might
possibly invalidate the low estimates of $b$. A straightforward Monte Carlo experiment
could be designed to examine this issue. Simulating random draws for the two periods
of the Great Inflation (GI) and Great Moderation (GM) with their different variances of
the innovations, which are here assumed to be normally distributed, it turned out that
the pasted growth rate series (may not necessarily exhibit but) can exhibit non-normal
behaviour of the type measured in the empirical series. It follows that if one wants to
rule out that the fat-tail results in the literature are spurious, they have to be checked
for a possible influence of the structural change.

To be clear enough, the last statement refers to the original output growth rates. It was
explained why we choose to work with these rather than filtered data, where nevertheless
our Monte Carlo tests take the serial correlation into account. Because of the problem
with the long period we subsequently concentrated on each of the two subsamples of
GI and GM. A nice finding was that either the normal distribution or the Laplacian
(for which $b = 1$) are rejected by the tests, but not both. The question of whether in
this sense the normal or Laplace distribution tends to prevail is answered with an “it
depends”, namely, on the specific period and on the kind of output data. It is only for
quarterly industrial production that normality is rejected in both subsamples. For its
monthly growth rates, normality is rejected in GI but not GM (the Laplacian is rejected
for GM but not GI), whereas for the quarterly GDP and firm output data it is just the
other way around. Unfortunately, the confidence intervals around the single estimates of
$b$ that we discussed are also relatively wide.

Regarding the issue of whether non-normality of output growth rates might be consid-
ered a new stylized fact to be incorporated in macroeconomic modelling, the investigation
can thus be quickly summarized as providing “mixed evidence”. Accordingly, a researcher
might pick out that sample period or kind of data which delivers the results that he or
she prefers. Nonetheless, if the main emphasis is put on quarterly (rather than monthly)
industrial production, a more positive conclusion could be that non-normality constitutes
a weak stylized fact, a statement on which perhaps a greater number of economists may
be able agree.

In any case, following the spirit of evaluating or even estimating a model by the
method of simulated moments (as indicated in footnote 19), a more ambitious analysis
of its dynamic properties would first have to bring the issue of fat tails up for a more
rigorous discussion and see what role it plays in the model and in the real world. It was the purpose of the present paper to lay some foundations for this.

Appendix

Data sources

Real GDP was obtained from the Bureau of Economic Analysis, at http://www.bea.gov/national/index.htm#gdp, and industrial production from the Board of Governors of the Federal Reserve System, at www.federalreserve.gov/releases/G17/table1_2.htm.

The firm sector output series was extracted from the database fmdata.dat in the zip file fmfp.zip, provided by Ray Fair for working with his macroeconometric model. It is a plain textfile downloadable from http://fairemodel.econ.yale.edu/fp/fp.htm. The acronym to identify the series is ‘Y’, as explained in Appendix A.4, Table A.2., of the script Estimating How The Macroeconomy Works by R.C. Fair, January 2004, which can be downloaded from http://fairemodel.econ.yale.edu/rayfair/pdf/2003a.pdf.

Estimation of $a$ and $m$ for an exponential power distribution

Suppose that the shape parameter $b$ has already been estimated beforehand as described by eq. (2) in the main text. Setting, in a ML estimation, the partial derivative of the log-likelihood function with respect to $m$ equal to zero, $\hat{m}$ can then be obtained as the solution of the implicit equation in $m$,

$$\sum_{t=1}^{T} |x_t - m|^{b-1} \text{sgn}(x_t - m) = 0$$

where $\text{sgn}$ is the sign function, $\text{sgn}(y) = 1$ ($0$, $-1$) if $y > 0$ ($y = 0$ or $y < 0$, respectively); cf. Mineo (2003, p. 112). On this basis, we can subsequently refer to Chiodi (1988), who has proposed the following expression as an unbiased estimate of $a$,

$$\hat{a} = \left[ \frac{\sum_{t=1}^{T} |x_t - \hat{m}|^{b}}{T - \hat{b}/2} \right]^{1/b}$$

(quoted from Mineo and Ruggieri, 2005, p. 4, eq. (9)).

Random variates from EP distributions

In general, the generation of pseudo random numbers drawn from an EP distribution involves draws from a Gamma distribution, which in turn requires some computational effort (see, e.g., Zhu and Zinde-Walsh, 2009, p. 91, or Li, 2011, Section 2, which both allow for an asymmetric shape also). For the class of standardized distributions with

---

20In our view, the discussion in de Grauwe (2012) and Ascari et al. (2012) was not yet rigorous enough.
shape $b > 1$ (besides $m = 0$, $a = 1$), Chiodi (1995, Section 4) set up a faster and easy-to-implement algorithm which has the advantage that it only needs the generation of uniformly distributed random numbers. A random number $z$ is here generated in the following two stages:

1. Repeat
   
   draw $U$ and $V$ from the uniform distribution on $[-1, +1]$
   
   and put $W = |U|^b + |V|^{b/(b-1)}$;
   
   Until $W \leq 1$.

2. Return $z = U \cdot [-b \ln(W)/W]^{1/b}$.

In the polar case $b = 1$, the random deviates $z$ can be obtained as the product $z = SE$, where $S$ is a random sign and $E$ is a random draw from the standard exponential distribution (which is defined for nonnegative real numbers $x$ and has density $\exp(-x)$). The deviates $E$ are most easily computed by drawing $U$ from the uniform distribution on $(0, 1]$ and putting $E = -\ln(U)$ (see, eg., Press et al., 1986, p. 278).

For the remaining case $b < 1$, we can readily set up a rejection procedure ourselves. With respect to $b_0 = 1$ and the parameters of the standardized EPD, $a = 1$ and $m = 0$, write $f(z; b) = f(z; a, b, m)$ and $g(z) = f(z; b_0)$. Let $[z_1, z_2]$ be a symmetric and sufficiently wide interval outside of which the density $f(\cdot; b)$ is negligibly low ($z_1 = -z_2$) and set $c = f(z_1; b)/f(z_1; b_0)$. Since except for $z$ close to zero $f(z; b_0)$ has a steeper slope than $f(z; b)$ (which is easily seen after taking logs), we have $f(z; b) \leq c g(z)$ for all $z \in [z_1, z_2]$.

In this sense the density $g(\cdot)$ dominates $f(\cdot; b)$ and we can make use of the following elementary algorithm (Devroye, 1986, pp. 41f):

1. Repeat
   
   draw $U$ from the uniform distribution on $[0, 1]$;
   
   repeat
   
   draw $V$ from the Laplace density $g(\cdot) = f(\cdot; b_0)$
   
   until $z_1 \leq V \leq z_2$;
   
   put $W = c g(V)/f(V; b)$;
   
   Until $U W \leq 1$.

2. Return $z = V$.

The distance between the curves $f(\cdot; b)$ and $c g(z)$ is relatively small, so that the procedure should be sufficiently efficient (at least much more efficient than simply adopting a constant dominating density $g$; Devroye, 1986, p. 43).

Of course, the draws thus obtained are iid. To take account of an autocorrelation $\rho$ put, in round $t$, $z_t = \rho z_{t-1} + \sqrt{1 - \rho^2} \tilde{z}$, where besides $|\rho| < 1$ it is supposed that $z_{t-1}$ is a draw from the previous round and $\tilde{z}$ a draw from the EP distribution, both of them with

---

21 The procedure can still be accelerated by suitable squeeze methods, at the price of a more complicated computer code. Since the original version is already fast enough, this does not seem worth the effort.
the same variance $\sigma^2$. It is easily seen that then $\text{Var}(z_t) = \sigma^2$ and $\text{Corr}(z_t, z_{t-1}) = \rho$. It is well-known that for normal distributions, $b=2$, $z_t$ is normally distributed, too. We know of no mathematical proof that establishes the analogous statement for general values of $b$. The property can, however, be confirmed by simulation studies, even for $b$ close to one, although (very) large samples are required for a satisfactory convergence of the sample density function towards the theoretical density (the smaller $b$ or the higher $\rho$, the larger the samples).

References


Li, S. (2011): Regression models with the error term following the asymmetric exponential power distribution. *Mimeo*, Department of Economics, Rutgers University.


