Peaks or tails -
What distinguishes financial data?

by Walter Krämer and Ralf Runde[1]

Abstract

We argue against the view that it is mostly the peaks of the empirical densities of stock returns (and of other risky returns as well) that set such data aside from "normal" variables. We show that peaks depend on sample size and on the way returns are standardized, and that for given data sets of stock returns, both higher peaks and lower peaks than in a standard normal case can be obtained.

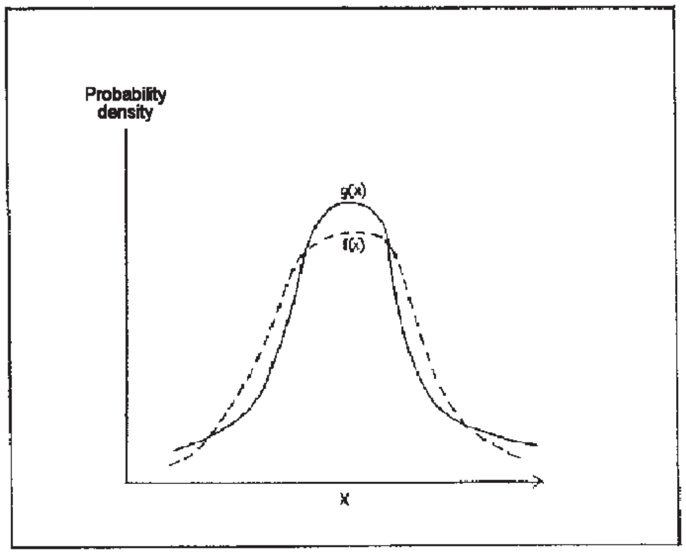
JEL-classification: C13, C14

[1]We are grateful to Deutsche Forschungsgemeinschaft (DFG) for financial support, and to Adrian Pagan for stimulating discussion and comments. Stock returns were obtained from Deutsche Finanzdatenbank (DFDB).

1 Peaks versus tails as indicators of leptocurtic stock returns

Starting with Mandelbrot (1963), the normality of changes of speculative prices, which since Bachelier (1900) had been considered a good first approximation, has been in much dispute. In particular, it is by now well known that returns on speculative assets typically exhibit higher than normal kurtosis as expressed in both higher peaks and fatter tails than can be found in normal data. Figure 1 (from Granger and Orr, 1972) is a typical diagram which has in various versions been used to illustrate this fact.

Figure 1: Typical densities of financial as compared to "normal" data



Below we argue against putting too much emphasis on peaks, and advocate to focus on the tails. As we demonstrate below, and as has long been known, the tails of a distribution are robust against the particular standardisation used, whereas the peak is not. For instance, if the second moment of the data does not exist, then the usual standardisation, i.e. dividing the returns by their empirical standard deviation, will always produce sharp peaks in the empirical densities which are getting ever sharper as the sample size increases (in the limit, all empirical probability mass is concentrated in arbitrary intervals around the origin).

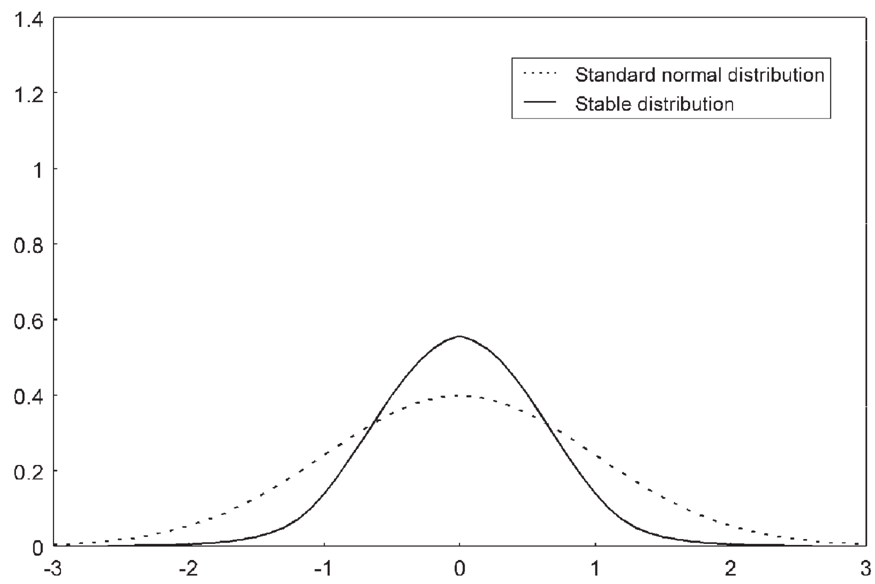
While such peaks correctly point towards deviations from normality and are therefore often correctly suggested as rough indicators of excess kurtosis (see e.g. Pagan 1994, 1996), their exact magnitude is an artifact of sample size, so a different standardisation like dividing the raw data by their interquartile or interdecile range appears preferable from a statistical point of view: as sample size increases, the empirical density then converges to a nondegenerate limit, and any finite-sample properties of the empirical density can unambiguously be viewed as estimates for some underlying population properties.

However, if standardisation is done by interquantile ranges rather than by standard deviations, many empirical return distributions have smaller peaks than the standard normal density, as we demonstrate with German stock returns below. The message is: return-histograms based on conventional standardisation of returns are a good indicator of nonnormality, but a bad estimator of the true densities of returns. If a good estimator of the return density is used, the peak in the empirical density is no longer a reliable indicator of nonnormality.

2 Normalisation with non-existing second moments

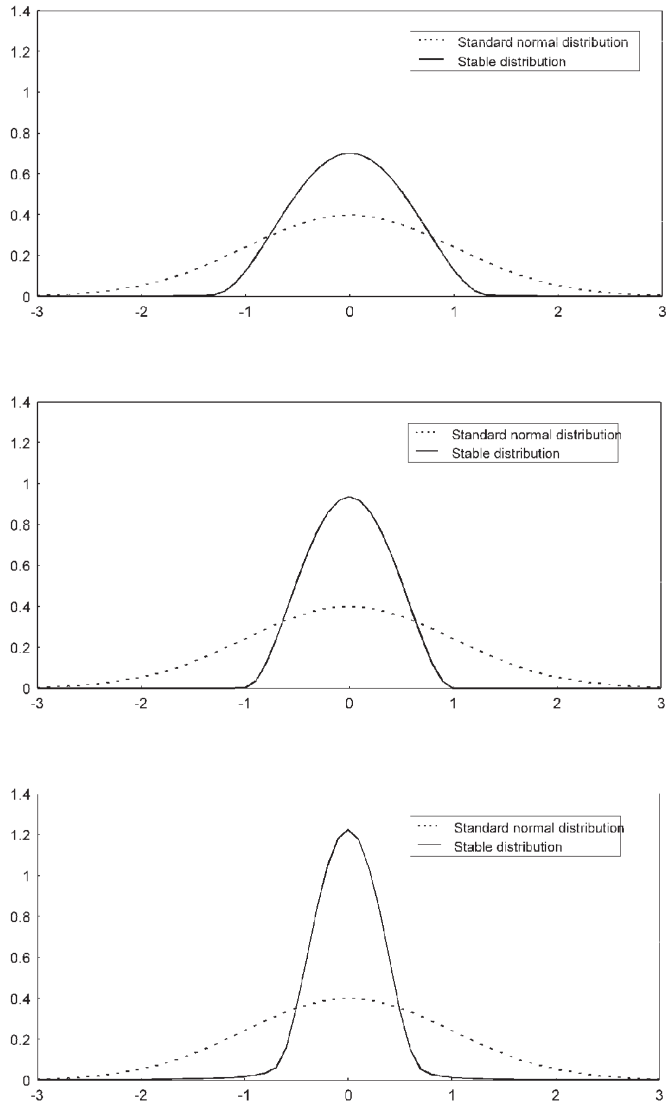
First, we consider the effect of dividing the raw data by their empirical standard deviation when the population second moment is infinite. Figure 2 shows the estimated density of $n = 100$ independent symmetric stable random variables, $X_1, \ldots, X_{100}$, with location parameter 0, scale parameter 1 and characteristic exponent (tail index) $\alpha = 1.5$ (the stable variates were generated along the lines of Chambers et al., 1976). With a sample of that size, the peak of the empirical density is 0.557 - slightly higher than that of the standard normal distribution. The density estimate was obtained by averaging 500 empirical densities, each obtained by a non-parametric kernel-estimate with a biweight-kernel and a bandwidth of 0.7 (for details see e.g. Silverman, 1986).

Figure 2: Standard normal distribution versus standardized empirical stable distribution with sample size n = 100



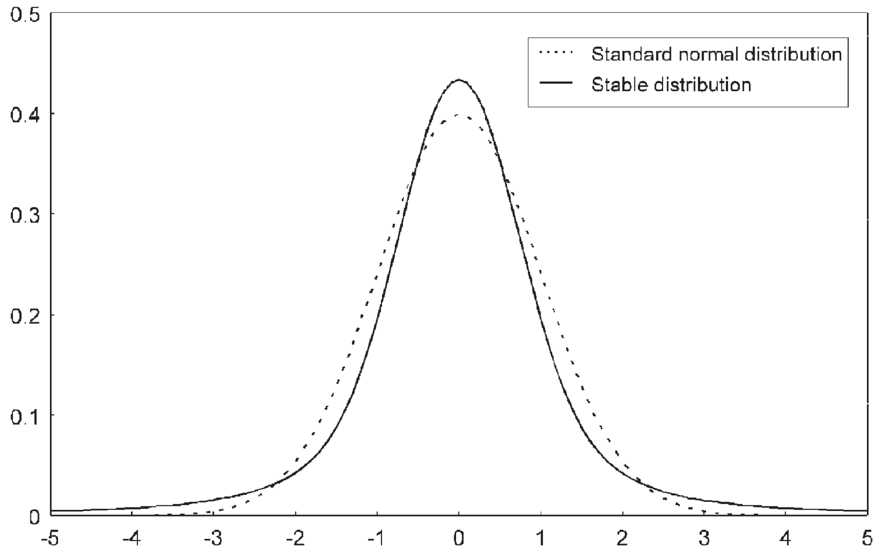
The peak increases rapidly as sample size increases (figure 3), surpassing the standard normal peak by ever wider margins.

Figure 3: Standard normal distribution versus standardized empirical stable distribution with sample size n = 1000, 5000 and 10000



However, quite a different scenario emerges when standardisation is done differently. Figure 4 shows the corresponding empirical density of independent symmetric stable random variables for sample size 10000, where the variates were divided by their empirical interquartile range rather than by their empirical standard deviation. Contrary to figure 3, the peak is now only slightly above the normal peak, and at the same hight as for smaller sample sizes (the respective figures for $n = 100$, $n = 1000$ and $n = 5000$ are virtually identical to figure 4 and are therefore not explicitely given here).

Figure 4: Standard normal distribution versus standardized stable distribution, n = 10000



3 Application to German stock returns

Next we consider different standardisations of "real" data. Figures 5, 6 and 7 show the empirical density of $n = 7993$ daily returns of the German stock price index DAX, from Jan. 4, 1960 to Jan. 31, 1992. In figure 5, returns are standardised by the empirical standard deviation, in figure 6, returns are standardized to have an interdecile range of 2.564 and in figure 7, returns are standardized to have an interquartile range of 1.35 (the interdecile and interquartile ranges, respectively, of the standard normal distribution).

Figure 5: Standard normal distribution versus DAX returns (normalisation with empirical standard deviation)

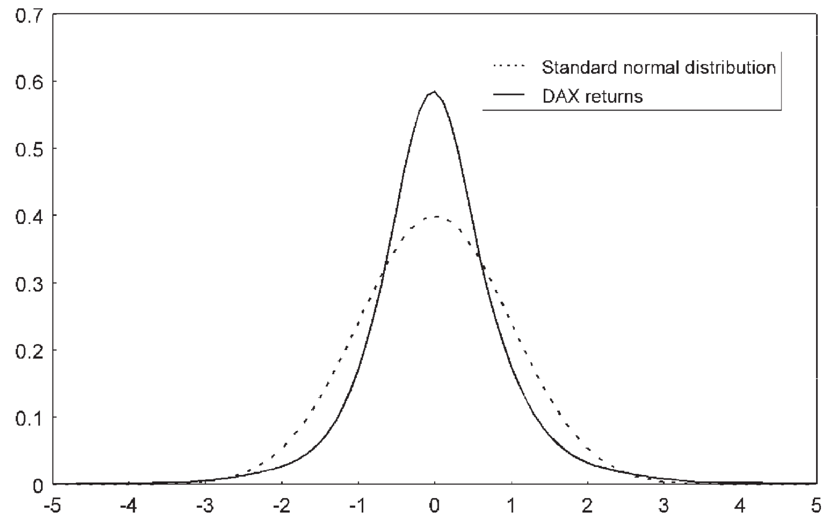
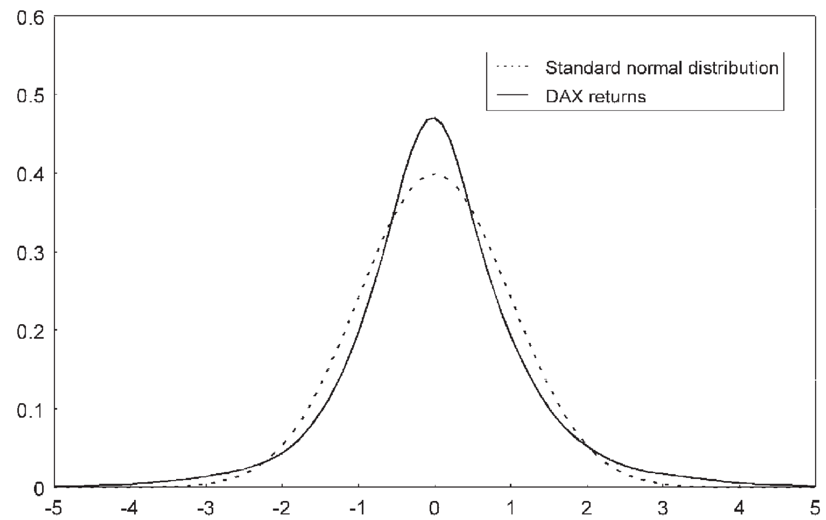
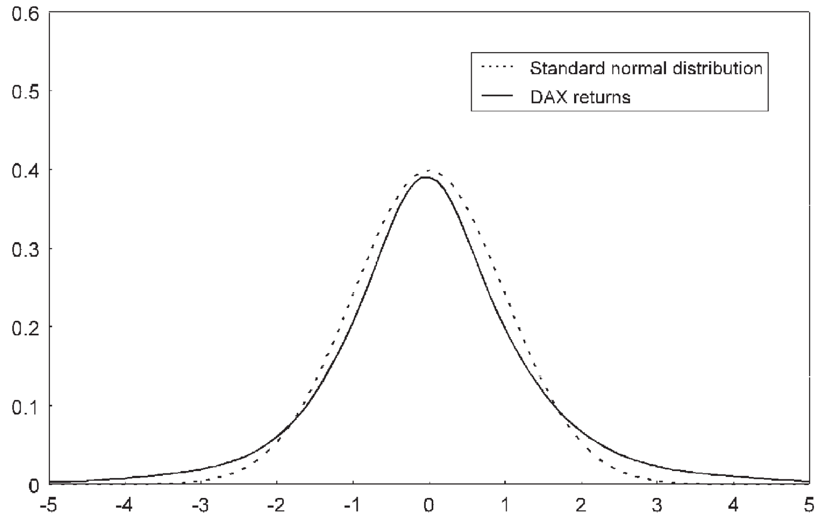


Figure 6: Standard normal distribution versus DAX returns (normalisation with interdecile range)



The figures show that higher-than-normal peaks are an artifact of the standardisation procedure: The conventional standardisation based on sample standard deviation produces a peak well above the normal (figure 5), but when the interdecile range is used for standardisation, the peak is getting smaller (figure 6), and it falls below the peak of the standard normal when the interquartile range is used (figure 7).

Figure 7: Standard normal distribution versus DAX returns (normalisation with interquartile range)



4 Conclusion

This note has shown that the peaks of empirical densities of stock returns are as much a function of the particular standardisation procedure as of the data itself, and that judging "perverse return behaviour" by the peaks is only valid when returns are standardized by empirical standard deviations.

References

Bachelier, L. (1900): "Théorie de la spéculation." *Annales de l'École Normale Supérieure* 17, Ser. 3, 21 - 86.

Chambers, B.W.; Mallows, C.L. and **Stuck, B.W.** (1976): "A method for simulating stable random variables." *Journal of the American Statistical Association* 71, 340 - 344.

Granger, C. and **Orr, D.** (1972): "Infinite variance and research strategy in time series analysis." *Journal of the American Statistical Association* 67, 275 - 285.

Mandelbrot, B. (1963): "The variation of certain speculative prices." *Journal of Business* 36, 394 - 419.

Pagan, A. (1993): Discussion of D.J. Duffie and R. Engle: "Finance and ARCH", *ESEM* 1994, Uppsala.

Pagan, A. (1996): "The econometrics of financial markets." *Journal of Empirical Finance*, Vol. 3, No. 1 ....

Silverman, B.W. (1986): *Density estimation for statistics and data analysis*, Chapman and Hall, London.