TI 2001-070/2
Tinbergen Institute Discussion Paper

Portfolio Diversification Effects and Regular Variation in Financial Data

Namwon Hyung
Casper G. de Vries

Department of General Economics, Faculty of Economics, Erasmus University Rotterdam, Eurandom, and Tinbergen Institute

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Keizersgracht 482
1017 EG Amsterdam
The Netherlands
Tel.: +31.(0)20.5513500
Fax: +31.(0)20.5513555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31.(0)10.4088900
Fax: +31.(0)10.4089031

Portfolio Diversification Effects and Regular Variation in Financial Data

Namwon Hyung and Casper G. de Vries
Tinbergen Institute, Erasmus Universiteit Rotterdam and EURANDOM

June 21, 2001

Abstract

Portfolio risk is in an important way driven by 'abnormal' returns emanating from heavy tailed distributed asset returns. The theory of regular variation and extreme values provides a model for this feature of financial data. We first review this theory and subsequently study the problem of portfolio diversification in particular. We show that if the portfolio asset return distributions are regulary varying at infinity, then Feller's convolution theorem implies that the portfolio diversification is more effective than if the underlying distribution would be normal. This is illustrated by a simulation study and an application to S&P stock returns.

1 Introduction

Risk analysis and management are the bread and butter of the financial industry. For example, reinsurance companies are in the business of spreading the risk of insurance contracts, and risk managers at commercial banks perform a daily measurement of the tail risk of their trading portfolios, known as the Value-at-Risk (VaR) exercise. The portfolio structure influences the portfolio return performance and its risk characteristics. One can reduce the portfolio risk, if defined as the standard deviation, through diversification by virtue of the law of large numbers, since returns on individual assets are imperfectly correlated with each other. Apart from the standard deviation, the financial industry often employs so called downside risk measures to further characterize the asset risk, since it is widely recognized that large losses are more frequent than a normal distribution based statistic like the standard deviation suggests. Thus portfolio risk is in an important way driven by the 'abnormal' returns emanating from heavy tailed distributed asset returns. The theory of regular variation and extreme values provides a model for understanding this typical feature of financial data.

We start our essay by reviewing the theory on distributions with heavy tails. Subsequently, we study one diversification problem in more detail. The question we ask is how effective portfolio diversification can be for managing the downside risk. We show that if the portfolio asset return distributions are regularly varying at infinity, then Feller's convolution theorem implies that the portfolio diversification is more effective than if the underlying distribution were normally distributed. Under normality one can use a 'square-root rule' to calculate the effects of portfolio diversification. As it turns out, with heavy tailed distributed returns, a modified root-rule can be formulated. The results for the downside risk of portfolio diversification are illustrated by a simulation study. An application to S&P stock returns demonstrates the empirical relevance of the theory.

2 Review of Regular Variation and Extreme Value Theory

It is well known that many financial time series exhibit a fat-tailed return distribution. For example, this shows up in the higher than normal kurtosis of stock returns. There are several definitions of a "heavy tailed" distribution. One definition would be to take all distributions which generate a kurtosis above three. But this class is too broad for what we see in the data. It is not only the higher than normal kurtosis, but sequential moment plots typically reveal failure of the higher moments, see Embrechts et al. (1997, pp.311-315). Therefore we define the heavy tail feature more finely as the class of distributions which exhibit a power like behavior comparable to the Pareto distribution. The distributions in this class all have some of the higher moments unbounded.

Throughout the paper we define the return on an asset as the logarithmic price differential. Thus if P_t is the time t asset price, then $X_t = \ln(P_t/P_{t-1})$ is the one period return. Suppose $X_1, X_2, \ldots$ is a sequence of independent and identically distributed (i.i.d.) random variables with a common distribution function (d.f.) $F(x)$. We identify the class of distributions with heavy tails, as the distributions which are regularly varying at infinity:

Definition 1 *A distribution $F(x)$ varies regularly at minus infinity with tail index α if*

$$\lim_{t\to\infty} \frac{F(-tx)}{F(-t)} = x^{-\alpha},\ \alpha > 0,\ x > 0. \tag{1}$$

To interpret the property of regular variation, we review a number of its implications. One important implication of regular variation is the connection between heavy tails and the distribution of the extreme order statistics. Regular variation is a necessary and sufficient condition for the distribution of the minimum or maximum to be in the domain of the attraction of the Frechet limit law (extreme value distribution). Let $M_n = \max\{-X_1, -X_2, \ldots, -X_n\}$, i.e. M_n is the maximum loss in a sample of size n. The distribution function of the maximum loss M_n is

$$P\{M_n \le x\} = [1 - F(-x)]^n,\ x > 0. \tag{2}$$

Extreme value theory studies the limiting distribution of M_n (appropriately normalized) as n becomes large, while regular variation characterizes the tails as x becomes large. For the laws which satisfy (1) and a choice of the normalizing constant $a_n = -F^{-1}(1/n) > 0$, we have for loss x that as $n \to \infty$

$$\left[1 - F(-a_n^{-1}x)\right]^n \xrightarrow{w} G(x) = \exp\left(-x^{-\alpha}\right), \ \alpha > 0, x > 0, \tag{3}$$

where $\xrightarrow{w}$ denotes weak convergence. Note that (1) and the definition of a_n imply

$$\left[1 - F(-a_n^{-1}x)\right]^n = \left[1 - F(-a_n^{-1})\frac{F(-a_n^{-1}x)}{F(-a_n^{-1})}\right]^n \approx \left[1 - \frac{x^{-\alpha}}{n}\right]^n,$$

which explains the convergence towards the Frechet limit law $\exp(-x^{-\alpha})$. One can also show that starting from the limit law, the underlying distribution must be regularly varying at infinity (see e.g. Leadbetter et. al (1983) for details).

A second (indirect) implication of regular variation is that, to a first order approximation, all distributions have a tail comparable with the Pareto distribution:

$$F(-x) = Ax^{-\alpha}[1 + o(1)], \ \alpha > 0, \ A > 0, \tag{4}$$

as $x \to \infty$. This follows from the fact that all regularly varying distributions are in the domain of attraction of the Frechet limit law for the distribution of the maximum. Since this is also the case for the Pareto distribution $F(-x) = Ax^{-\alpha}$, all these distributions satisfy (4).

A third implication of regular variation is that the tail parts do not decay rapidly enough for the moments above α to be integrable. From (4) for large x

$$f(-x) \approx \alpha A x^{-\alpha-1},$$

so that the density declines at a power rate $x^{-\alpha-1}$ far to the left of the center of the distribution. This power is outweighed by the explosion of x^m in the computation of moments $m > \alpha$. There are other distributions which exhibit unbounded moments, but the regular variation requirement imposes a certain smoothness, such that eventually the tail shape is monotonic.

A fourth implication of regular variation is the simplicity of tail probabilities for convoluted data. Suppose the data $\{X_i\}$ are generated by a heavy tailed distribution which satisfies (4). Evidently (4) satisfies (1). From the Feller's Theorem (1971, VIII.8), the distribution of the sum satisfies

$$\Pr\left\{\sum_{i=1}^{k} X_i \le -x\right\} = kAx^{-\alpha}[1 + o(1)], \text{ as } x \to \infty. \tag{5}$$

Thus the multiperiod return distribution, recall the additivity of the logarithmic returns, has a tail shape which is identical to the tail shape of the one period returns, but which is scaled up by the length of the return horizon k. Risk managers in financial institutions often calculate the VaR over a one day investment horizon as the loss quantile x for which $F(-x) = \delta$, for a given desired risk level δ. If based on the expansion (4), the VaR is well approximated by $A^{1/\alpha}\delta^{1/\alpha}$ for sufficiently small δ. For regulatory purposes, the VaR must often be calculated over a multiperiod investment horizon apart from the one day exercise. The reason is that regulators are concerned about the impossibility of being able to immediately execute sales for large parts of the portfolio under adverse market conditions. Thus regulators are interested in the portfolio risk over a multiday investment period. If one relies on (5), the k-period VaR easily follows from the one period VaR, holding the risk level δ constant, by scaling the one period VaR with the factor $k^{1/\alpha}$. Thus a bank can go from the high frequency estimate to the low frequency estimate without having to reestimate the parameters by using this 'α-root rule'. If the normal model would be used, the appropriate scaling is by means of the standard 'square-root rule'. Assume that returns have bounded second moments. We have the following result from de Vries (2000): *At a constant risk level δ, increasing the time horizon k, increases the VaR estimate under the normal model by more, i.e. by $k^{1/2}$, than under the fat tail model, where the increase is only a factor $k^{1/\alpha}$, since $\alpha > 2$.*

We now briefly review the statistical procedures required for using heavy tailed distributions. An excellent treatise of extreme value theory is given in Embrechts et. al. (1997) who cover both the probability theory and the statistical matters. The primary objective of the estimation is to obtain quantiles x, denoted as VaR, at a given risk level δ. It is common to require estimates of the VaR not only near the boundary of the range of observed data, but also beyond the boundary, where the empirical distribution function is of no avail. Consider two loss probabilities δ and t with $\delta < \frac{1}{n} < t$, where n is sample size. Then from (4) the quantile x_δ associated with δ follows as below

$$x_\delta = x_t \left(\frac{t}{\delta}\right)^{1/\alpha} \left(\frac{1 + o\left(x_t^{-\alpha}\right)}{1 + o\left(x_\delta^{-\alpha}\right)}\right)^{1/\alpha}.$$

This suggests the following estimator: Ignore the higher order terms in the expansion, replace the probability t by the empirical distribution $\frac{m}{n}$, replace x_t by the $(m+1)$-th descending order statistic, and use an estimator $\widehat{1/\alpha}$ for the inverse of the tail index $1/\alpha$

$$\hat{x}_\delta = X_{(m+1)} \left(\frac{m}{n\delta}\right)^{\widehat{1/\alpha}}. \tag{6}$$

There are two random variables in this statistic: $X_{(m+1)}$ and $\widehat{1/\alpha}$. Since the latter appears in the power, this statistic drives the properties of the VaR estimator as long as $m/n \to 0$ as $n \to \infty$. For this reason we can concentrate on the properties of the tail index estimator alone.

The most popular estimator are Hill's (1975) estimators of the tail index α

and the scale A:

$$\widehat{\alpha} = \left(\frac{1}{m} \sum_{i=1}^{m} \ln \left(X_{(i)} \right) - \ln \left(X_{(m+1)} \right) \right)^{-1}, \tag{7}$$

and

$$\hat{A} = \frac{m}{n} (X_{(m+1)})^{\hat{\alpha}}, \tag{8}$$

where $X_{(i)}$ is the i-th descending order statistic. These estimators can be interpreted as conditional maximum likelihood estimators. If in (4) there were no higher order terms, so that the distribution is a Pareto distribution, then the Hill estimator (7) is the maximum likelihood estimator. With second order terms present, the Pareto approximation is only good in the tail area. Hence the statistic is calculated on a subset of the most extreme order statistics. The difficulty then is the determination of this subset. The choice of the sample threshold m is not a trivial problem, which was only recently solved by means of a bootstrap procedure, see Danielson et. al. (2000). In the bootstrap the asymptotic mean squared error is minimized via a control variate type device. To guarantee convergence in distribution care must be taken in bootstrapping with subsamples sizes only[1].

3 Diversification Effects and Simulation Study

In the previous section we showed that if the data $\{X_i\}$ are generated by a heavy tailed distribution for which

$$\Pr \{X_i \leq -x\} = Ax^{-\alpha}[1 + o(1)],\ \alpha > 0,\ A > 0, \tag{9}$$

as $x \to \infty$, then for the convolution

$$\Pr \left\{ \sum_{i=1}^{k} X_i \leq -x \right\} = kAx^{-\alpha}[1 + o(1)], \text{ as } x \to \infty. \tag{10}$$

We now use Feller's theorem for deriving the benefits from cross-sectional portfolio diversification. Suppose the following one factor model applies

$$R_i = \beta_i R + Q_i, \tag{11}$$

where R is the return on the market portfolio and Q_i is the idiosyncratic risk of the return on asset i. This is the well known Capital Asset Pricing Model from

[1] Related to this is the reason for subsampling in endpoint estimation, see Shao and Tu (1995, pp.123-124); for a continuous d.f. the largest observation is by necessity a downward biased estimate of the endpoint. This bias can be corrected via a subsample bootstrap procedure, since a comparison between the subsample resample endpoints and the full sample endpoint allows one to gauge the extent of the bias.

finance. Consider a portfolio of k assets with weights w_i, $w_i > 0, \sum_{i=1}^{k} w_i = 1$. We focus on equally weighted portfolios $w_i = 1/k$. Let $\bar{\beta} = \frac{1}{k}\sum_{i=1}^{k} \beta_i$. Suppose the Q_i are cross-sectionally *i.i.d.* distributed and satisfy (9). Conditional on the market return, using (9), the downside risk of a single asset portfolio is

$$\begin{aligned}
\Pr\{R_i \leq -x | R\} &= \Pr\{Q_i + \beta_i R \leq -x | R\} \\
&= \Pr\{Q_i \leq -x - \beta_i R | R\} \\
&= A(x + \beta_i R)^{-\alpha}[1 + o(1)]. \qquad (12)
\end{aligned}$$

The diversification benefits from the equally weighted portfolio against the downside risk then follow as

$$\begin{aligned}
\Pr\left\{ \frac{1}{k}\sum_{i=1}^{k} R_i \leq -x \middle| R \right\} &= \Pr\left\{ \sum_{i=1}^{k} Q_i \leq k(-x - \bar{\beta}R) \middle| R \right\} \\
&= kA(k[x + \bar{\beta}R])^{-\alpha}[1 + o(1)] \\
&= k^{1-\alpha}A(x + \bar{\beta}R)^{-\alpha}[1 + o(1)] \qquad (13)
\end{aligned}$$

Thus diversification reduces the downside risk at a fixed VaR level whenever $\alpha > 1$ (assuming $\bar{\beta} \approx \beta_i$). If $\alpha > 2$, then diversification is working even better than if the returns would be normally distributed.

We investigate the diversification effects further by means of a simulation study. In this study we examine the diversification benefit under the Student-t distribution. For this distribution the degrees of freedom equals the tail index. The Student-t was simulated with respectively degrees of freedom $d = 1, 2, 3$ and 6. For the fixed levels of quantiles, we calculate the loss probabilities at different levels of aggregation. The number of observations was taken 100,000 and each experiment was repeated 1,000 times. The results from the Monte Carlo experiments are reported in Tables 1 and 2.

In Table 1 we report the empirical loss probabilities for fixed quantiles. The chosen quantiles are respectively 5%, 1%, 0.5% and 0.05% losses for the non-aggregated series. The values in the first row of Table 1 are reasonably close to the pre-set probability levels. We do not report the standard errors of the simulation to save on space. The probabilities are decreasing monotonically as the level of aggregation k increases. By the law of large numbers, as more series are averaged the distribution is degenerating to the sample mean and hence the probabilities beyond a certain quantile are all decreasing. The effect of diversification is less pronounced when the degrees of freedom d is smaller. For example comparing $k = 1$ to $k = 10$, the loss probability is decreasing from 5% to 0.684% if $d = 2$, but to 0.002% if $d = 6$. Fama and Miller (1972, p. 270) discuss the case of sum stable distributions, and note that for $\alpha < 1$ diversification actually increases the dispersion. For the hairline case $d = 1$, diversification does not help but does not hurt either. This explains the invariance in the first four columns from the Table 1.

To experiment for the case where one does not know the true data generating mechanism, we compare the aggregation effects calculated from two proxy models, when the true data are generated by the Student law. The first proxy model uses the semi-parametric presumption that the distribution is heavy tailed, the second model is based on the fully-parametric assumption of normality. In case one works from the presumption of regular variation, we have from (10) the following first order approximation to the diversification effect

$$P\left\{\frac{1}{k}\sum_{i=1}^{k}X_i \leq -x\right\} \approx k^{1-\alpha}Ax^{-\alpha}. \tag{14}$$

By means of a Taylor expansion at infinity, one calculates the following parameter values for the Student law $\alpha = d$ and $A = \frac{\Gamma((d+1)/2)}{\Gamma(d/2)}d^{\frac{d-2}{2}}/\sqrt{\pi}$ and where $\Gamma(\cdot)$ denotes the gamma function. In the experiment , we assume these parameters are estimated without error. The columns labelled *1ST* in Table 2 give the results for this proxy model.

The other proxy model assumes erroneously that the X_i are i.i.d. normally distributed. We do this for the reason that the normal law is so often used as the workhorse model in finance, even though it does not capture the characteristic tail feature of the data. Under this model, one calculates the loss probability for the averaged returns as follows

$$P\left\{\frac{1}{k}\sum_{i=1}^{k}X_i \leq -x\right\} \approx P\left\{Z \leq -\sqrt{k}\frac{x}{\sqrt{Var\left[X\right]}}\right\} \tag{15}$$

where Z denotes a random variable with the standard normal distribution. For the centered Student t-distribution, $E[X] = 0$ and $Var[X] = d/(d-2)$, provided the degrees of freedom is larger than two. Again, we assume these parameters are estimated without error. The results are reported in columns labelled *NOR*.

The first column *SIM* of Table 2 reproduces the results from Table 1, and give the theoretically correct values.[2] Table 2 shows that the accuracy of the normal model is very poor when the percentage loss x is large, since the normal model underestimates the true probability of the tail part. The exponential decline of the tails of the normal model is too rapid to measure the diversification effects of the Student-t. If the threshold value of x is chosen from the center of the distribution, an approximation via the normal model fits reasonably well. Note that the crossing point of the normal and fat tail distributions shifts towards the center as the aggregation level increases. For example, if $k = 2$ the normal model overestimates the probability of loss for the quantile 2.353. However this normal model underestimates the probabilities when $k \geq 3$ for the quantile $s_{5.0}^{(3)} = 2.353$. For the larger quantiles such as $s_{1.0}^{(3)} = 4.541, s_{0.5}^{(3)} = 5.841$ and $s_{0.05}^{(3)} = 12.941$ the normal model shows poor accuracy even when $k = 1$, since the tail feature is already dominant in this area. Overall the first order approximation based fat tail model provides a good approximation to the true probability even at a moderate level of aggregation. Furthermore, the heavy tail model becomes more accurate at higher levels of aggregation and smaller quantiles.

[2] Since we reran the experiment for the purpose of Table 2, entries differ slightly from those that are reported in Table 1.

4 Empirical Analysis of Diversification for Stock Returns

We provide an application to portfolio diversification effects for a portfolio of stocks. We analyze daily returns (close-to-close data), including cash dividends, for companies listed on the S&P 100 index in March of 2001. The data were obtained from the Datastream. The data span runs from January 2, 1980, through March 6, 2001, giving a sample size of n = 5,526. Thus more than 20 years of data are considered, including the 1987 Crash. To see the effects of portfolio diversification, we selected 15 stocks arbitrarily. In Table 3 the list of selected series is given, and the summary statistics for each stock return series are presented. On an annual basis the returns are around 10% (multiply μ_1 by a factor 2.5), have similar second moments, and exhibit considerably higher than normal kurtosis. This latter feature is also captured by the estimates for the tail index α in Table 4. We estimate the tail index from the semi-parametric subsample bootstrap method proposed by Danielsson et al. (2000).[3] In Table 4, we report the estimate of α, its standard error, the estimate of the scale parameter A, and the optimal number of order statistics $\hat{m}$ calculated from the bootstrap procedure. From Table 4 we see that the tail index estimates range between 1.8 and 4.4, but appear fairly concentrated around 3. The scale parameter estimates, however, differ considerably.

For the analysis of cross-diversification effects, we averaged the series of stock returns cross-cumulatively using the particular ordering from Table 3. Thus the first k series from Table 3 are averaged for the analysis of k-convolution effects. This particular aggregation method was used to enable a comparison with the Monte Carlo study from the previous section. At first we take the parameters α and A from the first stock, and assume these apply to all other stocks as well; this resembles most closely the set-up of the Monte Carlo experiment. But as the estimates from Table 4 indicate, both parameters differ across the different stocks, and we relax the presumption of parameter equality later. Due to the difference in parameters, it is evident that if the stocks in Table 3 had been ordered differently, the diversification effects reported under the assumption of parameter equality would be different. Table 5 presents tail parameter estimates for the averaged series. One clearly sees that aggregation affects the estimates of the scale A, but the tail index estimates are relatively constant, which is in line with the Table 4 results and formula (13).

In Table 6 we analyze the effects of convolution for the tail probabilities. The numbers in row *EMP* are the probabilities estimated by the empirical distribution function. The numbers in the rows labelled *NOR (I)* give the probabilities from the normal model based formula (15), using the mean and variance estimates reported in Table 4. We report the estimated probability using the averaged series in rows *NOR (II)*. The row *1ST (I)* gives the probability esti-

[3] We choose the subsample size as $n_1 = n/2$. Estimation was performed by searching over the minimum MSE of the subsample bootstrap estimates by varying $\hat{m}_1$ from 10 to $n_1/2$. We drew 1,000 resamples in the bootstrap procedure.

mates by the heavy tail model using the first order approximation as in (14), and using the estimates for the tail parameters for stock 1, ALCOA, from Table 4, and assuming that these estimates apply to all other stocks in the portfolio. From the Table 6, we find that the events beyond the threshold loss $s = 0.05$ are already quite extreme cases. The empirical probability of these events is only 0.7% when $k = 1$. The normal model clearly underestimates the probability of these events. In particular for $k = 1$ at the 0.05, 0.1 and 0.2 loss levels, the *1ST (I)* heavy tail based estimates are much closer to the observed probabilities under *EMP,* in comparison to the normal based estimates *NOR (I).*

Next, we discuss the aggregated series $k = 5, 10, 15$. The calculation in rows *1ST (I)* assume that all stocks are identically and independently distributed. As is clear from Table 4, the 15 return series do not have identical distributions cross-sectionally due to variations in the scale parameter A. For this reason we used an alternative method. First, if we relax the homogeneity of A_i but retain the equality of teh tail indexes, then we have the more general convolution formula

$$\Pr\{X + Y \leq -s\} \approx (A_x + A_y)\, s^{-\alpha},$$

where $\Pr\{X \leq -s\} \approx A_x s^{-\alpha}$ and $\Pr\{Y \leq -s\} \approx A_y s^{-\alpha}$, c.f. (14). Under the heterogeneity of A_i, the equivalent of equation (14) reads

$$\Pr\left\{\frac{1}{k}\sum_{i=1}^{k} X_t^{(i)} \leq -s\right\} \approx k^{-\alpha}\left(\sum_{i=1}^{k} A_i\right) s^{-\alpha}. \tag{16}$$

The estimated probabilities on the basis of (16) are reported in rows *1ST (II)* of Table 6. As can be seen from comparing *1ST (I)* to *1ST (II)*, the probabilities in the column *1ST (I)* improve orders of magnitude by using the heterogeneity of the A_i.

As a benchmark, we report the estimated tail probability using the averaged series itself. Of course, this method is even better since it can directly capture the dependence and heterogeneity of parameters:

$$\Pr\left\{\frac{1}{k}\sum_{i=1}^{k} X_t^{(i)} \leq -s\right\} \approx \tilde{A} s^{-\tilde{\alpha}} \tag{17}$$

where $\tilde{A}, \tilde{\alpha}$ are the estimated tail parameters using $\bar{X}_t = \frac{1}{k}\sum_{i=1}^{k} X_t^{(i)}$, see Table 5. The estimated values in rows *1ST (III)* obtained by this direct method are reasonably close to the empirical probabilities in rows *EMP*. The differences between the numbers in row *1ST (I)* and *1ST (III)* in Table 6 can in large part be explained by the wide variation in the A_i, but relatively stable values of α_i. Comparing rows *1ST (II)* and *1ST (III),* one notices that the A heterogeneity goes a long way towards explaining the aggregation effects, but not all is explained by the differences in scale.

In future work we intend to relax the assumption of independence in (14). When there exists a strong dependence between returns, the actual diversification effects are smaller than (14). As explained by (11) each individual stock return contains common components such as the market risk. The idiosyncratic risk may be diversified fully in arbitrarily large portfolio, but the strong degree of cross-sectional dependence induced by common risk cannot be diversified. This may explain why the numbers in column *1ST(II)* are further from the entries under *EMP* compared to the entries calculated with the direct method in column *1ST(III)*.

5 Conclusion

This paper first reviewed the theory of regular variation and extreme values, since heavy tailed distributions provide a good model for understanding the recurrent sizable losses in the financial market place. Subsequently, we study the problem of portfolio diversification in particular. We show that if the asset return distributions are regularly varying at infinity, then Feller's convolution theorem implies that the portfolio diversification can be easily calculated, comparing with the simplicity of the normal based calculations. This is illustrated by a simulation study and an application to S&P stock returns. As each stock return has a distribution with a heavy tail, the prediction by the normal model is quite poor the for quantiles located far in the tail area. We showed that it is important to account for differences in scales when computing the diversification effects of portfolio investment. Per contrast, differences in tail shape were not large and did not seem to matter much for the diversification effect.

References

[1] Danielsson, J., L. de Haan, L. Peng and C.G. de Vries (2000). Using a bootstrap method to choose the sample fraction in tail index estimation, *Journal of Multivariate Analysis* 76, 226-248.

[2] Embrechts, P., Kluppelberg, C., and T. Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag, Berlin.

[3] Fama, E.F., and M.H. Miller (1972). *The Theory of Finance*, Dryden Press, Hinsdale.

[4] Feller, W. (1971, Vol. II). *An Introduction to Probability Theory and Its Applications*, Wiley, New York.

[5] Geluk, J, L. Peng, and C.G. de Vries (2000). Convolutions of heavy tailed random variables and applications to portfolio diversification and MA(1) time series, *Advances in Applied Probability* 32, 1011-1026.

[6] Hill, B.M. (1975) A simple general approach to inference about the tail of a distribution, *Annals of Statistics*, 3(5), 1163-1173.

[7] Jansen, D. and C.G. de Vries (1991). On the frequency of large stock returns: Putting booms and busts into perspective, *Review of Economics and Statistics*, 73, 18-24.

[8] Leadbetter, M.R., Lindgren G. and H. Rootzén (1983). *Extremes and related Properties of Random Sequences and Processes*, Springer-Verlag, Berlin.

[9] Shao, J. and W. Tsu (1995). *The Jacknife and Bootstrap*, Springer-Verlag, New York.

[10] de Vries, C.G. (2000). Second order tail effects, in W.S. Chan, W.K. Li and H. Tong (eds), *Statistics and Finance: An interface*, Imperial College Press, Singapore, 153-165.

Table 1. The Probability of Loss for Student's t-Distribution

	$d=1$				$d=2$			
Quantile	$s^{(1)}_{5.0}$	$s^{(1)}_{1.0}$	$s^{(1)}_{0.5}$	$s^{(1)}_{0.05}$	$s^{(2)}_{5.0}$	$s^{(2)}_{1.0}$	$s^{(2)}_{0.5}$	$s^{(2)}_{0.05}$
	6.314	31.821	63.657	636.619	2.920	6.695	9.925	31.598
$k=1$	4.996	1.000	.501	.0499	4.990	1.000	.4991	.0501
2	4.998	1.001	.500	.0499	3.124	0.536	.2611	.0253
3	4.999	0.999	.499	.0498	2.238	0.362	.1747	.0167
4	4.997	0.998	.499	.0497	1.723	0.271	.1311	.0125
5	4.997	0.999	.499	.0497	1.390	0.216	.1046	.0101
10	4.996	0.998	.499	.0495	0.684	0.107	.0519	.0051
15	4.995	0.999	.499	.0497	0.444	0.070	.0341	.0033

	$d=3$				$d=6$			
Quantile	$s^{(3)}_{5.0}$	$s^{(3)}_{1.0}$	$s^{(3)}_{0.5}$	$s^{(3)}_{0.05}$	$s^{(6)}_{5.0}$	$s^{(6)}_{1.0}$	$s^{(6)}_{0.5}$	$s^{(6)}_{0.05}$
	2.353	4.541	5.841	12.941	1.943	3.143	3.707	5.959
$k=1$	5.002	.9977	.4989	.0496	5.000	1.000	.4989	.04956
2	2.329	.3257	.1491	.0129	1.557	.1383	.0515	.00270
3	1.279	.1509	.0675	.0058	0.549	.0264	.0085	.00039
4	0.777	.0846	.0377	.0032	0.209	.0068	.0021	.00009
5	0.507	.0539	.0240	.0020	0.085	.0021	.0006	.00001
10	0.115	.0127	.0058	.0005	0.002	.0000	.0000	.00000
15	0.047	.0055	.0025	.0002	0.000	.0000	.0000	.00000

Quantile $s^{(d)}_x$ denotes the x% quantile from Student t-distribution with d degrees of freedom, and.k denotes the number of convolutions.

Table 2. The Probability of Loss for Student's t-Distribution and Approximations

Quantile	$s_{5.0}^{(3)} = 2.353$			$s_{1.0}^{(3)} = 4.541$		
k	*SIM*	*NOR*	*1ST*	*SIM*	*NOR*	*1ST*
1	5.002	8.715	8.464	.9983	.4374	1.178
2	2.329	2.735	2.116	.3266	.0105	.2944
3	1.279	0.931	0.940	.1514	.0003	.1308
4	0.778	0.329	0.529	.0849	.0000	.0736
5	0.507	0.119	0.339	.0538	.0000	.0471
10	0.115	0.001	0.085	.0126	.0000	.0118
15	0.047	0.000	0.038	.0055	.0000	.0052
Quantile	$s_{0.5}^{(3)} = 5.841$			$s_{0.05}^{(3)} = 12.941$		
k	*SIM*	*NOR*	*1ST*	*SIM*	*NOR*	*1ST*
1	.4991	.0373	.5533	.04973	.00000	.05088
2	.1495	.0001	.1383	.01308	.00000	.01272
3	.0674	.0000	.0615	.00577	.00000	.00565
4	.0378	.0000	.0346	.00325	.00000	.00318
5	.0239	.0000	.0221	.00206	.00000	.00204
10	.0058	.0000	.0055	.00053	.00000	.00051
15	.0025	.0000	.0025	.00026	.00000	.00023

Quantile $s_x^{(3)}$ denotes the x% quantile from the Student t-distribution with $d = 3$ degrees of freedom, and k denotes the number of convolutions. The entries in column *SIM* are calculated from 1,000 simulations with sample size 100,000, the numbers in column *NOR* are from the normal model (15), and the entries in column *1ST* are based on the fat tail model (14).

Table 3. Selected Stocks and Summary Statistics

Series	Name	μ_1	μ_2	μ_3	μ_4
1	ALCOA	5.54	1.94	-0.26	13.40
2	AT & T	4.28	1.73	-0.35	7.75
3	BLACK & DECKER	2.04	2.24	-0.33	10.57
4	CAMPBELL SOUP	6.30	1.75	0.28	9.07
5	DISNEY (WALT)	6.63	1.95	-1.30	29.82
6	ENTERGY	4.53	1.63	-0.97	23.65
7	GEN.DYNAMICS	5.76	1.81	0.27	10.25
8	HEINZ HJ	6.58	1.59	0.11	6.35
9	JOHNSON & JOHNSON	6.92	1.63	-0.32	9.46
10	MERCK	7.56	1.58	-0.03	6.32
11	PEPSICO	7.39	1.77	-0.04	7.82
12	RALSTON PURINA	7.02	1.63	0.70	15.40
13	SEARS ROEBUCK	4.88	1.96	-0.25	16.83
14	UNITED TECHNOLOGIES	6.11	1.68	-0.10	6.84
15	XEROX	1.02	2.19	-1.78	33.77

Observations cover 01/01/1980 - 03/06/2001, giving 5526 daily observations. The sample means (μ_1) are multiplied by 10,000 and the standard errors (μ_2) are multiplied by 100. The μ_3 and μ_4 denote the sample Skewness and Kurtosis, respectively.

Table 4. Left Tail Parameter Estimates

Series	$\hat{\alpha}^{-1}$ $(s.e.)$	$\hat{\alpha}$ $(s.e.)$	$\hat{A}$	$\hat{m}$
1	.275 (.028)	3.633 (0.365)	0.140	99
2	.375 (.022)	2.670 (0.153)	2.407	303
3	.311 (.027)	3.211 (0.275)	1.126	136
4	.226 (.113)	4.428 (2.214)	0.020	4
5	.395 (.018)	2.533 (0.117)	4.946	469
6	.565 (.019)	1.769 (0.059)	52.98	884
7	.310 (.025)	3.223 (0.261)	0.512	153
8	.294 (.021)	3.397 (0.242)	0.177	197
9	.301 (.017)	3.322 (0.193)	0.222	296
10	.259 (.022)	3.858 (0.326)	0.039	140
11	.269 (.032)	3.713 (0.441)	0.104	71
12	.320 (.023)	3.127 (0.227)	0.427	190
13	.310 (.020)	3.223 (0.212)	0.593	231
14	.229 (.028)	4.361 (0.537)	0.010	66
15	.460 (.021)	2.174 (0.101)	21.18	459

The estimate of α is by the bootstrap method of Danielsson et. al. (2000) with 1,000 resamples. Standard errors are in parenthesis. The values in $\hat{A}$ are scaled up by 1,000,000. We also report the estimated optimal number of order statistics $\hat{m}$.

Table 5. Tail Parameters of the Averaged Series

k	$\hat{\alpha}^{-1}$ $(s.e.)$	$\hat{\alpha}$ $(s.e.)$	$\hat{A}$	$\hat{m}$
1	.275 (.028)	3.633 (0.365)	0.140	99
5	.349 (.022)	2.868 (0.181)	0.425	251
10	.378 (.019)	2.644 (0.132)	0.611	404
15	.357 (.018)	2.798 (0.145)	0.324	374

The estimate of α is by the bootstrap method of Danielsson et. al. (2000) with 1,000 resamples. The values in $\hat{A}$ are multiplied by 1,000,000. Standard errors are in the parenthesis. The k denotes the number of individual stocks included in the averaged series.

Table 6. Lower Tail Probabilities in Percentages

s		$k = 1$	$k = 5$	$k = 10$	$k = 15$
	EMP	6.7873	1.7738	.90498	.92308
	NOR (I)	9.8286	0.1941	.00222	.00003
0.025	*NOR (II)*	-	1.8426	.57558	.52000
	1ST (I)	9.2833	0.1340	.02160	.00743
	1ST (II)	-	1.6507	.96343	.29956
	1ST (III)	-	1.6716	1.0540	.98295
	EMP	.70588	.21719	.12670	.14480
	NOR (I)	.49007	.00000	.00000	.00000
0.05	*NOR (II)*	-	.00149	.00002	.00001
	1ST (I)	.74812	.01080	.00174	.00060
	1ST (II)	-	.13302	.07764	.02414
	1ST (III)	-	.22901	.16856	.14137
	EMP	.07240	.03620	.01810	.01810
	NOR (I)	.00001	.00000	.00000	.00000
0.10	*NOR (II)*	-	.00000	.00000	.00000
	1ST (I)	.06029	.00087	.00014	.00005
	1ST (II)	-	.01072	.00626	.00195
	1ST (III)	-	.03137	.02696	.02033
	EMP	.01810	.01810	.00000	.00000
	NOR (I)	.00000	.00000	.00000	.00000
0.20	*NOR (II)*	-	.00000	.00000	.00000
	1ST (I)	.00486	.00007	.00001	.00000
	1ST (II)	-	.00086	.00050	.00016
	1ST (III)	-	.00430	.00431	.00292

The entries in rows *EMP* are the probabilities from the empirical distribution, the numbers in rows *NOR(I)* are the probabilities from the normal model (15), and the numbers in row *1ST(I)* are the probabilities from the fat tail model (14). The numbers in rows *1ST(II)* are the probabilities calculated from (16), allowing for the differences in scale. The numbers in rows *1ST(III)* are the probabilities calculated directly from the parameters of averaged series itself. Similarly, the normal based approximation using the averaged series is in rows *NOR(II)*. The k denotes the number of individual stocks included in the averaged series, and s is the loss quantile.