Sub-additivity re-examined: the case for Value-at-Risk*

Jon Danielsson[†]   Bjorn N. Jorgensen[‡]   Mandira Sarma[§]   Casper G. de Vries[¶]

February 28, 2005

Abstract

This paper studies the issue of sub-additivity of Value-at-Risk (VaR) for heavy tailed asset returns. Using the notion of "regular variation" to define heavy tailed distribution, we establish that for heavy tailed asset return distributions with well defined mean, VaR is sub-additive in the tail region, the most relevant region for risk management. This is further demonstrated with the help of Monte Carlo simulation of 95% and 99% VaR for asset returns following three categories of bivariate distributions. Our results provide a new dimension into the ongoing debate over non sub-additivity of VaR.

KEY WORDS: Value-at-Risk, sub-additivity, regular variation, tail index, heavy tailed distribution.

JEL Classification: G00, G18

---

1 Introduction

Value-at-Risk (VaR) has been widely accepted as a tool for financial risk management. The Basel Committee's recommendation for VaR in 1996 (Basel Committee on Banking Supervision, 1996) and recently proposed Basel II norms (Basel Committee on Banking Supervision, 2003) have heightened the importance of VaR as a market risk measure. Following the guidelines of the Basel Committee, financial regulators all over the world have adopted VaR for designing capital adequacy standard for banks and financial institutions. Apart from financial regulators like central banks and Securities Exchange regulators, financial firms have adopted VaR for internal risk management and allocation of resources.

Despite its global recognition as a measure of financial risk, VaR has been criticised on certain theoretical grounds. Artzner et al. (1997) and Artzner et al. (1999) have criticised VaR for not satisfying the conditions of a "coherent" risk measure. The axioms of "coherency" comprise four mathematical properties, viz., monotonicity, positive homogeneity, translation invariance and sub-additivity. VaR, being a tail-quantile, is not sub-additive in general, and therefore, is not coherent. VaR is sub-additive under the restrictive assumption of normally distributed asset returns, which is a rare situation. since most often asset returns are found to be heavy tailed rather than normally distributed.

In this paper we study the issue of non sub-additivity of VaR in the tail region of heavy tailed asset returns. Using the concept of "regular variation" to define heavy tails, we show that for heavy tailed distributions, VaR is sub-additive in the tail region. Thus, although VaR is not sub-additive for a generic distribution, at the tail region of heavy tailed distribution VaR does satisfy sub-additivity. Keeping in view that financial returns are often heavy tailed, coupled with the fact that only the tail region is most relevant for risk management and not the entire distribution, this provides an interesting insight into the ongoing debate on the suitability of VaR as a risk measure. This is specially relevant for stress testing that require estimation of extreme quantiles, lying far out in the tail region.

This paper is organised as follows: Section 2 discusses the concept of sub-additivity, along with a brief review of the sub-additivity debate. Section 3 discusses the concept of "regular variation" and defines heavy-tailed distributions as those with "regularly varying" tails. In Section 4 we discuss sub-additivity of VaR at the tail region and establish that VaR is sub-additive in the tail area for distributions with well-defined first moment. This is further demonstrated in Section 5 with the help of Monte Carlo simulation to display the sub-additivity of 95% and 99% VaR for simulated asset returns belonging to three categories of bivariate distributions. Section 6 concludes the paper.

2 Sub-additivity

Let X and Y be two financial assets. A risk measure ρ is sub-additive if the following is true.

$$\rho(X + Y) \leq \rho(X) + \rho(Y) \tag{1}$$

Thus, the risk measure of the sum of two assets is bounded above by the sum of their individual

risks.[1] The property of sub-additivity can be motivated by many practical considerations. Sub-additivity ensures that the diversification principle of modern portfolio theory holds. A sub-additive measure would always generate a lower risk measure for a diversified portfolio than a non-diversified portfolio. In terms of internal risk management, sub-additivity also implies that the overall risk of a financial firm can be added up to be equal to or less than the sum of the risks of individual departments of the firm. This appears to give an appealing dimension into the idea of integrated risk management. Further, in the absence of sub-additivity, a financial firm with risk $X + Y$ would calculate a smaller regulatory capital of $\rho(X) + \rho(Y)$, leading to underestimation of risk reserve.

It can be proved easily by constructing suitable examples that VaR violates sub-additivity property (Artzner et al., 1999; Acerbi & Tasche, 2001; Acerbi et al., 2001). The lack of sub-additivity of VaR has been severely criticised, and several alternative measures have been proposed to replace VaR such as *tail conditional expectation* (TCE) and *worst conditional expectation* (WCE) by Artzner et al. (1999) and *expected shortfall* (ES) by Acerbi et al. (2001).

It is however, argued that "imposing sub-additivity for all risks (including dependent risks) is not in line with what could be called best practice" (Dhaene et al., 2003).[2] The measure of global risk may not be a priori smaller than the sum total of local risks. The following example attempts to illustrate this point.

Consider a hypothetical economy where there are only two banks, viz. B1 and B2. Suppose that B1 has only one investment project P1 and B2 has only one investment project P2, P1 and P2 being independent, each project having exactly the same pay off structure as follows:

Pay-off	Probability
High profit (H)	$\frac{1}{4}$
Low profit (L)	$\frac{1}{4}$
High loss (-H)	$\frac{1}{4}$
Low loss (-L)	$\frac{1}{4}$

Thus, the probability of failure of a bank is $\frac{1}{2}$.

Under this scenario, it is easy to show that

$$\Pr\{\text{At least one bank fails}\} = \frac{3}{4}$$
$$\Pr\{\text{Systemic failure}\} = \frac{1}{4}$$

Now, assume that the two banks merge. Then, the hypothetical economy has just one bank, investing in two independent projects P1 and P2 with the pay-off structure from each project

---

[1] "...a merger does not create extra risk" (Artzner et al., 1999). However it depends. Merger usually creates positive diversification effect but might increase the systemic risk. See example that follows.

[2] Dhaene et al. (2003) have also argued that the axioms of "coherence" leads to a very restrictive set of risk measures that cannot be used in practical situations.

as given above. Under this scenario it is easily shown that

$$\Pr\{\text{Bank fails}\} = \frac{6}{16}$$
$$\Pr\{\text{Systemic failure}\} = \frac{6}{16}$$

Thus, the probability of bank failure $\left(\frac{6}{16}\right)$ is less in the second scenario than that in the first scenario $\left(\frac{3}{4}\right)$. This implies that the diversification effect of the merger has reduced the risk of the bank failure.

As far as the global risk is concerned, in the first scenario, systemic breakdown occurs when both the banks fail and in the second scenario this occurs when the merged bank fails. In this example we see that the probability of systemic failure is higher in the second scenario $\left(\frac{6}{16}\right)$ than that in the first scenario $\left(\frac{1}{4}\right)$! Thus, the merger has led to an increase in the global risk, despite the benefit of diversification.[3]

This simple example illustrates that the diversification does not necessarily lead to a reduction in the global risk. From this point of view sub-additivity may not depict the complex nature of the financial markets.

As the debate over the requirement of sub-additivity (or, for that matter, the axioms of coherence) continues, we take a fresh look at VaR for heavy tailed asset return distribution. We present a new dimension into the debate by establishing that VaR is sub-additive in the tail regions of heavy tailed distributions.

3   Heavy tailed asset returns and Regular Variation

Empirical studies have established that the distribution of speculative asset returns tend to have heavier tails than the normal distribution tails (Mandelbrot, 1963; Pagan, 1996; Engle, 1982; Jansen & de Vries, 1991). Heavy tailed distributions are often defined in terms of higher than normal kurtosis. However, the kurtosis of a distribution may be high if either the tails of the cdf are heavier than the normal or if the center is more peaked, or both. Further, it is not only the higher than normal kurtosis, but also failure of higher moments that define heavy tails.

In this paper we define heavy tailed distribution as one characterised by the failure of the moments of order m (> 0) or higher. Such distributions have tails that exhibit a power type behaviour like the Pareto distribution, as commonly observed in finance. Such tail behaviour can be mathematically defined by using the notion of "regular variation", as defined below[4] .

Definition: A cdf $F(x)$ varies regularly at minus infinity with tail index $\alpha > 0$ if

$$\lim_{t \to \infty} \frac{F(-tx)}{F(-t)} = x^{-\alpha} \quad \forall x > 0 \tag{2}$$

---

[3]One can easily generalise this example to incorporate more complex financial entities, and still come up with similar argument.

[4]For an encyclopedic treatment of regular variation, see Bingham et al. (1987); Resnick (1987).

This implies that, to a first order approximation, all distributions have a tail comparable to the Pareto distribution:

$$F(-x) = Ax^{-\alpha}[1 + o(1)], \ x > 0, \ \text{ for } \alpha > 0 \text{ and } A > 0 \tag{3}$$

This implies, for large x,

$$f(-x) \approx \alpha A x^{-\alpha-1} \ x > 0, \ \text{ for } \alpha > 0 \text{ and } A > 0 \tag{4}$$

so that the density declines at a power rate $x^{-\alpha-1}$ far to the left of the centre of the distribution which contrasts with the exponentially fast declining tails of the Gaussian distribution. This power is outweighed by the explosion of x^m in the computation of moments of order $m \geq \alpha$. Thus, moments of order $m \geq \alpha$ are unbounded and therefore these distributions display heavy tailed behaviour. The power α is called the tail index and it determines the number of bounded moments. It is readily verified that Student–t distributions, among others, vary regularly at infinity, has degrees of freedom equal to the tail index and satisfies the above approximation. Likewise, the stationary distribution of the popular GARCH(1,1) process has regularly varying tails, see de Haan et al. (1989).

Further, to a second order approximation, the tail of a regularly varying cdf can be approximated as

$$F(-x) = Ax^{-\alpha} \left[1 + Bx^{-\beta} + O\left(x^{-\beta} \right) \right], \text{ as } x \to \infty \tag{5}$$

4 Sub-additivity of VaR in the tail

In this section, we examine sub-additivity of VaR at the tail region of heavy tailed distributions defined as above. We consider only the lower tail.[5]

In the following, we assume that X and Y are two asset returns, each having a regularly varying tail with tail index $\alpha > 0$. In this paper we consider only the case where the two asset returns have equal tail index and equal or unequal tail coefficients. If the tail coefficients are equal then the asset returns have identical tails while the tails are non-identical if the tail coefficients are different. Thus, we assume that the source of non-identical tail behaviour arises from the differing values of the tail coefficients.[6]

We consider the case of independent X and Y in Section 4.1 and the case where X and Y are dependent in Section 4.2.

---

[5]For defining the upper tails of a heavy tailed distributions, we may use the notion of "regular variation" at plus infinity and carry on with an analogous proof.

[6]Some empirical studies have found that most asset returns distributions tend to display equal tail coefficients but do differ considerably with respect to their scale coefficients. See, eg. Hyung & de Vries (2002).

4.1 Case of independent assets

Proposition 1 *Suppose that X and Y are two independent asset returns both having regularly varying tails with index $\alpha > 0$ and scale coefficients $A > 0$ and $B > 0$. Then for $\alpha > 1$ VaR is sub-additive in the tail region, regardless of whether or not $A = B$.*

Proof:
See Appendix A.1.

Thus, if we assume that $\alpha > 1$, so that the mean of the assets are well defined, then sub-additivity of VaR is established. The case of $\alpha \leq 1$, or the case of unbounded mean is very rare in finance, and therefore the assumption of $\alpha > 1$ is not unreasonable for financial assets.

4.1.1 Second order approximation

The second order tail approximation (5) approximates a much larger area in the tail than the first order approximation.

We can show that even a second order approximation of the tails, when such an approximation is valid, can display VaR sub-additivity in some cases.

For example, suppose that X and Y are two independent asset returns, each having regularly varying tails with the second order approximation as in (5) with identical tail indexes and identical tail coefficients.

Application of the Bruijn's theory of asymptotic inversion (Bingham et al., 1987) (pages 28-29) leads to the following results

$$VaR_p(X) \approx A^{\frac{1}{\alpha}}p^{-\frac{1}{\alpha}}\left[1+\frac{B}{\alpha}A^{-\frac{\beta}{\alpha}}p^{\frac{\beta}{\alpha}}\right] \tag{6}$$

$$VaR_p(Y) \approx A^{\frac{1}{\alpha}}p^{-\frac{1}{\alpha}}\left[1+\frac{B}{\alpha}A^{-\frac{\beta}{\alpha}}p^{\frac{\beta}{\alpha}}\right] \tag{7}$$

$$VaR_p(X)+VaR_p(Y) \approx 2A^{\frac{1}{\alpha}}p^{-\frac{1}{\alpha}}\left[1+\frac{B}{\alpha}A^{-\frac{\beta}{\alpha}}p^{\frac{\beta}{\alpha}}\right] \tag{8}$$

Proposition 2 *If X and Y are independent returns with regularly varying tails, each following the second order approximation (5) with identical tail index and identical scale coefficients, then $VaR_p(X + Y)$ can be approximated by the following expression.*

$$VaR_p(X+Y) \approx (2A)^{\frac{1}{\alpha}}p^{-\frac{1}{\alpha}}\left[1+\frac{B}{\alpha}\left(\frac{p}{2A}\right)^{\frac{\beta}{\alpha}}+\frac{\alpha+1}{2}E(X^2)\left(\frac{p}{2A}\right)^{\frac{2}{\alpha}}\right]+E(X) \tag{9}$$

Proof:
See Appendix A.2.

Proposition 3 *If X and Y are independent returns with regularly varying tails, each following the second order approximation (5) with identical tail index and identical scale coefficients, and if $E(X) = 0$, $\beta < 2$ and $B > 0$, then VaR is sub-additive in the tail region.*

Proof:

Using results from (6) to (9), and using the conditions $E(X) = 0$, $\beta < 2$ and $B > 0$, it is easy to show that

$$VaR_p(X+Y) - VaR_p(X) - VaR_p(Y) \approx A^{\frac{1}{\alpha}} p^{-\frac{1}{\alpha}} \left[\left(2^{\frac{1}{\alpha}} - 2\right) + \frac{B}{\alpha} A^{-\frac{\beta}{\alpha}} p^{\frac{\beta}{\alpha}} \left(2^{\frac{1}{\alpha} - \frac{\beta}{\alpha}} - 2\right) \right]$$
$$< 0$$

Thus under these specific assumptions, even a second order tail approximation leads to VaR sub-additivity.

4.2   Case of dependent assets

It is well known that financial assets are not necessarily independent. In this section we consider the case where the assets are assumed to be dependent.

Suppose that X_1 and X_2 are two assets, not necessarily independent. The simplest way to model the dependence between X_1 and X_2 would be to introduce the dependence through a common market factor, as in the case of a single index market model, given below

$$X_i = \beta_i R + Q_i, \ i = 1, 2 \tag{10}$$

where R denotes the return of the market portfolio, β_i the market risk and Q_i the idiosyncratic risk of asset X_i. In this model, the Q_is and R are independent of each other; further Q_is are themselves independent of each other. Thus, the only source of cross-sectional dependence between X_1 and X_2 is the common market risk. The security specific risks Q_i are independent of each other and therefore can be diversified away.

Since R and Q_i are independent, we can use Feller's convolution theorem to approximate the tails of X_1 and X_2, depending upon the tail behaviour of R, Q_1 and Q_2. We can further use it to approximate the tail of $X_1 + X_2$. Thus, under such a model, we can proceed in a similar manner as in the case of independent asset returns. To illustrate this, we present below one particular case, viz., the case where R, Q_1 and Q_2 have regularly varying tails with the same tail index α, but with different tail coefficients.

Proposition 4 *Suppose that asset returns X_1 and X_2 can be modelled by the single index market model, where R, Q_1 and Q_2 all have regularly varying tails with tail index $\alpha > 0$ and tail coefficients $A_r > 0$, $A_1 > 0$ and $A_2 > 0$ respectively. If $\alpha > 1$, then under the assumption that the distribution of R is symmetric, VaR is sub-additive in the tail region.*

Proof:

See Appendix A.3

In general, the single index market model (10) may not describe the true nature of the dependence between X_1 and X_2 since Q_i's may not be cross sectionally independent, although each one of them may be independent from the common market factor R. For example, apart from the market risk, the assets X_1 and X_2 may be dependent on an industry specific risk, depicted by the movement of an industry specific index S, also known as "sectoral index"

in finance. Such industry specific factor may lead to dependence between Q_1 and Q_2. We may model such cross sectional dependence by generalising the model (10) by incorporating a sector specific factor S.

$$X_i = \beta_i \ R + \tau_i \ S + Q_i, \ i = 1, 2 \tag{11}$$

where R is the market factor, S is the industry specific factor and Q_i is the idiosyncratic risk of the asset X_i. In this model Q_i is independent of R and S. Further, Q_i are cross sectionally independent. In this model, τ_i is the industry specific risk of the asset X_i. If S has a regularly varying tail with scale coefficient A_s and tail index α, then under the assumption of symmetric tails for R and S, it can be shown in the similar manner as in Appendix A.3, that

$$
\begin{aligned}
VaR_p(X_1) &\approx p^{\frac{1}{\alpha}} \left(|\beta_1|^\alpha A_r + |\tau_1|^\alpha A_s + A_1 \right)^{\frac{1}{\alpha}} \\
VaR_p(X_2) &\approx p^{\frac{1}{\alpha}} \left(|\beta_2|^\alpha A_r + |\tau_2|^\alpha A_s + A_2 \right)^{\frac{1}{\alpha}} \\
VaR_p(X_1 + X_2) &\approx p^{\frac{1}{\alpha}} \left(|\beta_1 + \beta_2|^\alpha A_r + |\tau_1 + \tau_2|^\alpha A_s + A_1 + A_2 \right)^{\frac{1}{\alpha}}
\end{aligned}
$$

Proceeding similarly as in the case of case of the single index market model above (Proposition 4), one case easily establish that

$$VaR_p(X_1 + X_2) \leq VaR_p(X_1) + VaR_p(X_2)$$

One can generalise the model by introducing more factors in the dependence structure between X_1 and X_2 and establish sub-additivity in an analogous manner.

5   Simulation results

In order to demonstrate empirically the above results, we have carried out a set of simulations. In this section we present the results from the simulations from three categories of bivariate distributions: Student's t, jump process and bivariate GARCH. We simulate from two sample sizes $N = 10^3$ and 10^6, the former is chosen to mimic the real world sample sizes, and the latter the asymptotics. The number of simulations from each sample size is $S = 10000$ and 200. In the simulations, p denotes the significance level of VaR and n denotes the number of cases where

$$VaR_p(X_1 + X_2) > VaR_p(X_1) + VaR_p(X_2)$$

5.1   Student's t-distribution

The Student's t-distribution is widely applied in modeling financial returns. It has a regularly varying tail with tail index equal to its degrees of freedom and tail coefficient as below:

$$A = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \frac{\nu^{\frac{\nu-1}{2}}}{\sqrt{\nu\pi}}$$

where ν is the degrees of freedom. Further, for $\nu > 2$ the tail of Student's t-distribution satisfy the second order approximation (5) with $\beta = 2$, and $B = -\frac{\nu^2}{2}\frac{\nu+1}{\nu+2}$.

Suppose that X_1 is N draws from $t_{(\nu_1)}$ and X_2 is N draws from $t_{(\nu_1)}$, both being IID. We then provide a dependence structure as follows. Suppose ρ is a correlation coefficients. Consider the Choleski decomposition of the covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} = A'A$$

Then the data matrix is given by

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} A'$$

Thus, the correlation between the two columns of X is ρ at least when the second moment is defined. The results from simulating these data are presented in Tables 1 and 2.

As shown in these tables, n, the number of times when sub-additivity fails is very high when $\nu_1 = 1$ and/or $\nu_2 = 1$. These are the cases when the first moments of X_1 and X_2 are not well defined. Existence of well-defined mean is a necessary assumption in our analysis for sub-additivity to hold.

When the degrees of freedom of the Student's t-variates is higher than 1, then the first moment is well defined. As shown in Tables 1 and 2 for well defined first moment, the violation of sub-additivity is negligible or zero.

5.2 Jump process

x_1 and x_2 are N draws from $N(0,1)$, where each process is subject to the occasional jump:

$$x_{i,j} = \begin{cases} x_{i,j} & \text{with probability} & a \\ b+c, & c\tilde{\,}U(0,d) & \text{with probability} & 1-a \end{cases}$$

in our case

$$x_{i,j} = \begin{cases} x_{i,j} & \text{with probability} & 0.995 \\ 10+c, & c\tilde{\,}U(0,0.2) & \text{with probability} & 0.005 \end{cases}$$

we then introduce what we call *joint event prob*, denoted by q, which is the probability that on days when the x_1 jumps, x_2 also Jumps.

Table 3 presents the results of the simulation from the jump process.

5.3 BEKK GARCH

we estimate the parameters of a bivariate GARCH model, specifically the BEKK form with daily data from Microsoft and Goldman Sachs over four years. We then simulate from this

9

model. Using these simulated results, we estimate VaRs of the individual returns and their sum. In Table 4 we present the results from the simulation. It is seen from this table that the number of sub-additivity failure is close to zero.

6 Conclusion

In this paper we take a fresh look at the issue of sub-additivity of VaR. We argue that for heavy tailed asset return distributions VaR is sub-additive in the tail region. We define heavy tailed distribution as one characterised by the failure of moments of order $m > 0$ or higher. Using the notion of "regular variation" to describe such a tail behaviour, we establish that for distributions with well defined mean ($m > 1$), VaR is sub-additive at the tail region of the distribution. Thus in the relevant region for risk management, i.e., the tail region, VaR is sub-additive. We further demonstrate our results with the help of a set of Monte Carlo simulation of 99% and 95% VaR for asset returns simulated from three categories of bivariate distributions: Student's t-distribution, jump processes and bivariate GARCH processes.

Our results provide a new insight into the sub-additivity debate and offers VaR to be still applicable for risk management despite its lack of sub-additivity in general.

References

Acerbi, C. & Tasche, D. (2001), Expected Shortfall: a natural coherent alternative to Value at Risk, Abaxbank Working Paper arXiv:cond-mat/0105191.

Acerbi, C., Nordio, C. & Sirtori, C. (2001), Expected Shortfall as a Tool for Financial Risk Management, Abaxbank Working Paper arXiv:cond-mat/0102304.

Artzner, P., Delbaen, F., Eber, J. M. & Heath, D. (1997), 'Thinking coherently', *Risk* **10**(11), 68–71.

Artzner, P., Delbaen, F., Eber, J. M. & Heath, D. (1999), 'Coherent Measures of Riak', *Mathematical Finance* **9**, 203–228.

Basel Committee on Banking Supervision (1996), Ammendment to the capital accord to incorporate market risks, Committee Report 24, Basel Committee on Banking Supervision, Basel, Switzerland.

Basel Committee on Banking Supervision (2003), Consultative Document The New Basel Capital Accord, Committee report, Basel Committee on Banking Supervision, Basel, Switzerland.

Bingham, N. H., Goldie, C. M. & Teugels, J. L. (1987), *Regular Variation*, Cambridge University Press, Cambridge.

de Haan, L., Resnick, S. I., Rootzen, H. & de Vries, C. G. (1989), 'Extremal Behaviour of Solutions to a Stochastic Difference Equation with Applications to Arch-Processes', *Stochastic Processes and their Applications* pp. 213–224.

Dhaene, J., Goovaerts, M. J. & Kaas, R. (2003), 'Economic Capital Allocation Derived from Risk Masures', *North American Actuarial Journal* **7**(2), 44–56.

Engle, R. (1982), 'Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation', *Econometrica* **50**, 987–1007.

Feller, W. (1971), *An Introduction to Probability Theory and Its Applications*, Vol. II, John Wiley & Sons.

Garcia, R., Renault, E. & Tsafack, G. (2003), Proper Conditioning for Coherent VaR in Portfolio Management, Unversite de Montreal, CIRANO and CIREQ Working Paper.

Hyung, N. & de Vries, C. G. (2002), 'Portfolio Diversification Effects and Regular Variation in Financial Data', *Journal of German Statistical Society* **86**, 69–82.

Jansen, D. W. & de Vries, C. G. (1991), 'On the Frequency of Large Stock Returns: Putting Booms and Busts into Perspective', *The Review of Economics and Statistics* **73**(1), 18–24.

Loeve, M. (1963), *Probability Theory*, 3rd edn, Springer, New York.

Mandelbrot, B. B. (1963), 'The variation of certain speculative prices', *Journal of Business* **36**, 392 – 417.

Pagan, A. (1996), 'The econometrics of financial markets', *Journal of Empirical Finance* **3**, 15 – 102.

Resnick, S. I. (1987), *Extreme Values, Regular Variation and Point Process*, Springer-Verlag.

Appendix A

A.1 Proof of Proposition 1

Suppose that X and Y are two independent asset returns both having regularly varying tails such that

$$\Pr\{X \leq -x\} \approx A\, x^{-\alpha}$$
$$\Pr\{Y \leq -x\} \approx B\, x^{-\alpha}$$

Then the VaR at p-level for X, denoted by $VaR_p(X)$, is given by

$$\Pr\{X \leq VaR_p(X)\} \approx p$$
$$A\,[VaR_p(X)]^{-\alpha} \approx p$$
$$VaR_p(X) \approx \left[\frac{A}{p}\right]^{\frac{1}{\alpha}}$$

Similarly,

$$VaR_p(Y) \approx \left[\frac{B}{p}\right]^{\frac{1}{\alpha}}$$

We consider the following cases:

I. $A = B$[7] In this case,

$$VaR_p(X) + VaR_p(Y) \approx 2\left[\frac{A}{p}\right]^{\frac{1}{\alpha}}$$

Using Feller's convolution theorem (Feller, 1971)(page 278), we have

$$\Pr\{X + Y \leq VaR_p(X+Y)\} \approx p$$
$$2\,A\,[VaR_p(X+Y)]^{-\alpha} \approx p$$
$$VaR_p(X+Y) \approx 2^{\frac{1}{\alpha}}\left[\frac{A}{p}\right]^{\frac{1}{\alpha}}$$

Thus it follows

$$VaR_p(X+Y) - [VaR_p(X) + VaR_p(Y)] \approx \left[\frac{A}{p}\right]^{\frac{1}{\alpha}}[2^{\frac{1}{\alpha}} - 2]$$

Following cases may be considered

---

[7]Garcia et al. (2003) have examined this by using a different approach than ours.

(a) $\alpha = 1$: in this case $VaR_p(X+Y) = VaR_p(X) + VaR_p(Y)$ and therefore VaR is additive.

(b) $0 < \alpha < 1$: in this case $VaR_p(X+Y) > VaR_p(X) + VaR_p(Y)$ and hence VaR is super-additive.

(c) $\alpha > 1$: in this case $VaR_p(X+Y) < VaR_p(X) + VaR_p(Y)$ and hence VaR is sub-additive.

II. $A \neq B$

In this case, Feller's theorem gives

$$\Pr\{X + Y \leq -x\} \approx (A+B)\, x^{-\alpha}$$

Thus we get the VaR as:

$$VaR_p(X+Y) \approx \left[\frac{A+B}{p}\right]^{\frac{1}{\alpha}}$$

This gives

$$VaR_p(X+Y) - VaR_p(X) - VaR_p(Y) \approx \left(\frac{1}{p}\right)^{\frac{1}{\alpha}}\left[(A+B)^{\frac{1}{\alpha}} - A^{\frac{1}{\alpha}} - B^{\frac{1}{\alpha}}\right]$$

Following cases may be considered:

(a) $\alpha > 1$: In this case, $\frac{1}{\alpha} < 1$. Using C_α inequality (Loeve, 1963), (p. 155), it directly follows that

$$(A+B)^{\frac{1}{\alpha}} \leq A^{\frac{1}{\alpha}} + B^{\frac{1}{\alpha}}$$

VaR is sub-additive in this case

(b) $\alpha < 1$: In this case, $\frac{1}{\alpha} > 1$. Hence the C_α inequality gives

$$(A+B)^{\frac{1}{\alpha}} \leq 2^{\frac{1}{\alpha}-1}(A^{\frac{1}{\alpha}} + B^{\frac{1}{\alpha}})$$

Thus VaR is not necessarily sub-additive in this case.

A.2 Proof of Proposition 2

Suppose that X and Y are two independent asset returns, each having regularly varying tails with the following identical second order approximation

$$\Pr\{X \leq -x\} = \Pr\{X > x\} = Ax^{-\alpha}\left[1 + Bx^{-\beta} + o(x^{-\alpha})\right]$$

$$\Pr\{Y \leq -x\} = \Pr\{Y > x\} = Ax^{-\alpha}\left[1 + Bx^{-\beta} + o(x^{-\alpha})\right]$$

$$where\ \alpha > 2,\ \beta > 0,\ A > 0\ and\ B \in R^1 - \{0\}$$

Following the second order convolution result from Dacorogna et al (1998) and de Vries (1999), we have,

$$\Pr\{X + Y > x\} \quad = \quad \Pr\{X + Y \le -x\} \approx 2Ax^{-\alpha}\left[1 + Bx^{-\beta} + \alpha E(X)x^{-1} + \frac{1}{2}\alpha(\alpha+1)E(X^2)x^{-2}\right]$$

where $E(X) = E(Y)$ and $E(X^2) = E(Y^2)$

Let $p = \Pr\{X + Y \le -x\}$. Then,

$$
\begin{aligned}
p &\approx 2Ax^{-\alpha}\left[1 + Bx^{-\beta} + \alpha E(X)x^{-1} + \frac{1}{2}\alpha(\alpha+1)E(X^2)x^{-2}\right] \\
\frac{p}{2A} &\approx x^{-\alpha}\left[1 + Bx^{-\beta} + \alpha E(X)x^{-1} + \frac{1}{2}\alpha(\alpha+1)E(X^2)x^{-2}\right] \\
y &\approx x^{-\alpha}\left[1 + Bx^{-\beta} + \alpha E(X)x^{-1} + \frac{1}{2}\alpha(\alpha+1)E(X^2)x^{-2}\right] \quad where\ y = \frac{p}{2A} \quad (12) \\
x &\approx y^{-\frac{1}{\alpha}}\left[1 + Bx^{-\beta} + \alpha E(X)x^{-1} + \frac{1}{2}\alpha(\alpha+1)E(X^2)x^{-2}\right]^{\frac{1}{\alpha}} \\
&= y^{-\frac{1}{\alpha}}g\left(x(y)\right) \\
&= y^{-\frac{1}{\alpha}} + y^{-\frac{1}{\alpha}}\{g\left(x(y)\right) - 1\} \\
&= y^{-\frac{1}{\alpha}} + \epsilon(y)\ say \quad (13)
\end{aligned}
$$

Clearly,

$$\frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}} \to 0 \ as\ x \to \infty\ since, as\ x \to \infty, g\left(x(y)\right) \to 1$$

Using (13) in (12), we have the following:

$$
\begin{aligned}
y &\approx \left(y^{-\frac{1}{\alpha}} + \epsilon(y)\right)^{-\alpha}\left[1 + B\left(y^{-\frac{1}{\alpha}} + \epsilon(y)\right)^{-\beta} + \alpha E(X)\left(y^{-\frac{1}{\alpha}} + \epsilon(y)\right)^{-1} + \frac{1}{2}\alpha(\alpha+1)E(X^2)\left(y^{-\frac{1}{\alpha}} + \epsilon(y)\right)^{-2}\right] \\
&= y\left(1 + \frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right)^{-\alpha}\left[1 + By^{\frac{\beta}{\alpha}}\left(1 + \frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right)^{-\beta} + \alpha E(X)y^{\frac{1}{\alpha}}\left(1 + \frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right)^{-1} + \frac{1}{2}\alpha(\alpha+1)y^{\frac{2}{\alpha}}\left(1 + \frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right)^{-2}\right]
\end{aligned}
$$

Thus,

$$1 \approx \left(1 + \frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right)^{-\alpha}\left[1 + By^{\frac{\beta}{\alpha}}\left(1 + \frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right)^{-\beta} + \alpha E(X)y^{\frac{1}{\alpha}}\left(1 + \frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right)^{-1} + \frac{1}{2}\alpha(\alpha+1)E(X^2)y^{\frac{2}{\alpha}}\left(1 + \frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right)^{-2}\right]$$

$$(14)$$

Now, let

$$
\begin{aligned}
u &= \frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}, and \\
f(u) &= (1 + u)^{-\alpha}
\end{aligned}
$$

Taylor's expansion of $f(u)$ around $u = 0$ gives

$$
\begin{aligned}
f(u) &= f(0) + uf'(u)|_{u=0} + \frac{u^2}{2!}f''(u)|_{u=0} + \frac{u^3}{3!}f'''(u)|_{u=0} + \dots \\
&= 1 - \alpha u + \alpha(\alpha+1)\frac{u^2}{2} - \alpha(\alpha+1)(\alpha+2)\frac{u^3}{6} + \dots
\end{aligned}
$$

Thus, to a first order approximation,

$$
\begin{aligned}
f(u) &\approx 1 - \alpha u \\
\left(1 + \frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right)^{-\alpha} &\approx 1 - \alpha\frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}
\end{aligned} \tag{15}
$$

Using (15) in (14), we have,

$$
\begin{aligned}
1 &\approx \left(1 - \alpha\frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right)\left[1 + By^{\frac{\beta}{\alpha}}\left(1 - \beta\frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right) + \alpha E(X)y^{\frac{1}{\alpha}}\left(1 - \frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right)\right. \\
&\quad \left. + \frac{1}{2}\alpha(\alpha+1)E(X^2)y^{\frac{2}{\alpha}}\left(1 - 2\frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right)\right] \\
&= 1 + By^{\frac{\beta}{\alpha}}\left(1 - \beta\frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right) + \alpha E(X)y^{\frac{1}{\alpha}}\left(1 - \frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right) + \frac{1}{2}\alpha(\alpha+1)E(X^2)y^{\frac{2}{\alpha}}\left(1 - 2\frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}}\right) \\
&\quad - \alpha\frac{\epsilon(y)}{y^{-\frac{1}{\alpha}}} - \alpha B\epsilon(y)y^{\frac{\beta}{\alpha}+\frac{1}{\alpha}}\left(1 - \beta\epsilon(y)y^{\frac{1}{\alpha}}\right) - \alpha^2 E(X)\epsilon(y)y^{\frac{2}{\alpha}}(1 - \epsilon(y)y^{\frac{1}{\alpha}}) \\
&\quad - \frac{1}{2}\alpha^2(\alpha+1)E(X^2)\epsilon(y)y^{\frac{3}{\alpha}}(1 - 2\epsilon(y)y^{\frac{1}{\alpha}}) \\
\alpha\epsilon(y)y^{\frac{1}{\alpha}} &\approx By^{\frac{\beta}{\alpha}} + \alpha E(X)y^{\frac{1}{\alpha}} + \frac{1}{2}\alpha(\alpha+1)E(X^2)y^{\frac{2}{\alpha}} - \epsilon(y)\left[B\beta y^{\frac{\beta}{\alpha}+\frac{1}{\alpha}} + \alpha E(X)y^{\frac{2}{\alpha}}\right. \\
&\quad \left. + \alpha(\alpha+1)E(X^2)y^{\frac{3}{\alpha}} + \alpha By^{\frac{\beta}{\alpha}+\frac{1}{\alpha}} + \alpha^2 E(X)y^{\frac{2}{\alpha}} + \frac{1}{2}\alpha^2(\alpha+1)E(X^2)y^{\frac{3}{\alpha}}\right] \\
&\quad + (\epsilon(y))^2\left[\alpha\beta By^{\frac{\beta}{\alpha}+\frac{2}{\alpha}} + \alpha^2 E(X)y^{\frac{3}{\alpha}} + \alpha^2(\alpha+1)E(X^2)y^{\frac{4}{\alpha}}\right] \\
\epsilon(y) &\approx \frac{B}{\alpha}y^{\frac{\beta}{\alpha}-\frac{1}{\alpha}} + E(X) + \frac{1}{2}(\alpha+1)E(X^2)y^{\frac{1}{\alpha}}
\end{aligned} \tag{16}
$$

As $x \to \infty$, the other terms $\to 0$ faster. Using (16) in (13),

$$
\begin{aligned}
x &\approx y^{-\frac{1}{\alpha}} + \frac{B}{\alpha}y^{\frac{\beta}{\alpha}-\frac{1}{\alpha}} + E(X) + \frac{1}{2}(\alpha+1)E(X^2)y^{\frac{1}{\alpha}} \\
&\quad y^{-\frac{1}{\alpha}}\left[1 + \frac{B}{\alpha}y^{\frac{\beta}{\alpha}} + \frac{\alpha+1}{2}E(X^2)y^{\frac{2}{\alpha}}\right] + E(X)
\end{aligned}
$$

substituting $y = \frac{p}{2A}$,

$$
x \approx \left(\frac{p}{2A}\right)^{-\frac{1}{\alpha}}\left[1 + \frac{B}{\alpha}\left(\frac{p}{2A}\right)^{\frac{\beta}{\alpha}} + \frac{\alpha+1}{2}E(X^2)\left(\frac{p}{2A}\right)^{\frac{2}{\alpha}}\right] + E(X)
$$

Appendix A.3 Proof of Proposition 4

Suppose that R has a regularly varying tail with index α and Q_i $i = 1, 2$ has a regularly varying tail with index α. Further, suppose that R has a symmetric distribution. Thus, to a first order approximation,

$$
\begin{aligned}
\Pr\{R \leq -x\} &\approx A_r \, x^{-\alpha} \\
\Pr\{R \geq x\} &\approx A_r \, x^{-\alpha} \\
\Pr\{\beta_i R \leq -x\} &= \Pr\{R \leq -\frac{x}{\beta_i}\}
\end{aligned}
$$

If $\beta_i > 0$ then

$$
\Pr\left\{R \leq -\frac{x}{\beta_i}\right\} \approx A_r \beta_i^\alpha x^{-\alpha}
$$

If $\beta_i < 0$ then

$$
\begin{aligned}
\Pr\{\beta_i R \leq -x\} &= \Pr\{-|\beta_i| R \leq -x\} \\
&= \Pr\{|\beta_i| R \geq x\} \\
&\approx A_r |\beta_i|^\alpha \, x^{-\alpha}
\end{aligned}
$$

Thus

$$
\Pr\{\beta_i R \leq -x\} \approx A_r |\beta_i|^\alpha \, x^{-\alpha}, \quad \beta_i \in R
$$

For the individual assets Q_1 and Q_2

$$
\Pr\{Q_i \leq -x\} \approx A_i \, x^{-\alpha}, \; i = 1, 2
$$

By Feller's convolution theorem

$$
\begin{aligned}
\Pr\{X_i \leq -x\} &\approx |\beta_i|^\alpha A_r \, x^{-\alpha} + A_i x^{-\alpha} \\
p &\approx x^{-\alpha} \left(A_i + |\beta_i|^\alpha A_r\right) \\
x &\approx p^{-\frac{1}{\alpha}} \left(A_i + |\beta_i|^\alpha A_r\right)^{\frac{1}{\alpha}}
\end{aligned}
$$

Similarly

$$
\Pr\{X_1 + X_2 \leq -x\} \approx |\beta_1 + \beta_2|^\alpha A_r \, x^{-\alpha} + A_1 x^{-\alpha} + A_2 x^{-\alpha}
$$

Thus,

$$
\begin{aligned}
VaR_p(X_1) &\approx p^{-\frac{1}{\alpha}} \left(A_1 + |\beta_1|^\alpha A_r\right)^{\frac{1}{\alpha}} \\
VaR_p(X_2) &\approx p^{-\frac{1}{\alpha}} \left(A_2 + |\beta_2|^\alpha A_r\right)^{\frac{1}{\alpha}} \\
VaR_p(X_1 + X_2) &\approx p^{-\frac{1}{\alpha}} \left[\left(A_1 + A_2 + |\beta_1 + \beta_2|^\alpha A_r\right)^{\frac{1}{\alpha}}\right]
\end{aligned}
$$

To establish the sub-additivity we proceed as follows.

$$
\begin{aligned}
VaR_p(X_1 + X_2) \;\approx\;\; & p^{-\frac{1}{\alpha}}\left[(A_1 + A_2 + |\beta_1 + \beta_2|^\alpha A_r)^{\frac{1}{\alpha}}\right] \\[4pt]
\leq\;\; & p^{\frac{1}{\alpha}}\left[A_r\left(|\beta_1| + |\beta_2|\right)^\alpha + (A_1 + A_2)\right]^{\frac{1}{\alpha}} \\
& \text{Using Triangular inequality} \\[4pt]
=\;\; & p^{\frac{1}{\alpha}}\left[A_r\left(|\beta_1| + |\beta_2|\right)^\alpha + \left((A_1 + A_2)^{\frac{1}{\alpha}}\right)^\alpha\right]^{\frac{1}{\alpha}} \\[4pt]
\leq\;\; & p^{\frac{1}{\alpha}}\left[A_r\left(|\beta_1| + |\beta_2|\right)^\alpha + \left(A_1^{\frac{1}{\alpha}} + A_2^{\frac{1}{\alpha}}\right)^\alpha\right]^{\frac{1}{\alpha}} \\
& \text{Using } C_\alpha \text{ inequality for } \alpha > 1 \\[4pt]
=\;\; & p^{\frac{1}{\alpha}}\left[\left(A_r^{\frac{1}{\alpha}}|\beta_1| + A_r^{\frac{1}{\alpha}}|\beta_2|\right)^\alpha + \left(A_1^{\frac{1}{\alpha}} + A_2^{\frac{1}{\alpha}}\right)^\alpha\right]^{\frac{1}{\alpha}} \\[4pt]
\leq\;\; & p^{\frac{1}{\alpha}}\left[\left(\left(A_r^{\frac{1}{\alpha}}|\beta_1|\right)^\alpha + \left(A_1^{\frac{1}{\alpha}}\right)^\alpha\right)^{\frac{1}{\alpha}} + \left(\left(A_r^{\frac{1}{\alpha}}|\beta_2|\right)^\alpha + \left(A_2^{\frac{1}{\alpha}}\right)^\alpha\right)^{\frac{1}{\alpha}}\right], \\
& \text{Using Minkowski's Inequality for } \alpha > 1 \\[4pt]
=\;\; & p^{\frac{1}{\alpha}}\left(A_r|\beta_1|^\alpha + A_1\right)^{\frac{1}{\alpha}} + p^{\frac{1}{\alpha}}\left(A_r|\beta_2|^\alpha + A_2\right)^{\frac{1}{\alpha}} \\[4pt]
=\;\; & VaR_p(X_1) + VaR_p(X_2)
\end{aligned}
$$

Thus, for $\alpha > 1$, VaR is sub-additive.

Table 1 Simulation from Student's t-distribution: realistic sample size

N	S	p	ν_1	ν_2	ρ	n	$\frac{n}{S}$
1000	10000	0.01	1	1	0.000	4088	0.4088
1000	10000	0.05	1	1	0.000	4611	0.4611
1000	10000	0.01	1	1	0.500	4308	0.4308
1000	10000	0.05	1	1	0.500	4578	0.4578
1000	10000	0.01	2	1	0.000	95	0.0095
1000	10000	0.05	2	1	0.000	0	0.0000
1000	10000	0.01	2	1	0.500	867	0.0867
1000	10000	0.05	2	1	0.500	126	0.0126
1000	10000	0.01	1	3	0.000	95	0.0095
1000	10000	0.05	1	3	0.000	0	0.0000
1000	10000	0.01	1	3	0.500	867	0.0867
1000	10000	0.05	1	3	0.500	126	0.0126
1000	10000	0.01	3	3	0.000	3	0.0003
1000	10000	0.05	3	3	0.000	0	0.0000
1000	10000	0.01	3	3	0.500	165	0.0165
1000	10000	0.05	3	3	0.500	5	0.0005
1000	10000	0.01	4	4	0.000	0	0.0000
1000	10000	0.05	4	4	0.000	0	0.0000
1000	10000	0.01	4	4	0.500	48	0.0048
1000	10000	0.05	4	4	0.500	1	0.0001
1000	10000	0.01	6	6	0.000	0	0.0000
1000	10000	0.05	6	6	0.000	0	0.0000
1000	10000	0.01	6	6	0.500	17	0.0017
1000	10000	0.05	6	6	0.500	0	0.0000

Table 2 Simulation from Student's t-distribution: very large Sample size

N	S	p	ν_1	ν_2	ρ	n	$\frac{n}{S}$
1000000	200	0.01	1	1	0.000	110	0.550
1000000	200	0.05	1	1	0.000	109	0.545
1000000	200	0.01	1	1	0.500	103	0.515
1000000	200	0.05	1	1	0.500	96	0.480
1000000	200	0.01	2	1	0.000	0	0.000
1000000	200	0.05	2	1	0.000	0	0.000
1000000	200	0.01	2	1	0.500	0	0.000
1000000	200	0.05	2	1	0.500	0	0.000
1000000	200	0.01	1	3	0.000	0	0.000
1000000	200	0.05	1	3	0.000	0	0.000
1000000	200	0.01	1	3	0.500	0	0.000
1000000	200	0.05	1	3	0.500	0	0.000
1000000	200	0.01	3	3	0.000	0	0.000
1000000	200	0.05	3	3	0.000	0	0.000
1000000	200	0.01	3	3	0.500	0	0.000
1000000	200	0.05	3	3	0.500	0	0.000
1000000	200	0.01	4	4	0.000	0	0.000
1000000	200	0.05	4	4	0.000	0	0.000
1000000	200	0.01	4	4	0.500	0	0.000
1000000	200	0.05	4	4	0.500	0	0.000
1000000	200	0.01	6	6	0.000	0	0.000
1000000	200	0.05	6	6	0.000	0	0.000
1000000	200	0.01	6	6	0.500	0	0.000
1000000	200	0.05	6	6	0.500	0	0.000

Table 3 Simulation from Jump processes

N	S	p	q	n	$\frac{n}{S}$
1000	10000	0.01	0.000	2504	0.2504
1000	10000	0.01	0.050	4811	0.4811
1000	10000	0.05	0.000	0	0.0000
1000	10000	0.05	0.050	0	0.0000
1000000	200	0.01	0.000	90	0.4500
1000000	200	0.01	0.050	1	0.0050
1000000	200	0.05	0.000	0	0.0000
1000000	200	0.05	0.050	0	0.0000

Table 4 Simulation from the BEKK GARCH processes

N	S	p	n	$\frac{n}{S}$
1000	10000	0.01	10	0.0010
1000	10000	0.05	2	0.0002
1000000	200	0.01	0	0.0000
1000000	200	0.05	0	0.0000