

# Proper Conditioning for Coherent VaR in Portfolio Management\*

René Garcia

Université de Montréal, CIRANO and CIREQ

Éric Renault

University of North Carolina at Chapel Hill, CIRANO and CIREQ

Georges Tsafack

Université de Montréal, CIRANO and CIREQ

First Draft: November 2003, This Version: May 2005

## Abstract

Value at Risk (VaR) is a central concept in risk management. As stressed by Artzner et al. (1999), VaR may not possess the subadditivity property required to be a coherent measure of risk. The key idea of this paper is that, when tail thickness is responsible for violation of subadditivity, eliciting proper conditioning information may restore VaR rationale for decentralized risk management. The argument is threefold. First, since individual traders are hired because they possess a richer information on their specific market segment than senior management, they just have to follow consistently the prudential targets set by senior management to ensure that decentralized VaR control will work in a coherent way. The intuition is that if one could build a fictitious conditioning information set merging all individual pieces of information, it would be rich enough to restore VaR subadditivity. Second, in this decentralization context, we show that if senior management has access ex-post to the portfolio shares of the individual traders, it amounts to recovering some of their private information. These shares can be used to improve backtesting in order to check that the prudential targets have been enforced by the traders. Finally, we stress that tail thickness required to violate subadditivity, even for small probabilities, remains an extreme situation since it corresponds to such poor conditioning information that expected loss appears to be infinite. We then conclude that lack of coherency of decentralized VaR management, that is VaR non-subadditivity at the richest level of information, should be an exception rather than a rule.

*Keywords:* Value at risk, Decentralized risk management, Coherent measures of risk, Subadditivity of VaR, Heavy-tail distributions, Stable distributions

*JEL classification:* G1, C1, C4

---

\*The authors gratefully acknowledge financial support from the Institut de Finance Mathématique de Montréal (IFM2), the Fonds québécois de la recherche sur la société et la culture (FQRSC), the Social Sciences and Humanities Research Council of Canada (SSHRC), and the Network of Centres of Excellence MITACS. The first author also thanks Hydro-Québec and the Bank of Canada for financial support. Address for correspondence: Département de Sciences Économiques, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal (Québec) H3C 3J7, Canada.

# 1 Introduction

Value at risk (VaR) - the amount of money such that there is typically a 95% or 99% probability of a portfolio losing less than that amount over a certain horizon, has become a central concept in risk management. Financial institutions, regulators as well as non-financial corporations use this method to measure financial risk. Although the VaR risk measure seems to agree with a concept of maximum loss popular with practitioners, it is not a coherent measure of risk, as stressed by Artzner et al.(1999), since it is not subadditive. Subadditivity appears as a natural requirement in a decentralization context within a financial institution. Suppose the supervisory unit in charge of portfolio management, say in a pension fund, wants to decentralize the management of certain parts of the portfolio to specialists of market segments. If the unit wants to impose a global VaR amount on the whole portfolio, subadditivity will allow to decentralize its VaR constraint into several VaR constraints, one per specialist. The supervisory unit will then be assured that the VaR of the global risk will not surpass the sum of the individual VaRs.

Artzner et al.(1999) give examples of the nonsubadditivity of VaR and conclude that it creates severe aggregation problems since it does not behave nicely with respect to the addition of risks, even independent ones. A notable exception is the case where all risks are jointly normally distributed, since the quantiles satisfy subadditivity as long as probabilities of exceedence are smaller than 0.5. In this paper, we provide an analysis of the robustness of subadditivity for a portfolio VaR under various departures from the joint normality case.

The main idea is that, when tail thickness is responsible for violation of subadditivity, eliciting proper conditioning information may restore VaR rationale for decentralized risk management, in spite of spurious evidence of the contrary. This provides an argument for the robustness of VaR coherency in practice, that is for decentralized portfolio management often called Rent-a-Trader system. The specialists are hired because they have access to specific information on which they condition their portfolio decisions. Therefore, the central management unit possesses only a subset of the conditioning information which belongs to each specialist. Naturally, in such a context, a distribution appears always more leptokurtic to the central unit than to the specialist. Because of a lack of information, VaR may appear non subadditive to the central management unit, but without bad consequences for the actual risk incurred. We are then able to show that decentralized portfolio management with a VaR allocation to each specialist will work despite evidence to the contrary. VaR is therefore decentralizable if specialists obey their VaR requirements. Of course, central management will still want to assess VaR exceedence. We distinguish the case where they have access to some information from the case where they have to rely on unconditional information only.

We first provide an illustration where traders have access to private information, which is unobservable to both central management and the other traders, but where they communicate their individual portfolio shares to central management. We discuss ways for central management to improve backtesting of VaR requirements in this context.

In the absence of conditioning information such as the make-up of portfolios, central management can only rely on thicker tailed distributions to assess VaR exceedence. Therefore, we analyze VaR subadditivity in the context of portfolios with heavy-tailed return distributions. We use stable distributions to characterize the unconditional return distributions of traders' portfolios. By appealing to a mixture representation property of stable distributions, we link formally the Rent-a-Trader system to stable distributions. More generally, as shown in early work by Mandelbrot (1963) and Fama (1965), stable distributions accommodate well heavy-tailed financial series, with the consequence that it produces measures of risk based on the tails of the distribution, such as value at risk, which are more reliable (see in particular Mittnik and Rachev, 1993, and Mittnik, Paoletta and Rachev, 2000). We generally investigate the robustness of VaR subadditivity in the context of portfolios with stable distributions.

We show that for stable distributions, the subadditivity property is maintained for realistic levels of the probability of exceedence given that the loss has a finite expectation. The result stands even if the segments of the portfolio exhibit skewness, as long as this skewness is not different among the segments. We then study the robustness of VaR subadditivity in the more general case of heavy-tailed distributions whose tails decay like power functions. We show that VaR remains subadditive, at least for sufficiently small probabilities, again when the expected loss is finite. In this respect we are inclined to conclude that non-subadditive VaR remains an extreme situation.

The rest of the paper is organized as follows. Section 2 details the Rent-a-Trader system whereby portfolio management is decentralized to specialists. We show that deconditioning by central management always increase tail fatness and spuriously make VaR look non-subadditive. We provide a simple illustration with private signals to traders and show how the transmission of information in the form of portfolio shares helps to assess risk more accurately. Section 3 specializes the discussion of deconditioning to stable distributions and extends the analysis of the robustness of VaR subadditivity to heavy-tailed distributions in general. Section 4 concludes. Proofs of the various propositions are collected in Appendices.

## **2 Decentralized VaR control in a Rent-a-Trader System**

In many financial institutions, a central management unit delegates the investment decisions for parts of the portfolio to specialists either inside or outside the firm. These specialists

know very well specific segments of the financial markets and their specific information is perceived to add value to the general portfolio management process. The usual way to proceed is for senior management to allocate some VaR targets to the various traders. The main questions which arise in such a context is to determine if the VaR of the institutional portfolio is respected as long as the specialists respect their own VaR, all these VaRs being set at a certain level of confidence. This question translates directly into the subadditivity property of the VaR risk measure but in a particular way. Evaluating VaR exceedence in a context of decentralization raises an additional issue since the same information is not shared by every segment of the portfolio. In fact, the very usefulness of decentralization is precisely to try to exploit the informational advantage of the specialists.

In this section we set up the problem accounting for the specific information of the individual traders and the control problem of central management. The main message is that even though senior management might perceive violations of subadditivity given its coarse information set, these violations do not prevent controlling an aggregate VaR through decentralized management. The rationale for this apparent paradox is that violations of subadditivity disappear when conditioning on the finer information set made up of all the specific information sets of the individual traders. Of course, this finer information is not available to central management and aggregate VaR control can only be achieved insofar as individual traders use their private information, while consistently following the prudential rules imposed by the senior management. When it comes to backtesting of VaR exceedences by the individual traders, individual portfolio shares, which are usually available to senior management, could serve as a partial substitute to traders' private information sets.

## 2.1 VaR Decentralization with Differential Information

Senior management is interested in the value at risk  $VaR_p(X)$  associated to the random value  $X$  of its portfolio:

$$P[X \leq -VaR_p(X)] = p. \quad (2.1)$$

This value  $X$  will be conformable to a VaR requirement  $VaR_p^*$  if and only if:  $VaR_p(X) \leq VaR_p^*$ , that is if:

$$P[X \leq -VaR_p^*] \leq p. \quad (2.2)$$

But  $X$  is the aggregation of the net results  $X_j$  of  $n$  traders  $j = 1, \dots, n$ :

$$X = \sum_{j=1}^n X_j \quad (2.3)$$

Trader  $j$  has obtained result  $X_j$  by building a portfolio  $\theta_j(I_j)$ , which is a function of his private information  $I_j$ .

Let us consider that trader  $j$  has received from senior management a target  $S_j$  in terms of VaR, that is:

$$VaR_p(X_j | I_j) \leq S_j \tag{2.4}$$

where:

$$P[X_j \leq -VaR_p(X_j | I_j) | I_j] = p. \tag{2.5}$$

Note that  $S_j$  is a given number chosen by senior management. Typically, the quantity  $VaR_p(X_j | I_j)$ , which depends on private information, cannot be observed at the central level. Therefore, senior management cannot check directly that the requested target (2.4) has been met or equivalently that:

$$P[X_j \leq -S_j | I_j] \leq p. \tag{2.6}$$

In this respect, the purpose of backtesting will be twofold. First, as usual, senior management must check on a time series of portfolios returns that  $X = \sum_{j=1}^n X_j$  fulfills the VaR requirement (2.2). It is quite natural for central management to imagine that this requirement will be ensured by the enforcement of targets  $S_j$  if and only if these targets are chosen ex-ante in order to fulfill:

$$\sum_{j=1}^n S_j \leq VaR_p^*. \tag{2.7}$$

Second, even though (2.6) cannot be observed, senior management should be interested to seek valuable information about individual trader  $j$  behavior. Of course, historical observation allows him to test for a weak consequence of (2.6), that is:

$$P[X_j \leq -S_j] \leq p. \tag{2.8}$$

But, for the targets to appear credible, a tighter control should be performed. The purpose of the next section is to show that, under a set of natural assumptions, both goals may be met. In other words, it will be true that the enforcement of targets conformable to (2.7) will ensure (2.2). Moreover, senior management will have at its disposal some relevant information for a more powerful test of (2.6) than only through its weak consequence (2.8).

## 2.2 Proper Conditioning for Subadditive VaR

The first crucial assumption amounts to consider that, even though sub-additivity of VaR is not guaranteed at the senior management level, there exists a latent information  $I$ , nesting all individual information sets, such that the (unfeasible) conditioning by this information would restore subadditivity of VaR:

*Assumption A1: There exists  $I \supset \cup_{j=1}^n I_j$  such that*

$$VaR_p(\sum_{j=1}^n X_j | I) \leq \sum_{j=1}^n VaR_p(X_j | I) \quad (2.9)$$

The larger is the conditioning information set, the more likely is the proximity to normality that is required for subadditivity. However, there is a bound to this data augmentation strategy insofar as one wants to get something relevant for decentralization of risk management. This upper bound for latent information is actually produced by assumption A2:

*Assumption A2: For all  $j = 1, \dots, n$ :*

$$VaR_p(X_j | I) \leq VaR_p(X_j | I_j) \quad (2.10)$$

It is worth noticing that assumption A2 is fulfilled in particular if for all  $j$ , the conditional distributions of  $X_j$  given  $I$  or  $I_j$  coincide. In other words, latent information other than  $I_j$  is irrelevant for forecasting the result  $X_j$  of trader's  $j$  investment. This latter condition, a kind of cross-sectional equivalent to a non-causality assumption (external information does not cause  $X_j$  given  $I_j$ ) is fairly natural if one imagines trader  $j$  as an expert of his market segment. Trader  $j$  has at his disposal all the relevant information for his market segment.

However, assumption A2 is more general than this special case of cross-sectional non-causality. It only means that the part of latent information that trader  $j$  does not observe does not affect his perceived potential loss with probability  $p$ . In particular, we have:

**Proposition 2.1:** *Assumption A2 implies that,  $I_j$  almost certainly:*

$$X_j \leq -VaR_p(X_j | I_j) \Leftrightarrow X_j \leq -VaR_p(X_j | I) \quad (2.11)$$

Under assumption A2, taking into account the larger latent information does not change,  $I_j$  almost certainly, the occurrence of VaR exceedence for  $j$ . Note that the converse is not true in general even though we have, by the law of iterated expectations:

$$I_j \subset I \implies P[X_j \leq -VaR_p(X_j | I) | I_j] = p = P[X_j \leq -VaR_p(X_j | I_j) | I_j] \quad (2.12)$$

But the equality of probabilities does not imply the equality of events. The most important result of this section is stated in the following proposition.

**Proposition 2.2:** *Under assumptions A1 and A2,  $\sum_{j=1}^n S_j \leq VaR_p^*$  and  $VaR_p(X_j | I_j) \leq S_j$  for all  $j$  implies that:*

$$P[\sum_{j=1}^n X_j \leq -VaR_p^*] \leq p \quad (2.13)$$

Proofs for these two propositions are provided in Appendix 1.

In other words, the VaR target  $VaR_p(X_j | I_j) \leq S_j$  imposed to each specialist  $j = 1, \dots, n$  ensures that the VaR of the portfolio  $\sum_{j=1}^n X_j$  will not exceed  $\sum_{j=1}^n S_j$ . It is worth reminding that this result has been obtained while VaR may not be subadditive for senior management, that is  $VaR_p(\sum_{j=1}^n X_j)$  may exceed  $\sum_{j=1}^n VaR_p(X_j)$ . This convenient result has basically been obtained thanks to additional conditioning which has restored sub-additivity without introducing additional perceived risk thanks to assumptions A1 and A2.

As already mentioned, assumption A2 is also useful for the purpose of backtesting, or more precisely for ex-post control of the risk taking behaviour of the specialists. Senior management would like to check that the announced target  $S_j$  has been respected by specialist  $j$ , that is:

$$VaR_p(X_j | I_j) \leq S_j \quad (2.14)$$

While it cannot be checked directly since the specialist's information set  $I_j$  cannot be observed, assumption A2 is telling central management that a necessary condition for the target to be respected is:

$$VaR_p(X_j | I) \leq S_j, \quad (2.15)$$

that is :

$$P[X_j \leq -S_j | I] \leq p. \quad (2.16)$$

Although senior management cannot observe the information set  $I$ , it has access to some partial information such as the specialists' actions. Let us assume, at it is often the case in practice, that the portfolio composition  $\theta_k(I_k)$  selected by each trader  $k = 1, \dots, n$  is available to central management. Since by definition:

$$\theta_k(I_k) \in I_k \subset I$$

inequality (2.13) implies, by application of the law of iterated expectations, that:

$$P[X_j \leq -S_j | \theta_k(I_k), k = 1, \dots, n] \leq p \quad (2.17)$$

This condition can actually be tested by senior management from an econometric model of conditional probability distributions, including for instance ARCH effects. Then, the control over trader  $j$  behaviour appears a priori much more powerful than the solely unconditional control  $P[X_j \leq -S_j] \leq p$  that could have been performed without resorting to assumption A2. An example is given in the next section.

### 2.3 A Simple Illustration of a Rent-a-Trader System

We provide an illustration of the general propositions of the previous sections in a simple setting. We assume that there exists two traders and that each can choose a portfolio made up of one risk-free asset and two risky funds. The returns of the two risky funds depend on two state variables  $s_1$  and  $s_2$ . State variable  $s_1$  is observable to trader 1, but unobservable to trader 2 and to central management. Similarly, trader 2 is the only one to observe  $s_2$ . The two state variables are assumed to be *i.i.d. Bernoulli* ( $\lambda$ ). This setting fulfills both assumptions A1 and A2. Subadditivity is maintained when conditioning on both  $s_1$  and  $s_2$ , and the example is constructed to ensure cross-sectional non-causality between traders 1 and 2.

We can write the returns as  $\tilde{R}_1 = s_1 R_1^1 + (1 - s_1) R_1^0$ , and  $\tilde{R}_2 = s_2 R_2^1 + (1 - s_2) R_2^0$ , where  $R_1^1, R_2^1, R_1^0$  and  $R_2^0$  are mutually independent with  $R_1^1$  and  $R_2^1$  following the same probability distribution  $N(\mu_1, \bar{\sigma}^2)$  and  $R_1^0$  and  $R_2^0$  following  $N(\mu_0, \bar{\sigma}^2)$ . Moreover, the unconditional mean  $\mu$  of the two distributions is assumed to be equal to the risk-free rate<sup>1</sup>. These assumptions imply that, without any information on the state variables, a risk averse agent will only invest in the risk free asset. Therefore, central management will have an incentive to hire traders 1 and 2, who have private information on state variables  $s_1$  and  $s_2$  respectively. In this context, if each trader forms his portfolio such that the VaR requirement imposed by central management is satisfied conditionally to any possible value of his private information, then the VaR requirement of the global portfolio will be satisfied and the *apparent* violation of subadditivity will not matter.

We further assume that each trader communicates his portfolio shares to central management. We show how, based on this information, central management can recover statistically the parameters of the conditional distributions of the traders' portfolio returns and assess whether traders have respected the VaR requirement or not.

---

<sup>1</sup>It is important to realize that funds 1 and 2 have the same conditional means  $\mu_1$  in state 1 and  $\mu_0$  in state 0 and the same conditional variance  $\bar{\sigma}^2$  in any state. They differ only by the realization of the states, which do not necessarily coincide. For example, fund 1 could be in state 0 when fund 2 is in state 1.

### 2.3.1 Trader behavior

We start by computing the optimal mean-variance portfolio of traders 1 and 2 given their private information on  $s_1$  and  $s_2$  respectively. We assume for simplicity that the VaR of their optimal portfolio is always below  $S_j^2$ , the target set by central management. A first result is that a trader will never invest in the risky fund for which he does not have any information<sup>3</sup>. Moreover, traders will always include a non-zero share of their respective risky fund in their optimal portfolio along with the risk-free asset. In the good state, they will hold a long position, in the bad state they will sell the risky fund short<sup>4</sup>.

We can write the portfolio return of trader 1 as  $\tilde{R}^1 = R_f + \theta_{11} (\tilde{R}_1 - R_f) + \theta_{12} (\tilde{R}_2 - R_f)$ . The expectation and the variance conditional on the state are:

$$\begin{aligned} E(\tilde{R}^1 | s_1 = i) &= R_f + \theta_{11} (\mu_i - R_f) \\ \text{Var}(\tilde{R}^1 | s_1 = i) &= \theta_{11}^2 \bar{\sigma}^2 + \theta_{12}^2 \sigma^2 \end{aligned} \quad (2.18)$$

Normalizing initial wealth to one, and denoting the risk aversion coefficient of trader 1 by  $\gamma_1$ , the optimal portfolio is solution of:

$$\text{Max}_{\theta_1} \left\{ (R_f + \theta_{11} (\mu_i - R_f)) - \frac{\gamma_1}{2} (\theta_{11}^2 \bar{\sigma}^2 + \theta_{12}^2 \sigma^2) \right\}$$

with  $\theta_1 = (\theta_{11}, \theta_{12})$ . The solution  $\hat{\theta}_1 = (\hat{\theta}_{11}, \hat{\theta}_{12})$  is given by

$$\begin{aligned} \hat{\theta}_{11} &= (\mu_i - R_f) / \gamma_1 \bar{\sigma}^2 \\ \hat{\theta}_{12} &= 0 \end{aligned} \quad (2.19)$$

The proportion invested in the risk-free asset is  $1 - \hat{\theta}_{11}$ . Trader 1 never invests in fund 2 for which he has no information. Intuitively, investing in asset 2 without any information on  $s_2$  will mean investing in an asset which has the same mean as the risk-free asset but a higher variance. Therefore, a simple second-order stochastic dominance argument implies that any risk-averse agent will not invest in asset 2. Additionally, in the bad state ( $\mu_1 < R_f$ ), we have  $\hat{\theta}_{11} < 0$ , which means that the trader is short selling fund 1, for which he has information. In the good state ( $\mu_0 > R_f$ ),  $\hat{\theta}_{11} > 0$  and trader 1 takes a long position in fund 1. Figure 1 illustrates these results.

<sup>2</sup>When the VaR constraint of trader 1 (resp. 2) binds, it can be shown that he may have to invest a nonzero part in asset 2 (resp. 1).

<sup>3</sup>In a general framework, Merton (1987) assumes this result and justifies his assumption by the fact that the portfolios held by actual investors contain only a small fraction of the thousand of traded securities available. In our setting, the result follows directly from the private information held by the traders.

<sup>4</sup>In a framework with only two assets (a risk-free asset and a risky asset with the same unconditional expected return), Sentana (1998) assumes and rationalizes the fact that the wealth invested in the risky asset is a linear function of the information that the investor has on this asset.

### 2.3.2 Subadditivity and Coherence of VaR when properly conditioned

In this context, we show that VaR may appear non subadditive to central management, while in fact it remains subadditive conditionally to the private information of the traders. To show this, we assume for simplicity that both traders have the same risk aversion  $\gamma$  and that their initial wealth is normalized to one. We further assume that both information variables  $s_1$  and  $s_2$  are in the good state ( $s_1 = s_2 = 0$ ). Therefore, it follows that  $\hat{\theta}_{11} = \hat{\theta}_{22} = \theta > 0$ , traders' portfolios are  $\bar{R}^1 = (1 - \theta) R_f + \theta \bar{R}_1$  and  $\bar{R}^2 = (1 - \theta) R_f + \theta \bar{R}_2$  and the aggregate portfolio at central management level is  $\bar{R}^1 + \bar{R}^2 = 2(1 - \theta) R_f + \theta (\bar{R}_1 + \bar{R}_2)$

First, let us start by the apparent non subadditivity of VaR for central management. Proposition 2.3 below shows that even for small probability  $p$ , the VaR subadditivity of the central portfolio can be apparently seen as violated due to the asymmetric information between traders and central management.

**Proposition 2.3:** *For any mixture probability values  $\lambda < 1/2$ , at the level  $p = \Pr(\bar{R}_1 + \bar{R}_2 \leq \mu_1 + \mu_0)$  we have*

$$VaR_p(\bar{R}^1 + \bar{R}^2) > VaR_p(\bar{R}^1) + VaR_p(\bar{R}^2)$$

and

$$p = \lambda^2 \Phi(-\bar{\mu}/\sqrt{2\bar{\sigma}}) + \lambda(1 - \lambda) + (1 - \lambda)^2 \Phi(\bar{\mu}/\sqrt{2\bar{\sigma}})$$

where  $\Phi$  denotes the cumulative standard normal distribution.

Proof: see Appendix 1.

However, conditionally to information on both states, all returns become normal and then it is well-known that for all  $p \leq 0.5$ , subadditivity follows. Therefore, with  $\lambda < 1/2$ , at the level  $p = \Pr(\bar{R}_1 + \bar{R}_2 \leq \mu_1 + \mu_0)$ , we have shown that without conditioning on private information, decentralization may appear as increasing the risk of the aggregate portfolio as measured by the VaR, while in fact it does not since conditionally to the private information of traders, returns become normal and subadditivity prevails. In the next section, we provide central management with some information, namely the composition of the traders' portfolios.

### 2.3.3 Backtesting VaR requirements

In this simple model, a central unit can test condition (2.17), which can be written  $P[X_j \leq -S_j | \theta_k(I_k), k = 1, 2] \leq p$  for all  $j$ , when the shares of the portfolios held by traders 1 and 2 are known. It is important to realize that in this setting, knowing the portfolio

composition of traders 1 and 2 is equivalent to knowing their private information  $s_1$  and  $s_2$ . Indeed, if trader 1 takes a long position in risky fund 1 it means that  $s_1$  is in the good state and inversely if he short-sells fund 1. Similarly, the position of trader 2 will be fully revealing. We can write  $s_j = 1_{\{\theta_{jj} < 0\}}$ . Since each private information completely defines the return distribution of the fund in which the corresponding trader invests and given that the two informations  $s_1$  and  $s_2$  are independent, the condition to test is exactly  $P[X_j \leq -S_j | \theta_j(s_j)] \leq p$  for  $j = 1, 2$ .

By the conditional normality of the return distributions, central management needs only to infer the means and variances in the good and bad states for each trader in order to test if each trader obeys his VaR limit with probability  $p$ . In section 2, we assume that central management knows the underlying model. In practice, central management must estimate the model based on time series of returns and compute *VaR* conditionally on past returns. In this dynamic setting, assumption 2 can be rewritten as:

$$VaR_p [X_j | I_t, I_\tau, \tau \leq t] \leq VaR_p [X_j | I_{jt}, I_\tau, \tau \leq t]. \quad (2.20)$$

In other words, we will assume that all past information becomes common knowledge. Propositions equivalent to Propositions 2.1 and 2.2 can be derived in a dynamic context. We provide below an example of a filtering procedure that central management can follow to recover the needed information about the returns distributions based on the available information on portfolio shares.

To start, central management can assume that conditional returns of a particular trader are normal and follow a switching-GARCH model for a given number of states:

$$\begin{cases} R_{t+1} = \mu_{s_t t} + \sigma_{s_t t} \varepsilon_{t+1} \\ \sigma_{s_t t}^2 = \omega_{s_t} + \alpha_{s_t} \sigma_{s_{t-1} t-1}^2 + \beta_{s_t} \left( R_t - \mu_{s_{t-1} t-1} \right)^2 \end{cases} \quad (2.21)$$

where  $\varepsilon_{t+1} \sim N(0, 1)$  and  $s_t$  is a Markovian process taking two values 1 or 2 with transition matrix  $\Lambda$ . By using the optimal portfolio share in a mean-variance context as in (2.19), central management can use the information on the portfolio shares and the publicly available risk-free rate to estimate the model:

$$\begin{aligned} R_{t+1} &= R_{ft} + \gamma \sigma_{s_t t}^2 \theta_t + \sigma_{s_t t} \varepsilon_{t+1} \\ \sigma_{s_t t}^2 &= \omega_{s_t} + \alpha_{s_t} \sigma_{s_{t-1} t-1}^2 + \beta_{s_t} \left( R_t - R_{ft-1} - \gamma \sigma_{s_{t-1} t-1}^2 \theta_{t-1} \right)^2 \end{aligned} \quad (2.22)$$

This model can be estimated by using for example the procedure described in Gray (1996), which applies the filtering algorithm of Hamilton (1989) in a GARCH model. In this procedure, to avoid an infinite recursion, the state-dependent conditional variance  $\sigma_{s_{t-1} t-1}^2$  is replaced by the state-independent conditional variance  $\sigma_{t-1}^2$  given by:

$$\begin{aligned}
\sigma_{t-1}^2 = & \lambda_{t-1} \left[ (R_{ft-1} + \rho\sigma_{1t-1}^2\theta_{t-1})^2 + \sigma_{1t-1}^2 \right] \\
& + (1 - \lambda_{t-1}) \left[ (R_{ft-1} + \rho\sigma_{2t-1}^2\theta_{t-1})^2 + \sigma_{2t-1}^2 \right] \\
& - \left[ \lambda_{t-1} (R_{ft-1} + \rho\sigma_{1t-1}^2\theta_{t-1}) + (1 - \lambda_{t-1}) (R_{ft-1} + \rho\sigma_{2t-1}^2\theta_{t-1}) \right]^2
\end{aligned} \tag{2.23}$$

where  $\lambda_t$  is the probability that  $s_t = 1$  conditional on the information available at time  $t$ . This probability is the by-product of the maximum likelihood estimation. The estimation also provides the means and variances of the trader's portfolio returns in both states. One can therefore compute the conditional VaR with probability  $p$  and compare it to the VaR requirement set by central management.

Our main argument until now has been to illustrate how conditioning and deconditioning could affect the VaR subadditivity of a portfolio in the context of decentralized management. In the example just developed, the unconditional distribution of returns was a mixture of normals. In the absence of information on the make-up of portfolios, central management can only rely on this thicker tailed distribution to assess VaR exceedence. In the next section, we analyze VaR subadditivity in the context of portfolios with heavy-tailed return distributions. This realistic setting for financial applications will provide a sense of how robust is the subadditivity property for most practical purposes.

### 3 Stable distributions, Deconditioning and VaR Subadditivity

In this section, we use stable distributions to characterize the unconditional return distributions of traders' portfolios. By appealing to a mixture representation property of stable distributions, we link formally the Rent-a-Trader system of the last section to stable distributions. More generally, as shown in early work by Mandelbrot (1963) and Fama (1965), stable distributions accommodate well heavy-tailed financial series, with the consequence that it produces measures of risk based on the tails of the distribution, such as value at risk, which are more reliable (see in particular Mitnik and Rachev, 1993, and Mitnik, Paoletta and Rachev, 2000). The main purpose of this section is to investigate the robustness of VaR subadditivity in the context of portfolios with stable distributions.

#### 3.1 The Stable Distribution

A random variable  $X$  is said to have a stable distribution if for any positive numbers  $a$  and  $b$ , there is a positive number  $c$  and a real number  $d$  such that

$$aX_1 + bX_2 \stackrel{d}{=} cX + d$$

where  $X_1$  and  $X_2$  are independent copies of  $X$ , and where  $\stackrel{d}{=}$  denotes equality in distribution. Then, it can be shown that there is a number  $\alpha \in ]0, 2]$ , called the index of stability, such that:  $c^\alpha = a^\alpha + b^\alpha$ . For sake of expositional simplicity, we will assume in what follows that  $\alpha$  is different from one. From the Lévy-Khinchine representation of the characteristic function, the characteristic function of a stable random variable can be written as:

$$E \exp(i\theta X) = \exp \left\{ -\sigma^\alpha |\theta|^\alpha \left[ 1 - i\beta (\text{sign } \theta) \tan \frac{\Pi\alpha}{2} \right] + i\mu\theta \right\} \quad (3.1)$$

The parameters  $\sigma$ ,  $\beta$  and  $\mu$  are unique ( $\beta$  is irrelevant when  $\alpha = 2$ ) and it is clear that  $\mu$  is a location parameter and  $\sigma$  is a scale parameter.

In other words, while (3.1) defines the  $S_\alpha[\sigma, \beta, \mu]$  stable distribution:

$$\begin{aligned} X &\rightsquigarrow S_\alpha[\sigma, \beta, \mu] \\ &\Leftrightarrow \frac{X - \mu}{\sigma} \rightsquigarrow S_\alpha[1, \beta, 0] \end{aligned} \quad (3.2)$$

Since the characteristic function is real if and only if  $\mu = \beta = 0$ , we see that  $\beta$  is a skewness parameter.  $S_\alpha[\sigma, \beta, \mu]$  is symmetric around  $\mu$  if and only  $\beta = 0$ .

$$X \rightsquigarrow S_\alpha[\sigma, \beta, 0] \Leftrightarrow (-X) \rightsquigarrow S_\alpha[\sigma, -\beta, 0]$$

The distribution  $S_\alpha(\sigma, \beta, 0)$  is said to be:

$$\left\{ \begin{array}{l} \text{Skewed to the right if } \beta > 0 \\ \text{and totally skewed if } \beta = 1. \\ \text{Skewed to the left if } \beta < 0 \\ \text{and totally skewed if } \beta = -1. \end{array} \right.$$

For  $-1 < \beta < 1$ , the support of  $S_\alpha(\sigma, \beta, \mu)$  is the whole real line while the support of the distribution  $S_\alpha(\sigma, 1, 0)$  (resp.  $S_\alpha(\sigma, -1, 0)$ ) is the positive half line (resp. the negative half line).

### 3.2 Linking the Rent-a-Trader System to Stable Distributions

In this section, we rationalize the use of stable distributions in the context of the Rent-a-Trader system of the last section. We argued that conditioning on the information of specialists can reduce the thickness of the tails and therefore make the conditional VaR subadditive whereas the unconditional VaR appeared non-subadditive. For stable distributions, we recall below that a symmetric stable distribution can be seen as a mixture of normal distributions or again as a mixture of stable distributions with less fat tails. Therefore, we appeal to this mixture representation to make the link with the decentralized portfolio management system setting discussed earlier.

**Proposition 3.1.:** Let  $X$  be a random variable with a symmetric  $\alpha$ - stable distribution,  $X \rightsquigarrow S_\alpha(\sigma, 0, 0)$ ,  $0 < \alpha < 2$ , then  $X$  can be seen as a mixture of normal distributions:

$$X|A \rightsquigarrow N(0, \sigma^2 A), \quad (3.1)$$

or more generally as a mixture of stable distributions:

$$X|A \rightsquigarrow S_{\alpha'}(\sigma A^{1/\alpha'}, 0, 0), 0 < \alpha < \alpha', \quad (3.2)$$

and the probability distribution of the mixing variable  $A$  is defined by its Laplace transform:

$$E(\exp(-\gamma A)) = \exp(-\gamma^{\alpha/\alpha'}) \quad (3.3)$$

Proof: See Proposition 1.3.1 p. 20 in Samorodnitsky and Taqqu (1994).

The main message of this proposition is that proper conditioning of a random variable with a symmetric stable distribution can produce either a normal random variable or a stable distribution with less fat tails (higher  $\alpha$ ). In the first case we recover the VaR subadditivity for a  $p$  less than 0.5, in the second we increase the range of validity of the subadditivity property as we will see in the next subsection.

To complete the argument, we need to discuss how realistic is the symmetry assumption for distributions produced by portfolio traders. A first observation is that the central unit does not need to give a benchmark to the traders in the context of a decentralization portfolio management system. Therefore, the distribution of interest is not the deviations of the trader's portfolio returns from the benchmark, which ought to be skewed to the right, but simply the raw returns of the trader's strategy. The right skewness of the latter returns is less of a necessity. Of course the central unit may ex-post want to assess the performance of the trader relatively by benchmarking with an index or other traders in similar markets but ex-ante it is enough to control for the risk taken by the trader. Therefore, stable distributions can appropriately be used to characterize the returns of the traders' portfolios when central management does not have any information to condition upon, such as the portfolio shares. Then, the next question is to what extent is subadditivity lost when portfolio returns follow stable distributions.

### 3.3 Very Heavy-tailed Distributions and Subadditivity

We will first treat the general case of Pareto-like distributions and show that they satisfy an asymptotic convolution stability property. This allows us to get asymptotic assessments of VaR subadditivity. In the second subsection, we specialize the results to the stable distribution and obtain under certain assumptions some finite assessments of VaR subadditivity.

### 3.3.1 The General Case

The probability distribution of a real random variable  $X$  is said to be “very heavy-tailed” if, for some positive  $\alpha$ :

$$E \left[ (\text{Max}(X, 0))^\delta \right] = +\infty \text{ for all } \delta > \alpha.$$

This is in particular the case of all distributions in the maximum domain of attraction of the Frechet distribution, because they are such that:

$$1 - F(x) = x^{-\alpha} L(x) \tag{3.1}$$

for some function  $L$  on  $]0, +\infty[$  slowly varying at  $(+\infty)$ . We focus here on the particular case when the function  $L$  converges at infinity to a positive constant. It is convenient to denote this positive limit as  $L(+\infty) = \sigma^\alpha$ . Then, the distribution function  $F$  has a power right tail:

$$\bar{F}(x) = 1 - F(x) \sim \left( \frac{\sigma}{x} \right)^\alpha \tag{3.2}$$

For expositional simplicity, we will assume in addition that  $F$  is continuous and strictly increasing in some neighborhood of infinity. This allows us to characterize the limit  $\sigma^\alpha = L(+\infty)$  from the tail behavior of a well-defined inverse function  $\bar{F}^{-1}$  of  $\bar{F}$ :

$$\sigma = \lim_{q \rightarrow 1} q^{1/\alpha} \bar{F}^{-1}(q) \tag{3.3}$$

Distributions functions conformable to (3.2) and (3.3) are often termed “Pareto-like” and include, besides the Pareto distribution itself, the Cauchy and Burr distributions as well as stable distributions with exponent  $\alpha < 2$ .

As far as subadditivity of VaR is concerned, the nice feature of Pareto like distributions is to fulfill an asymptotic convolution stability property. This property is fulfilled on the whole support by stable distributions themselves. More precisely, the following lemma is a straightforward application of Feller convolution theorem (Feller, 1971, p. 278):

**Lemma 3.2:** *If  $X$  and  $Y$  are two independent Pareto-like random variables, with the same exponent  $\alpha$ , that is:*

$$\begin{aligned} 1 - F_X(x) &\underset{x \rightarrow +\infty}{\sim} \left( \frac{\sigma_X}{x} \right)^\alpha \\ 1 - F_Y(x) &\underset{x \rightarrow +\infty}{\sim} \left( \frac{\sigma_Y}{x} \right)^\alpha \end{aligned}$$

where  $F_X$  and  $F_Y$  denote respectively the distribution functions of  $X$  and  $Y$ , then  $X + Y$  is also Pareto-like with the same exponent  $\alpha$  and:

$$1 - F_{X+Y}(x) \underset{x \rightarrow +\infty}{\sim} \frac{\sigma_X^\alpha + \sigma_Y^\alpha}{x^\alpha}$$

From now on, the Pareto-like distribution will be applied to losses. In other words, when  $X$  denotes the end-of-period wealth of interest, we maintain the assumption that  $(-X)$  follows a Pareto-like distribution. Therefore, with obvious notations, formula (3.3) becomes:

$$\lim_{p \rightarrow 0} VaR_p(X)p^{1/\alpha} = \sigma_X \quad (3.4)$$

Then, the asymptotic convolution stability allows us to use the convexity/concavity properties of the function  $\sigma \rightarrow \sigma^\alpha$  to get asymptotic assessments of VaR subadditivity. This is the result of proposition 4.1 below.

**Proposition 3.2**

*If  $X$  and  $Y$  are two independent random variables such that  $(-X)$  and  $(-Y)$  follow Pareto-like distributions with the same exponent  $\alpha$  :*

*(i) If  $\alpha > 1$  :*

*There exists  $p_0 \in ]0, 1]$  such that, for any  $p \in ]0, p_0[$  :*

$$VaR_p [X + Y] < VaR_p(X) + VaR_p(Y)$$

*(ii) If  $\alpha < 1$  :*

*There exists  $p_0 \in ]0, 1[$  such that, for any  $p \in ]0, p_0[$  :*

$$VaR_p [X + Y] > VaR_p(X) + VaR_p(Y)$$

*(iii) If  $\alpha = 1$  :*

$$\lim_{p \rightarrow 0} p [VaR_p (X + Y) - VaR_p(X) - VaR_p(Y)] = 0$$

Proof: See Appendix 2.

Proposition 3.2 shows that VaR is subadditive, at least for sufficiently small probabilities, when the possible loss has finite expectation ( $\alpha > 1$ ). In this respect, non-subadditive VaR remains quite an extreme situation<sup>5</sup>.

However, in the case of very heavy tail distributions ( $\alpha$  close to zero), violation of subadditivity may be arbitrarily extreme. We propose below a way to measure the degree of violation of subadditivity. While non-subadditivity means:

$$VaR_p (X + Y) > VaR_p(X) + VaR_p(Y), \quad (3.5)$$

---

<sup>5</sup>Recently, Ibragimov (2004) obtained a similar result with a different approach based on the analysis of majorization properties of linear combinations of random variables.

that is the loss of the portfolio  $(X + Y)$  can exceed  $VaR_p(X) + VaR_p(Y)$  with probability  $p$ , the question is: with what probability  $kp, k \geq 1$ , can we loose more than  $VaR_p(X) + VaR_p(Y)$  with probability  $p$ ? In other words, the question is with what probability  $kp, k \geq 1$ , can we have:

$$VaR_{kp}(X + Y) > (VaR_p(X) + VaR_p(Y)). \quad (3.6)$$

While violation (3.5) of subadditivity means that (3.6) is fulfilled with  $k = 1$ , it cannot be fulfilled with  $k = 2$  since, for any random variables  $X$  and  $Y$ :

$$VaR_{2p}(X + Y) \leq VaR_p(X) + VaR_p(Y) \quad (3.7)$$

Indeed:

$$P[X + Y \leq -VaR_p(X) - VaR_p(Y)] \leq 2p$$

$$\begin{aligned} & P[X + Y \leq -VaR_p(X) - VaR_p(Y)] \\ & \leq P[X \leq -VaR_p(X)] + P[Y \leq -VaR_p(Y)]. \end{aligned}$$

However, we can show:

**Proposition 3.3:**

For any  $k$  in  $]0, 2[$ , there exists  $\alpha_0$  in  $]0, 1[$  such that for any  $\alpha$  in  $]0, \alpha_0[$ , there exists  $p_0$  in  $]0, 1[$  such that for any  $p$  in  $]0, p_0[$ :

$$VaR_{kp}(X + Y) > VaR_p(X) + VaR_p(Y)$$

Proof: See Appendix 3.

In other words, when  $\alpha$  is arbitrarily small, the possible loss of  $X + Y$  can exceed  $VaR_p(X) + VaR_p(Y)$  with a probability  $kp$ , for  $k$  arbitrarily close to 2. Therefore, by iterating the argument, the loss of a portfolio  $\sum_{i=1}^n X_i$  can exceed  $\sum_{i=1}^n VaR_p X_i$  with a probability arbitrarily close to  $(np)$ . It is clear that such a non-subadditivity would definitely make unreliable any decentralized risk management strategy based on VaR targets assigned to the independent segments of the portfolio.

### 3.3.2 The Case of the Stable Distribution

The left tail behaviour of the stable distribution is characterized as follows.

**Proposition 3.4:** Let  $X \rightsquigarrow S_\alpha[\sigma, \beta, \mu]$  with  $0 < \alpha < 2$ .

Then:

$$\lim_{\lambda \rightarrow +\infty} \lambda^\alpha P[X < -\lambda] = C_\alpha \frac{1 - \beta}{2} \sigma^\alpha$$

$$\text{where } C_\alpha = \frac{1 - \alpha}{\Gamma(2 - \alpha) \cos(\frac{\Pi\alpha}{2})}.$$

By comparison with the notations of the previous section, we see that when  $X \rightsquigarrow S_\alpha(\sigma, \beta, \mu)$ , with  $-1 \leq \beta < 1$ , it is heavy-tailed on the left with  $\alpha$  as left tail power parameter and  $\sigma_X = \left[ C_\alpha \frac{(1-\beta)}{2} \right]^{1/\alpha} \sigma$  as left tail scale parameter. The following result expresses the convolution property for stable distributions.

**Proposition 3.5:** If  $X \rightsquigarrow S_\alpha[\sigma_1, \beta_1, \mu_1]$  and  $Y \rightsquigarrow S_\alpha[\sigma_2, \beta_2, \mu_2]$  are independent stable variables, then  $X + Y = Z \rightsquigarrow S_\alpha[\sigma, \beta, \mu]$  with:

$$\begin{cases} \sigma = [\sigma_1^\alpha + \sigma_2^\alpha]^{1/\alpha} \\ \beta = \frac{\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha}{\sigma^\alpha} \\ \mu = \mu_1 + \mu_2 \end{cases}$$

It is then clear that:

$$\begin{aligned} \sigma_X^\alpha &= C_\alpha \frac{(1 - \beta_1)}{2} \sigma_1^\alpha \\ \sigma_Y^\alpha &= C_\alpha \frac{(1 - \beta_2)}{2} \sigma_2^\alpha \\ \sigma_Z^\alpha &= C_\alpha \frac{(1 - \beta)}{2} \sigma^\alpha \end{aligned}$$

fulfill the convolution property:

$$\sigma_Z^\alpha = \sigma_X^\alpha + \sigma_Y^\alpha.$$

Therefore, all the results of the previous section can be applied to stable distributions. Besides their stability property, stable distributions are of interest since they allow to extend some results to the case of finite  $p$  (and not only  $p \rightarrow 0$ ). To see this, let us first remark that, thanks to the location-scale property:

$$\begin{aligned} X &\rightsquigarrow S_\alpha(\sigma, \beta, \mu) \\ \Rightarrow -VaR_p(X) &= \mu + \sigma \phi_{\alpha, \beta}^{-1}(p) \end{aligned}$$

where  $\phi_{\alpha, \beta}$  denotes the cumulative distribution function of  $S_\alpha[1, \beta, 0]$ .

A simple corollary of proposition 3.2 is then.

**Corollary 3.1:** For  $|\beta| < 1$  and  $p$  such that  $\phi_{\alpha, \beta}^{-1}(p) < 0$ , we have:

Let  $X, Y$  be two independent stable variables

$$\begin{cases} X \rightsquigarrow S_\alpha[\sigma_1, \beta, \mu_1] \\ Y \rightsquigarrow S_\alpha[\sigma_2, \beta, \mu_2] \end{cases}$$

(i) If  $\alpha > 1$ :

$$VaR_p [X + Y] < VaR_p(X) + VaR_p(Y).$$

(ii) If  $\alpha < 1$  :

$$VaR_p [X + Y] > VaR_p(X) + VaR_p(Y).$$

It is worth noticing that, by contrast with proposition 3.1, corollary 3.1 can surely be applied for values of the probability  $p$  which are not excessively small. For instance, if  $\beta = 0$ , corollary 3.1 applies for  $p < 1/2$ . The price to pay to get this stronger result is that we have assumed that  $X$  and  $Y$  have the same skewness parameter. This assumption, which was not needed to apply the convolution property, may appear overly restrictive to the point where only  $\beta = 0$  has some practical content. Hopefully, the subadditivity should not be lost for not too different skewness parameters.

## 4 Conclusion

In this paper, we have argued that, in the context of decentralized portfolio management, central management possesses only a fraction of information which belongs to each specialist. In such a context, a distribution appears always thicker to the central unit than to the specialist. Therefore, because of a lack of information, VaR may appear fallaciously non subadditive to the central management unit. We were then able to show that decentralized portfolio management with a VaR allocation to each specialist will work despite evidence to the contrary. Furthermore, we have shown that value at risk remains subadditive in many situations of practical interest. In the case of heavy-tail distributions, we have shown, at least for sufficiently small probabilities, that VaR remains subadditive when the possible loss has finite expectation. In this respect, non-subadditive VaR remains quite an extreme situation.

Even though we show that for all practical purposes VaR is not as incoherent a measure of risk as it is often argued, it remains that portfolio optimization practices using VaR as a simple substitute to variance, i.e. maximization of expected return under a VaR constraint) may generate perverse effects. In particular, there is a risk that a manager who is controlled only through a maximal loss level with probability  $(1 - p)$  will be enticed to expose himself to huge possible losses with probability  $p$ . To control for such a risk, one can add to VaR the expected loss beyond the VaR or consider a parameterized family of more limited possible risks. Ortobelli, Rachev and Schwartz (2000) provide a thorough analysis of optimal portfolio allocation with stable distributed returns, including with a safety-first optimal allocation problem, whereby investors maximize their expected wealth while minimizing at the same time the risk of loss. Efficient frontiers in the latter case are a function of the threshold VaR.

## References

- [1] Artzner, P., F. Delbaen, J.-M. Eber and D. Heath (1999), “Coherent Measures of Risk,” *Mathematical Finance*, 9, 3, 203-228.
- [2] Basak, S. and Shapiro, A. (2001) “Value-at-Risk-Based Risk Management: Optimal Policies and Asset Prices”, *Review of Financial Studies*, 14, 371-405.
- [3] Bollerslev, T. (1986) “Generalized Autoregressive Conditional Heteroskedasticity”, *Journal of Econometrics*, 31, 307-327
- [4] Duffie, D.J., Pan, J., (1997), “An Overview of Value at Risk”, *Journal of Derivatives*, 4, 7-49.
- [5] Engle, R. (1982) “Autoregressive Conditional Heteroskedasticity Models with estimation of Variance of United Kingdom Inflation”, *Econometrica*, 50, 987-1007.
- [6] Fama, E. (1965), “Behavior of Stock Market Prices”, *Journal of Business*, 38, 34-105.
- [7] Gray, S. F. (1996) “Modeling the conditional distribution of interest rates as a regime-switching process”, *Journal of Financial Economics*, 42, 27-62
- [8] Hamilton, J.D. (1989) “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle”, *Econometrica*, 57, 357-384.
- [9] Hamilton, J.D. (1994) *Time Series Analysis*, Princeton University Press.
- [10] Ibragimov, R. (2004), “On the Robustness of Economic Models to Heavy-Tailedness Assumptions,” Working Paper, Yale University.
- [11] Jorion, P. (2001) *Value at Risk: The New Benchmark for Managing Financial Risk*, McGraw-Hill, New York.
- [12] Mandelbrot, B. (1963), “The Variation of Speculative Prices”, *Journal of Business*, 36, 394-419.
- [13] Merton, R.C., (1987), “A Simple Model of Capital Market Equilibrium with Incomplete Information”, *Journal of Finance*, 42, 483-510.
- [14] Mittnik, S. and S. T. Rachev (1993), “Modeling Asset Returns with Alternative Stable Distributions”, *Econometric Reviews*, 12(3), 261-330.
- [15] Mittnik, S., M. S. Paolella, and S. T. Rachev (2000), “Diagnosis and treating the fat tails in financial returns data”, *Journal of Empirical Finance*, 7, 389-416.

- [16] Ortobelli, S., S. T. Rachev and E. Schwartz (2000), “The Problem of Optimal Asset Allocation with Stable Distributed Returns”, Working Paper, Anderson Graduate School of Management.
- [17] Samorodnitsky and M. S. Taqqu (1994), “Stable Non-Gaussian Random Processes - Stochastic Models with Infinite Variance,” Chapman & Hall, New York.
- [18] Sentana, Enrique (1998), “Least Squares Predictions and Mean-Variance Analysis”, CEMFI Working Paper.

## Appendix 1

### Proof of Proposition 2.1:

For  $A$  random event, we define the random variable:

$$\mathbf{1}_A = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{otherwise.} \end{cases}$$

Assumption A2 implies that:

$$\mathbf{1}_{[X_j \leq -VaR_p(X_j|I_j)]} \leq \mathbf{1}_{[X_j \leq -VaR_p(X_j|I)]}.$$

But these two random variables have the same conditional expectation given  $I_j$  since by definition:

$$\begin{aligned} E \left[ \mathbf{1}_{[X_j \leq -VaR_p(X_j|I_j)]} | I_j \right] &= \Pr \{ X_j \leq -VaR_p(X_j | I_j) | I_j \} \\ &= p \end{aligned}$$

and by the law of iterated expectations:

$$\begin{aligned} E \left[ \mathbf{1}_{[X_j \leq -VaR_p(X_j|I)]} | I_j \right] &= E \left[ E \left[ \mathbf{1}_{[X_j \leq -VaR_p(X_j|I)]} | I \right] | I_j \right] \\ &= E \left[ \Pr \{ X_j \leq -VaR_p(X_j | I) | I \} | I_j \right] \\ &= E [p | I_j] \\ &= p \end{aligned}$$

Therefore, these two random variables must coincide  $I_j$  almost surely. This achieves the proof of proposition 2.1.

### Proof of Proposition 2.2:

Since for all  $j$ ,  $VaR_p(X_j | I_j) \leq S_j$  and, by assumption A2:

$$VaR_p(X_j | I) \leq VaR_p(X_j | I_j)$$

we have:

$$p(I) = P \left[ \sum_{j=1}^n X_j \leq - \sum_{j=1}^n S_j | I \right] \leq P \left[ \sum_{j=1}^n X_j \leq - \sum_{j=1}^n VaR_p(X_j | I) | I \right]$$

But, by assumption A1:

$$VaR_p \left( \sum_{j=1}^n X_j | I \right) \leq \sum_{j=1}^n VaR_p(X_j | I).$$

Thus:

$$p(I) \leq P \left[ \sum_{j=1}^n X_j \leq -VaR_p \left( \sum_{j=1}^n X_j | I \right) | I \right] = p.$$

Since this inequality is identically true, for all possible values of the random information set  $I$ , we conclude by considering unconditional expectations that:

$$P \left[ \sum_{j=1}^n X_j \leq - \sum_{j=1}^n S_j \right] \leq p.$$

A fortiori:

$$P \left[ \sum_{j=1}^n X_j \leq -VaR_p^* \right] \leq p$$

**Proof of Proposition 2.3:**

**Proof of Proposition 2.3:** we have

$$\begin{aligned} \bar{R}^1 &= (1 - \theta) R_f + \theta \bar{R}_1 \\ \bar{R}^2 &= (1 - \theta) R_f + \theta \bar{R}_2 \text{ and} \\ \bar{R}^1 + \bar{R}^2 &= 2(1 - \theta) R_f + \theta (\bar{R}_1 + \bar{R}_2) \end{aligned}$$

Since  $R_f$  is nonrandom and  $\theta > 0$ , by using translation invariance and positive homogeneity of VaR, we have that

$$\begin{aligned} VaR_p \left( \bar{R}^1 + \bar{R}^2 \right) &> VaR_p \left( \bar{R}^1 \right) + VaR_p \left( \bar{R}^2 \right) \text{ is equivalent to} \\ VaR_p \left( \bar{R}_1 + \bar{R}_2 \right) &> VaR_p \left( \bar{R}_1 \right) + VaR_p \left( \bar{R}_2 \right), \text{ we will then show this second formulation.} \end{aligned}$$

Let  $\bar{\mu} \equiv \mu_1 - \mu_0 < 0$  and let us define the following variables

$$\begin{aligned} X_1 &\equiv R_1^1 - \mu_0 \sim N(\bar{\mu}, \bar{\sigma}^2) \\ X_2 &\equiv R_2^1 - \mu_0 \sim N(\bar{\mu}, \bar{\sigma}^2) \\ Z_1 &\equiv R_1^0 - \mu_0 \sim N(0, \bar{\sigma}^2) \\ Z_2 &\equiv R_2^0 - \mu_0 \sim N(0, \bar{\sigma}^2) \end{aligned}$$

we have

$$\begin{aligned} \bar{R}_1 &= s_1 X_1 + (1 - s_1) Z_1 + \mu_0 \\ \bar{R}_2 &= s_2 X_2 + (1 - s_2) Z_2 + \mu_0 \end{aligned}$$

and

$$\begin{aligned} p &= \Pr \left( \bar{R}_1 + \bar{R}_2 \leq \mu_1 + \mu_0 \right) \\ &= \Pr \left( s_1 X_1 + (1 - s_1) Z_1 + s_2 X_2 + (1 - s_2) Z_2 \leq \mu_1 - \mu_0 \right) \\ &= \Pr \left( s_1 s_2 (X_1 + X_2) + s_1 (1 - s_2) (X_1 + Z_2) + \right. \\ &\quad \left. + s_2 (1 - s_1) (X_2 + Z_1) + (1 - s_1) (1 - s_2) (Z_1 + Z_2) \leq \bar{\mu} \right) \\ &= \lambda^2 F_{X_1 + X_2}(\bar{\mu}) + \lambda (1 - \lambda) F_{X_1 + Z_2}(\bar{\mu}) + \\ &\quad + \lambda (1 - \lambda) F_{X_2 + Z_1}(\bar{\mu}) + (1 - \lambda)^2 F_{Z_1 + Z_2}(\bar{\mu}) \\ &= \lambda^2 F_{X_1 + X_2}(\bar{\mu}) + \lambda (1 - \lambda) + (1 - \lambda)^2 F_{Z_1 + Z_2}(\bar{\mu}) \end{aligned}$$

i.e.  $p = \lambda^2 F_{X_1+X_2}(\bar{\mu}) + \lambda(1-\lambda) + (1-\lambda)^2 F_{Z_1+Z_2}(\bar{\mu})$

**Remark:**

$p = 1/2$  for  $\bar{\mu} = 0$ ,

and  $p \rightarrow \lambda$ . To see this second result we can notice that  $F_{Z_1+Z_2}(\bar{\mu}) \xrightarrow{\bar{\mu} \rightarrow -\infty} 0$ , and

$$\begin{aligned} F_{X_1+X_2}(\bar{\mu}) &= \Pr(X_1 + X_2 \leq \bar{\mu}) \\ &= \Pr((X_1 - \bar{\mu}) + (X_2 - \bar{\mu}) \leq -\bar{\mu}) \\ &= \Pr(N(0, 2\sigma_1^2) \leq -\bar{\mu}) \\ &\xrightarrow{\bar{\mu} \rightarrow -\infty} 1 \end{aligned}$$

therefore

for  $\lambda < 1/2$ ,  $p$  can take any value into  $(\lambda, 1/2)$  as  $\bar{\mu}$  varies into  $(-\infty, 0)$

and for  $\lambda > 1/2$ ,  $p$  can take any value into  $(1/2, \lambda)$  as  $\bar{\mu}$  varies into  $(-\infty, 0)$

so while  $\lambda$  and  $\bar{\mu}$  vary  $p$  can take any value into  $(0, 1)$

Let us pose

$$\begin{aligned} U_1 &\equiv s_1 X_1 + (1 - s_1) Z_1 \\ U_2 &\equiv s_2 X_2 + (1 - s_2) Z_2 \end{aligned}$$

We can easily see that  $U_j = \tilde{R}_j - \mu_0$ ,  $j = 1, 2$ . Therefore the proposition is equivalent to  $VaR_p(U_1 + U_2) > VaR_p(U_1) + VaR_p(U_2)$ . By definition of  $p$ , we have  $VaR_p(U_1 + U_2) = -\bar{\mu}$ , and since  $U_1 \stackrel{d}{=} U_2$  then  $VaR_p(U_1) = VaR_p(U_2)$  and the problem becomes  $VaR_p(U_1 + U_2) > 2VaR_p(U_1)$  i.e.  $VaR_p(U_1) < -\bar{\mu}/2$  or equivalently  $\Pr(U_1 \leq \bar{\mu}/2) < p$ .

we have

$$\begin{aligned} \Pr(U_1 \leq \bar{\mu}/2) &= \Pr(s_1 X_1 + (1 - s_1) Z_1 \leq \bar{\mu}/2) \\ &= \lambda F_{X_1}(\bar{\mu}/2) + (1 - \lambda) F_{Z_1}(\bar{\mu}/2) \end{aligned}$$

let  $\varepsilon$  be a  $N(0, 1)$  random variable and  $\Phi$  its c.d.f., we have:

$$\begin{aligned} F_{X_1+X_2}(\bar{\mu}) &= \Pr(2\bar{\mu} + \sqrt{2}\bar{\sigma}\varepsilon < \bar{\mu}) \\ &= \Phi(-\bar{\mu}/\sqrt{2}\bar{\sigma}); \end{aligned}$$

$$\begin{aligned} F_{X_1}(\bar{\mu}/2) &= \Pr(\bar{\mu} + \bar{\sigma}\varepsilon < \bar{\mu}/2) \\ &= \Phi(-\bar{\mu}/2\bar{\sigma}); \end{aligned}$$

$$\begin{aligned} F_{Z_1+Z_2}(\bar{\mu}) &= \Pr(\sqrt{2}\bar{\sigma}\varepsilon < \bar{\mu}) \\ &= \Phi(\bar{\mu}/\sqrt{2}\bar{\sigma}); \end{aligned}$$

and

$$\begin{aligned}
F_{Z_1}(\bar{\mu}/2) &= \Pr(\bar{\sigma}\varepsilon < \bar{\mu}/2) \\
&= \Phi(\bar{\mu}/2\bar{\sigma}).
\end{aligned}$$

so

$$\begin{aligned}
p - \Pr(U_1 \leq \bar{\mu}/2) &= \lambda^2 \Phi(-\bar{\mu}/\sqrt{2\bar{\sigma}}) + \lambda(1-\lambda) + (1-\lambda)^2 \Phi(\bar{\mu}/\sqrt{2\bar{\sigma}}) \\
&\quad - \lambda \Phi(-\bar{\mu}/2\bar{\sigma}) - (1-\lambda) \Phi(\bar{\mu}/2\bar{\sigma}) \\
&= \left[ \underbrace{\Phi(-\bar{\mu}/\sqrt{2\bar{\sigma}}) + \Phi(\bar{\mu}/\sqrt{2\bar{\sigma}})}_{=1} - 1 \right] \lambda^2 + \\
&\quad + \left[ -2\Phi(\bar{\mu}/\sqrt{2\bar{\sigma}}) + \underbrace{1 - \Phi(-\bar{\mu}/2\bar{\sigma})}_{=\Phi(\bar{\mu}/2\bar{\sigma})} + \Phi(\bar{\mu}/2\bar{\sigma}) \right] \lambda \\
&\quad + [\Phi(\bar{\mu}/\sqrt{2\bar{\sigma}}) - \Phi(\bar{\mu}/2\bar{\sigma})]
\end{aligned}$$

i.e.

$$\begin{aligned}
p - \Pr(U_1 \leq \bar{\mu}/2) &= [\Phi(\bar{\mu}/\sqrt{2\bar{\sigma}}) - \Phi(\bar{\mu}/2\bar{\sigma})] [1 - 2\lambda] \\
&> 0 \quad \text{since } \lambda < 1/2 \text{ and } \Phi(\bar{\mu}/\sqrt{2\bar{\sigma}}) > \Phi(\bar{\mu}/2\bar{\sigma})
\end{aligned}$$

which means  $\Pr(U_1 \leq \bar{\mu}/2) < p$ , and therefore the proposition follows.

**Q.E.D**

## Appendix 2

**Proof of Proposition 3.2:** Let us denote  $Z = X + Y$ .

(i) If  $\alpha > 1$  :

$$\sigma_Z = [\sigma_X^\alpha + \sigma_Y^\alpha]^{1/\alpha} < \sigma_X + \sigma_Y$$

Consider a real  $r$  such that:

$$\sigma_Z < r < \sigma_X + \sigma_Y$$

Since  $\sigma_Z = \lim_{p \rightarrow 0} VaR_p(Z)p^{1/\alpha}$ ,

there exists  $p_1$  such that, for any  $p \in ]0, p_1[$ :

$$VaR_p(Z)p^{1/\alpha} < r.$$

For similar reasons, there exists  $p_2$  such that, for any  $p \in ]0, p_2[$

$$r < VaR_p(X)p^{1/\alpha} + VaR_p(Y)p^{1/\alpha}.$$

Therefore, with  $p_0 = \min(p_1, p_2)$  we have for any  $p \in ]0, p_0[$  :

$$VaR_p(Z) < VaR_p(X) + VaR_p(Y)$$

(ii) If  $\alpha < 1$

$$\sigma_Z = [\sigma_X^\alpha + \sigma_Y^\alpha]^{1/\alpha} > \sigma_X + \sigma_Y.$$

A similar argument shows that there exists  $p_0$  such that, for any  $p \in ]0, p_0[$ :

$$VaR_p(Z) > VaR_p(X) + VaR_p(Y).$$

(iii) If  $\alpha = 1$

$$\sigma_Z = \sigma_X + \sigma_Y$$

that is:

$$\lim_{p \rightarrow 0} p VaR_p(Z) = \lim_{p \rightarrow 0} p VaR_p(X) + \lim_{p \rightarrow 0} p VaR_p(Y).$$

**Proof of Proposition 3.3:**

As a function of  $\alpha \in ]0, 1[$ , the function  $\frac{[\sigma_X + \sigma_Y]^\alpha}{\sigma_X^\alpha + \sigma_Y^\alpha}$  is continuous and increasing, from the value  $1/2$  for  $\alpha = 0$  to the value  $1$  for  $\alpha = 1$ . Then, for  $k$  given in  $]0, 2[$ , there exists  $\alpha_0$  such that, for any  $\alpha$  in  $]0, \alpha_0[$  :

$$\frac{[\sigma_X + \sigma_Y]^\alpha}{\sigma_X^\alpha + \sigma_Y^\alpha} < \frac{1}{k},$$

that is:

$$\sigma_X + \sigma_Y < k^{-1/\alpha} [\sigma_X^\alpha + \sigma_Y^\alpha]^{1/\alpha} = k^{-1/\alpha} \sigma_Z.$$

Since  $\sigma_X = \lim_{p \rightarrow 0} p^{1/\alpha} \text{VaR}_p(X)$ ,  $\sigma_Y = \lim_{p \rightarrow 0} p^{1/\alpha} \text{VaR}_p(Y)$   
and  $\sigma_Z = \lim_{p \rightarrow 0} (kp)^{1/\alpha} \text{VaR}_{kp}(z)$ , we have the existence of  $p_0$  such that for any  $p$  in  $]0, p_0[$ :

$$p^{1/\alpha} \text{VaR}_p(X) + p^{1/\alpha} \text{VaR}_p(Y) < k^{-1/\alpha} (kp)^{1/\alpha} \text{VaR}_{kp}(Z)$$

that is  $\text{VaR}_p(X) + \text{VaR}_p(Y) < \text{VaR}_{kp}(Z)$ .

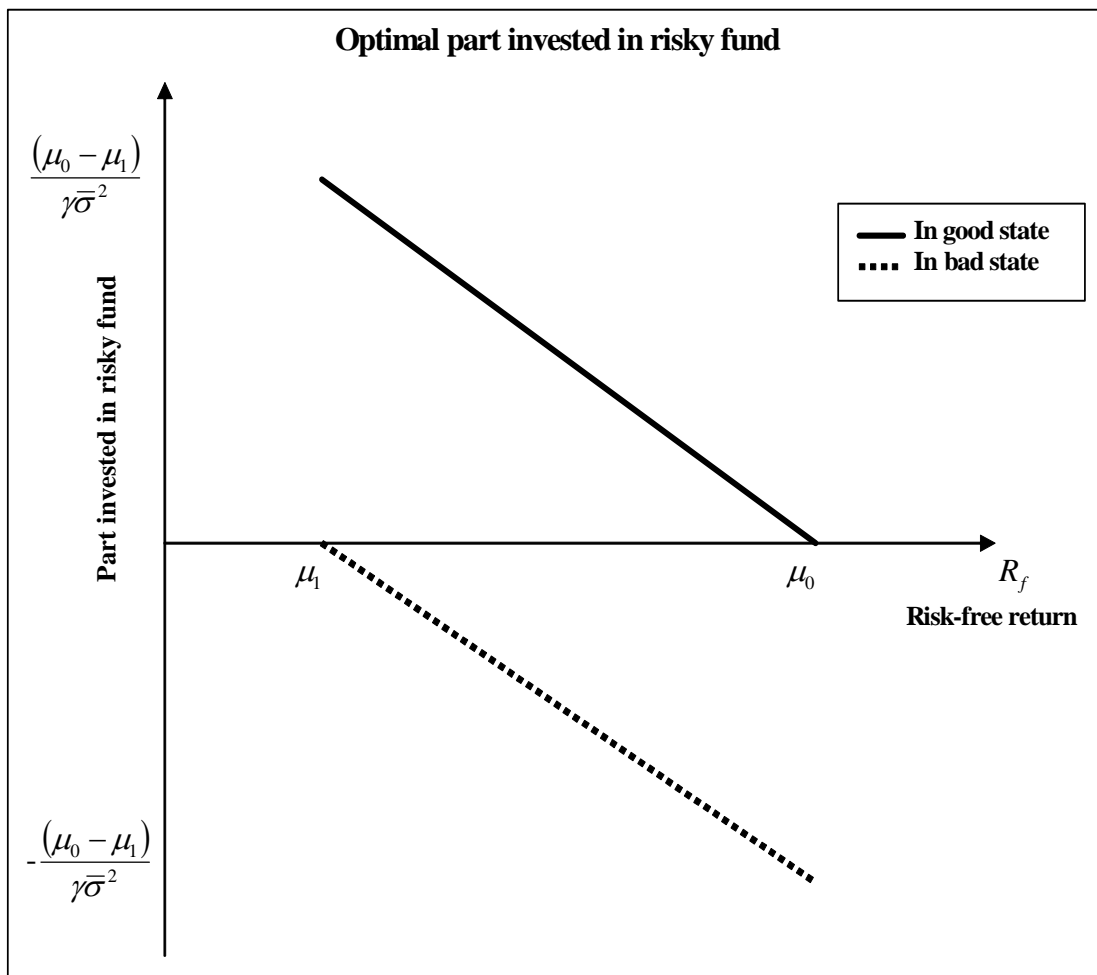


Figure 1: Optimal Share Invested in Risky Fund as Function of the Risk-free return