

# The economic value of controlling for large losses in portfolio selection

Alexandra Dias\*

School of Management  
University of Leicester

## Abstract

During financial crises equity portfolios have suffered large losses. Methodologies for portfolio selection taking into account the possibility of large losses have existed for decades but their economic value is not well established. This article investigates the economic value in reducing the probability of large losses in portfolio selection. We combine mean-variance analysis with semi-parametric estimation of potential portfolio large losses. We find that strategies that reduce the probability of large losses outperform efficient minimum variance portfolios, especially when semi-parametric estimation is used. Our results are robust to transaction costs.

*JEL classification:* G11; G14

*Keywords:* portfolio selection, portfolio tail probability, extreme losses, risk management.

---

\* *Correspondence address:* Alexandra Dias, School of Management, University of Leicester, University Road, LE1 7RH Leicester, UK. Tel: +44(0)116 252 5019. Fax: +44(0)116 252 3949. E-mail: Alexandra.Dias@le.ac.uk

# 1 Introduction

Portfolio selection with controlled downside risk is a problem of acute practical interest. Large losses in financial portfolios are more frequent and larger than expected under the classical Markowitz (1952) framework. This is due to the non-normality of asset returns and has been recognized since Mandelbrot (1963). Portfolios composed using the classical “normal” mean-variance portfolio optimization are subject to potential large losses originated by the fat-tailedness of asset returns. Our contribution is to evaluate the economic value of taking into account the possibility of large losses in portfolio selection.

In the existing literature one stream of research investigates the effect of the inclusion of higher moments in portfolio allocation. A problem with this approach is that it requires the estimation of possibly many higher-order cross-moments; see Martellini and Ziemann (2010) for a recent reference. Another approach focuses on constraining the portfolio downside risk via VaR or another risk measure; see for instance Alexander and Baptista (2002). This literature focuses mostly on probabilistic properties and estimation methods rather than on the economic significance of using large losses as a criterion in portfolio selection.

The concept of limiting downside risk goes back to Roy (1952) who introduced into portfolio selection the principle of safety-first. Roy used the first two moments of the assets return distribution to limit the probability

of a disastrous loss. The study of portfolio selection for safety-first investors used the assumption of normally distributed asset returns for some time. An essentially distribution free approach was taken by Arzac and Bawa (1977) who used VaR as a downside risk measure. A later paper on portfolio allocation with safety-first without the normality assumption is Gouriéroux et al. (2000) who use a non-parametric estimate of the full distribution of the asset returns. Jansen et al. (2000) concentrates on estimating the portfolio fat-tail distribution using the safety-first principle combined with statistical extreme value theory to limit downside risk.

These criteria for portfolio selection, based on the tail properties of the asset return distribution, often choose a corner solution, meaning that they put most weight on the asset with the thinnest tail. This has been observed by, for instance, Jansen et al. (2000), Hartmann et al. (2004) and Poon et al. (2003). The theoretical explanation for this is in a result from Geluk and de Hann (1987). They show that the tail-heaviness of the convolution of heavy-tailed variables is determined mainly by the variable with the heaviest tail. Corner solutions are a serious drawback in the use of heavy tail modeling with the safety-first principle. Hyung and de Vries (2007) attempt to overcome the corner solution problem by using a second order expansion at infinity of the asymptotically Pareto tail probability. We take a simpler route. We choose the portfolio with the thinnest tail among the set of possible efficient minimum variance portfolios. With this criterion, on the one

hand we do not lose the diversification effect of mean-variance portfolios in “normal” times; on the other hand we keep the probability of large losses under control for the abnormal “heavy-tailed” market times.

In order to evaluate the economic value of controlling for the probability of a portfolio large loss we consider a mean-variance investor who is not willing to ignore the risk of a large loss. We suppose that the possible investment alternatives are the set of possible efficient minimum variance portfolios formed in a universe of assets. We opt for using minimum variance portfolios because there is evidence (see for instance Brandt (2009)) that due to the difficulty of estimating the mean portfolio return, minimum variance portfolios can actually perform better than tangency mean-variance portfolios.

Our investor chooses from among these minimum variance portfolios the one with the smallest probability of incurring a large loss. With this approach we avoid the corner solution problem arising from using the safety-first principle with extreme value modeling. At the same time our approach inherits the diversification effect from the mean-variance construction.

For the estimation of the probability of a portfolio large loss we use a semi-parametric estimator introduced by de Haan and de Ronde (1998). The advantage of this estimator is that it does not assume any particular parametric family for the dependence structure between large losses on the assets that form the portfolio. This is a clear statistical advantage when

estimating the probability of events for which the number of observations is by definition small.

The data for our analysis consist of monthly returns on the constituents of the Dow Jones Industrial Average. We consider as a benchmark an investor who chooses the global minimum variance portfolio among the possible investment alternatives. For comparison we also consider a strategy where the minimum VaR portfolio is chosen. To evaluate the appropriateness of the semi-parametric estimator we use as an alternative a standard extreme value parametric tail estimator.

Our results indicate that the strategies that control for the probability of a large loss outperform the pure global minimum variance portfolio strategy, where the performance is measured by the Sharpe and Sortino ratios. We also find that an investor would pay a positive fee to change from the global minimum variance strategy to a strategy which controls for the tail probability. Further, the use of the semi-parametric estimator outperforms both the use of a standard extreme value parametric tail estimator and the minimum VaR strategies. These results are robust to transaction costs.

The organization of this paper is as follows. In Section 2 we describe the methodology that we use in our study. In Section 3 we present our empirical study and the results obtained. Section 4 concludes the paper.

## 2 Methodology

Our methodology for evaluating the economic value of controlling for large losses in portfolio allocation is to calculate the effect of taking into account the probability of a large loss on the performance of portfolios. Our universe of assets is the set of stocks that form the Dow Jones Industrial Average (DJIA). Starting in October 1991, using data since February 1973, every month we form portfolios with all the possible combinations of three different assets. We use portfolios of three assets for computational ease and because the portfolio diversification effect is largely attained with few assets as reported in Solnik (1995).

We present the results of using four strategies. The first is a benchmark strategy. For each three asset combination we calculate the global minimum variance efficient portfolio. From among all the global minimum variance portfolios we choose the one with the smallest variance. For the second strategy, we estimate the probability of a large loss over one month for each of the minimum variance portfolios using a semi-parametric estimator. The strategy consists of choosing the minimum variance portfolio with the smallest probability of a large loss. We call this strategy the minimum large loss strategy. In the third strategy, we compute the tail-heaviness of each minimum variance portfolio using a parametric tail-index estimator, then we choose the portfolio with the thinnest losses tail (lower tail-index). Finally, in the fourth strategy we choose the portfolio with the minimum VaR. In

this strategy we assume multivariate normality. The returns distribution is usually misspecified under this assumption but the VaR estimator has a closed form expression. If this strategy were to perform better we could conclude that specification error is less of a problem than estimation error.

For all strategies we assume that short-selling is not allowed and that the investor can not hold cash in the portfolio. These assumptions seem reasonable in many practical cases. We keep the weights of the portfolios constant for one month and after one month we recalculate the four portfolios according to each strategy using the previous 224 months of data. We continue with the four strategies until June 2010. We obtain 224 out-of-sample monthly returns for each strategy. We evaluate the performance of each strategy implied by these returns and compare the results.

## 2.1 The global minimum variance portfolio (Strategy 1)

Our benchmark strategy is the classical Merton (1972) global minimum variance portfolio on the efficient mean-variance boundary. For a number of  $n$  securities, the minimum variance portfolio weights are the elements of the vector  $w_\sigma$ ,

$$w_\sigma = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}, \quad (1)$$

where  $\mathbf{1}$  is the  $n \times 1$  unit vector and  $\Sigma$  is the  $n \times n$  variance-covariance matrix of the portfolio asset returns. The reasons why we choose as a benchmark the minimum variance portfolio are first that we want to avoid our results

of being driven by the expected value of the asset returns, and second that our goal is exactly to measure the economic value of minimizing the risk of the portfolio.

## **2.2 Estimation of the probability of a large loss (Strategy 2)**

In our second strategy we follow a semi-parametric approach to estimate the probability of a portfolio large loss. Concretely, we use a semi-parametric estimator for estimating the probability of large events, introduced by de Haan and de Ronde (1998) in the context of Extreme Value Theory. We choose to use this semi-parametric estimator for two main reasons. Firstly, it is a multivariate estimator in the sense that it uses information from all the univariate asset returns. By comparison, a univariate estimator uses only the characteristics of the portfolio return (univariate) distribution. As a consequence, we expect better portfolio performance using the semi-parametric estimator. Secondly, unlike multivariate parametric estimators, it does not assume any particular family for the dependence structure of the returns. Due to the small number of observations of large losses this is an obvious advantage over fully-parametric estimators. With few observations the choice of a parametric family for the dependence structure of large losses would have to rely mainly on economic arguments and very little on the robustness of statistical inference. Due to the semi-parametric nature of the estimator increasing the number of assets in the portfolio does not make the

estimation more difficult.

Here we summarize briefly how to estimate the probability of portfolio large losses using this estimator. For a detailed statistically orientated exposition and implementation see Dias (2009).

Denote by  $R_i$  the random variable representing the monthly returns on asset  $i$ . The portfolio return  $R_p$  is the weighted average of the individual asset returns,  $R_p = w_1R_1 + w_2R_2 + \dots + w_dR_d$ , where  $d$  is the number of assets in the portfolio and  $w_i$  is the weight of asset  $i$ . We want to estimate the probability of a portfolio return loss being larger than a given value  $L$ <sup>1</sup>. There is a set  $C$  of possible (multivariate) asset returns whose realization implies a portfolio return loss larger than  $L$ .  $C$  is a set of  $d$ -dimensional asset returns. Estimating the probability of a portfolio loss larger than  $L$  is equivalent to estimating the probability of having a ( $d$ -dimensional) realization of asset returns in  $C$ .

Given that portfolio large losses are not often observed there are few observations from historical data in the set  $C$ . This is the main challenge in estimating tail probabilities. We approach it using a theoretical result from de Haan and Resnick (1977). This result shows that, under conditions usually seen in practice, there is a relation between the probability of a portfolio large loss and the probability of a portfolio small loss. As small

---

<sup>1</sup>Typically  $L$  corresponds to a quantile of the portfolio losses distribution of 90% or higher.

losses are not uncommon in a portfolio, we can estimate its probability more reliably. de Haan and Resnick (1977) show that the two probabilities are related by a multiplicative constant that we can also estimate. If we know the relation between the probability of small and large losses then we can obtain an estimate for the probability of a portfolio large loss.

In this approach, in order to estimate the probability of a portfolio large loss, we standardize the individual asset returns. Given that we are interested in tail observations, the standardization is done with the parameters of the distribution of the asset maximum return. Specifically, the standardized returns, for each asset  $i$ , are given by  $\tilde{R}_i = (1 + \gamma(R_i - b)/a)^{1/\gamma}$ , where  $\gamma$  is the shape parameter (tail index), and  $b$  and  $a$  are respectively location and scale parameters of asset  $i$  returns. The parameters  $\gamma$ ,  $b$  and  $a$  must be estimated. We follow here de Haan and de Ronde (1998) for estimating them. The estimators are reported in the Appendix. The set of asset returns belonging to  $C$ , after being standardized, is denoted by  $cA$ , where  $c$  is a positive scalar and  $A$  is a  $d$ -dimensional set. The standardized asset returns in  $cA$  correspond to portfolio large losses, and if  $c$  is large, the returns in  $A$  correspond to “not so large” portfolio losses.

Historical observations in  $A$  are not scarce and we can estimate its probability,  $\hat{p}(A)$ . For this, we use the estimator  $\hat{p}$  proposed by de Haan and Resnick (1993),

$$\hat{p}(A) = \frac{1}{k} \sum_{t=1}^n \mathbb{I}(\tilde{\mathbf{R}}_t \in A), \quad (2)$$

where  $\tilde{\mathbf{R}}_t$  are all the observed ( $d$ -dimensional) standardized asset returns,  $n$  is the number of historical observations,  $\mathbb{I}(\cdot)$  represents the indicator function and  $k$  is the number of upper-order statistics used to estimate the shape parameter  $\gamma$ . We use  $k = 0.2 \times n$ . For a discussion on the choice of  $k$  in the context of portfolio large returns see Dias (2009) and references therein.

Given the standardization performed on the asset returns, the expected value of  $\tilde{R}_i$  is one. Hence, if the boundary<sup>2</sup> of the set  $A$  contains the unit vector then the set  $A$  has a significant proportion of the historical observations and we can obtain a reliable estimate of its probability. This motivates the definition of the constant  $c$  given in Dekkers et al. (1989).  $c$  is uniquely defined as the scalar such that the unit vector is on the lower boundary of  $A$ ; see Appendix.

From de Haan and Resnick (1977) we have the following estimator for the probability of a portfolio large loss:

$$\hat{p}(C) = \frac{k}{nc} \hat{p}(A). \quad (3)$$

### **2.3 Parametric estimation of the portfolio tail heaviness (strategy 3)**

The semi-parametric estimator described in the previous section takes into account the characteristics of all the univariate asset returns distributions.

---

<sup>2</sup>For a precise mathematical exposition see Dekkers et al. (1989).

We want to investigate whether there is economic value in using information from the individual assets to estimate the probability of a portfolio large loss. To do this we compare the performance of the portfolios found using the semi-parametric estimator (strategy 2) with the performance using a univariate maximum likelihood estimator of the portfolio tail-heaviness (strategy 3). This estimator uses information from the univariate distribution of the portfolio returns and not directly the individual assets distributional characteristics. We call this the minimum tail-index strategy.

We use the so-called peaks-over-threshold (POT) method which is perhaps the most popular statistical method for estimating the tail heaviness of a random variable. For an application in finance see for instance Longin and Solnik (2001). The POT method is based on a theoretical result from Balkema and de Haan (1974) and Pickands (1975).

Let  $R$  denote the portfolio returns variable with distribution function  $F_R$ . For simplicity we concentrate here on the tail of the positive values of the variable. Apart from a change of sign everything works analogously for the negative values.

The upper end point of the density function associated with  $F_R$  is denoted by  $x_F$ . In the case of a normally distributed random variable, for instance,  $x_F = +\infty$ . The extremes of the variable are defined in terms of the exceedances over a threshold  $u < x_F$ . The excess distribution function of the random variable  $R$  over the threshold  $u$  is  $F_u(x) = P(R - u < x | R > u)$ ,

$x \geq 0$ . For statistical purposes we use the following result from extreme value theory that gives an approximation for the distribution function of the exceedances. Balkema and de Haan (1974) and Pickands (1975) show that there exists a unique non degenerate limit distribution,  $G_{\xi,\beta}$ , such that  $\lim_{u \rightarrow x_F} \sup_{0 < x < x_F - u} |F_u(x) - G_{\xi,\beta}(x)| = 0$ .<sup>3</sup>

The limit distribution  $G_{\xi,\beta}$  is the generalized Pareto distribution and has the form

$$G_{\xi,\beta}(x) = 1 - (1 + \xi x/\beta)^{-1/\beta} \quad (4)$$

where  $x \geq 0$  if  $\xi \geq 0$  and  $0 \leq x \leq -\beta/\xi$  if  $\xi < 0$ .  $\xi$  is the tail heaviness parameter and  $\beta$  is a parameter depending on  $u$ . Distributions with a power tail (heavy tailed distributions as the Student- $t$  for instance) have  $\xi > 0$ , distributions with exponential tail (thin tailed distributions as the normal distribution) have  $\xi = 0$  and distributions with a finite tail (as the uniform distribution for instance) have  $\xi < 0$ . We estimate the parameters of the distribution by maximum likelihood. More details about the statistical procedure can be found for instance in Embrechts et al. (1997).

## 2.4 The minimum VaR portfolio (strategy 4)

In the spirit of mean-variance analysis, Alexander and Baptista (2002) define the mean-VaR efficient frontier where VaR replaces the standard deviation as measure of risk. In this setup the portfolio weights are obtained by

---

<sup>3</sup>This result is also valid for non-i.i.d. processes; see Leadbetter et al. (1983).

minimizing the VaR for a target portfolio expected return. Alexander and Baptista (2002) assume that the returns on the assets have a multivariate normal distribution and derive a closed form solution for the minimum VaR portfolio analogous to the minimum variance portfolio. We use the global minimum VaR portfolio as a benchmark for strategies where tail risk is considered.

Let  $n \geq 2$  be the number of securities. The  $n \times 1$  vector of expected returns of each asset is denoted by  $\mu$  and  $\Sigma$  is the  $n \times n$  variance-covariance matrix of the asset returns. In order to give the minimum VaR portfolio we need to define the following constants:  $A = \mathbf{1}^T \Sigma^{-1} \mu$ ,  $B = \mu^T \Sigma^{-1} \mu$ ,  $C = \mathbf{1}^T \Sigma^{-1} \mathbf{1}$  and  $D = BC - A^2$ , where  $\mathbf{1}$  is the  $n \times 1$  unit vector. Let  $\Phi$  be the distribution function of a univariate standard normal variable. The minimum VaR portfolio at the  $100t\%$  confidence level, with  $t \in (1/2, 1)$  and  $t > \Phi(\sqrt{D/C})$ , is given by

$$m_t = g + h \left( \frac{A}{C} + \sqrt{\frac{D}{C} \left( \frac{(t^*)^2}{(t^*)^2 C - D} - \frac{1}{C} \right)} \right), \quad (5)$$

where  $g$  and  $h$  are the  $n \times 1$  vectors  $g = (1/D)[B(\Sigma^{-1} \mathbf{1}) - A(\Sigma^{-1} \mu)]$  and  $h = (1/D)[C(\Sigma^{-1} \mu) - A(\Sigma^{-1} \mathbf{1})]$ . The point  $t^*$  is such that  $\Phi(-t^*) = 1 - t$ .

## 2.5 Measurement of the economic value

To measure the value of incorporating the probability of large losses in portfolio selection we compare the performance of the strategies that use the tail probability to that of the minimum variance strategy.

Chamberlain (1983) shows that if asset returns have an elliptical distribution then the mean-variance approximation of the expected utility is exact for all utility functions. Even without assuming ellipticity we can use quadratic utility as a second order approximation to the investor's true utility function. In this case the realized utility at time  $t$  is

$$U(W_t) = W_{t-1}R_{p,t} - \frac{\alpha W_{t-1}^2}{2}R_{p,t}^2,$$

where  $W_t$  is the investor's wealth at time  $t$ ,  $\alpha$  is his absolute risk aversion and  $R_{p,t}$  is the portfolio return for period  $t$ . For comparison purposes between strategies we assume that  $\alpha W_t$  is constant which implies a constant relative risk aversion  $\gamma = \alpha W_t / (1 - \alpha W_t)$ . We use the values of  $\gamma = 1$  and  $\gamma = 10$  in our analysis.

We estimate the expected utility using the average realized utility. The value of a strategy is estimated by equating the average realized utilities for the two alternative strategies. We suppose that holding a portfolio other than the minimum variance portfolio subject to expenses  $\delta$  yields the same average utility as holding a minimum variance portfolio. An investor should be indifferent between these two strategies. Hence,  $\delta$  is the maximum fee that an investor would be willing to pay to switch from a minimum variance portfolio to the alternative strategy. If  $\delta$  is expressed as a fraction of the initial wealth invested,  $\delta$  solves the following equation:

$$\sum_{t=1}^T (R_{s,t} - \delta) - \frac{\gamma}{2(1 + \gamma)} (R_{s,t} - \delta)^2 = \sum_{t=1}^T R_{p,t} - \frac{\gamma}{2(1 + \gamma)} R_{p,t}^2, \quad (6)$$

where  $R_{p,t}$  and  $R_{s,t}$  denote the returns on the minimum variance portfolio and on the alternative strategy portfolio, respectively.

### 3 Data and empirical results

The data used in our study is collected from Datastream and covers the period from February 1973 until June 2010. From the 30 components of DJIA we use the data from the 24 stocks for which prices are available for the whole period. We have approximately 37 years of monthly returns on 24 stocks which translates into 448 monthly return observations per stock.

In each month, from November 1991 until June 2010, we compute all the minimum variance portfolios for the 2,024 possible combinations of 3 different stocks from the 24 stocks available. From these 2,024 minimum variance portfolios we choose three portfolios, one according to each of the first three strategies: the portfolio with the smallest variance, the portfolio with the lowest probability of a portfolio loss larger than 0.10 (using the semi-parametric estimator), the portfolio with the smaller loss tail-index (using the classic parametric estimator). For the fourth strategy we calculate the 2,024 minimum 90% VaR portfolios and choose the one with smallest VaR.

After one month we calculate the monthly portfolio returns and we adjust the composition of each of the four portfolios, using the previous 224 months of data. We repeat the procedure until June 2010, obtaining 225 portfolios

and 224 portfolio returns for each strategy. The analysis of these (out-of-sample) portfolio returns follows.

### 3.1 Portfolio returns analysis

We split the period from November 1991 until June 2010 into four subperiods with equal length. We perform the analysis and report the results for each of these subperiods as well as for the entire period.

The annualized mean realized return is reported in Table 1. We observe that the average return varies considerably across the different subperiods. The highest average for the entire sample is obtained by the minimum large loss strategy as well as in three of the four subperiods.

We do not report the standard deviation but rather the semi-standard deviations<sup>4</sup> because we are interested in breaking down the negative and the positive portfolio risk and performance. For all four strategies the results are good in the sense that the negative semi-standard deviation is smaller than the positive semi-standard deviation for the entire sample and for most of the subperiods. The negative semi-standard deviation for the minimum large loss and minimum tail-index strategies are higher than for the minimum variance strategy. This is consistent with results from Dittmar (2002), Harvey and Siddique (2000) and Mitton and Vorkink (2007) where investors are shown to be willing to accept a lower expected return and higher volatility

---

<sup>4</sup>Recall that the positive semi-variance is estimated by  $\sigma^{+2} = 1/T \sum_{i=1}^T ([R_i - E(R)]^+)^2$ , where  $T$  is the number of observations.

compared to the mean-variance benchmark in exchange for higher skewness.

In terms of formal measures of performance we report the Sharpe ratio and the Sortino ratio. The riskless rate used to calculate the Sharpe ratio is the one-month US-Treasury bond rate. The minimum large loss strategy attains the best Sharpe ratios in all cases except for the period 03/2001–10/2005. These results are confirmed by the Sortino ratios (calculated with a minimum acceptable return equal to the riskless rate). Given that the Sortino ratio penalizes specifically the returns below the minimum acceptable return our results indicate that reducing the probability of large portfolio losses have an effect on controlling for downside risk.

[Insert Table 1]

Although the minimum large loss strategy does not have the lowest negative semi-standard deviation, neither does it produce the lowest returns. We see in Table 2 that the minimum tail-index strategy has the lowest returns in the entire sample and in two of the subperiods. Also, the minimum tail-index strategy attains a minimum return lower than the minimum large loss strategy in four out of the five periods (including the entire sample). This is an indicator of the greater power of the semi-parametric estimation compared with the classic parametric estimation of the loss tail in terms of portfolio selection. We observe that, with the exception of the last subperiod, the maximum return is substantially larger using the minimum large loss strategy. On the one hand, this is not surprising when compared with

the minimum variance strategy where by construction the positive part of the volatility is also minimized. On the other hand, this result shows that minimizing the probability of a large loss instead of the variance has also an effect on the positive portfolio returns.

We report in Table 2 the realized frequency of negative returns. Differences between the different strategies are not so great. The strategy with the lowest frequency is the minimum large loss strategy mainly due to a substantially lower realized frequency in the subperiod 11/1991–06/1996.

We also list in Table 2 the skewness and the kurtosis of the realized portfolio returns. It is noticeable that the minimum large loss strategy has the most positive skewness, which agrees with the fact that it has the largest maximum return. Also in line with this fact is the high kurtosis that the minimum large loss strategy attains in the second and fourth subperiods. The lowest kurtosis is attained by the minimum variance and minimum VaR strategies in the subperiod 11/1991–06/1996. Note that this subperiod has the smaller difference between maximum and minimum realized portfolio returns.

[Insert Table 2]

### **3.2 Economic value**

We evaluate the economic value of each strategy by calculating, using equation (6), the fee that an investor with constant risk aversion would be willing

to pay to change from the minimum variance strategy an alternative strategy.

The results reported in Table 1, columns  $\delta_1$  and  $\delta_{10}$ , suggest that incorporating tail information in portfolio selection implies substantial differences in the out-of-sample portfolio returns. There are also differences between the strategies. We first look at the results for an investor with  $\gamma = 1$ . For the entire sample the minimum large loss and minimum VaR strategies present gains relative to the minimum variance strategy of 2.62% and 0.75% respectively. These values are economically significant in accordance with the literature. In the first subperiod the three strategies produce a positive fee where the minimum large loss strategy has the highest gain and the VaR strategy the lowest. In the second period the VaR strategy presents a loss of 2.06%. In the third subperiod only the VaR strategy has a gain. In the last subperiod only the minimum large loss strategy has a gain, 4.98%. The relative results between strategies for an investor with a constant risk aversion  $\gamma = 10$  are similar to an investor with  $\gamma = 1$ .

Overall, the minimum large loss strategy performs best, followed by the minimum VaR strategy. The poorer performance of the minimum tail index strategy is probably due to loss of information. The minimum tail-index strategy models the tail of the univariate portfolio return distribution while the other two strategies use information from all the individual stocks in the portfolio.

### 3.3 Portfolio weights analysis

In order to analyze the portfolio weights we define two variables: the minimum portfolio weight in each portfolio and the maximum portfolio weight in each portfolio. We calculate the average and the standard deviation of the minimum and of the maximum weight for the entire sample and for the four subperiods. The results, reported in Table 3, show that the minimum large loss strategy has the lower minimum and the larger maximum in all subperiods. Safety-first strategies, based only on minimizing the tail probability, are known to produce portfolios where most of the weight is in the stock with smaller loss probability (see for instance Jansen et al. (2000), Hartmann et al. (2004) and Poon et al. (2003)). Here this does not happen. Our two stage procedure of choosing the minimum variance portfolio with the smallest probability of a large loss avoids corner solutions for the portfolio weights. As a consequence the portfolio keeps both the variance diversification effect and a small probability of a large loss.

[Insert Table 3]

### 3.4 Robustness to transaction costs

We evaluate the effect of transaction costs on the performance of the different portfolios by supposing that there is a fixed transaction cost  $c$  on each traded dollar for any stock. We assume that  $c$  includes transaction fees, commissions and bid-ask spread. The total transaction cost is a function of

the turnover rates of all the stocks in the portfolio.

Denote the portfolio weights at the beginning of month  $t$ , after choosing the new portfolio, by the  $d \times 1$  vector  $w_t$  and the assets' returns on month  $t$  by the  $d \times 1$  vector  $R_t$ , where  $d$  is the number of assets in the portfolio. The weight of stock  $i$  in the portfolio during month  $t$  is  $w_{i,t}$  and the return on stock  $i$  on month  $t$  is  $R_{i,t}$ . The portfolio return in month  $t$  is  $R_t w_t$ .

The turnover rate in month  $t$ , after choosing the new portfolio, is

$$TO_t = \sum_{i=1}^d \left| w_{i,t+1} - w_{i,t} \frac{1 + R_{i,t}}{1 + R_t w_t} \right|.$$

In Table 4 we report the average turnover rate for the four strategies. We observe that the minimum VaR and minimum variance strategies have the lowest turnover rates. The minimum large loss strategy has the highest turnover rates. That means that this strategy will be more expensive to implement due to higher transaction costs. Hence, we analyze the effect of transaction costs on the economic value of each strategy.

The total transaction cost in month  $t$  per dollar invested is  $cTO_t$ . The portfolio return net of transaction costs on month  $t$  is  $R_t w_t - cTO_t$ . We assume a transaction cost of  $c = 20$  basis points and use (6) to calculate the fee that an investor would be willing to pay to change from the minimum variance strategy to one of the others. The results are given in Table 4. The minimum large loss strategy continues to have gains in all periods but one, the same as the case without transaction costs. The minimum tail-index and minimum VaR strategies have a positive fee only in the first subperiod

11/1991–06/1996.

We conclude that the minimum large loss strategy presents the highest gains when changing from the minimum variance strategy even when transaction costs are taken into account.

[Insert Table 4]

## 4 Conclusion

It is well known that the mean-variance approach to portfolio selection does not take into account the large losses that are observed in the market. Even when constraining the portfolio selection by minimizing the downside risk, if multivariate normality is assumed, the potential portfolio large losses are still most likely to be underestimated. To overcome this shortcoming researchers have included higher moments in portfolio selection, or estimated portfolio large losses using extreme value theory. The first approach quickly becomes unfeasible because a large number of cross-moments has to be estimated; the second approach usually yields corner solutions where most of the portfolio weight is on the asset with the thinnest tail. These approaches leave still unanswered the question of the economic value of taking large losses into account in portfolio selection.

Our study shows that minimum variance portfolios with lower probability of portfolio large losses outperform minimum variance portfolios. We

find that with this methodology we avoid corner solutions and that an investor is willing to pay a positive fee to change from a minimum variance to a minimum variance with low probability of large losses strategy.

Using a semi-parametric estimator for the probability of a portfolio large loss leads to a best ex post performance when compared with portfolio selection using a classical parametric (univariate) portfolio tail estimator. This result is explained by the fact that the semi-parametric estimator uses directly information from all the portfolio components while the classical estimator uses only information from the univariate portfolio distribution.

The minimum VaR strategy performance lies between the other two tail estimation portfolio strategies. On the one hand VaR uses information from all the assets in the portfolio. On the other hand, it underestimates the probability of large losses because it assumes (in our study) multivariate normality.

Using the semi-parametric estimator for the probability of a portfolio large loss produces strategies with higher turnover rates associated with higher transaction costs. Nevertheless, when transaction costs are included our results show that an investor would still be willing to pay a positive fee of 24 to 49 basis points per year to change from minimum variance to a minimum large loss strategy.

## 5 Appendix

Here we present the estimators used to compute the parameters  $a$ ,  $b$ ,  $\gamma$ ,  $c$  and  $k$  involved in the estimation of the probability of a portfolio large loss in Section 2.2.

Consider a finite random sample with size  $n$  of univariate asset returns  $R_1, R_2, \dots, R_n$ . If we denote the ordered sample returns as  $R_{(n)} \leq R_{(n-1)} \leq \dots \leq R_{(1)}$  then  $R_{(k)}$  is called the  $k$ th upper-order statistic.

### 5.1 The shape parameter (tail index)

Define the function,

$$M_r(R) := \frac{1}{k} \sum_{i=1}^k (\log R_{(i)} - \log R_{(k+1)})^r .$$

for  $r = 1, 2$ . The so-called moment estimator of the shape parameter  $\gamma$  (Dekkers et al. (1989)), is given by

$$\hat{\gamma} := M_1(R) + 1 - \frac{1}{2} \left( 1 - \frac{M_1(R)^2}{M_2(R)} \right)^{-1} . \quad (7)$$

This is a consistent estimator of the shape parameter. Under an additional technical condition (Dekkers et al. (1989)) we have that if  $\gamma \geq 0$  then  $\sqrt{k}(\hat{\gamma} - \gamma)$  has asymptotically a normal distribution with mean zero and variance  $1 + \gamma^2$ . For the case  $\gamma < 0$  (less common in finance applications) the distribution of the statistic  $\hat{\gamma}$  has a more complex expression (see Dekkers et al. (1989)).

## 5.2 The normalizing constants

To estimate the normalizing constants  $a$  and  $b$ , the parameters of the univariate extreme value distribution, we use estimators studied by Dekkers et al. (1989). Define the functions  $\bar{t}$ ,  $\rho_1$  and  $\rho_2$  as

$$\bar{t} = t \wedge 0, \quad \rho_1(t) = \frac{1}{1 - \bar{t}} \quad \text{and} \quad \rho_2(t) = \frac{2}{(1 - \bar{t})(1 - 2\bar{t})}.$$

Then, the estimators for  $a$  and  $b$  are

$$\hat{b} := R_{(k+1)} \tag{8}$$

$$\hat{a} := \frac{R_{(k+1)} \sqrt{3M_1(R)^2 - M_2(R)}}{\sqrt{3(\rho_1(\hat{\gamma}))^2 - \rho_2(\hat{\gamma})}}. \tag{9}$$

## 5.3 The scaling constant $c$ and the set $A$

To calculate the probability of observing large joint movements in a set  $C$ , for some loss  $L$ , we need to estimate the scaling constant  $c$  and the set  $A$ . We have to impose a condition in order to have  $c$  and  $A$  uniquely defined. We need  $A$  to have enough observations  $\tilde{\mathbf{R}}$  so that we can use the non-parametric estimator (2). Given that  $\tilde{\mathbf{R}}$  are standardized observations with unit mean, this happens if we impose the requirement that the unit vector  $\mathbf{1}$  is on the boundary of  $A$  (Dekkers et al. (1989)).

Given a set  $C$ , for some  $L$ , we can always define a function  $f^*(\mathbf{R})$  such that

$$C = \{\mathbf{R} \mid f^*(\mathbf{R}) \geq 1\},$$

Each point  $\mathbf{R}$  in the set of large losses can be written as the transform

of a point  $\tilde{\mathbf{R}}$  by the inverse mapping of  $\tilde{\mathbf{R}} := \left( \mathbf{1} + \hat{\gamma}(\mathbf{R} - \hat{\mathbf{b}})/\hat{\mathbf{a}} \right)^{1/\hat{\gamma}}$ ,

$$\mathbf{R} = \mathbf{a} \frac{\tilde{\mathbf{R}}^{\hat{\gamma}} - \mathbf{1}}{\hat{\gamma}} + \mathbf{b}.$$

At this point we assume that the function  $f^*$  is defined for the case when all returns take the same value. This assumption should not be too restrictive in practice. Hence, from the definition of the function  $f^*$  there exists a value  $x$  such that  $\mathbf{R} = (x, x, \dots, x)$  is solution of the equation  $f^*(\mathbf{R}) = 1$  which is equivalent to the existence of a value  $c$  such that  $c \cdot \mathbf{1}$  is a solution of the equation

$$f^* \left( \mathbf{a} \frac{(c \mathbf{1})^{\hat{\gamma}} - \mathbf{1}}{\hat{\gamma}} + \mathbf{b} \right) = 1.$$

For this  $c$  the unit vector is on the boundary of  $A$ . Since we have only estimates of  $\mathbf{a}$ ,  $\mathbf{b}$  and  $\gamma$  we can find only an estimate  $\hat{c}$  of  $c$  as the solution of  $f^*(\hat{\mathbf{a}}((\hat{c} \mathbf{1})^{\hat{\gamma}} - \mathbf{1})/\hat{\gamma} + \hat{\mathbf{b}}) = 1$ .

Finally, given that  $cA$  is the set of asset returns belonging to  $C$  after being standardized, we define

$$\hat{A} := \frac{1}{\hat{c}} \left( \mathbf{1} + \hat{\gamma} \frac{C - \hat{\mathbf{b}}}{\hat{\mathbf{a}}} \right)^{1/\hat{\gamma}}$$

as the estimator of  $A$ .

## References

- Alexander, G. J. and Baptista, A. M., 2002. Economic implications of using a mean-VaR model for portfolio selection: a comparison with mean-variance analysis. *Journal of Economic Dynamics & Control* 26, 1159–1193.
- Arzac, E. R. and Bawa, V. S., 1977. Portfolio choice and equilibrium in capital markets with safety-first investors. *Journal of Financial Economics* 4, 277–288.
- Balkema, A. A. and de Haan, L., 1974. Residual life time at great age. *The Annals of Probability* 2, 792–804.
- Brandt, M., 2009. Portfolio choice problems, in: Hansen, L. P. and Aït-Sahalia, Y. (Eds.), *Handbook of Financial Econometrics*, Vol 1. North-Holland, Amsterdam, pp. 267–336.
- Chamberlain, G., 1983. A characterization of the distributions that imply mean-variance utility functions. *Journal of Economic Theory* 29, 185–201.
- de Haan, L. and de Ronde, J., 1998. Sea and wind: Multivariate extremes at work. *Extremes* 1, 7–45.
- de Haan, L. and Resnick, S., 1977. Limit theory for multivariate extremes. *Z. Wahrsch. Verw. Gebiete* 40, 317–337.
- de Haan, L. and Resnick, S., 1993. Estimating the limit distribution of multivariate extremes. *Communications in Statistics. Stochastic Models* 9(2), 275–309.
- Dekkers, A. L. M., Einmahl, J. H. J., and de Haan, L., 1989. A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics* 17(4), 1833–1855.
- Dias, A., 2009. Semi-parametric estimation of portfolio large losses. Working paper, Warwick Business School, University of Warwick.
- Dittmar, R., 2002. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross-section of equity returns. *Journal of Finance* 57, 369–403.
- Embrechts, P., Klüppelberg, C., and Mikosch, T., 1997. *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, Berlin.
- Geluk, J. and de Hann, L., 1987. Regular variation, extensions and Tauberian theorems. *CWI Tract*, vol. 40, CWI, Amsterdam.
- Gourieroux, C., Laurent, J. P., and Scaillet, O., 2000. Sensitivity analysis of values at risk. *Journal of Empirical Finance* 7, 225–246.

- Hartmann, P., Straetmans, S., and de Vries, C. G., 2004. Asset market linkages in crisis periods. *Review of Economics and Statistics* 86, 313–326.
- Harvey, C. and Siddique, A., 2000. Conditional skewness in asset pricing tests. *Journal of Finance* 55, 1263–1295.
- Hyung, N. and de Vries, C. G., 2007. Portfolio selection with heavy tails. *Journal of Empirical Finance* 14, 383–400.
- Jansen, D., Koedijk, K. J., and de Vries, C. G., 2000. Portfolio selection with limited down-side risk. *Journal of Empirical Finance*, 7, 247–269.
- Leadbetter, M. R., Lindgren, G., and Rootzén, H., 1983. *Extremes and Related Properties of Random Sequences and Processes*. Springer Verlag, New York.
- Longin, F. and Solnik, B., 2001. Extreme correlation of international equity markets. *Journal of Finance* 56(2), 649–676.
- Mandelbrot, B. B., 1963. The variation of certain speculative prices. *Journal of Business* 36, 394–419.
- Markowitz, H. M., 1952. Portfolio selection. *Journal of Finance* 7, 77–91.
- Martellini, L. and Ziemann, V., 2010. Improved estimates of higher-order comoments and implications for portfolio selection. *Review of Financial Studies* 23(4), 1467–1502.
- Merton, R. C., 1972. An analytic derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis* 7, 1851–1872.
- Mitton, T. and Vorkink, K., 2007. Equilibrium underdiversification and the preference for skewness. *Review of Financial Studies* 20, 1255–1288.
- Pickands, J., 1975. Statistical inference using extreme order statistics. *The Annals of Statistics* 3, 119–131.
- Poon, S. H., Rockinger, M., and Tawn, J., 2003. Nonparametric extreme value dependence measures and finance applications. *Statistica Sinica* 13, 929–953.
- Roy, A. D., 1952. Safety first and holding of assets. *Econometrica* 20, 431–449.
- Solnik, B. H., 1995. Why Not Diversify Internationally Rather than Domestically? *Financial Analysts Journal* 51, 89–94.

Table 1: Ex post performance of each strategy

Period	Obs.	$\mu$	$\sigma^-$	$\sigma^+$	$SR$	$SoR$	$\delta_1$	$\delta_{10}$
Panel A: Minimum variance strategy								
Entire sample	224	0.0675	0.1001	0.1103	0.189	0.266	-	-
11/1991-06/1996	56	0.0774	0.0606	0.0853	0.268	0.409	-	-
07/1996-02/2001	56	0.1365	0.1127	0.1465	0.461	0.701	-	-
03/2001-10/2005	56	0.0682	0.0883	0.1029	0.139	0.201	-	-
11/2005-06/2010	56	-0.0073	0.1261	0.0964	-0.333	-0.398	-	-
Panel B: Minimum large loss strategy (semi-parametric)								
Entire sample	224	0.0986	0.1130	0.1474	0.315	0.491	2.62	2.37
11/1991-06/1996	56	0.2085	0.0657	0.1234	1.182	2.130	12.08	11.94
07/1996-02/2001	56	0.1743	0.1254	0.1798	0.557	0.910	3.02	2.74
03/2001-10/2005	56	-0.0189	0.1366	0.1212	-0.350	-0.450	-8.55	-8.84
11/2005-06/2010	56	0.0452	0.1111	0.1572	-0.016	-0.026	4.98	4.73
Panel C: Minimum tail-index strategy (parametric)								
Entire sample	224	0.0667	0.1214	0.1284	0.154	0.215	-0.30	-0.49
11/1991-06/1996	56	0.1623	0.0621	0.1059	0.941	1.594	7.86	7.80
07/1996-02/2001	56	0.1436	0.1319	0.1544	0.453	0.656	0.46	0.31
03/2001-10/2005	56	0.0038	0.1374	0.1165	-0.234	-0.298	-6.39	-6.67
11/2005-06/2010	56	-0.0307	0.1373	0.1318	-0.395	-0.517	-2.62	-2.83
Panel D: Minimum VaR strategy								
Entire sample	224	0.0756	0.0997	0.1119	0.240	0.341	0.75	0.74
11/1991-06/1996	56	0.1112	0.0467	0.0840	0.653	1.114	3.21	3.26
07/1996-02/2001	56	0.1145	0.1283	0.1484	0.323	0.466	-2.06	-2.16
03/2001-10/2005	56	0.0888	0.0721	0.1084	0.299	0.493	1.98	2.01
11/2005-06/2010	56	-0.0078	0.1262	0.0964	-0.336	-0.402	-0.05	-0.05

For the entire sample and four subperiods we summarize in this table the ex post performance of each of the strategies. We report the annualized realized mean return ( $\mu$ ), the annualized realized negative (positive) semi-volatility ( $\sigma^-$  ( $\sigma^+$ )), the realized Sharpe ratio ( $RS$ ), the realized Sortino ratio ( $SoR$ ) and the average annualized percentage fee ( $\delta_\gamma$ ) that an investor with constant risk aversion of  $\gamma = 1$  or  $\gamma = 10$  would be willing to pay to change from the minimum variance strategy to each of the other strategies considered.

Table 2: Descriptive statistics of the portfolio returns for each strategy

Period	Min	5%	Med	95%	Max	$p^-$	$m^3$	$m^4$
Panel A: Minimum variance strategy								
Entire sample	-0.878	-0.562	0.093	1.176	4.445	0.428	-0.421	1.485
11/1991-06/1996	-0.502	-0.416	0.054	0.840	1.042	0.446	-0.060	-0.833
07/1996-02/2001	-0.743	-0.626	0.171	1.530	4.445	0.357	-0.121	0.109
03/2001-10/2005	-0.841	-0.454	0.086	1.312	2.024	0.464	-0.498	3.030
11/2005-06/2010	-0.878	-0.631	0.060	1.033	1.619	0.446	-0.963	1.555
Panel B: Minimum large loss strategy (semi-parametric)								
Entire sample	-0.811	-0.645	0.104	1.786	14.800	0.401	0.408	2.431
11/1991-06/1996	-0.545	-0.475	0.295	1.278	2.183	0.285	-0.082	-0.030
07/1996-02/2001	-0.786	-0.706	0.180	2.039	9.249	0.392	0.214	1.008
03/2001-10/2005	-0.811	-0.708	-0.003	1.412	2.318	0.499	-0.259	-0.262
11/2005-06/2010	-0.779	-0.599	0.077	1.561	14.800	0.428	1.533	6.860
Panel C: Minimum tail-index strategy (parametric)								
Entire sample	-0.948	-0.551	0.106	1.581	5.379	0.419	-0.485	1.927
11/1991-06/1996	-0.536	-0.432	0.268	1.106	2.128	0.339	-0.173	0.047
07/1996-02/2001	-0.876	-0.678	0.284	1.905	3.319	0.375	-0.542	0.447
03/2001-10/2005	-0.948	-0.557	0.017	1.284	3.200	0.464	-1.025	4.754
11/2005-06/2010	-0.823	-0.690	-0.007	1.484	5.379	0.499	0.163	0.950
Panel D: Minimum VaR strategy								
Entire sample	-0.879	-0.580	0.090	1.292	4.321	0.424	-0.427	1.620
11/1991-06/1996	-0.422	-0.350	0.094	0.817	1.219	0.392	0.125	-0.408
07/1996-02/2001	-0.797	-0.681	0.183	1.551	4.321	0.357	-0.323	0.236
03/2001-10/2005	-0.734	-0.397	0.027	1.410	2.706	0.499	0.299	1.583
11/2005-06/2010	-0.879	-0.631	0.059	1.022	1.620	0.446	-0.965	1.580

We report in this table the descriptive statistics of the realized portfolio returns: annualized minimum, median and maximum realized returns (Min, Med, Max), annualized realized 5% and 95% quantiles (5%, 95%), realized frequency of negative returns ( $p^-$ ), realized skewness ( $m^3$ ) and realized kurtosis ( $m^4$ ).

Table 3: Portfolio allocation analysis

Period	Minimum weight		Maximum weight	
	Mean	Stdev	Mean	Stdev
Panel A: Minimum variance strategy				
Entire sample	0.190	0.043	0.506	0.023
11/1991–06/1996	0.164	0.030	0.477	0.018
07/1996–02/2001	0.150	0.020	0.507	0.016
03/2001–10/2005	0.204	0.032	0.532	0.005
11/2005–06/2010	0.241	0.003	0.508	0.004
Panel B: Minimum large loss strategy (semi-parametric)				
Entire sample	0.081	0.106	0.615	0.151
11/1991–06/1996	0.007	0.005	0.639	0.071
07/1996–02/2001	0.101	0.106	0.635	0.185
03/2001–10/2005	0.128	0.114	0.590	0.177
11/2005–06/2010	0.090	0.113	0.595	0.143
Panel C: Minimum tail-index strategy (parametric)				
Entire sample	0.139	0.057	0.586	0.096
11/1991–06/1996	0.096	0.053	0.615	0.106
07/1996–02/2001	0.124	0.047	0.603	0.086
03/2001–10/2005	0.148	0.052	0.589	0.107
11/2005–06/2010	0.189	0.032	0.537	0.063
Panel D: Minimum VaR strategy				
Entire sample	0.190	0.045	0.521	0.012
11/1991–06/1996	0.143	0.003	0.515	0.002
07/1996–02/2001	0.156	0.021	0.526	0.014
03/2001–10/2005	0.219	0.028	0.532	0.007
11/2005–06/2010	0.242	0.002	0.511	0.003

The table illustrates the distribution of the assets weights in each portfolio. We present the mean and the standard deviation of the minimum and maximum asset weights in each portfolio. Stdev stands for standard deviation.

Table 4: Turnover and robustness to transaction costs

Period	Turnover	$\delta_1$	$\delta_{10}$
Panel A: Minimum variance strategy			
Entire sample	0.048	-	-
11/1991–06/1996	0.043	-	-
07/1996–02/2001	0.079	-	-
03/2001–10/2005	0.047	-	-
11/2005–06/2010	0.022	-	-
Panel B: Minimum large loss strategy (semi-parametric)			
Entire sample	0.927	0.49	0.24
11/1991–06/1996	0.794	10.32	10.18
07/1996–02/2001	1.185	0.32	0.045
03/2001–10/2005	0.920	-10.47	-10.76
11/2005–06/2010	0.808	3.06	2.82
Panel C: Minimum tail-index strategy (parametric)			
Entire sample	0.575	-2.57	-2.76
11/1991–06/1996	0.612	5.25	5.18
07/1996–02/2001	0.527	-2.16	-2.31
03/2001–10/2005	0.726	-8.35	-8.63
11/2005–06/2010	0.432	-4.41	-4.62
Panel D: Minimum VaR strategy			
Entire sample	0.042	-1.46	-1.47
11/1991–06/1996	0.020	0.92	0.96
07/1996–02/2001	0.082	-4.63	-4.73
03/2001–10/2005	0.046	-0.14	-0.10
11/2005–06/2010	0.020	-1.89	-1.89

In this table we report the average turnover rate and the average annualized percentage fee ( $\delta_\gamma$ ) that an investor with constant risk aversion of  $\gamma = 1$  or  $\gamma = 10$  would be willing to pay to change from the minimum variance strategy to each of the other strategies considered.