VU Research Portal

A note on optimal estimation from a risk management perspective under possibly mis-specified tail behavior

Lucas, A.

1997

document version

Early version, also known as pre-print

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Lucas, A. (1997). *A note on optimal estimation from a risk management perspective under possibly mis-specified tail behavior*. (Research Memorandum; No. 1997-56). FEWEB.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Serie Research Memoranda

A Note on Optimal Estimation from a Risk Management Perspective
under Possibly Mis-specified Tail behavior

André Lucas

Research Memorandum 1997-56

November 1997

.

vrije Universiteit *amsterdam*

A Note on Optimal Estimation from a Risk Management Perspective under Possibly Mis-specified Tail Behavior

André Lucas*

*Financial Sector Management, Free University, De Boelelaan 1105,
1081HV Amsterdam, the Netherlands, email: alucas@econ.vu.nl*

Many financial time-series show leptokurtic behavior, i.e., fat tails. Such tail behavior is important for risk management. In this paper I focus on the calculation of Value-at-Risk (VaR) as a downside-risk measure for optimal asset portfolios. Using a framework centered around the Student t distribution, I explicitly allow for a discrepancy between the fat-tailedness of the true distribution of asset returns and that of the distribution used by the investment manager. As a result, numbers for the over-estimation or under-estimation of the true VaR of a given portfolio can be computed. These numbers are used to rank several well-known estimation methods for determining the unknown parameters of the distribution of asset returns. Minimizing the absolute (percentage) mismatch between the nominal and actual or true VaR leads to the choice of a Gaussian maximum quasi-likelihood estimator, i.e., a least-squares type estimator. The maximum likelihood estimator has a less satisfactory behavior. Outlier robust estimators perform even worse if the required confidence level for the VaR is high. An explanation for these results is provided.

Keywords: Value-at-Risk; leptokurtosis; downside-risk; optimal asset allocation; model mis-specification; minimax optimality; robustness; risk management; quasi-likelihood.

JEL codes: G11; C13; C44.

1. Introduction and summary

Uncertainty is the key ingredient of most financial and economic decision problems. For example, an investment manager trying to design a solid investment policy has to come up with a consistent set of forecasts for future

*I have benefitted from discussions with Guus Boender, Bernard Compaijen, Cees Dert, Patrick Groenendijk, Pieter Klaassen, and Peter Vlaar. Only I am responsible for any remaining errors or obscurities.

returns on alternative investment opportunities. In certain cases, predictions of macro-economic developments are needed as well, e.g., inflation rates in case of pension funds with liabilities that are defined in real terms. Other sources of risk affecting the performance of financial policies include interest rate risk, exchange rate risk, credit risk, etc.

In order to characterize the risk associated with decision making under uncertainty, different risk measures have been proposed. In the framework of asset management, the most familiar risk measure is the standard deviation of portfolio returns, see, e.g., Markowitz (1959). The main drawback of the standard deviation as a measure of risk is that it is symmetric: extreme positive returns are treated the same way as extreme negative returns. As an alternative to the standard deviation, several downside-risk measures have been proposed. By far the most popular downside-risk measure used today is the Value-at-Risk (VaR) as introduced by J.P. Morgan. The VaR measures the maximum (dollar) loss on a portfolio over a given period of time given a certain confidence level. For example, if the VaR of a portfolio is 10 million dollar with a confidence level of 99% and a holding period of 10 days, this means that there is only a 1% probability that the portfolio will produce a loss of more than 10 million dollar if it cannot be liquidated within a 10 days period. The popularity of VaR is enhanced by the fact that regulatory institutions have adopted the use of VaR as a measure of risk in their supervisory policies. The Basle Committee on Banking Supervision (1996), for example, proposes to require banks to report VaR figures to the supervisory institution, e.g., the Central Bank, on a regular basis. Within certain limits, banks are allowed to construct their own models for computing such VaR measures.

Given the popularity of VaR and its importance for supervision, it is interesting to study the properties of alternative methods for computing VaR. All methods for calculating VaR have to deal with the fact that it is difficult to obtain reliable estimates of the (lower) quantiles of a distribution from a limited number of data points. Different methods have been proposed, ranging from the use of parametric models (often the normal distribution), via the use of semi-nonparametric models, to the employment of fully non-parametric methods like historical simulation. For a recent survey on issues involving VaR, see Jorion (1997). Evidently, the tail behavior of the distribution of asset returns plays a prominent role in the calculation of VaR. Given the fat-tailedness observed for many financial time-series, see, e.g., de Vries (1994) for foreign exchange markets or Campbell et al. (1997) for stock markets, blind use of the normal distribution or variants thereof in the present context is clearly suspect.

In the present paper, I focus on the effect of fat-tailedness on model-based

parametric VaR calculations. Following Lucas and Klaassen (1996), I use an integrated framework in which optimal financial policies and restrictions on allowable VaR are dealt with simultaneously. This slightly departs from the mainstream literature on VaR calculation, which usually treats the composition of the portfolio of assets as fixed. This is not fully satisfactory, as bounds on allowable VaR are set by supervisory institutions before investment managers determine their optimal financial policies. The direct use of restrictions on VaR in a model for determining the optimal asset allocation, allows for a more natural way to study the effect of alternative VaR estimation methods on the properties of optimal financial policies.

The aim of this paper is to rank different estimation methods for the unknown parameters in the distribution of asset returns in case this distribution is possibly mis-specified. Throughout this paper, I make heavy use of the Student t distribution as device for describing the behavior of asset returns. The Student t distribution has several advantages. First, it nests the normal distribution, which is frequently used in practice. Second, the Student t distribution is a fat-tailed alternative for the normal distribution that easily yields to tractable derivations. The qualitative insights obtained by using the Student t distribution readily carry over to other fat-tailed distributions. A drawback of the Student t distribution is that it does not describe volatility clustering, a stylized fact observed for many financial time-series. In the present paper, where I focus on a one-period model, this shortcoming is not very restrictive. Moreover, the results of the present paper are also relevant for cases with volatility clustering. This follows from the fact that at longer forecast horizons, the forecasting density begins to look more like the unconditional density of asset returns, which in the case of (generalized) autoregressive volatility clustering (ARCH and GARCH) is leptokurtic, see, e.g., Nelson (1990). This leptokurtic distribution can then be approximated by a Student t distribution.

The framework used in the present paper is as follows. Using a stylized one-period asset allocation model centered around the Student t distribution, I obtain optimal asset allocation strategies for the investment manager. These strategies can be used to obtain estimates of the true VaR if the distribution used by the manager in his optimization problem differs from the true distribution of asset returns. Naturally, in order to obtain definite results, an estimation method must be chosen to match the characteristics of the distribution used by the manager to the characteristics of the true distribution of asset returns. I consider the class of maximum (quasi-)likelihood estimators (MQL) based on the Student t distribution, see White (1982) and Gouridroux et al. (1984). This class contains several well-known estimators such as the least-squares estimator, the maximum likelihood estimator,

and some outlier robust estimators. Using the mismatch between the postulated and the actual VaR, these different estimators can be ranked according to their performance. Minimizing the maximum absolute percentage discrepancy between the postulated and the true VaR results in the choice of the least-squares estimator as the optimal estimator for VaR calculations in practical settings. The maximum likelihood estimator based on the degree of leptokurtosis specified by the investment manager, performs less satisfactory. The performance of the outlier robust estimators is even worse.

The explanation of these results follows from the fact that in the proposed framework, two quantities determine the VaR, namely: (1) the dispersion measure of the distribution (e.g., the variance-covariance matrix), and (2) the fat-tailedness of the distribution. If the manager uses the least-squares estimator, his dispersion measure is positively biased if reality is fat-tailed. This partially offsets the bias in the VaR estimate caused by neglected leptokurtosis. By contrast, if an outlier robust estimation method is used, the estimate of dispersion is much less affected by a mismatch between postulated and actual leptokurtosis. This follows directly from the properties of outlier robust estimators. Given the smaller bias in the dispersion estimate, the effect of the (mis-specified) degree of fat-tailedness will be greater, leading to larger discrepancies between actual and postulated VaR. The case of the maximum likelihood estimator lies, somewhere inbetween.

The paper is set up as follows. Section 2 describes the framework used, while Section 3 presents the numerical results. Section 4 concludes.

2. Framework

Throughout this paper, I use the Student t distribution for describing the behavior of asset returns. The Student t distribution is characterized by three parameters: the location μ , the precision Ω^{-1} , and the degrees of freedom parameter ν . The precise form of the t distribution is

$$f(r) = f(r; \mu, \Omega, \nu) = \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) |\pi\nu\Omega|^{1/2}} \left(1 + \frac{(r - \mu)'\Omega^{-1}(r - \mu)}{\nu}\right)^{-(\nu+n)/2}, \quad (1)$$

with $r \in \mathbb{R}^n$ denoting the vector of asset returns, with n denoting the number of asset categories. It is well known that the normal distribution is nested in the class of Student t distributions. This is easily seen by setting $\nu = \infty$.

Given the stochastic behavior of asset returns, I now turn to the decision problem faced by the investment manager. Consider a manager who has an amount of one dollar, which he can invest in the above n asset categories. The manager is not allowed to take short positions. Moreover, the manager has a

given risk tolerance in the sense that the probability that his portfolio has a return below r^{low} is sufficiently small. Given his risk tolerance, the manager tries to maximize his expected return. Formally, the decision problem is given by the following equations:

$$\max_{x \in \mathbb{R}^n} x' \mu_m \quad (2)$$

subject to

$$x_i \geq 0, \quad \sum_{i=1}^n x_i = 1, \quad P_m(x'(1+r) \leq 1+r^{low}) \leq \psi, \quad (3)$$

with P_m denoting the cumulative Student t distribution with parameters μ_m , Ω_m , and ν_m used by the manager, ψ denoting the required confidence level for the probabilistic constraint in (3), and x_i denoting the fraction or amount invested in asset category $i = 1, \dots, n$. The probabilistic requirement on the portfolio's return in (3) is called a shortfall constraint, see Leibowitz and Kogelman (1991) and Leibowitz et al. (1992). As an alternative to (2) and (3), we could use a general utility function in order to characterize the manager's preferences. The set-up of the decision problems as presented above, however, has the advantage that it is more closely linked to standard practice in investment management. The use of downside-risk measures to assess the performance of a given portfolio is widely accepted. Moreover, the investment problem as stated above is directly linked to VaR. For $r^{low} < 0$, the VaR per dollar invested is given by $-r^{low}$, with associated confidence level $1 - \psi$. For example, for $r^{low} = -0.10$ and $\psi = 0.01$ the manager looks for the portfolio with maximum expected return subject to the constraint that the VaR of this portfolio does not exceed 10% of the invested notional amount with a probability of 99%.

As fat tails are intimately linked to the occurrence of extreme returns, it is not surprising that the choice of the distribution P_m and its parameters μ_m , Ω_m , and ν_m , affect the form of the optimal portfolio. Lucas and Klaassen (1996) show that the direction of the effect heavily depends on the value of the required confidence level $1 - \psi$. Low levels of confidence lead to more aggressive asset allocations for fatter-tailed distributions, i.e., allocations with more volatile portfolio returns. The reverse holds for high levels of confidence. Evidently, "more aggressive" can become "too aggressive" if the fat-tailedness of P_m either exceeds or falls below that of the true distribution of asset returns. The mismatch can be quite substantial, see also Section 3. In order to capture the mismatch between true and postulated VaR, I introduce the true distribution of asset returns, P_t , which is characterized by the parameters μ_t , Ω_t , and ν_t .

If both the true distribution P_t and the distribution used by the manager P_m belong to the class of Student t distributions, it would be best to use the maximum likelihood estimator to fix μ_m , Ω_m , and ν_m directly at μ_t , Ω_t , and ν_t , respectively. In the present paper, however, I consider the situation where the value of ν_m is set by the manager *a priori*. This occurs, for example, if the manager uses the normal distribution while the true distribution of asset returns is possibly fat-tailed. Alternatively, the manager can use the full maximum likelihood estimator based on past data, while there is a structural break in the degree of leptokurtosis between the sample period and the planning period. In that case, there will also be a mismatch between the true degree of fat-tailedness (ν_t) and the postulated or estimated one (ν_m).

Given the value of ν_m , the manager tries to fit his postulated distribution P_m to the true distribution P_t . From the many different possibilities for achieving this, I consider the class of maximum quasi-likelihood (MQL) estimators based on the Student t distribution, see White (1982) and Gouriéroux et al. (1984). This class of estimators nests several well-known estimation principles, like the least-squares estimator and the maximum likelihood (ML) estimator. Also some outlier robust estimators are contained within this class. Note that the quasi-likelihood need not coincide with either the true likelihood or the likelihood postulated by the manager. For example, the manager might use the least-squares estimator to estimate the unknown parameters μ and Ω , irrespective of whether he believes that asset returns are fat-tailed or not. The MQL estimator solves the maximization problem

$$\max_{\mu, \Omega} E. [\log(f(r; \mu, \Omega, \nu_e))], \quad (4)$$

where $f(\cdot)$ is given in (1), and ν_e denotes the degrees of freedom parameter determining the MQL estimator used, see White (1982), Gouriéroux et al. (1984), and Hampel et al. (1986). The symbol $E(\cdot)$ denotes the expectations operator, taken with respect to either the true measure P_t , $E_t(\cdot)$, or the manager's distribution P_m , $E_m(\cdot)$. Denote the optimum values of μ and Ω following from (4) by μ_e and Ω_e , respectively. Note that we now have three sets of parameters (μ, Ω, ν) . The parameters indexed by m are the parameters as used by the manager in his optimization problem (2)-(3). The parameters indexed by t are the true parameters characterizing the distribution of asset returns. Finally, the parameters indexed by e are the parameters solving the score equations for the MQL estimator. For example, if $\nu_e = \nu_m$, $\mu_m = \mu_e$ and $\Omega_m = \Omega_e$ denote the ML estimates. Alternatively, if $\nu_e = \infty$, μ_e and Ω_e denote the least-squares estimates of the unknown parameters.

The first order conditions or quasi-score equations corresponding to (4)

are

$$E. \left[\frac{r - \mu_e}{1 + (r - \mu_e)' \Omega_e^{-1} (r - \mu_e) / \nu_e} \right] = 0, \quad (5)$$

and

$$E. \left[\frac{(1 - k/\nu_e)(r - \mu_e)(r - \mu_e)'}{1 + (r - \mu_e)' \Omega_e^{-1} (r - \mu_e) / \nu_e} \right] = \Omega_e. \quad (6)$$

As both P_m and P_t are spherically symmetric in $(r - \mu_m)$ and $(r - \mu_t)$, respectively, and the function in brackets in (5) is odd in $(r - \mu_e)$, it is evident that $\mu_e = \mu_m$ for P_m , and $\mu_e = \mu_t$ for P_t . Using this fact and considering the case of P_m , it is easy to verify that (6) is satisfied by the matrix $\Omega_e = c_m \Omega_m$ for some positive constant c_m . A similar constant c_t exists for P_t . The constants c_m and c_t depend on ν_e and on ν_m and ν_t , respectively. They can most easily be determined using numerical techniques. Note, however, that for P_m and $\nu_e = \nu_m$, $c_m = 1$, while for $\nu_e = \infty$, $c_m = \nu_m / (\nu_m - 2)$.

Given the results described above, the investment manager proceeds as follows. After postulating the degree of leptokurtosis ν_m and defining the estimation principle ν_e , he obtains estimates μ_e and Ω_e . Moreover, he can compute the constant c_m . From all these estimates, the manager computes the inputs for his optimization problem (2)-(3) as $\mu_m = \mu_e$ and $\Omega_m = \Omega_e / c_m$. Next, the manager solves the asset allocation problem and obtains an optimal asset allocation x_m .

Note that the estimation is carried out using the true data generating process. This is modeled by substituting the true probability measure P_t into (5) and (6). As a result, $\mu_e = \mu_t$ and $\Omega_e = c_t \Omega_t$, such that $\mu_m = \mu_t$ and $\Omega_m = (c_t / c_m) \Omega_t$. I distinguish the VaR per dollar invested postulated by the manager, $-r_m^{low} = -r_m^{low}$, from the true VaR, $-r_t^{low}$. If the VaR constraint of the asset allocation problem (2)-(3) is binding (as is the case for the examples considered in Section 3), the true VaR per dollar invested ($-r_t^{low}$) of the manager's optimal asset allocation x_m can be computed as follows:

$$\begin{aligned} \psi &= P_t(x_m'(1 + r) \leq 1 + r_t^{low}) \\ &= P_t(x_m' r \leq r_t^{low}) \\ &= P_t \left(\frac{x_m'(r - \mu_t)}{\sqrt{x_m' \Omega_t x_m}} \leq \frac{r_t^{low} - x_m' \mu_t}{\sqrt{x_m' \Omega_t x_m}} \right) \\ &= P_t \left(\frac{x_m'(r - \mu_t)}{\sqrt{x_m' \Omega_t x_m}} \leq \frac{r_m^{low} - (r_m^{low} - r_t^{low}) - x_m' \mu_t}{\sqrt{x_m' \Omega_t x_m}} \right) \\ &= P_t \left(\frac{x_m'(r - \mu_t)}{\sqrt{x_m' \Omega_t x_m}} \leq \frac{-(r_m^{low} - r_t^{low})}{\sqrt{x_m' \Omega_t x_m}} + \frac{r_m^{low} - x_m' \mu_m}{\sqrt{x_m' \Omega_m x_m}} \sqrt{\frac{c_t}{c_m}} \right) \end{aligned}$$

$$= P_t \left(\frac{x'_m(r - \mu_t)}{\sqrt{x'_m \Omega_t x_m}} \leq \frac{-(r_m^{low} - r_t^{low})}{\sqrt{x'_m \Omega_t x_m}} + P_m^{-1}(\psi) \sqrt{\frac{c_t}{c_m}} \right), \quad (7)$$

such that

$$\begin{aligned} (r_m^{low} - r_t^{low}) &= \left(P_m^{-1}(\psi) \sqrt{\frac{c_t}{c_m}} - P_t^{-1}(\psi) \right) \sqrt{x'_m \Omega_t x_m} \\ &= \left(P_m^{-1}(\psi) \left[\sqrt{\frac{c_t}{c_m}} - 1 \right] + [P_m^{-1}(\psi) - P_t^{-1}(\psi)] \right) \sqrt{x'_m \Omega_t x_m}, \end{aligned} \quad (8)$$

with P_t^{-1} and P_m^{-1} denoting the inverses of the **standard** Student t distributions with ν_t and ν_m degrees of freedom, respectively. Equation (8) shows that the mismatch between the actual and the postulated VaR per dollar invested depends on three factors. First, there is the effect of the bias in the estimate of Ω , reflected in the fraction c_t/c_m . Second, there is the direct effect of the degree of fat-tailedness of the manager's distribution compared to the true degree of leptokurtosis. Third, the bias in Ω has an effect on the optimal asset allocation that is chosen, x . As demonstrated in the next section, the right-hand side of (8) can take on either a positive or a negative sign, such that both under-estimation and over-estimation of the true VaR are possible.

3. Computational Results

In this section I present some numerical results corresponding to the findings of the previous section. As MQL estimators for the unknown parameters μ and Ω I consider the least-squares estimator, $\nu_e = \infty$, the maximum likelihood estimator $\nu_e = \nu_m$, and the Student t(5) MQL estimator, $\nu_e = 5$. The first two estimators are two well-known benchmarks. The third estimator can be interpreted as an outlier robust estimator. It is less sensitive to outliers than the least-squares estimator, while at the same time being reasonably efficient if asset returns happen to be normally distributed, see, e.g., Hampel et al. (1986) and Franses and Lucas (1997). In order to check the sensitivity of the results to different parameter configurations, I consider $\psi = 1\%$, 5% and $r_m^{low} = -5\%$, -10% . The nominal shortfall returns r_m^{low} are on an annual basis using monthly compounding, see also the data description below. The above parameter values were also considered in Lucas and Klaassen (1996). Note that $\psi = 1\%$ is most relevant for practical purposes in the context of risk management for banks given the Basle (1996) proposals.

Equation (8) reveals that the mismatch between actual and postulated VaR also depends on the asset mix x_m and the true precision of asset returns,

Ω_t^{-1} . In order to obtain reasonable figures for these quantities from an applied perspective, empirical data is used. From DATASTREAM, I obtain monthly total returns on the S&P500 index, monthly holding period returns for 10 year treasury bonds, and the interest rate for 1 month Eurodollar deposits. The value of Ω_t is set equal to $(1 - 2/\nu_t)$ times the variance-covariance matrix of these three time-series, such that the true distributions of asset returns all have the same variance. Based on all the parameter values mentioned above, the asset allocation problem (2)-(3) is solved for different combinations (ν_m, ν_t) using the optimization package AIMMS. The resulting optimal policies x_m are plugged into (8), giving the mismatch between postulated and actual VaR. Table 1 presents some percentage mismatches for several combinations (ν_m, ν_t) .

The first thing to note in Table 1 is that the percentage mismatch in VaR per dollar invested can be quite large. For example, if the manager uses the normal distribution ($\nu_m = \infty$) and the least-squares estimator ($\nu_e = \infty$) while reality is leptokurtic ($\nu_t = 3$), the mismatch ranges from -45 to +31 per cent of the postulated VaR. Also note that for $(\nu_m, \nu_t, \nu_e, \psi, r_m^{low}) = (\infty, 3, 5, 1\%, -5\%)$, the true VaR is about 2.5 times larger than the postulated VaR. Obviously, if the postulated and the true degrees of freedom parameters coincide, there is no mismatch, irrespective of the estimation methodology characterized by the choice of ν_e .

The results for the least-squares estimator $\nu_e = \infty$ corroborate the findings in Lucas and Klaassen (1996). It turns out that for this case the required confidence level $1 - \psi$ for the VaR plays a crucial role for the direction of the mismatch. High levels of confidence (99%) result in an under-estimation of VaR if the distribution used by the manager has thinner tails than the true distribution. The reverse holds for low levels of confidence (95%). Note that this result no longer holds if different estimators for the unknown parameters μ and Ω are considered.

For fixed values of ν_m , the mismatch appears to be monotonously increasing or decreasing in ν_t . Moreover, for $\nu_e = \infty$ and $\nu_e = 5$ the mismatch is also monotone in ν_m for fixed ν_t . By contrast, for the ML estimator, $\nu_e = \nu_m$, the mismatch sometimes follows an inverted U-shaped pattern in ν_m for fixed values of ν_t , e.g., for $\psi = 5\%$ and $r_m^{low} = -10\%$.

Except for $\nu_m = \infty$, the numbers in Table 1 for high levels of confidence ($\psi = 1\%$) are much smaller in absolute magnitude for the least-squares estimator than for the ML estimator ($\nu_e = \nu_m$) and the MQL $t(5)$ estimator ($\nu_e = 5$). The conclusion is more mixed for $\psi = 5\%$, again stressing the importance of the choice of the VaR's confidence level. Note that for $\psi = 5\%$ the direction of the bias for $\nu_e = \infty$ and for $\nu_e = \nu_m$ or $\nu_e = 5$ is opposite, excepting the Gaussian case $\nu_m = \infty$ for the ML estimator $\nu_e = \nu_m$.

Table 1

The table contains the percentage mismatch between the postulated value-at-risk (VaR) per dollar invested ($-r_m^{low}$) and the true VaR ($-r_t^{low}$), i.e., $100 \cdot (r_t^{low}/r_m^{low} - 1)$. The values of r_m^{low} mentioned in the table are annualized (using monthly compounding), while the percentage mismatches are based on the corresponding monthly shortfall returns. ν_t denotes the true degrees of freedom parameter of asset returns. ν_m is the degrees of freedom parameter used by the manager to determine the optimal asset allocation. ν_e denotes the degrees of freedom parameter of the quasi-likelihood used for estimating the unknown parameters. The constant $1 - \psi$ denotes the confidence level for the VaR. For example, the entry 15 for $\nu_e = \nu_m = 10$, $\nu_t = 3$, $r_m^{low} = -10\%$, and $\psi = 5\%$, means that if the manager uses a Student t distribution with 10 degrees of freedom while in reality asset returns follow a Student t distribution with 3 degrees of freedom, and if the manager uses the maximum likelihood estimator for $\nu_m = 10$ to estimate mean returns and covariances, and if the postulated VaR per dollar invested at a 95% confidence level is 0.10 dollar on an annual basis, then the true VaR per month per dollar invested is approximately 15 per cent higher than the postulated VaR of $1 - (0.9)^{1/12} \approx 0.874$ dollar cents.

ν_m	$\nu_e = \infty$								$\nu_e = \nu_m$							
	$r_m^{low} = -5\%$				$r_m^{low} = -10\%$				$r_m^{low} = -5\%$				$r_m^{low} = -10\%$			
	ν_t				ν_t				ν_t				ν_t			
	3	5	10	∞	3	5	10	∞	3	5	10	∞	3	5	10	∞
$\psi = 5\%$																
3	0	40	52	57	0	29	38	41	0	-22	-34	-44	0	-16	-25	-32
5	-34	0	10	14	-25	0	7	10	20	0	-11	-20	15	0	-8	-15
10	-42	-10	0	4	-31	-7	0	3	20	7	0	-6	15	5	0	-4
∞	-45	-13	-4	0	-33	-10	-3	0	-45	-13	-4	0	-33	-10	-3	0
$\psi = 1\%$																
3	0	-1	-14	-27	0	-1	-10	-20	0	-49	-73	-90	0	-35	-53	-66
5	1	0	-12	-26	1	0	-9	-19	59	0	-30	-52	43	0	-22	-38
10	15	13	0	-14	11	10	0	-10	90	30	0	-22	66	22	0	-16
∞	31	29	15	0	22	21	11	0	31	29	15	0	22	21	11	0

ν_m	$\nu_e = 5$							
	$r_m^{low} = -5\%$				$r_m^{low} = -10\%$			
	ν_t				ν_t			
	3	5	10	∞	3	5	10	∞
$\psi = 5\%$								
3	0	-18	-28	-36	0	-13	-20	-26
5	20	0	-11	-20	15	0	-8	-15
10	34	12	0	-9	25	9	0	-7
∞	47	23	10	0	34	17	7	0
$\psi = 1\%$								
3	0	-45	-68	-85	0	-33	-50	-62
5	59	0	-30	-52	43	0	-22	-38
10	106	35	0	-25	77	26	0	-19
∞	153	70	29	0	111	51	21	0

As supervisors require the reporting of high confidence level VaRs, i.e., $\psi = 1\%$, the remaining discussion of Table 1 is focused on the case $\psi = 1\%$. In that case, the direction of the bias in the VaR is the same for all estimators considered. The magnitude of the bias, however, differs quite substantially. As mentioned before, the least-squares estimator results in the best overall performance in this case. The robust MQL **t(5)** estimator has the worst overall performance, especially if the true degree of fat-tailedness is underestimated by the investment manager. In that case the optimal asset allocation can turn out far too risky compared to required VaR bounds. Vice versa, if the true degree of leptokurtosis is underestimated, allocation policies are too prudent (with the ML estimator performing slightly worse than the MQL **t(5)** estimator).

To conclude the discussion of Table 1, it is interesting to relate the findings for the ML estimator ($\nu_e = \nu_m$) to recent results obtained by Vlaar (1997). Using a context of interest rate risk management for banks, Vlaar finds that the VaR based on MQL **t(5)** estimates of the parameters in his model is violated far too often during the sample. Stated differently, the VaR is severely under-estimated using the given estimation method. By contrast, the use of a least-squares procedure, i.e., Gaussian MQL, results in much less violations of the estimated VaR measure. These results can be explained perfectly by looking at Table 1. As Vlaar, consider the ML estimator ($\nu_e = \nu_m$) for high levels of confidence ($\psi = 1\%$). If $\nu_t = 3$, the bias in VaR using $\nu_m = 5$ is almost twice as large as the bias for $\nu_m = \text{co}$. Vlaar remarks that the degrees of freedom parameter can fall far below 5 if it is estimated freely, suggesting that the $\nu_t = 3$ column could be reasonable for the data considered. Given this remark, the numbers in Table 1 help to explain both the violations of the VaR estimates and the ranking of VaR estimates based on the Gaussian and the **t(5)** quasi-likelihood for the data set used by Vlaar.

If Table 1 is extended using more combinations (ν_m, ν_t, ν_e), an overall performance measure can be constructed for ranking different estimation methods from a VaR perspective in the context of potential model misspecification. Table 2 presents the maximum absolute percentage mismatch for different values of ν_e, ψ , and r_m^{low} .

Using a minimax criterion, i.e., minimizing (over the set of estimators) the maximum mismatch per estimator, results in the choice of either the MQL **t(10)** or the least-squares estimator. For high levels of confidence ($\psi = 1\%$), the least-squares estimator is clearly the most robust under model misspecification from a VaR estimation perspective. Note that $\psi = 1\%$ is most relevant for practical purposes given the Basle (1996) proposals for banking supervision. Moreover, although the least-squares estimator for $\psi = 5\%$ is next to worst using the above minimax criterion, the differences across

Table 2

The table contains the maximum absolute percentage mismatch between the postulated value-at-risk (VaR) per dollar invested ($-r_m^{low}$) and the true VaR ($-r_t^{low}$), i.e., $\max_{\nu_m, \nu_t} |100 \cdot (r_t^{low}/r_m^{low} - 1)|$. Values are taken from an extended version of Table 1, where the values of ν_m , ν_t , and ν_e considered lie in the set $\{3, 4, \dots, 10\} \cup \{\infty\}$. See also the note to Table 1. Minimax entries per row are in boldface.

ψ	r_m^{low}	ν_e									
		ν_m	∞	10	9	8	7	6	5	4	3
5%	-5%	45	57	27	30	34	37	42	47	53	61
5%	-10%	33	41	20	22	24	27	30	34	39	45
1%	-5%	90	34	127	131	135	140	146	153	160	171
1%	-10%	66	25	93	95	99	102	106	111	117	124

estimators are less striking than for $\psi = 1\%$. The performance of the robust estimators for $\psi = 1\%$ is dramatic. This may seem strange at first sight, as these estimators are designed to be less sensitive to distributional misspecification, see, e.g., Hampel et al. (1986). The puzzle is solved by defining the parameters of interest. Although the robust estimators for μ and Ω may be less sensitive to model mis-specification, this does not necessarily hold for the VaR. To be more precise, consider (8). If $\nu_t \neq \nu_m$ for $\psi = 1\%$, the second term in parentheses is positive, while the first term is negative. Less bias in the estimate of Ω implies that the ratio c_t/c_m lies closer to zero, thus reducing the off-setting effect of the first term in (8) with respect to the second term. This explains the large biases in VaR for robust estimators of (μ, Ω) . The non-robustness of the least-squares estimator for Ω makes the off-setting effect of the first term in (8) much larger, thus making the least-squares estimator a robust estimator from a VaR perspective in this case.

4. Conclusion

In this paper I have considered the effect of mis-specified tail behavior on Value-at-Risk (VaR). It appeared that the least-squares estimator is most robust from a VaR estimation perspective in the framework considered in the paper. The maximum likelihood estimator and a robust estimator based on a Student t(5) quasi-likelihood, perform less satisfactory. Paradoxically, the robustness of the least-squares estimator from a VaR point of view follows from its non-robustness as an estimator for the dispersion of asset returns. The least-squares estimator compensates for the possibly mis-specified tail behavior by picking a larger dispersion measure, thus using an alternative way

to increase the probability on extreme returns. The compensating effect for the maximum likelihood estimator and the robust estimator is much smaller, resulting in larger biases for estimates of VaR.

References

- Basle Committee on Banking Supervision (1995): "Proposal to issue a supplement to the Basle capital accord to cover market risks," Report nr. 24, Bank for International Settlements, Basle. Available on the Web at <http://www.bis.org>.
- Campbell, J.Y., A.W. Lo, and A.C. MacKinlay (1997): *The Econometrics of Financial Markets*. Princeton: Princeton University Press.
- de Vries, C.G. (1994): "Stylized facts of nominal exchange rate returns," in *The Handbook of International Macroeconomics*, F. van der Ploeg (ed.), Cambridge: Blackwell, 348-389.
- Franses, P.H., and A. Lucas (1997): "Outlier robust cointegration analysis," VU Research Memorandum 1997-45, Free University Amsterdam. Available on the Web at <http://www.econ.vu.nl/vakgroep/bfs/alm/lucas/publications.html>.
- Gourikroux, C., A. Monfort, and A. Trognon (1984): "Pseudo maximum likelihood methods: theory," *Econometrica*, 52, 681-700.
- Hampel, F.R., E.M. Ronchetti, P.J. Rousseeuw, and W.A. Stahel (1986): *Robust Statistics. The Approach Based on Influence Functions*. New York: Wiley.
- Jorion, Ph. (1997): *Value at Risk*. Chicago: IRWIN.
- Leibowitz, M.L., and S. Kogelman (1991): "Asset allocation under shortfall constraints," *Journal of Portfolio Management* winter, 18-23.
- Leibowitz, M.L., S. Kogelman, and L.N. Bader (1992): "Asset performance and surplus control: a dual shortfall approach," *Journal of Portfolio Management* winter, 28-37.
- Lucas, A., and P. Klaassen (1996): "The effect of fat tails on optimal asset allocations," VU Research Memorandum 1996-47, Free University Amsterdam. Available on the Web at <http://www.econ.vu.nl/vakgroep/bfs/alm/lucas/publications.html>.
- Markowitz, H.M. (1959) *Portfolio Selection: Efficient Diversification of Investments*. New York: Wiley.
- Nelson, D.B. (1990): "Stationarity and persistence in the GARCH(1,1) model," *Econometric Theory*, 6, 318-334.
- Vlaar, P.J.G. (1997): "The effect of the dynamics in the term-structure of interest rates for bank risk management: an analysis of value-at-risk models," (in Dutch) Central Bank Research Report WO&E 512, De Nederlandsche Bank, Amsterdam.
- White, H. (1982): "Maximum likelihood estimation of misspecified models," *Econometrica*, 50, 1-25.