

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.

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 Gouriéroux et al. (1984). This class contains several well-known estimators such as the least-squares estimator, the maximum likelihood estimator,

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 .

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, $-r_t^{low} = -r_m^{low}$, 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:

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

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 $\nu_e = \nu_m$.

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 $\nu_m = \infty$. 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

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>.
- Gouriéroux, 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.