

The Hidden Risks of Optimizing Bond Portfolios under VaR

Peter Winker

Dietmar Maringer

June 2004

Abstract

Value at risk (VaR) has become a standard measure of portfolio risk over the last decade. It even became one of the corner stones in the Basel II accord about banks' equity requirements. Nevertheless, the practical application of the VaR concept suffers from two problems: how to estimate VaR and how to optimize a portfolio for a given level of VaR? For the first problem, several approaches have been suggested including the historical simulation method. The optimization problem can be tackled using recent advances in heuristic optimization algorithms. However, our application to bond portfolios shows that a solution to the two aforementioned problems gives raise to a third one: the actual VaR of bond portfolios optimized under a VaR constraint might exceed its nominal level to a large extent. Thus, optimizing bond portfolios under a VaR constraint might increase risk. This finding is of relevance not only for investors, but even more so for bank regulation authorities.

JEL: C15, C16, G11, G28

Keywords: VaR, risk, portfolio optimization, heuristic optimization

Peter Winker, Dietmar Maringer

Faculty of Economics, Law and Social Sciences, Nordhäuser Strasse 63, 99089 Erfurt,
Germany.

email: {peter.winker, dietmar.maringer}@uni-erfurt.de

The authors are grateful to S. Heng, M. Kalkbrenner, C. Kreuter and participants of conferences in Sydney and Neuchâtel for valuable comments.

1 Motivation

Value at risk (VaR) has become a standard measure of portfolio risk over the last decade. Given a portfolio with an initial value of V_0 , the VaR for a given probability α until time τ are the losses which will not be exceeded until τ with probability α . The success of this concept might be attributed to three causes. First, this risk measure is highly intuitive and closely related to investors' goals. Second, VaR does not depend on any specific assumptions about return distributions or risk aversion. Third, and this might have been the crucial factor, VaR has been imposed on banks and other financial institutions by the Basel II accord about banks' equity requirements.¹ Consequently, VaR can be considered as a standard instrument in assessing portfolio risk and credit risk.²

Unfortunately, the implementation of VaR is hampered by two major problems. First, though easy to interpret, it turns out to be at least as difficult to estimate as any other more "traditional" risk measure. Second, when used as a constraint in portfolio optimization, the resulting optimization problem cannot be dealt with using standard routines.

In order to deal with the first problem, three different approaches have been suggested and are used in practice:³ (i) using parametric models by assuming certain distributional properties of asset returns, (ii) historic simulation, i.e. using an empirical distribution of asset returns based on past returns, and (iii) Monte Carlo approaches which typically combine (i) and (ii). These three methods are also explicitly sanctioned by the rules imposed in the Basel II accord,⁴ i.e. banks are free to choose and implement one of the methods for assessing the VaR of their asset portfolios. The quality of the risk measure obtained through the different methods depends on the extent to which the underlying assumptions are satisfied. In particular, for (i) the distributional assumption (often normality) has to be met by the data, and for (ii) the joint distribu-

¹See Basel Committee on Banking Supervision (2003). For a critique of the Basel II proposals, see e.g., Danielsson et al. (2001).

²For an introduction to VaR and alternative risk measures for assessing credit risk, see, e.g., Sauners & Allen (2002).

³See, e.g., Jorion (2000).

⁴Basel Committee on Banking Supervision (2003), §490 (c) states: "No particular type of VaR model (e.g. variance-covariance, historical simulation, or Monte Carlo) is prescribed. However, the model used must be able to capture adequately all of the material risks exposure of the institution's equity portfolio."

tion of asset returns should be stable such that the past empirical distribution provides a reasonable approximation to the future distribution.

When returns are perfectly normally distributed the myopic portfolio optimization problem under a VaR constraint is equivalent to the one under a constraint on the portfolio's volatility.⁵ However, the distribution of asset returns is not well described by a normal distribution. In particular, fat tails and excess peakedness are typical features of asset returns difficult to deal with using parametric distributions. Hence, the use of empirical distributions is an accepted or even favoured alternative both in theory⁶ and practice.⁷ Nevertheless, we will also provide evidence on the stability of return distributions required for using empirical distributions in VaR measures. Furthermore, we will provide evidence to what extent the portfolio selection process is influenced by the method chosen to estimate risk.

The second problem results from the functional form of the risk constraint when using VaR in an optimization context.⁸ This problem arises naturally from requirements imposed by the Basel II accord on institutional investors such as banks. According to the minimum capital requirements, banks have to underlie each investment with a certain amount of equity depending on the risk class of the assets considered. Therefore, banks might want to (i) lock as little equity as possible, or (ii) construct a yield maximizing portfolio that is (just) possible given the existing equity with regard to the VaR limit. We focus on the second case assuming that banks' equity is fixed in the short run. Hence, utility based considerations whether slightly higher VaR might result in significantly higher returns or utility, respectively, are not relevant in this case as the VaR limit is regarded as binding in the short run.

Not least because of the Basel II accords, but also because of its popularity in the industry, VaR has attracted a considerable amount of theoretical research. It was found that VaR has some undesired properties as it is not what Artzner et al. (1999) call a *coherent risk measure*. In particular, the lack of sub-additivity is seen as a major shortcoming: depending on the included assets' distributions, a portfolio can display a higher VaR than its components. Hence, a risk reduction could be achieved when splitting the portfolio and assessing the individual securities – which contradicts the usual

⁵See De Giorgi (2002) and also the results in Maringer & Winker (2003).

⁶See, e.g. Jorion (2000), Pritsker (1997) or Lucas & Klaasen (1998).

⁷cf. footnote 4.

⁸In this context, see also Basak & Shapiro (2001) and Alexander & Baptista (2003).

principle of risk diversification that should be reflected by the risk measure.⁹ As a consequence, the displayed risk (in terms of VaR) could be lowered when unbundling portfolios or assessing a bank's divisions separately rather than as a whole. This and other shortcomings are meanwhile well-known and are already addressed in most textbooks on risk management and VaR;¹⁰ a more detailed presentation can therefore be left out in the sense of brevity.

As a consequence, a number of modifications on VaR and alternative risk measures related to VaR have been suggested. Arguably the most prominent amongst them is the Conditional Value-at-Risk (CVaR) or Expected Shortfall, which indicates the expected loss on the α worst days where the VaR limit is reached or exceeded. CVaR can be shown to have more desirable theoretical properties¹¹ yet is not as popular as the VaR for various reasons, and it is unlikely that it (or some other variant) will replace VaR in the near future. It therefore appears desirable to understand VaR and its influencing factors as good as possible. In this light and with the institutional requirement to use VaR as prime risk measure, this contribution will concentrate on VaR.

The portfolio optimization problem with a given constraint on the Value at Risk is formalized in Section 2.1. In contrast to the mean–variance approach, quadratic programming will not provide a solution in this case. Likewise, other standard optimization methodologies would demand either simplifications of the problem specification or are designed for a Expected Shortfall setting.¹² Nevertheless, recent advances in the application of optimization heuristics to portfolio optimization problems¹³ allow to tackle this problem efficiently. We make use of a modified version of Memetic Algorithms which is described in Section 2.3.

After providing a solution approach to the two problems of portfolio optimization under VaR, we realize a third problem with important implications for investors and bank regulation. When an investor is required to have enough equity to cover the loss in her investments that might occur within a given (short) period and with a given (low) probability, than she will have an incentive to construct a portfolio that maximizes the expected (utility of the) return that meets the shortfall constraints with the given amount of equity. Such a behaviour corresponds to the standard profit maximization

⁹See also Szegö (2002).

¹⁰See, e.g., Crouhy et al. (2001).

¹¹See Pflug (2000) on the coherence of VaR and CVaR.

¹²See, e.g., Rockafellar & Uryasev (2000) (extended in Krokmal et al. (2001)) or Uryasev (2000).

¹³See, e.g., Gilli & Küllezi (2002*b*) or Maringer & Winker (2003).

assumption if we assume that there are cost of equity. When choice is left to the portfolio manager / investor, she will have an incentive to choose the method that allows highest returns, especially when this method is supposedly more reliable than other methods and is used for ex post evaluation.¹⁴ In our setting this method turns out to be the historic simulation. While this method is useful to provide an ex post assessment of historic value at risk and also expected VaR for a given portfolio, it fails in a portfolio optimization setting. In fact, the optimization procedure results in high return portfolios just meeting the VaR constraint on the historic data. However, the actual VaR of these portfolios out of sample turns out to be much higher than its nominal level. In fact, optimizing portfolios under a VaR constraint typically results in portfolios with a VaR much higher than the defined constraint. This effect can be described as the hidden risk of optimizing portfolios under VaR. To our knowledge, this is the first paper to provide empirical evidence on this issue based on optimized portfolios.

The rest of this paper is organized as follows. Section 2 introduces the optimization problem, the data used for the empirical analysis, and the optimization heuristic to solve the complex portfolio optimization problem under VaR. In Section 3, we summarize the main findings. First, we provide some statistics on the distribution of asset returns and their stability, before turning to the actual VaR of the optimized portfolios. Section 4 concludes.

2 Model

2.1 The Optimization Problem

The investor for our problem has an initial endowment of V_0 that can be either invested in bonds or kept as cash; without loss of generality, the rate of return of the latter is assumed to be zero. Given that the losses until time τ must not exceed a (fixed) value of $\underline{\delta}^{VaR} \cdot V_0$ with a given probability of α , and that this VaR constraint is the only constraint, a manager of a bond portfolio will be inclined to find a combination that has maximum expected yield that does not violate this VaR constraint.

¹⁴cf. Basel Committee on Banking Supervision (2003) §§149 and 151 – 152.

The optimization model can therefore be written as

$$\max_{n_i} E(r_P) = \sum_i \frac{n_i \cdot L_i \cdot D_{i,0}}{V_0} \cdot r_i$$

s.t.

$$n_i \in \mathbb{N}_0^+ \forall i$$

$$\sum_i n_i \cdot L_i \cdot D_{i,0} \leq V_0$$

$$\text{prob} \left(V_\tau \leq V_0 \cdot \left(1 - \underline{\delta}^{VaR} \right) \right) = \alpha$$

where L_i and $D_{i,0}$ are lot size (in CHF) and current clean price (in per cent), respectively, of bond i , and r_i is its yield to maturity. n_i is the number of lots kept in the portfolio which has to be non-negative. Moreover, the cash position must also be non-negative. V_τ is the value of the portfolio at time τ (consisting of the value of the bonds including accrued interest from time 0 to τ) plus cash.

For estimating V_τ , we apply the following methods:

- Assuming *normal distribution*, the VaR constraint can be rewritten as

$$E(V_\tau) - u_\alpha \cdot \sigma_{V_\tau} \geq V_0 \cdot \left(1 - \underline{\delta}^{VaR} \right)$$

where u_α is the respective quantile of the standard normal distribution. The expected value for V_τ and its volatility are alternatively estimated from past observations either in a standard way (“plain vanilla” or “pv” henceforth) or with weighted values where more recent observations contribute stronger. The latter version turned out advantageous for stock portfolios in a similar setting¹⁵ with decay factor of 0.99 which is applied here, too. The weights are therefore $w_s = \frac{0.99^{(S+1)-s}}{\sum_{t=1}^S 0.99^t}$ where the simulations are ordered chronologically and $s = 1$ is the simulation based on the oldest, $s = S$ on the most recent of the S observations.

- Assuming *empirical distribution*, the VaR constraint can be rewritten as

$$\sum_{s=1}^S b_s \leq \alpha \quad \text{with } b_s = \begin{cases} \frac{1}{S} & \text{if } V_s \leq V_0 \cdot \left(1 - \underline{\delta}^{VaR} \right) \\ 0 & \text{otherwise} \end{cases}$$

¹⁵See Maringer (2004).

where $V_{s,\tau}$ is one out of S simulations for the wealth at time τ based on historic (in sample) observations. To parallel the weighted version of the normal distribution, the alternative is to weigh the b_s 's correspondingly:

$$\sum_{s=1}^S b_s \leq \alpha \quad \text{with } b_s = \begin{cases} w_s = \frac{0.99^{(S+1)-s}}{\sum_{t=1}^S 0.99^t} & \text{if } V_s \leq V_0 \cdot (1 - \underline{\delta}^{VaR}) \\ 0 & \text{otherwise} \end{cases}$$

where, again, simulation $s = S$ is based on the most recent observation.

For the main computational study presented in the following sections, the investor will be endowed with $V_0 = \text{CHF } 1,000,000$, and the VaR constraint will be that the next day's wealth will not be below 990,000 (i.e., $\underline{\delta}^{VaR} = 0.01$) with a probability of $\alpha = [0.025; 0.05; 0.1]$. These default probabilities α are higher than those usually applied in practice with respect to the available data: when T observations are available, then by definition $\tau = \alpha \cdot T$ observations constitute the losses at or below the VaR level. When α is lower, more observations ought to be available and hence the T ought to be higher to have a sufficient number of shortfall observations. When keeping $\tau = 5 (= 0.025 \cdot 200)$ as a lower bound (as we do in our study), shortfall probabilities of $\alpha = 0.01$ and 0.001 would demand $T = \frac{\tau}{\alpha} = 500$ and $5,000$ observations, respectively; with more assets in the portfolio, the problem is even intensified as higher τ (and therefore also T) ought to be used in order to avoid immediate data fitting. Usually (and for our data), the necessary (or desired) number of observations exceeds the number of actually observed data. This gap is often circumvented by generating artificial data. However, as we want to make a clear distinction between actually observed (empirical) data and generated data (based on some kind of assumption of their distribution), this study prefers higher values for α that are manageable with the available empirical data. In preliminary studies, longer data series were used and alternative values for α and $\underline{\delta}^{VaR}$ were investigated. The findings confirmed the qualitative results reported for the main study and are therefore omitted in the sense of brevity.

2.2 Data

The computational study is based on the fixed coupon bonds quoted on the Swiss stock exchange in local currency, i.e. CHF. From all quoted bonds, we chose randomly 42 Swiss and 113 foreign issuers, though it was sought that no industry sector or issued

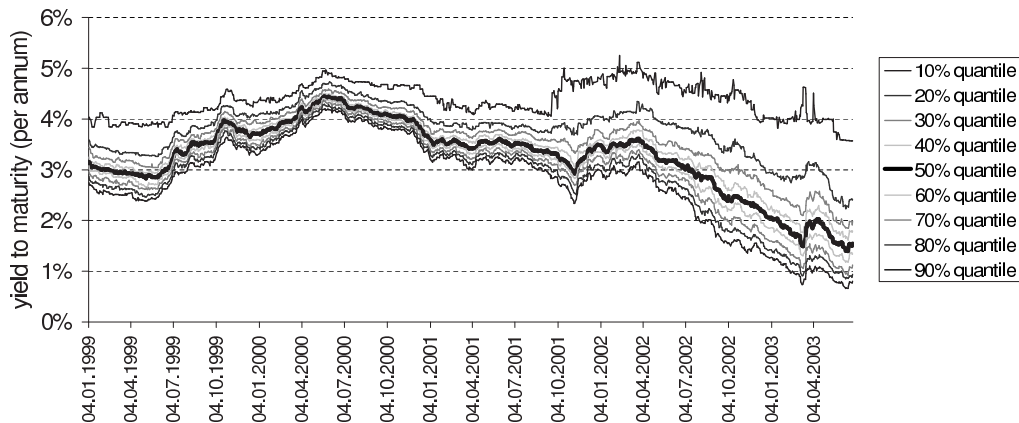


Figure 1: Quantiles for bonds' yield to maturity over time.

volume is over- or under represented. For these bonds, we have daily (clean) closing prices (when traded) for the period January 1999 through June 2003. All included bonds have a time to maturity of at least two years (typically five years) and the median issued amount is CHF 100,000,000 and CHF 200,000,000 for domestic and international bonds, respectively.

In particular for the earlier part of this time series, thin trading causes many missing data. When there was no current quote, the most recently quoted clean price plus exact accrued interest was used for current bond valuation from which the current yield to maturity was estimated. As can be seen from figure 1, the median yield to maturity of the bonds moved from about 3% per annum (in 1999) up to approximately 4.3% (in the year 2000) and later on down to approximately 1.4% per year (in 2003). As can also be seen from this figure, the top 10% of bonds with highest returns have yields of at least 4% per year. Calculating the statistics for yields of the individual bonds by considering the previous 200 trading days where available (which for most bonds is from December 1999 onwards), we find that the median for the standard deviations of the individual bonds' yields is 0.27%, and the lower and upper 10% percentiles for these individual standard deviations are 0.15% and 0.51%, respectively. The distributions of these individual yields are slightly left-skewed: the median skewness is -0.09 , the 10% percentiles are -0.87 and $+0.79$. Also, these yields have slightly negative excess kurtosis (median: -0.67 ; 10% percentiles: -1.36 and $+1.08$, respectively) which might partly be contributed to the (not uncommon) way of using recently paid clean prices when the value of a portfolio or asset has to be estimated yet no current prices

are available. It appears noteworthy, that the median excess kurtosis is positive from summer 2001 until 2002 and negative otherwise.

From this data set, random selections of bonds were drawn by first choosing a random date and then selecting $N = 10$ (20) different bonds. Any of these selections was accepted only if a minimum number of different quotes within the in sample as well as the out of sample time frames were observed (in sample frame: chosen date plus 200 in sample days; out of sample frame: the subsequent 100 trading days). For both values of N , 250 of such case sets were generated independently.

2.3 Optimization Method

Due to the type of the risk constraint combined with the integer constraint on the number of traded lots and the non-negativity constraint, the optimization problem cannot be solved analytically, but it can be approached with heuristic optimization techniques (HO). The recent literature holds several examples for successful applications of HO to portfolio optimization, including optimization under different risk measures (e.g., Dueck & Winker (1992)), cardinality constraints and integer constraints (e.g., Chang et al. (2000) or Maringer & Kellerer (2003)), index tracking (e.g. Gilli & K llezi (2002a)) or optimization under VaR constraints (e.g. Gilli & K llezi (2002b) or Maringer (2004)).

For our optimization problem, we use a modified version of *Memetic Algorithms*¹⁶ where principles of heuristic local search are combined with evolutionary search strategies. The basic idea of heuristic local search is to start with a random (and usually sub-optimal, yet valid) solution and suggest slight modification for this solution. These modifications are kept if they improve the current solution — or at least do not downgrade it beyond a certain threshold.¹⁷ During the first iterations, this threshold is rather generous (and therefore easily allows escaping local optima) but becomes more strict in due course (and therefore promotes hill-climbing search during the last iterations within a neighborhood that is supposedly close to the global optimum). Whereas traditional local search heuristics use just one agent that represents the current solution, Memetic Algorithms enhance this local search strategy by introducing a whole population of agents (each representing a solution) which merely perform independent local search, but which also “compete” and “cooperate” in regular intervals. In an

¹⁶See Moscato (1989).

¹⁷This concept is known as *Threshold Accepting* algorithm. See Winker (2001).

evolutionary fashion, these interactions are to ultimately eliminate inferior solutions (competition) and to combine different solutions (cooperation). The modified version of the Memetic Algorithm used here has proved useful and reliable in Maringer & Winker (2003) where stock portfolios are optimized under Value at Risk and where the algorithm and its characteristics are presented in more detail.

The implementation was done on two standard Pentium IV computers using Matlab 6. The different values for the shortfall probability α , the considered methods for estimating a portfolio's VaR, and the number of different case sets resulted in 6,000 different optimization problems for the main computational study. Each of these was solved repeatedly and independently, and the best found solution of any of the runs was used for the subsequent analyses. Depending on the problem size and distributional assumptions, the computational time ranged approximately from 10 to 20 seconds per run.

3 Results

3.1 Distribution

The decision of whether to estimate the VaR with the normal (or any other parametric) rather than the empirical distribution depends on how well the main properties of the observed data for the assets (or at least, via the CLT, the resulting portfolios) can be captured with the parametric distribution. For the given data set, the portfolio values appear far from normally distributed: regardless of the method for VaR estimation, there is hardly any optimized portfolio where the null hypothesis of normal distributed price changes cannot be rejected at the usual 5% level of significance both based on a standard Jarque-Bera test (as can be seen in Table 1) and the Kolmogorov-Smirnow test. Looking at the bond prices the null is rejected for virtually any of the assets in the data set — the details can therefore be omitted in the sense of brevity. The main reasons for the high rate of rejection are the leptokurtic and highly peaked distributions in the portfolios: even when taking into account that the higher moments do not necessarily exist (and therefore the Jarque-Bera test, using skewness and kurtosis, might not be appropriate) and calculating the Selector statistics¹⁸ the picture remains more or less unchanged (cf. Table 2).

¹⁸See Schmid & Trede (2003).

method	$N = 10, \alpha = \dots$			$N = 20, \alpha = \dots$		
	2.5%	5.0%	10.0%	2.5%	5.0%	10.0%
in sample						
empirical	4	3	3	4	4	1
empirical, weighted	4	2	2	3	2	2
normal	4	3	3	4	4	3
normal, weighted	4	3	4	3	3	3
out of sample						
empirical	24	26	22	25	23	21
empirical, weighted	26	22	26	23	25	24
normal	25	25	23	29	26	22
normal weighted	22	23	25	23	23	23

Table 1: Number of portfolios (out of 250): Do not reject H_0 : normal distribution, 5% level of significance

method	$N = 10, \alpha = \dots$			$N = 20, \alpha = \dots$		
	2.5%	5.0%	10.0%	2.5%	5.0%	10.0%
in sample						
empirical	6	3	3	8	0	0
empirical, weighted	3	2	3	4	0	0
normal	3	3	3	0	0	0
normal, weighted	3	3	3	0	1	0
out of sample						
empirical	13	10	10	19	14	11
empirical, weighted	15	11	9	15	16	16
normal	14	12	14	19	18	21
normal weighted	13	11	12	14	19	15

Table 2: Number of portfolios (out of 250): Do not reject H_0 : normal distribution, 5% level of significance based on the Selector Statistics test for leptokurtosis

At first sight, this seems to confirm the view that the normality assumption in the optimization process might be inadequate and that the use of empirical distributions might be the better choice: For most of the portfolios (see Table 3) and for an even higher share of the included assets, the hypothesis of same in and out of sample distributions cannot be rejected, hence using past realizations for estimates of future outcomes appears legitimate.

method	$N = 10, \alpha = \dots$			$N = 20, \alpha = \dots$		
	2.5%	5.0%	10.0%	2.5%	5.0%	10.0%
empirical	200	191	199	178	169	174
empirical, weighted	204	199	200	176	173	170
normal	199	195	195	175	176	175
normal, weighted	199	199	195	178	177	172

Table 3: Number of portfolios (out of 250) the H_0 : same i.s. and o.o.s. distribution (100 out of sample days) *cannot* be rejected at the 5% significance level

To test whether the distributions are stable and allow reliable estimates of the VaR, we repeatedly generated random weights for any portfolio in the two case sets where the integer and the budget constraints are the only restrictions. Then, the share of portfolios with out of sample losses higher than the expected VaR is determined. As can be seen from Table 4 for the first out of sample day, the use of the empirical distributions allows for estimations of the VaR such that the frequency of larger losses corresponds more or less to the respective confidence level. Under the normality assumption, higher values for α result in overly cautious estimations of the VaR — violations of which occur less often than expected. In particular for higher values of α , the empirical distribution produces more reliable results than the normal distribution. This relative advantage remains unaffected when longer out of sample periods are used for evaluation; the respective statistics are therefore omitted in the sense of brevity.

To test the statistical significance of the deviations between the expected shortfall probability, α , and the actually observed percentage of violations of the VaR limit as reported in Table 4, $\tilde{\alpha}$, a likelihood ratio test can be employed by computing the test

method	$N = 10, \alpha = \dots$			$N = 20, \alpha = \dots$		
	2.5%	5.0%	10.0%	2.5%	5.0%	10.0%
empirical	2.6 %	4.2%	9.1%	3.2%	5.6%	10.2%
empirical, weighted	2.4%	4.1%	8.2%	3.0%	5.4%	9.9%
normal	2.9%	4.1%	6.4%*	3.2%	5.3%	8.1%
normal, weighted	2.8%	4.0%	6.2%*	3.1%	5.1%	8.0%

Table 4: Percentage of portfolios with random asset weights exceeding the estimated VaR limit on the first out of sample day for the two case sets with a confidence level of α (* difference statistically significant at the 5% level)

statistics¹⁹

$$LR = -2 \cdot \ln \left(\frac{\alpha^x \cdot (1 - \alpha)^{(P-x)}}{\tilde{\alpha}^x \cdot (1 - \tilde{\alpha})^{(P-x)}} \right) \quad (1)$$

where P is the number of portfolios and $x = \tilde{\alpha} \cdot P$. The p value for this test statistic is then determined from a χ_1^2 distribution. As can be seen, statistically significant deviations between accepted shortfall probability and actually realized portion of shortfalls occur only for the portfolios with $N = 10$ and $\alpha = 10\%$. For all other instances, these deviations (though mostly larger for the normal than the empirical distribution) are statistically not significant.

3.2 The Hidden Risks in Optimized Portfolios

Unlike portfolios without optimization, the value of portfolios that are optimized under the empirical distribution will fall significantly more often below $V_0 \cdot E(\delta^{VaR})$, the expected VaR²⁰, than the chosen confidence level α . On the first out of sample day (Table 5) the actual frequency of excessive shortfalls will be 1.5 to three times the frequency originally expected (depending on α and case set). When the same portfolios are optimized under the normal distribution, however, the frequency will be underestimated only for small α 's, for high confidence levels, on the other hand, the frequency will be overestimated, i.e., the VaR is estimated too cautiously. The assumption of the

¹⁹See also Kupiec (1995).

²⁰Due to the specification and the chosen assets, the critical VaR, the out of sample data were compared to, is set to $E(\delta^{VaR}) \leq \underline{\delta}^{VaR}$, the loss actually expected with the planned probability of α .

normal distribution leads (for both optimized and random portfolios) to more cautious estimates of the VaR when α is high. The extreme leptokurtosis of the actual distributions cannot be captured by the normal distribution, and as a result it is hardly possible to get reliable estimates for the VaR limit: For large values of α , the VaR limit is estimated too far away from the expected value, for lower values, however, the actually realized values are within the bandwidth of their accepted value.²¹ Based on the LR statistics presented in equation (1), the empirical distribution virtually always leads to highly significant deviations between accepted shortfall probability, α , and actual percentage of shortfalls – whereas under the normal, virtually all of the actually realized shortfall probabilities are within the accepted range.

method	$N = 10, \alpha = \dots$			$N = 20, \alpha = \dots$		
	2.5%	5.0%	10.0%	2.5%	5.0%	10.0%
empirical	5.2%**	7.6%	16.8%***	8.8%***	10.8%***	16.9%***
empirical weighted	6.0%***	8.0%*	15.2%***	8.4%***	10.8%***	16.1%***
normal	3.2%	3.6%	7.2%	4.4%	6.8%	8.4%
normal weighted	3.2%	4.0%	6.0%**	3.6%	5.6%	6.8%

Table 5: Percentage of optimized portfolios exceeding the estimated VaR, $V_0 \cdot E(\delta^{VaR})$, on the first out of sample day for the two case sets with a confidence level of α (difference statistically significant at the 5% (*), 2.5% (**), and 1% (***) level)

The smaller α , the more only extreme outliers contribute to the shortfalls — the estimated frequencies for the first out of sample day are therefore more sensible to the chosen sample. Table 6 therefore takes into account larger out of sample periods, namely the first 50 and 100 out of sample trading days for the $N = 10$ and the $N = 20$ case sets, respectively. The basic conclusion from the first out of sample day that has been drawn for the “empirically” optimized portfolios, however, remains unchanged: the actual percentage of cases where the VaR is violated is significantly higher than the accepted level of α . For the optimization results under the normal distribution, the frequencies of shortfalls increase; resulting figures are closer to or below α when α is large, yet exceeding it when α is low. In the light of the results from the previous’ sec-

²¹For very small values of α the opposite can be observed: the VaR is underestimated, and the limit is violated too often. With respect to the data set, however, tests with smaller values of α than the ones presented were not possible, a more detailed discussion of these effects has therefore be left to future research.

tions, one can conclude that the specification errors of the normal distribution become more obvious in these cases. In short, longer out of sample periods reinforce the specification errors made by the normal distribution and increases the differences between accepted and actually realized shortfall frequencies. Nonetheless, the deviations under the normal are dramatically smaller than those caused by using the supposedly more accurate use of the empirical distribution.

method	$N = 10, \alpha = \dots$			$N = 20, \alpha = \dots$		
	2.5%	5.0%	10.0%	2.5%	5.0%	10.0%
$T_{oos} = 50$						
empirical	7.3%***	9.9%***	17.8%***	9.8%***	13.5%***	19.7%***
empirical weighted	7.2%***	9.8%***	16.4%***	9.3%***	12.9%***	18.8%***
normal	5.6%***	6.8%	8.9%	7.2%***	8.9%***	11.6%
normal weighted	5.3%**	6.4%	8.5%	6.8%***	8.4%**	11.0%
$T_{oos} = 100$						
empirical	7.8%***	10.6%***	18.6%***	10.4%***	14.0%***	20.1%***
empirical weighted	7.8%***	10.5%***	17.2%***	10.1%***	13.4%***	19.2%***
normal	6.2%***	7.5%	9.6%	8.0%***	9.6%***	12.1%
normal weighted	6.0%***	7.2%	9.3%	7.7%	9.2%***	11.6%

Table 6: Average percentage of the first T_{oos} out of sample days where the loss exceeds the expected VaR for optimized portfolios (differences statistically significant at the 2.5% (***) and 1% (***) level)

The advantage of the empirical over the normal distribution that had been identified for non-optimized portfolios and the statistical properties of the actual distribution, seems therefore lost and in some cases even reverted into the opposite when a VaR constraint is used in the optimization process. Despite its specification errors, the normal distribution seems to cause less problems than empirical distribution that has been shown to be closer to reality for the single assets and non-optimized portfolios.

The major reason for this is that VaR is a quantile risk measure and therefore focuses on the number of shortfalls rather than their magnitude.²² This can be exploited when empirical distributions are used. When optimizing under an empirical distribution, a number of excessive losses beyond the specified VaR limit will contribute equally to the confidence level α as would the same number of small losses; the optimization process will therefore “favor” those losses that come with high yields. Since

²²See also Artzner et al. (1999).

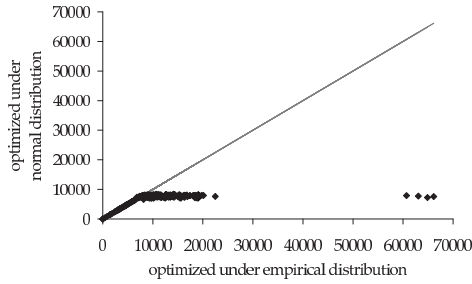
it is usually the high yield bonds that exhibit massive losses in the past, these bonds will be given high weights. The problems arising from this effect are reinforced when the high yield of a bond comes from a small number of high losses rather than several small losses: a loss beyond the specified VaR limit will be considered a rare event, and the loss limit estimated with the confidence level α will be distinctly below the accepted limit, i.e., $E(\delta^{VaR}) \ll \underline{\delta}^{VaR}$. Out of sample, this expected limit might turn out to be too optimistic and is therefore violated too often, hence the observed percentage of days with out of sample losses beyond the expected VaR is distinctly higher than the originally accepted level of α .

In addition, there is a hidden danger of data fitting for the empirical distribution: Slight in sample violations of the specified VaR limit of $\underline{\delta}^{VaR}$ can (and will) sometimes be avoided by slight changes in the combination of assets' weights that have only a minor effect on the portfolio yield. As a consequence, there might be more cases close to the specified VaR than the investor is aware of since they are just slightly above the limit and therefore do not count towards the level α ; out of sample, however, this hidden risk causes more shortfalls than expected.²³

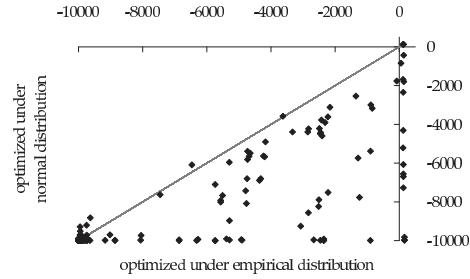
Both effects become more apparent from the scatter plots in Figure 2 where the results for portfolios optimized under empirical distributions are directly compared to the results when optimized under normal distribution. The magnitude of extreme losses shows up when the risk is measured in terms of volatility: "empirical" portfolios accept a standard deviation of up to CHF 20,000 and, on rare occasions, even more. When optimizing under the normality assumption, the definition of VaR imposes an implicit upper limit on the volatility of $\sigma_{V_\tau} \leq \frac{E(V_\tau) - V_0 \cdot (1 - \underline{\delta}^{VaR})}{u_\alpha}$, which, for $\alpha = 0.1$, is below CHF 10,000 for any portfolio in the case set. The volatility will be (approximately) the same regardless of the assumed distribution only if the "normal" portfolios have low volatility; when the optimal portfolios under normality actually make use of the specified risk limit (in sample), then their empirical counterparts are very likely to accept large variations in the respective portfolio's value (see Figure 2(a)).

Figure 2(b) demonstrates that, at the same time, there is a considerable number of portfolios which, when optimized under empirical distributions, are expected to have a smaller loss than one would expect for the same portfolio when optimized under

²³Because of the peakedness and the discussed effect, that $E(\delta^{VaR}) < \underline{\delta}^{VaR}$ for larger values of α , this effect of data fitting does not show as often as for assets with other empirical distributions (see Maringer (2004)).



(a) Standard deviation of next day's value of the portfolio



(b) Loss expected with a probability of α

Figure 2: Expected standard deviation and VaR of optimized portfolios for the $N = 20$ case set with $\alpha = 0.1$ (gray: 45° line)

normal distribution. When α is chosen rather large, the peakedness of the empirical distribution results in a VaR limit closer to the portfolio's expected value than predicted when the normal distribution is assumed: the rare, yet extreme in sample losses are perfectly ignored by the empirical distribution. If these extreme losses are rare enough, it might even happen that given a sufficiently large confidence level the estimated VaR limit will be a gain rather than a loss. This can be observed already for some portfolios in the $\alpha = 0.1$ case. Under the normal distribution, on the other hand, they do show up. Under empirical distributions, the investor will therefore be more inclined to accept extreme (in sample) losses without violating the risk constraint in sample, under the normal distribution, the investor will be more reluctant. This explains why a portfolio optimized under the empirical distribution will have a higher expected yield than the same portfolio optimized under the normality assumption. Figure 3 illustrates these differences. The larger the set of available assets, the more is the investor able to make use of this fact. Not surprisingly, the deviations between accepted α and actual percentage of out of sample shortfalls therefore increases, when N is larger, i.e., the investor has a larger set of alternatives to choose from (see Tables 5 and 6).

The consequences of these effects are twofold: First, the "empirical" optimizer underestimates the chances for exceeding the VaR limit since the scenarios where the limit is narrowly not exceeded in sample have a fair chance of exceeding it out of sample — hence the percentage of cases or days with losses beyond $E(\delta^{VaR}) \cdot V_0$ is higher than α , i.e., the expected percentage. Second, since the "empirical" optimizer does accept extreme losses in sample, she has a good chance of facing them out of sample as well. The "empirical" investor will therefore not only encounter losses

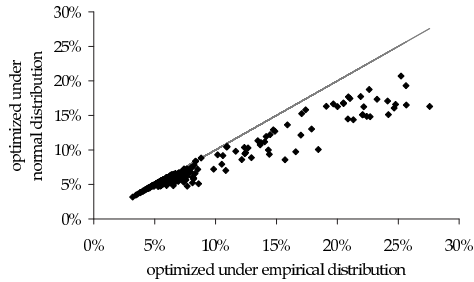


Figure 3: Expected portfolio yield depending on the distribution assumption for the $N = 20$ case set and $\alpha = 0.1$ (gray: 45° line)

exceeding the estimated VaR limit more frequently than the “normal” investor, the “empirical” investor’s losses will also be higher.

To what extent the deficiencies of empirical distribution are exploited in the optimization process depends on several aspects where the number of the in sample observations or simulations certainly is a very crucial one. Long time series, however, are not always available nor can they be reliably generated,²⁴ in addition the stability of the distribution becomes a major issue, and including more historic data might bring only diminishing contributions when weighted values (or alternative prediction models such as GARCH models) are used. Detailed tests of these aspects, however, were not possible with the available data and have therefore to be left to future research.

4 Conclusion

During the last years, Value at Risk has become an industry standard and has been imposed by the Basel II accords on equity requirements. Meanwhile, the literature has pointed out several shortfalls and theoretical caveats that are associated with the nature of this risk measure. This paper adds another aspect to this discussion: the pitfalls that might come when VaR is used not only for evaluation purposes of assets or portfolios, but already as an explicit constraint on the risk in the portfolio optimization process itself. Meant as answer to the problem of estimating the amount of capital that is at stake with a given confidence level, the suggested solution causes serious new problems.

²⁴The problem of small sample sizes becomes even more apparent when, e.g., credit portfolios are considered instead of publicly traded assets.

Our findings from an empirical study are that exact methods for estimating the risk (such as the use of empirical distributions) favor portfolios that actually have serious hidden risk by exploiting the nature and definition of the risk constraint; on the other side, inexact methods (such as the normality assumption) have less hidden risk, but are error prone because of their specification errors. Hence, neither approach appears capable of justifying VaR as a sole risk measure in the context of portfolio optimization.

The advantages of the VaR concept when evaluating a given portfolio *ex post* are diminished or even disappear when VaR replaces the “traditional” constraint in an optimization procedure. It is therefore noteworthy that these “hidden risks” are not the result from including securities with some fancy return distribution that obviously exploit VaR’s not being a coherent risk measure. The hidden risks identified in our computational study emerge only, when VaR with empirical distributions enters the decision process: While it can be confirmed that (in principle) empirical distributions are superior in measuring the VaR, it could be rejected that this would imply that empirical distributions would make a superior choice for the optimization process.

With VaR having become a key figure for assessing risk, this paper focuses on this risk measure and investigates whether there are additional potential shortcomings aside from already known problems due to its incoherence. However, the presented evidence brings up a series of new questions that ought to be investigated. One central issue is to identify under what circumstances the disadvantages of the (otherwise inferior) empirical distributions can be avoided within a Basel II framework. As this problem is far from straightforward, it cannot be answered reliably with our results so far and would be beyond the scope of this paper. Future research is therefore needed to find whether additional measures, including the inclusion of higher moments, additional risk constraints, and upper and lower bounds on asset weights, might be a remedy to some of the current shortcomings of VaR as a risk constraint.

References

- Alexander, G. J. & Baptista, A. M. (2003), ‘Portfolio performance evaluation using value at risk’, *The Journal of Portfolio Management* **29**(4), 93–102.
- Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1999), ‘Coherent measures of risk’, *Mathematical Finance* **9**(3), 203–228.

- Basak, S. & Shapiro, A. (2001), 'Value-at-risk-based risk management: Optimal policies and asset prices', *The Review of Financial Studies* **14**(2), 371–405.
- Basel Committee on Banking Supervision (2003), 'Consultative document: The new Basel capital accord, bank for international settlement', available from: www.bis.org.
- Chang, T.-J., Meade, N., Beasley, J. & Sharaiha, Y. (2000), 'Heuristics for cardinality constrained portfolio optimisation', *Computers and Operations Research* **27**(13), 1271–1302.
- Crouhy, M., Galai, D. & Mark, R. (2001), *Risk Management*, McGraw-Hill, New York et al.
- Daniélsson, J., Embrechts, P., Goodhart, C., Keating, C., Muennich, F., Renault, O. & Shin, H. S. (2001), An academic response to Basel II, Special Paper Series 130, LSE Financial Markets Group, London.
- De Giorgi, E. (2002), A note on portfolio selection under various risk measures, Working paper, Risklab, ETH Zürich.
- Dueck, G. & Winker, P. (1992), 'New concepts and algorithms for portfolio choice', *Applied Stochastic Models and Data Analysis* **8**, 159–178.
- Gilli, M. & Këllezi, E. (2002a), A global optimization heuristic for portfolio choice with VaR and Expected Shortfall, in E. Kontoghiorghes, B. Rustem & S. Siokos, eds, 'Computational Methods in Decision-making, Economics and Finance', Kluwer, pp. 165–181.
- Gilli, M. & Këllezi, E. (2002b), The threshold accepting heuristic for index tracking, in P. Pardalos & V. Tsitsiringos, eds, 'In Financial Engineering, E-Commerce, and Supply Chain', Applied Optimization Series, Kluwer, pp. 1–18.
- Jorion, P. (2000), *Value at Risk*, 2nd edn, McGraw-Hill, Chicago et al.
- Krokhmal, P., Palmquist, J. & Uryasev, S. (2001), Portfolio optimization with conditional value-at-risk objective and constraints, Working paper, University of Florida.
- Kupiec, P. (1995), 'Techniques for verifying the accuracy of risk measurement models', *Journal of Derivatives* **3**(2), 73–84.

- Lucas, A. & Klaasen, P. (1998), 'Extreme returns, downside risk, and optimal asset allocation', *The Journal of Portfolio Management* **25**(1), 71–79.
- Maringer, D. (2004), 'Distribution assumptions and risk constraints in portfolio optimization', *Computational Management Science* **forthcoming**.
- Maringer, D. & Kellerer, H. (2003), 'Optimization of cardinality constrained portfolios with a hybrid local search algorithm', *OR Spectrum* **25**(4), 481–495.
- Maringer, D. & Winker, P. (2003), Portfolio optimization under different risk constraints with modified memetic algorithms, Discussion Paper No. 2003-005E, Faculty of Economics, Law and Social Sciences, University of Erfurt.
- Moscato, P. (1989), On evolution, search, optimization, genetic algorithms and martial arts: Towards memetic algorithms, Report 790, Caltech Concurrent Computation Program.
- Pflug, G. (2000), Some remarks on the value-at-risk and the conditional value-at-risk, in S. Uryasev, ed., 'Probabilistic Constrained Optimization: Methodology and Applications', Kluwer, pp. 278–287.
- Pritsker, M. (1997), 'Evaluating value at risk methodologies: Accuracy versus computational time', *Journal of Financial Services Research* **12**(2/3), 201–242.
- Rockafellar, R. T. & Uryasev, S. (2000), 'Optimization of conditional value at risk', *Journal of Risk* **2**(3), 21–41.
- Sauers, A. & Allen, L., eds (2002), *Credit Risk Measurement: New Approaches to Value-at-Risk and Other Paradigms*, 2nd edn, Wiley, New York, NY.
- Schmid, F. & Trede, M. (2003), 'Simple tests for peakdness, fat tails and leptokurtosis based on quantiles', *Computational Statistics & Data Analysis* **43**(1), 1–12.
- Szegö, G. (2002), 'Measures of risk', *Journal of Banking & Finance* **26**(7), 1253–1272.
- Uryasev, S. (2000), 'Conditional value-at-risk: Optimization algorithms and applications', *Financial Engineering News* **14**.
- Winker, P. (2001), *Optimization Heuristics in Econometrics. Applications of Threshold Accepting*, John Wiley & Sons, Ltd., Chichester et al.