

Insch Quantrend

VaR: What is it really good for?

CHRISTOPHER CRUDEN and PURNER SCHNEIDER, INSCH
CAPITAL MANAGEMENT

December 2011

VaR: What is it really good for? We look for the answer by examining various value at risk (VaR) measures for a sample of 121 CTA programmes with track records running from the start of 2000 to July 2011. VaR is a well-known risk measurement tool, especially applied to measure past realizations of market risk.

Central in setting capital requirements for banks under the Basel Accord, VaR acts also as a risk management tool. However, is it a valid risk management tool? The answer lies in VaR's predictive power. How useful are VaR measures in predicting future performance rather than describing past performance?


VaR as a measure of market risk

VaR methods were invented in the late 1980s and flourished in the 1990s, boosted by events like the October 1987 crash and the LTCM failure. Accompanied by risk scenarios and stress-testing, VaR became the regulatory norm for banks, as well as a standard reporting measure for managed funds.

The last financial crisis proved VaR's weaknesses as a risk measurement and management tool, when the observed losses appeared to be millionth year events by the existing VaR measures. Subsequently VaR fell out of favour with practitioners. The initial success of VaR was occasioned by its relative ease of calculation and its conceptual intuitiveness.


A general definition of VaR is found in Philippe Jorion's FRM Handbook: "VaR is the maximum loss over a target horizon such that there is a low, prespecified probability that the actual loss will be larger." This is not a formula; just a definition. Myriad formulae have been proposed over time, guaranteeing that no precise measure of VaR exists. A brief description of the most popular methods is provided below.

One drawback of the VaR definition above is that it doesn't provide information on the behaviour of losses that exceed the VaR threshold. A solution to this problem is providing VaR accompanied by the expected tail loss or ETL. ETL is simply the expected value of the losses that fall below VaR over a certain period. While conceptually easy to grasp, ETL may be calculated easily or with great difficulty, depending on the VaR method applied.



COMMENTARY

High Frequency Trading
How the market developed and
where it is headed



COMMENTARY

Trading Hedge Fund Shares
Evolution of a now established
secondary market

A third and most informative way of using VaR consists in plotting the VaR surface, which involves representing the behaviour of VaR over a range of probability thresholds and holding periods. Other tools complement VaR: scenario analysis, stress-testing, reverse scenario analysis. One particular case is the Worst-Case Scenario, which answers the following question: Given that a worst period loss will occur with certainty, how bad will the loss be?

Not a coherent risk measure

VaR is not a coherent risk measure (coherence requires monotonicity, sub-additivity, homogeneity and translational invariance). We will spare you the details, other than mentioning that the lack of sub-additivity (in particular) implies that VaR does not always reflect portfolio diversification.

The methods of calculating VaR (and occasionally the ETL) can be classified by several criteria: parametric or non-parametric, static or dynamic, based on historical or simulated distributions. There are also hybrid methods, which combine parametric and non-parametric approaches. Parametric approaches are more powerful than the non-parametric ones, as they make use of the additional information contained in an assumed distribution function (normal, Student t, lognormal, extreme value, etc.). However, they are also the most vulnerable to error if the assumed function does not adequately fit the data.

VAR has additional shortcomings such as assuming risk-neutrality, by using objective rather than subjective probabilities. This means VAR does not account for investors' risk preferences. Some practitioners suggest using implied volatility measures instead of historical volatilities in estimating normal VAR.

The typical example of erroneous VaR application is the use of normal VaR to estimate the extreme expected losses of stock market returns – as empirical evidence holds that stock returns are leptokurtic, fat-tailed, and auto-correlated. As a result, normal VaR underestimates the severity of losses in individual securities.

Several approaches have been proposed to resolve the issues of fat tails and asymmetric distribution. One example is using the Student t distribution – which has fatter tails than the normal. Another example is the Cornish-Fisher modified VaR, which improves on normal VaR by modifying the alpha quantile of the Gaussian function to incorporate the impact of skewness and kurtosis. It is believed that the Cornish-Fisher approach works well for distributions that are not far from the normal.

Using mirrored returns

Adding mirrored returns to the historical distribution is a procedure that doubles the number of observations without changing the variance of the distribution. For this symmetric distribution, normal VaR may apply. By changing the skew of the distribution however, VaR calculated in this way can no longer be used for strategies that involve the management of skewness and of the downside tail.

Normal VaR (and variations : lognormal, Student t, etc.) are attractive because they produce straightforward formulas for both VaR and ETL, and require the estimation of few parameters: mean and standard deviation for the normal, degrees of freedom for the Student t distribution. Nevertheless, the normal distribution has an insurmountable drawback: it is based on the central limit theorem. According to this theorem, the distribution of sample means drawn from any distribution converges to the normal distribution as the number of samples increases.

At the tail, extreme value theorems apply. Two theories co-exist: Peaks Over Threshold, or Generalized Pareto Distribution (GPD) Approach, which describes the distribution of returns beyond a predefined threshold, and Extreme Value Distribution, which models the distribution of maxima and minima. Both of these theories rule out the normal, the lognormal and the Student t. However, the problem with these theories is that they require estimations of shape and scale parameters; the shape of the VaR surface changes considerably with these parameters.

The VaR obtained with the GPD approach is crucially influenced by the choice of threshold used in fitting the tail. There are several ways of choosing the appropriate threshold, such as the quintile-quintile plot. Extreme value theories are especially useful for pricing insurance contracts. Among non-parametric methods the most straightforward is historical VaR, which simply orders past returns and returns the fifth percentile lowest return.

Another non-parametric approach is the Kernel VaR, a generalization of historical VaR. The main step of the procedure is the construction of a smooth Kernel distribution. The Kernel technique assumes that each observation has some signal and some noise (each observation could have a point nearby). The procedure consists in placing blocks of optimal bandwidth at each observation, in order to extract the finer structure of the data, while minimising the overall noise. The advantage of Kernel VaR over historical VaR is its smoothness.

Powerful theorem

Chebyshev's inequality (attributed to his friend Irénée-Jules Bienaymé) is a very powerful theorem stating that no more than $1/(1+k^2)$ of the observations are more than k standard deviations below the mean - for any type of distribution. Chebyshev's inequality leads naturally to a 'Chebyshev' VaR estimate, useful as a sanity check for any other VaR results. The 5% VaR is obtained for $k = 4.36$. As always, there is a caveat also in using Chebyshev's inequality to determine VaR: in order to make inferences, we need to assume stationarity of the mean and variance of the distribution.

The most important shortcoming of VaR methods so far is their reliance on past performance. Forward-looking measures of VaR can be obtained through simulation techniques, stress-testing, worst-case scenarios, bootstrapping, or dynamic forecasting of volatility.

Monte Carlo simulation in general is based on generating a large number of samples of normal (or other) distributions, along with

correlations and other data dependencies. Historical simulation (bootstrapping) consists of generating scenarios by sampling historical returns associated with assumed risk factors. It has the disadvantage of failing to condition forecasts on the actual states of markets. Weighted historical simulation (e.g. exponential smoothing) reduces the weight of returns located farther in the past. Filtered historical simulation (Adesi et al., 2000) scales returns by their current conditional volatility forecasts, to ensure independent identically distributed returns.

Conditional volatility approaches come in various forms: GARCH, ARMA, CAVIAR, and so on. These methods incorporate various types of time-varying auto-regressive conditional volatility. Conditional volatility approaches accommodate empirical evidence in that:

- Monthly stock returns are negatively skewed and fat-tailed;
- Conditional means and volatilities are variable over times; and
- Returns are not independent and identically distributed.

Their drawback is the need to impose a parametric relation between variance and VaR. We just mention (without explaining) that many other approaches exist, such as mixtures of distributions, stochastic volatility models, stable Levy processes, elliptical and hyperbolic approaches, Box-Cox, Johnson, Generalized Error Distribution approach.

CTAs at work

If VaR does have predictive value, then it is useful in pricing derivatives, by use of delta, gamma and related approximations, or analytical and algorithmic methods for modelling non-linear risks. It is also useful in hedging downside risks, valuing portfolios, pricing insurance contracts, and rating managed investments.

Financial analysts read VaR with a more or less critical eye as an indicator of a manager's style and skill.

The financial crisis showed plainly the weaknesses of VaR in predicting extreme losses. Some researchers and practitioners claimed that VaR may still provide a helpful picture of everyday risk, that is, the risk taken in normal conditions. If true, VaR would still be something, as opposed to nothing.

As opposed to individual securities, characterized by negative skewness and excess kurtosis, managed return investments seek positive skewness and limited downside risk. In consequence, the most appropriate VaR methods may be different from the ones adequate in assessing the downside risk of equities.

The Insch Quantrend research on traditional equity and bond funds has shown that the performance of these funds is vulnerable to market conditions. CTAs differ from traditional funds by trading in a different universe, and by applying trading strategies aimed to reduce downside risk and increase the upside potential. Which method, if any, is the most informative of future performance, future drawdowns, or even future VaR, for CTAs?

To paraphrase Keynes, it's better to be roughly wrong than

precisely wrong. That is why we restrict our calculations to five straightforward measures of monthly 5% VaR:

1. Historical VaR
2. Kernel VaR
3. Normal VaR
4. Cornish-Fisher VaR
5. Chebyshev VaR

We gave up estimating the Generalized Pareto Distribution VaR because it is very sensitive to the choice of threshold and the threshold is a different one for each CTA manager. We would first need to locate the threshold that applies to each CTA and then use it to fit each CTA's tail distribution. We may do this in a future issue.

We examined the monthly return series of 121 CTA programmes, selected for their track record since January 2000 (at least). The data sources are Bloomberg and on-line CTA databases.

CTA programmes run at various leverage levels. To eliminate the effect of leverage, we scaled all the monthly returns by the programmes' standard deviations over the entire period. Therefore, in our results the programmes are geared down so that their monthly standard deviations equal 1%. This finding stands for empirical evidence that CTAs successfully manage downside risk, and achieve positive skewness and non-fat tails.

As expected, Chebyshev VaR is the most extreme VaR measure, at about four times standard deviation. This measure shows the fifth percentile loss for any distribution that the returns may have assumed during the periods analysed. Interestingly, the normal VaR average is situated below the historical VaR average at any time. The Cornish-Fisher VaR, obtained by applying a skewness and kurtosis correction to normal VaR is also at any point in time above the normal VaR.

Our survey found average skewness ranges from 0.18 and 0.35, while average excess kurtosis hovers around zero (like that of a normal distribution). Recall that these values are based on returns that are scaled down so that their monthly standard deviation equals 1%. Obtaining the true values requires scaling up the returns to reflect their real leverage.

So far, so good. Next, let's see how useful VaR is in predicting future:

- Average returns
- Standard deviations
- Largest drawdowns
- VaR

We are interested to learn whether past VaR values contain information regarding the future performance of CTA programmes. Therefore we make 43 rolling regressions for each VaR measure, to regress future three-year rolling statistics on past five years' VaR. The regressor and the regressand are not simultaneous. For example, at the end of June 2008, we read the realized VaR for

the last five years' observations of monthly returns, and check if it predicts average returns over the next three years, that is, from end of June 2008 to end of June 2011.

As mentioned above, average monthly returns, standard deviations, drawdowns, and VaR are calculated based on scaled returns, to eliminate the impact of leverage.

First, we note the stability of five-year rolling VaR by each measure. Unlike the case of stock markets, the VaR of CTAs appears to be unaffected by the financial crisis. The line for the Cornish-Fisher VaR is almost invisible in the graph, hidden below the historical VaR line. We may conclude that Cornish Fisher VaR is the best fit for historical VaR. The Kernel VaR is between Chebyshev and normal VaR, at about three times standard deviation. We consider it thus to be a conservative VaR measure.

VaR's predictive power

Three-year rolling average returns hovered close to 0.25% per month – scaled at 1% monthly standard deviation. The average was slightly lower during the financial crisis years. Past measures of VaR have a fractional yet positive impact on future returns in this estimation window. This implies that managers that had lower tail risk in absolute value in the past five years achieve better average returns over the future three years.

The estimates have very low standard errors, therefore they can be taken at face value. The possibility that this finding may be only a time frame pattern is a bigger worry than statistical significance. We cannot say for sure that the empirical linear dependence that we found can be generalized for other time periods.

Predictive power on standard deviation

Three-year monthly rolling standard deviation ranges from 0.8% to 1.1%, lower during the financial crisis and higher in the subsequent recovery. Recall that the returns were scaled by the whole period's average standard deviation, from January 2000 to July 2011.

Past measures of VaR also have a positive impact on future monthly standard deviation of returns. This implies that managers that had lower values at risk in absolute value in the past five years have higher realized standard deviations over the future three years. Could this mean that managers that had low tails risk in the past enhance their risk-taking in future? We refrain from drawing a conclusion. Again the estimates have very low standard errors, but they could also be a characteristic of the particular time frame.

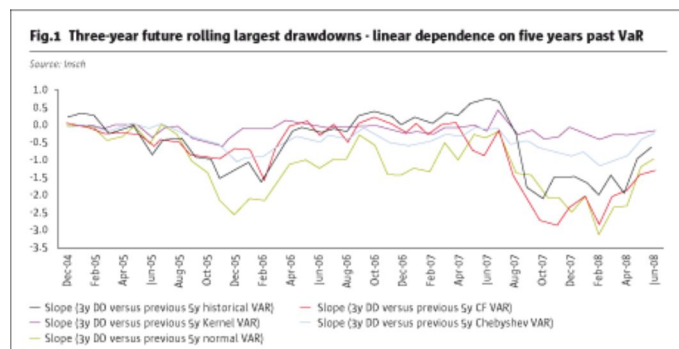
Predictive power on future drawdowns

Three-year rolling largest drawdowns are on average approximately double the average standard deviation.

The impact of past measures of VaR on future drawdowns is sometimes positive, but most times negative. Negative values imply that managers with past large VaRs in absolute value have less severe drawdowns in the next three years. This is somewhat counter-intuitive, and again it may be a time-frame characteristic.

Predictive power on future realized VaR

Three-year historical VaR ranges from one time to 1.5 times standard deviation, similar to the five-year historical VaR. Interestingly, VaR was lower in absolute value during the financial crisis years. We also checked the impact of past measures of VaR on future realized historical VaR. The relationship is ambiguous, positive in the first years and negative afterwards. Maybe the financial crisis has brought a behaviour shift? We judge the finding to be inconclusive.



In the end, we are not convinced by the predictive power of VaR. We are not

convinced either whether more sophisticated VaR measures would yield better results. We cannot conclude that VaR is useful as a risk management tool. However, in the process of estimating various measures of VaR we gained understanding on the style and main performance features of CTAs, such as positive skewness, no excess kurtosis, and VaR measures that are relatively stable over time and are not affected by stock market distresses.

Conclusion

In spite of its deficiencies, VaR is one of the standard risk measures, due mainly to its conceptual simplicity and regulatory importance. According to some, VaR works only for liquid securities over short periods in normal markets; it goes without saying that it does not predict catastrophic outcomes, unless these are somehow incorporated in the distributional assumptions.

We state once more the possible ways to exploit VaR as a risk management tool: if VaR has any predictive power on performance, its utility is obvious for investment management. If only the tail behaviour can be predicted, then VaR application is still useful in hedging tail risk and in pricing derivatives and insurance contracts.

If there is no predictive power in any of the VaR methods, then VaR is nothing but a past performance descriptor. In that case VaR is hopeless in signalling any danger to come – be it catastrophic or everyday risk.

We do not dismiss the VaR utility completely. There is a lot to learn in the process of constructing VaR. For instance, the failure of VaR for equities, equity indices and funds during the most recent financial crisis was quite enlightening on how serious the market conditions were.

The topic of our research is timely, as both financial markets are now crossing one of the worst months since the collapse of Lehman Brothers. Our Risk to Revenue analysis of traditional funds shows that these funds' losses closely follow stock markets.

Consensus is that market losses cannot be modelled by the simple and static VaR measures that we applied in this report, as distributions during periods of tail losses are not normal and non-stationary. Nevertheless, while traditional funds suffered abnormal losses and even became inactive during the financial crisis, the established CTAs in our sample continued to post normal-times performance and normal-times tail losses through the management of downside risk. While losses were still realized, they were not driven by a significant shift in the distribution of returns.

CTA managers use hedging strategies and instruments to protect against extreme losses. That is why basic VaR approaches are not completely useless measures in assessing the tail risk, as they are in assessing the extreme downside risk of individual securities or traditional funds. VaR may or may not be useful to predict the performance of CTAs – we have found weak evidence of its predictive power.

Further, we realise the caveats of our empirical analysis: on the one hand, our results rely on the reliability of self-reported CTA results. On the other hand, our screening of established CTAs with a long track record may lead to survivorship bias (as do all the widely employed indices). Our research opens expansion tracks like enlarging the universe of CTAs, correcting for survivorship bias and replicating the analysis for different time frames, and not in the least, applying the VaR analysis to traditional funds and hedge funds.

Christopher L. Cruden is CEO of *Insch Capital Management AG* and *Purnur Schneider, FRM*, is Associate Director of *InschQuantrend Ltd*.