ABSTRACT

Any model is a simplified version of reality, and thus always contains the possibility that the simplifying assumptions do not well match, or properly account for, the true governing processes of the world. We refer to model risk as the risk of not accurately estimating the probability of future losses due to a failure of a risk model. Sources of model risk in risk measurement models include differences between the assumed and actual distribution and errors in the logical framework of the model. Thus, successfully estimating the risk to one’s portfolio requires not only accurate estimates for the inputs to the model, but also a detailed understanding of the distributions from which the risky outcomes arise. Understanding how a model might fail is the first step towards rectifying the problem of model risk. We demonstrate a measurement of this model risk associated with the problem of parameter uncertainty in a value-at-risk model. Oftentimes the parameters of a distribution may be estimated with low precision, or there may be disagreement about the governing processes on which to base a risk model. Existing standard risk models do not adequately handle this parameter uncertainty. We show how these traditional methods for handling parameter uncertainty often fail, and provide a technique for quantifying risk more accurately.
INTRODUCTION

Model risk comes primarily from two possible shortcomings in the model development process. These shortcomings come out in phrases such as “models are only as good as the assumptions they are based on”; and “models are only as good as the data behind them.” Commonly, these risks express themselves as misspecified distributions and parameter uncertainty. Consider the well-known Black-Scholes options pricing model, one of the most successful pricing models in finance. Even so, it falls prey to both of these problems. First it assumes a normal distribution on asset returns. Thus in a world filled with fat tails and volatility smiles, a naïve Black-Scholes user will underpay for all options, misprice at-the-money compared with out-of-the-money options, and quickly go broke. Second, the price generated from a Black-Scholes equation is only as valid as the estimate used for the asset volatility. This crucial parameter must be estimated from historical volatility or forecasted from a volatility model. The uncertainty behind this volatility estimate can explode on those with too much faith in their model outputs.

Furthermore, tertiary risks abound. These can be as basic as implementation risk (say, a trader uses Black-Scholes to value an American put) to regime shifts (volatility smiles suddenly appearing) to unknown or unanticipated risks (such as worsening liquidity placing stress upon heavily leveraged arbitrageurs such as Long Term Capital Management).

In this chapter we focus on the model risk contained in value-at-risk (VaR) models, and primarily on the risk of parameter uncertainty. We build upon the ideas discussed in Hsu and Kalesnik (2009) and demonstrate that properly accounting for parameter uncertainty can result in posterior distributions that do a superior job of capturing return characteristics without resorting to exotic distributions that are difficult to work with and may not properly describe asset return behavior in any case.

The issues of misspecified distributions have received tremendous attention in the last decade since the collapse of Long Term Capital Management. In particular, the distributions used in standard VaR analysis do not adequately capture the frequency of extreme shocks to asset prices (kurtosis) or the size of those shocks (negative skew). This topic even went mainstream with the publication of Nassim Taleb’s 2007 book, The Black Swan: The Impact of the Highly Improbable, which discussed in detail the manifestations of fat tails. The difficulty of addressing these problems in risk management has received further attention from authors such as Derman (1996), Hendricks (1996), and Nocera (2009).
More advanced practitioners of VaR methods have begun to employ more esoteric distributional assumptions around asset returns, modeling them with Levy, Cauchy, or other fat-tailed stable Paretian distributions in an effort to capture this kurtosis and a negative skew. Lucas (2000) is but one of the practitioners who have published journal articles discussing these techniques. However, when VaR estimates are produced using more esoteric distributional assumptions, the process loses much of its intuitive appeal. The parameter choice for these distributions can be complicated and difficult to agree upon. It becomes increasingly difficult to express committee views in the parameter choice. Additionally, these models fall prey to a bias towards certainty inherent in the modeling process: as the level of complexity of the model grows, almost always our estimation of the remaining uncertainty shrinks more quickly than the uncertainty itself. This is particularly true when the output of the model is so incredibly simple—a single dollar figure, expressing the maximum likely loss.

**UNCERTAINTY AROUND THE MEAN ESTIMATE**

Our approach can be illustrated with an example of disagreement around something as simple as mean return. Suppose a five-member investment committee is attempting to measure VaR on their equities index portfolio at the beginning of 2009. Given the tremendous volatility through the bear market of 2008, members of the committee could hold widely differing views on possible returns for 2009. Three members of our committee believe that the unprecedented monetary stimulus and likely fiscal stimulus package will, over the course of the year, right the economy and rescue the United States from a prolonged recession. They believe that, given the sharply discounted valuation ratios, expected returns for 2009 will come in at 20 percent. On the other hand, two members of the committee believe that the worst is yet to come. They see the economy entering into a multi-year recession with no improvement on the horizon, and see room for valuation levels to fall even further. The bearish group believes that equities will continue downwards with a substantially negative return of −20 percent. To illustrate the point, both sides will agree that volatility will average 15 percent (later we will look at volatility uncertainty).

Given this agreement, how should the committee model the market risk faced by their portfolio? We will consider four different methods to quantify the disparate views amongst the committee members. We will then
utilize these assumptions in a log-normal distribution to determine the VaR for the portfolio. In the following notation, \( N(\mu, \sigma) \) is a normal distribution with mean and standard deviation of \((\mu, \sigma)\).

1. We could use the majority opinion number of +20 percent as our mean estimate. Our distribution is: \( \ln r_1 \sim N(20\%, 15\%) \).

2. We could use the mean estimate of the committee members for our expected return. This gives a distribution of: \( \ln r_2 \sim N(4\%, 15\%) \).

3. To protect against a worst-case scenario, the committee could decide to assume an expected return of \(-20\%\) and use \( \ln r_3 \sim N(-20\%, 15\%) \).

4. Finally, the committee could explicitly model the uncertainty in the expected mean return. This gives a return distribution of:

\[
\ln r_4 \sim \begin{cases} 
\text{prob} = \frac{2}{5}, & N(-20\%, 15\%) \\
\text{prob} = \frac{3}{5}, & N(-20\%, 15\%) 
\end{cases}
\]

We plot the four ex ante distribution functions in Figure 25.1. Note how explicitly modeling the mean uncertainty captures the bimodal views of the group while assigning a far lower probability to the mean outcome (which no individual expects to occur) than using the second scenario. From the

**Figure 25.1 Probability Density Function for Equity Returns with Uncertainty of the Mean**
distribution we can also see that scenario 4 results in a distribution with large negative kurtosis, even though the individual distributions had kurtosis of zero before blending them together. Thus the mean uncertainty results in thinner tails rather than the fat tails modeled by Cauchy or Levy distributions. However, we can see that the parameter uncertainty leads to decidedly non-log-normal distributions, even when the starting distributions are themselves log-normal.

The risk statistics presented in Table 25.1 give further insight into these distributions. The first four columns characterize the ex ante distributions. The fifth column shows the VaR at the 5 percent confidence level for the equity portfolio. The sixth column shows the expected loss conditional on observations in the lowest 5 percent tail of the distribution. And the seventh column shows the results of an investment decision rule: if we wish to cap our expected losses at $−25$ percent with a 5 percent probability, this column shows the maximum percentage of our assets to be invested in equities, with the remainder in a zero-return cash fund.

The differences in the first three distributions all come as a result of parameter selection. The first is clearly suboptimal because it ignores information; the majority rule approach fails to take into account the range of possible outcomes and will underestimate the VaR and often be overinvested. Scenario 2 shows that using the mean estimate delivers an improved risk assessment—the 5 percent VaR level and expected shortfall both increase dramatically. However, given this setup, the investment allocation decision does not change. Scenario 2 is still a naïve approach, and underestimates both the VaR and the expected shortfall. Note that both scenario 2 and scenario 4 have the same expected mean, but in scenario 4 the resulting standard deviation becomes considerably larger by modeling the mean uncertainty.

Although the worst-case assumption shown in scenario 3 shows the most similar risk characteristics to the mean uncertainty model, it is still suboptimal

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Mean</th>
<th>Volatility</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>5% VaR</th>
<th>Expected Shortfall</th>
<th>Max % Invested (5% chance of 25% loss)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20.00%</td>
<td>15.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>4.59%</td>
<td>10.42%</td>
<td>100.00%</td>
</tr>
<tr>
<td>2</td>
<td>4.00%</td>
<td>15.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>18.70%</td>
<td>23.70%</td>
<td>100.00%</td>
</tr>
<tr>
<td>3</td>
<td>-20.00%</td>
<td>15.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>36.05%</td>
<td>39.95%</td>
<td>69.36%</td>
</tr>
<tr>
<td>4</td>
<td>4.00%</td>
<td>24.68%</td>
<td>-20.44%</td>
<td>-72.89%</td>
<td>31.13%</td>
<td>36.11%</td>
<td>80.30%</td>
</tr>
</tbody>
</table>
because it results in a significant under allocation to equities. This 11 percent under allocation costs the portfolio 44 basis points in annual expected return.

The table also shows numerically what was strikingly depicted in the figure: the mean uncertainty approach leads to negative skewness and large negative kurtosis. However, we can also see that the negative kurtosis is dominated by the increase in the volatility from 15 to 24 percent. This suggests that exotic explicit modeling of fat tails may not be as important as properly modeling parameter uncertainty.

**UNCERTAINTY AROUND THE VARIANCE ESTIMATE**

Perhaps we have a different situation, where a group of investors may agree on the mean return but differ as to expected future volatility. In this section we illustrate how modeling uncertainty in our standard deviation estimates can improve VaR forecasts. We take the same approach as before, with a five-member investment committee. This time they all agree on an expected annual return of 10 percent, but three members expect a relatively quiet market with volatility of 12 percent while two members forecast a much higher volatility of 25 percent. Again we compare four approaches to modeling VaR in this situation. In the first, the majority rule decision would model volatility at 12 percent; in the second, we average the individual votes and utilize a volatility of 17.25 percent; in the third, we apply the worst-case scenario of 25 percent; and the fourth directly models the uncertainty in the variance parameter estimate in the same manner as that used in the previous section for the mean return.

Figure 25.2 shows the ex ante probability density functions for each of the four scenarios, and Table 25.2 shows the risk characteristics. As expected, scenario 1 and scenario 3 provide the least and most, respectively, conservative risk assessments. The interesting comparison is between the use of the average forecast, scenario 2, and directly modeling the uncertainty in scenario 4. We see that scenario 4 provides a higher standard deviation than that in scenario 2. However, the difference between these two is nowhere near as large as the earlier case where we modeled uncertainty around the mean expected return (an increase in volatility from 15 to 24.6 percent, compared with a rather mild increase from 17.2 to 18.3 percent). Also, uncertainty in the variance estimate produces high positive kurtosis instead of the negative kurtosis produced by uncertainty in the mean estimate. This transforms the starting log-normal distributions into a fat-tailed distribution, and it is this increase in kurtosis that drives the major differences in risk assessment between scenario 2 and scenario 4.
We next conduct a historical test of how explicitly modeling uncertainty in the mean return would have impacted VaR estimates. Our test is based on a setup similar to our example of mean uncertainty. We follow an investment committee of five members who estimate VaR for their investment in the S&P 500 index from 1987 through June 2009. The committee is divided into two camps. Three members are believers in mean reversion; they estimate the future return for their equity investment based on a comparison of the past 36-month returns to a long-run average equity premium. The second group of two members believe in trend following, and base their forecasted return on the average past three-month return. At times these expectations will be very similar, and at times they will differ dramatically.
This gives us a chance to demonstrate how the amount of disagreement affects the quality of VaR calculations. For simplicity, both groups agree to estimate the expected volatility by using the historical three-year average. We should note that neither of these two strategies are being proposed as a “best-implementation” type of strategy. Rather, they have been chosen as contrasting forecast models and their ability to demonstrate how the level of disagreement in mean return estimates can influence the performance of VaR models.

For each month in our time series, we calculate four estimates of 5 percent VaR: an estimate based on the mean reversion forecast, an estimate based on the trend following forecast, an estimate based on the weighted average forecast, and an estimate based on our technique of explicitly modeling the uncertainty in the expected return. We then judge the quality of the forecasts based on the frequency of violations.

In Figure 25.3, we see the history of the four VaR estimates based on actual market returns. Although the forecasts vary quite a bit through time, each of the models delivers a very similar average forecast of between −6.25 and −6.62 percent. The trending forecast, being based on a much shorter amount of data, produces much more volatile estimates. Yet all four measures are highly correlated and rather slow moving through time.

Table 25.3 displays these model performance numbers. Across all months in our sample, we see that both the mean reversion forecast and the trending forecast have historically underestimated VaR. Their resulting violations of
7.78 and 8.89 percent significantly exceed the 5 percent target. However, the interesting comparison is between the third estimate, taking the average forecast, and the fourth, which involves directly modeling the mean uncertainty. Both of these measures still underestimate the 5 percent VaR bound, but perform considerably better than either of the individual forecasts. While the average estimate produces a 6.67 percent violation frequency, the mean uncertainty model results in a 5.56 percent violation—still exceeding, but much closer to, the 5 percent target and a full percentage point better than the average estimate.

The second part of Table 25.3 demonstrates how the strength of disagreement in the forecasts influences the quality of the VaR estimates. We rank each month based on the absolute difference between the mean reversion estimate and the trending estimate. We then split the sample into two, with those in the half with the greatest difference considered “High Disagreement” periods and those in the second half labeled “High Agreement” periods.

When our mean reversion and trending models have relatively similar forecasts, little can be gained from modeling this disagreement. We see that the mean reversion estimate slightly underestimates, and the trending forecast overestimates, the VaR figures. And the average forecast and mean uncertainty forecast are equally good (and quite good in this historical test) at 5.15 percent. However, the months where there is a high disagreement in the forecasts show a different result. The two naïve forecasts both perform poorly at 11.19 percent violations, and the average forecast produces an 8.21 percent violation rate. However, the mean uncertainty model still scores quite well with a 5.97 percent violation rate—very good, considering the alternative.

Note that the periods of agreement tend to be quiet market periods of lower volatility and higher returns. In these circumstances, the penalty for a poor VaR estimate is generally relatively light. It is in the periods of high disagreement, characterized by higher volatility, greater negative returns, and turning points in the market, that the penalties for underestimating VaR

<table>
<thead>
<tr>
<th></th>
<th>Mean Reversion</th>
<th>Trending</th>
<th>Average</th>
<th>Mean Uncertainty</th>
<th>Expected Shortfall</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Months</td>
<td>7.78%</td>
<td>8.89%</td>
<td>6.67%</td>
<td>5.56%</td>
<td>−8.33%</td>
</tr>
<tr>
<td>High Disagreement</td>
<td>11.19%</td>
<td>11.19%</td>
<td>8.21%</td>
<td>5.97%</td>
<td>−9.15%</td>
</tr>
<tr>
<td>High Agreement</td>
<td>4.41%</td>
<td>6.62%</td>
<td>5.15%</td>
<td>5.15%</td>
<td>−7.52%</td>
</tr>
</tbody>
</table>
are at their greatest. These periods include the market crash in October 1987, the entire 12 months from June 2008 to May 2009, and much of the crash in the technology bubble of 2000 to 2001. These are the times when we would most like to have reliable estimates, and these are exactly the times when the relative performance of the mean uncertainty model is at its best.

The final column of Table 25.3 shows the expected shortfalls estimated by the mean uncertainty model. It predicts an average shortfall of −8.33 percent across all months, with a significantly greater shortfall of −9.15 percent in the High Disagreement months compared to an expected shortfall of −7.52 percent in the High Agreement months. This also shows the greater value of precise VaR estimates during times of higher market uncertainty.

CONCLUSION

We have presented a way to directly model uncertainty in key parameters commonly used in value-at-risk models. Such disagreements are commonplace in financial markets at large as well as within organizations. We show that by directly incorporating diverse views into our risk models, the resulting VaR estimates exhibit superior performance characteristics to other commonly used techniques. Not only do the resulting ex ante distributions better capture the assumptions underlying the investment, they also result in more accurate VaR predictions in our test case for the S&P 500 returns. Particularly when the level of disagreement is high, leading to high levels of parameter uncertainty, we see that directly modeling this parameter uncertainty performs exceptionally well.

Standard risk management approaches fail to consider parameter uncertainty, which has led to improper risk management. Blind faith in parameter estimates has too often led to blind faith in the resulting VaR outputs, and when these estimates are too often exceeded, the proposed solution is commonly to fatten up the tails by using exotic distributions. We show, however, that directly modeling the uncertainty in mean and variance returns using standard log-normal distributions can result in posterior distributions with high degrees of skewness and kurtosis. If we accept a simple world of time-varying expected returns and variances, the resulting uncertainty around these constantly shifting parameters places us squarely in this world of interesting and effective posterior distributions.

REFERENCES


NOTES

1. The amount of expected mean reversion is parameterized so that both groups have, on average, the same expected return for their equity investment.

[AQ1]: Vitali Kalesnik added as 2nd author in manuscript