VaR Exceedances at Large Financial Institutions: 2002–2014

Research Paper 007

March 5, 2015

info@northstarrisk.com

VaR Exceedances at Large Financial Institutions: 2002–2014

Large financial institutions include information about their risk models in their regulatory filings. The data reported by these institutions raises serious question about the accuracy of their risk models.

Introduction

Value at risk (VaR) is one of the most widely used risk statistics in finance. VaR is used by both large and small financial institutions. As a risk statistic, VaR is far from perfect, but it has a lot to recommend it. The concept of VaR is easy to understand. VaR can be applied to different asset classes and security types. It allows us to boil down the risk of an entire portfolio, no matter how complex, to a single number. Importantly for the purposes of this paper, VaR is easy to backtest.

Large financial institutions are required to report VaR for their trading operations by both national and international regulators. In this paper, we focus on VaR statistics reported in United States Securities and Exchange Commission (SEC) filings. The precise methodology varies from firm to firm, but most firms report 1-day VaR at either the 95% or 99% confidence interval. In recent years, most firms have also included a backtesting summary, which indicates how accurate their VaR models were in the most recent period.

In this paper, we include data from Bank of America, Bear Stearns, Citigroup, Credit Suisse, Goldman Sachs, JPMorgan Chase, Lehman Brothers, and Morgan Stanley. In their regulatory filings these firms make clear that VaR is not the only statistic that they use to manage the risk of their portfolios. They rightfully point out a number of well-known shortcomings of VaR. While VaR is far from perfect, their backtest results are often far outside the expected range. This was true before, during and after the financial crisis. We suggest a number of reasons why firms might report biased or inaccurate numbers in their regulatory filings. Comments made in their filings suggest that some of these firms are still using very simple models and there may be some confusion about how VaR models should be backtested.

Review of VaR Backtesting

If the 95% VaR of a portfolio is $400, then we expect the portfolio will lose $400 or less in 95% of scenarios, and lose more than $400 in 5% of scenarios. We can define VaR for any confidence level, but 95% and 99% have become extremely popular choices in practice. The time horizon also needs to be specified for VaR. On trading desks, with liquid portfolios, it is common to measure VaR over one-day.

If an actual loss equals or exceeds the predicted VaR threshold, that event is known as an exceedance. Another way to explain VaR is to say that for a one-day 95% VaR, the probability of an exceedance event on any given day is 5%, 5% = 1 − 95%.

info@northstarrisk.com

Exhibit 1 provides a graphical representation of VaR at the 95% confidence level. The exhibit shows the probability density function for the returns of a portfolio. Because VaR is being measured at the 95% confidence level, 5% of the distribution is to the left of the VaR level, and 95% is to the right.
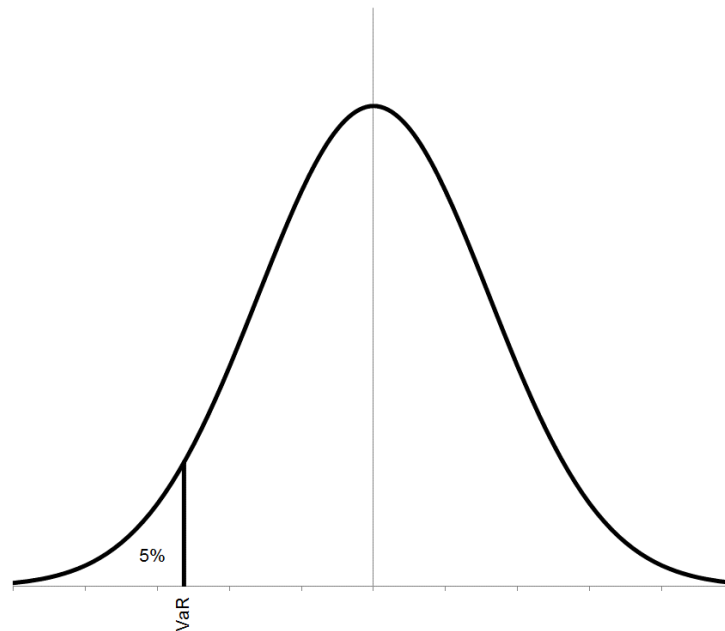
Exhibit 1: 95% VaR

The probability of a VaR exceedance should be conditionally independent of all available information at the time the forecast is made. In other words, if we are calculating the 95% VaR for a portfolio, then the probability of an exceedance should always be 5%. The probability shouldn't be different because today is Tuesday, because it was sunny yesterday, or because your firm has been having a good month. Importantly, the probability should not vary because there was an exceedance the previous day, or because risk levels are elevated.

That VaR exceedances should be conditionally independent is widely accepted by risk managers, but there has been some confusion on this point in the past. Early VaR models were often slow to react to changing market conditions. They understated risk when market volatility was increasing, and overstated risk when market volatility was decreasing. The result: when volatility was increasing there would be too many exceedances, and when volatility was decreasing there would be too few. While this may have been what many risk managers observed in practice, this was never viewed as optimal. Fortunately, modern VaR models can easily adapt to changing market conditions. Modern VaR models should perform well in both calm markets and highly volatile markets.

Backtesting VaR models is extremely straightforward. If we are measuring 1-day 95% VaR, then over 100 days we expect to see 5 exceedances. Of course, exceedance events occur randomly. Over any given 100 day period, we may see slightly more or less than 5 exceedances. While the precise number of exceedances is impossible to predict, we can still say something about the likelihood of observing a particular number of exceedances. In fact, if exceedances are conditionally independent then the

number of exceedances will follow a binomial distribution. If we denote the probability of an exceedance by *p*, then the probability of observing *k* exceedances over *n* days is

$$P[K = k] = \binom{n}{k} p^k (1-p)^{n-k}$$

Equation 1

For 1-day 95% VaR, the probability of seeing exactly 5 exceedances over 100 days is 18.0%. The probability of seeing 4 or 6 is slightly less, 17.8% and 15.0%, respectively. At the 95% confidence level, we should not be surprised if we see 4, 5 or 6 exceedances. By contrast, the probability of seeing 0 exceedances or 11 exceedances are both less than 1%. We should be very surprised if we see 0 or 11 exceedances.

A common assumption for backtesting is that there is no trading. If our VaR forecast is going to be useful for managing risk, we need to make our forecast before the start of the trading day. We want to establish the risk of the portfolio, before trading begins. Unfortunately, intraday trading will increase or decrease the risk of the portfolio making exceedances more or less likely. Does this intraday trading tend to increase or decrease exceedances? On the one hand, many firms will seek to reduce their risk if they experience significant losses early in the trading day, making VaR exceedance less likely. On the other hand, many firms close out or seek to neutralize books at the end of the trading day. This reduces end-of-day risk relative to intraday risk, making exceedance more likely. On balance it is difficult to say which of these effects dominates.

How to deal with intraday trading is an open question in risk management. In a perfect world, VaR statistics would anticipate trading. In reality, making a VaR model that anticipates trading is extremely difficult, and rarely expected. If we could anticipate trading and expected trading to reduce VaR, making an adjustment for trading could be viewed as anti-conservative. Would we even want to make an adjustment in that case? In practice, even without any adjustments our VaR forecasts might be robust to small amounts of trading. In order to isolate the impact of trading, some firms use non-trading P&L to backtest their VaR models. The non-trading P&L represents the P&L that a firm would have experienced if it had simply held the portfolio used to calculate the VaR statistic for the entire trading day. In the end, we have no strong opinion about how to handle intraday trading, but we strongly believe that intraday trading should not be used as an excuse for VaR model failure. If exceedances are consistently too frequent or too infrequent, firms should, at the very least, calculate and report non-trading P&L to ensure that this is the only source of model failure.

Data Summary & Analysis

In this paper we analyze daily VaR estimates reported to the United States Securities and Exchange Commission (SEC). The SEC requires large publicly traded financial institutions to report quantitative information about the market risk of their trading operations. Most firms satisfy this requirement by reporting daily VaR estimates.

The following tables provide a summary of the number of reported VaR exceedances reported by several large financial institutions between 2002 and 2014. The appendix includes data for each year, but in this section we have divided the data into three time periods, which we refer to as pre-crisis, crisis

and post-crisis. The firms are indicated by their tickers[1]. Not all firms reported for the entire period. The number of years for which we were able to obtain data is reported in the second row of each table. Also, the firms typically report the number of exceedances for their fiscal year, which may vary from firm to firm. Because we do not have the same number of data points for each firm, and the data points we do have are not perfectly aligned, we should be careful about any comparison between firms. This data was collected from 10-K and 10-F filings from the Security and Exchange Commission's EDGAR database.

All firms report a 1-day VaR, but some use a 95% confidence level and some use a 99% confidence level. Two firms, JP Morgan Chase and Morgan Stanley switched from 99% to 95% during the sample period. Morgan Stanley switched in 2006. Because of this, in the pre-crisis period, 2002–2006, we have only included 2002–2005 for Morgan Stanley. In general, we are encouraged by this switch from 99% to 95%. As risk managers it is tempting to push further out into the tail, but as we increase the confidence level our VaR estimates become less certain and more difficult to backtest. The choice is ultimately subjective, but based on our experience 95% provides a better balance between severity and accuracy.

The last rows of each table contain: the total number of exceedances observed, n; the expected number of exceedances, E[n]; the probability of observing n or less exceedances, P[$x \leq n$]; and, in the crisis table, the probability of observing n or more exceedances, P[$x \geq n$]. If we have 4 years of data for a firm and they are reporting VaR at the 95% confidence level, then we would expect to see 50 exceedances, 50 = [4 years] x [252 days/year] x (1 – 95%). The probability of seeing more or less than the observed number of exceedances is calculated using the binomial distribution as in Equation 1.[2] We've highlighted the number exceedance in either blue (significantly fewer than expected), red (significantly more than expected), or yellow (close to what we would expect).

If firms report backtesting results based on both trading and non-trading P&L, we tend to use the result that is mentioned first or mentioned most prominently. When in doubt, we tried to give a firm the benefit of the doubt, and use the result that was closest to the expected outcome.

The summary of exceedances pre-crisis is shown in Exhibit 2. In total, across 7 firms, there were just 12 reported exceedance events, compared to the 141 we would have expected. The number of reported exceedances is extremely low. For most of the firms there is less than a 1 in 10,000 chance of observing so few exceedances. Bear Stearns is by far the worst offender. They reported no exceedances over 4 years, compared to the 50 we would have expected. The odds of this happening are 1 in 2.85 x 10^{22}. This is exceedingly unlikely. In the life of the universe (~1.38 x 10^{10} years) we would not expect to see this happen even once by chance. Bank of America, appears to have the best result, but this has more to do with the small number of data points and the high degree of uncertainty associated with the 99% confidence level. Even then, there is only a 1.89% chance of observing this outcome.

---

[1] BAC = Bank of America, BSC = Bear Stearns, C = Citigroup, CS = Credit Suisse, GS = Goldman Sachs, JPM = JPMorgan Chase, LEH = Lehman Brothers, and MS = Morgan Stanley.
[2] For more information on the mechanics of backtesting VaR statistics, see *Mathematics and Statistics for Financial Risk Management*, 2nd Edition by Michael B. Miller, 2014, John Wiley & Sons, Inc.

2002–2006	BAC	BSC	CS	GS	JPM	LEH	MS
# years	3	4	4	3	5	1	4
conf. level	99%	95%	99%	95%	99%	95%	99%
n	2	0	3	4	2	1	0
E[n]	8	50	10	38	13	13	10
P[$x \leq n$]	1.89%	0.00%	0.95%	0.00%	0.03%	0.00%	0.00%

Exhibit 2: VaR Exceedances, Pre-Crisis

While the firms may have seemed very similar prior to the crisis, during the crisis the results were all over the place. As shown in Exhibit 3, five of the firms had far too many exceedances. The incredibly high number of exceedances reported by these firms is almost as improbable as the incredibly low number that was reported in the pre-crisis period. Maybe the financial crisis was a 1 in 200 year event, but the 34 exceedances that Credit Suisse reported should only happen once every 2.57×10^{14} years. Interestingly, Lehman —arguably the most problematic firm in the group—had far too *few* exceedances. Only Goldman Sachs was close to what we would expect.

2007–2008	BAC	BSC	CS	GS	JPM	LEH	MS
# years	2	1	2	2	2	1	2
conf. level	99%	95%	99%	95%	99%	95%	95%
n	16	27	34	23	11	4	33
E[n]	5	13	5	25	5	13	25
P[$x \leq n$]	100.00%	99.99%	100.00%	37.45%	99.45%	0.42%	95.03%
P[$x \geq n$]]	0.01%	0.02%	0.00%	70.20%	1.40%	99.88%	7.21%

Exhibit 3: VaR Exceedances, Crisis

In the wake of the financial crisis we might have expected firms to redouble their risk management efforts. At least in terms of VaR exceedances, it doesn't look like much has changed. Exhibit 4 shows the number of exceedances post-crisis. As in the pre-crisis period the number of exceedances is far less than expected.

2009–2014	BAC	C	CS	GS	JPM	MS
# years	6	3	5	6	6	6
conf. level	99%	99%	99%	95%	95%	95%
n	3	1	0	6	14	6
E[n]	15	8	13	76	76	76
P[$x \leq n$]	0.02%	0.43%	0.00%	0.00%	0.00%	0.00%

Exhibit 4: VaR Exceedances, Post-Crisis

Possible Explanations

In this section we suggest a number of explanations for the poor performance of the banks' VaR models in practice. Based on the publicly available data, it is impossible to say why the models have performed the way they have, but —as we will see in this section and the next— some explanations are more likely than others. Most of these explanations are very general in nature. They do not have anything to do with the specific requirements of SEC reporting. These explanations are likely to apply to other regulators and many would apply absent any regulatory requirement.

One possibility is that there is no problem. Maybe the models are useful in their current form and would not be more useful if they were more accurate. It's possible that the models are biased in a predictable fashion, and that risk managers simply correct for these biases when using the numbers produced by their VaR models. If a process for correcting the reported VaR did exist, why wouldn't the banks report the corrected numbers? Certainly banks, regulators and the public would benefit from VaR models that more accurately reflected the current risk of the banks' portfolios.

Another possibility is that the models cannot be made more accurate. In our experience VaR models are much more accurate than what we observe in the banks' public filings, both in calm markets and volatile markets, across various security types and trading strategies. This seems an unlikely explanation.

A third possibility is that the models could be made more accurate, but that the cost would be too great or the benefit too small. VaR models are widely used in finance and nothing in the bank's public statements suggests that they have strong objections to VaR. The small benefit argument seems very unlikely. It's possible that firms lack the expertise to build accurate risk models. It's possible that there has been a brain-drain away from big banks to hedge funds and technology firms in recent years, and that banks have failed to invest adequately in risk management expertise. A very real possibility is that the banks in our study are simply too big to manage. Banks may be burdened with legacy systems, they might struggle to gather the necessary data in a timely fashion, and they may be unable to turn the knowledge held by individuals within their firms into working models.

A fourth possibility is that risk managers are incentivized by their firms to report inaccurate numbers. We'll comment more on this possibility in the next section. As we'll see, it's difficult to differentiate between this possibility and lack of expertise.

A final possibility is that the banks themselves are incentivized by regulators to report inaccurate numbers. Under Basel II, for firms using the internal model approach, trading book capital charges are based on VaR adjusted by a multiplication factor. [3] If the firm has too many exceedances, the multiplication factor is increased. Firms may reasonably decide that it is better to slightly overstate VaR and hold slightly more capital if this significantly reduces the possibility of a large increase in capital due to a multiplier penalty. This bias could explain the pre- and post-crisis periods, but it is more difficult to reconcile with the results seen during the crisis.

In the end, if the models don't work as expected, then regulators and investors are left to guess as to why the models are failing and to what extent the models are inaccurate. In the end, if the models are not working as expected, then it is difficult, if not impossible, to use the models to accurately gauge the risk of individual financial institutions and the risk that these firms pose to the financial system as a whole.

Confusion in the Ranks

For most of the firms, it appears that the VaR models are reacting too slowly to changing market environments. We see too few exceedances pre- and post-crisis and too many during the crisis. In their regulatory filings many of the firms indicate that their VaR models are based on straight historical

---

[3] The VaR used for Basel related charges is currently the 10-day 99% VaR, not the 1-day 95% or 99%, which has been the focus of this paper. Still, if firms use the same or similar risk models to calculate different VaR statistics, then we might see similar biases and errors across various risk statistics.

simulation, with very long windows. Historical simulation is widely used in textbooks, and may have been popular in the past when computing power was more limited. However, slightly more complicated models —both parametric and non-parametric— which place more weight on recent data are widely used today and can react very quickly to changing market environments.

In their regulatory filings, many of the firms provide information about how they calculate and backtest VaR. Take for example the following two quotes from Bank of America and Goldman Sachs about the methodology they use to calculate VaR:

> "We use one VaR model … that uses a historical simulation approach based on a three-year window of historical data. …Management continually reviews, evaluates and enhances our VaR model so that it reflects the material risks in our trading portfolio."
> —Bank of America 10-K, 12/31/2013

> "We sample from 5 years of historical data ... The historical data is weighted… [The model] gives greater importance to more recent observations and reflects current asset volatilities, which improves the accuracy of our estimates of potential loss. …our VaR would increase with increasing market volatility…"
> —Goldman Sachs 10-K, 12/31/2013

Bank of America seems to be using straight historical simulation, while Goldman Sachs is using a model that weights more recent data more heavily. Not all firms provide this detail, but it appears that a number of firms are using basic historical simulation, and that firms weighting more recent data more heavily are the exception. Goldman Sachs claims that its methodology "reflects current asset volatilities", yet they still see too few exceedances pre- and post-crisis.

We find it hard to understand why large financial firms would stick with straight historical simulation. Based on our own experience, even if the firms were using straight historical simulation, the inaccuracy of the models still seems far too great. This strengthens our conviction that slow models alone cannot fully account for the discrepancies in the data presented in the last section.

Bank of America indicates that it continually "enhances" its VaR model. Other firms make similar statements indicating that their methodology changes over time. If the models were accurate in the short-term, we could easily gauge the impact of any changes. When the models are highly inaccurate, though, it is very difficult to discern the impact of any changes. It is even more difficult to guess how these changes might manifest themselves in the next crisis.

Other comments found in the filings seem to indicate confusion about how VaR backtesting even works. Take for example Citigroup:

> "…[VAR is] estimated, at a 99% confidence level … Citi would thus expect the number of days where trading losses exceed its VAR to be less than two or three exceptions per year."
> —Citigroup 10-K, 12/31/2011

At the 99% confidence level a firm would expect to see 2.5 exceedance per year, so we understand where the "two or three" comes from, but we don't understand why Citigroup would expect to see "less than two or three". Morgan Stanley is closer to the mark when it states that

> "…for a 95%/one-day VaR, the expected number of times that trading losses should exceed VaR during the year is 13, and, in general, if trading losses were to exceed VaR more than 21 times in a year, the adequacy of the VaR model could be questioned."
> —Morgan Stanley 10-K, 12/31/2011

but, while Morgan Stanley mentions the problem of too many exceedances, they fail to mention the problem of too few.

We suspect that many managers consider VaR exceedances to be a negative, and really don't mind too few exceedances. Exceedances are associated with large losses, and financial firms do not like to lose money. This may lead managers to conclude that less exceedance are associated with better performance. Unfortunately this is not the case. A firm can be highly profitable in a year when it experience the expected number of exceedances. A firm can also lose lots of money with very few exceedances.

This last point is very important. If a firm doubles the size of all of its positions, then, going forward, both the VaR and P&L will be twice what they would have been otherwise, but the number of exceedances will be exactly the same. The number of exceedances doesn't in any way indicate how much risk a firm is taking. The number of exceedances doesn't indicate that risk is too high or too low. The number of exceedances in no way indicates that a firm is doing a good or bad job at managing risk.

A related misconception has to do with capital requirements. If VaR is used to set capital requirements, then overstating VaR will lead a firm to set aside more capital. While setting aside more capital might be a good idea, overstating VaR is not the best way to achieve this outcome. If the VaR of a firm is $1 and it believes that capital reserves equal to 2x VaR would be adequate, then the firm should set aside $2 = 2 x 1$. If the firm wants to be more conservative, they could either overstate their VaR or increase the multiple. For example stating that the VaR is $1.5 or increasing the multiplier to 3x would both result in $3 of reserves being set aside: 2 x $1.5 = 3 x $1 = $3. If both strategies result in the same capital requirement, then clearly increasing the multiplier is preferable, since it preserves the accuracy of the VaR model. Overstating VaR in an attempt to be conservative is both unnecessary and counterproductive.

VaR is not a perfect risk measure and all of the firms claim to use other risk models in order to get a more complete picture of the risks they are facing. Some of these additional risk measures are discussed in their regulatory filings. For example, many of the firms include discussions of stress testing. Unfortunately we cannot backtest stress test predictions. In general, we are left to wonder: if their VaR models are not accurate, how accurate are their other risk models?

Summary

Value at risk is one of the most widely used statistics in financial risk management. Large financial institutions regularly report VaR for their trading operations. Unfortunately, in any given year, the number of VaR exceedances tend to be either far too high or far too low. This makes it very difficult to rely on the VaR levels reported by these firms. Regulators and investors are deprived of a potentially very useful tool for evaluating the risk being taken by these large financial institutions. We are left to wonder about the ability of these firms to accurately measure and manage risk.

Appendix

The following tables show the number of exceedances by year, along with the corresponding confidence levels. This data was collected from 10-K and 10-F filings from the United States Security and Exchange Commission's EDGAR database. In the exceedances table, we have used blue to indicate results that are extremely low, red for results that are extremely high, and yellow for results that are close to what we would expect.

Exceedances	BAC	BSC	C	CS	GS	JPM	LEH	MS
2002						0		0
2003		0		0		2		0
2004	0	0		1	1	0		0
2005	2	0		0	0	0		0
2006	0	0		2	3	0	1	1
2007	14	27		9	10	8	4	15
2008	2			25	13	3		18
2009	0			0	0	1		1
2010	0			0	2	0		0
2011	0			0	3	3		4
2012	0		0	0	0	3		0
2013	0		0	0	0	2		1
2014	3		1		1	5		0

Level	BAC	BSC	C	CS	GS	JPM	LEH	MS
2002						99%		99%
2003		95%		99%		99%		99%
2004	99%	95%		99%	95%	99%		99%
2005	99%	95%		99%	95%	99%		99%
2006	99%	95%		99%	95%	99%	95%	95%
2007	99%	95%		99%	95%	99%	95%	95%
2008	99%			99%	95%	99%		95%
2009	99%			99%	95%	95%		95%
2010	99%			99%	95%	95%		95%
2011	99%			99%	95%	95%		95%
2012	99%		99%	99%	95%	95%		95%
2013	99%		99%	99%	95%	95%		95%
2014	99%		99%		95%	95%		95%

References & Further Reading

Campbell, Sean D. 2006. "A Review of Backtesting and Backtesting Procedures". *Journal of Risk*, Winter.

Contreras, Patricio. 2010. "Is VAR a useful tool in volatile markets?". *Risk Magazine*.

Miller, Michael B. 2014. *Mathematics and Statistics for Financial Risk Management*, 2nd Edition. Hoboken, NJ: John Wiley & Sons, Inc.

About Northstar Risk Corp.

Northstar Risk provides software and consulting services to help hedge funds manage risk, evaluate performance, and effectively communicate with investors. Northstar's platform seamlessly integrates with back office and order management systems to provide real-time risk and performance analytics, all in one easy-to-use application.

For more information, visit www.northstarrisk.com or e-mail info@northstarrisk.com.