Article from:

Risks & Rewards

August 2009 – Issue 54

A BLACK SWAN TEST

By David Ingram

Many commentators have suggested that firms need to do stress tests to examine their vulnerability to adverse situations that are not within the data set used to parameterize their risk models. We suggest the adoption of a terminology to describe stress tests and also a methodology that can be adopted by any risk model user to test and communicate a test of the stability of model results. This method can be called a Black Swan test. The terminology would be to set one Black Swan equal to the most adverse data point. A one Black Swan stress test would be a test of a repeat of the worst event in the data set. A two Black Swan stress test would be a test of experience twice as adverse as the worst data point. So for credit losses for a certain class of bonds, if the historical period worst loss was 2 percent, then a 1BLS stress test would be a 2 percent loss, a 4 percent loss a 2BLS stress test, etc. A company could report the results of their stress tests as:

Tests show that the company can withstand a 3.5BLS stress test for credit and a 4.2BLS for equity risk and a simultaneous 1.7BLS credit and equity stress.

Similar terminology could be used to describe a test of model stability. A 1BLS model stability test would be performed by adding a single additional point to the data used to parameterize the model. So a 1BLS model stability test would involve adding a single data point equal to the worst point in the data set. A 2BLS test would be adding a data point that is twice as bad as the worst point.

For the model stability test, the model with the alternate parameterization would then be used to re-determine the risk metrics that are the primary purpose of the model.

This methodology and terminology gives a way that firms can consistently test and communicate tests to the management, board and maybe someday to those with a real need for the information, the shareholders.

The power of the idea is the complete simplicity of it and hopefully the clarity with which it can be communicated to various audiences.

So now for an example: first the most simple example, looking at the risk of a holding of an S&P 500 index equity position of $100 million. If we use the history from the past 25 years we find that the worst year was 2002 when a loss of 22.1 percent occurred. For simplicity, we will also use the simple assumption of normally distributed returns (just for the illustration—I am not recommending that this is a completely valid assumption), then we get the following:

	Column 1 Historical (1983 - 2007)	Column 2 1BLS Test	Column 3 2BLS Test	Column 4 Historical (1984 - 2008)
Average	13.8%	12.4%	11.6%	11.4%
Std Dev	15.6%	16.8%	19.0%	18.5%
Worst Year	-22.1%	N/A	N/A	-37.0%
VaR @				
5.0%	-11.8%	-15.2%	-19.7%	-18.9%
2.0%	-18.1%	-22.1%	-27.5%	-26.5%
1.0%	-22.4%	-26.6%	-32.7%	-31.5%
0.5%	-26.3%	-30.8%	-37.4%	-36.1%
0.4%	-27.4%	-32.1%	-38.9%	-37.5%
0.2%	-31.0%	-35.9%	-43.2%	-41.7%

Column 1 shows the extension of the historical data using the assumption of a normal distribution of returns for the mean of 13.8 percent and standard deviation of 15.6 percent that were determined from the historical data. Columns 2 and 3 show the 1BLS and 2BLS model stability tests, respectively. For comparison, Column 4 shows the same thing as Column 1, but for the period starting and ending one year later.

In this case, the 1BLS stress test would be the 22.1 percent loss of 2002. That makes the 2BLS stress test a 44.2 percent loss. The actual 25-year results including 2008 brings in the 37.0

"… SET ONE BLACK SWAN EQUAL TO THE MOST ADVERSE DATA POINT."

percent loss of 2008 and drops off the 22.6 percent gain of 1983 that was included in the 1983 to 2007 historical series.

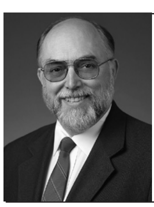
So the actual results of 2008 turned out to be a 1.67 BLS event. My suggestion is also that we should substitute that way of characterizing a new adverse event instead of the commonly used reference to the implied probability of the prior risk models, which would have said that 2008 was a 1/1800-year event.

Using this terminology, firms could report their resiliency in terms of what multiple of a 22.1 percent loss (1BLS) they could withstand. So a firm that consisted of just that $100 million equity position, a fixed liability of equal size and capital of $30 million could be said to be able to withstand a 136 percent Black Swan stress.

While regulators and creditors might be interested in company failure, investors generally have a much lower threshold for pain. This terminology could also be used to communicate volatility to the market. This could be done with what I would call the one-quarter Black Swan tests results. With a one-quarter Black Swan stress test, firms would report what multiple of 1BLS would result in a 25 percent drop in profits or a 25 percent drop in surplus. This would replace the current reporting of purely random stress tests. So in the case of the equity position, let's assume that the liability was guaranteed 3 percent, resulting in an expected profit of 10.8 percent. A 25 percent drop in profit would occur if S&P 500 return was at 11.1 percent positive return. This is, of course, much less than 0BLS and would be reported as such. A 25 percent drop in surplus would result from an S&P return of -4.5 percent, which would be reported as a 0.2BLS stress.

The advantage of using the Black Swan terminology in this case is that there is some implied probability to the discussion. Nothing specific, but saying that something is just 45 percent as bad as the worst experienced, or 0.45BLS, implies a pretty high degree of likelihood, while a 10 percent drop just presents a puzzle to the reader.

It would be quite easy for some party to determine a reasonable value for a 1BLS test for each major risk where firms are exposed based upon total market or total industry type statistics. Companies could use those benchmark type Black Swan tests and they could additionally show their own Black Swan test calibrated to their own results. The standardized Black Swan tests could also help with the issue that arises when firms develop their own distributions of losses using a process that drops out their actual worst experience because of an assumption that the circumstances that led to that historical data point will never, ever be repeated. The Black Swan test does not imply that we know how the next "once in a lifetime" loss will be but that we do know that it will likely be at least as large as the largest we have previously experienced.



David Ingram, FSA, CERA, FRM, PRM is an ERM advisor to insurers at Willis Re in New York, N.Y. He can be reached at david.ingram@willis.com.