

Canadian Investment Review

Backtesting and Stress Testing Risk Management Systems

by

Peter Christoffersen
McGill University and CIRANO
www.christoffersen.ca

Canadian Investment Review Seminar © 2003 by Peter Christoffersen

Key Issues

- Issue 1: Choosing a risk metric: VaR versus Expected Shortfall (ES).
- Issue 2: Choosing a statistical model: Historical Simulation (HS) versus dynamic volatility modeling.
- Issue 3: Backtesting risk models: Lessons from some real-life VaRs.
- Issue 4: Stress testing: Introducing a coherent framework.

Canadian Investment Review Seminar © 2003 by Peter Christoffersen

Issue 1: Choosing a Risk Metric

- VaR has become the industry benchmark for risk-reporting because it captures an important aspect of risk, namely how bad thing can get in the next K days with probability p
- However, it is not clear how K and p should be chosen.
- The VaR assumes that portfolio is constant across the next K days, which is unrealistic when K is larger than a day or a week.
- A key drawback is that the magnitude of the potential losses is ignored in VaR.

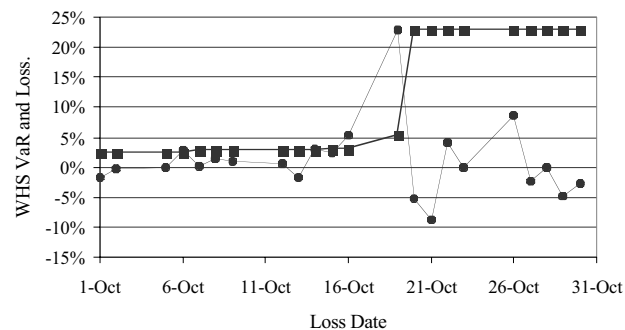
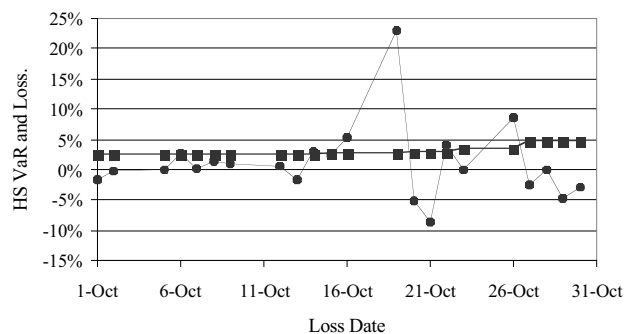
Expected Shortfall

- Expected Shortfall (ES) or TailVaR captures the loss magnitude.
- Expected Shortfall is defined as the expected value of losses in the cases where the VaR is violated.
- The expected value implies that the magnitude of the large negative returns now matter!
- Can capture nonlinearities (e.g. options) and non-normality in returns.

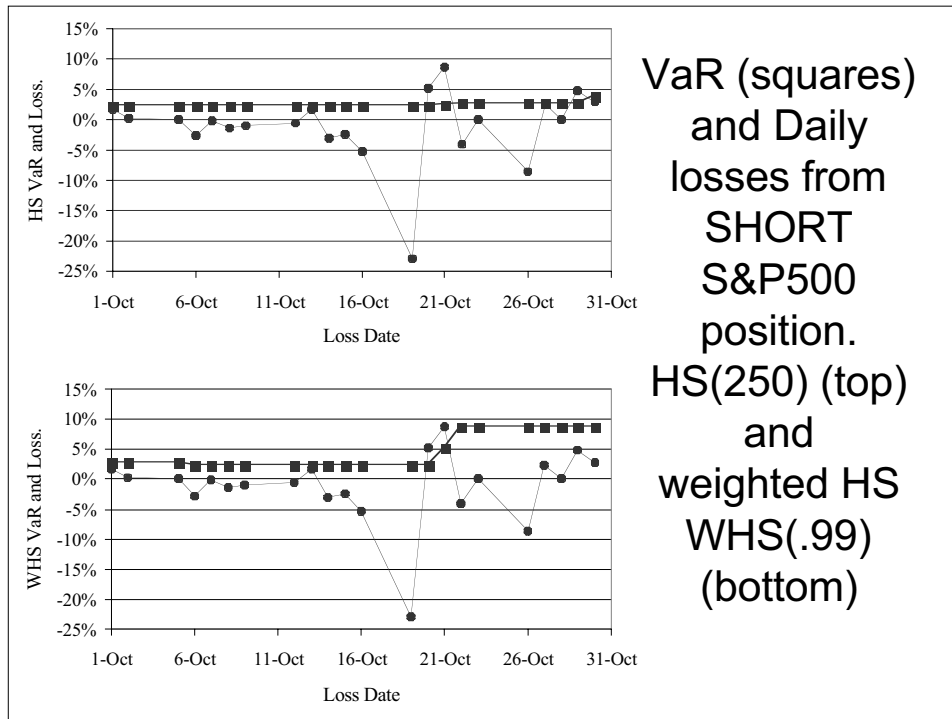
Issue 2: Calculating VaR

- Historical Simulation (HS) calculates VaR putting equal weight on the last m historical returns. HS is widely used in practice because of
 - 1) The ease with which it is implemented
 - No parameters have to be estimated
 - 2) Its model-free nature.
 - HS does not rely on a parametric model.
- Bad models lead to bad model-based VaRs. But what about good models?
- The choice of m is key and not easy to make.
- HS does not react to dramatic events.
- Weighted HS alternative is not foolproof.

Canadian Investment Review Seminar © 2003 by Peter Christoffersen



VaR (squares)
and Daily
losses from
LONG
S&P500
Position.
HS(250) VaR
(top) and
weighted HS
WHS(.99) VaR
(bottom)



VaR (squares) and Daily losses from SHORT S&P500 position. HS(250) (top) and weighted HS WHS(.99) (bottom)

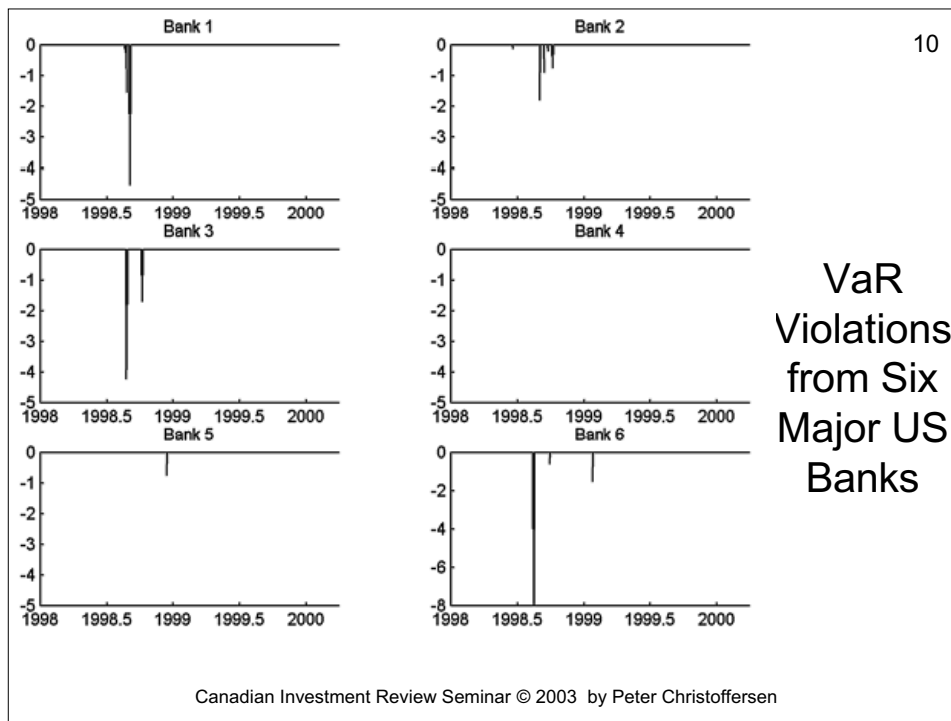
We Need a Volatility Model

- Portfolio return variance is correlated over time.
- A dynamic variance model estimated on portfolio returns would have picked up the increase in market variance from the crash regardless of whether the crash meant a gain or a loss to us.
- We therefore should build volatility dynamics into our risk model.
- Key challenge at longer horizons: What drives long term volatility?

Issue 3: Backtesting VaRs

- The next figure shows the violations (measured in excess return standard deviations) of the VaR in six large U.S. commercial banks from January 1998 to March 2001.
- The return is daily and the VaR is 1-day, 1%.
- One bank (4) does not have any violations at all, and in general the banks have fewer violations than expected.
- Thus the banks on average report a VaR which is higher than it should be. They are on average conservative.

Canadian Investment Review Seminar © 2003 by Peter Christoffersen



Clustered Violations

- The violations of each of Bank 1,2,3,5 fall within a short time span (Fall of 1998), and often on adjacent days.
- This clustering of VaR violations is a serious sign of risk model misspecification.
- Most banks rely on Historical Simulation, which is very slow at updating the VaR when market volatility increases.
- The clustering of violations across banks is important from a regulator perspective.
- It motivates the use of backtesting techniques.

Canadian Investment Review Seminar © 2003 by Peter Christoffersen

Backtesting VaRs

- A VaR_{t+1}^p measure promises that the actual return will only be worse than the VaR_{t+1}^p forecast a fraction p of the days in the sample.
- If we observe a time series of past ex-ante VaR forecasts and past ex-post returns, we can define the “hit sequence” of VaR violations as

$$I_{t+1} = \begin{cases} 1, & \text{if Loss} > VaR_{t+1}^p \\ 0, & \text{if Loss} < VaR_{t+1}^p \end{cases}$$

Canadian Investment Review Seminar © 2003 by Peter Christoffersen

Conditional Backtesting

- Consider regressing the hit sequence on the vector of “suspected misspecification” variables, X_t known on day t .
- In a simple linear regression, we would have

$$I_{t+1} = b_0 + b_1'X_t + e_{t+1}$$

- The hypothesis of a correct VaR model is

$$H_0 : b_0 = p, b_1 = 0$$

- F-test allows for constructive, conditional backtesting.

Issue 4: Stress Testing

- What do we do if our historical data set contains no large losses?
- We can generate extreme scenarios of the portfolio returns and then assess the resulting VaR or ES from the risk model.
- This is referred to as stress testing, as we are *stressing* the model by exposing it to data different from the data used when specifying and estimating the model.

The Problem with Stress Testing

- VaR and ES are proper probabilistic statements:
 - What is the loss such that I will lose more only 1% of the time (VaR)?
 - What is the expected loss when I violate my VaR (ES)?
- But standard stress testing does not tell the risk manager anything about the probability of the scenario happening, thus it is not clear what the risk management action (if any) should be.

Coherent Stress Testing

- If we are willing to assign scenario probabilities, then stress testing can be very useful.
- To be explicit, consider a simple example of one stress scenario, which we define as a probability distribution $f_{stress}(*)$ of the portfolio returns.
- We can simulate portfolio returns from the risk model, call it $f(*)$ and we also simulate from the scenario distribution $f_{stress}(*)$ using the prespecified scenario probability.

Combining the Distributions

- Once we have simulated data from the combined data set, we can calculate the VaR or ES risk measure on the combined data.
- If the risk measure is viewed to be inappropriately high then the portfolio can be rebalanced.
- Backtesting can be done on the extended data set as well.

Choosing Scenarios

- Simulate shocks which are more likely to occur than the historical data base suggests.
- Simulate shocks that have never occurred but could.
- Simulate shocks reflecting the possibility that current statistical patterns could break down.
- Simulate shocks which reflect structural breaks which could occur.

Summary: Key Points

- The Value-at-Risk measure is ok but Expected Shortfall is better.
- The Historical Simulation and Weighted Historical Simulation techniques do not properly capture the variance dynamics in returns.
- Conditional Backtesting can catch inadequate models in a constructive fashion.
- Coherent Stress Testing using prespecified scenario probabilities enables proper risk management actions to be taken.

Background Material

Elements of Financial Risk Management

By Peter Christoffersen

Academic Press / Elsevier Science

2003