Risk Measurement for Financial Institutions

Dr. Graeme West
Financial Modelling Agency
www.finmod.co.za
graeme@finmod.co.za

July 13, 2006

Contents

Chapter 1

Risk types and their measurement

There are various types of risk. A common classification of risks is based on the source of the underlying uncertainty.

1.1 Market Risk

By market risk, we mean the potential for unexpected changes in value of a position resulting from changes in market prices, which results in uncertainty of future earnings resulting from changes in market conditions, (e.g., prices of assets, interest rates).

These pricing parameters include security prices, interest rates, volatility, and correlation and inter-relationships.

Over the last few years measures of market risk have evolved to become synonymous with stress testing, measurement of sensitivities, and VaR measurement.

1.2 Credit Risk

Credit risk is a significant element of the galaxy of risks facing the derivatives dealer and the derivatives end-user. There are different grades of credit risk. The most obvious one is the risk of default. Default means that the counterparty to which one is exposed will cease to make payments on obligations into which it has entered because it is unable to make such payments.

This is the worst case credit event that can take place. From this point of view, credit risk has three main components:

- Probability of default - probability that a counterparty will not be able to meet its contractual obligations
- Recovery Rate - percentage of the claim we will recover if the counterparty defaults
- Credit Exposure - this related to the exposure we have if the counterparty defaults

But this view is very naive. An intermediate credit risk occurs when the counterparty's creditworthiness is downgraded by the credit agencies causing the value of obligations it has issued to decline in value. One can see immediately that market risk and credit risk interact in that the contracts into which we enter with counterparties will fluctuate in value with changes in market prices, thus affecting the size of our credit exposure. Note also that we are only exposed to credit risk on contracts in

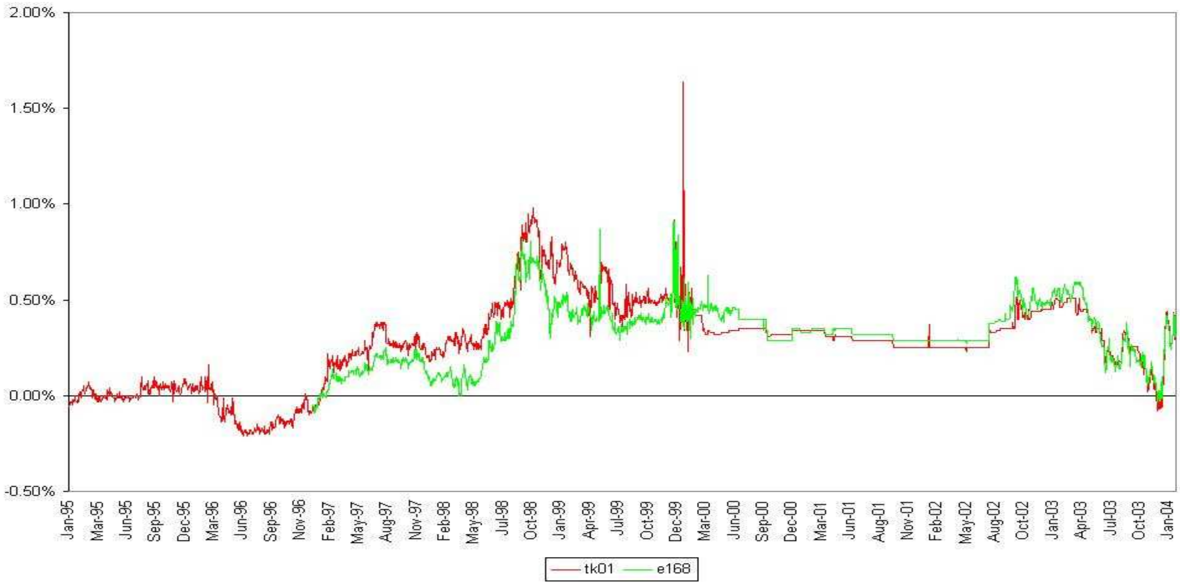


Figure 1.1: Widening spreads in ytm for bonds of different creditworthiness

which we are owed some form of payment. If we owe the counterparty payment and the counterparty defaults, we are not at risk of losing any future cash flows.

The effect of a change in credit quality can be very gradual. In the graph, we have the time series of the difference of the ytm's of the tk01 and e168 to the r153. These bonds have maturity 31 Mar 2008 and 1 Jun 2008 and 31 Aug 2010 respectively, with annual coupons of 10%, 11% and 13% respectively. Thus, they are (or should be) very similar bonds. There are differences in creditworthiness however, and this distinction has become more apparent with the ANC government and the stated intention of the privatisation of the parastatals. Previously, NP government guarantees of the performance of the parastatals was implicit.

Credit risk is one source of market risk, but is not always priced properly.

1.3 Market imperfections

Within credit markets, two important market imperfections, adverse selection and moral hazard, imply that there are additional benefits from controlling counterparty credit risk, and from limiting concentrations of credit risk by industry, geographic region, and so on. This piece is adapted from (Duffie & Singleton 2000).

Adverse selection

Suppose, as is often the case with a simple loan, that a borrower knows more than its lender, say a bank, about the borrower's credit risk. Being at an informational disadvantage, the bank, in light of the distribution of default risks across the population of borrowers, may find it profitable to limit borrowers' access to the bank's credit, rather than allowing borrowers to select the sizes of their own loans without restriction. An attempt to compensate for credit risk by specifying a higher average interest rate, or by a schedule of interest rates that increases with the size of the loan, may have unintended consequences. A disproportionate fraction of borrowers willing to pay a high interest rate on a loan are privately aware that their own high credit risk makes even the high interest rate attractive. An interest rate so high that it compensates for this adverse selection could mean that almost no borrower finds a loan attractive, and that the bank would do little or no business.

It usually is more effective to limit access to credit. Even though adverse selection can still occur to some degree, the bank can earn profits on average, depending on the distribution of default risk in the population of borrowers. When a bank does have some information on the credit quality of individual borrowers (that it can legally use to set borrowing rates or access to credit) the bank can use both price and quantity controls to enhance the profitability of its lending operations. For example, banks typically set interest rates according to the credit ratings of borrowers, coupled with limited access to credit.

In the case of an over-the-counter derivative, such as a swap, an analogous asymmetry of credit information often exists. For example, counterparty A is typically better informed about its own credit quality than about the credit quality of counterparty B. (Likewise B usually knows more about its own default risk than about the default risk of A.) By the same adverse-selection reasoning described above for loans, A may wish to limit the extent of its exposure to default by B. Likewise, B does not wish its potential exposure to default by counterparty A to become large. Rather than limiting access to credit in terms of the notional size of the swap, or its market value (which in any case is typically zero at inception), it makes sense to measure credit risk in terms of the probability distribution of the exposure to default by the other counterparty.

Moral hazard

Within banking circles, there is a well known saying: "If you owe your bank R100,000 that you don't have, your are in big trouble. If you owe your bank R100,000,000 that you don't have, your bank is in big trouble." One of the reasons that large loans are more risky than small loans, other things being equal, is that they provide incentives for borrowers to undertake riskier behaviour. If these big bets turn out badly (as they ultimately did in many cases) the risk takers can walk away. If the big bets pay off, there are large gains.

An obvious defence against the moral hazard induced by offering large loans to risky borrowers is to limit access to credit. The same story applies, in effect, with over-the-counter derivatives. Indeed, it makes sense, when examining the probability distribution of credit exposure on an OTC derivative, to use measures that place special emphasis on the largest potential exposures.

1.4 Liquidity risk

Liquidity risk is reflected in the increased costs of adjusting financial positions. This may be evidenced by bid-ask spreads widening; more dramatically arbitrage-free relationships fail or the market may disappear altogether. In extreme conditions a firm may lose its access to credit, and have an inability to fund its illiquid assets.

There are 2 types of liquidity risk:

- Normal or usual liquidity risk - this risk arises from dealing in markets that are less than fully liquid in their standard day-to-day operation. This occurs in almost all financial markets but is more severe in developing markets and. specialist OC instruments.
- Crises liquidity risk - liquidity arising because of market crises e.g. times of crisis such as 1987 crash, the ERM crisis of 1992, the Russian crisis in August 1998, and the SE Asian crisis of 1998, we find the market had lost its normal level of liquidity. One can only liquidate positions by taking much larger losses.

1.5 Operational risk

This includes the risk of a mistake or breakdown in the trading, settlement or risk-management operation. These include

- trading errors
- not understanding the deal, deal mispricing
- parameter measurement errors
- back office oversight such as not exercising in the money options
- information systems failures

An important type of operational risk is management errors, neglect or incompetence which can be evidenced by

- unmonitored trading, fraud, rogue trading
- insufficient attention to developing and then testing risk management systems
- breakdown of customer relations
- regulatory and legal problems
- the insidious failure to quantify the risk appetite.

1.6 Legal risk

Legal risk is the risk of loss arising from uncertainty about the enforceability of contracts.

Its includes risks from:

- Arguments over insufficient documentation
- Alleged breach of conditions
- Enforceability of contract provisions - regards netting, collateral or third-party guarantees in default of bankruptcy

Legal risk has been a particular issue with derivative contracts. Many banks found their swaps contracts with London Boroughs of Hammersmith and Fulham voided when the courts of England upheld the argument that the borough management did not have the legal authority to deal in swaps.

Chapter 2

Infamous risk management disasters

In all the cases we discuss the institution was exposed to risks, supposedly without management being aware of them. But in many cases senior management were aware of the weaknesses in their risk-control systems, (or should have been) but failed to act. And very often the risks would have been picked up even with the simplest VaR implementation.

2.1 Wall street crash of 1987

When the portfolio insurance policy comprises a protective put position, no adjustment is required once the strategy is in place. However, when insurance is effected through equivalent dynamic hedging in index futures and risk free bills, it destabilises markets by supporting downward trends. This is because dynamic hedging involves selling index futures when stock prices fall. This causes the prices of index futures to fall below their theoretical cost-of-carry value. Then index arbitrageurs step in to close the gap between the futures and the underlying stock market by buying futures and selling stocks through a sell program trade.

2.2 Metallgesellschaft

Metallgesellschaft is a huge German industrial conglomerate dealing in energy products. From 1990 to 1993 they sold long-term forward contracts supplying oil products (the equivalent of 180 million barrels of oil) to their consumers. In order to hedge the position, they went long a like number of oil futures.

However, futures are short term contracts. As each future expired, they rolled it over to the next expiry. Of course, this exposed the company to basis risk.

The price of oil decreased. Thus they made a loss on the futures position and a profit on the OTC forwards. The problem is that losses on futures lead to margin calls whereas the profits on the forwards were still a long time from being realised. In fact, there were $1 billion margin calls on the futures positions.

The management of Metallgesellschaft were unwilling to continue to fund the position. They fired all the dealers, closed out all the futures positions, and allowed the counterparties to the forwards to walk away.

A loss of $1.3 billion was incurred. The share price fell from 64DM to 24DM.

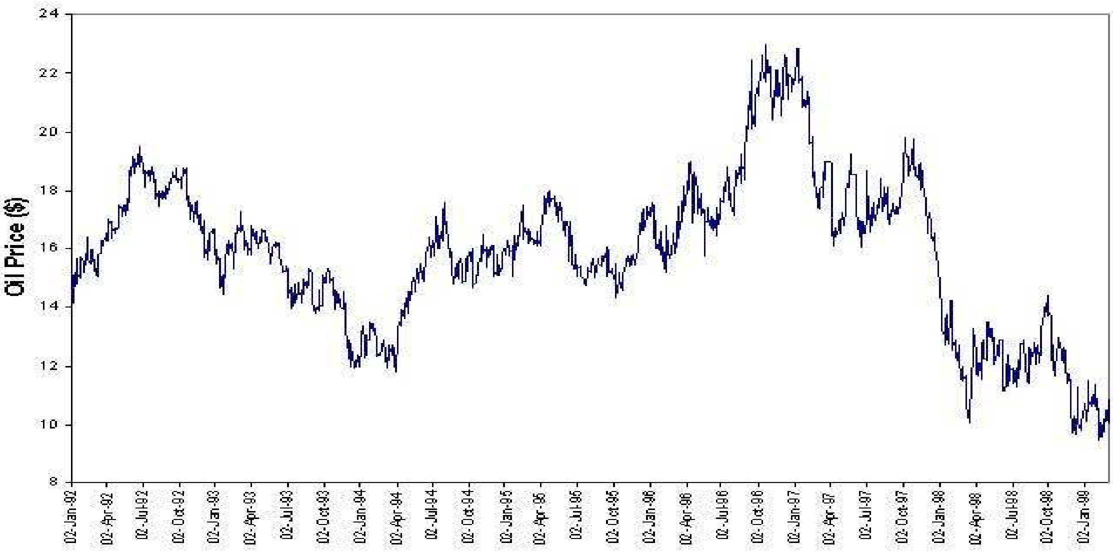


Figure 2.1: The price of oil (Dubai)

2.3 Kidder Peabody

In April 1994, Kidder announced that losses at its government bond desk would lead to a $210 million charge against earnings, reversing what had been expected to be the firm's largest quarterly profit in its 129-year history. The company disclosed that Joseph Jett, the head of the government bond desk, had manufactured $350 million in "phantom" trading profits, in what the Securities and Exchange Commission later called a "merciless" exploitation of the firm's computer system. Kidder's internal report on the incident concluded that the deception went unnoticed for two years due to "lax supervision". Mr. Jett, who denied that his actions were unknown to his superiors, was found guilty of recordkeeping violations by an administrative judge.

2.4 Barings

This is probably the most famous banking disaster of all. Early in 1995, the futures desk of Baring's in Singapore was controlled by Nic Leeson, a 28 year old trader.

He had long and unauthorised speculative futures positions on the Nikkei.

However, the Nikkei fell significantly. This was in no small part due to the Kobe earthquake of 17 January. He was faced with huge margin calls, which for a while he funded by taking option premia. However, eventually there was no more funds, and he absconded on 23 February 1995. The bank officially failed 26 February 1995, with a loss of $1 billion.

Leeson was sentenced to 6 and a half years in prison. The bank was sold for 1 pound to ING.

The reasons for failure: outright operational failure, tardiness of exchanges, tardiness of the Bank of England.

In a survey of 1997 (Paul-Choudhury 1997):

- 75% of risk managers believed that their organisation could not suffer a Barings-type disaster.

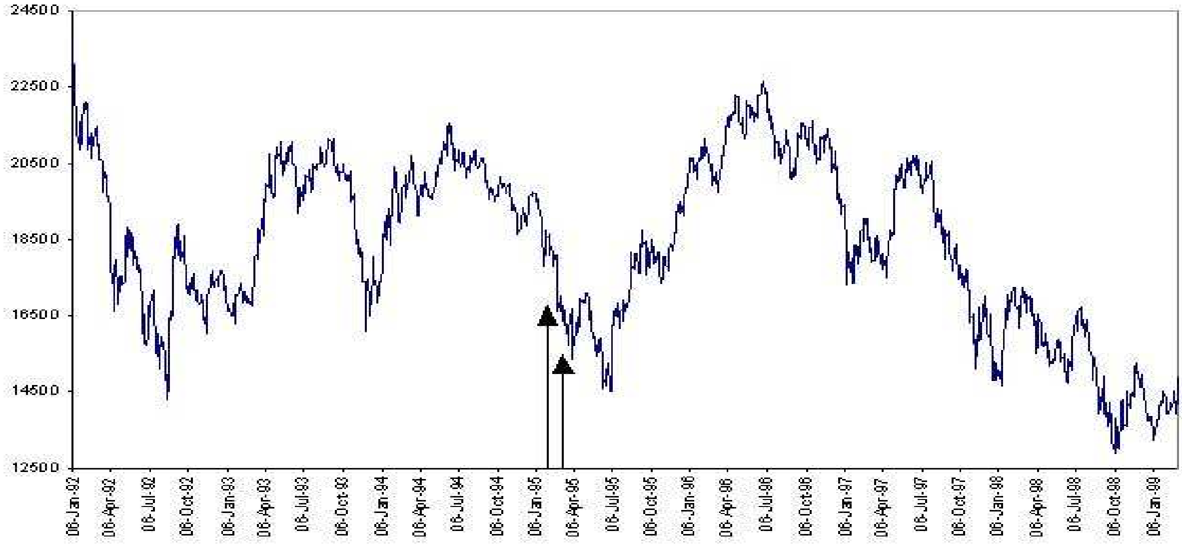


Figure 2.2: The Nikkei index

- 75% of traders believed that their organisation could suffer a Barings-type disaster.
- 85% of traders believed that they could hide trades from their risk manager.

2.5 US S&L Industry

In the 1980s, Savings-and-Loan institutions were making long term loans in housing and property at a fixed rate, and taking short term deposits such as mortgage payments. In the face of market volatility and changes in the shape of the interest rate term struture, the US Congress made the mistake of deregulating the industry. This allows moral hazard.

One consequence of this deregulation was that savings-and-loan institutions has access to extensive credit through deposit insurance, while at the same time there was no real enforcement of limits on the riskiness of savings-and-loans investments. This encouraged some savings-and-loans owners to take on highly levered and risky portfolios of long-term loans, mortgage-backed securities, and other risky assets. Many went insolvent.

2.6 Orange County

The investment pool was invested in highly leveraged investments. The dealer Bob Citron insisted that MtM was irrelevant because a hold to maturity strategy was followed. This nonsense was believed for some time, but the eventual outcome was bankruptcy.

2.7 LTCM

Long Term Capital Management failed spectacularly in 1998. This was a very exclusive hedge fund whose partners included Myron Scholes, Robert Merton and John Meriwether. Their basic play, all over the world, was on credit spreads narrowing - thus, they were typically long credit risky bonds and short credit-safe bonds.

What were the reasons?

- Widening credit spreads and liquidity squeeze after the Russian default of 1998 - subsequent talk of a market in liquidity options by Scholes, amongst others.
- Very large leverage, which increased as the trouble increased, and as liquidity dried up. In other words, not enough long term capital.
- Excessive reliance on VaR without performing stress testing. They were caught out as a new paradigm was emerging: as VaR inputs are always historical, none of what was happening was an 'input' to the VaR model.
- Model risk - too many complex plays. Infatuation with sexy deals, which were retained as the portfolio was reduced. This reduced the liquidity even further.

LTCM was bailed out under rather suspicious circumstances by a consortium of creditors organised by Alan Greenspan of the Federal Reserve Bank. The exact conditions and motives for this are still not known - involvement by the legislators increases moral hazard going forward. It was argued that failure of LTCM could destabilise international capital markets. See (Kolman April 1999), (The Financial Economists Roundtable October 6, 1999), and the standard book on the subject, (Lowenstein 2000).

2.8 Allied Irish Bank

More recently, Allied Irish Banks PLC disclosed in February that a rogue trader accumulated almost \$700 million in losses over a five-year period. The losses, incurred at its U.S. foreign exchange operation Allfirst, caused the company to reduce its 2001 net income by over \$260 million (about 38%). The Wall Street Journal claimed that Allfirst had a 25-year-old junior employee monitoring currency trading risk, an assertion that the bank denied. Bank officials believe that John Rusnak avoided the company's internal checks by contracting out administration to banks that were complicit in the fraud.

2.9 National Australia Bank

The NAB options team made a loss in October 2003, right around the time they were expecting their performance bonuses, and rather than jeopardise them they tried to push the loss forward and wait for an opportunity to trade out of it - then decided to bet on the Australian dollar dropping. With the Australian dollar charging ahead in late 2003, they were left with a \$180m loss within weeks.

One of the dealers was open with the press. His most serious claim was that the bank's risk-management department had been signing off on the losses for months. "We were already over the limits for a number of months and the bank knew about it... It has been going on and off for a year and consistently every day since October. It was signed off every day by the risk-management people."

This is a direct contradiction of the bank's claims that the \$180 million loss was the result of unauthorised trades that had been hidden from senior management.

A former options trader wrote: "I can tell you that NAB have been doing dodgy trading stuff for much longer than a few months. The global FX options market has been waiting for them to blow up for years. No-one is surprised by this at all, except the fact that it took so long."

The risk management situation at NAB seemed very poor. Chris Lewis was the senior KPMG auditor who had headed a due diligence team to advice whether the bank should buy Homeside in Florida;

this advise was in the affirmative. As auditor he also signed the 2000 accounts and claimed they were "free of material mis-statement", when in fact the bank was about to lose $3.6 billion from mortgage servicing risk at Homeside, which wasn't even mentioned in the annual report.

Lewis was hired as the head of risk soon afterwards! It is clearly a conflict to have auditors who spend years convincing themselves everything is okay and then go and take over the reigns of internal audit at the same client, as there is a lack of fresh perspective. Not to mention that his competency was in question.

Chapter 3

Value at Risk

We will focus for the remainder of this course on measuring market risks. It is only measurement of this type of risk, that has evolved to a state of near-finality, from a quantitative point of view. The standard ways of measuring market risks is via VaR or a relative thereof, stress testing, and sensitivities.

VaR was the first risk management tool developed that took into account portfolio and diversification effects.

VaR is the largest loss on a portfolio that will be experienced to a given high level of confidence, over a specified holding period, based on a distribution of value changes.

So, if the 10 day 95% VaR is R10m then over the next 10 days, the portfolio will

- with 95% probability, either make a profit, or a loss less than R10m.
- with 95% probability, will have a p&l of more than -R10m.
- with 5% probability, will make a loss of more than R10m.
- with 5% probability, will have a p&l of less than -R10m.

This does not mean that the 'risk' is R10m - the whole porfolio could vapourise, and the loss will presumably be more than R10m.

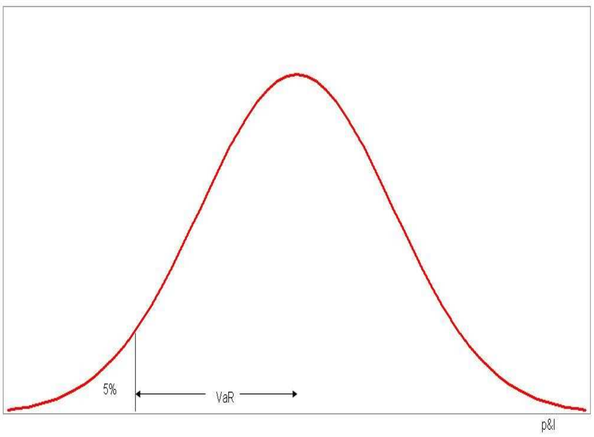


Figure 3.1: A typical p&l distribution with tail

The term 10 days above is known as the 'holding period'.

The term 95% is known as the 'confidence level'.

Thus, a formal definition:

Definition 1 *The N-day VaR is x at the α confidence level means that, according to a distribution of value changes, with probability α, the total p&l over the next N days will be $-x$ or more.*

Are the following consistent?

- 1 day 95% VaR of R10m
- 1 day 99% VaR of R5m

Certainly not. As our confidence increases the VaR number must increase. So, we might have

- 1 day 95% VaR of R10m
- 1 day 99% VaR of R20m

Are the following consistent?

- 1 day 95% VaR of R10m
- 1 day 99% VaR of R20m
- 10 day 99% VaR of R20m

Certainly not. More likely we would have:

- 10 day 99% VaR of R40m, say.

So, the following may very well be consistent:

- 1 day 95% VaR of R10m
- 1 day 99% VaR of R20m
- 10 day 99% VaR of R40m

Changing the holding period or the confidence level changes the reported VaR, but not the reality.

What are the factors driving the VaR of a position?

- size of positions - should be linear in size. But in extreme cases the size of the position affects the liquidity.
- direction of positions - not linear in direction eg. a call.
- riskiness of the positions - more speculative positions and/or more volatility should contribute to an increase in VaR.
- the combination of positions - correlation between positions.

'Distribution of value changes' - distribution needs to be determined, either explicitly or implicitly, and sampled. Current VaR possibilities:

- The Variance-Covariance approach, in other words the classic RiskMetrics© approach, or a variation thereof.
- Various historical simulation approaches - 'history repeats itself'.
- Monte Carlo simulation.

The choice of VaR method can be a function of the nature of the portfolio. For fixed income and equity, a variance-covariance approach is probably adequate. For plain vanilla options a simple enhancement of VCV such as the delta-gamma approach is often claimed to be suitable (the author disagrees), but if there are more exotic options, a more advanced full revaluation method is required such as historical or Monte Carlo.

The fundamental problem we are faced with is how to aggregate risks of various positions. They cannot just be added, because of possible interactions (correlations) between the risks.

In making the decision of which method to use, there is a tradeoff between computational time spent and the 'accuracy' of the model. It should be noted in this regard that traders will attempt to game the model if their limits or remuneration is a function of the VaR number and there are perceived or actual limitations to the VaR calculation. Thus (as already mentioned) limits on VaR need to be supplemented by limits on notionals, on the sensitivities, and by stress and scenario testing.

3.1 RiskMetrics©

In the following examples we compute VaR using standard deviations and correlations of financial returns, under the assumption that these returns are normally distributed. In most markets the statistical information is provided by RiskMetrics, but in South Africa, for example, the data is provided a day late. This is unsatisfactory for immediate risk management. Thus the institution should have their own databases of RiskMetrics type data.

The RiskMetrics assumption is that standardised returns are normally distributed given the value of this standard deviation. This is of course the fundamental Geometric Brownian Motion model.

α% VaR is derived via $-z_\alpha$ times the standard deviation of returns, which is given by $\frac{\sigma}{\sqrt{250}}$, where σ is the annualised volatility of returns. Here $z_.$ is the inverse of the cumulative normal distribution, so, for example, if $\alpha = 95\%$ then $-z_{0.95} = -1.645$. Thus, a 'bad' outcome, for a portfolio which is positive valued, would be a negative stock return of $-z_\alpha \frac{\sigma}{\sqrt{250}}$, and the VaR is

$$V\left(1 - \exp\left(-z_\alpha \frac{\sigma}{\sqrt{250}}\right)\right) \tag{3.1}$$

If the portfolio is negative valued, the bad outcome would be a positive stock return of $z_\alpha \frac{\sigma}{\sqrt{250}}$, and so the VaR is

$$V\left(1 - \exp\left(z_\alpha \frac{\sigma}{\sqrt{250}}\right)\right) \tag{3.2}$$

Note here we have two negatives, giving us a positive VaR value.

We will call this approach the 'RiskMetrics full precision' method. For another possibility, note that by Taylor series $1-e^x \approx -x \approx e^{-x}-1$. Hence, for either a long or short position, VaR is approximately given by

$$|V| z_\alpha \frac{\sigma}{\sqrt{250}} \tag{3.3}$$

We will call this the 'standard RiskMetrics simplification'. Indeed, when reading (J.P.Morgan & Reuters December 18, 1996) it is very problematic to know at any stage which method is being referred to. Unfortunately, the standard simplification method does not have much theoretical motivation: prices are not normally distributed under any model - it is returns that are typically modelled as being normal.

Example 1 *You hold 2,000,000 shares of SAB. Currently the share is trading at 70.90 and the volatility of the return of SAB, measured historically, is 24.31%.*

What is your 95% VaR over a 1-day horizon on 23-Jan-04?

Your exposure is equal to the market value of the position in ZAR. The market value of the position is $2,000,000 \cdot 70.90 = 141,800,000$.

The VaR of the position is $2,000,000 \cdot 70.90 \cdot \left(1 - \exp\left(-z_{0.95}\frac{24.31\%}{\sqrt{250}}\right)\right) = 3,540,616.$

Now suppose we have a portfolio. Here the covolatility matrix Σ is measured in returns. Then $\sigma(R)$ is the volatility of the return of the portfolio, and is found as $\sigma(R) = \sqrt{w'\Sigma w}$ as in classical portfolio theory. Here w_i are the proportional value weights, with $\sum_{i=1}^{n} w_i = 1$. So VaR can be measured directly. The assumption is again made that the return R is normally distributed, and the formulae for VaR are as before.

Example 2 *You hold 2,000,000 shares of SAB and 500000 shares of SOL. SOL is trading at 105.20 with a volatility of 32.10%. The correlation in returns is 4.14%. What is your 95% VaR over a 1-day horizon on 23-Jan-04?*

This time, the MtM is $2,000,000 \cdot 70.90 + 500000 \cdot 105.20 = 194,400,000$.

The daily standard deviation in returns are $\sigma_1 = 1.54\%$ *and* $\sigma_2 = 2.03\%$. *The value weights are* $w_1 = 72.94\%$ *and* $w_2 = 27.06\%$. *The correlation in returns is* $\rho = 4.14\%$. *Thus, using the portfolio theory formula for the standard deviation of the returns of a portfolio,*

$$\sigma(R) = \sqrt{w_1^2\sigma_1^2 + 2w_1w_2\rho\sigma_1\sigma_2 + w_2^2\sigma_2^2} \tag{3.4}$$

which is equal to 1.27% in this case. The VaR calculation proceeds as before, yielding a VaR of 4,015,381.

This is a good opportunity to introduce the concept of undiversified VaR. We calculate the VaR for each instrument on a stand-alone basis: $\text{VaR}_1 = 3,540,616$ and $\text{VaR}_2 = 1,727,495$, for a total undiversified VaR of 5,268,111. The fact that the VaR of the portfolio is actually 4,015,381 is an illustration of portfolio benefits.

RiskMetrics provides users with $1.645\sigma_1$, $1.645\sigma_2$, and ρ. One has to take care of the factor 1.645: whether to leave it in or divide it out, according to the required application. As has been indicated, this information is certainly provided in the South African environment, but it is a day late. It is not difficult to calculate these numbers oneself, using the prescribed methodology. The EWMA method with $\lambda = 0.94$ is the method prescribed by RiskMetrics for the volatility and correlation calculations.

If we make the standard RiskMetrics simplification, a neat simplifying trick is possible. Suppose $\sigma(R)$ and Σ are daily measures. Then

$$\text{VaR} = |V|z_\alpha\sigma(R) = \sqrt{V^2}z_\alpha\sqrt{w'\Sigma w} = z_\alpha\sqrt{W'\Sigma W} \tag{3.5}$$

where $W_i = w_iV$ is the value of the i^{th} component.

Calculating VaR on a portfolio of cash flows usually involves more steps than the basic ones outlined in the examples above. Even before calculating VaR, you need to estimate to which risk factors a particular portfolio is exposed. The RiskMetrics methodology for doing this is to decompose financial instruments into their basic cash flow components. We use a simple example - a bond - to demonstrate how to compute VaR. See (J.P.Morgan & Reuters December 18, 1996, §1.2.1).

Example 3 *Suppose on 25-Jun-03 we are long a r150 bond. This expires 28-Feb-05, with a 12.00% coupon paid, with coupon dates 28-Feb and 31-Aug. How do we calculate VaR using the standard RiskMetrics simplification?*

The first step is to map the cash flows onto standardised time vertices, which are 1m, 3m, 6m, 1y, 2y, 3y, 4y, 5y, 7y, 9y, 10y, 15y, 20y and 30y (J.P.Morgan & Reuters December 18, 1996, §6.2). We will suppose we have the volatilities and correlations of the return of the zero coupon bond for all of these time vertices.

The actual cash flows are converted to RiskMetrics cash flows by mapping (redistributing) them onto the RiskMetrics vertices. The purpose of the mapping is to standardize the cash flow intervals of the instrument such that we can use the volatilities and correlations of the prices of zero coupon bonds that are routinely computed for the given vertices in the RiskMetrics data sets. (It would be impossible to provide volatility and correlation estimates on every possible maturity so RiskMetrics provides a mapping methodology which distributes cash flows to a workable set of standard maturities).

The RiskMetrics methodology (J.P.Morgan & Reuters December 18, 1996, Chapter 6) for mapping these cash flows is not completely trivial, but is completely consistent. We linearly interpolate the risk free rates at the nodes to risk free rates at the actual cash flow dates. Likewise we linearly interpolate the price volatilities at the nodes to price volatilities at the actual cash flow dates.

However, there is another method of calculating the price return volatility of the interpolated node. If A and C are known, and B is interpolated between them,

$$\sigma_B = \sqrt{w^2\sigma_A^2 + 2w(1-w)\rho_{A,C}\sigma_A\sigma_C + (1-w)^2\sigma_C^2} \tag{3.6}$$

Here the unknown is w; the above can be reformulated as a quadratic, where $w \in [0,1]$ is the smaller of the two roots of the quadratic $\alpha x^2 + \beta x + \gamma$ with

$$\begin{aligned} {}^{1}\alpha &= \sigma_A^2 + \sigma_C^2 - 2\rho_{A,C}\sigma_A\sigma_C \\ \beta &= 2\rho_{A,C}\sigma_A\sigma_C - 2\sigma_C^2 \\ \gamma &= \sigma_C^2 - \sigma_B^2 \end{aligned}$$

Thus we have a portfolio of cash flows occurring at standardised vertices, for which we have the price volatilities and correlations.

Using the formula $VaR = z_{95\%}\sqrt{W'\Sigma W}$ we get the VaR of the bond.

When the relationship between position value and market rates is nonlinear, then we cannot estimate changes in value by multiplying 'estimated changes in rates' by 'sensitivity of the position to changing rates'; the latter is not constant (i.e., the definition of a nonlinear position).

Recall that for equity option positions

$$\begin{aligned} \delta V &\approx \frac{\partial V}{\partial S}\,\delta S + \tfrac{1}{2}\frac{\partial^2 V}{\partial S^2}\,(\delta S)^2 \\ &= \Delta\,\delta S + \tfrac{1}{2}\Gamma\,(\delta S)^2. \end{aligned}$$

The RiskMetrics analytical method approximates the nonlinear relationship via a Taylor series expansion. This approach assumes that the change in value of the instrument is approximated by its delta (the first derivative of the option's value with respect to the underlying variable) and its gamma (the second derivative of the option's value with respect to the underlying price). In practice, other greeks such as vega (volatility), rho (interest rate) and theta (time to maturity) can also be used to improve the accuracy of the approximation. These methods calculate the risk for a single instrument purely as a function of the current status of the instrument, in particular, its current value and sensitivities (greeks).

We present two types of analytical methods for computing VaR - the delta and delta-gamma approximation. In either case, the valuation is a monotone function of the underlying variable, and so a level

[1] Don't try to solve these quadratics in excel. Spurious answers are typical because these numbers are typically very small. Precision in vba or better is fine though.

					= provided by RiskMetrics							
Valuation Date	25-Jun-03		RM Vertice	S'nd	rfr	daily vols		Correlations				
Settlement date	30-Jun-03		30	30.00	13.500%	0.022%		100.00%	68.08%	31.57%	7.57%	38.85%
Bond	r150		90	91.00	12.625%	0.032%		68.08%	100.00%	67.07%	44.17%	59.36%
			180	182.00	11.500%	0.078%		31.57%	67.07%	100.00%	87.70%	60.08%
S-a coupon	6.000%		360	365.00	10.250%	0.135%		7.57%	44.17%	87.70%	100.00%	53.84%
Maturity	28-Feb-05		720	730.00	9.170%	0.274%		38.85%	59.36%	60.08%	53.84%	100.00%
		Term (days)	Flow	PV flow	rfr	daily vols	weights	Diagonal matrix of volatilities				
NCD	31-Aug-03	67	0.06000	0.058588	0.129693	0.000279	0.253828	0.022%	0	0	0	0
NCD	28-Feb-04	248	0.06000	0.055660	0.110492	0.000986	0.492049	0	0.032%	0	0	0
NCD	31-Aug-04	433	0.06000	0.053257	0.100488	0.001612	0.651509	0	0	0.078%	0	0
Maturity	28-Feb-05	614	1.06000	0.903245	0.095132	0.002301	0.226528	0	0	0	0.135%	0
				1.070751				0	0	0	0	0.274%
								Covariance Matrix				
	Mapped values at vertices							4.67E-08	4.7E-08	5.3E-08	2.2E-08	2.3E-07
	30	91	182	365	730			4.72E-08	1E-07	1.7E-07	1.9E-07	5.2E-07
	0.014871	0.043717						5.32E-08	1.7E-07	6.1E-07	9.3E-07	1.3E-06
			0.027388	0.028273				2.21E-08	1.9E-07	9.3E-07	1.8E-06	2E-06
				0.034698	0.018560			2.3E-07	5.2E-07	1.3E-06	2E-06	7.5E-06
				0.204610	0.698634							
	0.014871	0.043717	0.027388	0.267581	0.717194	1.070751						
PV of mapped flows												
Variance of map		0.0004876%										
Confidence		99%										
Normal multiple		2.326341928										
VaR (1 day)		0.00513682			= provided by RiskMetrics							
VaR (10 day)		0.016244061			= provided by you							

Figure 3.2: The RiskMetrics method for cash flows (for example, coupon bonds)

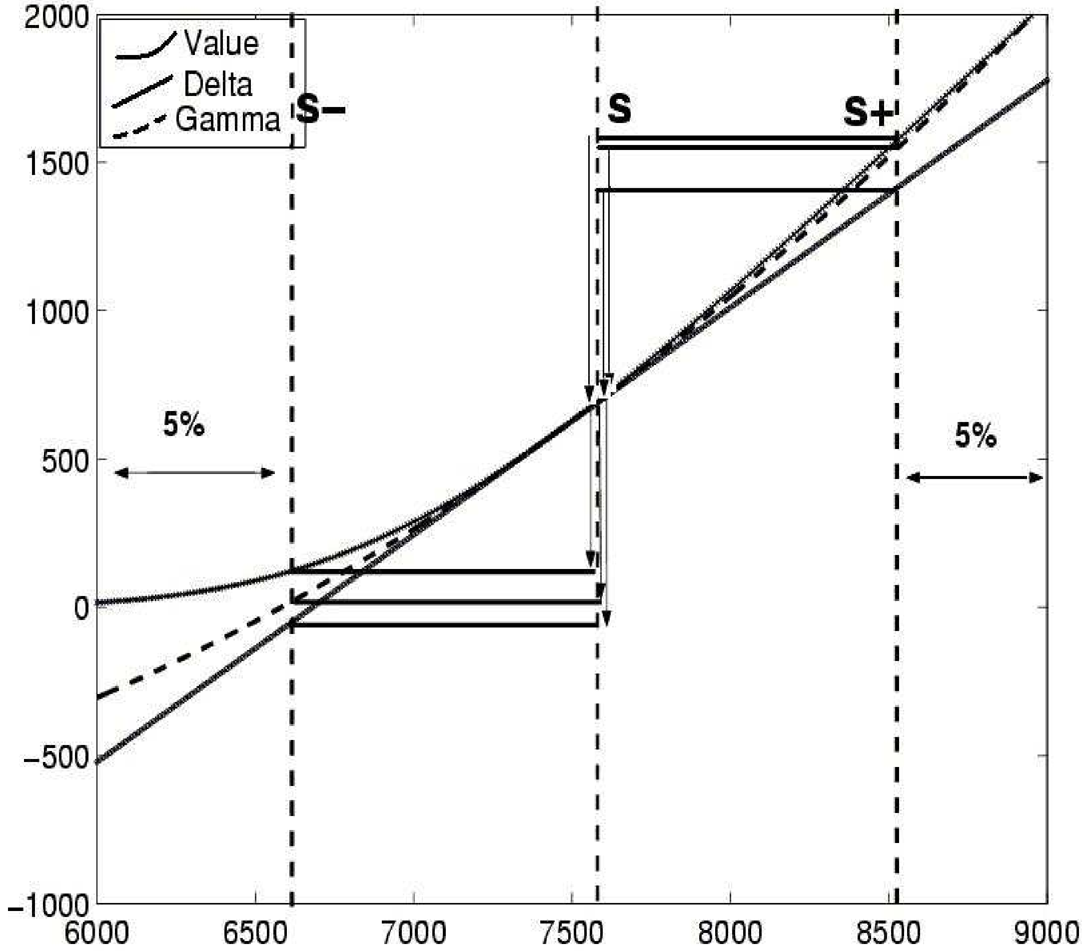


Figure 3.3: A comparison of value, the delta approximation, and the delta-gamma approximation

of confidence of that variable can be translated into the same level of confidence for the price. Note the assumption that only the underlying variable can change; other variables such as volatility are fixed.

Note from the diagram that if we are long the equity call option then the delta-gamma method gives

$$\mathrm{VaR}(V) = \Delta(S - S^-) - \tfrac{1}{2}\Gamma(S - S^-)^2$$

and if we are short the equity option then

$$\mathrm{VaR}(V) = \Delta(S^+ - S) + \tfrac{1}{2}\Gamma(S^+ - S)^2.$$

where S^- is that down value of the stock which corresponds to the confidence level required, and S^+ is that up value of the stock which corresponds to the confidence level required. The delta method would be given by the first order terms only.

The role of gamma here is quite intuitive - long gamma ensures additional profit under any market move, and so reduces the risk of the long position, conversely, it increases the risk of a short position.

Because of the sign differences in whether we are long or short, care needs to be taken with aggregation using this approach. Furthermore, this approach ignores the fact that there are other variables which impact the value of the position, such as the volatility in the case of an equity option, which in reality needs to be estimated and measured frequently. Since there is a correlation between price changes and changes in volatility, this missing factor can be significant.

Furthermore, for these methods to have any meaning at all, the price of the derivative must be a monotone function of the price of the underlying.[2]

Example 4 *Let us consider an OTC European call option on the ALSI40, expiry 17-Mar-05, strike 10,000, with the current valuation date being 15-Jan-04. The RiskMetrics method will focus on price risk exclusively and therefore ignore the risk associated with volatility (vega), interest rate (rho) and time decay (theta risk).*

The spot of the ALSI40 is 10,048, and the dividend yield for the term of the option is estimated as 3.00%. The risk free rate for the term of the option is 8.40% and the SAFEX volatility for the term is 20.50%. As mentioned, we will assume that these values do not change, and we will use the SAFEX volatility without any considerations for the skew, and allowing for this blend of exchange traded models and otc models.

The value of the position is 1,185.41, the delta is 0.638973, and the gamma 0.000158.

The daily volatility σ *of the ALSI40 is 1.30%. Thus* $S^- = Se^{-1.645\sigma} = 9,835.98$*, and* $S^+ = Se^{1.645\sigma} = 10,264.59$*. Hence, with the delta method, if we are long then*

$$VaR(V) = \Delta(S - S^-) = 135.47$$

and if we are short then

$$VaR(V) = \Delta(S^+ - S) = 138.39$$

and with the delta-gamma method, if we are long then

$$VaR(V) = \Delta(S - S^-) - \tfrac{1}{2}\Gamma(S - S^-)^2 = 131.91$$

and if we are short then

$$VaR(V) = \Delta(S^+ - S) + \tfrac{1}{2}\Gamma(S^+ - S)^2 = 142.11.$$

The delta approximation is reasonably accurate when the spot does not change significantly, but less so in the more extreme cases. This is because the delta is a linear approximation of a non linear relationship between the value of the spot and the price of the option. We may be able to improve this approximation by including the gamma term, which accounts for nonlinear (i.e. squared returns) effects of changes in the spot.

Note that in this example, how incorporating gamma changes VaR relative to the delta-only approximation.

The main attraction of such a method is its simplicity, however, this is also the problem. This approach ignores other effects such as interest rate and volatility exposure, and fits a normal distribution to data which is known not to be normally distributed. As such it will underestimate the frequency of large moves and should underestimate the 'true VaR'. This method is really only suitable for the simplest portfolios.

Despite the number of possibly tenuous assumptions, RiskMetrics performs satisfactorily well in back-testing. (Pafka & Kondor 2001) claim that this is an artifact of the choice of the risk measure: firstly that the forecasting horizon is one day, and secondly that the significance level is 95%. The first factor allows even fairly crude volatility models to perform well, and secondly the fact that the significance level is not too high means that the fat tail effect is not too severe.

[2]For example, how would one use methods like this to do calculations involving barrier options, which have traded in the South African market?

For more complicated portfolios - ones with several instruments, including options, the RiskMetric approach involves moment matching. For this, see (Zangari 1996), (Mina & Ulmer April 1999), (Pichler & Selitsch 2000) and (Hull 2002, Chapter 16).

3.2 Classical historical simulation

The historical method is a full revaluation method. The revaluation of the entire portfolio is calculated for each of the last N days, as if the evolution in market variables that occurred on each of those days was to reoccur now. Thus,

$$\ln \frac{x^i}{x(t)} = \ln \frac{x(i)}{x(i-1)} \tag{3.7}$$

or

$$x^i = x(t) \cdot \frac{x(i)}{x(i-1)} \tag{3.8}$$

would be the market factor update formula for the variable x. Here t denotes the current day, i one of the past business days, $t - N + 1 \leq i \leq t$, and $i - 1$ the business day before that.

Example 5 *Suppose on 22-Jan-04 we are long a r153 bond. We perform 400 historical simulations on the ytm. We then apply the bond pricing formula ie. full revaluation to the ytms so obtained to get the all in price of the bond on 28-Jan-04. This is FOUR business days after 22-Jan-04, which is three days after the next business date.*

We get from full revaluation 400 bond prices: a minimum of 1.2316, a 5^{th} *percentile of 1.2387, an average of 1.2440, a* 95^{th} *percentile of 1.2497, and a maximum of 1.2600. Thus if we are long the bond then the 95% VaR is 0.0054 per unit, and if we are short then the 95% VaR is 0.0057 per unit.*

We could also simply determine the appropriate percentile ytm and calculate the AIP there, to get the same results. However, this does not help as soon as we start aggregation.

Example 6 *Let us consider an OTC European call option on the ALSI40; expiry 20-Mar-03, strike 12000, with the current valuation date being 19-Jun-02. We perform 400 historical simulations on the spot, on the risk free rate, and on the atm volatility, valuing on 20-Jun-02. We stress the dividend yield in the reverse direction of the spot stress in such a manner that the monetary value of the dividends is constant.*

We get from full revaluation 400 option prices: a minimum of 294.59, a 5^{th} *percentile of 400.20, an average of 468.22, a* 95^{th} *percentile of 551.17, and a maximum of 754.32. Thus if we are long the option then the 95% VaR is 68.02, and if we are short then the 95% VaR is 82.95.*

3.3 Historical simulation with volatility adjusting

This method was first proposed in (Hull & White 1998), and has become quite prevalent academically. It has not been widely implemented in the industry, although it is starting to gain some prominence in South Africa. One of the main critisms of the historical method is that the returns of the past can be inappropriate for current market conditions. For example, if our window of N days is an almost entirely quiet period, and there is currently a very sudden spike in volatility, the historical method would still be using the 'quiet data', and the new volatility regime would only be factored in gradually, one day at a time.

The Historical V@R method accurately reflects the historical probability distribution of the market variables, which is an attractive feature especially in markets where the normality assumption is far from reality. However, historical V@R's main disadvantage is that it incorporates no volatility

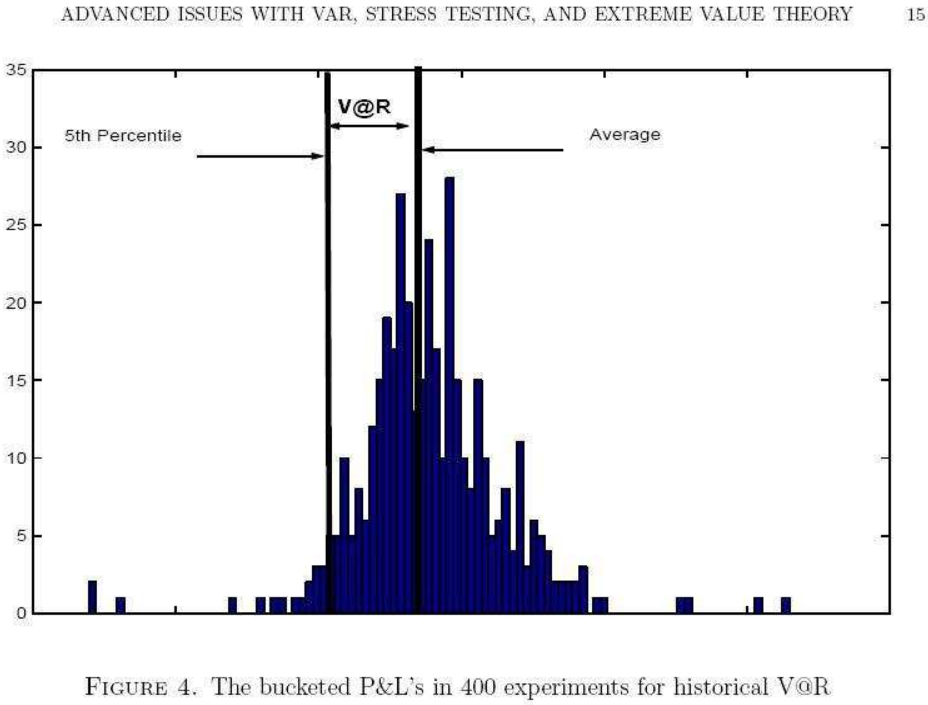


Figure 3.4: The bucketed values of the instrument in 400 experiments for historical V@R

updating, in that it assumes the distribution of returns is stationary. As we will see, the Hull-White method is a modification of the historical method that overcomes this difficulty.

If the current volatility of a market variable is 30% per day and a month ago was 15%, the returns a month ago understate the returns we expect to see now. The Hull-White method adjusts the historical data on each market variable to reflect the difference between the old historical volatility of the market variable and its current historical volatility, so in the above example, it doubles the return that was observed.

The basic idea of the volatility adjusting is that we should only compare standardised variables, which have been standardised by dividing by their volatility. Thus

$$\frac{1}{\sigma(t)} \ln \frac{x^i}{x(t)} = \frac{1}{\sigma(i-1)} \ln \frac{x(i)}{x(i-1)}$$

or

$$x^i = x(t) \cdot \left(\frac{x(i)}{x(i-1)} \right)^{\frac{\sigma(t)}{\sigma(i-1)}}$$

would be the experimental values for the factor x, indexed by the value i where $t - N + 1 \leq i \leq t$. The volatility is historical volatility, unless a reliable implied volatility is available, in which case it is implied volatility.

An appropriate method for implied volatility updating is required. If exactly the same strategy is to be used one will need to measure and adjust by the volatility of volatility. But mathematically one cannot use any historical volatility calculation scheme - such as EWMA for example - as implied volatility does not follow a (Geometric Brownian Motion) random walk, but is mean reverting. We prefer just to use straight historical for implied volatility. Thus, for implied volatility σ_I:

$$\sigma_I^i = \sigma_I(t)\frac{\sigma_I(i)}{\sigma_I(i-1)}$$

and we have three market factor update formulae:

$$x^i = x(t) \cdot \left(\frac{x(i)}{x(i-1)}\right)^{\frac{\sigma(t)}{\sigma(i-1)}} \tag{3.9}$$

$$x^i = x(t) \cdot \left(\frac{x(i)}{x(i-1)}\right)^{\frac{\sigma_I(t)}{\sigma_I(i-1)}} \tag{3.10}$$

$$\sigma_I^i = \sigma_I(t)\frac{\sigma_I(i)}{\sigma_I(i-1)} \tag{3.11}$$

where

- (3.9) is used where the variable x is available and implied volatility is not, so an historical volatility is calculated;
- (3.10) is used where the variable x is available and a reliable estimate of implied volatility is too (for example, a futures level);
- (3.11) is used on an implied volatility variable.

Very often implied volatilities are suspicious, due to being stale or illiquid, and in this case, the historical volatility should be preferred i.e. (3.9) should be preferred to (3.10).

Example 7 *Let us consider the same OTC European call option on the ALSI40 as before: expiry 20-Mar-03, strike 12,000, with the current valuation date being 19-Jun-02. We perform 400 historical simulations with volatility adjusting on the spot, on the risk free rate, and simple historical simulations on the atm volatility, valuing on 20-Jun-02. We stress the dividend yield as previously.*

We get from full revaluation 400 option prices: a minimum of 316.66, a 5^{th} *percentile of 406.19, an average of 466.72, a* 95^{th} *percentile of 540.30, and a maximum of 717.58. Thus if we are long the option then the 95% VaR is 60.53, and if we are short then the 95% VaR is 73.59.*

In a personal communication, Alan White says "I always liked that [the Hull-White] scheme. My view is that the various approaches form a continuum in which different methods are used to characterize the distributions in question. The parametric approach tries to match moments. It can evolve quickly but fails to capture many of the details of the distributions. The historical simulation assumes the sample distribution is the population distribution. This captures the details of the distribution but evolves too slowly if the distribution is not stationary. We attempted to marry these two approaches."

An example of a time series of p&l's and (minus) the VaRs under various methods is as in Figure 3.5. This shows, for example, how the historical method is completely inadequate: at the time of a market crash, the VaR measure does not jump, as logically it should.

3.4 Monte Carlo method

The second alternative offered by RiskMetrics, structured Monte Carlo simulation, involves creating a large number of possible rate scenarios and performing full revaluation of the portfolio under each of these scenarios.

VaR is then defined as the appropriate percentile of the distribution of value changes. Due to the required revaluations, this approach is computationally far more intensive than the analytic RiskMetrics approach. The two RiskMetrics methods - analytic and Monte Carlo - differ not in terms of how

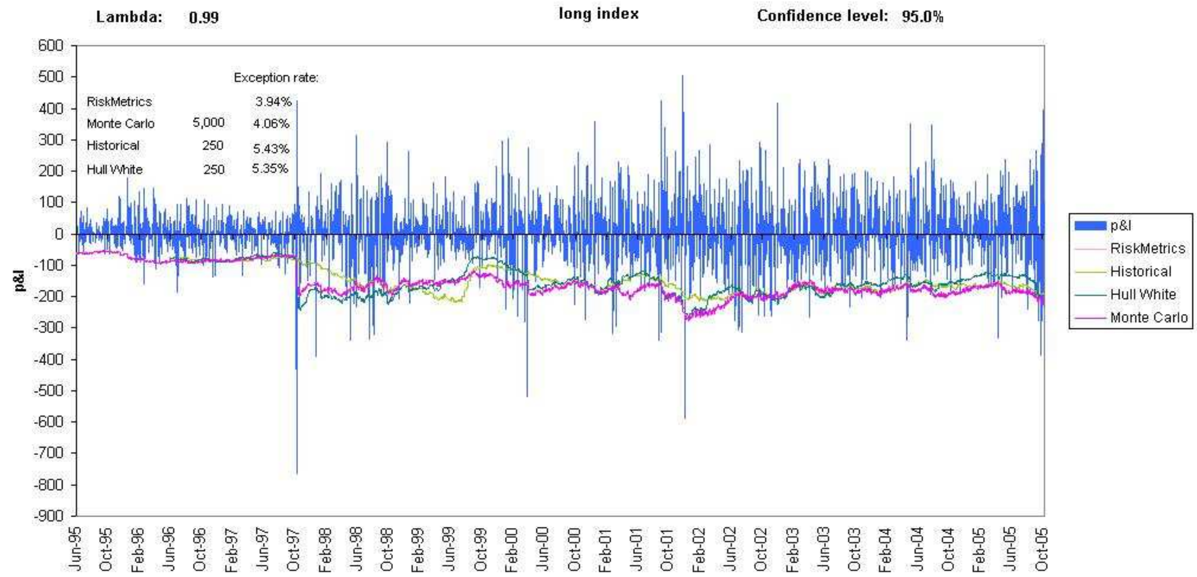


Figure 3.5: p&l's and VaR for a long ALSI40 position

market movements are forecast (since both use the RiskMetrics volatility and correlation estimates) but in how the value of portfolios changes as a result of market movements. The analytical approach approximates changes in value, while the structured Monte Carlo approach fully revalues portfolios under various scenarios.

The RiskMetrics Monte Carlo methodology consists of three major steps:

- Scenario generation, using the volatility and correlation estimates for the market factors which drive our portfolio, we produce a large number of future price scenarios in accordance with the lognormal models.
- For each scenario, we compute instrument (molecule) values and then portfolio values.
- We report the results of the simulation, either as a portfolio distribution or as a particular risk measure.

Other Monte Carlo methods may vary the first step by creating returns by (possibly quite involved) modelled distributions, using pseudo random numbers to draw a sample from the distribution. The next two steps are as above. The calculation of VaR then proceeds as for the historical simulation method. Indeed, this is very similar to the historical method except for the manner in which experiments are created.

The advances in RiskMetrics Monte Carlo is that one overcomes the pathologies involved with approximations like the delta-gamma method.

The advances in other Monte Carlo methods over RiskMetrics Monte Carlo are in the creation of the distributions. However, to create experiments using a Monte Carlo method is fraught with dangers. Each market variable has to be modelled according to an estimated distribution and the relationships between distributions (such as correlation or less obvious non-linear relationships, for which copulas are becoming prominent). Using the Monte Carlo approach means one is committed to the use of such distributions and the estimations one makes. These distributions can become inappropriate; possibly in an insidious manner. To build and 'keep current' a Monte Carlo risk management system requires continual re-estimation, a good reserve of analytic and statistical skills, and non-automatic decisions.

Example 8 *Suppose we hold an r153 bond on 22-Jan-04. What is the VaR?*

The close was 9.10%. We estimate the annual volatility of the yield to be 12.25%. Using excel/vba, we first create uniformly distributed random numbers U then transform them into normally distributed random numbers Z by using the inverse of the cumulative normal distribution.[3] *We then determine our new yields: $y(T+1) = y(T)\exp\left(\frac{\sigma}{\sqrt{250}}Z\right)$. We then apply the bond pricing formula for 28-Jan-04 to get the new all in prices. We then work out the VaR, by examining averages and percentiles, in the usual way. The 95% VaR is about 6,417 (long) and 6,473 (short) per unit.*

Choleski decomposition

Suppose we are interested in a portfolio with more than one security, or more generally, more than one source of random normal noise. Let us start with the case where we have two such random variables. We cannot simply take two random number generators and paste them together, unless the underlyings are independent. However, typically there will be a measured or estimated correlation between the two random variables, and this needs to appear in the random numbers generated.

If the two stocks were uncorrelated, we could have

$$r_1 = a_1 Z_1, \quad r_2 = a_2 Z_2$$

With the correlation, we want the Z_1 to influence r_2. Thus the appropriate setup is

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} a_{1,1} & 0 \\ a_{2,1} & a_{2,2} \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}. \tag{3.12}$$

or $r = AZ$. Thus

$$rr' = AZZ'A' \tag{3.13}$$

and so

$$\Sigma = AA' \tag{3.14}$$

by taking expectations. Thus, A is found as a type of lower-triangular square root matrix of the known variance-covariance matrix Σ. The most common solution (it is not unique) is known as the Choleski decomposition. All that has been said is valid for any number of dimensions, and simple algorithms for calculating the Choleski decomposition are available (Burden & Faires 1997, Algorithm 6.6).

In the case of two variables, it is convenient to explicitly note the solution. Here

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{bmatrix} \tag{3.15}$$

and

$$A =^{4} \begin{bmatrix} \sigma_1 & 0 \\ \sigma_2\rho & \sigma_2\sqrt{1-\rho^2} \end{bmatrix} \tag{3.16}$$

A theoretical requirement here is that the matrix Σ be positive semi-definite. The covariance matrix is in theory positive definite as long as the variables are truly different ie. we do not have the situation that one is a linear combination of the others (so that there is some combination which gives the 0 entry). If there are more assets in the matrix than number of historical data points the matrix will be rank-deficient and so only positive semi-definite. Moreover, in practice because all parameters are estimated, and in a large matrix there will be some assets which are nearly linear combinations of others, and also taking into account numerical roundoff, the matrix may not be positive semi-definite at all (Dowd 1998, §2.3.4). However, this problem has recently been completely solved (Higham 2002), by mathematically finding the (semi-definite) correlation matrix which is closest (in an appropriate norm) to a given matrix, in particular, to our mis-estimated matrix.

[3] In excel this is given by the function norminv.

[4] Note the error in (Dowd 1998, Chapter 5 §2.2).

Example 9 *We reconsider the example in Example 2. You hold 2,000,000 shares of SAB and 500000 shares of SOL. SOL is trading at 105.20 with a volatility of 32.10%. The correlation in returns is 4.14%. What is your 95% VaR over a 1-day horizon on 23-Jan-04?*

Using excel/vba, we first extract pairs of uniformly distributed random numbers U_1, U_2, then transform them into pairs of normally distributed random numbers Z_1, Z_2 by using the inverse of the cumulative normal distribution. We then apply the Choleski decomposition:

$$r_1 = \frac{\sigma_1}{\sqrt{250}} Z_1, \quad r_2 = \frac{\sigma_2}{\sqrt{250}} (\rho Z_1 + \sqrt{1-\rho^2} Z_2) \tag{3.17}$$

and determine our new prices: $S_1(T+1) = S_1(T)\exp(r_1)$, $S_2(T+1) = S_2(T)\exp(r_2)$. We then work out the portfolio MtF's, and then work out the VaR, by examining averages and percentiles, in the usual way. The 95% VaRs are 3,979,192 and 4,058,332.

A possible example of the first few calculations is shown in Table 3.1. The calculation would typically use 10,000 calculations or more.

Rnd(1)	Rnd(2)	Cumnorm inverse(1)	Cumnorm inverse(2)	Correlated return(1)	Correlated return(2)	MtF(1)	MtF(2)	New portfolio MtM
0.7055	0.5334	0.5404	0.0839	0.0083	0.0022	71.49	105.43	195,696,457
0.5795	0.2896	0.2007	-0.555	0.0031	-0.011	71.12	104.04	194,258,373
0.3019	0.7747	-0.519	0.7545	-0.008	0.0149	70.34	106.78	194,061,564
0.014	0.7607	-2.197	0.7086	-0.034	0.0125	68.55	106.53	190,354,401
0.8145	0.709	0.8946	0.5506	0.0138	0.0119	71.88	106.46	196,994,225
0.0454	0.414	-1.692	-0.217	-0.026	-0.006	69.08	104.59	190,454,297
0.8626	0.7905	1.0922	0.8081	0.0168	0.0173	72.10	107.04	197,719,253

Table 3.1: The first few experiments under a bivariate Monte Carlo run

Example 10 *Consider on 22-Jan-04 a portfolio which consists of long 110,000,000 r153 and short 175,000,000 tk01. The closes of these instruments are 9.10% and 9.41% respectively, with AIP for 27-Jan-04 being 1.2434199 and 1.0524115 respectively, and delta -5.464 and -3.433 respectively. Thus this is an almost delta neutral portfolio; the risks associated should be quite small.*

We have $\sigma_1 = 12.25\%$, $\sigma_2 = 15.18\%$, and $\rho = 91.25\%$.

Using excel/vba, we first extract pairs of uniformly distributed random numbers U_1, U_2, then transform them into pairs of normally distributed random numbers Z_1, Z_2 by using the inverse of the cumulative normal distribution. We then apply the Choleski decomposition:

$$r_1 = \frac{\sigma_1}{\sqrt{250}} Z_1, \quad r_2 = \frac{\sigma_2}{\sqrt{250}} (\rho Z_1 + \sqrt{1-\rho^2} Z_2) \tag{3.18}$$

and determine our new yields: $y_1(T+1) = y_1(T)\exp(r_1)$, $y_2(T+1) = y_2(T)\exp(r_2)$. We then apply the bond pricing formula to get the new all in prices. We then work out the VaR, by examining averages and percentiles, in the usual way. The 95% VaRs are 296,964 and 291,291.

Another very effective and computationally very efficient way around this problem is to reduce the dimensions of the problem by using principal component analysis or factor analysis. Principal component analysis is a topic on its own, and has become very prevalent in financial quantitative analysis. See (Dowd 1998, Box 3.3).

3.5 Summary

Here we summarise the essential features of the competing methods.

	RiskMetrics	Historical	Hull-White	Monte Carlo
Revaluation	analytic	full	full	full
Distributions	normal	actual	quasi-actual	created
Tails	thin	actual	quasi-actual	created
Intellectual effort	moderate	very low	low	very high
Model risk	enormous	moderate	low	high
Computation time	low	moderate	moderate	high
Communicability	easy	easy	moderate	very difficult

In the same correspondence as previously, Alan White says "In my experience in North America the historical simulation approach has won the war. Some institutions use the parametric approach (particularly those with large portfolios of exotics) but they appear to be in the minority. I don't know anyone who uses the approach we suggested despite its advantages. Perhaps the perception is that the improvement in the measures does not compensate for the cost of implementing the procedure. Just maintaining the historical data base seems to tax the capabilities of many institutions.

As for software vendors [implementing the Hull-White method], my sense is that this market segment (VaR systems) is now a mature market with thin profits for the software companies. It seems unlikely to me that they will be implementing many changes in this environment."

Chapter 4

Stress testing and Sensitivities

4.1 VaR can be an inadequate measure of risk

VaR is generally used as a quantitative measure for how severe losses could be. Yet, significant catastrophes have been evident, even in instances where state-of-the-art VaR computations have been deployed. For example, VaR calculations were conducted prior to the implosion in August 1998 by Long-Term Capital Management (LTCM). What went wrong with LTCM risk-forecasts? They may have placed too much faith on their exquisitely tuned computer models. Sources say LTCMs worst-case scenario was only about 60% as bad as the one that actually occurred. In other words, stress testing was inadequate. In fact, it seems that stress-testing was almost non-existent at LTCM; most risk-measurement was done using VaR methods. The problem has, at its basis, LTCMs inability to accurately measure, control and manage extreme risk. It is extreme risk that LTCMs VaR calculations could not accurately estimate, and it is extreme risk that needs to be measured in stress testing.

Stress tests can provide useful information about a firm's risk exposure that VaR methods can easily miss, particularly if VaR models focus on "normal" market risks rather than the risks associated with rare or extreme events. Such information can be fed into strategic planning, capital allocation, hedging, and other major decisions.

Stress testing is essential for examining the vulnerability of the institution to unusual events that plausibly could happen (but have not previously happened, so are not 'inputs' to our VaR model) or happen so rarely that VaR 'ignores' them because they are in the tails. Thus, market crashes are typically washed into the tails, so VaR does not alert us to their full impact. Thus stress testing is a necessary safeguard again possible failures in the VaR methodology.

Scenarios should take into account the effects that large market moves will have on liquidity. Usually a VaR system will assume perfect liquidity or at least that the existing liquidity regime will be maintained.

The results of scenario analysis should be used to identify vulnerabilities that the institution is exposed to. These should be actioned by management, retaining those risks that they see as tolerable. (It is impossible to remove all risks because by doing so the rewards will also disappear.)

4.2 Stress Testing

There are various types of stress analysis (Wee & Lee 1999), (Fender, Gibson & Mosser November 2001):

- The first type uses scenarios from recent history, such as the 1987 equity crash. We can ask what the impact would be of some historical market event, such as a market crash, repeating itself.

- Institution-specific scenario analysis. Identify scenarios based on the institution's portfolio, businesses, and structural risks. This seeks to identify the vulnerabilities and the worst-case loss events specific to the firm.
- Extreme standard deviation scenarios. Identify extreme moves and construct the scenarios in which such losses can occur. For example, what will be the losses in a 5 - 10 standard deviation event?
- Predefined or set-piece scenarios that have proven to be useful in practice. The risk manager should also be able to create plausible scenarios.
- Mechanical-search stress tests, also called sensitivity stress tests (Fender et al. November 2001), (Hosoya & Shimizu December 2002). This can be performed fairly mechanically. Key variables are moved one at a time and the portfolio is revalued under those moves. What results is a vector or matrix of portfolio revaluations under the market moves.

  Any market modelling required for these purposes is usually fairly routine. For example, when stressing the 'price level' of the equity market, individual stocks may be stressed in a manner consistent with the Capital Asset Pricing Model (Sharpe 1964) ie.

  $$\frac{dS}{S} = \alpha + \beta \frac{dI}{I}$$

  where S denotes the stock and I the index, and the CAPM parameters α and β are of the stock w.r.t. to index. This ensures that the volatility dispersion within the portfolio is modelled. Furthermore, α can be eliminated from the above equation because it is usually insignificant when compared with the size of stress that we are interested in.
- Quantitative evaluation of distributions of tail events and extreme value theory. Based on observed historical market events, quantify the impact of a series of tail events to evaluate the severity of the worst case losses. This approach also evaluates the distribution of tail events to determine if there are any patterns that should be used for scenario analysis.

4.3 Other risk measures and their uses

- Stress testing.
- Greeks or sensitivities - the favourite of dealers, because it is on this basis that they will manage the book.
  - aggregate delta of an equity option portfolio,
  - aggregate rho of a fixed income portfolio,
  - gamma or convexity.
- key rate duration - shifting pieces of the yield curve, or even other term structures such as volatility. This is therefore a type of scenario analysis.
- Cash ladders - asset liability management.
- Stop-loss limits.

Many of these measures are much 'lower-level' than V@R. They provide information only about a limited subset of a portfolios risk, and illuminate the specific contributors to risk. It gives the directional impact of each risk - 'the feel of the risks'.

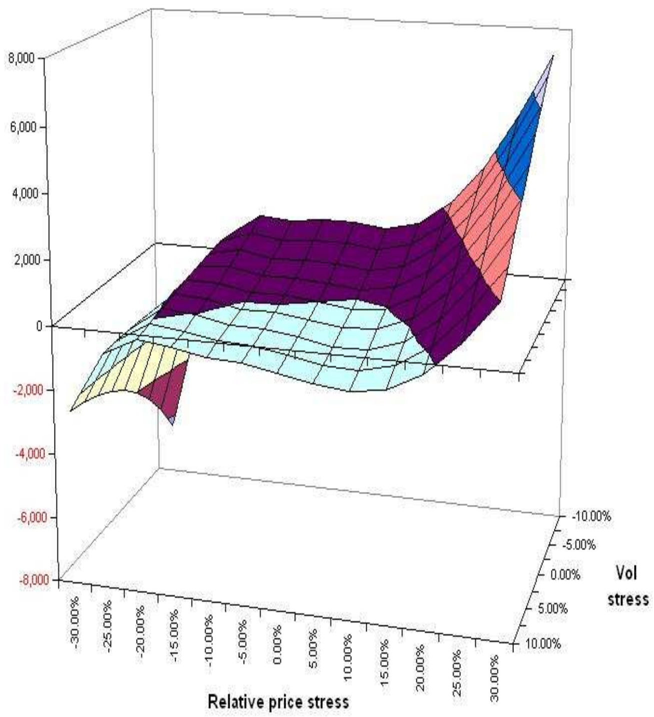


Figure 4.1: A price-volatility stress matrix

This will help a risk manager understand where the risks come from and what can be done to lower it (if necessary). VaR can be (although it does not have to be) a type of 'black box'. In this regard, it should be noted that there is a definite distinction between risk measurement and risk management. Risk measurement is a natural consequence of the ability to price instruments and manage the data associated with that pricing, and is best performed with a blend of analytic and IT skills. Risk management is the process of considering the business reasons and intuitions behind the risk measures and then acting upon them. Here business skills and plain obstinacy are most appropriate.

All risk measures can be equally valid and can be aimed at different audiences, for example:

- Greeks are for the dealer,
- stress and VaR for management,
- VaR (supplemented by stress testing) are for the regulator.

The uses of these risk measures can be inferred in a logical manner in terms of the limit hierarchy that is placed on the business.

- Individual dealers, following explicit strategies, should have their limits set via the Greeks,
- stress and VaR limits should be placed from the desk level, up to the level of the entire institution,
- the stress and VaR figures for the entire institution are provided to the regulator.

4.4 Calculating analytic Greeks

4.4.1 Interest rate instruments

By this we mean instruments such as JIBAR instruments, bonds, FRAs and swaps, all of which do not have a volatility input.

(1) Suppose we have any set of fixed cash flows, such as a JIBAR instrument or a bond. A FRA also falls into this category, because of the way in which it can be mapped as a fixed long and a fixed short cash flow. See (West 2006, §1.7). Thus, we have

$$V_{\text{flows}} = \sum_{i=1}^{n} c_i e^{-r_i \tau_i}$$

for some cash flows $c_1,\ c_2,\ \ldots,\ c_n$ and some terms $\tau_1,\ \tau_2,\ \ldots,\ \tau_n$. $r_1,\ r_2,\ \ldots,\ r_n$ are the NACC rates for the terms $\tau_1,\ \tau_2,\ \ldots,\ \tau_n$.

Let us now calculate $\frac{dV}{d\overline{r}}$, the derivative w.r.t. parallel shifts in the yield curve. Clearly

$$\frac{dV}{d\overline{r}} = \sum_{i=1}^{n} -\tau_i c_i e^{-r_i \tau_i} \tag{4.1}$$

We will also be interested in the sensitivity of instruments to movements in particular parts of the yield curve - we will want to bucket the ρ exposures. In this case it is straightforward: for each bucket, we simply take the summation over the flows that occur in that particular bucket.

We can also calculate the decay in value due to time:

$$\frac{dV}{d\tau} = \sum_{i=1}^{n} -r_i c_i e^{-r_i \tau_i} \tag{4.2}$$

Note that in this derivative, τ is a variable which increases; in order to obtain the theta, we would negate this quantity.

(2) Let us consider a just issued swap. The value of the fixed payments is

$$V_{\text{fix}} = R \sum_{i=1}^{n} \alpha_i Z(t, t_i) \tag{4.3}$$

where R is the agreed fixed rate (known as the swap rate), n is the number of payments outstanding, and α_i is the length of the i^{th} 3 month period on an actual/365 basis. This valuation formula holds whether or not today t is a reset date.

Although notionals are not exchanged at termination, let us imagine that they are. Then the value of the fixed payments become

$$V_{\text{fix}} = R \sum_{i=1}^{n} \alpha_i Z(t, t_i) + Z(t, t_n) \tag{4.4}$$

Given that the notional on the floating leg is now paid, the floating leg is just a floating rate note, so it has a value of 1: it is not exposed to any interest rate risks.

Now let us suppose that the swap is under way. suppose the period under way is of length α_1 and the rate has been fixed at time $t_0 < t$ to be r_1^{fix}. Effectively the floating leg is now a fixed payment at t_1 and the creation of a notional 1 floating rate note at that time. Hence it has value

$$V_{\text{float}} = Z(t, t_1) \left[1 + \alpha_1 r_1^{\text{fix}}\right] \tag{4.5}$$

Finally, suppose that the swap is forward starting; suppose the forward observation date is $t_0 > t$, payments will commence at t_1. Then

$$V_{\text{float}} = Z(t, t_0) \tag{4.6}$$

as the floating payments are just the creation of a floating rate note at time t_0.

Thus we have rewritten both the fixed and the floating legs of the swap as a portfolio of known cash flows. Calculation of greeks and bucketing will be as before.

(3) For options, bucketing of the risks will need to be performed.

4.4.2 Equity Forwards

Possibly the only case where analytic Greeks for equity derivatives can be used is for forwards:

$$V_{\text{forward}} = S - Q - e^{-r\tau} X$$

where Q is the present value of dividends. See (West 2006, §8.5). Then $\Delta = 1$ and $\rho = \tau e^{-r\tau} X$.

4.5 Numeric greeks for equity instruments

4.5.1 What is a numeric sensitivity?

Where analytic formulae hold true these can be used for calculation of sensitivities. For the most part however, analytic Greek formulae do not hold true. In this case, we need to estimate the sensitivities numerically.

See Figure 4.2. The horizontal axis is spot, so the gradient of the curve is the delta curve. At a spot value of 90, delta is the slope of the soft line, which is tangent to the graph at 90. Alternatively, we could estimate the delta to be the gradient of the darker line. This is the line which goes through the point $(80, V(80))$ and $(100, V(100))$. This line has a gradient of $\frac{V(100) - V(80)}{100 - 80}$. Providing we have a formula for the value function $V(\cdot)$ this is easily calculated, even if we do not have a formula for the slope.

This numerical Greek has been calculated with what we will call a 'twitch' of 10 (Rands, or points). The Greek is $\Delta = \frac{V(S+\epsilon) - V(S-\epsilon)}{2\epsilon}$ where here ϵ is 10. Typically we will make ϵ much smaller. Also, we will make the changes relative, not absolute: a twitch of 10 doesn't make much sense when the spot is 5! So we calculate as follows:

$$\begin{aligned}\Delta &= \frac{V(S+\epsilon S) - V(S-\epsilon S)}{(S+\epsilon S) - (S-\epsilon S)} \\ &= \frac{V(S+\epsilon S) - V(S-\epsilon S)}{2\epsilon S}\end{aligned}$$

In this example, it seems that only S is moving. In reality, when S moves, other variables might move too, and that needs to be taken into account in this calculation. Such numeric Greeks are sometimes called 'shadow' Greeks e.g. shadow gamma (Taleb 1997).

Numeric Greeks are necessary when there is a skew

The skew can be summarised with the statement that, for otherwise identical options, at different strikes there will be different volatilities in use, which are then plugged into the relevant Black (-Scholes or SAFEX-) equation. Furthermore, when stock price moves there will be a simultaneous

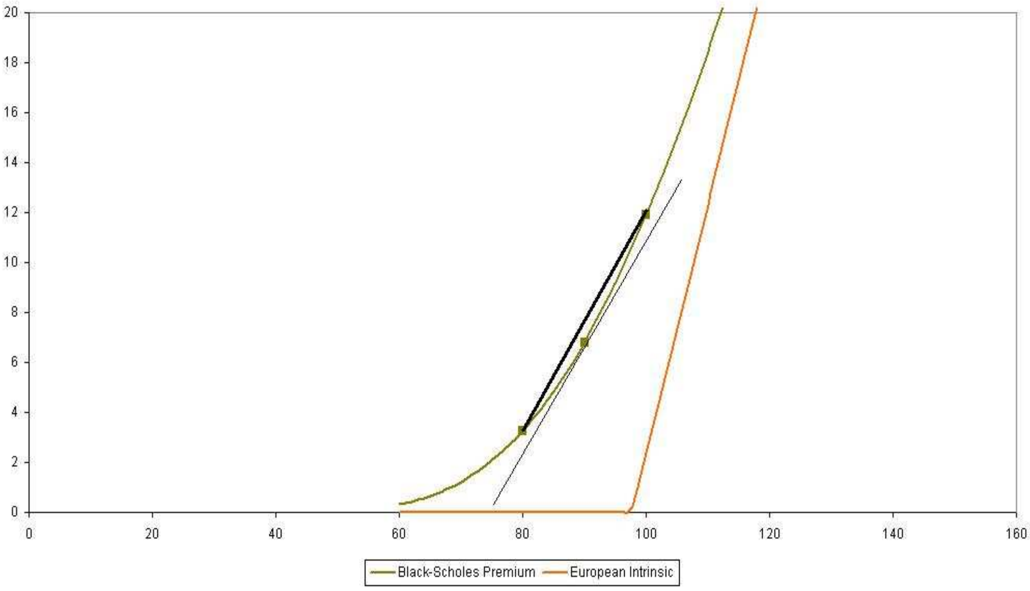


Figure 4.2: Numerical calculation of the delta of an option

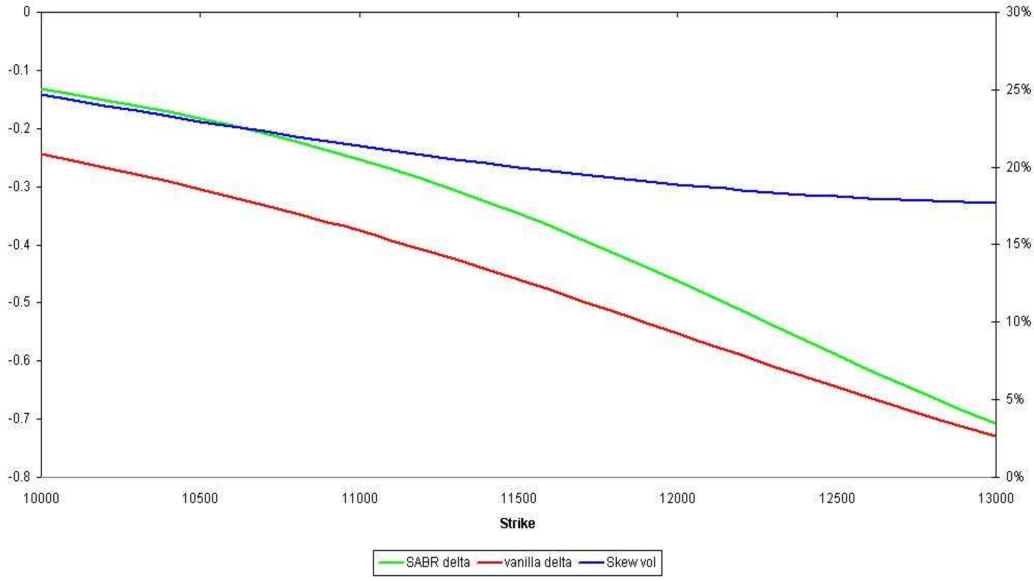


Figure 4.3: The delta under a flat model and under a stochastic volatility model for different strikes

move in the skew. Thus it is erroneous to model volatility as constant. The extent to which volatility moves as a consequence of spot/futures moving can be exactly modelled within a stochastic volatility model, for example.

If the futures level moves, then the skew model predicts the skew volatility will also move. The skew model will price the option taking into account the movement in the futures level AND the movement in the skew volatility. This could have a dramatic effect on the delta.

Numeric Greeks are necessary when there is a dividend yield

Another example where analytic Greeks fail is where an equity option price includes a dividend yield. The typical calculation of Greeks such as Delta would proceed under the assumption that the dividend yield remains constant, as we have seen in the Black-Scholes equation. However, in reality this is not the case. All other things being equal, when the stock price moves up, the dividend yield will move down. In fact, for moderate moves in stock (as is the case here) it would be better to model that the present value of short term dividends (that means, perhaps, the dividends during the life of the option) remains constant. The impact of making the erroneous assumption that the dividend yield is a constant can be quite material.

For example, if we assume that the present value of dividends is the constant, then $\Delta = \eta N(\eta d_1)$ (although now of course assuming that there is no skew)! Thus the error is of order $e^{-q\tau}$, which is very material where the dividends are significant or the term is long.

4.5.2 Algorithms for calculating numeric sensitivities

(a) Delta, $\Delta = \frac{\partial V}{\partial S}$, where S denotes the price of the underlying. Numerically, Delta is found using central differences as

$$\Delta = \frac{V((1+\epsilon)S) - V((1-\epsilon)S)}{2\epsilon S} \tag{4.7}$$

Δ gives the change in value of V for a one point (rand or index point) change in S.

Because Δ cannot be aggregated across different underlyings, it is not as useful as ΔS. From the above, $\epsilon\Delta S$ is to first order the profit on V for a change of ϵS to the value of S. Thus $\pm.01\Delta S$ is the profit on a 1% move up/down in the underlying. ΔS is the rand equivalent delta, and can be aggregated across different underlyings.

(b) Gamma, $\Gamma = \frac{\partial^2 V}{\partial S^2}$. Numerically, Γ is found as

$$\Gamma = \frac{V((1+\epsilon)S) - 2V(S) + V((1-\epsilon)S)}{\epsilon^2 S^2} \tag{4.8}$$

$\epsilon\Gamma S$ is the approximate change in Δ requirements for a change of ϵS to the value of S i.e. the number of additional S needed for rebalancing the hedge.[1] This rebalancing will cost about $\epsilon\Gamma S^2$.

ΓS^2 is used for measuring the notional cost of rebalancing the hedge, and is the rand equivalent Gamma. $\pm 0.01\Gamma S^2$ is the notional cost of rebalancing the hedge under a 1% move up/down in S.

(c) Vega is defined to be $\frac{\partial V}{\partial \sigma_{\text{atm}}}$ ie. the sensitivity to changes in the level of the SAFEX atm term structure. This is found numerically using central differences as

$$\mathcal{V} = \frac{V(\sigma_{\text{atm}} + \epsilon) - V(\sigma_{\text{atm}} - \epsilon)}{2\epsilon} \tag{4.9}$$

where σ_{atm} denotes the entire SAFEX atm volatility term structure of the index; the shift of ϵ is made parallel to the entire term structure.

$\pm 0.01\mathcal{V}$ gives the profit on a 100bp move up/down in the SAFEX atm term structure.

(d) Theta, $\Theta = \frac{\partial V}{\partial t}$, is an annual measure of the time decay of V. We calculate

$$\Theta = \frac{V(t+\epsilon) - V(t)}{\epsilon} \cdot 365 \tag{4.10}$$

Whether we now multiply by $\frac{1}{250}$ to get the time value gain on V per trading day, or by $\frac{\text{nbd}(t)-t}{365}$ to get the time value between date t and the next business day, or simply by $\frac{1}{365}$, is largely a matter of choice.

[1] Since, by Taylor series, $\Delta(S+\epsilon S) = \Delta(S) + \Gamma\epsilon S + \cdots$.

Also, there are two different possible assumptions about what happens to S as time moves forward. The one is that S remains constant, so θ is calculated without moving S. The other is that S rolls up the forward curve, so in the shadow calculation we replace S with $Se^{r\epsilon}$, and also take into account any dividend payments.

(e) Rho, $\rho = \frac{\partial V}{\partial r}$, where r denotes the input risk free rate NACC. These are found from a standard NACC yield curve. In the case of multiple input risk free rates, ρ is the sensitivity to a simultaneous (parallel) shift in the entire term structure. ρ is found numerically using central differences as

$$\rho = \frac{V(r+\epsilon) - V(r-\epsilon)}{2\epsilon} \tag{4.11}$$

$\pm 0.01\rho$ gives the profit on a 100bp NACC parallel move up/down in the term structure.

4.5.3 Taylor series

Consider a first order (delta-gamma-rho-vega-theta) Taylor series expansion as follows:

$$dV \quad \simeq \quad \Delta\,\delta S + \tfrac{1}{2}\Gamma\,(\delta S)^2 + \rho\,\delta r + \mathcal{V}\,\delta\sigma_{\text{atm}} + \theta\,\delta t$$

This expansion allows us to attribution our p&l according to the sensitivities, and this enables us to analyse what type of bets the dealer is making.

For a simple instrument, such as an equity derivative, we can attribute the p&l as

- $\Delta_{t-1}\,(S_t - S_{t-1})$ is the p&l due to delta;
- $\frac{1}{2}\Gamma_{t-1}\,(S_t - S_{t-1})^2$ is the p&l due to gamma;
- $\rho_{t-1}\,(r_t - r_{t-1})$ is the p&l due to rho;
- $\mathcal{V}_{t-1}\,(\sigma_{\text{atm},t} - \sigma_{\text{atm},t-1})$ is the p&l due to vega;
- $\theta_{t-1}\,\delta t$ is the p&l due to theta;
- the remainder is the p&l due to error in the Taylor series expansion.

At regular intervals we should check that the error term is not material. Of course, we can attribute a percentage to this error term, which should not be more than a couple of percent. After all, the error term is a measure of how well the Taylor series expansion is fitting the actual p&l. As expected, for more complicated products, these errors can be more material, and the method should not be used. Alternatively, a higher order Taylor series expansion could be derived and the appropriate attribution recalculated.

Another possible occasion when the attribution will be less satisfactory is during market turbulence, when the moves δS, δr, etc. are large.

Chapter 5

The regulatory environment

5.1 Backtesting

If the relevant regulatory body gives its approval, a bank can use its own internal VaR calculation as a basis for capital adequacy, rather than another more punitive measure that must be used by banks that do not have such approval.

The internal models approach is the most desirable method for determining capital adequacy.[1] The quantification of market risk for capital adequacy is determined by the bank's own VaR model. However, given a free hand, a bank could simply understate their VaR. Hence, there is a need to test the model to see if the stated VaR is consistent with the p&l that actually occurs. This is the purpose of backtesting, which applies statistical tests to see if the number of exceptions that have occurred is consistent with the number of exceptions predicted by the model.

To find the market risk charge under VaR, we first determine the 10 trading day (or two week) VaR at the 99% confidence level, call it VaR^{10}. The Market Risk Charge at time t is

$$\max\left\{\mathrm{VaR}^{10}_{t-1}, k\cdot \mathrm{Ave}\left(\mathrm{VaR}^{10}_{t-1},\ \mathrm{VaR}^{10}_{t-2},\ \ldots, \mathrm{VaR}^{10}_{t-60}\right)\right\} \tag{5.1}$$

where k can be as low as 3 but may be increased to as much as 4 if backtesting proves unsatisfactory ie. backtesting reveals that a bank is overly optimistic in the estimates of VaR. This provision is clearly to prevent gaming by the bank - under-reporting VaR numbers (in the expectation that they will not backtest) in order to lower capital requirements.

The factor $k \approx 3$ comes from thin air - the so-called hysteria factor. Legend has it that it arose as a compromise between the US regulatory authorities (who wanted $k = 1$) and the German authorities (who wanted $k = 5$) (Dowd 1998).

In principle the correct approach would be to measure the VaR at the holding period and confidence level that maps to the preferred probability of institutional survival (eg. 1 year and 99.75% for an A rated bank) and then use $k = 1$ (Dowd 1998).

The 10-day VaR is actually calculated using a 1-day VaR and the square root of time rule. In this case the Market Risk Charge at time t is

$$\sqrt{10}\max\left\{\mathrm{VaR}^{1}_{t-1}, k\cdot \mathrm{Ave}\left(\mathrm{VaR}^{1}_{t-1},\ \mathrm{VaR}^{1}_{t-2},\ \ldots, \mathrm{VaR}^{1}_{t-60}\right)\right\} \tag{5.2}$$

where the VaR numbers are now with daily horizon.

[1] We are specifically referring here to market risk capital requirements. To obtain regulatory approval to use an internal model, the institution also has to have in place tested measures for credit risk and operational risk. That is not the subject of this document. For an overview of the entire regulatory process, known as Basel II, see (Basel Committee on Banking Supervision 2003), (Basel Committee on Banking Supervision 2004*b*).

The actual regulatory capital requirement will also include a Specific Risk Charge for issuer-specific risks, such as credit risks. In Basle II operational risk charges are included.

In an interesting study done in 1996, at the time these proposals were being made into regulations, (Dimson & Marsh 1996) found that the internal models approach was the only method that was consistently sufficient to safeguard the capital of banks in times of stress. They also found that a method they call net capital at risk, which can be viewed as a misapplication of a VaR-type approach, was by far the worst of the several methods examined. This would be, for example, a pure net delta approach to VaR - as would occur if a South African bank was to use a delta equivalent position in the TOP40 for their entire set of domestic equity-based positions. Thus, they conclude that the internal models approach only makes sense with stringent quality control.

An exception is a day on which the loss amount was greater than the VaR amount. If we are working with a 95% confidence, and if the model is accurate, then on average we should have an exception on 1 day out of 20.

Even though Capital Adequacy is based on 99% VaR with a 10-day holding period, backtesting is performed on VaR with a daily horizon, and can be performed at other confidence levels. There is no theoretical problem with this, and the advantage of using daily VaR is that a larger sample is available and so statistical tests have greater power. Of course the horizon cannot be less than the frequency of p&l reporting, and this is almost always daily.

Backtesting is a logical manner of providing suitable incentives for use of internal models.

The question arises as to whether to use actual or theoretical p&l's. It is often argued that VaR measures cannot be compared against actual trading outcomes, since the actual outcomes are contaminated by changes in portfolio composition, and more specifically intra-day trading. This problem becomes more severe the longer the holding period, and so the backtesting framework involves one-day VaR.

To the extent that backtesting is purely an exercise in statistics, it is clear that the theoretical p&l's should be used for an uncontaminated test. However, what the regulator is really interested in is the solvency of the institution in reality, not in a theoretical world! Thus there are arguments for both approaches, and in fact backtesting has been a requirement for approved VaR models since the beginning of 1999, on both an actual (traded) and theoretical (hypothetical) basis. This is to ensure that the model is continually evaluated for reasonability. Backtesting for approved models occurs on a quarterly basis with one year of historical data (250 trading days) as input.

Extensive backtesting guidelines are given in the January 1997 Basle accord (Basel Committee on Banking Supervision 1996*b*).

Because of the statistical limitations of backtesting the Basle committee introduced a three zone approach:

- Green zone: the test does not raise any concerns about the model. The test results are consistent with an accurate model.
- Yellow zone: the test raises concerns about the model, but the evidence is not conclusive. The test results could be consistent with either an accurate or inaccurate model.

  The capital adequacy factor (k-factor) will be increased by the regulator. The placement in the yellow zone (closer to green or red) should guide the increase in a firms capital requirement. The following recommendations are made (Basel Committee on Banking Supervision 1996*b*):

Number of exceptions	Zone	Scaling factor (k-factor)
0-4	Green	3
5	Yellow	3.4
6	Yellow	3.5
7	Yellow	3.65
8	Yellow	3.75
9	Yellow	3.85
10+	Red	4 (model withdrawn)

The basic idea is that the increase in the k-value should be at least sufficient to return the model to the 99% standard in terms of capital requirements. Nevertheless, some game theory is possible here, at least in principle. To obtain exact answers in this regard requires additional distributional assumptions which may not hold in reality.

This is the most difficult case, but the burden of proof in these situations will be on the institution to demonstrate that their model is sound. This is achieved through decomposition of exceptions, documentation of each exception, and provision of backtesting results at other confidence intervals, for example.

- Red zone: the test almost certainly raises concerns about the model. The test results are almost certainly inconsistent with an accurate model.[2] The k-factor is increased to 4, and approval for the existing model is almost certainly withdrawn.

Because we are taking a sample from a distribution, the sample is subject to error. Based on the sample, we test if the model is valid using standard statistical hypothesis testing. Recall that there are two types of errors associated with statistical tests:

- Type I error: rejecting a valid model,
- Type II error: accepting an invalid model.

As is well known in statistics, it is impossible to control the size of these errors simultaneously.

For each day in history we determine whether or not an exception occurred, so we have $p_1,\ p_2,\ \ldots,\ p_N$ being 0-1 Boolean variables, where the vector starts on the first day on which backtesting was performed.

Under the null hypothesis, if the model is valid, the total $\sum_{i=1}^{N} p_i$ is distributed according to the binominal distribution, with a total number of experiments being N and the failure probability p being for example 0.05 ie. 1 minus the VaR confidence level. The failure probability has to be reasonable, so that the model test can be significant. For example, if $p = 0.0001$, we probably won't have any exceptions, but we also won't be able to accept the null hypothesis at any significance. Thus it is typical to choose a confidence of 95% or 99%. For regulatory purposes the required probability is 99%.

For a binomial distribution with sample size N and failure probability p, and X being the total number of failures, we have

$$P[X = i] = \binom{N}{i} p^i (1-p)^{N-i} \tag{5.3}$$

and so

$$P[X \leq x] = \sum_{i=0}^{x} \binom{N}{i} p^i (1-p)^{N-i} \tag{5.4}$$

[2] One of the reasons that (Basel Committee on Banking Supervision 1996*b*) give for why this can occur is because the volatility and correlation estimates of the model are old, and have been outdated by a major market regime shift, thus causing a large number of exceptions. The requirements of models is that parameter estimates be not more than three months old. This is almost laughable. An even half-decent model should have automated parameter updating on a daily basis.

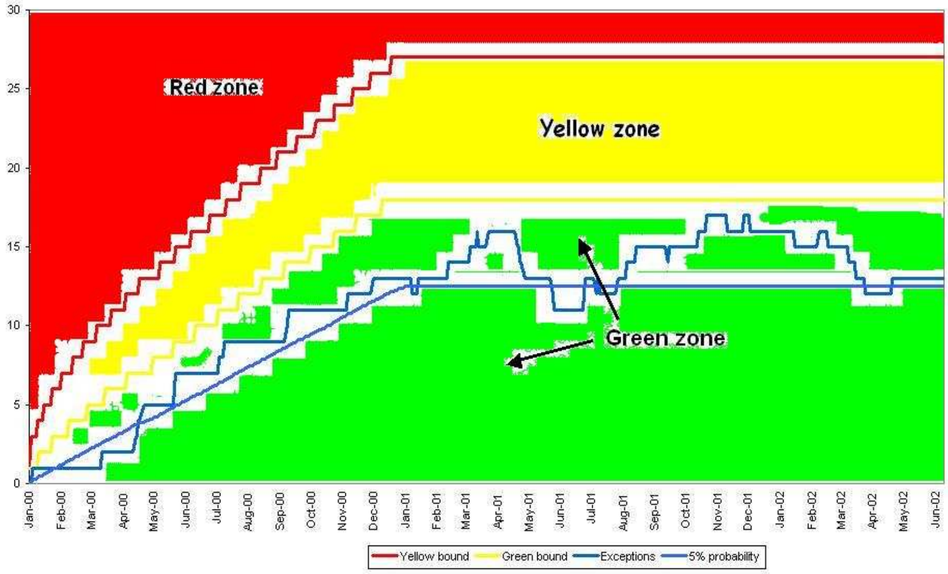


Figure 5.1: Diagrammatic representation of the three zone approach for 95% VaR

We reject the null hypothesis if the number of exceptions is larger than the test level. There are two test levels which correspond to the yellow and red zones. The yellow zone is determined by the Basle Accord as having a 5% Type I error. Thus, if the VaR model is valid, there is only a 5% chance of it being in the yellow zone - in other words, bad luck in terms of the number of exceptions. The red zone is defined by the Basle Accord as having a 0.01% Type I error. Thus if the VaR model is valid, there is only a 0.01% chance that it will have this number of exceptions. This is very generous, because one concludes that once a possibly suspect (but not absurd) model has been approved, only a very poor performance subsequently will lead to approval being withdrawn. The Basle regulations ensure that Type I errors are almost impossible, but it allows for a significant proportion of Type II errors.

On a sample of $N = 250$ observations, with a failure probability of 0.01, the yellow zone starts with 5 exceptions and the red zone starts with 10 exceptions. Note that 2.5 exceptions are expected.

Presumably, in order to qualify as an approved model, it starts off life in the green zone. Thus, at the time of inception, the daily VaR at the 99% confidence level had at most 4 exceptions. If, at any quarter end, the previous 250 observations had from 5 to 9 exceptions, the model is reclassified into the yellow zone. If it had 10 or more exceptions, it is reclassified into the red zone.

Warning: in many statistics textbooks normal distribution approximations are available for the binomial distribution. These are not to be used here, because they only apply for binominals where the failure probability is 'not too extreme', for example, (Underhill & Bradfield 1994) specifies that the normal approximation is only valid if $0.1 < p < 0.9$. We are interested in the case where $p = 0.05$, $p = 0.01$ or perhaps $p = 0.0001$.

5.2 Other requirements for internal model approval

Banks that use the internal models approach for meeting market risk capital requirements must have in place a rigorous and comprehensive stress testing program. The stress scenarios need to cover a range of factors that can create extraordinary p&l's in trading portfolios, or make the control of risk in those portfolios very difficult. These factors include low-probability events in all major types of risks, including the various components of market, liquidity, credit, and operational risks.

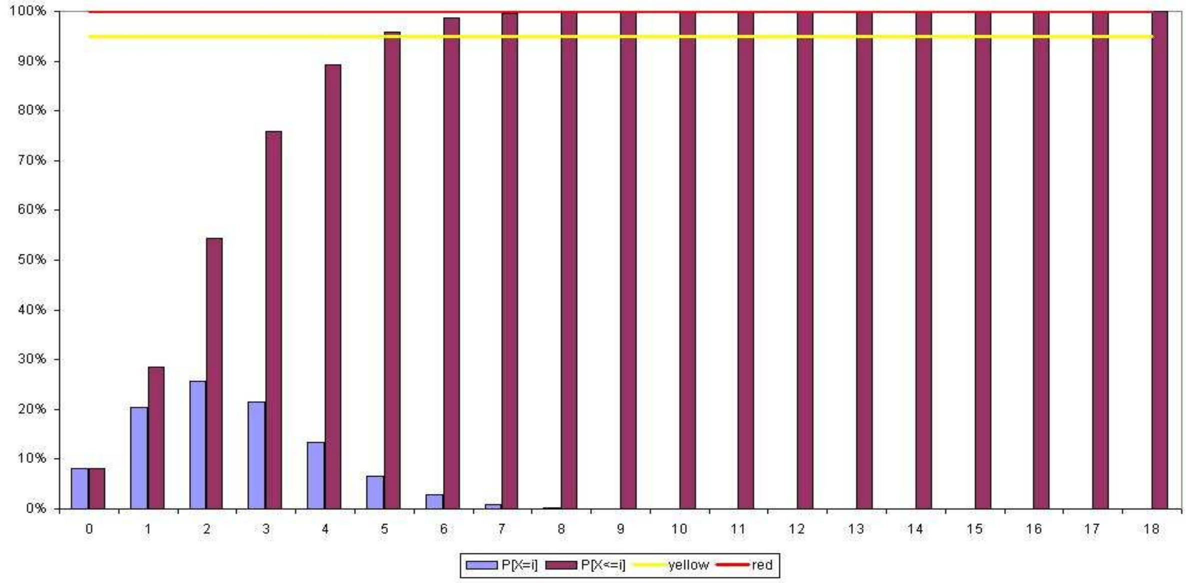


Figure 5.2: The yellow zone starts with 5 exceptions and the red zone starts with 10 exceptions for 99% VaR.

The institution must be able to provide information as follows (Basel Committee on Banking Supervision 1996*a*):

- Information on the largest losses experienced during the reporting period.
- Stress testing the current portfolio against past periods of significant disturbance, incorporating both the large price movements and the sharp reduction in liquidity associated with these events.

  Stress testing the sensitivity of the bank's exposure to changes in the assumptions about volatilities and correlations. Due consideration should be given to the sharp variation that at times has occurred in periods of significant market disturbance. The 1987 equity crash, for example, involved correlations within risk factors approaching the extreme values of 1 or -1 for several days at the height of the disturbance.
- Use of scenarios developed by the bank itself to capture the specific characteristics of its portfolio. Banks should provide supervisory authorities with a description of the methodology used to identify and carry out the scenarios as well as with a description of the results derived from these scenarios.

If the testing reveals particular vulnerability to a given set of circumstances, the national authorities would expect the bank to take prompt steps to manage those risks appropriately (eg. by hedging against that outcome or reducing the size of its exposures).

5.3 Credit and operational risk measures

5.3.1 Basel I

(Basel Committee on Banking Supervision 1988). This was finalised in 1988, effective 1992.

We need to set a minimum level of risk capital, which is a percentage of the total risk weighted assets (RWA).

Capital adequacy is to have capital of at least 8% of the total risk weighted assets. Of this, at least 4% needs to be tier I capital (equity and disclosed reserves); and the rest tier II capital (undisclosed reserves, revaluation reserves, provisions, etc.). Tier III capital (short term subordinated debt) can be used as capital for market risk only.

For each counterparty, the approach to determining the RWAs is to perform a reduction algorithm to the assets under consideration.

(1) We first consider the on-balance sheet assets. These are typically outright loans. For each loan, we multiply the amount outstanding by the relevant risk capital weight in Figure 5.3. (This is only some of the weights, see (Basel Committee on Banking Supervision 1988) for the full list.) This gives the credit risk charge for that asset.

Product	Weight
Cash	0%
Claim on central bank in domestic currency	0%
Claims on OECD[3]banks	20%
Loans guaranteed by OECD	20%
Mortgages	50%
Claims on private sector	100%
Claims on non-OECD banks longer than 1y	100%

Figure 5.3: Basel I risk capital weights

(2) Next we first consider the off-balance sheet assets. We want to define, for each asset, a equivalent value that makes it 'loan-like', that is, we homogenise the assets.

Some off-balance sheet items are very much 'loan-like' already. For example, a guarantee is very much like a loan, so it carries a risk capital weight of 100%. A performance bond carries a weight of 50%.

(3) Finally, we come to derivatives. An important point to note about derivatives is that typically a bank will have many derivatives with the same counterparty, and that the notional of a derivative typically dwarfs its actual value. For any particular counterparty, we would like to take into account possible netting of all the outstanding instruments we have with that counterparty. The question of netting is whether or not the requisite "ISDA is in place". If it is, we define the net gross ratio, given by

$$\mathrm{NGR} = \frac{\max(0, \sum_j \mathrm{MtM}_j)}{\sum_j \max(0, \mathrm{MtM}_j)} \tag{5.5}$$

where the summation is taken across all instruments outstanding with that counterparty.

Now we define the exposure as

$$\max(0, \mathrm{MtM}) + [0.4 + 0.6\mathrm{NGR}]\,\mathrm{Notional}\,P$$

where P is the percentage seen in Figure 5.4.

(4) The above amounts are the 'loan-like' sizes of the off-balance sheet items. We then apply the weights seen in Figure 5.3 to those figures. Since players in the derivatives market tend to be higher credits, this figure is then again halved. This gives the credit risk charge for that asset.

(5) By addition across all counterparties we find the total risk weighted assets.

[3]Organisation for Economic Co-operation and Development. Currently Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Turkey, United Kingdom and United States.

Residual maturity	Interest rate	Forex, gold	Equity	Other metal	Commodities
$m \leq 1$y	0%	1%	6%	7%	10%
1y $< m \leq 5$y	0.5%	5%	8%	7%	12%
5y $< m$	1.5%	7.5%	10%	8%	15%

Figure 5.4: Multipliers

What are the weaknesses of this approach? There is inadequate differentiation between credit qualities and credit maturities. There are inadequate risk mitigation / portfolio construction incentives.

5.3.2 Basel II

The Basel Committee's goal was to finalise the New Accord by the fourth quarter of 2003 with implementation to take effect in G-10 countries in 2006/2007, and later in other countries (at the discretion of the Central Banks in those countries).

The (Basel Committee on Banking Supervision 2003) was a consultative document issued for comment. The final version of Basel II appeared as (Basel Committee on Banking Supervision 2004*b*).

Basel II is based on the so-called three pillars (Basel Committee on Banking Supervision 2003):

I. Further developing of capital regulation that encompasses minimum capital requirements, by increasing substantially the risk sensitivity of the minimum capital requirements.

   The current Accord is based on the concept of a capital ratio where the numerator represents the amount of capital a bank has available and the denominator is a measure of the risks faced by the bank and is referred to as risk-weighted assets. The resulting capital ratio may be no less than 8%. Under the proposed New Accord, the regulations that define the numerator of the capital ratio (i.e. the definition of regulatory capital) remain unchanged. Similarly, the minimum required ratio of 8% is not changing. The modifications, therefore, are occurring in the definition of risk-weighted assets, that is, in the methods used to measure the risks faced by banks.

   (a) The treatment of market risk arising from trading activities was the subject of the Basel Committee's 1996 Amendment to the Capital Accord. The proposed New Accord envisions this treatment remaining unchanged.

   (b) Substantive changes to the treatment of credit risk relative to the current Accord; with three options offered for the calculation thereof.

   (c) The introduction of an explicit treatment of operational risk, that will result in a measure of operational risk being included in the denominator of a bank's capital ratio; with three options offered for the calculation thereof. Operational risk is defined as the risk of losses resulting from inadequate or failed internal processes, people and systems, or external events.

II. Supervisory review of capital adequacy. Judgements of risk and capital adequacy must be based on more than an assessment of whether a bank complies with minimum capital requirements. This pillar seems to be a statement that empowers the supervisor in this respect. To quote (Basel Committee on Banking Supervision 2003): 'it is inevitable that a capital adequacy framework, even the more forward looking New Accord, will lag to some extent behind the changing risk profiles of complex banking organisations, particularly as they take advantage of newly available business opportunities. Accordingly, this heightens the importance of, and attention supervisors must pay to pillar two.'

   Stress testing falls under this pillar.

III. Public disclosure (market discipline). Again, from (Basel Committee on Banking Supervision 2003): 'The Committee has sought to encourage market discipline by developing a set of disclosure requirements that allow stakeholders to assess key information about a bank's risk profile and level of capitalisation.' For example, annual reports to shareholders will have to include fairly solid evidence, adhering to some prescribed general formats, rather than qualitative meanderings, about the methods of risk control.

    Public disclosure strengthens the incentives for a bank to behave prudently.

    Just how much control the regulator will have over the nature of this information will depend on the legal jurisdiction that applies to the particular regulator-bank relationship.

Countries not represented in the G-10 have more freedom in their implementation programme, and many have been busy evaluating the suitability of the new Framework for their jurisdiction. In order to further this process the committee convened a Working Group comprised of members from non-G10 countries to assess the whether and when of Basel II, and to provide practical suggestions to supervisors for the transition to the new Framework. This occurred in the first half of 2003 with the report being (Basel Committee on Banking Supervision 2004*a*). The document is not intended to be an interpretation of Basel II rules, but rather as practical advise for countries with differing resources, particular concentration risks, etc. (Basel Committee on Banking Supervision 2003), (Basel Committee on Banking Supervision 2004*b*). South Africa has an implementation target date of 1 January 2008.

The first pillar of Basel II deals with the new methods for calculation of exposures. For credit risk there are broadly two approaches:

(a) Standardised Approach: similar to Basel I in that banks are required to slot their credit exposures into supervisory categories. The categories can be determined as corresponding to the rating categories published by a rating agency. For example, we see in Figure 5.3 that any claim with a counterparty in the private sector has a risk weight of 100% under Basel I. In Basel II, if the counterparty is rated anything from AAA to AA-, then the risk weight is 20%, A+ to A- is 50%, etc. Below B- is 150%.[4]

    Another source of relief is that the risk weight for mortgages has been reduced. Also, some recognition for protection from bought credit derivatives is available.

    However, this approach is now very punitive in terms of the amount of capital that needs to be held. Moreover, in South Africa the lack of many internationally rated corporates makes this method difficult to apply. Hence banks will be aiming for the internal ratings-based approach.

(b) The internal ratings-based approach.

    (i) for corporate, bank and sovereign exposures. The bank will use a model where the risk components are

        - Probability of default (PD), a percentage
        - Loss given default (LGD), a percentage
        - Exposure at default (EAD), in currency units
        - Maturity (M), in years

        The capital required will be proportional to

$$\mathrm{EAD} \cdot \mathrm{LGD} \cdot [s - \mathrm{PD}] \cdot f(M) \tag{5.6}$$

        where s is a parameter which is intended to cover unexpected losses at the 99.9% confidence level, and f a given increasing function with $f(1) = 1$.

        For the *foundational* internal ratings-based approach the PD is provided by the bank, the others are set by the supervisor. For the *advanced* internal ratings-based approach all inputs

[4]Lest this doesn't make sense, remember that we are going to multiply by 8% later on.

for an internal model are provided by the bank: these will be at least the factors listed above, there may be others. The loss given default is especially difficult for a bank to quantify, given again the small number of defaults historically, so banks here will typically pass this issue back to the regulator, and stick to the foundational model. Moreover, some research has shown that in emerging markets, banks tends to have far higher concentration risk levels, and hence this method could be more punitive than would first appear! (Risk Magazine June 2003). With the takeover of ABSA by Barclay's, ABSA made a firm commitment to implement the internal ratings-based approach, with the other big banks then making commitments to do likewise (Risk Magazine November 2005).

Small/medium enterprise exposures with an exposure of more than €1m will be treated as corporate exposures, with some firm-size adjustments.

(ii) for retail exposures. There is only one approach available. The bank provides all of PD, LGD and EAD; on a pooled basis, rather than on a name by name basis. The exposures are divided into three categories, which is broadly mortgages, credit cards, and other. The formula for capital required is the same as (5.6) except the maturity factor disappears for each of these categories; the s factor is calibrated differently for each of the three categories.[5] The capital required is then aggregated.

Small/medium enterprise exposures with an exposure of less than €1m will be treated as retail exposures.

(iii) for specialised lending. There are two options: this is treated as a corporate exposure. Commercial retail estate has a special category. Alternatively, a scale of five quality grades with risk weights for each grade is available.

(iv) equity exposures. There are two options: a corporate approach is taken, but LGD is mandated at 90%. Alternatively, a quarterly VaR type approach can be taken to model the potential decrease in the value of equity holdings.

In the internal rating based approach, the bank needs to calculate the PD (under either of the two variations). For this Basel II sets certain requirements in the calculation methodology and testing for PD estimates.

The requirement is that obligors are classified into risk buckets, and each risk bucket is assigned a PD. This PD must be a long-run average of the one-year realised default rates for that bucket. As such, it could vary significantly from a case-by-case model of obligor's PDs that will result from (KMV 2006) for example. Basel also enunciates a preference for a *through the cycle* approach i.e. the PD is based on long term statistics and is not overly dependent on where we are in the economic cycle.

Although there is great freedom in the definition of the risk buckets, the benchmarking and backtesting requirements of these buckets is quite stringent.

Calculation methods for operational risk:

1. Basic Indicator Approach: the measure is a bank's average annual gross income over the previous three years. This average, multiplied by a factor of 0.15 set by the Committee, produces the capital requirement.

2. Standardised Approach: similar, but banks must calculate a capital requirement for each business line. This is determined by multiplying gross income by specific supervisory factors determined by the Committee.

3. Advanced Measurement Approaches Fairly open ended specifications, stated as aimed to encourage the growth of the quantification of operational risk. Banks using such methods are permitted to recognise the risk mitigating impact of insurance.

[5] Both here and above the s factor has a modelled correlation within the group as a significant input.

The second two methods do not produce a significant saving over the first, so banks will typically opt for the first approach.

	Credit Risk			**Market Risk**		**Operational Risk**		
Basel 1988	Standardised approach			Standardised approach		-		
Basel 1996	Standardised approach			St'dised approach	Internal model (VaR, stress)	-		
Basel 2006++	St'dised approach	Foundational internal rating approach	Advanced Internal ratng approach	St'dised approach	Internal model (VaR, stress)	Basic indicator approach	St'dised approach	Advanced measurement

Chapter 6

Coherent risk measures

6.1 VaR cannot be used for calculating diversification

If f is a risk measure, the diversification benefit of aggregating portfolio's A and B is defined to be

$$f(A) + f(B) - f(A + B) \tag{6.1}$$

When using full revaluation VaR as the methodology for computing a risk measure, it's quite possible to get negative diversification. Pathological examples are possible, but the following example is not absurd:

Suppose one has a portfolio that is made up by a Trader A and Trader B. Trader A has a portfolio consisting of a sold put that is far out of the money, and has one day to expiry. Trader B has a portfolio that consists of a sold call that is also far out of the money, and also has one day to expiry. Using any historical VaR approach, say we find that each option has a probability of 4% of ending up in the money.

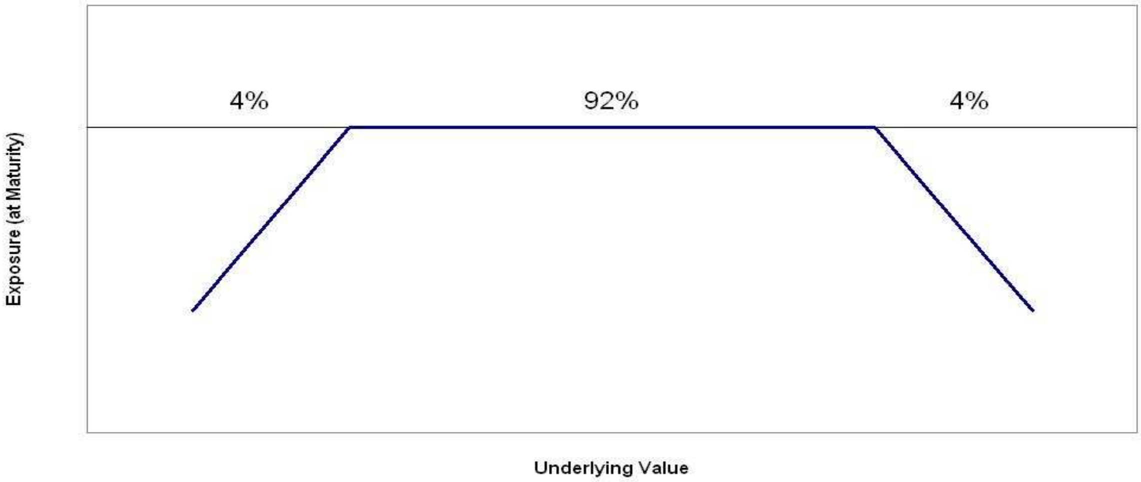
Trader A and B each have a portfolio that has a 96% chance of not losing any money, so each has a 95% VaR of zero.[1] However, the combined portfolio has only a 92% chance of not losing any money, so its VaR is non-trivial. Therefore we have a case where the risk of the combined portfolio is greater than the risks associated with the individual portfolios, i.e. negative diversification benefit if VaR is used to measure the diversification benefit. This example appears in (Artzner, Delbaen, Eber & Heath 1997).

What is so awkward about the lack of sub-additivity is the fact that this can give rise to regulatory arbitrage or to the break-down of global risk management within one single firm. This is also a serious concern for regulators. If regulation allows the capital requirement of a firm to be calculated as the sum of the requirements of its subsidiaries and if the requirements are based on VaR, the firm could create artificial subsidiaries in order to save regulatory capital.

6.2 Risk measures and coherence

This example introduces the concept of a "Coherent Risk Measure". If $f(A+B) \leq f(A)+f(B)$, where A and B denote portfolios, then f is said to be coherent (Artzner et al. 1997), (Artzner, Delbaen, Eber & Heath 1999). In fact a coherent risk measure needs to satisfy five properties (Artzner et al. 1999), as follows:

[1]To be precise, their VaR is actually a very small negative number! The average of their V_i's is negative, the $5^{th}\%$ is zero.



- translation invariance: $f(A + \alpha r) = f(A) - r$, where r is a reference risk free investment. (As David Heath has explained to me, this condition is simply there to ensure that the risk measure and the p&l measure is in the same numeraire, namely, currency.)
- Subadditivity: $f(A + B) \leq f(A) + f(B)$
- Positive homogeneity: for all $\lambda \geq 0$, $f(\lambda A) = \lambda f(A)$.
- Monotoneity: if $A \leq B$ then $f(A) \leq f(B)$.
- Relevance: if $A \neq 0$ then $f(A) > 0$.

The property we have focused on means 'a merger does not create extra risk', and is a natural requirement (Artzner et al. 1999).

In other words the risk measure f of a portfolio consisting of sub-portfolios A and B would always be less than or equal to the sum of the risk measure of portfolio A with the risk measure of portfolio B. The example above shows that full revaluation VaR is not coherent. It also means that as a conservative measure of risk, one can simply add the risks calculated for the various sub-portfolios, if the measure is coherent.

The earlier example is not a purely theoretical example. In practice, even on large and diverse portfolios, using VaR to calculate the diversification benefit does indeed occasionally lead to the case where this diversification is negative.

There is thus a need for practical and intuitive coherent risk measures. The basic example - originally presented in this country in (Eber 29-30 June 1999) - is that in the place of a VaR calculation we use a concept known as Expected Tail Loss (ETL) or Expected Shortfall (ES). It is easiest to understand in the setting of a historical-type VaR calculation, let us say 95% VaR. It would entail instead of taking the 5^{th} percentile of the p&l's to yield a VaR number, take the average of the p&l's up to the 5^{th} percentile to yield an ES number.

Looking at the 5th percentile we end up with a VaR number which basically represents the best outcome of a set of bad outcomes on a bad day. Using ES we look at an average bad outcome on a bad day. This ES number turns out to be a coherent risk measure (Artzner et al. 1997), (Eber 29-30 June 1999), (Acerbi & Tasche 2001), and therefore guarantees that the diversification is always positive. As stated in the abstract of (Acerbi & Tasche 2001), "Expected Shortfall (ES) in several variants has been proposed as remedy for the deficiencies of Value-at-Risk (VaR) which in general is not a coherent risk measure".

A readable account of these and related issues is (Acerbi, Nordio & Sirtori 2001).

One should report both VaR and ES, but use only ES to calculate and report diversification.

Note that because standard deviation is sub-additive the standard RiskMetrics simplification is coherent:

$$\begin{aligned}\sigma^2(X+Y) &= \sigma^2(X)+\sigma^2(Y)+2\sigma(X)\sigma(Y)\rho \\ &\leq \sigma^2(X)+\sigma^2(Y)+2\sigma(X)\sigma(Y) \\ &= (\sigma(X)+\sigma(Y))^2\end{aligned}$$

and so

$$\sigma(X+Y) \leq \sigma(X)+\sigma(Y)$$

and hence

$$\mathrm{VaR}(X+Y) \leq \mathrm{VaR}(X)+\mathrm{VaR}(Y),$$

which is the definition of subadditivity. The usual RiskMetrics VaR is also subadditive (and hence coherent), but this is a mathematical exercise for masochists - it is not easy at all. According to (Breuer, Krenn & Pistovčák 2002) to guarantee sub-additivity of (presumably parametric) VaR, the value of the portfolio has to be a linear function of risk factors whose changes are elliptically distributed.

6.3 Measuring diversification

The diversification benefit of portfolio P_0 is equal to

$$f\left(\sum_{i=1}^{n} P_i\right)+f(P_0)-f\left(\sum_{i=0}^{n} P_i\right)$$

where f denotes ES and $P_1,\ P_2,\ \ldots,\ P_n$ are the (original) portfolios against which the diversification is measured.

6.4 Coherent capital allocation

The intention is to allocate capital costs in a coherent manner. This sounds like quite an otherworldly exercise, but one can make the task quite concrete and ask: of my risk number (such as ES), how much (as a percentage, say) is due to each of my positions? Then, given my capital adequacy charges (which may or may not be calculated via an approved internal model!) I can allocate as a cost the charges in those proportions to each of those desks.

Each desk can break down their own charges amongst their dealers, and management can decide where the greatest risk managment focus needs to lie.

(Denault 2001) has developed a method of allocating the risk capital costs to the various subportfolios in a fair manner, yielding for each portfolio, a risk appraisal that takes diversification into account. We wish to thank Freddie Delbaen, who contributed significantly to that paper, for clarifying certain issues.

The approach of (Denault 2001) is axiomatic, starting from a risk measure which is coherent in the above sense. We may specialise the results of (Denault 2001) to the case of the coherent Expected Shortfall risk measure in which case his results become quite concrete.

An allocation method for risk capital is then said to be coherent if

- The risk capital is fully allocated to the portfolios, in particular, each portfolio can be assigned a percentage of the total risk capital.
- There is 'no undercut': no portfolio's allocation is higher than if they stood alone. Similarly for any coalition of portfolios and coalition of fractional portfolios.
- 'Symmetry': a portfolio's allocation depends only on its contribution to risk within the firm, and nothing else.
- 'Riskless allocation': a portfolio that increases its cash position will see its allocated capital decrease by the same amount.

All of these requirements have precise mathematical formulations.

A coherent allocation is to be understood as one that is fair and credible.

One should not be surprised to be told that this is a game theoretic problem where the portfolios are players, looking for their own optimal strategy. (Denault 2001) applies some results from game theory to show that the so-called Aumann-Shapley value from game theory is an appropriate allocation (it is a Nash equilibrium in the theory of cooperative games). Further, some results (fairly easy to derive in this special case) from (Tasche July 1999) on the differentiability of Expected Shortfall show that the Aumann-Shapley value is given by

$$K_i = -\mathbb{E}[X_i | X \leq q_\alpha] \tag{6.2}$$

where X_i denotes the (random vector of) p&l's of the i^{th} portfolio, $X = \sum_j X_j$ is the vector of p&l's of the company, and q_α is the α percentile of X.

Capital allocation			ETL
Bank			23,254,108
Business A			86.42%
	Desk T		4.33%
		Book T1	0.09%
		Book T2	4.24%
	Desk E		81.47%
		Book E1	72.39%
		Book E2	0.49%
		Book E3	1.34%
		Book E4	7.25%
	Desk R		0.90%
		Book R1	-0.25%
		Book R2	0.001%
		Book R3	1.14%
	Broker		-0.28%
Business B			10.32%
	Desk B1		9.46%
	Desk B2		0.87%
Business C			3.26%

Figure 6.1: Coherent capital allocation using ETL

Hence, as a percentage of total capital, the capital cost for the i^{th} portfolio is

$$\frac{\mathbb{E}[X_i | X \leq q_\alpha]}{\mathbb{E}[X | X \leq q_\alpha]} \tag{6.3}$$

In the context of any historical or Monte Carlo type VaR model, this fraction is easy to calculate:

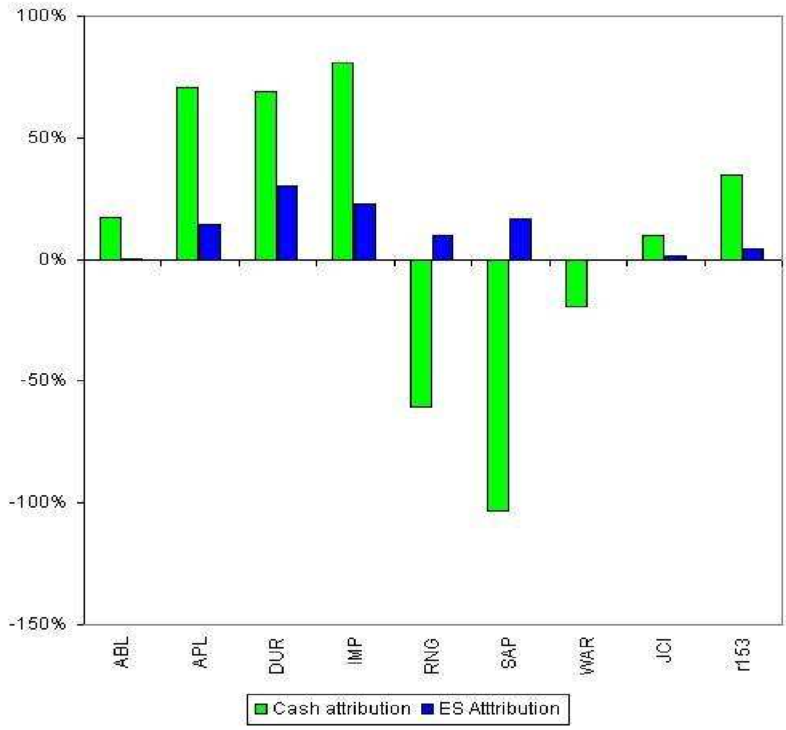


Figure 6.2: Capital allocation and cash allocation to a long and short equity portfolio. Note, for example, how the shorts in RNG and SAP raise cash without material contribution to risk, while the short in WAR appears to have no impact on risk.

- The denominator is the average of the $1-\alpha\%$ worst p&l's of the entire bank,
- The numerator is the average of the p&l's that correspond to the same experiments as in the denominator.

An example of how this might transpire is in Figure 6.1 and Figure 6.2.

6.5 Greek Attribution

Consider a first order (delta-gamma-rho-vega-theta) Taylor series expansion as follows:

$$dV \simeq \Delta\,\delta S + \tfrac{1}{2}\Gamma\,(\delta S)^2 + \rho\,\delta r + \mathcal{V}\,\delta\sigma_{\text{atm}} + \theta\,\delta t$$

This expansion allows us to attribution our p&l according to the sensitivities, and this enables us to analyse what type of bets the dealer is making.

For a simple instrument, such as an equity derivative, we can attribute the p&l as

- $\Delta_{t-1}\,(S_t - S_{t-1})$ is the p&l due to delta;
- $\frac{1}{2}\Gamma_{t-1}\,(S_t - S_{t-1})^2$ is the p&l due to gamma;
- $\rho_{t-1}\,(r_t - r_{t-1})$ is the p&l due to rho;
- $\mathcal{V}_{t-1}\,(\sigma_{\text{atm},t} - \sigma_{\text{atm},t-1})$ is the p&l due to vega;

- $\theta_{t-1}\,\delta t$ is the p&l due to theta;
- the remainder is the p&l due to error in the Taylor series expansion.

At regular intervals we should check that the error term is not material. Of course, we can attribute a percentage to this error term, which should not be more than a couple of percent. After all, the error term is a measure of how well the Taylor series expansion is fitting the actual p&l. As expected, for more complicated products, these errors can be more material, and the method should not be used. Alternatively, a higher order Taylor series expansion could be derived and the appropriate attribution recalculated.

Another possible occasion when the attribution will be less satisfactory is during market turbulence, when the moves δS, δr, etc. are large.

Now suppose that we wish to decompose a VaR or ES measure into exposures to the various Greeks. Assume that we are using a historical-type or Monte Carlo method for calculating our VaR or ES. Then we can consider the p&l's generated by the various historical or Monte Carlo experiments, as follows:

$$dV_i = \Delta(S_i - S) + \tfrac{1}{2}\Gamma(S_i - S)^2 + \rho(r_i - r) + \mathcal{V}(\sigma_{\mathrm{atm}\,i} - \sigma_{\mathrm{atm}}) + \theta\delta t + \epsilon_i$$

where S is the original spot, S_i is the i^{th} spot experiment, etc. The p&l's due to delta are the $\Delta(S_i - S)$, etc.

Of course, our risk measure is calculated by looking at the tail of this distribution of p&l's.

Long position		Short position	
full VaR	30.09	full VaR	40.51
full ETL	38.77	full ETL	65.35
greek VaR	30.98	greek VaR	39.90
greek ETL	37.63	greek ETL	64.24
delta attribution	126.86%	delta attribution	78.77%
gamma attribution	-26.67%	gamma attribution	18.11%
rho attribution	0.00%	rho attribution	0.04%
sv attribution	-8.80%	sv attribution	4.74%
theta attribution	5.67%	theta attribution	-3.36%
error atttribution	2.94%	error atttribution	1.71%

Figure 6.3: Coherent greek attribution using ES

Again we should check that the ϵ_i are not material. This time it is even easier than before: the capital attribution method is attributing a percentage to this error term, which should not be more than a couple of percent.

Bibliography

Acerbi, C., Nordio, C. & Sirtori, C. (2001), 'Expected shortfall as a tool for financial risk management'. *http://www.gloriamundi.org/var/wps.html

Acerbi, C. & Tasche, D. (2001), 'On the coherence of expected shortfall'. *http://www-m1.mathematik.tu-muenchen.de/m4/Papers/Tasche/shortfall.pdf

Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1997), 'Thinking coherently', *Risk* **10**(11), 68–71.

Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1999), 'Coherent measures of risk', *Math. Finance* **9**(3), 203–228. *http://www.math.ethz.ch/~delbaen

Basel Committee on Banking Supervision (1988), 'International convergence of capital measurement and capital standards (July 1988, updated to April 1998)'. *http://www.bis.org/publ/bcbsc111.pdf

Basel Committee on Banking Supervision (1996*a*), 'Amendment to the capital accord to incorporate market risks'. *http://www.bis.org/publ/bcbs24.htm

Basel Committee on Banking Supervision (1996*b*), 'Supervisory framework for the use of 'backtesting' in conjunction with the internal models approach to market risk capital requirements'. *http://www.bis.org/publ/bcbs22.htm

Basel Committee on Banking Supervision (2003), 'Consultative document: Overview of the new Basel capital accord'. *http://www.bis.org/publ/cp3ov.pdf

Basel Committee on Banking Supervision (2004*a*), 'Implementation of Basel II: Practical considerations'. *http://www.bis.org/publ/bcbs109.pdf

Basel Committee on Banking Supervision (2004*b*), 'International convergence of capital measurement and capital standards: A revised framework'. *http://www.bis.org/publ/bcbs107.pdf

Breuer, T., Krenn, G. & Pistovčák, F. (2002), Stress tests, maximum loss, and value at risk, *in* B. Britzelmaier, S. Geberl, M. Menichetti & H.-R. Kaufmann, eds, 'Regulierung oder Deregulierung der Finanzmrkte: Strategieimplikationen fr Finanzdienstleister', Physica-Verlag (Springer).

Burden, R. L. & Faires, J. D. (1997), *Numerical Analysis*, sixth edn, Brooks Cole.

Denault, M. (2001), 'Coherent allocation of risk capital', *Journal of Risk* **4**(1), 1–34.

Dimson, E. & Marsh, P. (1996), 'Stress tests of capital requirements', *The Wharton School* **96-50**.

Dowd, K. (1998), *Beyond Value at Risk: The new science of risk management*, Wiley series in Frontiers of Finance, John Wiley and Sons.

Duffie, D. & Singleton, K. (2000), 'Credit risk pricing and risk measurement for financial institutions'. Graduate School of Business, Stanford University.

Eber, J.-M. (29-30 June 1999), 'Measures of risk for effective risk management'. Talk, Risk Conferences, Advanced techniques and strategies for risk measurement and management of derivatives, Johannesburg.

Fender, I., Gibson, M. & Mosser, P. (November 2001), 'An international survey of stress tests', *Current Issues in Economics and Finance* **7**(10).

Higham, N. J. (2002), 'Computing the nearest correlation matrix - a problem from finance', *IMA Journal of Numerical Analysis* **22**, 329–343.
*http://www.ma.man.ac.uk/ nareports/narep369.pdf

Hosoya, M. & Shimizu, T. (December 2002), 'Implications of a macro stress test on financial stability: Summary of the second census on stress tests', *Market Review, Bank of Japan Financial Markets Department* **2002-E-4**.

Hull, J. (2002), *Options, Futures, and Other Derivatives*, fifth edn, Prentice Hall.

Hull, J. & White, A. (1998), 'Incorporating volatility up-dating into the historical simulation method for V@R', *Journal of Risk* **1**(Fall).

J.P.Morgan & Reuters (December 18, 1996), *RiskMetrics - Technical Document*, fourth edn, J.P.Morgan and Reuters, New York.
*http://www.riskmetrics.com

KMV (2006), 'Moody's KMV'.
*http://www.moodyskmv.com/

Kolman, J. (April 1999), 'LTCM speaks', *Derivatives Strategy* .
*http://www.derivativesstrategy.com/magazine/archive/1999/0499fea1.asp

Lowenstein, R. (2000), *When Genius Failed: The Rise and Fall of Long Term Capital Management*, Random House.

Mina, J. & Ulmer, A. (April 1999), 'Delta-gamma four ways', *CreditMetrics Monitor* .
*http://www.riskmetrics.com/deltaow.html

Pafka, S. & Kondor, I. (2001), 'Evaluating the RiskMetrics methodology in measuring volatility and value-at-risk in financial markets'.
*http://www.colbud.hu/pdf/Kondor/riskm.pdf

Paul-Choudhury, S. (1997), 'This year's model', *Risk* **10**, 18–23.

Pichler, S. & Selitsch, K. (2000), 'A comparison of analytical VaR methodologies for portfolios that include options', *Model Risk* .

Risk Magazine (June 2003), 'South Africa special report'.

Risk Magazine (November 2005), 'South Africa special report'.

Sharpe, W. F. (1964), 'Capital asset prices - a theory of market equilibrium under conditions of risk', *Journal of Finance* pp. 425–442.

Taleb, N. (1997), *Dynamic Hedging: managing vanilla and exotic options*, John Wiley and Sons.

Tasche, D. (July 1999), 'Risk contributions and performance measurement', *Technische Universität München* **TUM-M9909**.

The Financial Economists Roundtable (October 6, 1999), 'The financial economists roundtable statement on Long-Term Capital Management and the report of the presidents' working group on financial markets'.
*http://www.luc.edu/orgs/finroundtable/statement99.html

Underhill, L. & Bradfield, D. (1994), *IntroSTAT*, Juta and Company.

Wee, L.-S. & Lee, J. (1999), 'Integrating stress testing with risk management', *Bank Accounting and Finance* **Spring**, 7–19.

West, G. (2006), 'The Mathematics of South African Financial Markets and Instruments: Lecture notes'.
*http://www.finmod.co.za/safm.pdf

Zangari, P. (1996), 'A VaR methodology for portfolios that include options', *RiskMetrics Monitor* **First Quarter**.