Rule Base LTV and Correction Function for Concentration Risk

YIRAN ZHANG
(M.Sc., HU BERLIN)

A THESIS SUBMITTED FOR THE DEGREE OF MASTER OF PHILOSOPHY
DEPARTMENT OF ACCOUNTING AND FINANCE
MANCHESTER BUSINESS SCHOOL
2011
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contents</td>
<td>I</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2 Literature Review</td>
<td>4</td>
</tr>
<tr>
<td>2.1 Loan Haircut and Value at Risk</td>
<td>4</td>
</tr>
<tr>
<td>2.2 Concentration Measure</td>
<td>7</td>
</tr>
<tr>
<td>2.3 Bond</td>
<td>12</td>
</tr>
<tr>
<td>3 VaR and Herfindahl Index in a Gaussian framework</td>
<td>15</td>
</tr>
<tr>
<td>3.1 Definition and Methodology</td>
<td>15</td>
</tr>
<tr>
<td>3.1.1 Back Testing Framework</td>
<td>15</td>
</tr>
<tr>
<td>3.1.2 Asset return and risk</td>
<td>18</td>
</tr>
<tr>
<td>3.1.3 Value at Risk LTV</td>
<td>19</td>
</tr>
</tbody>
</table>
4 Student-t VaR LTV and Arbitrage Pricing Theory

4.1 Definition and Methodology

4.1.1 Actual Portfolio Return

4.1.2 Diversification Effect

4.1.3 Simplification of DF under Arbitrage Pricing Theory

4.1.4 Performance Evaluation

4.2 Data and Implementation

4.3 Result and Discussion
5 Multifactor framework

5.1 Definition and Methodology .................................................. 61
  5.1.1 Further Simplification of Herfindahl Index .......................... 62
  5.1.2 Multiple Factor Model ..................................................... 64
5.2 Data and Implementation ....................................................... 70
5.3 Result and Discussion .......................................................... 72
5.4 Final comparison ............................................................... 79

6 Bond LTV

6.1 Definition and Methodology ................................................... 82
6.2 Data and Implementation ....................................................... 85
6.3 Result and Discussion .......................................................... 90

7 Summary

List of tables

List of figures
Chapter 1

Introduction

The guideline of Basel committee has defined a standard collateral valuation methodology as well as an advanced one. Eligible banks are also allowed to develop their own-estimated model to calculate loan haircut. The goal of this thesis is - with a certain analogy to the Basel committee rules - to design a transparent and simple method for setting loan limit, which is determined by the market risk of the underlying collateralized portfolio. The loan haircut (or one minus LTV, loan to value) is closely connected to VaR (Value at Risk), a risk measure widely used in the finance industry. It is well known that VaR estimation is very sensitive to model assumptions and time consuming to calculate, especially in the multivariate case. Here, we strive to produce the VaR estimate with a simple Rule-Based system that is robust and perform well with back testing. Initially, model based VaR LTV is used as a benchmark to which the Rule-Based LTV will be evaluated. Later, the legitimacy of such model based LTV will also be investigated.

There are two ways to evaluate the performance of the RB LTV. The first criterium is the
coefficient of regressing the RB LTV against the VAR LTV. The ideal value of the slope of the linear regression should be close to 1 and the intercept being 0. The second criterium is then back testing. By comparing the portfolio LTV to the actual portfolio return, we can assess the reliability of the RB LTV system. Define the scenario where actual loss exceeds LTV as a breach. We expect the percentage of breaches to be equal to the risk appetite, on which the LTV calculation is based.

While the Lombard loan collateralized portfolio may consist of a wide range of assets, this thesis focused on only two major asset classes: equity and bond. In particular, in the four main studies of this thesis, the first three are based on equity portfolios sampled from a universe of 50 and 500 international stocks, while the forth study is based on bond portfolios sampled from 500 international bonds. The aim, methodology scope of the four studies are summarised below:

1) VaR and Herfindahl Index in a Gaussian framework This chapter produced two sets of LTVs; VaR LTV based on normal distribution, and the Rule-Based LTV based on a lookup table. The lookup table assigns a VaR according to the volatility of single stock, which when summed together according to investment weights gives the Rule-Based LTV for the portfolio. Since, in this simple Rule-Based system, the correlation and concentration effect is omitted, we used the Herfindahl Index to make adjustments to the Rule-Based LTV.

2) Student-t VaR LTV and Arbitrage Pricing Theory In this chapter, the Student-t distribution is used for calculating the VaR LTV. An analytical expression for a Diversification Factor (DF) is used to replace the crude Herfindahl Index. The Abitrage Pricing Theory
framework is used here to simplify the calculation of asset correlation.

3) Multifactor framework This chapter adopts a new framework based on a multifactor model proposed by Bluhm (2004). The Diversification Factor now has two components; an average correlation and a volatility modified Herfindahl index. By using the multi-factor model, the portfolio market risk can now be analysed according to region and industry concentration.

4) Bond LTV This last chapter calculates LTV for a collaterised bond portfolio. The bond return depends on yield difference, which can be further broken down to swap rate and credit default swap spread. Here, several factors such as rating, duration, region and industry are used to measure bond portfolio concentration risk.
Chapter 2

Literature Review

2.1 Loan Haircut and Value at Risk

Under Basel II guidance[1], banks have two ways to calculate a loan hair cut: the standard supervisory haircuts or an own-estimated haircuts. These haircuts depend on the different types of instruments, transactions as well as their marked-to-market frequencies. A further alternative to these two methods is for the banks to use VaR models for calculating potential price volatility. This haircut will later be used to calculate the exposure of a loan as well. In paragraph 150 of the BIS document "Basel II: Revised international capital framework(comprehensive version)”, it is defined that the portfolio haircut is the weighted average of the individual haircuts. For the own estimation procedure, Basel II has laid out specific requirements for the risk factors regarding bond: issuer, rating, time to maturity and modified duration. There are also other quantitative criteria for an own-estimation for loan hair cut:
1). A 99th percentile, one-tailed confidence interval is required.

2). Holding period should be determined in reference to the instrument types. For example, equities, which belongs to the category "other capital market transactions", has a holding period of 10 business days.

3). A minimum of historical observation period of 1 year for historical VaR is required. And the dataset needs to be updated at least once every three months.

4). There is no prescription for particular models. For example, the banks are free to choose from historical simulations and Monte Carlo simulation.

Therefore, it seems quite natural for banks to build up their collateral valuation methodology based on a Value at Risk model. Although challenged by the new risk measure like expected shortfall, Value at Risk is still one of the most widely used market risk measurement techniques. Acerbi and Tasche [2] has discussed the incoherent shortcomings of Value at Risk as the risk measure. It suffers critique since it’s not a coherent risk measure when the model does not assume Gaussian distribution. Another shortcoming is that a variance-covariance Value at Risk measure requires per-determined distribution assumptions for the underlying model, which needs to be clarified to avoid confusion. It is still a preferred as best practise methodology by many bank practitioners because of many of its benefits: simplicity, wide applicability to any financial instruments, etc. And most importantly, VaR still statisfies the coherent property when we assume Gaussian distribution for the asset returns. Sometimes when it is difficult to assess risk for the whole company, the best practice is to built separate risk assessment for each segment and sum them up in a "bottom-up" manner.
The performance of a Value at Risk model can be evaluated by backtesting. It compares the forecast of potential loss with the real profit and loss data. Cassidy and Gizycki (1997) discussed several back testing methodologies when using data from Australian banks. Under the capital adequacy arrangements proposed by the Basle Committee, each bank must meet a capital requirement expressed as the higher of: (i) an average of the daily VaR measures on each of the preceding sixty trading days, adjusted by a multiplication factor; and (ii) the bank’s previous day’s VaR number. The determination of the multiplication factor is based on the exceptions, or breaches of the backtesting results. The underlying assumption is that, if the percentage of breaches are larger than one minus the confidence level, it is indicated that the model has problems. Hence the higher the breach rate, the higher the multiplication factors. They have tested two aspects of the breach: the time between two breaches and the percentage of breaches. The authors have concluded that the statistics based on the first aspect of breach has little power and doesn’t reject VaR models, which underestimate risk. The test statistics based on the second aspect improves on testing power, yet it requires a large sample size to have significant test power. They have also conducted the following test: variance testing on the variance of losses estimated by VaR and the variance of the real P&L figure; normality test on the distribution of P&L and a risk-tracking test which assess the correlation between VaR estimate and the magnitude of the daily P&L. The authors concluded that although these tests can provide useful information to risk managers and bank practitioners, they fail to meet the precision level as laid out in the regulatory treatment by Basel Committee.

In the following studies, we are going to apply relatively simple backtesting methodologies to evaluate the performance of the LTV system and focus on developing a simple and
practical system.

## 2.2 Concentration Measure

Concentration risk is a more thoroughly researched in credit risk area than in that of market risk. Abundant literatures can be found in concentration measure in credit portfolio management. Hence we will start with concentration in credit risk and later derive a similar methodology for the market risk aspect. It is noted that the underlying model for internal ratings-based risk weights of Basell II, namely the Asymptotic Single-Risk Factor model assumes a perfectly fine-grained portfolio. Thus it rules out the possibilities of name or sector concentration, which is not the case in reality. Yet this assumption serves as one important basis for the bottom-up approach of risk aggregation. An informal survey conducted by BIS[5] shows that concentration is often dealt with in the form of "diversification" among bank experts. It then warrants capital relief relative to the capital requirement defined in Pillar 1 of Basel II. Some banks will employ simple concentration measures like Herfindahl-Hirschman index while others use more sophisticated internal model.

Luetkebohmert(2008)[6] has described several ad hoc concentration measures, which include concentration ratio, Lorenz Curve, Gini Coefficient as well as Herfindahl index. There are also other model-based approaches such as granularity adjustments, which address the issue of name concentration. It is first proposed by Gordy(2004)[7] and later refined by Gordy and Luetkebohmert(2006)[8]. The major spirit of granularity adjustment is to construct a Taylor series to approximate the difference between true loss quantile
and the conditional loss quantile provided by IRB model. Although several revised form of the granularity adjustment has been proposed for improvement, two shortcomings still prevail: 1. the asymptotic nature of granularity adjustment requires the portfolio size not to be too small, or the accuracy will suffer. 2: The different model basis of IRB formulae and granularity adjustment makes it impossible to compare them and test the adjustment.

Sector concentration, on the other hand, violates the assumption of a single-factor model. In reality, often the different constituents of a credit portfolio are driven by several sector factors instead of one sole systemic factor. An average of sector correlations plays one important role here. If all sectors are perfectly correlated, there is no difference between the results obtained from a single-factor model and a multi-factor model. There are several works aiming at investigating the gap between the single-factor model used in Basel II IRB and a multi-factor model as well as bridging this gap. An ideal solution is to develop a multi-factor model, which does not rely heavily on Monte Carlo Simulation and can be expressed in close-form. Pykhtin(2004)[9] proposed an analytical methodology to calculate VaR and Expected Shortfall under the multi-factor Merton framework. This technique is then tested and refined by several literatures. Although it provides an good alternative to Monte Carlo simulation, it is still too mathematically complicated to be widely used. It also requires more input parameters when comparing to other models. Garcia Cespedes et al(2005)[10] derived a scaling factor which can convert the economic capital obtained from a single-factor model to one from of a multi-factor model. Their methodology has the advantage of being intuitive and can be simply tabulated for application. Yet this model requires recalibration of asset correlations and there are still some difference between its underlying assumptions and those of the Basel II IBR. Nevertheless, we feel that their
work is most relevant and base our study on a similar framework to theirs. A detailed summary of their work is presented below.

Caspedes et al have investigated the capital requirements under the Basel II Capital Accord and proposed a relatively simple and intuitive methodology to adjust for the diversification problem of the capital requirement. They introduced a capital diversification factor, which transforms the capital requirement from a one-factor model to the diversified capital based on a multi-factor model. This allows them to express the diversified capital as the sum of stand-alone capital requirements. The capital diversification can be decomposed further to different sources. This feature is then very desirable for risk management purpose.

They have modified the credit risk single factor model in Basel II by introducing different sectors and different risk factors. Instead of modeling one obligor’s creditworthiness with one systemic factor, they model each obligor’s default probability with a sector random variable, to which this obligor belongs. Then these sector variables will be connected by an underlying systemic factor.

\[
Y_j = \sqrt{\rho_k}Z_k + \sqrt{1 - \rho_k}\varepsilon_j \tag{2.1}
\]

\[
Z_k = \sqrt{\beta_k}Z + \sqrt{1 - \beta}\eta_k \tag{2.2}
\]

\(\rho_k\) is the factor load for sector k, and \(\beta\) the single correlation parameter for all factors. The risk capital calculated using this modified model is considered as the stand-alone capital
for each sector.

\[ VaR_k(\alpha) = \sum_{j \in \text{Sector}_k} LGD_j \times EAD_j \times N \left( \frac{N-1(PD_j) - \sqrt{\rho_k z^\alpha}}{\sqrt{1-\rho_k}} \right) \]  \hspace{1cm} (2.3)

The overall required capital for a portfolio would be the sum of all stand-alone capitals, assuming perfect correlation among sectors.

\[ C^{1f} = \sum_{k=1}^{K} C_k \]  \hspace{1cm} (2.4)

They then defined a capital diversification factor DF, which can convert added-up stand-alone capitals to the economic capital calculated using a multi-factor model.

\[ C^{mf} = DF \sum_{k=1}^{K} C_k \]  \hspace{1cm} (2.5)

It consists of two parameters: the cross-sector correlation \( \beta \) and the capital diversification index CDI, which is based on the relative size of each sector’s capital. The latter is an application of Herfindahl Index to the stand-alone capital of each sector.

\[ CDI = \sum_k \frac{C_k}{(C^{1f})^2} = \sum_k w_k^2 \]  \hspace{1cm} (2.6)

\( C_k \) is the capital contribution for sector \( k \) and \( C^{1f} \) being the capital requirement obtained from single-factor model. If losses are normally distributed, a closed-form expression can be obtained for the diversification factor.

\[ DF^N = \sqrt{(1-\hat{\beta})CDI + \hat{\beta}} \]  \hspace{1cm} (2.7)
\( \hat{\beta} \) is the correlation of credit loss correlation. In the practice, where the distribution is not normal, the diversification factor is approximated by the close-form expression under Gaussian distribution. The authors carried out numerical calculation using this approximation and obtained reasonable results. They have also applied this model to capital allocation and risk contribution. They assign marginal diversification factor for each sector and further decompose the source of diversification to overall portfolio diversification, sector size and sector correlation.

\[
C^{mf} = \sum_{k=1}^{K} DF_k C_k
\]  \hspace{1cm} (2.8)

\[
DF_k = DF + \Delta DF_{Size} + \Delta DF_{Corr}
\]  \hspace{1cm} (2.9)

Here they introduced the concept of average sector correlation, which greatly facilitate the derivation process.

To test this approximation methodology, the authors have calculated risk capital based on both multi-factor model and the DF-adjusted single-factor model. They used Monte Carlo simulation to calculate for 3000 portfolios and obtained a close fit between these 2 models. The authors have also varied the parameters for calculating diversification factor to produce a DF surface. Their empirical results confirmed the validity of the theoretical expression of DF under Gaussian distribution. Finally, they have also argued that the benefit of this model lies in its intuitiveness as well as simplicity. Its parameters can be used as risk and concentration indicators. And the model itself can be easily tabulated and used as a basis for regulatory rules or capital allocation.
2.3 Bond

In this asset class, market risk and credit risk are inter-related since the credit risk will be also affecting the fixed income securities secondary market prices. Yield spread between emerging market bond and the bonds which is considered virtually risk free with similar maturity the additional compensation for investors to hold this riskier bond. It reflects the credit worthiness of the bond issuer. During market turbulence, Cunningham, Dixon and Hayes [11] have found that increased correlation in emerging market yield spreads. For example, in 1998 during the Russian Crisis, high correlations among more than 10 emerging markets are observed. Also they recorded strong directional co-movement during this period. Combining credit rating and the yield spread, the credit riskiness of bonds can be better explained. Cautions should apply since most ratings published for Moody are corporate bond ratings. The sovereign default experience can differ from the corporate one. If yield spread is used alone, then risk preference and liquidity risk are ignored. The possibility of recovery after default and the implications of debt structure for pricing are also not considered.

Daul, Sharp and Sorensen (2010) [12] decompose the local bond return to several factors. Apart from the important factors like yield curve, interest volatility and credit spread, they also consider currency exchange rate, carry interest rate and inflation component. If the security is denominated in a foreign currency, then the return needs to be adjusted by the currency return as well. They decompose this currency component into two items: forward premium, which is the expected interest rate difference and the currency surprise, the unexpected currency movement. This splitting is advantageous for the portfolio manager.
to have realistic hedge goals. The carry return is the holding period return even though the bond market conditions stay the same. Dual et al have divided the carry return further into two categories: the coupon return, which is due to accrued interest rate and coupon payment and the roll-down return, the change of clean prices when yield curve stays the same. Although complex mathematical models may be applied to treat the roll-down return, it does not have strong influence over the bond return. Yield curve change has significant meanings to bond investors. It is the major driver for bond return and the short rate presented by yield curve is an important policy tool. There are two kinds of yield curve change: parallel shift and slope change. Using the traditional method, effective duration and average change in yield curve, which is calculated by averaging the rate change at each chosen node, the return due to yield curve parallel shift can be calculated. Key rate duration addresses the bond return due to yield curve slope change. To capture the total yield curve influence, shift return and shape return should be combined together. The key rate duration model has advantage over the traditional mode, which uses twist and butterfly effect. It also does not require a stationary covariance structure of interest rate changes like the principal component models. Credit component of bond return is the extra return compensation for default risk bond investors are facing. In literatures, the US Treasury is considered virtually risk free. The difference between risk free bond yield and other bonds are the credit risk spread. In their paper, Daul, Sharp and Sorensen propose to model the return due to credit risk by multiplying spread duration and option-adjusted spread. They also distinguish between the sector spread and idiosyncratic spread, which can only be explained by company or country specific news. To price an option-embedded bond, interest rate volatility is required. By constructing a volatility curve based on
cap price and setting several nodes on the volatility curve, an average shift of market implied volatility curve can be calculated. In this case, a bonds vega is considered as the volatility sensitivity. The volatility return can be obtained by multiplying vega and average volatility shift. For those inflation-linked bond, it is also necessary to consider the inflation component. Daul et al use the same methodology by constructing an inflation curve, taking the average inflation curve shift and computing the return due to inflation with the inflation duration.
Chapter 3

VaR and Herfindahl Index in a Gaussian framework

3.1 Definition and Methodology

3.1.1 Back Testing Framework

"Lombard Credit" typically involves debtor pledging assets for the lombard loan. The word Lombard is an Italian region where such loans were originated. Nowadays, a Lombard portfolio refers to the portfolio of securities a client pledged for loan as the collateral. Typically, this loan will be used to finance investment in the capital market. From a lenders perspective, it is important to recover the loss of loan in the event of default with the collateral. Since the value of asset portfolio presented as collateral also fluctuates, it is important to consider the potential maximum loss within a given time horizon of the
collateral portfolio. The original collateral value less the possible maximum loss represents the value a lender may recover, at the event of default. This loan to value \((LTV)\) ratio reflects the amount of “hair-cut” a lender will apply when granting loan that is pledged against a portfolio of assets. To decide on the hair-cut (one minus Loan to Value), Value at Risk calculation is considered as the best candidate. \(^1\) It measures the maximum loss on a portfolio at a confidence level \(\alpha\), in a given holding period. In our study here, the holding period is 10 days, which is closely related to Basels 10-day Value at Risk.

Yet the drawbacks of this approach are obvious. First of all, it only concerns market risk; secondly, it is too time-consuming to do it for all clients; finally, this procedure is not very transparent to clients either.

For practical operation, it is cumbersome to calculate VaR for each collateral portfolio; volatility estimate, for example, will be different depending on the estimation procedure and the sample period used. Moreover, not all asset in the portfolio will have the same amount of history. It is difficult to maintain transparency and consistency for all borrower without lengthy procedure and documentation. Hence, Credit Suisse uses a Rule-Based system which a simplified procedure for calculating VaR for a large number of collateral portfolios, which is highly efficient and transparent. It consists of two steps. Firstly, instead of calculating the portfolio LTV directly, it considers every single constituent of the collateral portfolio. With a pre-determined lookup table, which is calibrated to the risk drivers affecting the asset value, the stand alone LTV for each asset is determined. The portfolio LTV is approximated by the weighted average of the stand alone LTV, omitting

\(^1\)The portfolio VaR may not be the perfect solution. Since certain risks e.g. legal risk and country risk are also difficult to map in the VaR calculation.
the diversification effect between assets.

The implicit assumption of the first step is that all the correlation coefficient equals to one\(^2\). Hence the second step is to study the portfolio’s concentration risk and compensate the diversification effect, which is omitted in the first step. To achieve this, we define a concentration measure and assess its relevance. Factors for concentration will be identified and used to decide how much the LTV should be adjusted. The final result, after correction for the concentration effect, is called the Adjusted Rule-Based LTV.

In this chapter, the actual portfolio is used to evaluate the Rule-Based LTV and its efficiency. The evaluation is done via a back testing procedure. For each portfolio, there are two LTV values: the actual portfolio VaR and the unadjusted/adjusted Rule-Based LTV. Ideally when the two LTV values of the same group of portfolios are plotted against each other, the Rule-Based System LTV should have a linear relationship with portfolio VAR LTV. The spread of this linear relationship can be used to assess the efficiency of the Rule-Based System. Moreover, when compared to the realized portfolio return, the number of breaches also indicates the performance of Rule-Based systems. Hence the back testing consists of two evaluations: the result of approximation and the potential loss from historical back testing.

\(^2\)Note that we are considering volatility here. If we are interested in variance, the coefficient should equals to zero.
3.1.2 Asset return and risk

Let \( r_{i,t} \) denote 10-day return for stock \( i \) at time \( t \) (\( t \) is measured at 10-day interval).

Return on asset \( i \) at time \( t \) is defined as

\[
\begin{align*}
  r_{i,t} &= \ln \left( \frac{P_t}{P_{t-1}} \right) \\
  P_t &= P_{t-1} \exp(r_{i,t})
\end{align*}
\]

If \( r_{i,t}^\alpha \) denote the \( \alpha\% \) lower quantile of the returns distribution (s.t. \( r_{i,t}^\alpha < 0 \)), then the \( LTV \) per dollar of asset value is, dividing both side of (3.2) by \( P_{t-1} \),

\[
LTV_{i,t}(\alpha) = \exp(VaR_{i,t}^\alpha) = \exp(r_{i,t}(\alpha)).
\]

For convenience, we will drop the reference value \( \alpha \) and the superscript \( \alpha \) from \( LTV \) and \( VaR \) from now on since it is understood that their definitions are based on a specific value for \( \alpha \). In this study, we will first focus on \( \alpha = 1\% \) scenario. Later the confidence level \( \alpha \) will vary between a certain range.

Here, we assume that return volatility is the key and only driver for asset Value at Risk. Let \( \sigma_{i,t} \) be the volatility of return on asset \( i \) at time \( t \) estimated using historical returns and exponential weighted moving average smoothing as follows: The reason for choosing this methodology is given in 3.1.3.

\[
\sigma_{i,t} = \frac{1 - \lambda}{1 - \lambda^{m+1}} \sum_{a=0}^{m} \lambda^a r_{i,t-a}^2
\]

where \( m \) is the total number of return observations available prior to \( t \), \( \lambda=0.94 \) is the decay factor.\(^3\)

\(^3\)The value of the decay factor is suggested by Fleiming, Kirby and Lstdiek.[13]
Similarly, covariance for returns of stock $i$ and stock $j$ at time $t$ can be written as:

$$\sigma_{ij,t} = \rho_{ij,t} \sigma_{i,t} \sigma_{j,t} = 1 - \frac{\lambda}{1 - \lambda^{m+1}} \sum_{a=0}^{m} \lambda^a r_{i,t-a} r_{j,t-a}$$  \hspace{1cm} (3.5)

In our setting, each portfolio consists of $N$ stocks. Each stock has weight $w_i$, which is held constant throughout the whole period, assuming that the investor will keep rebalancing the portfolio.

### 3.1.3 Value at Risk LTV

The idea of Value at Risk was introduced by RiskMetrics [14] in 1996, which reduces the riskiness of return distribution of portfolio to a single number. Since then, the methodology used by RiskMetrics is therefore considered as a benchmark model, the discussion and improvements of which were addressed by hundreds of papers in the following decade. Christoffersen, Hahn and Inoue [15] used a conditional moment framework to compare four VaR model volatility specifications: RiskMetrics, GARCH, Implied volatility, Reprojected volatility. They found the main difference between these models lies in the measurement of volatility used.

Christoferksen et al. tested four VaR models with variance specification of the 4 models using daily returns of S&P 500 from November 1985 to October 1994 with different level of significance $\alpha$. The authors found that at 1% level of significance, none of the four models for 1-day VaR are rejected from the null hypothesis that the model specification is appropriate, although the result deteriorates as $\alpha$ increases. On a pairwise comparison, RiskMetrics model is universally preferred with 1% level of significance.
Hence, in this study we generate the 1% percentile of portfolio return using the one step RiskMetrics VaR model is sufficient and appropriate. Moreover, since the time step in (3.1) is measured at 10-day interval, we omit the dynamics of daily returns when modelling the 10-day VaR.

Given the return scenario vector \( r_{p,t} = (r_{1,t}, \ldots, r_{i,t}) \) and the weight \( w_{p,t} = (w_{1,t}, \ldots, w_{i,t}) \) for stock \( i \), the return for portfolio \( p \) can be calculated as:

\[
r_{p,t} = \sum_{i=1}^{N} r_{i,t} \times w_{i}
\]  

Consider the portfolio volatility of portfolio \( p \):

\[
\sigma_{p,t}^2 = \sum_{i} w_{i}^2 \sigma_{i,t}^2 + \sum_{i} \sum_{j} w_{i} w_{j} \rho_{ij,t} \sigma_{i,t} \sigma_{j,t}, i \neq j
\]  

The correlation coefficient \( \rho_{ij} \) between stock \( i \) and stock \( j \) is defined in equation (3.5).

Under the Gaussian assumption, \( r_{i,t}^G \) in equation (3.3) can be computed analytically. Let \( r_{i,t}^G \) with superscript \( G \) denote the Gaussian \( \alpha \)% lower quantile, We have dropped the time subscript for convenience,

\[
r_{p}^{G} (\alpha) = \sigma_{p} \Phi_{G}^{-1} (\alpha) + \mu_{p}
\]  

where \( \Phi_{G}^{-1} (\cdot) \) denote the inverse of normal cumulative density function and \( \mu_{p} = \Sigma_{i} w_{i} \mu_{i} \) denotes the mean portfolio return. Assuming that the stock returns follow a martingal process at the 10-day horizon, the mean returns are zero. The estimate and forecast of the short-term return could be extremely noisy. Empirical evidence suggests that zero mean assumption leads to the best volatility and risk forecast.

From equation (3.3), the Loan to Value based on VaR model for portfolio \( p \) is given by:

\[
LTV_{p}^{VAR} = \exp(r_{p}^{G} (\alpha))
\]
3.1.4 Rule-based System LTV

The main feature of the Rule-Based system is that the empirical distribution of asset return of the entire asset class is used to calculate Value at Risk. The empirical distribution need not be the historical distribution of the asset concerned but generally calibrated to historical asset distribution of the entire equity using, for example, return volatility as the risk driver. It is important that this historical returns sample covers a wide range of stressed and calm periods that asset \(i\) may or may not have experienced. This historical empirical distribution, in the form of a lookup table, is used to calculates stand alone VaR for individual assets based on its volatility. The weighted average lookup value \(LTV^L\) of constituent assets become the unadjusted portfolio \(LTV^U\). Next, correlation among assets in the portfolio is then re-introduced to assess portfolio concentration and degree of diversification. Here the definition of concentration and its relevance jointly decide how much \(LTV^U\) should be adjusted upwards.

If we divide the historical cross-sectional volatility estimates into buckets, \((\sigma_1, \sigma_2, \ldots, \sigma_{l-1}, \sigma_l, \ldots]\), then

\[
r_i^E(\alpha) = \Phi_{l}^{-1}(\alpha) \quad \text{with } \sigma_i \in (\sigma_{l-1}, \sigma_l].
\]

(3.10)

\(r_i^E\) is the empirical \(\alpha\)% lower quantile \(\Phi_{l}^{-1}\) of all historical returns that fall into volatility bucket \(l\). Since the value of this quantile is simply “looked up” from a historical table, the \(LTV\) derived on this basis is called “look-up” \(LTV\), and is calculated according to (3.3) as
\[ LTV_i^E = \exp \left( r_i^E \right). \] (3.11)

To facilitate our theoretical comparison of the empirical versus Gaussian approaches, we might what to refer to an Gaussian implied volatility, \( \hat{\sigma}_i^G \), such that (3.10) can be expressed in the same way as (3.8):

\[ r_i^E (\alpha) = \hat{\sigma}_i^G \Phi^{-1}_G (\alpha) + \mu_i. \] (3.12)

Finally, the unadjusted portfolio \( LTV_p^U \) for portfolio \( p \) is given by:

\[ LTV_p^U = LTV_p^{\Sigma,E} = \sum_{i=1}^{N} w_i LTV_i^E = \sum_{i=1}^{N} w_i \exp (r_i^E) \] (3.13)

Recall the Value at Risk under Gaussian assumption in equation (3.8). Under the Gaussian assumption, the relationship between URB LTV calculated in (3.11) and (3.12) and the weighted average of individual Gaussian Value at Risk of each stock is:

\[ \sum_{i=1}^{N} w_i r_i^E (\alpha) = \sum_{i=1}^{N} w_i (\hat{\sigma}_i \Phi^{-1}_G (\alpha) + \mu_i) \approx \sum_{i=1}^{N} w_i (\sigma_i \Phi^{-1}_G (\alpha) + \mu_i) = \sum_{i=1}^{N} w_i \sigma_i \Phi^{-1}_G (\alpha) + w_i \mu_i \] (3.14)

\[ sum_{i=1}^{N} w_i \sigma_i \Phi^{-1}_G (\alpha) + w_i \mu_i = \sum_{i=1}^{N} w_i \sigma_i \Phi^{-1}_G (\alpha) + \mu_p \] (3.15)

From (3.8) and (3.15) it is clear that the source of difference between the empirical lookup LTV and Gaussian portfolio LTV lies in the first item of RHS. Recall (3.7), which gives the calculation of portfolio volatility:
$$\sigma_p = \sqrt{\sum_i w_i^2 \sigma_i^2 + \sum_i \sum_j w_i w_j \rho_{ij} \sigma_i \sigma_j, i \neq j}$$

When correlation is 1 for all assets, the portfolio volatility becomes:

$$\sigma_p = \sqrt{\sum_i w_i^2 \sigma_i^2 + \sum_i \sum_j w_i w_j \sigma_i \sigma_j, i \neq j} \sigma_p = \sqrt{\left(\sum_i w_i \sigma_i\right)^2} = \left(\sum_i w_i \sigma_i\right)^2$$

equations (3.8) and (3.15) give the same value.

When correlation is smaller than 1, $\sigma_p$ will be smaller than $\sum_i w_i^2 \sigma_i^2$ and hence the Gaussian VAR LTV is higher than URB LTV.

The difference between these LTVs is shown in figure 1 for different values of correlation and average volatility. To illustrate the relationship between equation (3.8) and (3.15), we setup a simple portfolio with only two stocks with identical return mean and volatility. The only variable is the correlation coefficient. The three curves are the difference between VAR LTV and RB LTV with different volatility. They all converge to zero as the correlation reaches 1. All three curves exhibit the form of exponential function. This feature is used to device the correction form for adjustment. To correct for the omitted diversification in (3.15), compensation should be added to URB LTV. In this study, we used concentration as a proxy of correlation. The more concentrated is the portfolio with respect to our chosen factors, the higher the correlation among assets in the portfolio.
3.1.5 Concentration Risk

Concentration here means heavy dependence on certain factors. The literature on concentration risk has most frequently considered geographical and industrial concentration. We will start with a single factor geographical concentration and extend it to a two factor model later in the next chapter.

The Herfindahl-Hirschmann index was used originally to assess the degree of monopoly in a market. Kelly(1981)[16] has conducted an overview of the usage of Herfindahl Index and concluded that it was broadly adopted to compute the degree of concentration. Let $N_p$ be the total number of stocks in portfolio $p$, and $N_{p,k}$ denote the number of stocks belonging to country $k$ in portfolio $p$. Let $M_p$ be the number of countries existing in portfolio $p$, Herfindahl index is defined as:

$$H_p = \sum_{k=1}^{M_p} s_{p,k}^2$$  \hspace{1cm} (3.16)
where \( s_{p,k} \) is the investment share of geographical factor \( k \) in portfolio \( p \):

\[
 s_{p,k} = \sum_{ic=1}^{N_{p,k}} w_{ic}
\]  

(3.17)

Here the Herfindahl index ranges between \( 1/M_p \) to 1; the former indicates fully diversified case while the latter implies a dominant concentration on one country. In implementation, the Herfindahl index is normalized as follows:

\[
 H^* = \frac{H_p - 1/M_p}{1 - 1/M_p}
\]  

(3.18)

The normalized Herfindahl index now ranges from 0 to 1. The larger the value of \( H^* \), the more concentrated the portfolio becomes.

The next step is to determine the relevance of this concentration measure, i.e. how well this concentration describes the correlation. The form of correction should also be specified. In this study, the relevance and the form of correction will be empirically driven. This will be explained in details in section 2.3.1.

### 3.1.6 Performance evaluation

Since we are interested in how closed the RB LTV is matched to the target: portfolio Gaussian VAR LTV. \(^4\) We can employ a simple linear regression here:

\[
 LTV_{p,t}^{VAR} = \alpha + \beta LTV_{p,t}^{U} + \varepsilon_{p,t}
\]  

(3.19)

\(^4\)The reason for choosing the Gaussian distribution here is because of its linear correlation. Although the final goal is to find a methodology which withstand the backtesting, it should also outperform the benchmark which is the Gaussian LTV.
If Rule-Based system LTV is efficient and replicates the Gaussian VAR LTV exactly, $\alpha$ should equals 0 and $\beta$ equals 1. The variance of $\varepsilon$, in this case, should approach zero as well. Moreover, a scatterplot can also provide some intuitive insight to the relation between these two LTV systems. A perfect replication of the target will yield a 45 degree straight line in the plot.

### 3.2 Data and Implementation

#### 3.2.1 Data Description

Table 3.1 summarizes the 50 stocks selected from the MSCI world index constituents list. There are 10 countries of origin and 5 stocks from each country. To better account for the concentration effect, region is used instead of country, since a region like Europe instead of France or Italy may better describe the common characteristics of stock market conditions.

In order to obtain as many observations as possible, stock price are collected over a 14-year period from 01.02.1996 to 01.02.2010. All together, there are 365 return observations per stock when they are calculated at a 10-day interval.

On each trading day, the adjusted price, which removes the effect of corporate actions such as dividend or stock split, is used to compute the 10-day return. Each price is given in its local currency to avoid the confounding effect of currency exchange rate fluctuation.

To better test the model, the whole period of data is divided into two subpeiods. Excluding the first 2 years used in EWMA volatility calculation, the following 6 years from
Table 3.1: MSCI Stock Information.

<table>
<thead>
<tr>
<th>No. of Stocks</th>
<th>Country</th>
<th>Region</th>
<th>Currency</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Austria</td>
<td>Europe</td>
<td>Euro</td>
</tr>
<tr>
<td>5</td>
<td>Brazil</td>
<td>Latin America</td>
<td>Brazilian Real</td>
</tr>
<tr>
<td>5</td>
<td>China</td>
<td>Far East</td>
<td>Chinese Yuan</td>
</tr>
<tr>
<td>5</td>
<td>Germany</td>
<td>Europe</td>
<td>Euro</td>
</tr>
<tr>
<td>5</td>
<td>Japan</td>
<td>Far East</td>
<td>Japanese Yen</td>
</tr>
<tr>
<td>5</td>
<td>South Africa</td>
<td>Africa</td>
<td>South African Rand</td>
</tr>
<tr>
<td>5</td>
<td>Sweden</td>
<td>Europe</td>
<td>Swedish Krona</td>
</tr>
<tr>
<td>5</td>
<td>Thailand</td>
<td>Far East</td>
<td>Thai Baht</td>
</tr>
<tr>
<td>5</td>
<td>Turkey</td>
<td>Middle East</td>
<td>Turkish Lira</td>
</tr>
<tr>
<td>5</td>
<td>United States</td>
<td>North America</td>
<td>US Dollar</td>
</tr>
</tbody>
</table>

1998 to 2003 are used for calibration of the lookup table, parameter estimation and the in-sample fit whiles the last 6 years from 2004 to 2010 are used for the out-of-sample performance evaluation. In the out-of-sample period, the RB LTV are based on lookup table calibrated in the in-sample period and will be regressed against the VAR LTV with up-to-date information. The unadjusted LTV will be compared to the VaR LTV first and the result will be examined to motivate the exact form of correction function for correlation.

### 3.2.2 Implementation

The program to carry out all the calculations is coded in Matlab. The procedure is described as following:
After reading the daily adjusted price, volatilities at 10-day interval of each stock on each single trading day are calculated using the exponential weighted moving average method, as defined in equation (3.4). The numbers of 10-day periods used to calculate volatility is 50, which equates the first 2 years. Therefore only 12 years out the original 14 are left for subsequent computation.

For a portfolio \( p \) consisting of 5 stocks, let \( r_{p,t} = (r_{1,t}, \ldots, r_{5,t}) \) be a 5 \times 1 vector of asset returns. We use the variance-covariance matrix \( \Sigma_t \) for the \( N \)-dimensional multi-normal distribution as defined in equation (3.5) to calculate the portfolio volatility, \( \sigma_{p,t} \) and then the analytical portfolio VaR. Multivariate normal distribution is considered here. The mean vector of \( r_t \) is assumed to be a vector of zeros. The covariance matrix \( \Sigma \)

1000 combinations of a 5-asset-portfolio are produced along with the randomly assigned weight for each component. They are assumed to be constant throughout the whole sample period.

There are 315 trading days and 1000 portfolios on each trading day. For each portfolio, the VAR LTV, Rule-Based LTVs unadjusted for concentration as well as the Herfindahl indices are calculated, after which the relevance of concentration is used to produce the adjusted RB LTV. On each day, 1000 LTVs for each type are stored for the final efficiency test.

The calibrated LTV lookup table is summarized in table 3.2. Instead of dividing the volatilities by arbitrarily bins, we use the quantile to construct lookup table slots. In this way, we can also make sure that in each bin there are enough data observations. As shown in the table, the higher the volatility, the lower the LTV value is assigned. For
better presentation, the volatilities in the table are annulized. The observations for each

<table>
<thead>
<tr>
<th>Quantile</th>
<th>Volatility</th>
<th>LTV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 5%</td>
<td>$\leq 0.0421$</td>
<td>0.9546</td>
</tr>
<tr>
<td>5%-10%</td>
<td>0.0421-0.0515</td>
<td>0.9176</td>
</tr>
<tr>
<td>10%-25%</td>
<td>0.0515 - 0.0650</td>
<td>0.8829</td>
</tr>
<tr>
<td>25%-50%</td>
<td>0.0650 - 0.0856</td>
<td>0.8449</td>
</tr>
<tr>
<td>50%-75%</td>
<td>0.0856 -0.1180</td>
<td>0.7871</td>
</tr>
<tr>
<td>75%-90%</td>
<td>0.1180 - 0.1625</td>
<td>0.7087</td>
</tr>
<tr>
<td>90%-95%</td>
<td>0.1625-0.1881</td>
<td>0.6667</td>
</tr>
<tr>
<td>Above 95%</td>
<td>$\geq 0.1881$</td>
<td>0.5281</td>
</tr>
</tbody>
</table>

LTV in both in-sample and out-of-sample periods are presented in scatter plots. The regression on efficiency test is run for the whole period. The performance of the RV LTV system will be assessed based on the regression test results.

### 3.3 Result and Discussion

#### 3.3.1 In sample evaluation and corrections

Table 3.3 provides some summary statistics. Although Unadjusted RB LTV has smaller spread between minimum and maximum values, it has a larger variance than the VaR LTV. The unadjusted Rule-Based LTV also has lower mean and median values than their VaR counterparts.
Table 3.3: Descriptive statistics of LTVs (In sample)

<table>
<thead>
<tr>
<th>Statistics</th>
<th>URB LTV</th>
<th>VAR LTV</th>
<th>VaR minus URB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>80.41%</td>
<td>85.65%</td>
<td>5.24%</td>
</tr>
<tr>
<td>Variance</td>
<td>0.46%</td>
<td>0.33%</td>
<td>0.08%</td>
</tr>
<tr>
<td>Min</td>
<td>52.82%</td>
<td>43.35%</td>
<td>-12.54%</td>
</tr>
<tr>
<td>Max</td>
<td>93.57%</td>
<td>98.78%</td>
<td>27.71%</td>
</tr>
<tr>
<td>Median</td>
<td>81.44%</td>
<td>86.90%</td>
<td>4.99%</td>
</tr>
</tbody>
</table>

In the scatter plots presented in figure 3.2, it can be seen that the spread of VAR LTV (from 0.4 to 1, 60%) is a bit wider than Unadjusted RB LTV (from 0.53 to 0.95, 52%). The lower end is flat for the latter because of some highly concentrated portfolios dominated by very volatile stocks. As shown in Table 3.2, there is only one lookup value for all volatilities higher than 18.81%. When all the portfolio constituents are highly volatile or when one or two stocks in the portfolio have the dominant weight, the portfolio LTV will only oscillate slightly around the last lookup value, 0.5281. These portfolios with the same LTV values are responsible for the flat lower bound of Unadjusted RB LTV.

The visible 45 degree trend validates the correlation between two LTV systems. The regression result of (3.19) is shown in Table 3.4. The non-negative alpha shows that the Gaussian VAR LTV can not be approximated by the unadjusted RV LTV alone.

Moreover, the minimum value of URB LTV appears to be higher than that of VAR LTV. Correlation between individual constituents is considered in VAR LTV, but left out in the calculation of URB LTV.

In this study, the concentration among assets is approximated by the geographic concen-
Figure 3.2: Scatter plot unadjusted VAR LTV v.s. RB LTV (In Sample)

Concentration measured by Herfindahl Index. Figure 3.3 is the histogram of Herfindahl index values. Although the weights of the portfolios are randomly selected, they are chose in some way that a full range of Herfindahl index, as calculated in (3.18), are produced.

Since our interest here is to use the measure for concentration to correct the unadjusted Rule-Based LTV, we test the relationship between the two LTVs and the Herfindahl index. In particular the relationship between LTV difference and Herfindahl index is checked and plotted in Figure 3.4. The correlation coefficient of LTV difference and the Herfindahl index is used as the relevance coefficient.

\[ R = \rho_{H,\sigma} = \frac{\text{cov}(H, \text{LTVDIFF})}{\sigma_H \sigma_{\text{LTVDIFF}}} \]  
(3.20)

\[ \text{LTVDIFF}_{p,t} = \text{LTV}^\text{VAR}_{p,t} - \text{LTV}^U_{p,t} \]  
(3.21)

Deciding the compensation function is the final step in the calibration of the URB LTV.
Figure 3.3: Histogram of Herfindahl Index

Figure 3.4: Scatter plot Herfindahl v.s. LTV difference (right), correlation = -0.3190
Given both the concentration level and its relevance, the last part of LTV adjustment is based on the exponential form of compensation as reflected in the curves in Figure 3.1.

\[ C_p = -0.05R(1 - \beta)exp(-2H_p^*) \] (3.22)

The first number 0.05 is approximately the average of the LTV difference for in sample period. The second term is the relevance for concentration measure. The third component \((1-\beta)\) scales down the compensation that is already considered in the unadjusted RB LTV. The last exponential component incorporates concentration (or more correctly diversification) into the compensation. The adjusted Rule-Based System LTV for portfolio \(p\) can be finally obtained as:

\[ LTV_{p,t}^{RB} = LTV_{p,t}^{U} + C_{p,t} \] (3.23)

The efficiency regression test for the adjusted result is also shown in Table 3.4. Based on the result, the adjusted RB LTV is closer to the VAR LTV. The intercept \(\alpha\) is smaller and the slope \(\beta\) is closer to one. The R-square value is also larger as well. Figure 3.2 and Figure 3.5 demonstrates the scatter plot of VAR LTV v.s. URB/RB LTV in the sample period, respectively. Both graphs exhibit improvements. In the scatter plot, the spread on the y-axis is lifted up slightly. In the kernel density plot, the shape of RB LTV becomes smoother and being moved to the right marginally.

<table>
<thead>
<tr>
<th>period</th>
<th>(\alpha)</th>
<th>(\beta)</th>
<th>R-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted</td>
<td>0.2299*(17.35)</td>
<td>0.7793*(76.32)</td>
<td>82.97%</td>
</tr>
<tr>
<td>Adjusted</td>
<td>0.2188* (20.96)</td>
<td>0.7863* (79.23)</td>
<td>84.19%</td>
</tr>
</tbody>
</table>
3.3.2 Out of Sample Forecast

The out-of-sample period starts from January 2004 and ends in December 2009. Figure 3.6 and Figure 3.7 the scatter plot of two LTVs in the period. The spread of the plot is smaller in out-of-sample period than in sample period. Table 3.5 illustrates the regression result and LTV difference analysis.

Table 3.5: Regression results for out-of-sample period

<table>
<thead>
<tr>
<th>period</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>R-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted</td>
<td>0.1600</td>
<td>0.8494</td>
<td>87.83%</td>
</tr>
<tr>
<td>Adjusted</td>
<td>0.1513</td>
<td>0.8526</td>
<td>88.66%</td>
</tr>
</tbody>
</table>

As stated in Table 3.6, the differences between VAR LTV and URB LTV become smaller, which has an average value of 3.18%. The mean LTV values of both systems are also higher.
Figure 3.6: 3 Scatter plot VAR LTV v.s. URB LTV (Out-of-sample)

Figure 3.7: 3 Scatter plot VAR LTV v.s. RB LTV (Out-of-sample)
Table 3.6: Descriptive statistics of LTVs (Out-of-Sample)

<table>
<thead>
<tr>
<th>Statistics</th>
<th>URB LTV</th>
<th>VAR LTV</th>
<th>VaR minus URB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>85.10%</td>
<td>88.28%</td>
<td>3.18%</td>
</tr>
<tr>
<td>Variance</td>
<td>0.36%</td>
<td>0.29%</td>
<td>0.04%</td>
</tr>
<tr>
<td>Min</td>
<td>53.33%</td>
<td>47.25%</td>
<td>-8.86%</td>
</tr>
<tr>
<td>Max</td>
<td>95.46%</td>
<td>98.43%</td>
<td>13.00%</td>
</tr>
<tr>
<td>Median</td>
<td>86.23%</td>
<td>89.89%</td>
<td>3.11%</td>
</tr>
</tbody>
</table>

than those of in sample period. One reason could be attributed to the returns during
asian financial crisis and internet bubbles in the first period are more extreme than those
during the subprime crisis in the second half. Table 3.7 supports this assumption with the
descriptive statistics of stock return volatilities. Its obvious that the 10-day returns are
less volatile in the out-of-sample period. Although the stock market experienced severe
downturn during the subprime crisis, the short-term price drops are not as drastic as
during the asian financial crisis and the internet bubbles. And hence the better result in
the out-of-sample period.

Table 3.7: Summary statistics of return volatility

<table>
<thead>
<tr>
<th>Stock Volatility</th>
<th>In Sample</th>
<th>Out-of-Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>9.62%</td>
<td>7.43%</td>
</tr>
<tr>
<td>Variance</td>
<td>0.22%</td>
<td>0.11%</td>
</tr>
<tr>
<td>Min</td>
<td>0.30%</td>
<td>0.80%</td>
</tr>
<tr>
<td>Max</td>
<td>35.97%</td>
<td>33.39%</td>
</tr>
<tr>
<td>Median</td>
<td>8.56%</td>
<td>6.89%</td>
</tr>
</tbody>
</table>

In the out-of-sample forecasting, the VAR LTVs are still calculated using up-to-date stock
return information. For RB LTV, the lookup table is calibrated in the in-sample period. The stock volatility decides which distribution it will use to estimate from the tail value. Yet the distribution and the tail shape, on which the lookup table is based, do not contain any new information out-of-sample period. If the lookup table updates with the new stock returns history, the 1% quantile value could be even higher since the distribution in the out-of-sample has thinner tails, and further narrowing the difference between two LTVs.

After adjustment by compensation stated in (3.22), R square from the regression in (3.19) improves. The parameters estimated are slightly better than those in Table 3.5 for URB LTV. Alpha is more close to 0 and beta increases marginally. The scatter plot of RB LTV v.s. VAR LTV is shown in Figure 3.7. The flat upper end of figure is improved.

It can be noted that the VAR LTV is on average higher than RB LTV. This lies in the fact that the VaR model assumes normal distribution, which is widely criticized as a distribution assumption for asset returns. In the next study, we will try another distribution assumption, which could be a better fit for the reality and derive an adjustment for diversification based on model.
Chapter 4

Student-t VaR LTV and Arbitrage

Pricing Theory

4.1 Definition and Methodology

4.1.1 Actual Portfolio Return

Motivated by the result in last chapter, it is advisable to change the distribution assumption in VaR calculation. Since the normal distribution underestimates the tail, it is possible that a distribution which includes the fat tail would produce better results. In the literature student-t distribution is used quite often to approximate the stock return distribution which is characterised by heavy tail. Thus it is explored that student-t based VAR LTV will provide larger VAR and a more conservative LTV. This could reduce the difference between LTV from the empirical constructed lookup table and model-based...
We would like to examine the real portfolio return data here before changing the assumption of the distribution. We will also try to fit both a normal distribution and a student-t distribution to the data and assess the eligibility. The real portfolio return data is plotted in the form of histogram. The distribution of empirical portfolio returns and fitted distributions for the two periods are presented in Figure 4.1 and 4.2. To justify the use of normal distribution used in VAR generation, a normal distribution is fitted to the data. In both in sample and out-of-sample periods, the normal distribution fails to describe the empirical data distribution properly. The empirical data distributions have higher kurtosis and fatter tails, which is not captured by Gaussian distribution. The VAR system concerns specifically severe loss scenarios; therefore it motivates the search for a better fitted distribution to generate VAR values.

Figure 4.1: Empirical data plot and fitted distribution for the in sample period
Here we assume that the standardised return $r$ follows a student-t distribution:

$$f(u) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{u^2}{\nu}\right)^{-1/2(\nu+1)}$$  \hspace{1cm} (4.1)

where $\nu$ is the degree of freedom and $\Gamma$ is the Gamma function. We will also use the t-location-scale distribution here with the mean $\mu$ and scale the variance $\sigma^2$.

Figure 4.3 and 4.4 also explain partially why VAR LTV is higher than RB LTV on average. When VAR LTV is decided by the tail of normal distribution, the Rule-Based LTV is based on empirical distribution, which has a fatter tail. Therefore, assuming that there are enough observations in each LTV lookup table group, the 1% quantile value should be lower in Rule-Based system theoretically.
Figure 4.3: Tail of empirical data plot and fitted distribution for the in-sample period

Figure 4.4: Tail of empirical data plot and fitted distribution for the out-of-sample period
To improve this deficiency, it necessitates a better distribution for Value at Risk calculation. The t-location-scale distribution provides a better fit for the data. At the negative return end, normal distribution assign lower probabilities to extreme low returns than the empirical and the t location-scale distribution. Table 4.1 summarize the results of fitting distributions to the data. It is shown that despite the high kurtosis, the distributions are only slightly skewed. Now the portfolio VaR defined in (3.8) becomes:

\[
r_p^{ST}(\alpha) = \sigma_p \Phi_{ST}^{-1}(\alpha, \nu) + \mu_p \tag{4.2}
\]

In contrast to the Gaussian case in the previous chapter, we have to make some modelling decisions about the parameter \(\nu\) in the student-t distribution. In this study we set \(\nu_i = \nu_j = \nu\), i.e. all stock return have the same number of degree of freedom. First of all, it is practically convenient. Instead of a different estimate for each stock, in the calibration for the client portfolio, we have a fixed constant. When we estimate the number of degree of freedom \(\nu\), for the out-of-sample period, we find the all the \(\nu\) estimates are close to 5. This result also applies to the estimate for the whole period.

More importantly, when all assets have the same number of degree of freedom, we can
make use of the linear correlation structure in equation (3.8) to calculate the portfolio
standard deviation, and use equation (3.8) as a basis to analyze concentration risk and
diversification effect in the following section. Under the linear dependence structure, the
portfolio volatility equals to the weighted average of constituent stock volatility if all
correlation equals 1. Otherwise, portfolio volatility will be smaller and portfolio Value
at Risk will be smaller in absolute value. Since we have a better knowledge about the
marginal distribution and a linear dependence structure will facilitate the analysis of
relationship between VAR LTVs and RB LTVs, we will propose to use copulas to model
the dependence structure. Explained by McNeil, Frey, Embrechts (2005)[17], copulas are
compatible with Value at Risk measure, because they are also expressed in quantiles.
Moreover, copulas allow us to use more sophisticated marginal model while keeping the
dependence structure simple. In our case, we can keep the desirable features of the linear
dependence structure and improve the Value at Risk result by using a t-location-scale
as the marginal distribution. Such Gaussian Copulas with arbitrary margins are called
meta-Gaussian distribution. Li [18] (2001) has proposed a meta-Gaussian model with
exponential marginals to predict default times of companies.

Since the Rule-Based LTV is constructed from the empirical return and volatility, the
lookup table constructed in the previous chapter is still valid. In this result, we strive to
correct URB LTV using in a more systematic way. The details will be explained in the
following section.
4.1.2 Diversification Effect

Let \( \Sigma^p(\alpha) \) be the portfolio VAR at \( \alpha \) confidence level and is derived from the weighted average of individual VaRs.

McNeil, Frey, Embrechts.(2005)[17] has discussed the subadditivity property of VaR, i.e., the sum of stand-alone VaR is smaller than the portfolio VaR. Recall the VaR expression under Gaussian: 

\[
r^G_p(\alpha) = \sigma_p \Phi^{-1}_G(\alpha) + \mu_p
\]

There are three prerequisites for this property to hold: 1) The constituents of the portfolio have a symmetric distribution, 2) The dependence structure among the assets should also be symmetric, 3) The asset return should not have heavy-tailed distribution.

Here, we want to demonstrate that, under the Gaussian framework, the sum of constituent \( LTV \) is smaller than the \( LTV \) based directly on portfolio return volatility. Equality holds only if \( \rho_{ij} = 1 \) for all \( i \) and \( j \).

**Theorem 1** If \( \rho_{ij} \neq 1 \) for all \( i \) and \( j \), then under multivariate normality of asset returns,

\[
r^G_P \geq r^{\Sigma G}_P \quad \text{for } \Phi^{-1}_G(\alpha) < 0
\]

\[
r^{\Sigma G}_P = \sum_{i=1}^{N} w_i r^G_i
\]

where \( \Phi^{-1}_G \) is the inverse of a cumulative distribution function of the normal distribution.

This result immediately implies that

\[ LTV^G_P \geq LTV^{\Sigma G}_P \]
Proof.

\[ r^G_P = \sum_{i=1}^{N} w_i r^G_i \]

\[ = \sum_{i=1}^{N} w_i (\sigma_i \Phi^{-1}_G(\alpha) + \mu_i) \]

\[ = \Phi^{-1}(\alpha) \sum_{i=1}^{N} w_i \sigma_i + \mu_P \]

(by eqn (3.6))

\[ = \Phi^{-1}(\alpha) \left( \sqrt{\left( \sum_{i=1}^{N} w_i \sigma_i \right)^2 + \mu_P} \right) \]

\[ \leq \Phi^{-1}(\alpha) \left( \sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i \neq j} w_i w_j \sigma_i \sigma_j + \mu_P \right) \]

for \( \rho_{ij} < 1 \) and \( \Phi^{-1}(\alpha) < 0 \)

\[ = \sigma^G_P \Phi^{-1}(\alpha) + \mu_P \]

(by eqn (3.7))

\[ = r^G_P \]

(by eqn (3.8))

\[ \square \]

From equation (3.7), the Gaussian volatility of portfolio can be written as

\[ \sigma^G_{P,t} = \sqrt{\sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i \neq j} w_i w_j \sigma_i \sigma_j \rho_{ij}} \]

\[ = \sqrt{\sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i \neq j} w_i w_j \sigma_i \sigma_j + \sum_{i \neq j} w_i w_j \sigma_i \sigma_j \rho_{ij} - \sum_{i \neq j} w_i w_j \sigma_i \sigma_j} \]

\[ = \sqrt{\left( \sum_{i=1}^{N} w_i \sigma_i \right)^2 \left( 1 + \frac{\sum_{i \neq j} w_i w_j \sigma_i \sigma_j \rho_{ij} - \sum_{i \neq j} w_i w_j \sigma_i \sigma_j}{\sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i \neq j} w_i w_j \sigma_i \sigma_j} \right)} \]

\[ = \sum_{i} w_i \sigma_i \sqrt{1 + \frac{\sum_{i \neq j} w_i w_j \sigma_i \sigma_j \rho_{ij} - \sum_{i \neq j} w_i w_j \sigma_i \sigma_j}{\sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i \neq j} w_i w_j \sigma_i \sigma_j}}. \]  

(4.5)
Then we can define the diversification factor following Cespedes et al:

**Definition 2 (Diversification Factor)**

\[
DF = \sqrt{1 + \frac{\sum_{i \neq j} w_i w_j \sigma_i \sigma_j \rho_{ij} - \sum_{i \neq j} w_i w_j \sigma_i \sigma_j}{\sum_i w_i^2 \sigma_i^2 + \sum_{i \neq j} w_i w_j \sigma_i \sigma_j}}
\]

(4.6)

Given that \(\rho_{ij}\) is typically less than 1, the DF is typically less than 1.

With the definition of diversification factor, the volatility of portfolio return in equation (4.5) can be written as

\[
\sigma_P^G = DF \sum_{i=1}^{N} w_i \sigma_i.
\]

(4.7)

If \(\mu_{i,t} = 0\) for all \(i\) and \(\mu_P = 0\), we have

\[
r_P^G = DF r_P^\Sigma^G
\]

And:

\[
LTV_p^G = \exp(r_P^G)
\]

\[
= \exp \left( DF \sum_{i=1}^{N} w_i \sigma_i \Phi^{-1}_G(\alpha) \right)
\]

\[
= \left[ \exp \left( \Phi^{-1}_G(\alpha) \sum_{i=1}^{N} w_i \sigma_i \right) \right]^{DF}
\]

Introducing the empirical LTV derived from the lookup table in equation (3.13) and the Gaussian implied volatility in (3.10), we get

\[
LTV_p^G = LTV_p^\Sigma, E \left[ \frac{\exp \left( \Phi^{-1}_G(\alpha) \sum_{i=1}^{N} w_i \sigma_i \right)}{\sum_{i=1}^{N} w_i \exp \left( \hat{\sigma}_i \Phi^{-1}_G(\alpha) \right)} \right]^{DF_p}
\]
If we are prepared to interchange the summation and the exponential term in the denominator, by that assuming \[ \sum w_i \exp(\hat{\sigma}_i \Phi^{-1}(\alpha)) \approx \exp(\sum_{i=1}^{N} w_i \hat{\sigma}_i \Phi^{-1}(\alpha)) \], then we have an approximated relationship between the VAR LTV and RB LTV:

\[
LTV^G_p = (LTV_{p\Sigma,E})^{DF_p} \tag{4.8}
\]

Now it only remains to simplify the calculation of DF without using the correlation matrix. To achieve that, a further simplification is required.

### 4.1.3 Simplification of DF under Arbitrage Pricing Theory

#### one dimensional case

Ross (1976)[19] first proposed the Arbitrage Pricing Theory model, which explains the stock return using several factors. Since the factors are quite intuitive, we can adopt it to easily extend the concentration into the 2-dimensional case. Moreover, the additive form of Arbitrage Pricing Theory framework adds to the ease of calculation. Under this framework, the return on a single stock \( i \) can be explained by summation of several factors:

\[
r_i = a_i + \sum_{G} b_{i,G} d_{i,G} I_G + e_i, \quad \sum_{G} d_{i,G} = 1 \tag{4.9}
\]

where \( a_i \) is the intercept, \( b_{i,G} \) the stock \( i \)'s sensitivity to factor \( G \), \( I_G \) the return of geographical index. \( d_{i,G} = 0, 1 \) is the indicator binary variable for stock \( i \), only when stock \( i \) belongs to geographical region \( G \), \( d_{i,G} = 1 \).
The variance of stock return \( i \) can then be expressed as:

\[
\sigma_i^2 = \sum_G d_{i,G}^2 b_{i,G}^2 \sigma_G^2 + \sigma_{e,i}^2 \tag{4.10}
\]

with \( \sigma_{e,i}^2 \) being the variance of the individual stock residual error term, \( \sigma_G^2 \) the variance of each index. It is assumed that factors and error term are uncorrelated: \( \text{cov}(G, c) = 0 \) and the factors are independent from each other: \( \text{cov}(G_1, G_2) = 0 \).

The covariance between stock \( i \) and stock \( j \) can also be expressed by the variance of Geographical index:

\[
\sigma_{i,j} = \sum_G b_{i,G} d_{i,G} b_{j,G} d_{j,G} \sigma_G^2 \tag{4.11}
\]

At the portfolio level, the portfolio return can also be attributed to several factor returns.

\[
r_p = a_p + \sum_G b_{p,G} I_L + e_p \tag{4.12}
\]

\[
\sigma_p^2 = \sum_L b_L^2 \sigma_L^2 + \sigma_{p,e}^2 \tag{4.13}
\]

with \( a_p = \sum_i w_i a_i \) being portfolio intercept, \( b_{p,G} = \sum_i w_i b_{i,G} d_{i,G} \) the portfolio sensitivity to factor G.

\( \sigma_{c,p}^2 \) is the variance of the portfolio residual error term. Under the independent assumption, it can be calculated as following:

\[
\sigma_{p,e}^2 = \sum_i w_i^2 \sigma_{c,i}^2 \tag{4.14}
\]
There is no cross term in equation (4.14) because of the uncorrelated error terms: \( \text{cov}(e_i, e_j)^2 = 0 \).

**Two dimensional case**

In addition to the geographical index \( G \), industry sector index \( S \) is also considered. The return and variance of a single stock is then given as follows:

\[
\begin{align*}
    r_i &= a_i + \sum_G b_{i,G}d_{i,G}I_G + \sum_S b_{i,S}d_{i,S}I_S + e_i \\
    \sigma^2_i &= \sum_G b_{i,G}^2d_{i,G}^2\sigma_G^2 + \sum_S b_{i,S}^2d_{i,S}^2\sigma_S^2 + \sigma_{i,e}^2
\end{align*}
\]

\( \sigma_{i,e}^2 \) is the variance of the individual stock error term, \( \sigma_G^2 \) the variance of each geographical index \( G \) and \( \sigma_S^2 \) the variance of each industrial index \( S \).

Similarly, it is assumed that factors and error term are uncorrelated: \( \text{cov}(G, e) = 0 \), \( \text{cov}(S, e) = 0 \). There is also no correlation between factors and across factors: \( \text{cov}(G_1, G_2) = 0 \), \( \text{cov}(S_1, S_2) = 0 \), \( \text{cov}(G, S) = 0 \).

The covariance between stock returns:

\[
\sigma_{i,j} = \sum_G b_{i,G}d_{i,G}b_{j,G}d_{j,G}\sigma_G^2 + \sum_S b_{i,S}d_{i,S}b_{j,S}d_{j,S}\sigma_S^2
\]

At the portfolio level:

\[
    r_p = a_p + \sum_G b_{p,G}I_G + \sum_S b_{p,S}I_S + e_p
\]
\[ \sigma_p^2 = \sum_G b_{p,G}^2 \sigma_G^2 + \sum_S b_{p,S}^2 \sigma_S^2 + \sigma_{p,e}^2 \]  \hspace{1cm} (4.19)

To further simplify the calculation, assume average sensitivities. This radical simplification suggests that all stocks belonging to the same factor will receive the same impact on the return. While error can be introduced, we believe that the benefit of reduced computing difficulty outweighs this drawback, since the ultimate goal is to find a practical model for daily bank operations. The covariance between portfolio constituent stocks can be approximated by:

\[ \tilde{\sigma}_{i,j} = \sum_G b_{d,G}^2 \sigma_G^2 + \sum_S b_{d,S}^2 \sigma_S^2 \]  \hspace{1cm} (4.20)

then approximated diversification factor becomes:

\[ \text{ADF} = \sqrt{1 + \frac{\sum_{i \neq j} w_i w_j (\tilde{\sigma}_{i,j} - w_i \sigma_i w_j \sigma_j)}{\sum_i w_i^2 \sigma_i^2 + \sum_{i \neq j} w_i w_j \sigma_i \sigma_j}} \]  \hspace{1cm} (4.21)

Therefore, equation (4.8) becomes:

\[ LTV_P^G = (LTV_P^{\Sigma,E})^{ADF_p} \]  \hspace{1cm} (4.22)

Now the concentration can be adjusted by a simpler form, which is based on a more solid mathematical ground. We test this method on the same set of data and assess how it has improved the results comparing to the first study.
4.1.4 Performance Evaluation

In this chapter, we introduce a new criterium to assess system efficiency. Apart from replicating the Value at Risk system, we also want to minimize the loss from the lombard loan from the resale of collateral portfolio in the event of default. Define the event when the real portfolio P&L is worse than the VaR value as a breach. Therefore it is advisable to compare the RB LTV with realized portfolio return and assess the breach ratio. Ideally, the breach should not exceed the pre-set risk appetite $\alpha$. We would also like to investigate the average LTV value, which represents the benefit the bank will provide to clients. The higher the average LTV, the more competitive the offer of the bank is. In the mean time, the bank should also look to minimize its risk, which is embedded in the breach rate.

4.2 Data and Implementation

In this chapter, we have used the same data set as in the last chapter. We have the 10-day return of 50 stocks from 01.02.1996 to 01.02.2010. They can be then further divided into geographical and industrial subgroups. The overview of industries presented in the dataset is given by Table 4.2. 12-year long data, excluding the first 2 years for EWMA calculation, is divided into 2 periods: in sample and out-of-sample. 1000 Portfolio is built based on the 50 stock returns.

To fit the stock return, we use the "fitdistr" package provided in R. The mean, variance and degree of freedom of a t-distribution are estimated using maximum likelihood estimation. Because when constructing the lookup table, all return data are mixed together so only
<table>
<thead>
<tr>
<th>Industry</th>
<th>No. of Stocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automotive</td>
<td>5</td>
</tr>
<tr>
<td>Basic Material</td>
<td>13</td>
</tr>
<tr>
<td>Conglomerate</td>
<td>5</td>
</tr>
<tr>
<td>Consumer good</td>
<td>3</td>
</tr>
<tr>
<td>Electronics</td>
<td>1</td>
</tr>
<tr>
<td>Financial</td>
<td>9</td>
</tr>
<tr>
<td>Health Care</td>
<td>3</td>
</tr>
<tr>
<td>Industrial</td>
<td>2</td>
</tr>
<tr>
<td>Materials and Construction</td>
<td>1</td>
</tr>
<tr>
<td>Retail</td>
<td>1</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>2</td>
</tr>
<tr>
<td>Transportation</td>
<td>2</td>
</tr>
<tr>
<td>Utilities</td>
<td>3</td>
</tr>
</tbody>
</table>
one degree of freedom needs to be estimated for all stock returns.

Because it is difficult to get the return on the appropriate industry and geographical indices, we construct the geographical and industrial return $I_G$ and $I_S$ by the stock returns belonging to the same geographical region and industry sector. These indices are simply the equally weighted average of the constituent stock returns. Based on these indices, the variance $\sigma_G^2$, $\sigma_S^2$, and sensitivities $b_G$, can be calculated.

Finally, the two LTVs will be compared to the actual realized portfolios returns. If the portfolio returns is even smaller than the LTVs, it is defined as a breach. The numbers of breaches will be investigated and controlled for the average level of LTV value.

4.3 Result and Discussion

In this study, the evaluation of new t-based VAR LTV and the adjusted RB LTV based on Arbitrage Pricing Theory is carried out again in efficiency regression tests and back testing.

Again the data are reported in two periods: in sample and out of sample. In both periods, we can see the regression result has improved using the new adjustment coefficient. Presented in Table 4.3, all the numbers reported in the brackets are t-statistics. They suggest that the fitting parameters are significant, rejecting the null hypothesis that the parameters equal to zero. In the in sample period, the ADC adjusted RB LTV is improved than the unadjust RB LTV: the intercept $\alpha$ is reduced from 0.1338 to 0.0120 and the slope $\beta$ increases from 0.8730 to 0.9883. Yet in the out-of-sample period, the RB LTV is over-
adjusted. The ADC adjustment has changed the intercept (from -0.0094 to -0.1647) and slope (from 1.0321 to 1.1816 ) further away from the ideal value 0 and 1 respectively. The adjusted R-Square also increase from 80.36% to 84.76% and from 85.54% to 88.29% respectively.

Table 4.3: Regression results of RB LTV with Approximated Diversification Coefficient (ADC) adjustment on student-t VAR LTV

<table>
<thead>
<tr>
<th>Rule-Based LTV</th>
<th>In sample</th>
<th>Out-of-Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADC</td>
<td>α</td>
<td>β</td>
</tr>
<tr>
<td>Unadjusted</td>
<td>0.1338* (15.56)</td>
<td>0.8730* (80.41)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>0.0120* (13.8)</td>
<td>0.9883* (93.76)</td>
</tr>
</tbody>
</table>

When looking at Figure 4.5, the trend line is closer to 45 degree, showing that the adjustment has brought RB LTV closer to the target: student-t based VAR LTV. And all the scatter plots for adjusted RB LTV have smoother upper and lower ends. Again the truncated distribution of Rule-Based LTV is due to the setup of lookup table. All portfolios with extremem small or large volatilities will receive the same lookup value and thus the shape of distribution. Another important evaluation is the back testing. There are two parameters to look at when gauging the performance of LTV: the breach percentage – the losses the bank will suffer and the average LTV the benefit that the bank may offer to its clients.

Ideally the LTV should have a breach rate similar to the target level. And the average level of LTV is also another important indicator of the LTV system performance. Although generally the breach rate increase with the average LTV level, the more efficient system can offer higher average LTV while reaching the target breach rate. The results shown in Table 4.4 indicate that the student-t based VAR LTV performs better than the unadjust
Figure 4.5: Scatterplot of the APT LTV v.s. Meta-Gaussian LTV
RB LTV. It has a low breach rate when compared to the target risk appetite rate and offers the higher LTV in average than the URB LTVs. On the other hand, the unadjusted RB LTV is too conservative, that is, its breach rate is far too low when compared to the target rates. Also, in lower risk appetite case (e.g. 0.75%), the URB LTV generates more breaches than the VAR LTV. This could be explained by the precision chosen to create the lookup table. The more data points are located in one lookup bucket, the more accurate the tail value will be. Therefore, with lower risk appetite, the URB LTV is no longer robust. The same behavior can be observed for the RB LTV; it performs well with significance level of 2% or 1% but produce less LTV than VAR LTV with lower significance level. As banks often work with a confidence level of 99.5% or 99.9%, this make it not an ideal candidate.

<table>
<thead>
<tr>
<th>RA (breach rate target)</th>
<th>Breach t-VAR</th>
<th>Breach URB</th>
<th>Breach RB</th>
<th>Average VAR LTV</th>
<th>Average URB LTV</th>
<th>Average RB</th>
</tr>
</thead>
<tbody>
<tr>
<td>2%</td>
<td>1.42%</td>
<td>0.91%</td>
<td>1.88%</td>
<td>0.8623</td>
<td>0.8391</td>
<td>0.8640</td>
</tr>
<tr>
<td>1%</td>
<td>0.40%</td>
<td>0.35%</td>
<td>0.89%</td>
<td>0.8283</td>
<td>0.8092</td>
<td>0.8381</td>
</tr>
<tr>
<td>0.75%</td>
<td>0.16%</td>
<td>0.27%</td>
<td>0.73%</td>
<td>0.8155</td>
<td>0.8027</td>
<td>0.8106</td>
</tr>
<tr>
<td>0.5%</td>
<td>0.03%</td>
<td>0.03%</td>
<td>0.20%</td>
<td>0.7982</td>
<td>0.7629</td>
<td>0.7977</td>
</tr>
</tbody>
</table>

To further investigate the sources of difference, we would like to go back to the lookup bin and examine the implied volatility defined in equation (3.10). Based on the model assumption, if we change the inverse CDF from Gaussian to student-t distribution with the number of degree of freedom of 5, we will have the student-t implied volatility \( \hat{\sigma}^t_i \).

Theoretically, the \( \alpha \)% quantile of return from the return bin as the average VaR should correspond to the mean value of volatility lookup bin. As shown in Figure 4.6, this is not the case with the student-t implied volatility. In most bins, the student-t implied volatilities do not align with the average value of volatility in that bin. This indicates that the student-t distribution has fatter tails than the empirical distribution, which can results
in over-estimation of the risk. Using the same reasoning, the Gaussian implied volatility shown in Figure 4.7 is always larger than the mean value of the volatility bin, owing to the fact that normal distribution does not assign enough probability at the tails.

Figure 4.6: student-t implied volatility in lookup bins (red lines are for lookup bucket and brown lines are the implied volatility)

The back testing result shown in Table 4.4 and Figure 4.8 also reveals that the RB LTV adjusted using ADC is indeed improved in both aspects. Its breach rates are quite close to the target value, as shown in Figure 4.9. However, in Figure 4.8, it is shown that at the same average LTV level, the RB LTV will generate more breaches than the VAR LTV based t-student distribution. It’s still not as efficient enough. Moreover, we have
Figure 4.7: Gaussian implied volatility in lookup bins (red lines are for lookup bucket and blue lines are the implied volatility)
made audacious assumption and simplification in equation 4.20, which can also introduce error in the calculation. Therefore, the adjustment based Arbitrage Pricing Theory is not a good candidate for LTV correction and we would like to incorporate other advanced models in the next chapter.

Figure 4.8: Back testing result of the APT LTV v.s. Meta-Gausian LTV
Figure 4.9: Actual breach rate v.s. Target Rate
Chapter 5

Multifactor framework

5.1 Definition and Methodology

In the previous chapters, we have defined two ways of determining LTVs: The VAR LTV and the RB LTV (i.e. Rule-Based LTV that is based on the lookup table):

\[ LTV^G_P = \exp (r^G_P) \]  
\[ r^G_P (\alpha) = \sigma^G_P \Phi^{-1}_G (\alpha) + \mu_P \]  
\[ \mu_P = \sum_{i=1}^{N} w_i \mu_{i,t} \]  
\[ \sigma^G_P = \sqrt{\sum_{i} w_i^2 \sigma_{i}^2 + \sum_{i \neq j} w_i w_j \sigma_{i} \sigma_{j} \rho_{ij}} \]

where \( r^G_P \) is the VaR of portfolio return based on Gaussian distribution of portfolio \( p \) at the confidence level \( \alpha \); (We drop \( \alpha \) for the sake of convenience); \( \Phi^{-1}_G \) the inverse Gaussian function; \( \sigma_p \) and \( \mu_p \) the variance and mean of the portfolio return, respectively. \( w_i \) is the
weight of each constituent stock \( i \) and \( \rho_{ij} \) denotes the linear correlation between asset \( i \) and \( j \).

\[
LTV_{P}^{\Sigma,E} = \sum_{i=1}^{N} w_{i} LTV_{i}^{E} = \sum_{i=1}^{N} w_{i} \exp \left( r_{i}^{E} \right) \tag{5.5}
\]

\( LTV_{i}^{E} \) is the lookup value of stock \( i \) obtained from the lookup table where \( r_{i}^{E} \) is the log lookup value. The superscript \( \Sigma \) indicates that this portfolio LTV is obtained by weighted averaging the individual stand-alone LTV.

To help the investigation of the relationship between equation (5.2) and (5.5) we introduce the third method, which is the sum of stand-alone VAR LTVs.

\[
LTV_{P}^{\Sigma,G} = \sum_{i=1}^{N} w_{i} LTV_{i}^{G} = \sum_{i=1}^{N} w_{i} \exp \left( r_{i}^{G} \right) \tag{5.6}
\]

In the chapter 4, we try to simplify the correlation between stocks by using Arbitrage Pricing Theory framework. The empirical result in section 4.3 suggests that APT is not an ideal solution. Therefore, in the following section, we want to break down the diversification factor into two parts that relate better to the risk factors.

### 5.1.1 Further Simplification of Herfindahl Index

Firstly, we simplify the Diversification Factor in equation (4.6):

**Theorem 3** (Simplication of Diversification Factor)

\[
DF = \sqrt{1 + (1 - HHI^{\sigma(N)}) \left( \bar{\rho} - 1 \right)}
\]
where $HHI^{\sigma(N)}$ is the $N$-asset modified Herfindahl-Hirschman diversification index with respect to volatilities,

$$HHI^{\sigma(N)} = \frac{\sum_{i=1}^{N} w_i^2 \sigma_i^2}{\left(\sum_{i=1}^{N} w_i \sigma_i\right)^2}$$

**Proof.** Suppose that there exists a constant $\bar{\rho}$ such that

$$\sum_{i \neq j} w_i w_j \sigma_i \sigma_j \rho_{ij} = \sum_{i \neq j} w_i w_j \sigma_i \sigma_j \bar{\rho}$$

$$\bar{\rho} = \sum_{i \neq j} \frac{w_i w_j \sigma_i \sigma_j}{\sum_{i \neq j} w_i w_j \sigma_i \sigma_j} \rho_{ij} = \sum_{i \neq j} \nu_{ij} \rho_{ij}. \tag{5.7}$$

where $0 \leq \nu_{ij} \leq 1$ for all $i \neq j$ and $\sum_{i \neq j} \nu_{ij} = 1$. Since $-1 \leq \rho_{ij} \leq 1$, $-1 \leq \bar{\rho} = \sum_{i \neq j} \nu_{ij} \rho_{ij} \leq 1$. Therefore, $\bar{\rho}$ is a weighted average of $\rho_{ij}$'s, which still satisfies the properties of being a linear correlation coefficient. Then substitute equation (5.7) into (4.6),

$$DF = \sqrt{1 + \frac{\sum_{i \neq j} w_i w_j \sigma_i \sigma_j (\bar{\rho} - 1)}{\sum_i w_i^2 \sigma_i^2 + \sum_{i \neq j} w_i w_j \sigma_i \sigma_j}}$$

$$= \sqrt{1 + \frac{\bar{\rho} - 1}{\sum_i w_i^2 \sigma_i^2 / (\sum_i w_i \sigma_i)^2}}.$$

Note that

$$\frac{\sum_i w_i^2 \sigma_i^2}{\sum_{i \neq j} w_i w_j \sigma_i \sigma_j} = \frac{\sum_i w_i^2 \sigma_i^2}{(\sum_i w_i \sigma_i)^2 - \sum_i w_i^2 \sigma_i^2}$$

$$= \frac{\sum_i w_i^2 \sigma_i^2 / (\sum_i w_i \sigma_i)^2}{1 - \sum_i w_i^2 \sigma_i^2 / (\sum_i w_i \sigma_i)^2}$$

$$= \frac{HHI^{\sigma(N)}}{1 - HHI^{\sigma(N)}}.$$

Note that Herfindahl-Hirschmann Index (HHI) is usually used for checking investment weights diversification. Here the HHI defined in Theorem 4.1 measures the diversification
over volatility of different magnitude. Hence,

\[ DF = \sqrt{\frac{1 + \frac{\bar{\rho} - 1}{HHI^{\sigma(N)}}}{1 - HHI^{\sigma(N)}} + 1} = \sqrt{1 + (\bar{\rho} - 1)(1 - HHI^{\sigma(N)})}. \]

Now the DF is decomposed into two components: the average correlation coefficient, which takes into account the correlation between assets; the N-asset HHI, which penalize the LTV for concentration on securities with high volatility. This is an improvement from the methodology in chapter 4. Because when using the ADC adjustment, we only compensate the URB LTV but the penalty is missing. The DF defined in Theorem 3 also allows us the flexibility to construct Herfindahl Index based on different types of concentration. This will be explored in the next section, 5.1.2.

### 5.1.2 Multiple Factor Model

Proposed by Bluhm, Overbeck. (2007) [20], the multi-factor model provides a better way to explain correlation.

Now we assume the standardized return of asset \( i \) is driven by a multifactor model as shown below

\[ \tilde{r}_{i,t} = \beta_i \sum_{k=1}^{K} a_{i,k} \psi_{k,t} + \epsilon_{i,t}, \quad a_{i,k} = \{0, 1\} \quad (5.8) \]

where \( \psi_{k,t} \sim N(0, 1) \) is the kth factor return, due to e.g. regional or industry effect, \( a_{i,k} \) is a \((0, 1)\) indicator to turn on or off the kth factor return, \( \beta_i \) is the factor loading for return \( i \),

64
which can be viewed as the equivalent of the sensitivity of CAPM and $\epsilon_{i,t} \sim N \left(0, (1 - \beta)^2\right)$ is an idiosyncratic factor which is independent from $\psi_{k,t}$.

As an illustration, we assume there are only two factors capturing the regional and industry effect. Moreover, we assume that the regional factor and the industrial factor consists of $N_R$ regions and $N_I$ industries, respectively such that the number of factors is $K = N_R + N_I$. This model, of course can be later expanded to include more risk factors like company size, etc.

Then the systematic composite factor can be written as following, dropping the time subscript temporarily for convenience,

$$\sum_{k=1}^{K} a_{i,k} \psi_k = \lambda_i \left( \psi_{R(i)} + \psi_{I(i)} \right), \quad (5.9)$$

where $\psi_{R(i)}$ and $\psi_{I(i)}$ denote, respectively, the regional factor and industry factor which asset $i$ belongs to, and $\lambda_i$ is a normalizing factor which guarantees the unit variance of systematic composite factor, i.e.

$$\lambda_i = \left( \sqrt{\text{var} (\psi_{R(i)} + \psi_{I(i)})} \right)^{-1}$$

$$= \left( \sqrt{\text{var} (\psi_{R(i)}) + \text{var} (\psi_{I(i)}) + 2\text{cov} (\psi_{R(i)}, \psi_{I(i)})} \right)^{-1}$$

$$= \left( \sqrt{2 + 2\rho_{R(i)I(i)}} \right)^{-1}$$

where $\rho_{xy}$ denote the correlation between the returns of region $x$, $\psi_{R(x)}$, and industry $y$, $\psi_{I(y)}$.

Now we define sector $s = (x, y)$ represents all possible pair of region and industry for $x = 1, 2, \cdots, N_R$ and $y = 1, 2, \cdots, N_I$ so that the number of unique sectors is $S = N_R \times N_I$. 65
Theorem 4 The cross-correlation among asset returns at time \( t \) can be written as

\[
\tilde{\rho}_{ij,t} = \beta_i \beta_j \lambda_i \lambda_j \left[ \rho_{R(i)R(i),t} + \rho_{I(i)R(j),t} + \rho_{R(i)I(j),t} + \rho_{I(i)I(j),t} \right], \tag{5.10}
\]

**Proof.** This result follows immediately from the return equations for asset \( i \) and \( j \) below as defined in equations (5.8) and (5.9),

\[
\tilde{r}_{i,t} = \beta_i \lambda_i \left( \psi_{R(i),t} + \psi_{I(i),t} \right) + \epsilon_{i,t} \\
\tilde{r}_{j,t} = \beta_j \lambda_j \left( \psi_{R(j),t} + \psi_{I(j),t} \right) + \epsilon_{j,t}.
\]

Under multiple factor model in equation (5.8), the average correlation coefficient defined in Theorem 4.1, \( \bar{\rho} \) in equation (5.7) can be substituted as

\[
\bar{\rho} = \sum_{i \neq j} w_i w_j \sigma_i \sigma_j \tilde{\rho}_{ij} = \sum_{i \neq j} v_{ij} \tilde{\rho}_{ij},
\]

where \( \tilde{\rho}_{ij} \), the average correlation coefficient based on the multi-factor model is given in equation (5.10). Benefit from Theorem 4, we can map the correlation between assets to certain factors. The calculation of the correlation matrix for \( N \) assets for each portfolio will be reduced to the correlation matrix of \( S \) risk factors. This correlation matrix can be re-used for all portfolios. In most occasions, an annual or semi-annual update can be sufficient.

Next, we will examine several possibilities to construct Herfindahl Index and investigate the relationship between them.
Theorem 5 (Upper bound for $DF$) Under multiple factor model in equation (5.8),

$$HHI^\alpha(N) \leq HHI^\alpha(S)$$  \hfill (5.11)

where $HHI^\alpha(N)$ is the $N$-asset HHI and $HHI^\alpha(S)$ stands for the sector HHI, and if

$$\tilde{DF} = \sqrt{1 + (\tilde{\rho} - 1) (1 - HHI^\alpha(N))}$$

$$\tilde{DF}^{UB(S)} = \sqrt{1 + (\tilde{\rho} - 1) (1 - HHI^\alpha(S))}$$

then

$$\tilde{DF} \leq \tilde{DF}^{UB(S)}$$

Proof. Given there are $N$ assets in portfolio $P$, let $N_s$ denote the number of assets in sector $s$ and $w_{s,i}$ denote the weight of asset $i$ in sector $s$. Then the weight of sector $s$ can be calculated as the sum of weights on all assets in sectors.

$$w_s = \sum_{i=1}^{N_s} w_{s,i}$$

The $N$-assets HHI is defined in Theorem 3. The $S$-sector HHI based on $N_R$ regions and $N_I$ industries can be defined as

$$HHI^{\bar{\sigma}(S)} = \frac{\sum_{s=1}^{S} w_s^2 \bar{\sigma}_s^2}{\left( \sum_{s=1}^{S} w_s \bar{\sigma}_s \right)^2},$$  \hfill (5.12)

where

$$\bar{\sigma}_s = \frac{\sum_{i=1}^{N_s} w_{s,i}}{w_s} \sigma_{s,i},$$

which is the weighted average of $N_s$ asset’s volatilities in sector $s$ as the volatility of sector $s$. In the above setup, we can easily show that

$$\sum_{s=1}^{S} \sum_{i=1}^{N_s} w_{s,i}^2 \sigma_{s,i}^2 = \sum_{j=1}^{N} w_j^2 \sigma_j^2.$$  \hfill (5.13)
\( \sigma_{s,i} \) denotes the volatility of asset \( i \) in sector \( s \), \( s = (x, y) \) indicates a unique region and industry pair for \( x = 1, 2, \cdots, N_R \) and \( y = 1, 2, \cdots, N_I \) and \( S \) is the number of the unique sectors, i.e. \( S = N_R \times N_I \). \( N \) is the asset number in the portfolio

First, we prove that the denominators in the expressions of \( HHI^{\sigma(N)} \) and \( HHI^{\bar{\sigma}(S)} \) are equivalent to each other. From equation (5.12),

\[
\sum_{s=1}^{S} w_s \bar{\sigma}_s = \sum_{s=1}^{S} w_s \left( \sum_{i=1}^{N_s} \frac{w_{s,i}}{w_s} \sigma_{s,i} \right) = \sum_{s=1}^{S} \sum_{i=1}^{N_s} w_{s,i} \sigma_{s,i} = \sum_{j=1}^{N} w_j \sigma_j.
\]

The last equality holds from equation (5.13).

Second, from the numerators in \( HHI^{\sigma(N)} \) and \( HHI^{\bar{\sigma}(S)} \),

\[
\sum_{s=1}^{S} w_s^2 \bar{\sigma}_s^2 = \sum_{s=1}^{S} w_s^2 \left( \sum_{i=1}^{N_s} \frac{w_{s,i}}{w_s^2} \sigma_{s,i}^2 \right) \geq \sum_{s=1}^{S} \sum_{i=1}^{N_s} w_{s,i}^2 \sigma_{s,i}^2 = \sum_{s=1}^{S} \sum_{i=1}^{N_s} w_{s,i}^2 \sigma_{s,i}^2 = \sum_{j=1}^{N} w_j^2 \sigma_j^2.
\]

Therefore, from equation (5.14) and (5.15),

\[
HHI^{\sigma(N)} \leq HHI^{\bar{\sigma}(S)}
\]

and

\[
\tilde{DF}_P \leq \tilde{DF}^{UB(S)}_P.
\]
It is shown that the fully-pledged N-asset $HHI$ is smaller than or equal to the sector based $HHI$. Hence, if we adopt the simplified version of Herfindahl Index, we know in what kind of influence will the error introduce due to the simplification. Since the sector based $HHI$ is the upper bound for N-asset $HHI$, the sector based DF will also never be smaller than the N-asset $HHI$. Thus the DF adjustment based on sectors will not be as conservative as those based on N-asset $HHI$. The discrepancy between these two $HHIs$ decreases as the sector number $S$ increases.

![Figure 5.1: Convergence of sector $HHI$ (red line) to N-asset $HHI$ (blue line). The x-axis is the number of sectors, $S$ and the y-axis the difference between these two $HHIs$.](image)

Their relationship is shown in Figure 5.1, the discrepancy quickly decreases starting from 10 sectors. It remains at a low level between 10 and 40 sectors, which is quite common in most practical cases. The difference drop to a trivial level after the sector number increase over 40. This result indicates that it might be sufficient to include 2 risk factors with five sectors each. We will check for the validity of this assumption in the result and discussion section 5.3.
5.2 Data and Implementation

In this study, we changed to a bigger data universe. A total of 12 years of stock 10-day returns are used, dating from 1998 to 2010. Subject to data availability a total of 489 stock data are used. We have grouped them to 10 regions and 10 industries accordingly. Following the study of Duellmann and Masschelein(2006)[21], we use the Global Industry Classification Scheme to group the industry data. A short summary will be found in Table 5.1. The division of data is similar to the previous chapters. The first two years (1998-1999) will be used to calculate EWMA volatility and the following ten years will be divided into two sample periods: in sample (2000 - 2004) and out-of-sample (2005-2010).

In this study, we retrieved the MSCI country index and industry index for the same period. There are return data for 34 industry branches based on the Global Industry Classification Scheme and 55 countries in total. Based on these data, we have constructed again 10 industry indices and 10 region indices using weighted averaging. To check for data validity, we include a cross correlation table in Appendix 1. For example, West Europe is highly correlated with USA ($\rho = 0.48$) and Eastern Europe ($\rho =0.71$), which is consistent with the common economic fact. Now we have four ways of calculating concentration: the geographical concentration, which is based on regions; the industrial concentration based on industry branches; a two-dimensional concentration considering both region and industry and finally the sector concentration, which consider every single combination of geographical and industrial factors as a sector. A sector can also be interpreted as one single box in Table 5.1. There are 100 sectors possible, and in this study, there are all together 89 sectors available. In the following section, we will examine the performance
of four concentration measures.

Table 5.1: summary of top 500 equity data

<table>
<thead>
<tr>
<th>Stock data</th>
<th>East Europe</th>
<th>Far East</th>
<th>Far East D</th>
<th>Middle East</th>
<th>North America</th>
<th>Oceanic</th>
<th>South Africa</th>
<th>South America</th>
<th>West Europe</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer Discretionary</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>14</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>20</td>
<td>20</td>
<td>46</td>
</tr>
<tr>
<td>Consumer Staples</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>11</td>
<td>1</td>
<td>4</td>
<td>20</td>
<td>12</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Energy</td>
<td>8</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>22</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Financials</td>
<td>4</td>
<td>22</td>
<td>15</td>
<td>13</td>
<td>3</td>
<td>9</td>
<td>38</td>
<td>104</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Care</td>
<td>1</td>
<td>1</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industrial</td>
<td>1</td>
<td>9</td>
<td>6</td>
<td>1</td>
<td>11</td>
<td></td>
<td>1</td>
<td>24</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>IT</td>
<td>11</td>
<td>4</td>
<td>1</td>
<td>20</td>
<td></td>
<td></td>
<td>9</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Materials</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>16</td>
<td>4</td>
<td>1</td>
<td>8</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Telecommunication</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>9</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Utilities</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>9</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Grand Total</td>
<td>24</td>
<td>74</td>
<td>33</td>
<td>4</td>
<td>132</td>
<td>11</td>
<td>4</td>
<td>35</td>
<td>172</td>
<td>489</td>
</tr>
</tbody>
</table>

To obtain the \( \beta_i \) for each stock, the excess return of stock \( i \) is regressed on the excess return of a market portfolio, which is approximated by the world factor from MSCI GEM2 Indicies. The risk free rate is assumed to be zero.

Also, to rule out the errors introduced by distribution assumptions in VAR LTV, we calculate the empirical portfolio volatility using the portfolio return \( r_p \) in Exponential smoothing:

\[
\sigma_{p,t}^E = \frac{1 - \lambda}{1 - \lambda^{m+1}} \sum_{a=0}^{m} \lambda^a r_{p,t-a}^2
\]  

Using \( \sigma_{p,t}^E \) we can now calculate the student-t based VAR LTV and also apply the lookup table scheme here to have a portfolio lookup LTV. We will examine the performance of RB LTV by using these two target values.
5.3 Result and Discussion

We will first examine the regression result of using the fully-pledged 2-dimensional concentration measure, then we will look at the all 4 version in back testing. The regression result has shown that adjustments based multifactor-ADC and the new data has also improved the approximation of RB LTV. The regression results are different from the 500 stock universe than the 50 stock universe. As shown by Table 5.2 the results in out-of-sample period does not suffer from over-correction as in the previous chapter. The improvement in in-sample period is also greater than using the ADC correction. Apart from apparent improvement of the intercept and the slope in both periods, R squares of the regressions have also raised nearly 9%. The regression results have confirmed that the modified Herfindahl index has enhanced the approximation for LTV.

Figure 5.2: Scatterplot of the Empirical-Gaussian LTV v.s. URB LTV in sample
When looking at the graphics, we can see that in both Figure 5.4 and Figure 5.5, the approximation is better when both empirical LTV and RB LTV are high. The diagonal line pattern is apparent in the region between 0.7 and 1. When empirical portfolio LTVs decrease lower than 0.6, there is no corresponding penalty for the RB LTV. This part of difference can’t be explained by the linear correlation. We will see soon in the back testing part, how severe this error will be.

Table 5.2: Regression results of RB LTV with modified Herfindahl Index with 2-dimensional concentration

<table>
<thead>
<tr>
<th>Rule-Based LTV</th>
<th>In sample</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Out-of-Sample</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ADC</td>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>R-Square</td>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>R-Square</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted</td>
<td>0.1232* (13.28)</td>
<td>0.9008* (78.33)</td>
<td>76.14%</td>
<td>0.659* (-9.41)</td>
<td>0.9518* (10.04)</td>
<td>85.03%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted</td>
<td>0.0070* (11.4)</td>
<td>0.9830* (94.82)</td>
<td>86.58%</td>
<td>-0.0650* (-12.96)</td>
<td>1.0559* (93.52)</td>
<td>86.86%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When looking at Table 5.3 and 5.4, we can see that the t-based VAR LTV is still too
Figure 5.4: Scatterplot of the Empirical-Gaussian LTV v.s. URB LTV out-of-sample

Figure 5.5: Scatterplot of the Empirical-Gaussian LTV v.s. RB LTV out-of-sample
conservative, even when we change to a new data set. It offers the lowest breach rate and average LTV in both in sample and out-of-sample period. The Unadjusted Rule-Based LTV is conservative in the in sample period. This is consistent with the assumption of full concentration and zero diversification. Yet in the out-of-sample period, the URB LTV has produced breach rate that is very similar to the target rate, before being adjusted. This could be attributed to the fact that the distribution of stock returns in the out-of-sample period has fatter tail than that in the in sample period. A suggestion of improvement is to update the lookup table more often during financial market turmoil. Finally we can see that the fully-pledged RB LTV based on 2-dimensional Herfindahl index also suffers from over-correction. Although it offers the highest average LTV value, its breach rate is 3-4 times larger than the target rate. One possible explanation can be the non-linear correlation between assets. The construction of region and industry indices by equally weighted averaging can also introduce errors.

Table 5.3: Backtesting results with different Risk Appetite (RA) in sample

<table>
<thead>
<tr>
<th>RA (breach rate target)</th>
<th>Breach t-VAR</th>
<th>Breach URB</th>
<th>Breach RB</th>
<th>Average t-VAR</th>
<th>Average URB LTV</th>
<th>Average RB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>0.10%</td>
<td>0.43%</td>
<td>2.06%</td>
<td>0.8117</td>
<td>0.8302</td>
<td>0.879</td>
</tr>
<tr>
<td>0.75%</td>
<td>0.04%</td>
<td>0.28%</td>
<td>1.66%</td>
<td>0.7986</td>
<td>0.8215</td>
<td>0.8725</td>
</tr>
<tr>
<td>0.5%</td>
<td>0.01%</td>
<td>0.16%</td>
<td>1.08%</td>
<td>0.7796</td>
<td>0.8059</td>
<td>0.8597</td>
</tr>
<tr>
<td>0.25%</td>
<td>0.00%</td>
<td>0.10%</td>
<td>0.72%</td>
<td>0.7457</td>
<td>0.7858</td>
<td>0.846</td>
</tr>
<tr>
<td>0.1%</td>
<td>0.00%</td>
<td>0.05%</td>
<td>0.46%</td>
<td>0.6977</td>
<td>0.7655</td>
<td>0.8305</td>
</tr>
</tbody>
</table>

Table 5.4: Backtesting results with different Risk Appetite (RA) out-of-sample

<table>
<thead>
<tr>
<th>RA (breach rate target)</th>
<th>Breach t-VAR</th>
<th>Breach URB</th>
<th>Breach RB</th>
<th>Average t-VAR</th>
<th>Average URB LTV</th>
<th>Average RB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>0.16%</td>
<td>1.01%</td>
<td>3.98%</td>
<td>0.8303</td>
<td>0.8593</td>
<td>0.8986</td>
</tr>
<tr>
<td>0.75%</td>
<td>0.03%</td>
<td>0.78%</td>
<td>3.40%</td>
<td>0.8185</td>
<td>0.8524</td>
<td>0.8935</td>
</tr>
<tr>
<td>0.5%</td>
<td>0.00%</td>
<td>0.45%</td>
<td>2.81%</td>
<td>0.8013</td>
<td>0.8405</td>
<td>0.8859</td>
</tr>
<tr>
<td>0.25%</td>
<td>0.00%</td>
<td>0.23%</td>
<td>1.82%</td>
<td>0.7705</td>
<td>0.8231</td>
<td>0.8715</td>
</tr>
<tr>
<td>0.1%</td>
<td>0.00%</td>
<td>0.10%</td>
<td>1.31%</td>
<td>0.7268</td>
<td>0.8069</td>
<td>0.8595</td>
</tr>
</tbody>
</table>
Because of the flexibility of multi-factor model, we can orchestrate another three kinds of adjustments: one dimensional region or industry based adjustment as well as N-asset Herfindahl Index. The plots of breach rate v.s. average LTV value are presented in Figure 5.6 and 5.7. In both plots, the t-based VAR LTV lies at the bottom of the plot, which presents low breach rate and low average LTV value. The URB LTV lines are located between t-VAR LTV and the cluster of RB LTVs in both plots. The full pledged RB LTV and N-asset RB LTV lie very close to each other in the right upper area, which represents high mean LTV and high breach rate. The RB LTVs based on 1-dimensional Region or Industry Herfindahl index have a breach rate similar to the target in the in sample period while providing relative high mean LTV values. However, their performance worsen in the out-of-sample period in terms of breach rates. When only examining the deviation from the target breach rate, we can see that the one dimensional region based Herfindahl index is closest to the benchmark. When comparing it to the result of last chapter, it is improved because of the new correction function. Another one dimensional Herfindahl Index based on industry lies slightly above. The two fully-pledged ADC which are based on sectors and assets have over compensated the URB LTV, leading to unsatisfactory performance. The region-based 1-dimensional concentration should be adopted to adjust URB LTV. Note that it still need to be calibrated to produce the desired breach rate since it is now higher than the target. This is also a common practise in banking operation.
Figure 5.6: Breach rate v.s. Average LTV in sample (RB LTV: 2-dimensional Herfindal Index, Sector: N-asset Herfindahl Index)

Figure 5.7: Breach rate v.s. Average LTV out-of-sample (RB LTV: 2-dimensional Herfindal Index, Sector: N-asset Herfindahl Index)
Figure 5.8: Actual breach rate v.s. Target Rate in sample (RB LTV: 2-dimensional Herfindal Index, Sector: N-asset Herfindahl Index)

Figure 5.9: Actual breach rate v.s. Target Rate out-of-sample (RB LTV: 2-dimensional Herfindal Index, Sector: N-asset Herfindahl Index)
5.4 Final comparison

To fully investigate the implication of using different correction functions, we compare the back testing results of correction function based on Arbitrage Pricing Theory and those based multi-factor model. We didn’t include the ad-hoc correction function in the first study since no back testing was carried out. Although the APT correction and multifactor correction use different data sets, we still feel the comparison is valid: we are not evaluating the absolute value of only one set of risk measure but a combination of two risk measures. Therefore with the back testing results carried out earlier we can assess the performance of these two different correction functions. For the same reason we are excluding the regression result here because the regression result is also data dependant.

We collected the results from APT correction and the multi-factor correction. The following Tables 5.5 and 5.6 give a final view of the back testing result. A desirable feature of the correction function is the one which produce high average LTV and maintain low breach rate in the mean time. According to this criterium the Fully-Blown correction function and Sector-based correction function, which have the highest average LTV values and breach rate, appear not to be ideal candidates. For those one-dimensional correction, as already discussed in the section 5.3, the correction function based on Region-Herfindahl index is preferred over that based on Industry-Herfindal index. In Figure 5.10 and 5.11\(^1\) we can see that the Region-Herfindahl Index correction performs also better than the APT based correction. The former produce a lower breach rate than the latter and has

\(^{1}\)A direct comparison of the Multi-factor Region correction and the APT correction could be difficult since in the average LTV v.s. Breach Rate plot these two curves lie far away from one another. We extended the short end of multi-factor modles(we add the risk appetite of 0.25% and 0.1%)
a higher average LTV level. This implies that after calibration, the correction based on Region-Herfindahl is most preferred out of all correction functions investigated in this paper.

Table 5.5: Cross comparison between different correction functions(1)

<table>
<thead>
<tr>
<th>RA (breach rate target)</th>
<th>Average Fully-Blown</th>
<th>Average Sector</th>
<th>Average Region</th>
<th>Average Industry</th>
<th>Average APT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00%</td>
<td>0.8888</td>
<td>0.8849</td>
<td>0.8769</td>
<td>0.8766</td>
<td>0.8367</td>
</tr>
<tr>
<td>0.75%</td>
<td>0.8830</td>
<td>0.8790</td>
<td>0.8681</td>
<td>0.8702</td>
<td>0.8269</td>
</tr>
<tr>
<td>0.5%</td>
<td>0.8728</td>
<td>0.8684</td>
<td>0.8593</td>
<td>0.8589</td>
<td>0.8111</td>
</tr>
</tbody>
</table>

Table 5.6: Cross comparison between different correction functions(2)

<table>
<thead>
<tr>
<th>RA (breach rate target)</th>
<th>Breach Fully-Blown</th>
<th>Breach Sector</th>
<th>Breach Region</th>
<th>Breach Industry</th>
<th>Breach APT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00%</td>
<td>3.02%</td>
<td>2.68%</td>
<td>1.98%</td>
<td>2.12%</td>
<td>0.87%</td>
</tr>
<tr>
<td>0.75%</td>
<td>2.53%</td>
<td>2.25%</td>
<td>1.66%</td>
<td>1.83%</td>
<td>0.73%</td>
</tr>
<tr>
<td>0.5%</td>
<td>1.94%</td>
<td>1.68%</td>
<td>1.21%</td>
<td>1.36%</td>
<td>0.20%</td>
</tr>
</tbody>
</table>
Figure 5.10: Cross comparison between different correction functions, Target Rate v.s Actual breach rate

Figure 5.11: Cross comparison between different correction functions, Average LTV v.s. Actual breach rate
Chapter 6

Bond LTV

6.1 Definition and Methodology

The risk management of a Bond portfolio is different from managing an equity portfolio. The factors influencing bond return and their correlation differ a lot from equity stock. In this chapter, we will investigate a simple system of deciding bond LTV without calculating bond VaR. Our task is to calibrate bond LTV and back test it on the real portfolio returns. As a first trail study, we will construct a unadjusted RB LTV based on a lookup table similar to those of equity LTV. We will not correct the correlation in bond portfolios and run regression tests. Instead, we will directly compare the URB LTVs to actual portfolio return and evaluate the performance of the LTV system by backtesting.

From the literature review we have established that for sovereign bonds, the interest rate yield curve plays an important role in affecting the bond return. For corporate bonds, interest rate risk and credit risk jointly determine the bond return. Consider a zero coupon
bond, which pays 1 unit at maturity date \( T \), its price at time \( t \) based on continuous discounting is given by:

\[
P_t = e^{y_t(T - t)} \tag{6.1}
\]

where \( y_t \) is the yield to maturity. It consists of two component, the interest rate risk part \( y_i^t \) and the credit risk part \( y_c^t \) :

\[
y_t = y_i^t + y_c^t \tag{6.2}
\]

Define the bond return following (3.1):

\[
r_t = \ln \frac{P_t}{P_{t-1}} = \ln[e^{y_t(T - t)} + y_{t-1}(T - t + 1)]
\]

\[
= -y_t(T - t) + y_{t-1}(T - t) + \frac{10}{365} y_{t-1} \quad (\frac{10}{365} \text{ is the 10-day time step})
\]

\[
\approx (y_{t-1} - y_t)(T - t) \quad (\text{ignore the small value } \frac{10}{365})
\]

\[
= (y_{i-1}^t - y_i^t + y_{c-1}^t - y_c^t)(T - t)
\]

\[
= (y_{i-1}^t - y_i^t)(T - t) + (y_{c-1}^t - y_c^t)(T - t)
\]

\[
= \Delta y_i^t(T - t) + \Delta y_c^t(T - t) \tag{6.3}
\]

Similar to (3.3), bond LTV can be derived based on Value at Risk scenario:
\[ LTV_{i,t}(\alpha) = \exp(r_{i,t}^\alpha) \]
\[ = \exp[\Delta y_{i,t}^i(\alpha)(T-t) + \Delta y_{i,t}^c(\alpha)(T-t)] \]
\[ = \exp[\Delta y_{i,t}^i(\alpha)(T-t)]\exp[\Delta y_{i,t}^c(\alpha)(T-t)] \]
\[ = LTV_{i,t}^i(\alpha)LTV_{i,t}^c(\alpha) \] (6.4)

Therefore, bond LTV can be also divided into two components. This finding motivates a two dimensional LTV lookup table. Similar to the procedures described in section 2.1.4, we identify implied interest rate volatility \(\sigma^{in}\) and CDS and build up lookup tables respectively.

Described in Brigo and Mercurio(2001)[22], interest rate Cap can be considered as a series of European call option on certain underlying interest rate with a agreed strike price, or Caplets. A simple Black model can be used to evaluate the pricing of this interest rate derivative. This pricing model is similar to the Black-Schole model for a equity option. Therefore, we can also obtain an implied volatility for the underlying rate. In our case, it is the implied volatility of interest rate.

For the interest rate component, we divide implied volatility calibrated from Caplet into buckets, \((0, \cdots, \sigma_{l-1}^{in}, \sigma_l^{in}, \cdots)\), then
\[ r_{i}^{in,E}(\alpha) = \Phi^{-1}_l(\alpha) \quad \text{with} \quad \sigma_i^{in} \in (\sigma_{l-1}^{in}, \sigma_l^{in}] \] (6.5)

where \(r_{i}^{in,E}\) is the empirical \(\alpha\) lower quantile \(\Phi^{-1}_l\) of all historical bond returns that fall into volatility bucket \(l\), when only interest rate risk is considered. It is possible to apply
curve fitting techniques to compute the implied interest rate volatility $\sigma^{in}$ to each bond with specific time to maturity $T-t$.

For the credit risk part, the risk driver is CDS $\delta_{i,t}$ is given by:

$$\delta_i = y_i - s_i \quad (6.6)$$

$s_i$ is the swap rate of the same currency, which considered to be risk free. The CDS will again be divided into buckets: $(0, \cdots, \delta_{l-1}, \delta_{l}, \cdots]$. The credit risk lookup value is given by:

$$r^{c,E}_i(\alpha) = \Phi^{-1}_l(\alpha) \quad \text{with} \quad \delta_{i} \in (\delta_{l-1}, \delta_{l}] . \quad (6.7)$$

Combining the two LTV component to yield the bond LTV:

$$LTV_i = LTV^{i,L}_i LTV^{c,L}_i = \exp(r^{i,E}_i)\exp(r^{c,E}_i) \quad (6.8)$$

Finally the portfolio unadjusted $LTV^U$, as defined in (3.13):

$$LTV^U_p = \sum_{i=1}^{N} w_i LTV_i$$

### 6.2 Data and Implementation

To calibrate bond LTV, we used a dataset consisting 500 bonds provided by Credit Suisse. The majority of these bonds are sovereign bonds. In addition, there are also corporate bonds from diverse sectors like banking, telecommunications etc. To make sure that the bond return is only attributable to interest risk and credit risk, we exclude those inflation-
linked bonds. For the remaining 481 bonds, we are looking at the period from 2000 to 2010. According to Table 6.1, about 85% of the bonds are denominated in US Dollar, Euro and Sterling Pounds. Because Caplet implied volatility and swap curves are most available in these currencies, we trimmed the data further to bonds denominated only in these three currencies.

Table 6.1: summary of top 500 bond data (inflation-linked bonds eliminated)

<table>
<thead>
<tr>
<th>Bond data</th>
<th>Sovereign</th>
<th>Financial</th>
<th>Multi-National</th>
<th>Telecom</th>
<th>Healthcare Services</th>
<th>Manufacturing</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUR</td>
<td>218</td>
<td>12</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>242</td>
<td></td>
</tr>
<tr>
<td>USD</td>
<td>111</td>
<td>15</td>
<td>5</td>
<td>1</td>
<td>132</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GBP</td>
<td>26</td>
<td>3</td>
<td>3</td>
<td></td>
<td>32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAD</td>
<td>22</td>
<td></td>
<td>3</td>
<td>132</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHF</td>
<td>8</td>
<td>2</td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JPY</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUD</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLN</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td></td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEK</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KOW</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOK</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRY</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MYR</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NZD</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZAR</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grand Total</td>
<td>419</td>
<td>38</td>
<td>21</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>481</td>
</tr>
</tbody>
</table>

Because the bond return data looks at all possible time to maturity $T - t$ yet the implied volatility as well as the swap rate are available only at certain maturities, e.g. $T = 1y, 3y, 5y, \ldots$, curve fitting technique is employed. Here we used the Nelson-Siegel model, which is preferred by many central banks. Yet please note that errors will be introduced
Figure 6.1: Implied interest rate volatility (%) v.s. Bond return(log)

The data has also confirmed that the implied interest rate volatility is positively related to the bond return fluctuation. We can see that when implied interest rate volatility increases, log bond returns oscillate around zero with a bigger range. We can see from Figure 6.1 that this trend is not monotone. The return are becoming more stable in volatility range between 6% and 8% as well as 13% and 14%. We also applied curve fitting techniques to smooth downside edge of bond return. On the other hand, since most of our bonds are sovereign bonds, the CDS data calculated from equation (6.2) are clustering around zero. It is difficult to discern the trends in Figure 6.2 between the CDS
Figure 6.2: CDS (%) v.s. Bond return (log)
range from -0.1 to 0.1, where most of the observations are located. By applying curve fitting techniques, we are able to extract a CDS lookup table with reasonable values as well, which will be presented in Table 6.3 in Section 5.3.

Figure 6.3: CDS (%) v.s. Implied interest rate volatility (bps)

Finally, the scatter plot between implied volatility and CDS in Figure 6.3 confirms that there is no pattern indicating correlation, when we are examining our dataset. Hence we treat these two risk factors as independent and combine the two lookup values using equation (6.8)
6.3 Result and Discussion

In this chapter, since we do not have VAR LTV, we are only concerned with comparing the URB LTV to the real portfolio return and evaluate its performance. The success of a Rule-Based system depends on the quality of the lookup table. Table 6.2 and table 6.3 presents the the two lookup tables of interest rate risk and credit risk.

<table>
<thead>
<tr>
<th>Quantile</th>
<th>Implied Volatility</th>
<th>LTV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 20%</td>
<td>$\leq 10.42%$</td>
<td>0.9719</td>
</tr>
<tr>
<td>20%-40%</td>
<td>10.42%-11.58%</td>
<td>0.9665</td>
</tr>
<tr>
<td>40%-60%</td>
<td>11.58%-12.69%</td>
<td>0.9654</td>
</tr>
<tr>
<td>60%-80%</td>
<td>12.69%-13.88%</td>
<td>0.9642</td>
</tr>
<tr>
<td>above 80%</td>
<td>$\geq 13.88%$</td>
<td>0.9631</td>
</tr>
</tbody>
</table>

More than 80% of the implied volatilities are between than 10% and 15%. Note that these volatilities are annualized. This also contributed to the lookup value at the higher quantile. Although the lookup values are still dropping as the volatility increases, the decrease is trivial. This is, in the high volatility section, only a small part of the return can be explained by interest rate. The reversed phenomenen can be observed in CDS lookup table. The majority of CDS data concentrated in the range between -0.0038 to 0.08. The difference between lookup values are also extremely small. Only when the spread increased up to 1.5%, the credit lookup value decreases relatively substaintially.

Figure 6.4 shown the breaches detected in the 3 systems: interest rate risk only, credit risk only and these two combined. It can be seen that credit risk system fails to perform.
Table 6.3: CDS Lookup Table

<table>
<thead>
<tr>
<th>Quantile</th>
<th>CDS</th>
<th>LTV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 50%</td>
<td>≤ -0.02%</td>
<td>0.9749</td>
</tr>
<tr>
<td>50%-70%</td>
<td>-0.02% - 0.01%</td>
<td>0.9747</td>
</tr>
<tr>
<td>70%-90%</td>
<td>-0.01% - 0.08%</td>
<td>0.9743</td>
</tr>
<tr>
<td>90%-95%</td>
<td>0.08% - 0.39%</td>
<td>0.9670</td>
</tr>
<tr>
<td>95%-99%</td>
<td>0.39% - 1.495%</td>
<td>0.9668</td>
</tr>
<tr>
<td>above 99%</td>
<td>≥ 1.495%</td>
<td>0.9514</td>
</tr>
</tbody>
</table>

Except for the first risk level 0.25, the breaches exceeded the set target at each risk appetite level: 0.5, 0.75 as well as 1%.

On the other hand, the 2-factor system overestimate the combined risks. Regardless of the risk level, the breaches are very near to zero, rendering this system fairly conservative. The interest rate risk only system yields the best results out of these three. Although its plotted line lies below the target line, its performance might improve after adjusting for the correlation error. This could be the next steps for a further study.

Table 6.4: Back Testing Result for LTV based on Bonds

<table>
<thead>
<tr>
<th>RA (breach rate target)</th>
<th>2-dimensional breach</th>
<th>Interest Breach</th>
<th>Credit Breach</th>
<th>Average 2 dimensional</th>
<th>Average Interest</th>
<th>Average Credit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00%</td>
<td>0.00%</td>
<td>0.32%</td>
<td>1.43%</td>
<td>0.9303</td>
<td>0.9644</td>
<td>0.9740</td>
</tr>
<tr>
<td>0.75%</td>
<td>0.00%</td>
<td>0.19%</td>
<td>1.43%</td>
<td>0.9363</td>
<td>0.9613</td>
<td>0.9740</td>
</tr>
<tr>
<td>0.5%</td>
<td>0.00%</td>
<td>0.12%</td>
<td>0.77%</td>
<td>0.9280</td>
<td>0.9580</td>
<td>0.9687</td>
</tr>
<tr>
<td>0.25%</td>
<td>0.00%</td>
<td>0.05%</td>
<td>0.15%</td>
<td>0.9125</td>
<td>0.9534</td>
<td>0.9571</td>
</tr>
</tbody>
</table>

The same performance can also be seen in Figure 6.5. In the plot of breach rate v.s. average LTV, the curve of interest risk URB LTV is lying lower than and to the right of the curve of credit risk. This indicates that the LTV based solely on interest risk can generate higher loan to value to customer and in the same time lower risk to the bank.
Figure 6.4: Actual breach rate v.s. Target Rate

Figure 6.5: Breach rate v.s. Average LTV
To put it more simply, based on this dataset and the lookup tables, the credit risk LTV is dominated by the interest risk LTV. From Table 6.1, we can see that more than 85% of the bonds are government sovereign bonds. Most of them are regarded as risk free or having low risk. Therefore, the CDS lookup value may not the accurate predictor of the market value movement. Again, the two-dimensional URB LTV is too conservative. It has a breach rate as close as to 0% and also provide a much lower average LTV than the other two. Note that although in the combination the credit risk LTV is not accurate and underestimate risk, the 2-dimensional average LTV is still relatively low. One reason is that this is due to the erroneous information contained in the CDS lookup value. This may be also explained by the fact the correlation among bonds are corrected so that the capital relief due to diversification is missing.
Chapter 7

Summary

We have conducted 4 studies in this paper to investigate the a Rule-Based system to determine loan hair cut, or Loan to Value (LTV) for collateral portfolio consisting of equity stocks and bonds. The Rule-Based system has two important components: the mechanism of calculating the stand-alone LTV for each portfolio constituent; the adjustment of diversification. The first component, the lookup table, stays the same through out the first three studies, which focus on the asset class equity; it also has a similar form in the fourth study of bonds. On the other hand, the second component, namely the diversification adjustment evolves from an empirical driven and rather ad-hoc manner to a Diversification Coefficient based a multi-factor modelling framework. In the forth study, we left out the diversification adjustment and devote the investigation to different risk drivers used to build up a lookup table.

The empirical favourable are positive and contain useful information. In the equity cases, we have seen close fit between the RB LTV and the target, VaR LTV. The backtesting
results also shows that comparing to a VaR LTV based on Gaussian assumption, the RB LTV is more cautious and the results also improves after diversification adjustment. In the bond LTV case, we found out the interest risk LTV produces good decent backtesting results while the credit risk LTV fails to perform. This can be explained by the fact that most of the data are sovereign bonds, which can be virtually considered as default risk-free.

The findings of this thesis can be concluded in twofolds. First of all, the advantages of such a Rule-Based system against a variance-covariance VaR system are its relative independence of distribution assumptions, ease to use and intuitiveness. The lookup table is virtually a historical VaR system and does not require an assumption for the return distribution. Hence the error introduced by misfitted distribution can be avoided. Complying to the regulatory requirements, the lookup table needs calibration only quarterly or monthly instead of daily. A large amount of computing effort, which is otherwise required in a variance-covariance VaR LTV system can be saved. The Diversification Coefficient adjustment is quite flexible. The aggregation of concentration risk of different factors can be achieved in an additive manner. The Diversification Coefficient methodology also condense the massive variance-covariance matrixes for all relevant equity into a much a much smaller matrix for the concentration factors. The last benefit of this Rule-Based system is that the risk drivers and concentration factors are quite intuitive. It could be easily explained to the clients without sophisticated statistical knowledge.

There exist also shortcomings of the RB LTV system. The performance of this system depends heavily on the calibration data quality, the choice of risk drivers and concentration factors. Due to the historical VaR nature of the lookup table scheme, it is importantl to
have as many as observations possible in the left tail. The dataset selection and calibration period is also of vital significance. For example, in the first study, the inclusion of the Asian Financial Crisis helps to predict extreme losses in the subprime crisis. Similarly, it would be inappropriate to calibrate the lookup table based on West European stocks and then use it to determine LTV for a portfolio consisting of stocks from developing countries like Brazil and Russia. The right choice of risk drivers guarantees a strong linkage between it and the asset return. For instance, credit default risk as a risk driver has low relevance to the return of most sovereign bonds. It should be also easy as well as feasible to calculate and there shouldn’t be too many risk drivers. Otherwise the Rule-Based system could be overcomplicated. The concentration factors should have high explanatory power to the asset return correlations. In our case, they should also be intuitive and with data available. Last but not least, the derivation of the Diversification Coefficient is based on the assumption of a multivariate normal distribution. The diversification adjustment is aimed to replicate the linear correlation between asset returns at best. We have seen in the back testing results of the third study that the fully-blown model tends to underestimate risk. Besides approximation error, another possible explanation could be that the non-linear part is missing and the full picture of correlation is not captured, hence the over-adjustment for diversification effects. The refinement of the Diversification Coefficient for equity and derivation of its counterpart for bond, remain as future possible extension of this thesis.

In summary, we have established in this paper a Rule-Based system to determine Loan to Value for equity and bond portfolios as collaterals for a Lombard Loan. The RB LTV in general can successfully level with and in some cases outperform the target, VaR LTV.
In compliance to the Basel II regulation, this RB LTV system could serve as a promising alternative as to the VaR system to calculate loan hair cut in the bank operation.
Bibliography


    Chapman & Hall/CRC financial mathematics series

    nomic capital, Deutsche Bundesbank Discussion Paper (series 2), no 9 and National

    with Smile, Inflation and Credit (2nd ed. 2006 ed.). Springer Verlag.
# List of Tables

3.1 MSCI Stock Information .................................................. 27

3.2 LTV lookup table .......................................................... 29

3.3 Descriptive statistics of LTVs (In sample) .......................... 30

3.4 Regression results for in sample period .............................. 33

3.5 Regression results for out-of-sample period ......................... 34

3.6 Descriptive statistics of LTVs (Out-of-Sample) .................... 36

3.7 Summary statistics of return volatility .............................. 36

4.1 Distribution statistics of portfolio return .......................... 42

4.2 MSCI Stock Information, Industry ..................................... 52

4.3 Regression results of RB LTV with Approximated Diversification Coefficient (ADC) adjustment on student-t VAR LTV .............................. 54

4.4 Backtesting results with different Risk Appetite (RA) ............. 56
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>summary of top 500 equity data</td>
<td>71</td>
</tr>
<tr>
<td>5.2</td>
<td>Regression results of RB LTV with modified Herfindahl Index with 2-</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>dimensional concentration</td>
<td></td>
</tr>
<tr>
<td>5.3</td>
<td>Backtesting results with different Risk Appetite (RA) in sample</td>
<td>75</td>
</tr>
<tr>
<td>5.4</td>
<td>Backtesting results with different Risk Appetite (RA) out-of-sample</td>
<td>75</td>
</tr>
<tr>
<td>5.5</td>
<td>Cross comparison between different correction functions (1)</td>
<td>80</td>
</tr>
<tr>
<td>5.6</td>
<td>Cross comparison between different correction functions (2)</td>
<td>80</td>
</tr>
<tr>
<td>6.1</td>
<td>summary of top 500 bond data (inflation-linked bonds eliminated)</td>
<td>86</td>
</tr>
<tr>
<td>6.2</td>
<td>Interest Rate Lookup Table</td>
<td>90</td>
</tr>
<tr>
<td>6.3</td>
<td>CDS Lookup Table</td>
<td>91</td>
</tr>
<tr>
<td>6.4</td>
<td>Back Testing Result for LTV based on Bonds</td>
<td>91</td>
</tr>
</tbody>
</table>
List of Figures

3.1 Correlation and portfolio volatility ........................................... 24
3.2 Scatter plot unadjusted VAR LTV v.s. RB LTV (In Sample) ............ 31
3.3 Histogram of Herfindahl Index .................................................... 32
3.4 Scatter plot Herfindahl v.s. LTV difference (right), correlation = -0.3190 . 32
3.5 3 Scatter plot VAR LTV v.s. RB LTV (In-sample) ........................ 34
3.6 3 Scatter plot VAR LTV v.s. URB LTV (Out-of-sample) ................... 35
3.7 3 Scatter plot VAR LTV v.s. RB LTV (Out-of-sample) ................... 35
4.1 Empirical data plot and fitted distribution for the in sample period .... 39
4.2 Empirical data plot and fitted distribution for the out-of-sample period ... 40
4.3 Tail of empirical data plot and fitted distribution for the in sample period 41
4.4 Tail of empirical data plot and fitted distribution for the out-of-sample period 41
4.5 Scatterplot of the APT LTV v.s. Meta-Gausian LTV ..................... 55
4.6 student-t implied volatility in lookup bins (red lines are for lookup bucket and brown lines are the implied volatility) ........................................ 57

4.7 Gaussian implied volatility in lookup bins (red lines are for lookup bucket and blue lines are the implied volatility) ........................................ 58

4.8 Back testing result of the APT LTV v.s. Meta-Gaussian LTV ............... 59

4.9 Actual breach rate v.s. Target Rate ................................................. 60

5.1 Convergence of sector $HHI$ (red line) to $N$-asset $HHI$ (blue line). The x-axis is the number of sectors, $S$ and the y-axis the difference between these two $HHIs$. ................................................................. 69

5.2 Scatterplot of the Empirical-Gaussian LTV v.s. URB LTV in sample ..... 72

5.3 Scatterplot of the Empirical-Gaussian LTV v.s. RB LTV in sample ..... 73

5.4 Scatterplot of the Empirical-Gaussian LTV v.s. URB LTV out-of-sample . 74

5.5 Scatterplot of the Empirical-Gaussian LTV v.s. RB LTV out-of-sample . 74

5.6 Breach rate v.s. Average LTV in sample (RB LTV: 2-dimensional Herfindal Index, Sector: N-asset Herfindahl Index) .......................... 77

5.7 Breach rate v.s. Average LTV out-of-sample (RB LTV: 2-dimensional Herfindal Index, Sector: N-asset Herfindahl Index) .......................... 77

5.8 Actual breach rate v.s. Target Rate in sample (RB LTV: 2-dimensional Herfindal Index, Sector: N-asset Herfindahl Index) .......................... 78
5.9 Actual breach rate v.s. Target Rate out-of-sample (RB LTV: 2-dimensional Herfindal Index, Sector: N-asset Herfindahl Index 78

5.10 Cross comparison between different correction functions, Target Rate v.s Actual breach rate 81

5.11 Cross comparison between different correction functions, Average LTV v.s.
Actual breach rate 81

6.1 Implied interest rate volatility (%) v.s. Bond return(log) 87

6.2 CDS (%) v.s. Bond return(log) 88

6.3 CDS (%) v.s. Implied interest rate volatility (bps) 89

6.4 Actual breach rate v.s. Target Rate 92

6.5 Breach rate v.s. Average LTV 92
<table>
<thead>
<tr>
<th>East Europe</th>
<th>Far East</th>
<th>Far East D</th>
<th>Middle East</th>
<th>North America</th>
<th>Oceanic</th>
<th>South Africa</th>
<th>South America</th>
<th>West Europe</th>
<th>Energy</th>
<th>Materials</th>
<th>Industrial</th>
<th>Consumer Discretionary</th>
<th>Consumer Staples</th>
<th>Health Care</th>
<th>Financials</th>
<th>IT</th>
<th>Telecommu-</th>
<th>Utilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>East Europe</td>
<td>1.00</td>
<td>0.44</td>
<td>0.44</td>
<td>-0.14</td>
<td>0.39</td>
<td>0.40</td>
<td>0.20</td>
<td>0.42</td>
<td>0.71</td>
<td>-0.06</td>
<td>0.22</td>
<td>-0.11</td>
<td>-0.18</td>
<td>-0.28</td>
<td>-0.28</td>
<td>0.23</td>
<td>0.16</td>
<td>-0.24</td>
</tr>
<tr>
<td>Far East</td>
<td>0.44</td>
<td>1.00</td>
<td>0.52</td>
<td>0.14</td>
<td>0.24</td>
<td>0.38</td>
<td>0.22</td>
<td>0.18</td>
<td>0.42</td>
<td>0.07</td>
<td>0.22</td>
<td>0.03</td>
<td>-0.33</td>
<td>-0.31</td>
<td>-0.01</td>
<td>0.14</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>Far East D</td>
<td>0.44</td>
<td>0.52</td>
<td>1.00</td>
<td>0.08</td>
<td>0.43</td>
<td>0.34</td>
<td>0.35</td>
<td>0.00</td>
<td>0.35</td>
<td>0.18</td>
<td>0.20</td>
<td>-0.26</td>
<td>-0.34</td>
<td>-0.40</td>
<td>-0.16</td>
<td>0.32</td>
<td>0.23</td>
<td>-0.25</td>
</tr>
<tr>
<td>Middle East</td>
<td>-0.14</td>
<td>0.14</td>
<td>0.08</td>
<td>1.00</td>
<td>0.19</td>
<td>0.15</td>
<td>-0.17</td>
<td>0.01</td>
<td>0.16</td>
<td>0.13</td>
<td>0.38</td>
<td>-0.25</td>
<td>-0.06</td>
<td>0.01</td>
<td>0.03</td>
<td>-0.47</td>
<td>0.20</td>
<td>-0.21</td>
</tr>
<tr>
<td>North America</td>
<td>0.39</td>
<td>0.24</td>
<td>0.43</td>
<td>0.19</td>
<td>1.00</td>
<td>0.26</td>
<td>0.52</td>
<td>0.11</td>
<td>0.49</td>
<td>0.05</td>
<td>0.26</td>
<td>-0.09</td>
<td>-0.35</td>
<td>-0.13</td>
<td>0.09</td>
<td>-0.03</td>
<td>0.18</td>
<td>-0.19</td>
</tr>
<tr>
<td>Oceanic</td>
<td>0.40</td>
<td>0.38</td>
<td>0.34</td>
<td>0.15</td>
<td>0.26</td>
<td>1.00</td>
<td>0.35</td>
<td>0.27</td>
<td>0.52</td>
<td>0.18</td>
<td>0.29</td>
<td>-0.27</td>
<td>-0.36</td>
<td>-0.69</td>
<td>-0.38</td>
<td>0.35</td>
<td>0.28</td>
<td>-0.20</td>
</tr>
<tr>
<td>South Africa</td>
<td>0.20</td>
<td>0.22</td>
<td>0.35</td>
<td>-0.17</td>
<td>0.52</td>
<td>0.35</td>
<td>1.00</td>
<td>-0.08</td>
<td>0.37</td>
<td>0.14</td>
<td>0.11</td>
<td>0.10</td>
<td>-0.62</td>
<td>-0.27</td>
<td>-0.09</td>
<td>0.19</td>
<td>0.13</td>
<td>0.01</td>
</tr>
<tr>
<td>South America</td>
<td>0.42</td>
<td>0.18</td>
<td>0.00</td>
<td>0.01</td>
<td>0.11</td>
<td>0.27</td>
<td>-0.08</td>
<td>1.00</td>
<td>0.49</td>
<td>-0.13</td>
<td>0.17</td>
<td>0.15</td>
<td>-0.13</td>
<td>-0.06</td>
<td>0.11</td>
<td>0.07</td>
<td>-0.18</td>
<td>-0.12</td>
</tr>
<tr>
<td>West Europe</td>
<td>0.71</td>
<td>0.42</td>
<td>0.55</td>
<td>0.16</td>
<td>0.49</td>
<td>0.53</td>
<td>0.37</td>
<td>0.49</td>
<td>1.00</td>
<td>-0.01</td>
<td>0.37</td>
<td>-0.13</td>
<td>-0.32</td>
<td>-0.22</td>
<td>-0.01</td>
<td>0.17</td>
<td>0.14</td>
<td>-0.34</td>
</tr>
<tr>
<td>Energy</td>
<td>-0.06</td>
<td>0.07</td>
<td>0.18</td>
<td>0.13</td>
<td>0.05</td>
<td>0.18</td>
<td>0.14</td>
<td>-0.13</td>
<td>-0.01</td>
<td>1.09</td>
<td>0.27</td>
<td>-0.41</td>
<td>-0.29</td>
<td>-0.41</td>
<td>-0.34</td>
<td>-0.26</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>Materials</td>
<td>0.22</td>
<td>0.22</td>
<td>0.20</td>
<td>0.38</td>
<td>0.26</td>
<td>0.29</td>
<td>0.11</td>
<td>0.17</td>
<td>0.37</td>
<td>0.27</td>
<td>1.00</td>
<td>-0.17</td>
<td>-0.42</td>
<td>-0.26</td>
<td>0.02</td>
<td>0.32</td>
<td>-0.05</td>
<td>-0.33</td>
</tr>
<tr>
<td>Industrial</td>
<td>-0.11</td>
<td>0.03</td>
<td>-0.26</td>
<td>-0.25</td>
<td>-0.09</td>
<td>-0.27</td>
<td>0.10</td>
<td>0.15</td>
<td>-0.13</td>
<td>-0.41</td>
<td>-0.12</td>
<td>1.00</td>
<td>0.02</td>
<td>0.41</td>
<td>0.25</td>
<td>0.04</td>
<td>0.25</td>
<td>0.07</td>
</tr>
<tr>
<td>Consumer Discretionary</td>
<td>-0.18</td>
<td>-0.33</td>
<td>-0.34</td>
<td>-0.06</td>
<td>-0.35</td>
<td>-0.36</td>
<td>-0.62</td>
<td>-0.13</td>
<td>-0.32</td>
<td>-0.29</td>
<td>-0.42</td>
<td>0.02</td>
<td>1.00</td>
<td>0.36</td>
<td>0.02</td>
<td>-0.03</td>
<td>-0.10</td>
<td>0.15</td>
</tr>
<tr>
<td>Consumer Staples</td>
<td>-0.28</td>
<td>-0.31</td>
<td>-0.40</td>
<td>0.01</td>
<td>-0.13</td>
<td>-0.69</td>
<td>-0.27</td>
<td>-0.06</td>
<td>-0.22</td>
<td>-0.41</td>
<td>-0.26</td>
<td>0.41</td>
<td>0.36</td>
<td>1.00</td>
<td>0.42</td>
<td>-0.36</td>
<td>-0.19</td>
<td>0.22</td>
</tr>
<tr>
<td>Health Care</td>
<td>-0.03</td>
<td>-0.01</td>
<td>-0.16</td>
<td>0.03</td>
<td>0.09</td>
<td>-0.38</td>
<td>-0.09</td>
<td>0.11</td>
<td>-0.01</td>
<td>-0.44</td>
<td>0.03</td>
<td>0.25</td>
<td>0.03</td>
<td>0.42</td>
<td>1.00</td>
<td>-0.25</td>
<td>-0.16</td>
<td>-0.11</td>
</tr>
<tr>
<td>Financials</td>
<td>0.23</td>
<td>0.14</td>
<td>0.32</td>
<td>-0.47</td>
<td>-0.03</td>
<td>0.35</td>
<td>0.19</td>
<td>0.07</td>
<td>0.17</td>
<td>-0.25</td>
<td>-0.28</td>
<td>0.08</td>
<td>-0.03</td>
<td>-0.36</td>
<td>-0.29</td>
<td>1.08</td>
<td>0.08</td>
<td>-0.15</td>
</tr>
<tr>
<td>IT</td>
<td>0.16</td>
<td>0.06</td>
<td>0.23</td>
<td>0.20</td>
<td>0.18</td>
<td>0.28</td>
<td>0.13</td>
<td>-0.18</td>
<td>0.14</td>
<td>-0.33</td>
<td>-0.05</td>
<td>-0.25</td>
<td>-0.10</td>
<td>-0.19</td>
<td>-0.16</td>
<td>0.09</td>
<td>1.00</td>
<td>-0.15</td>
</tr>
<tr>
<td>Telecommuni-</td>
<td>-0.24</td>
<td>0.05</td>
<td>-0.25</td>
<td>-0.21</td>
<td>-0.19</td>
<td>-0.20</td>
<td>0.01</td>
<td>-0.12</td>
<td>-0.34</td>
<td>-0.05</td>
<td>-0.33</td>
<td>0.25</td>
<td>0.15</td>
<td>0.22</td>
<td>-0.11</td>
<td>-0.16</td>
<td>1.00</td>
<td>-0.15</td>
</tr>
<tr>
<td>Utilities</td>
<td>-0.27</td>
<td>-0.29</td>
<td>-0.13</td>
<td>0.14</td>
<td>-0.03</td>
<td>-0.60</td>
<td>-0.14</td>
<td>-0.34</td>
<td>-0.38</td>
<td>0.03</td>
<td>-0.19</td>
<td>0.07</td>
<td>0.09</td>
<td>0.49</td>
<td>0.17</td>
<td>-0.44</td>
<td>0.13</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Appendix 1 Table for factor correlation.(2006-2007)