Global Risk Management

Ren-Raw Chen, Ph.D.

Library of Congress Control Number 1234567890
Chen, Ren-Raw Global Financial Risk Management – A Quantitative Guide /
Ren-Raw Chen
ISBN 1234567890123

Contents

Preface

I have done risk management research for more than two decades. In fact, my first industry work was at JP Morgan (15 Broad Street – the legendary building) in February of 1992 when Mr. Paul Morrison hired me to validate models used by JP Morgan at the time. Since then I had been mostly in the front office (including 1997 $\sim$ 1999 at Lehman as the desk quant in the Structured Credit Desk where Mr. Ken Ulmazaki was the head). In January of 2005, I joined Morgan Stanley Model Review Group and have been back on risk management till now.

In the summer of 2012, I was asked to teach this course for a group of talented EMBA students from the Peking University (our first MSGF program) and I have been gradually collecting my past notes. That is how this book came to existence. I am extremely grateful to all the past students who took this course from me and hence discovered the numerous mistakes in the original drafts. I particularly would like to thank the 2014 MSGF and 2015 MSQF classes who had made valuable suggestions and corrections to this book.

Ren-Raw Chen

New York, New York, April, 2016

Acknowledgments

I first need to thank my wife, Hsing-Yao, for her support for all these years. Without her taking care of all the work at home, I would not have achieved not only this book, but any of my research. Her unconditional love and support is the main reason I could accomplish anything at all.

I must thank Dr. Phelim Boyle and Dr. Louis Scott for their caring and support for all these years. Without their patience, encouragement, and kind guidance, this book would never be possible.

As mentioned in Preface, former students from MSQF and MSGF have helped in many tremendous ways in improving coverage, substances, and accuracy in this book. In particular, I would like to thank Yingqi Tian and Zhifang Sun (2015 MSQF) who have discovered important mistakes in earlier manuscripts.

Part I

Introduction

Chapter 1

Math Primer

1.1 Continuous and Discrete Returns

It is quite conventional (almost universal) to use returns to measure the performance (and risks) of a financial investment. Yet, no consensus has been reached on how they are calculated.

1.1.1 Definition

A discrete return is defined as:

$$r_t = \frac{P_{t+1} - P_t}{P_t} \tag{1.1}$$

where P is price of an asset and t is time. P_{t+1} and P_t are two consecutive prices (could be daily, weekly, monthly, etc.)

A continuous return is defined as:

$$r_t = \ln \frac{P_{t+1}}{P_t} = \ln P_{t+1} - \ln P_t \tag{1.2}$$

The two returns are very close to each other when the prices are observed frequently (e.g. daily) and start to deviate from each other over low frequencies (e.g. annually). Note that if the two consecutive dates are close to each other, like daily, then equation (1.1) can be approximated as dP/P which is identical to $d \ln P$, which is the definition of the continuous return by equation (1.2).

For example, we take the FaceBook stock:

FB Returns			
Date	Price	Disc.	Cont.
12/23/2008	26.93		
12/25/2008	26.51	-0.0156	-0.0157
12/26/2008	26.05	-0.0174	-0.0175
12/27/2008	25.91	-0.0054	-0.0054
12/30/2008	26.62	0.0274	0.0270

1.1.2 Average Return

It is common that we take average returns. We often compute returns of a certain frequency (e.g. daily) and store them in a database. Then we compute an average return over a particular time horizon (e.g. one year). In such a case, we need to know how an average is taken. There are two ways of "taking an average:

- geometrically and
- arithmetically

Theoretically, a geometric average is matched with discrete returns and an arithmetic average is matched with continuous returns, as demonstrated in the following equations:

$$\begin{aligned}\bar{r} &= \sqrt[n]{\left(\frac{P_{t+1}}{P_t}\right)\left(\frac{P_{t+2}}{P_{t+1}}\right)\cdots\left(\frac{P_{t+n}}{P_{t+(n-1)}}\right)} - 1 \\ &= \sqrt[n]{\left(\frac{P_{t+n}}{P_t}\right)} - 1\end{aligned}$$

and

$$\begin{aligned}\bar{r} &= \frac{\ln\left(\frac{P_{t+1}}{P_t}\right) + \ln\left(\frac{P_{t+2}}{P_{t+1}}\right)\cdots + \ln\left(\frac{P_{t+n}}{P_{t+(n-1)}}\right)}{n} \\ &= \frac{1}{n}\ln\left(\frac{P_{t+n}}{P_t}\right)\end{aligned}$$

In terms of FB, the average discrete return is $\sqrt[4]{26.62/26.51} - 1 = -0.00289$ and the average continuous return is $\ln[26.62/26.51]/4 = 0.002289$. We see that the two averages are so close to each other (indistinguishable at the 6th decimal place).

1.1.3 Annualization and Deannualization

Usually returns are reported in a per annum term, known as annualization. For example, The daily return of FB between 12/26/2008 and 12/27/2008 is -0.0054. This one-day return needs to be "annualized" into a per annum return. Given that there are approximately 252 trading days in a year, this number is multiplied by 252 to be -1.36 or -136%.

Reverse (deannualization) is used sometimes when we need raw returns. For example a 3-month return of 12% really is 3% for three months. Since someone else did the annualization, you must reverse it to get the raw return for the 3 months.

Standard deviations are annualized and deannualized as well. Instead of multiplying (and dividing) by the same adjustment factor (e.g. 252 for daily), it is multiplied by the square root of the factor (e.g. $\sqrt{252}$). In other words, 1% daily standard deviation is translated into 15.87% per annum standard deviation.

1.2 Linear Algebra

Matrix operations are important in performing calculations for risk management, as we are dealing with portfolios of large numbers of assets. Furthermore, Microsoft Excel now is equipped with matrix operation functions (via Ctrl-Shft-Enter as opposed to Enter) that can be easily matched with mathematical expressions. As a result, using matrices is extremely convenient and efficient.

Throughout this book, a matrix is symbolized as a bold non-italic letter. Usually a subscript is given as the dimension (row by column) of the matrix. Formally, we define $\mathbf{A}_{m\times n}$ as an m-row, n-column rectangular matrix. For example,

$$\begin{bmatrix} \square & \square & \square \\ \square & \square & \square \end{bmatrix}_{2\times 3}$$

is a matrix with 2 rows and 3 columns.

1.2.1 Addition/subtraction

Two matrices can be summed or subtracted only if they have identical dimensions.

1.2.2 Multiplication/Division

Two matrices can be multiplied only if the second subscript (i.e. column) of the first matrix matches with the first subscript (i.e. row) of the second matrix.

$$\begin{bmatrix} \square & \square \\ \square & \square \\ \square & \square \end{bmatrix}_{3\times 2} \begin{bmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \end{bmatrix}_{2\times 4} = \begin{bmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \end{bmatrix}_{3\times 4}$$

and the resulting matrix has the dimension of the first subscript of the first matrix (row) and the second subscript of the second matrix, as the example above shows.

The rule of multiplication is the ith row of the first matrix "sumproduct" by the jth column of the second matrix, as demonstrated below:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 6 & 5 \end{bmatrix} \times \begin{bmatrix} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{bmatrix} = \begin{bmatrix} 5 & 11 & 17 & 23 \\ 11 & 25 & 39 & 53 \\ 16 & 38 & 60 & 82 \end{bmatrix}$$

Figure 1.1: Matrix Multiplication

$5 = 1 \times 1 + 2 \times 2.$

Divisions are performed via matrix inversion. In other words, $\mathbf{A}_{m\times n} \div \mathbf{B} = \mathbf{A}_{m\times n} \times \mathbf{B}^{-1}$. Given that only square matrices can be inversed, matrix $\mathbf{B}_{n\times n}$ must be n rows and n columns.

1.2.3 Scaling

Any matrix can be scaled up or down by multiplying by a real number. When a matrix is multiplied by a real number, it is identical to each element of the matrix being multiplied by the real number.

1.2.4 Power

Matrices can be taken to a power, just like any real number. However, unlike scaling, it is NOT equal to each element raising to the power. Since, we do not use this function often, we choose to ignore it.

1.3 Calculus

Sensitivities of prices relative to risk factors are defined as partial derivatives. P&L analyses are leveraged upon total derivatives. As a result, some basic knowledge of calculus is helpful in understanding risk.

Under a single variable, partial derivatives are the same as total derivatives, as shown below:

$$y = f(x) = 3x^2 + 2x + 5$$
$$y' = \frac{dy}{dx} = 6x + 2$$

With two variables such as:

$$z = f(x, y) = 3x^3 + 6y^2 + 4x + 2y + 5$$

partial derivatives and total derivatives are not the same. The partials (∂) are:

$$\frac{\partial z}{\partial x} = 9x + 4$$
$$\frac{\partial z}{\partial y} = 12y + 2$$

The total derivative of z is (via Taylor's series expansion):

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy$$

In order to see the impact from x and y changes, we can divide dz by dx or dy, as follows:

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\frac{dy}{dx} = 9x + 4 + (12y + 2)\frac{dy}{dx}$$
$$\frac{dz}{dy} = \frac{\partial z}{\partial x}\frac{dx}{dy} + \frac{\partial z}{\partial y} = (9x + 4)\frac{dx}{dy} + (12y + 2)$$

where we can see the interaction between x and y. If x and y are unrelated ($dx/dy = 0$ and vice versa), then totals equal partials.

Partial derivatives measure sensitivities of the price a financial asset (z) with respective to risk factors (x and y). That is, how much is the price movement if one risk factor moves by a little while all else factors are held constant (usually 1 basis point, known as DV01 or PV01).

Total derivatives measure price movements when all risk factors are considered. As shown above, it needs the result of partial derivatives (i.e. Taylor's series expansion). Both risk factors (x and y) move by a little (1 basis point) and together is the total impact of price movement. Now we can regard dz and price change. In risk management, we try to explain why and how a price is moved up or down (i.e. price change).

Either total derivatives or partial derivatives measure sensitivities of the interested variable (i.e. P&L) with respect to risk factors. In other words, either d or ∂ represents a small change. In reality, as analytical derivatives (as those shown above) do not exist, numerical derivatives are a must. For the sake of convenience, d or ∂ is often replaced by Δ. For example ΔP represents the price change.

1.4 Statistics

Statistics are a crucial tool in modeling risk. As we shall see, the most common way to view risk is an asset's price fluctuation. And the easiest way to model price fluctuation is the standard deviation of the price (or more precisely return) distribution. A very common choice of the distribution is the normal (i.e. Gaussian) distribution which is bell-shaped, symmetrical, and no upper or lower limits.

1.4.1 Random Variable

A distribution is applied to a "random variable" because a random variable is a variable that we do not know its value. Hence, a set of (possible) values are assigned, which form a distribution. In other words, a distribution is a collection of all possible values of a random variable.

Usually we represent the price of a financial asset as a random variable as we do not know its value in the future. For example, we do not know the price of the IBM a week from now. Hence, we collect all possible values of the price and form a distribution. Usually we say that the possible prices follow a normal (Gaussian) distribution.

1.4.2 Stochastic Process

A stochastic process is a collection of the same random variables over time. For example, weekly future IBM stock prices form a stochastic process.

$$\frac{dS}{S} = \mu dt + \sigma dW$$

where $dW \sim N(0, dt)$. If dt is a week, then it is equal to $1/52 \sim 0.019230769$. This means that the weekly return of a stock is normally distributed with mean $\mu \times dt$ and variance $\sigma^2 \times dt$.

We shall see that this particular model (known as the Black-Scholes model) has no knowledge of time (known as stationarity). That is, all weekly returns of the stock no matter when have the same expected value and variance. While we can regard dW just as a normal random variable, we shall note that W is named after Norbert Wiener.[1]

1.4.3 dS and ΔS

Sometimes, we use ΔS for price change, as opposed to dS.

Sometimes, we just use Δ for price change.

1.4.4 Normal Distribution

Normal (Gaussian) distribution is used dominantly. Hence it is essential that one can obtain normal probabilities quickly. One method is to use the lookup table which is embedded in many investments texts.

The table only runs to the second digit for the critical value. To obtain a probability of, say, 1.2524, one usually uses a linear interpolation to approximate. Since 1.2524 is in between 1.25 and 1.26, we just take a weighted average of the two. $N(1.25) = 0.8944$ and $N(1.26) = 0.8962$. Hence, $N(1.2524) = (76\%) \times 0.8944 + (24\%) \times 0.8962 = 0.8948$. The other is to use the Excel function `NormSDist(`x`)` where x is the critical value in the Normal distribution. For example `NormSDist(1.2524)` $= 0.89478786$.

The table usually presents results for only positive x. This is because Normal is symmetrical and any negative x value can be flipped to acquire the positive value.

[1] Norbert Wiener (November 26, 1894 – March 18, 1964) was an American mathematician and philosopher. He was Professor of Mathematics at MIT.

Normal Probability Table

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Example: N(1.25)=0.8944 and N(-1.25)=1-N(1.25)=0.1056

Figure 1.2: Normal Probability Table

In other words, $N(-x) = 1 - N(x)$. For example, `NormSDist(1.64485)` is 0.95 and `NormSDist(-1.64485)` is $1 - 0.95 = 0.05$. Excel also provides an inverse function to obtain the critical value once given the probability – `NormSInv(`p`)`. For example `NormSInv(0.05)` is -1.64485.

While `NormSDist(`x`)` gives the probability of a standard normal (mean 0 and variance 1), `NormDist(`μ, σ^2, x`)`, on the other hand, gives a normal probability with mean and variance.

Excel also provides various other statistical functions which are very helpful.

Chapter 2

Overview

2.1 Introduction

Risk management is a crucial function in any corporation or business. And its success relies on both scientific tools and human judgments. The former can be standardized globally yet the latter depends on local cultures which vary from region to region, religion to religion, and regulation to regulation.

Hence, the focus of this book is on the scientific side of risk management, the part that can be globalized. Furthermore, while some of the tools introduced in this book can be applied to all types of companies and businesses, they are mostly used by financial companies. Examples given in this book are also mostly from the financial sector.

2.1.1 Business risks versus financial risks

In general, there are two broad types of risk any company or business faces the business risk (the left-hand-side of the balance sheet) and the financial risk (the right-hand-side of the balance sheet). The business risk refers to the uncertainty in the investments of the company, i.e. the assets. The financial risk refers to the uncertainty the debt capacity (in other words, the instability of its financing capability). For a financial institution, its assets consist of pre-dominantly financial securities and hence its business risk is also financial and the tools for managing the financial risk apply to managing its assets.

For an industrial company (i.e. non-financial), the business risk and the financial risk can be quite unrelated. Take Apple Inc. as an example, its assets are all the manufacturing facilities and its top-notch scientists, engineers, and workers.

Its business risk comes mainly from it can maintain its superiority in the cell phone and tablet businesses. The way to manage this risk is to keep innovating and being the dominant leader in these businesses. Hence the tools of managing the financial risk have absolutely no connection to how it manages its business risk. In fact, in the case of Apple Inc., the business risk is so important that it makes the financial risk entirely trivial.

But history has taught us that financial risk could be crucial for industrial companies. In 2001, a company named Excite@Home filed bankruptcy due to its failure to fulfill a convertible bond obligation of near $30 million. During the internet bubble (late 1990s to early 2000s), many internet companies issued convertible bonds as a new financial innovative tool to finance their large capital investments. Convertible bonds were popular then because mutual funds and especially pension funds can purchase (these companies cannot purchase stocks not listed in S&P 500 list of constituents) and benefit from their fast growth. Excite@Home was majority owned by AT&T (over 30%) and was the sole provider of internet services (combined with Comcast) in the tri-state area. It was perceived that Excite@Home to be one of the major internet services providers (especially because it was backed by AT&T). Hence, Excite@Homes bankruptcy was strictly a liquidity default (lack of cash) and not an economic default. In other words, Excite@Home was one typical example of how a successful industry company can bankruptcy due to bad management of its financial risk.

While the two sides of a companys risk must both be managed well to guarantee the companys survival, it has not been possible to integrate these two sides well until now. The new concept of enterprise risk has emerged recently that both business and financial risks are connected and need to be managed in an integrated manner. ERP (Enterprise Resource Planning) as a result becomes a standard for companies to follow. However, the focus of this book is the tools for financial risk management. Yet, as mentioned above, these tools can also be applied to many aspects of the business risk of the company. In fact, the overlap of the tools is why the reason ERP can be successfully applied in companies.

2.1.2 Financial risks

Financial risks are caused by fluctuations in prices of goods and financial securities generally called asset classes. The following are common asset classes:

- commodities: which are further categorized into:
- agriculture products such as corn, soy bean, oranges, sugar, cotton, etc.
  - metallurgical products such as gold, silver, copper, etc.

  - livestock products such as cattle, lean hogs, pork belly, etc.
  - energy products such as oil (various kinds), gas, etc.

- foreign currencies (euro, yen, yuan, etc.)
- equities
- interest rates (Treasuries, LIBOR , OIS , etc.)
- credit
- mortgage-backed securities

While each asset class (and its subclass) has its own risk, these risks are generally group into four major categories, defined by Basel Accords that are the most sophisticated bank regulation documentation available today. We shall describe the three Basel Accords later. According to the Accords, financial risk can be categorized as:

- Market risk
- Credit risk
- Liquidity risk
- Operational risk

Market risk

Market risk refers to risk that should be monitored and managed on a very frequent basis (at least daily) because as market conditions change, market prices move and hence profits and losses are generated. As a result, market risk is usually measured by the degree of fluctuations in prices know as volatility, or in technical terms the standard deviation of the price change (or return). In other words, the higher is the volatility (standard deviation), the higher is the risk. Market risk exists in every asset class, but some are easier to monitor than others. For example, equities are transacted very frequently (many times a day) so their market risk is easier to compute and measured. Fixed income securities, such as swaps, are not transacted frequently and hence the market risk is not easy to compute. As a result, some scientific methods are necessary to estimate the market risk.

Credit risk

Credit risk refers to losses due to credit events the most severe of which is bankruptcy. Credit default swaps (CDS) are the contract that provide the perfect hedge of the bankruptcy risk. CDS will be fully explored later in this book. In addition to the bankruptcy risk, market participants also transact bonds whose prices are based upon the probability of bankruptcy (known as credit spreads). As the probabilities move up or down, bond prices (or spreads) move up or down as well. Finally, pension funds and selected mutual funds are regulated to only purchase bonds with investment grades. Hence, if a bond is downgraded out of the investment group, it will be dumped in the marketplace immediately and huge losses can occur. This is known as the migration risk that must be also managed.

Liquidity risk

Liquidity risk refers to losses of asset values due to lack of trading in the market place. Usually, lack of trading introduces higher bid/offer spreads. In a severe situation, the market becomes one-sided and offer (or bid) disappears. Then the bid price can skyfall and large liquidity discounts should occur. During the 2008 crisis, we witnessed the exact same phenomenon. Banks were dumping subprime portfolios to the market after they realized that they were too slow in reacting to the defaults in subprime loans. Lack of buying (one-sided market) caused the prices of subprime portfolio to skyfall and resulted in the liquidity crisis.

Operational risk

Operational risk refers to losses due to human mistakes or frauds. In early September 2011, UBS announced that it had lost 2.3 billion dollars, as a result of unauthorized trading performed by Kweku Adoboli, a director of the bank's Global Synthetic Equities Trading team in London. On 16 April 2008, The Wall Street Journal released a controversial article suggesting that some banks might have understated borrowing costs they reported for the LIBOR during the 2008 credit crunch that may have misled others about the financial position of these banks. On 27 July 2012, the Financial Times published an article by a former trader which stated that LIBOR manipulation had been common since at least 1991. LIBOR underpins approximately \$350 trillion in derivatives. On 27 June 2012, Barclays Bank was fined \$200 million by the Commodity Futures Trading Commission, \$160 million by the United States Department of Justice and 59.5 million by the Financial Services Authority for attempted manipulation of the LIBOR and Euribor rates.

Collateral risk

Collateral risk is not a Basel defined risk but it has been under the spotlight after the 2008 crisis. Prior to the crisis, most transactions were done naked, which is that no collaterals were provided. During the crisis, such naked positions suffered huge loss of value. Hence, after the crisis, more and more transactions were done covered, which is that equivalent value of assets were provided as collaterals. As a result, banks now hold a large pool of assets from collaterals which go up and down in value constantly. Furthermore, banks now try to efficiently manage these collaterals by loaning them out (known as rehypothecation). As a result, managing the collateral risk is also part of the scope of risk management

2.1.3 Ways to measure and manage these risks

To manage risk effectively, we must quantify risk first. Profit or loss is a 0-1 event either you lose money or you make money. Risk represents the likelihood of losing money. As the likelihood rises, preventions must be taken (i.e. adjustments to the portfolio) in order to lose the minimum. Likewise, if the likelihood falls, more stakes can be taken to enhance the win. How to measure the likelihood of losing, i.e. risk, as a result, is a crucially important task. However, likelihood is just a concept. To have a concrete measure to represent it is highly difficult. Also, for different types of risk, the representations can also differ. In the following, we see some standard risk representations:

- market risk – VaR, stress test
- credit risk – JTD (jump to default), PD (probability of default), LGD (loss given default), EAD (exposure at default), EL (expected loss). UL (unexpected loss), ES (expected shortfall), EC (economic capital), CVA (credit value adjustment), and CVaR (credit value at risk)
- liquidity risk – liquidity weights, liquidity Value at Risk (LaR)
- operational risk – data mining (indicative)
- collateral management – rehypothecation

2.2 Review of Simple Hedging

The extreme risk management is to eliminate risk completely. However, in the concept of efficient market, no risk implies no return (or the minimal risk-free return).

As a result, the objective of risk management is not the removal of risk, but to control risk under a desired level while the objective of returns is maintained.

In standard derivatives text books, the introduction of hedging is via risk elimination known as the risk-free arbitrage. The famous Black-Scholes model is built on such a concept. There are two types of hedging static and dynamic. A static hedge is a buy-and-hold hedge, which is to engage a hedging trade and do nothing till the end of the hedging or investment horizon. A typical example is hedging with forwards or futures. Buying an asset and hedging with its futures completely eliminates the risk as the futures moves dollar-for-dollar with the underlying asset. As a result, there is no risk and in exchange there is no return. One could modify the hedge with an option where some the downside risk is eliminated but the upside potential is retained. Given that there is no free lunch, such a hedge is costly. In other words, an option hedge (usually puts) can be viewed as two separate hedges one is to eliminate risk which is free, like futures, and the other is to pay for upside returns.

A dynamic hedge is the same idea except that frequent rebalancing is required. That is, the hedge is meant for just a very short period of time (e.g. day) and at the end of the hedging period, re-hedging is necessary. Re-hedging implies buying and selling of hedging securities. The math of the hedging quantity (known and the hedge ratio) is much more complex than that of the static hedging. Furthermore, the computation of the hedge ratios require various financial models.

The pros of using the dynamic hedging are that it costs less and in many cases static hedges are not available. The con is that it relies on financial models that can be problematic and inaccurate.

The usual derivatives used for hedging (static or dynamic) are:

- options (puts)
- fowards and futures (dynamic)
- swaps

2.3 Market Risk

While there are many risk management tools, over the years a consensus has been reached that the concept of value at risk is most desirable. After the recent crisis, various enhancements have been proposed and the most important of which is stress testing.

2.3.1 Value at Risk

Value at Risk, abbreviated as VaR, was first developed by JP Morgan back in the early 90s. The basic idea is that fund managers need to know at a certain percentage (say 5%) how much money will he or she lose over the next day (or any investment horizon). In plain language, how much value is at risk over the next day?

To answer this question, one must have a distribution. Not surprisingly the most common choice is the normal distribution. The following diagram demonstrates where the 5% probability loss is in a normal distribution.

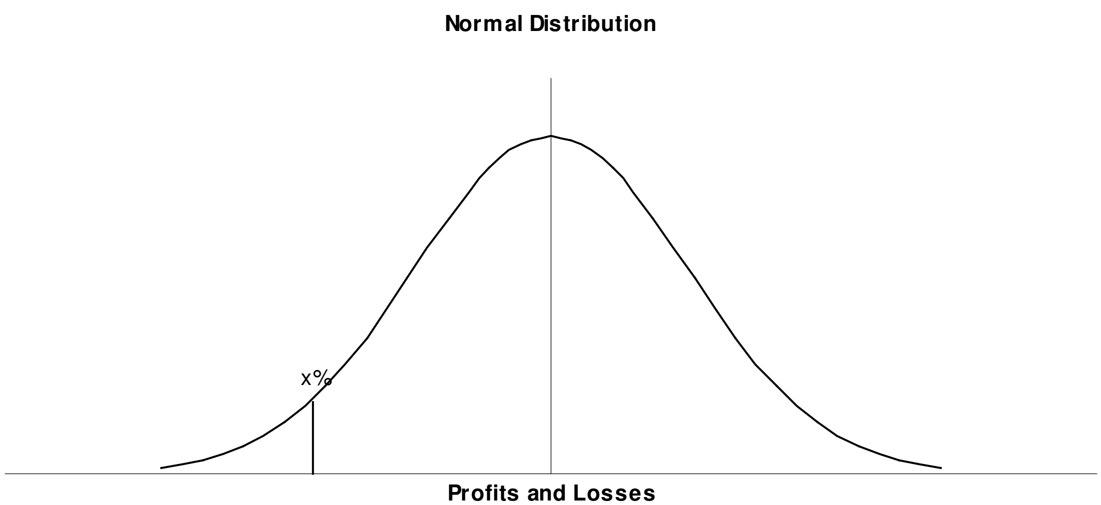


Figure 2.1: Normal Distribution

The normal distribution is often criticized as having too thin tails, which are contradictions to what researchers observe empirically. The following diagram compares a t distribution with the degrees of freedom of 2 and a normal distribution. As it is easily seen, the t distribution has much fatter tails than those of the normal. However, normal distributions are the only distribution that can be scaled by (the square root of) time. The details will be explored in the next chapter.

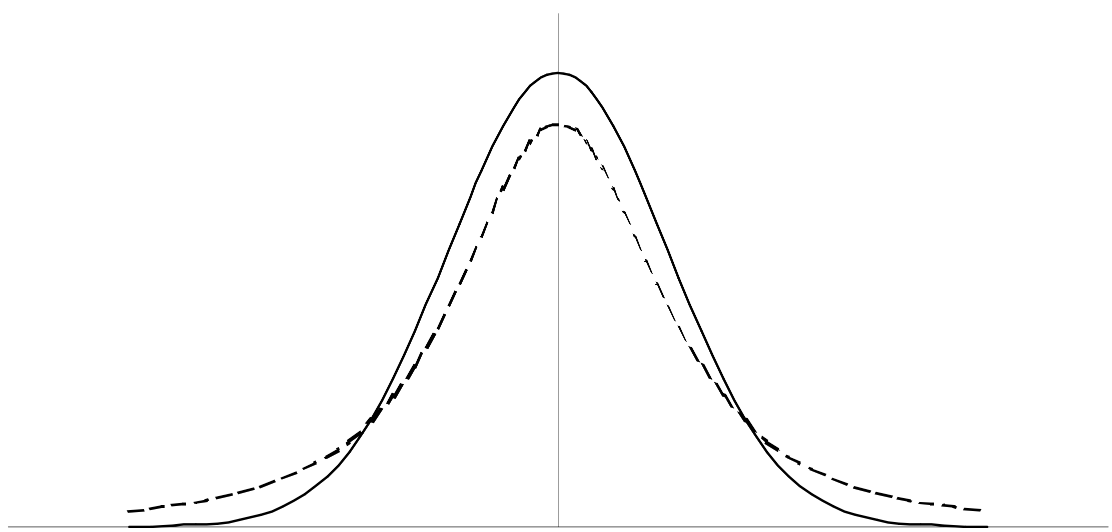


Figure 2.2: Fat-tailed (t with df=2) Distribution

Normal versus t (degrees of freedom = 2) Distributions There are three types of VaR now used by the industry:

- historical
- parametric
- factor-based

The historical VaR does not use the normal distribution but just form histograms from the past data. Historical VaR is model-free and completely dependent upon data. The 5% VaR is simply the best of the worst 5% of the data. For example, for 120 past observations, the 6-th observation from the worst is the 5% VaR.

The parametric VaR is based upon the normal distribution. Historical data are used to estimate mean and variance of the portfolio and then the normal distribution is used.

The factor-based VaR is the most comprehensive VaR of all. A linear factor model is used to estimate how various assets are related. Factor loadings are estimated and then a distribution is formulated.

2.3.2 Stress Test

According to wikipedia, regulators devise hypothetical future adverse economic scenarios to test banks. These established scenarios are then given to the banks in their jurisdiction and tests are run, under the close supervision of the regulator. They evaluate if the bank could endure the given adverse economic scenario, survive in business, and most importantly, continue to actively lend to households and business. If it is calculated that the bank can absorb the loss, and still meet the minimum bank capital requirements to remain in active business, they are deemed to have passed.

According to 2012 Stress Test Release by the Federal Reserve Bank on 23 February 2013, in the U.S. in 2012, an adverse scenario used in stress testing was all of the following:

- Unemployment at 13 percent
- 50 percent drop in equity prices
- 21 percent decline in housing prices.

A historical stress test if often performed to examine realistic (as opposed to hypothetically defined) stress scenarios. A historical stress test takes a very long history of data (minimally 10 years) and examine the worst losses. If the data history is not long enough, methods like benchmarking, extrapolations, indexing, etc. are used to estimate the worst losses. These worst losses are used as a guideline of how future potential losses can be. We shall discuss the details in the next chapter.

2.4 Credit Risk

Credit risk refers to losses occurred due to defaults. However, other derived credit risks such as spread risk and rating change risk must also be managed. These are discussed in details in a later chapter. Here, we first focus on two major sources of risk and various credit risk metrics.

2.4.1 Sources of credit risk

There are two major sources of credit risk banks need to manage well asset credit risk and counterparty credit risk. The first source refers to losses caused by defaults of the assets bank hold. Banks invest in various securities originated by various companies (e.g. corporate bonds), banks (swaps), and individuals (mortgage loans) which are all subject defaults. If these originators default, their securities will not have full values and hence banks suffer losses. Hence, each asset must be analyzed by the following two important credit metrics:

- loss given default (LGD)
- probability of default (PD)

An LGD is the amount of loss should a default occur and a PD is the likelihood of such loss occurring. The product of the two yields an expected loss (EL). Banks need to monitor its EL closely in order to keep its credit risk under control.

In the Appendix, PD term structures of various European nations over the crisis period are estimated and plotted to demonstrate how PDs skyrocketed during the crisis period. Not only were PDs shot higher, but the term structures also changed shapes during the crisis period.

The following diagram demonstrates empirically how in the past PD and LGD are correlated. This diagram is particularly important in that the highly positive correlation between PD and LGD reveals an understated risk when defaults happen. The Diagram indicates that as PD rises (firms are more likely to default) LGD rises as well (recoveries from defaults are little). As a result, EL is either very high when defaults are likely or very low when defaults are unlikely. We cannot estimate PD and LGD independently.

The second source of credit risk focuses on the loss due to counterparty default. Hence, securities that are subject to counterparty risks must be those that are transacted in the OTC (over the counter) market. When the counterparty defaults, naked (i.e. uncollateralized), in-the-money (i.e. receiving cash flows from the counterpary) positions will lose money and hence suffer losses. Most of such positions are swap positions. Swaps are the most popular contractual form in the OTC market. Popular swap contracts are:

- interest rate swap (IRS)
- foreign currency swap
- total return swap (TRS)

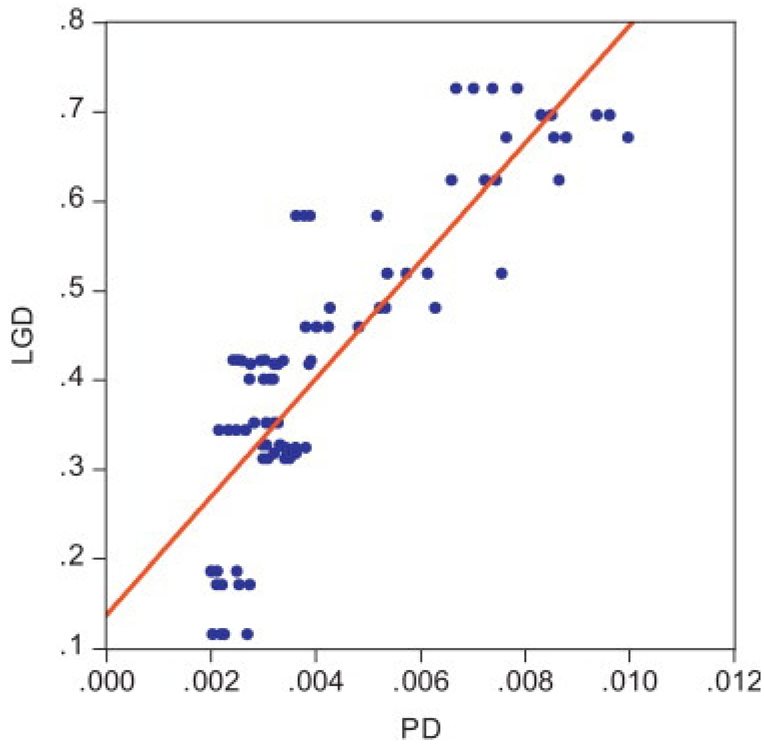


Figure 2.3: Negative Relation between PD and LGD

- credit default swap (CDS)

Before the 2008 crisis, counterparty credit risk was not quantified. Banks computed counterparty exposures but no effective management was in place. After the crisis, such risk has been quantified and incorporated into the cost of transacting, known as CVA (credit value adjustment). If a counterparty is credit riskier than the other counterparty, its CVA is higher and the cost of dealing this counterparty is higher. Traders who deal with this counterparty must be able to generate extra returns to offset the higher credit risk.

2.4.2 Credit risk metrics

The following diagrams depict the major credit risk metrics, which are:

- expected loss (EL)
- unexpected loss (UL)
- credit VaR
- expected shortfall (ES)
- economic capital (EC)
- jump to default or exposure at default (JTD/EAD)
- counterparty risk metrics
  - counterparty exposure
  - CVA (credit value adjustment)

Figure 16.2 depicts a typical loss distribution, usually highly positively skewed, and its related credit risk metrics. The expected loss (EL) is the mean value of this distribution, labeled by the left-most vertical bar in the Diagram. To find the mean value of a distribution one must carry out the convolution of loss function and probability density function. However, it is often approximated by PD times LGD as the Diagram demonstrates. The distribution also provides a credit VaR which is, similar to the market VaR, a tail critical value given a probability, labeled by the middle bar in the Diagram. Finally the unexpected loss (UL) is defined as the difference between the credit VaR (CVaR) and EL. Finally, expected shortfall (ES) is defined as the absolutely necessary capital to keep the company from default. Often, this is viewed as the worst-case loss (WCL), i.e. the worst tolerable loss which represents the case where all assets and counterparties default. Some risk management practices define ES and WCL differently, as there is no consensus over credit risk metrics. In other words, WCL can be viewed as JTD or EAD.

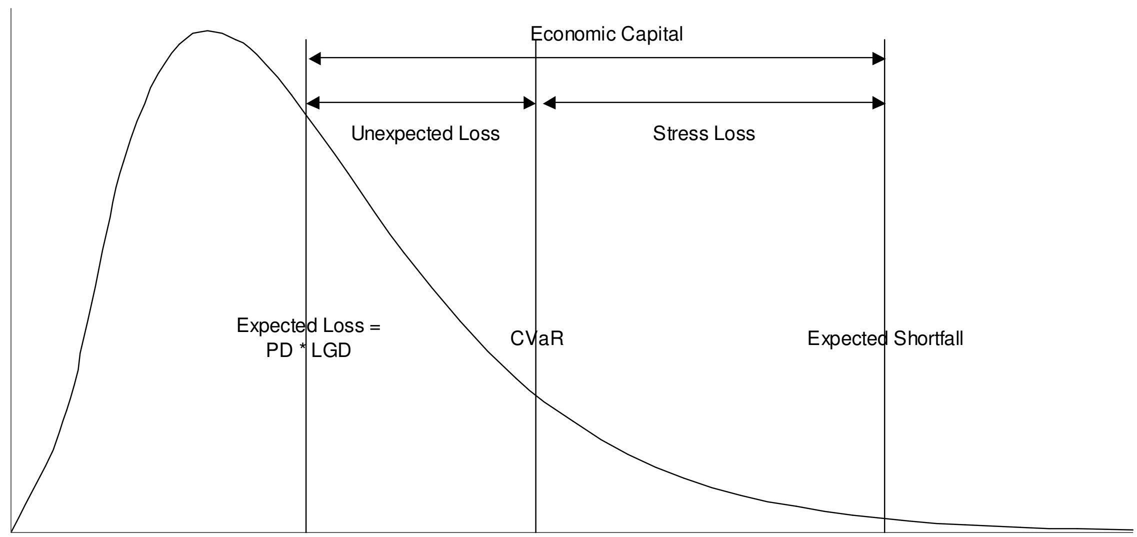


Figure 2.4: Credit Value at Risk

As far as counterparty risk goes, either a credit exposure is calculated and monitored, or a CVA charge is implemented and collected from each trading desk that is exposed to possible losses of counterparty default. These calculations will be discussed in details in the Credit Risk Chapters.

2.5 Liquidity Risk

The 2008 financial crisis is known as liquidity crisis and the call for liquidity quantification has been paramount. So far the models for measuring liquidity (not liquidity risk) are empirical and linear. The lack of theory (hence non-linearity) prevents the liquidity risk from being measured. In a later chapter, a model for liquidity quantification is presented. Liquidity is an old topic. It has been studied in accounting and market microstructure areas for a long time. Bank liquidity is closer to the accounting than to market microstructure which focuses on trading volume and bid-ask spreads. CPAs (certified public accountants) issue going concern audits to reveal their opinions if a firm can survive in a short run of not more than a year by looking at the firms short term liquidity. For a firm to survive in a short run, the only concern is the firms ability to meet its immediate cash flow obligations.

Such a liquidity-driven audit process ignores the economic nuance of the firm and can come to a different conclusion from an economically-driven default. In a crisis situation as the one we have been experiencing, such a point of view is more conservative as if a firm cannot survive the liquidity squeeze, the firm should default even though it is profitable. However, in a more normal situation where the liquidity squeeze is less eminent, such an audit is less conservative, which is against accountings Conservatism Principle. In a later chapter, we show that such a viewpoint has potentially a significant economic impact on the value of the firm. An in-depth case analysis shows that a firm that is subject to economic default and yet passes the myopic going concern audit will ultimately default and then results in a greater loss of economic value.

2.6 Operational Risk (Wiki)

Operational risk is the broad discipline focusing on the risks arising from the people, systems and processes through which a company operates. It can also include other classes of risk, such as fraud, legal risks, physical or environmental risks. A widely used definition of operational risk is the one contained in the Basel II regulations. This definition states that operational risk is the risk of loss resulting from inadequate or failed internal processes, people and systems, or from external events.[1] Operational risk management differs from other types of risk, because it is not used to generate profit (e.g. credit risk is exploited by lending institutions to create profit, market risk is exploited by traders and fund managers, and insurance risk is exploited by insurers). They all however manage operational risk to keep losses within their risk appetite - the amount of risk they are prepared to accept in pursuit of their objectives. What this means in practical terms is that organizations accept that their people, processes and systems are imperfect, and that losses will arise from errors and ineffective operations. The size of the loss they are prepared to accept, because the cost of correcting the errors or improving the systems is disproportionate to the benefit they will receive, determines their appetite for operational risk. The Basel II Committee defines operational risk as: The risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. However, the Basel Committee recognizes that operational risk is a term that has a variety of meanings and therefore, for internal purposes, banks are permitted to adopt their own definitions of operational risk, provided that the minimum elements in the Committee's definition are included. Basel II and various Supervisory bodies of the countries have prescribed various soundness standards for Operational Risk Management for Banks and similar Financial Institutions. To complement these standards, Basel II has given guidance to 3 broad methods of Capital calculation for Operational Risk:

- Basic Indicator Approach - based on annual revenue of the Financial Institution
- Standardized Approach - based on annual revenue of each of the broad business lines of the Financial Institution
- Advanced Measurement Approaches - based on the internally developed risk measurement framework of the bank adhering to the standards prescribed (methods include IMA, LDA, Scenario-based, Scorecard etc.)

The Operational Risk Management framework should include identification, measurement, monitoring, reporting, control and mitigation frameworks for Operational Risk.

2.7 Risk Management Modeling Building Blocks

To build a VaR model for a portfolio of various assets that carry very different risks is a big challenge. The first VaR model by JP Morgan (known as Riskmetrics) proposed cash flow mapping. That is, different assets are aggregated via their cash flows. Then the total VaR can be calculated using the aggregated cash flows. More recently, a delta method is used. A delta is the partial derivative of an asset with respective to a target risk factor. To use deltas, various assets must be priced using a consistent set of pricing models. Hence, the delta method is not possible unless all assets can be priced consistently. The advances in numerical methods (lattice and Monte Carlo) and computing powers now make the delta method successful. Using the delta method, positions can be aggregated and the total VaR can be computed. Moreover, using deltas can provide the important VaR decomposition an important concept of incremental VaR. Given that VaR is just a standard deviation, VaRs cannot be added. However, incremental VaRs can. In the next chapter, the details will be discussed.

2.7.1 Basic Models by Asset Class

To gain an integrated measure of all financial risks, properly evaluating various financial products is essential. In other words, valuation models must be employed for various asset classes:

- Equity – Black-Scholes/binomial, CAPM, local vol (implied binomial model)
- IR – Heath-Jarrow-Morton, Hull-White

- FX – Garman-Kolhegen (i.e. Black-Scholes)
- Commodities – Black, seasonality
- Mortgages (prepayment) – Andrew-Davidson
- ABS – loss timing function
- Credit – Jarrow-Turnbull, Duffie-Singleton, transition matrix, ad-hoc approaches

2.7.2 Risk Management Tools available

The production of risk numbers is a highly technical task. Hence, these risk numbers are usually produced by highly trained professionals. Large financial institutions can train internal personnel to perform the task. Smaller institutions can only buy standardized products off the shelf. In the following, we can see some popular products to choose.

- Oldest – Riskmetrics and Creditmetrics
- Enterprise – IBM, Oracle, SAP, etc.
- Valuation – Barra, Algo, etc.
- Consulting – Big 3, McKinsey, etc.
- Proprietary – large banks

2.8 Basel Accords

The Basel Committee on Banking Supervision is an international committee established by the Bank for International Settlements to formulate policy on prudential standards and best practices among financial regulators. The Basel Committee implemented the first Basel Capital Accord in 1988. Originally developed for internationally active banks in G10 countries, the Accord has now been implemented in over 100 countries for both large and small financial institutions, including credit unions. In 2000, the Basel Committee began consulting the financial services industry on a revision to the Basel Capital Accord. The purpose of the revision was to provide a more risk sensitive approach to capital adequacy. On June 26, 2004, the Basel Committee on Banking Supervision issued its revised framework, International Convergence of Capital Measurement and Capital Standards, a Revised Framework. This framework reflects the committees modifications, albeit limited in number, made in the Consultative Paper (CP3) issued in April 2003. In September 2011, the Basel Committee revised its capital standards in what is referred to as Basel III.

2.8.1 Basel I

The bank must maintain capital (Tier 1 and Tier 2) equal to at least 8% of its risk-weighted assets.

- 0% - cash, central bank and government debt and any OECD government debt
- 0%, 10%, 20% or 50% - public sector debt
- 20% - development bank debt, OECD bank debt, OECD securities firm debt, non-OECD bank debt (under one year maturity) and non-OECD public sector debt, cash in collection
- 50% - residential mortgages
- 100% - private sector debt, non-OECD bank debt (maturity over a year), real estate, plant and equipment, capital instruments issued at other banks OECD: Organisation for Economic Co-operation and Development

Tier 1 capital

Tier 1 capital is the core measure of a bank's financial strength from a regulator's point of view. It is composed of core capital, which consists primarily of

- common stock and
- disclosed reserves (or retained earnings), but may also include
- non-redeemable non-cumulative preferred stock.

The Basel Committee also observed that banks have used innovative instruments over the years to generate Tier 1 capital; these are subject to stringent conditions and are limited to a maximum of 15% of total Tier 1 capital.

Tier 2 capital (supplementary capital)

Tier 2 capital includes a number of important and legitimate constituents of a bank's capital base. These forms of banking capital were largely standardized in the Basel I accord, issued by the Basel Committee on Banking Supervision and left untouched by the Basel II accord. National regulators of most countries around the world have implemented these standards in local legislation. In the calculation of regulatory capital, Tier 2 is limited to 100% of Tier 1 capital.

- Undisclosed reserves
- Revaluation reserves
- General provisions/general loan-loss reserves
- Hybrid debt capital instruments
- Subordinated term debt

Tier 3 capital

Tertiary capital held by banks to meet part of their market risks, that includes a greater variety of debt than tier 1 and tier 2 capitals. Tier 3 capital debts may include a greater number of subordinated issues, undisclosed reserves and general loss reserves compared to tier 2 capital. Tier 3 capital is used to support market risk, commodities risk and foreign currency risk.

- Banks will be entitled to use Tier 3 capital solely to support market risks.
- Tier 3 capital will be limited to 250% of a banks Tier 1 capital and have a minimum maturity of two years.
- This means that a minimum of about 28% of market risks needs to be supported by Tier 1 capital that is not required to support risks in the remainder of the book;

  http://www.bis.org/publ/bcbs128b.pdf

Capital Ratios

Capital ratios are various percentage measures that measure the financial health of banks. The use of ratios makes regulation a lot easier and ratios can be applied to different sizes of banks. The important capital ratios are given below.

- Tier 1 capital ratio = Tier 1 capital / Risk-adjusted assets $\geq 6\%$
- Capital adequacy ratios (or total capital ratio) are a measure of the amount of a bank's core capital expressed as a percentage of its risk-weighted asset. Capital adequacy ratio is defined as: Total capital (Tier 1 and Tier 2) / Risk-adjusted assets $\geq 10\%$
- Leverage ratio = Tier 1 capital / Average total consolidated assets $\geq 5\%$
- Common stockholders equity ratio = Common stockholders equity / Balance sheet assets

The following is the balance sheet of Lehman Brothers Inc. as of 2002 .

Lehman Brothers Inc.				
as of 2002				
Assets			Liabilities	
Cash	2,265		Short-term Debt	123
Securities	70,881		Other Securities	50,352
Coll Ag'mt	101,149		Coll ST Financing	121,844
Receivables	21,191		Payables	12,758
Real Estate	138		Long-Term Debt	7,990
			Equity	3,152
Total	196,219		Total	196,219
million $				

Figure 2.5: Balance Sheet of Lehman Brothers

The following presents the balance sheet of Goldman from 2009   2011 (taken from Yahoo Finance). The equity ratio is 7.58%.

2.8.2 Basel II

Basel II is the second of the Basel Accords, (now extended and effectively superseded by Basel III), which are recommendations on banking laws and regulations issued by the Basel Committee on Banking Supervision. Politically, it was difficult to implement Basel II in the regulatory environment prior to 2008, and progress was generally slow until that year's major banking crisis caused mostly by credit default swaps, mortgage-backed security markets and similar derivatives. As Basel III was

Balance Sheet

Get Balance Sheet for: GO

View: **Annual Data** | Quarterly Data

All numbers in thousands

Period Ending	**Dec 30, 2011**	**Dec 30, 2010**	**Dec 30, 2009**
Assets			
Current Assets			
Cash And Cash Equivalents	308,061,000	281,874,000	219,233,000
Short Term Investments	-	-	-
Net Receivables	74,465,000	78,140,000	67,900,000
Inventory	-	-	-
Other Current Assets	-	-	-
Total Current Assets	-	-	-
Long Term Investments	517,547,000	523,259,000	532,341,000
Property Plant and Equipment	-	-	-
Goodwill	-	-	-
Intangible Assets	-	-	-
Accumulated Amortization	-	-	-
Other Assets	23,152,000	28,059,000	29,468,000
Deferred Long Term Asset Charges	-	-	-
Total Assets	**923,225,000**	**911,332,000**	**848,942,000**
Liabilities			
Current Liabilities			
Accounts Payable	682,943,000	659,208,000	624,064,000
Short/Current Long Term Debt	213,540,000	210,187,000	165,876,000
Other Current Liabilities	46,109,000	38,569,000	39,418,000
Total Current Liabilities	-	-	-
Long Term Debt	210,909,000	212,776,000	209,219,000
Other Liabilities	38,983,000	41,223,000	49,062,000
Deferred Long Term Liability Charges	-	-	-
Minority Interest	-	-	-
Negative Goodwill	-	-	-
Total Liabilities	**852,846,000**	**833,976,000**	**778,228,000**
Stockholders' Equity			
Misc Stocks Options Warrants	-	-	-
Redeemable Preferred Stock	-	-	-
Preferred Stock	3,100,000	6,957,000	6,957,000
Common Stock	8,000	8,000	8,000
Retained Earnings	58,834,000	57,163,000	50,252,000
Treasury Stock	(42,281,000)	(36,295,000)	(32,156,000)
Capital Surplus	45,553,000	42,103,000	39,770,000
Other Stockholder Equity	5,165,000	7,420,000	5,883,000
Total Stockholder Equity	**70,379,000**	**77,356,000**	**70,714,000**
Net Tangible Assets	**70,379,000**	**77,356,000**	**70,714,000**

Figure 2.6: Balance Sheet of Goldman Sachs

negotiated, this was top of mind, and accordingly much more stringent standards were contemplated, and quickly adopted in some key countries including the USA. The main message of Basel II is the following three pillars:

Pillar 1: Capital Adequacy

Min. of 8% (but now credit, market & operational)

Pillar 2: Supervisory Review

Supervisors responsible for ensuring banks have sound internal processes to assess capital adequacy

Pillar 3: Market Discipline

Enhanced disclosure by banks

Sets out disclosure requirements

2.8.3 Basel III

Basel III is a comprehensive set of reform measures, developed by the Basel Committee on Banking Supervision, to strengthen the regulation, supervision and risk management of the banking sector. These measures aim to:

- improve the banking sector's ability to absorb shocks arising from financial and economic stress, whatever the source
- improve risk management and governance
- strengthen banks' transparency and disclosures.

The main message of Basel III is:

- Increased overall capital requirement: Between 2013 and 2019, the common equity component of capital (core Tier 1) will increase from 2% of a banks risk-weighted assets before certain regulatory deductions to 4.5% after such deductions. A new 2.5% capital conservation buffer will be introduced, as well as a zero to 2.5% countercyclical capital buffer. The overall capital requirement (Tier 1 and Tier 2) will increase from 8% to 10.5% over the same period.
- Narrower definition of regulatory capital: Common equity will continue to qualify as core Tier 1 capital, but other hybrid capital instruments (upper Tier 1 and Tier 2) will be replaced by instruments that are more loss-absorbing and do not have incentives to redeem. Distinctions between upper and lower Tier 2 instruments, and all of Tier 3 instruments, will be abolished. All

non-qualifying instruments issued on or after 12 September 2010, and non-qualifying core Tier 1 instruments issued prior to that date, will both be derecognised in full from 1 January 2013; other non-qualifying instruments issued prior to 12 September 2010 will generally be phased out 10% per year from 2013 to 2023.

- Increased capital charges: Commencing 31 December 2010, re-securitisation exposures and certain liquidity commitments held in the banking book will require more capital. In the trading book, commencing 31 December 2010, banks will be subject to new stressed value-at-risk models, increased counterparty risk charges, more restricted netting of offsetting positions, increased charges for exposures to other financial institutions and increased charges for securitisation exposures.
- New leverage ratio: A minimum 3% Tier 1 leverage ratio, measured against a banks gross (and not risk-weighted) balance sheet, will be trialled until 2018 and adopted in 2019.
- Two new liquidity ratios: A liquidity coverage ratio requiring high-quality liquid assets to equal or exceed highly-stressed one-month cash outflows will be adopted from 2015. A net stable funding ratio requiring available stable funding to equal or exceed required stable funding over a one-year period will be adopted from 2018.

2.9 Types of Capital

2.9.1 Regulatory capital

The Basel Committee on Banking Supervision (BCBS), on which the United States serves as a participating member, developed international regulatory capital standards through a number of capital accords and related publications, which have collectively been in effect since 1988. In July 2013, the Federal Reserve Board finalized a rule to implement Basel III in the United States, a package of regulatory reforms developed by the BCBS. The comprehensive reform package is designed to help ensure that banks maintain strong capital positions that will enable them to continue lending to creditworthy households and businesses even after unforeseen losses and during severe economic downturns. This final rule increases both the quantity and quality of capital held by U.S. banking organizations. The Board also published the Community Banking Organization Reference Guide, which is intended to help small, non-complex banking organizations navigate the final rule and identify the changes most relevant to them. The capital ratio is the percentage of a bank's capital to its risk-weighted assets. Weights are defined by risk-sensitivity ratios whose calculation is dictated under the relevant Accord. Basel II requires that the total capital ratio must be no lower than 8%. Under the Basel II guidelines, banks are allowed to use their own estimated risk parameters for the purpose of calculating regulatory capital. This is known as the Internal Ratings-Based (IRB) Approach to capital requirements for credit risk. Only banks meeting certain minimum conditions, disclosure requirements and approval from their national supervisor are allowed to use this approach in estimating capital for various exposures.

2.9.2 Economic capital

As opposed to regulatory capital that is required by the governments or any regulatory body, economic capital the firm needs to ensure that its realistic balance sheet stays solvent over a certain time period with a pre-specified probability. Therefore, economic capital is often calculated as value at risk. The balance sheet, in this case, would be prepared showing market value (rather than book value) of assets and liabilities. Economic capital is the amount of risk capital which a firm requires to cover the risks that it is running or collecting as a going concern, such as market risk, credit risk, and operational risk. Firms and financial services regulators should then aim to hold risk capital of an amount equal at least to economic capital. It is the amount of money which is needed to secure survival in a worst case scenario. Firms and financial services regulators should then aim to hold risk capital of an amount equal at least to economic capital.

2.9.3 Risk-adjusted return on capital (RAROC)

An adjustment to the return on an investment that accounts for the element of risk. Risk-adjusted return on capital (RAROC) gives decision makers the ability to compare the returns on several different projects with varying risk levels. RAROC was popularized by Bankers Trust in the 1980s as an adjustment to simple return on capital (ROC). While there are many definitions of how RAROC should be calculated, one example provided by investopedia is: RAROC = (Revenue Expenses Expected Loss + Income from Capital) / Capital where Income from Capital = (Capital Charges) (risk-free rate) and Expected loss = average anticipated loss over the measurement period.

2.10 Appendix

2.10.1 Report on Excite@Homes bankruptcy

http://everything2.com/title/excite Excite (by then, Excite@Home) filed for bankruptcy on September 28, 2001. But let's not start there, let's start with 1999, when Excite was second only to Yahoo as a pure web portal. Management calculated that they could distinguish themselves by owning their own broadband pipes, and thereby surpass Yahoo and maybe even AOL in a few years. In a $6.7 billion deal, Excite merged with @home, the leading provider of cable internet services. They established contracts with large cable companies making Excite the 'broadband portal' for most Americans with cable modems in 2000 and 2001. However, things were turning ugly, as cash coming in from advertising contracts did not equal cash flowing out to employees and vendors. The company spent much of 2001 attempting to raise cash, selling divisions, selling convertible bonds (they raised $100 million from Promethean Capital this way), and closing unprofitable units. Unfortunately, they could not cut costs quickly enough to match the dropoff in advertising dollars. Even after firing half the staff, they were stuck with leases on the buildings those people had worked from. From 2000q2 to 2001q2, ad revenue dropped more than 60%. Lots of people were signing up for broadband, but the @home side of the business was too small to cover for the hemorrhaging Excite portal. The company needed to get cash from the outside, and they paid for it. Promethean Capital demanded immediate repayment on the convertible bonds only a few months after buying them, claiming that Excite@Home had substantially misrepresented the state of their business. During the firesale, they sold Blue Mountain Arts, their online gift card unit, for $35 million to Gibson Greetings. Keep in mind that they had paid $780 million for BMA in late 1999. The fortunes of the company were closely tied to AT&T, and they acted as AT&T's broadband internet provider. This allowed them to negotiate an $85 million investment from AT&T in June 2001, but that was peanuts compared to the company's burn rate, even after thousands of people had been fired. In August 2001, they fired their auditors, got a $50 million demand from Promethean, and had both Cox Cable and Comcast drop their service. Bankruptcy came the next month. If you visit Excite.com today, you will see a website that looks like the old Excite portal, but it is an unrelated service that purchased the URL and logo from Excite@Home after the bankruptcy filing. Business 2.0, August 28, 2001, USA Today, October 1, 2001, Various company press releases

Chapter 3

Hedging

3.1 What Is Hedging?

Hedging is the ultimate form of risk management. Hedging can be regarded as risk elimination. In other words, a perfect hedging should completely eliminate risk and results in a risk-free portfolio. While hedging sounds pleasant, it is neither practical, nor necessary. In reality, even if perfect hedging did exist, nobody would want it, as no risk usually implies no return. If portfolio theory has taught us anything, it would mean that we should maximize return at minimal risk.

Since hedging is risk elimination, it has become a perfect tool to develop pricing models. In an efficient market, the theory of hedging has been translated into pricing models. In other words, the cost of perfect hedging must be the price of the target security; otherwise arbitrage can take place and the efficient market is violated. Hence, by contradiction (which cannot exist), the cost of hedging must equal price of the target security. The most famous example is the Black-Scholes model developed in 1973.

Derivatives, namely options, futures (forwards), and swaps, are effective instruments to eliminate risk. More interestingly, if used effectively, they can be used to remove only the unwanted risk. This is known as the "static hedging". Since derivatives have maturity dates, once they expire, they need to be rolled over. Hence they cannot be used for long term hedging purposes. For long term hedging needs, one can pursue "dynamic hedging" strategies. A dynamic hedging strategy requires constant trading and re-balancing the portfolio. The hedging is perfect only for a very short period of time (a day, an hour, or a few minutes) and then it must be redone in order to remain perfectly hedged. As a result, dynamic hedging can be implemented only for securities that are very liquid. Over-the-counter (OTC) securities are generally not liquid and consequently are not good hedging securities. Since dynamic hedging requires very frequent trading, often it is automated and executed by computer programs, known as program trading.

In the following, we use simple examples to demonstrate static and dynamic hedging. Most of these materials can be found in options and futures text books.

3.2 Static versus Dynamic Hedging

3.2.1 Static Hedging

A static hedging strategy is a buy-and-hold strategy. A strategy is chosen, hedging security bought, and no action is necessary until the contract expires. The perfect example of such kind is the 'protective put". A protective put protects its underlying stock. Because a put and its underlying stock move in perfectly opposite directions, it provides a perfect hedge for the stock. By choosing the strike price of the put, one can set the protection level (i.e. minimum value) of the portfolio. Figure 3.1 describes graphically how a protective put portfolio pays out at the expiration date of the put. The positively sloped 45-degree dotted line is the underlying stock and the negative 45-degree dotted line is the protecting put. Together, they form a solid hockey stick function whose value will never fall below the strike price of the put.

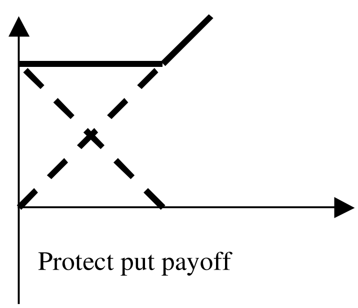


Figure 3.1: Protective Put

Another famous static hedging example is known as the "covered call". It is quite common that investors write (sell) naked call options trying to harvest the premiums. If the stock falls in value, then the call options will expire worthlessly and those who sell the calls can keep the premiums. However, such a strategy is highly risky. Should the stock price rise, the loss is theoretically unlimited. To hedge this risk, investors often buy the underlying stock. As the stock can be used for delivery for the call option when its price rises, it is a perfect hedge. Now the option is the target security and the stock is the hedging security. Figure 3.2 describes the payoff at the expiration date of the call option.

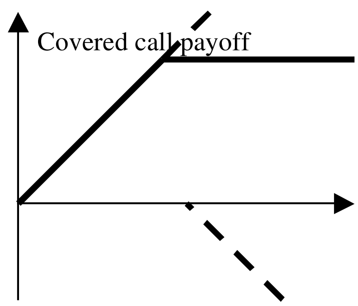


Figure 3.2: Covered Call

Buying the stock is expensive. An alternative is to buy another call option with a lower strike price. Since call prices are lower than stock prices, doing so can save cost of hedging. It goes without saying that this cheaper hedging is not as ideal as the hedging with stock. Figure 3.3 describes the result. As we can easily see that the payoff of this strategy (bull spread) is lower than the payoff of the covered call.

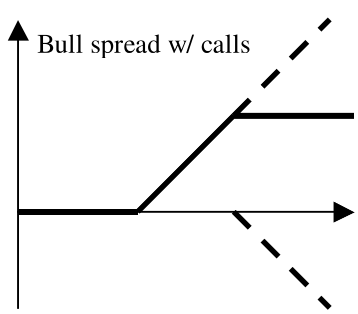


Figure 3.3: Bull Spread

3.2.2 Dynamic Hedging

A dynamic strategy, as mentioned, requires constant trading and re-balancing the portfolio. A perfect example would be the covered call given earlier. In a static strategy, buying the stock can provide an effective hedge but it is very costly. The reason why it is costly is that it keeps a very high payoff in the future (at the expiration of the option). If we do not need such a high payoff, then the covered call strategy is not desirable. One can certainly go with an alternative which is a bull spread to save cost (and give up high payoff). A really effective way is to do dynamic hedging using stock. This was first discovered by Fischer Black and Myron Scholes in 1973.

Assume the stock price follows a stochastic process (see Chapter 1) as follows:

$$\begin{aligned} &\frac{dS}{S} = \mu dt + \sigma dW \\ &\text{or} \\ &dS = \mu S dt + \sigma S dW \end{aligned} \tag{3.1}$$

This is a very simple process for the stock. What it says is that the stock has an expected return μ and volatility σ. dW is a Wiener process which is normally distributed with mean 0 and variance dt. Since the option is a derivative contract written on the stock, it must be a function of the stock price. Hence, we can measure the sensitivity of the option price change with respect to the stock price change:

$$\frac{\partial C}{\partial S}$$

For example, if $\partial C/\partial S$ is 0.25, which means any increase (or decrease) of the stock by \$1 will result in a 25 cents increase (or decrease) in the option (or vice versa that every \$1 change in option will result in \$4 change in the stock), then we can provide a very effective hedge by just doing a 1:4 hedge ratio. In other words, to protect a call option, we need to buy a quarter share of the stock. Of course the problem is that the hedge ratio keeps changing as the stock price keeps changing. Consequently stock shares must be bought and sold constantly to match the hedge ratio. As a result, such hedging requires constant trading.[1]

As we can easily see, there is nothing that prevents us from using the same methodology on other instruments. Indeed, such a method is proven to be universal. As long as there is an easy way to calculate $\partial X/\partial Y$, one can effectively hedge X with Y or vice versa. In the above case, Black-Scholes has derived a closed-form formula for $\partial C/\partial S$, which is "$N(d_1)$" (details to be shown later) which is a standard normal probability function.

[1]Note that the theory behind the strategy is actually quite profound. Black and Scholes first showed that such strategy is "self-financing". That is, such a strategy does not incur any cost. Later, this is proven to follow the Martingale Representation Theorem.

Remarks

There are three remarks. First, such hedging is prefect, which it eliminates risk completely. If it eliminates all risk, then the portfolio is risk-free and should earn the risk-free rate. If the portfolio requires no initial investment, then the portfolio should not produce any return. This "no free lunch" or "arbitrage free" argument is how pricing models are derived. In the case of the covered call, we can then derive the pricing model for the call option. We can equally apply the same hedging method for protective put and it will result in a put option model.

Second, if the risk is completely eliminated, then there is no purpose of doing any investment. One could easily just invest in the risk-free asset such as Treasury bills. As a result, the use of dynamic hedging is mostly for deriving pricing models, as opposed to finding the optimal hedging. Optimal hedging requires risk-return tradeoff which is not the purpose of dynamic hedging and hence is a different topic.

Third, when the market isn't perfect, that is, the dynamic hedging cannot fully eliminate risk. In this situation, the dynamic hedging becomes a useful way to eliminate the so-called "delta" risk. In other words, the dynamic hedging gives a "delta-neutral" portfolio. When the market isn't perfect, frictions in the market produce what is known as the "gamma risk". Without removing the delta risk, investors cannot see clearly how much the gamma risk is. By removing the delta risk (via being delta-neutral), investors can now see clearly and then manage the gamma risk. Gamma risk cannot be hedged and often it produces attractive returns. Hence, Wall Street traders make money by taking gamma risk and managing it well.

3.2.3 Other Hedging Examples

Swaps/Forwards

Swaps are a very effective tool to hedge risk. The most commonly used hedging strategy is in the area of interest rate risk. A standard vanilla swap is a contract that exchanges a set of fixed cash flows with a set of floating cash flows. These two legs of cash flows mimic fixed rate bond and floating rate bond. Given that floating rate bonds have almost no interest rate risk (i.e. duration), one can swap out interest rate risk by engaging in a fixed-floating interest rate swap.

Forwards are perfect hedges of the underlying assets. A forward contract can be regarded as a long call and a short put. Previously we see how each can perform a partial hedge. But together they become a forward contract which can guarantee a fixed cash flow (strike price) to the investor. In reality, there are very few liquid forward markets but one can create one with calls and puts.[2]

Futures

Futures contracts can be effective in providing effective hedging. One application is known as the "minimum variance hedging". The idea is to use futures to hedge the underlying. Often it is done with buying the underlying and shorting the futures.[3] Let the hedge ratio be h and hence the portfolio value is:

$$V = S - hF$$

Then the change of the portfolio value is:

$$dV = dS - hdF$$

Given that such a hedge is not perfect, there is variance of the portfolio, which can be easily calculated as:

$$\mathbb{V}[dV] = \sigma_S^2 + h^2\sigma_F^2 - 2h\rho\sigma_S\sigma_F$$

The goal here is to minimize this portfolio variance with the best choice of the hedge ratio h. To achieve that, we simply take the derivative of the equation with respect to the hedge ratio and set it to 0:

$$2h\sigma_F^2 - 2\rho\sigma_S\sigma_F = 0$$

which implies that:

$$\hat{h} = \rho\frac{\sigma_S}{\sigma_F}$$

It turns out that this result is identical to the regression coefficient of dS running against dF.

[2]Foreign exchange and short term interest rate forward markets are liquid.
[3]For commodity futures, it is difficult to short the underlying.

Part II

Market Risk

Chapter 4

Value At Risk

4.1 Introduction

Market risk is defined as the potential loss resulting from declining prices in the financial market. There are two dimensions to have a good overall grasp of the market risk risk factors and drivers risk exposures.

Market risk consists of the following broad stochastic market risk factors (also can be regarded as asset classes):

- interest rates
- prepayment speeds (mortgage rates)
- credit spreads
- FX rates
- commodities
- equities

The two drivers of the market risk exposures are:

- investment position and
- volatility (including variances and co-variances)

It of great importance that we understand how to manage the size of investment in each asset class (and hence the exposure to each risk factor) and the volatility of each risk factor. It is of equal importance that we understand how the risk factors move over time and learn how to predict their movements. It is an on-going battle that investors try to beat the markets while keep the risk under control.

Figure 4.10 is a simple organization chart that demonstrates how market risk is estimated.

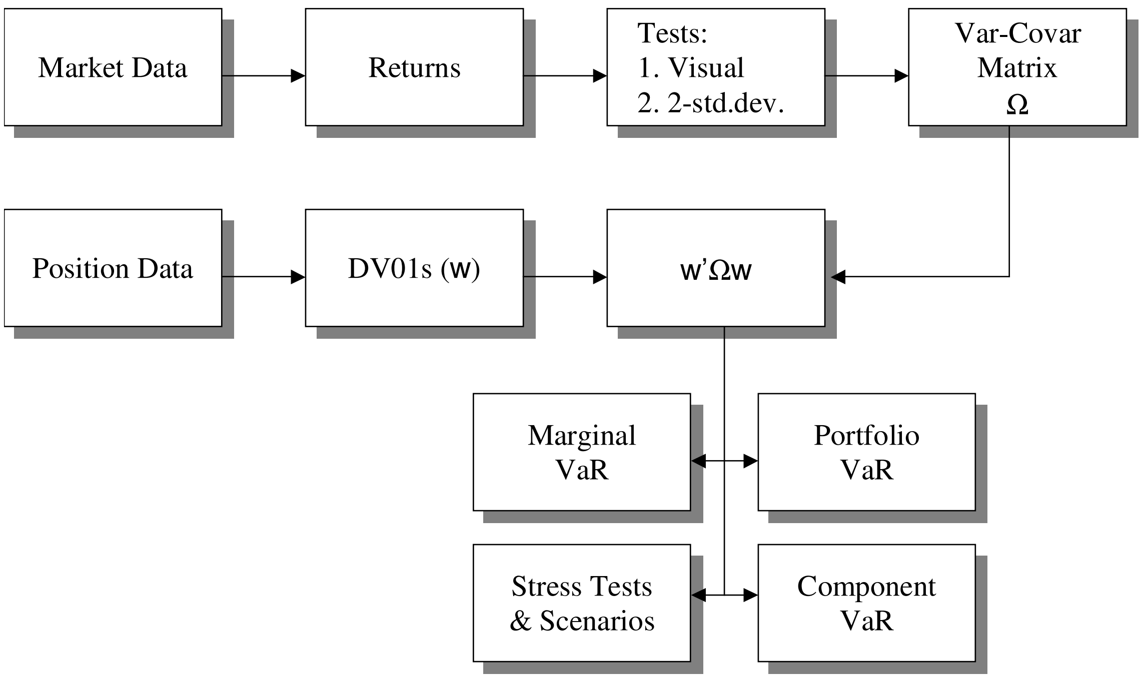


Figure 4.1: Risk Management Flow Chart

As we can see from the diagram, the output of this system is a set of market risk measures, known as VaR (value at risk) lying at the bottom of the diagram. There are three types of VaR. The portfolio VaR (far right) is the final aggregated risk number to be presented to the CEO of the company. Then, there is Absolute sub-portfolio VaR which is a number of VaR numbers of divisions of the company. As we shall demonstrate in details later, VaRs which are volatility based are not additive. Hence, the summation of all Absolute sub-portfolio VaRs is not the portfolio VaR. In order to force additivity, a series of Marginal sub-portfolio VaRs are created. These VaRs demonstrate how much each division of the company contributes to the total portfolio VaR.

Market Risk Management Toolbox:

- Market Risk Factors Identification
- Sensitivity Analysis
- VaR for Market Risk: Riskmetrics (JP Morgan) 94
- Stress Testing and Scenario Analysis
- Economic Capital Adequacy Level

Relative VaR

- Report reveals important differences between VaR and Relative VaR
- Global Equities has the stand-alone VaR (11%), but relative to its benchmark, the smallest relative VaR (1%)
- Global FI Portfolio has smallest stand-alone VaR (5%) but highest relative VaR (4%)

Marginal VaR measures how much risk a position adds to a portfolio or how much Portfolio VaR would change if a position is removed. Often the largest stand-alone risk positions are not the greatest contributors of risk hedges have a negative marginal VaR. Identifies which position to eliminate entirely in order is most effective to reduce risk.

Component VaR is closely related to Marginal VaR. The sum of all component VaRs add up to total Portfolio VaR. We should note that component VaR is like to increase position weight by $1 and measure the change in overall portfolio VaR multiply this change by position weighting.

4.1.1 An Example

Table 4.1 is a sample risk report. Note that Marginal VaRs add up to $946,078, not equal to the portfolio VaR. Although Asia ex Japan has the largest stand-alone VaR, it is the fourth largest contributor to risk, as measured by incremental VaR. The best 3 opportunities for reducing risk through hedges lie in US, Latam & Europe.

Diversification benefit = Overall Portfolio VaR – Sum of individual VaRs

Region	Market Value	Individual VaR	Marginal VaR	Component VaR	% Contribution
US	71.77	574,194	222,075	378,341	25.17%
S. Amer	10.26	512,944	220,114	369,626	24.59%
Europe	64.60	581,404	204,358	343,237	22.84%
Asia ex. Japan	12.69	589,734	196,046	317,346	21.11%
E. Europe	1.95	116,932	31,050	40,322	2.68%
Japan	19.57	195,694	48,012	30,068	2.00%
Africa	4.67	93,387	24,423	24,163	1.61%
Diversification		(1,161,186)			
Total	185.51	1,503,103	946,078	1,503,103	100.00%

Table 4.1: A Sample VaR Report

4.1.2 Limitations of VaR

Although VaR is the most widely used measure for market risk, it is far from perfect and presents many problems. First of all, VaRs are not additive. VaRs are transformations of standard deviations that are not additive. A fundamental assumption that future risk can be predicted from historical distribution is not realistic. Furthermore, it is vulnerable to regime shifts & sudden changes in market behavior. Finally, VaRs calculated by different methods have different limitations.

VaR gets me to 95% confidence. I pay my Risk Managers good salaries to look after the remaining 5% Dennis Weatherstone, former CEO, JP Morgan.

4.1.3 Stress Testing

Stress testing is designed to estimate potential economic losses in abnormal markets. It measures extreme market movements occur far more frequently than a normal distribution suggest. It provides a more comprehensive picture of risk. Table 4.2 presents a sample of what stress scenarios are.

4.1.4 Risk Reporting

An efficient risk reporting process is the final and the most important step of risk management. An efficient risk reporting process is the foundation for clear and

Standard Stress	Test Scenario
Implied Volatility	20%
Equity Index	10%
Yield Curve Twist	25 bps
Parallel Yield Curve Shift	100bps

Table 4.2: Stress Test Scenarios

timely communication of risk across enterprise (Corporate, Business Unit, Trading Desk). What makes a good risk report?

- Timely
- Reasonably accurate
- Highlight portfolio risk concentrations
- Include a written commentary
- Be concise

4.1.5 External Disclosure

From Chase Manhattans Annual Report: Chase conducts daily VaR backtesting for both regulatory compliance with the Basle Committee on Banking Supervision market risk capital rules and for internal evaluation of VaR against trading revenue.

During the year, a daily trading loss exceeded that years trading VaR on 2 days. This compares to an expected number of approximately 3 days. Considering the unsettled markets during the year, Chase believes its VaR model performed at a very high level of accuracy during the year.

4.1.6 The Basel Accord II

- Pillar 1: Capital Adequacy
  - Min. of 8% (but now credit, market & operational)
- Pillar 2: Supervisory Review

  - Supervisors responsible for ensuring banks have sound internal processes to assess capital adequacy

- Pillar 3: Market Discipline
  - Enhanced disclosure by banks
  - Sets out disclosure requirements

4.1.7 Capital Regulation

Figure 4.2 depicts the highlight of the capital adequacy ratio.

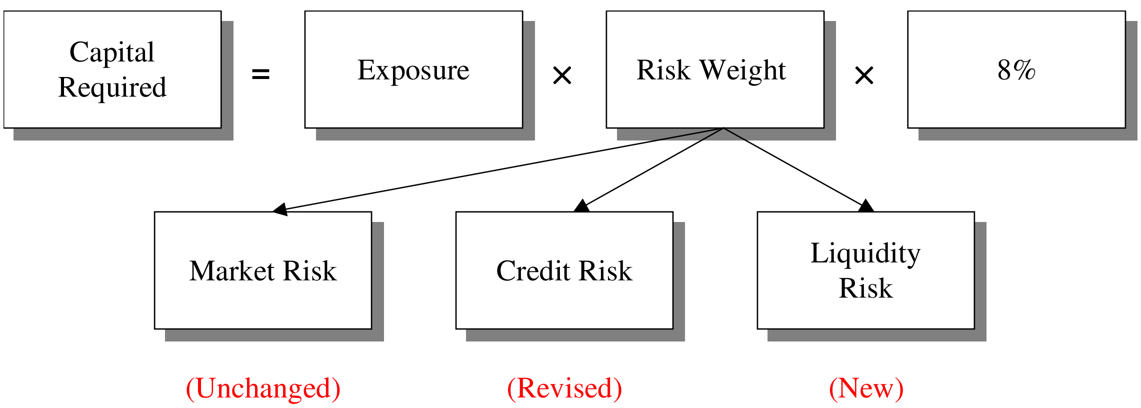


Figure 4.2: Capital Adequacy

Risk Weighted Assets are computed by:

- RWA = Banking Book RWA + Trading Book RWA
- Banking Book RWA = Position × Risk Weighting
- Trading Book RWA = Market Risk Capital / 8%

4.2 Major Types of VaR

The concept of Value at Risk (or VaR) began with a thought that had been developed in JP Morgan in early 1990s about how their money managers can sleep well at night and not be surprised the next morning. In other words, how could their managers know how much money was AT RISK over night. In 1998, Bank for International Settlements (BIS) used it for standard banking risk management (known as Basel Accord I).

VaR is to provide a notion (in dollar terms) of how much money can be possibly lost over a period of time (say over the next day (1-day VaR) or the next week (1-week VaR)) in the future. Managers choose a particular possibility, say 1%, and then VaR is a dollar amount associated with the possibility. To fulfill the different requests of such a wide variety of possibilities (from 0.1% VaR to 10% VaR), we need to estimate a tail distribution of the returns. There are three general ways to do so, known as:

- Historical VaR
- Parametric VaR
- Factor-based VaR

each of which will be discussed in details in this Lesson. Briefly speaking, a historical VaR is a histogram-based distribution. Past returns are collected and then a histogram is constructed and hence the tail distribution is obtained. The assumption of the historical VaR is that history will repeat itself and hence past distribution will be the distribution of the future.

A parametric VaR is to think that possible losses over a particular horizon in the future can be drawn form a Gaussian (normal) distribution. Hence, we need to estimate the two parameters of such a Gaussian distribution from the past data. As opposed to using the entire histogram, in parametric VaR, we only need two parameters from the past returns mean and variance. Then, we can construct a Gaussian distribution from which we can obtain a VaR by giving a desired left tail probability.

Finally, a factor-based VaR is to apply a model to estimate what the future dynamics of returns should look like. As opposed to the parametric VaR, a factor-based VaR adopts a full model that consists of a number of risk factors, each of which follows a dynamic random process. It is then not hard to imagine that the number of parameters in a factor-based VaR model is large and hence the work of estimating these parameters requires heavy econometric knowledge. As a demonstration, we graphically present the idea of VaR. An x% loss (left tail probability) corresponds to the level of r^* in the following diagram. Then r^* is multiplied by the total dollar position to compute VaR.

In addition to the calculation of the standard deviation calculation, VaR also includes stress tests, scenario analyses, and worst case analyses.

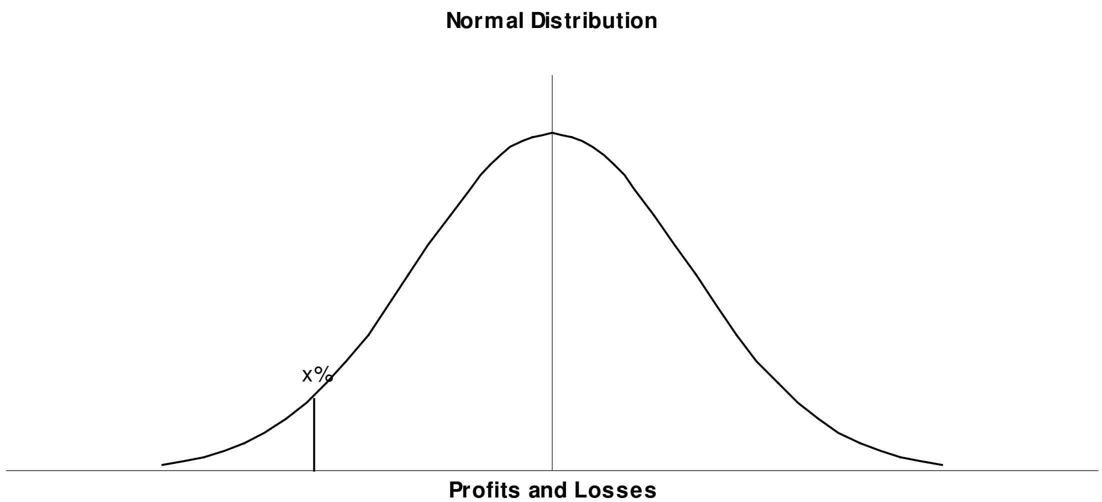


Figure 4.3: Value at Risk

4.2.1 Historical VaR

A historical VaR is one that is based upon only past return history and nothing else. A return history is called a histogram.

Histogram

A histogram is not necessarily Gaussian shape (usually not, known as the fat tails). To compute historical VaR is very simple. One simply sorts the possible losses in the past (say 1 year which is roughly 250 observations) and choose from the worst till the pre-set percentage (say 4% VaR) is met (say the 10th observation). An example is given in Figure 4.4.

Historical VaR is very simple. When the asset mix of a firm is simple (e.g. only stocks), historical VaR provides a somewhat reliable VaR number (if the process is stantionary that is past reflects future). When the asset mix is not so simple (e.g. stocks and bonds), then a histogram is easy to estimate (but may not be reliable to reflect the future).

However, historical VaR also suffers from a serious criticism. That is, historical VaR is not forward looking. Past distributions may not be accurate representations of the future. Other criticisms of historical VaR include:

- difficult to integrate VaRs of different securities (e.g. bond and stock)

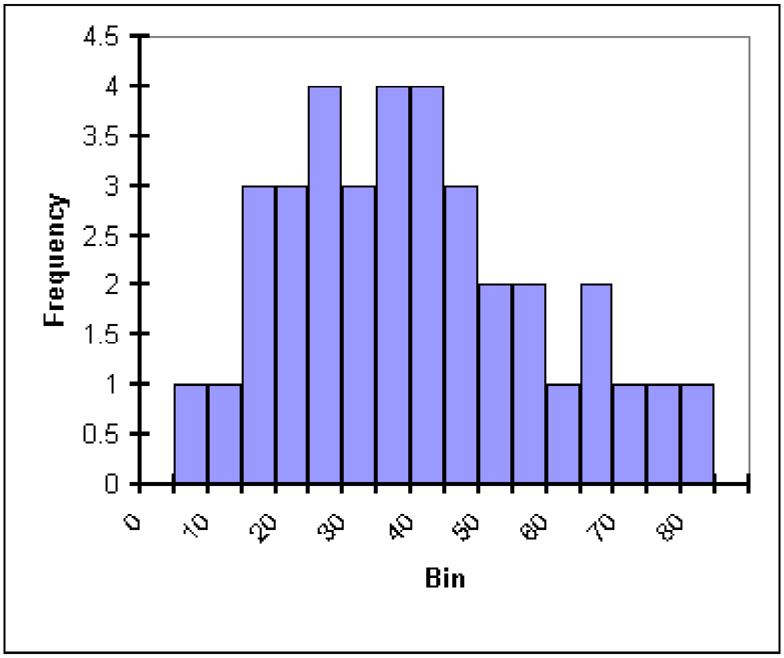


Figure 4.4: A Histogram Sample

- derivative assets have certain arbitrage-free relationship with its underlying asset
- empirically very unstable

Examples

An ideal example for the historical VaR is the VaR of a single stock. An example of FaceBook (FB) 2008/5/17 $\sim$ 2009/1/2 has 157 stock prices which turn into 156 returns. A histogram is summarized in Figure 4.5.

with the basic distribution statistics as in Table 4.3.

Historical VaR: there are 156 observations and 5% of that is 7.8 or 8. In other words, 5% historical VaR is the 8th worst observation which is 6.56%. This is one day VaR. If we invest in 36,010 shares (or \$1 million at 27.77 per share). Then the historical VaR is \$65,620.

Some arguments are made against the above VaR number. Industry argues that return VaRs are not convenient to use and furthermore cannot be combined

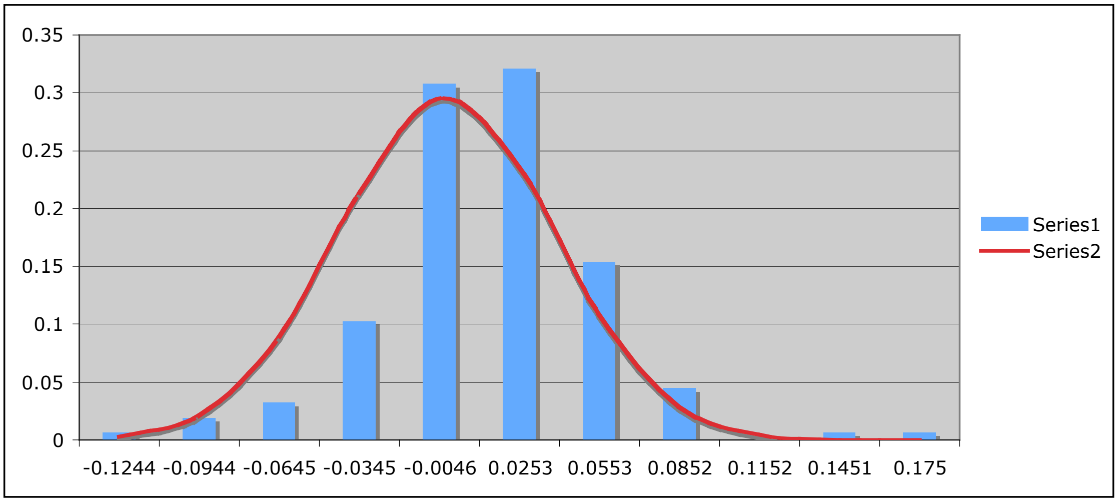


Figure 4.5: Facebook Example

Mean	-0.002049130
Std.dev	0.040425118
Skew	0.204800858
Extra kurt	2.686877978

Table 4.3: Facebook Distribution Statistics

(known as the additive property). Hence, it is more conventional for the industry to use price changes as opposed to returns. The resulting VaR is known as the dollar VaR.

For the same example, the FB price change (ΔS) histogram is plotted as follows (as opposed to returns that is $\Delta S/S$).

with the basic distribution statistics as in Table 4.4.

The 8th worst is \$1.88. Say the number of shares is 36,010 (or \$1 million at 27.77 per share). Then the historical VaR is \$67,699.

Historical VaR can work for any portfolio as long as there is data. In the following I combine FB with WMT (Walmart) for the same period. Given that both FB and WMT are traded, the portfolio value (which is one share of FB and one share of WMT) can be computed and observed daily. Hence, for the same 157

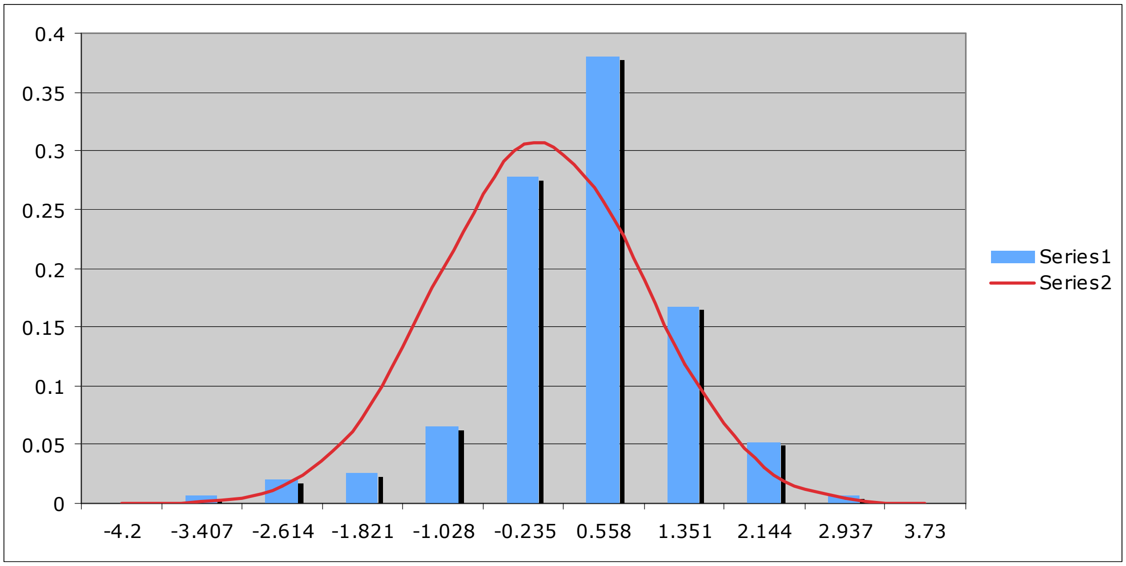


Figure 4.6: Facebook Example

Mean	-0.06705130
Std.dev	1.02337596
Skew	-0.50810290
Extra kurt	3.04349237

Table 4.4: Facebook Distribution Statistics

days, the return histogram for the portfolio is given in Figure 4.7.

with the basic distribution statistics in Table 4.5.

The eighth worst observation is 2.29%. This is the portfolio one-day VaR. Assuming \$1 million investment, the historical VaR is \$22,857.

Similarly, we can do the dollar VaR for the portfolio and the histogram is in Figure 4.8.

with the basic distribution statistics in Table 4.6.

The eighth worst observation is -2.14. This is the portfolio one-day VaR. Assuming \$1 million investment, the historical VaR is \$21,877.

The benefit of the historical VaR is that we do not need to compute the large

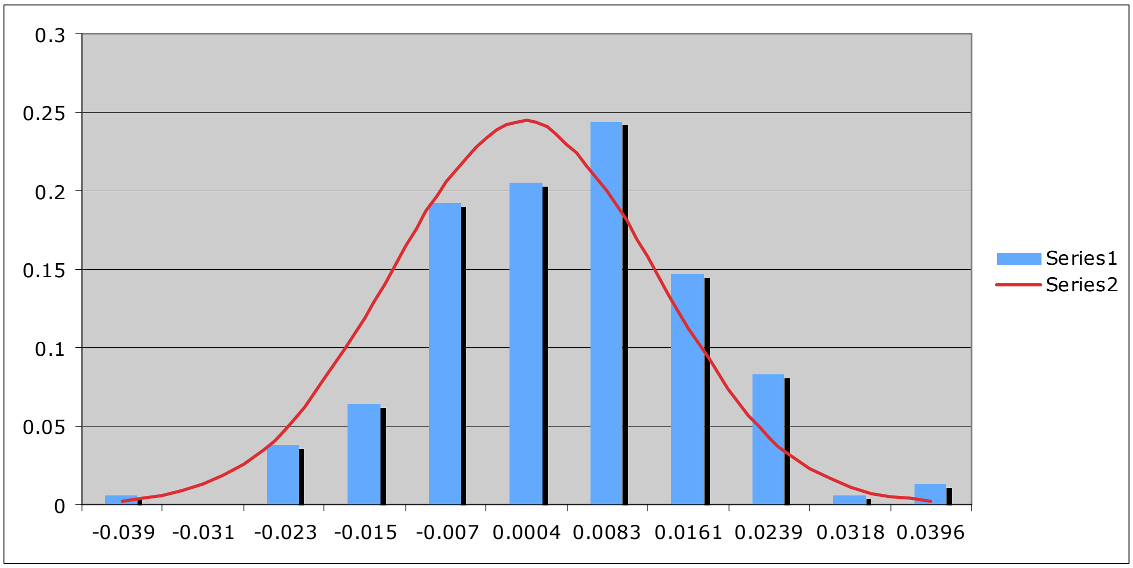


Figure 4.7: Facebook and Walmart

Mean	9.5066E-05
Std.dev	0.01273587
Skew	-0.08926460
Extra kurt	0.27394334

Table 4.5: Combined Distribution Statistics

variance-covariance matrix as the parametric VaR does (to be discussed in the next section). Historical VaR, we have seen, looks at only the 8th (i.e. the 5th percentile) observation of the history of the portfolio. However, to use the historical VaR, there must be data for all the constituents in the portfolio.

Also to be noted is that historical VaR does not suffer from the criticism of the fat tails. Given that it takes a chosen percentile from the left tail, both skewness and kurtosis are automatically incorporated and fat tails are supported.

The same principle can be applied to a stock (or any "primary" asset, or known as a "basic asset class") and its derivatives (or any derivative assets), such as an option like call or put. But often the derivatives are traded in an illiquid market (over the counter) and its prices are not always available. In this case, we can use a

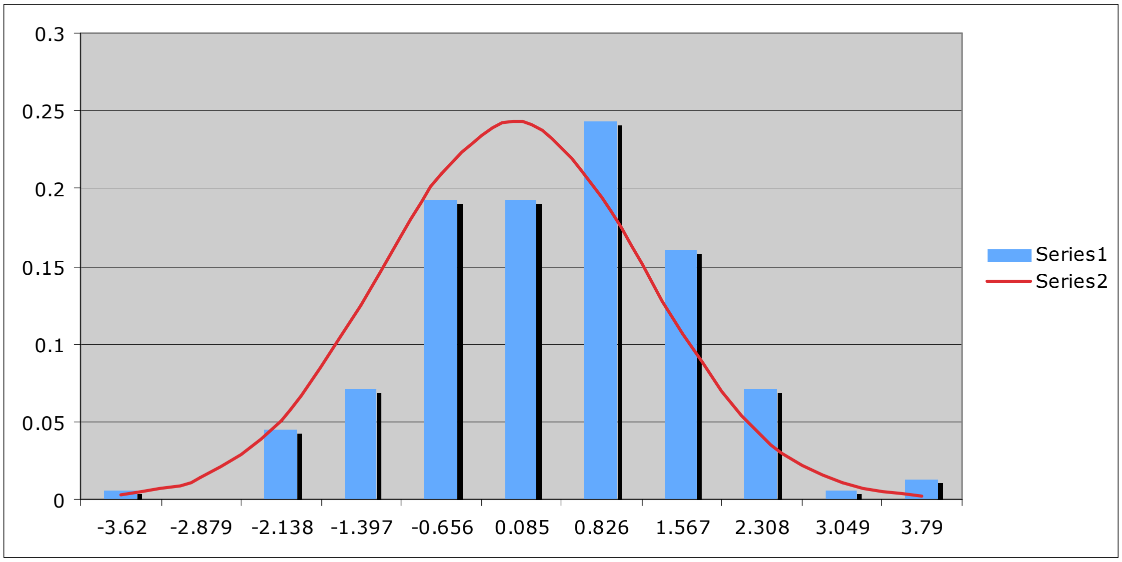


Figure 4.8: Facebook and Walmart (Dollar Change)

Mean	0.00923077
Std.dev	1.21440841
Skew	-0.08985380
Extra kurt	0.23078100

Table 4.6: Combined Distribution Statistics (Dollar Change)

model to "fill in" any missing prices there are in the history. The reason that we can do so is because there is a strong functional relationship between derivatives and their underlying assets. In the past several decades, numerous pricing models have been developed to accurately estimate the derivative prices from their underlying asset values. Later, we shall demonstrate several classical such pricing models. Here, we use the most famous Black-Scholes model for call option as a demonstration.

The Black-Scholes model for a call option on stock can be written as

$$C_t = S_t N(d_1) - e^{-r(T-t)} K N(d_2) \tag{4.1}$$

where S_t is the current stock price (assuming that t is the current time), K is the strike price, r is the constant risk-free rate (which can be relaxed to be random in order to match reality), σ is the volatility (which is standard deviation of stock returns), $T-t$ is the time till maturity of the option (assuming that T is the maturity time), $N(\cdot)$ is the standard normal probability, and

$$d_1 = \frac{\ln S_t - \ln K + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = d_1 - \sigma\sqrt{T-t} = \frac{\ln S_t - \ln K + (r - \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$$

The standard normal probability $N(\cdot)$ can be found in Excel as NORMSDIST(x) where x is either d_1 or d_2.

With model-generated option prices, we can do the historical VaR as follows. The following example is a histogram where the stock prices are collected historically and yet the call option prices are computed by the Black-Scholes model. The blue bars represent the distribution of the FB stock. The red bars represent the call option. The hollow bars represent the portfolio that consists of one stock and one option. The bin for each histogram is given as follows (10 equal intervals). Note that the histogram of the stock is the same as the one shown earlier.

bin	FB	FB call	Portfolio
1	-4.2	-4.1522	-8.3438
2	-3.407	-3.4974	-6.8968
3	-2.614	-2.8426	-5.4499
4	-1.821	-2.1878	-4.0029
5	-1.028	-1.533	-2.556
6	-0.235	-0.8782	-1.109
7	0.558	-0.2234	0.33793
8	1.351	0.43136	1.78488
9	2.144	1.08615	3.23183
10	2.937	1.74094	4.67878
11	3.73	2.39573	6.12573

Table 4.7: Facebook and Call

We can see that the distribution of option price changes is highly skewed. Certainly this is no surprising as options are leveraged derivative contracts.

The 8th worst price change (5% of the histogram) of the portfolio is -3.46. If the call option data are available, one could do a true historical VaR using actual option data.

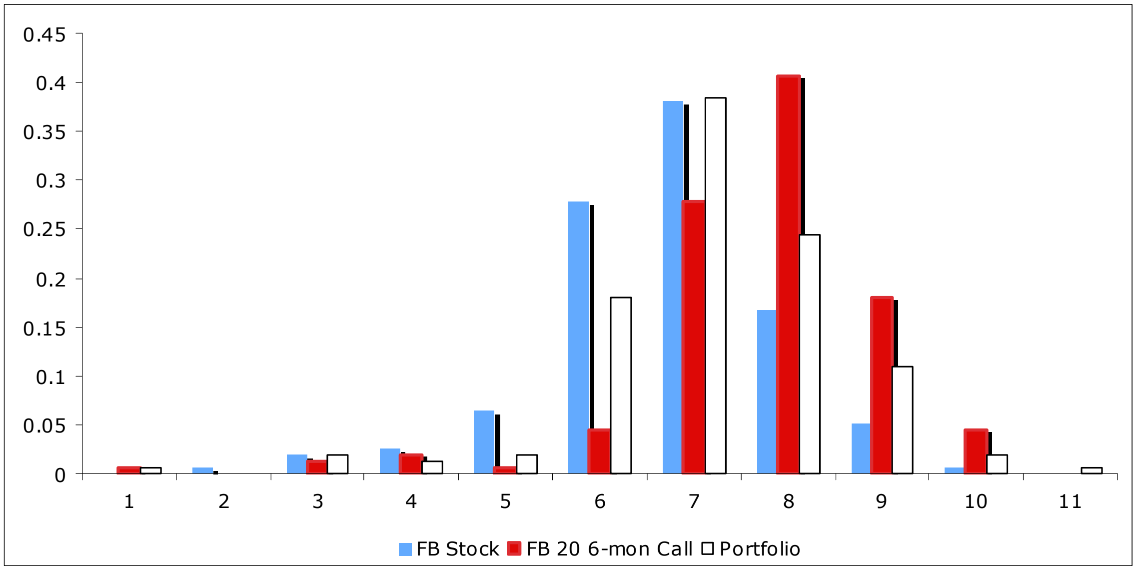


Figure 4.9: Facebook and Call

4.2.2 Parametric VaR

Once not all securities have historical data, historical VaR cannot apply. As a result, we can use parametric VaR. That is, we employ a model and no longer rely on history to compute VaR. The simplest parametric VaR model is the adoption of the Gaussian distribution.

Parametric VaR also has other advantages such as:

- all securities are modeled under a consistent framework (note that historical VaR combines securities values brute-forced, which could combine apples with oranges)

- parametric VaRs are scalable (square root of time)

An early concept of parametric VaR was one standard deviation value. In other words, if the market swings by one standard deviation, how much money would be lost. In a normal distribution, one standard deviation is the 16th percentile. In other words, one standard deviation drop in value can happen with 16% probability. Similarly, a two-standard deviation loss is 2.3% probability. If the manager wants to know what is 1% (or 5%) likely loss, under normal distribution, it is 2.326 (1.645) standard deviations. If one standard deviation is \$100,000, then with 1% probability the portfolio will lose more than \$233,000.

Example (continued)

Continue with the FB example, we know from the distribution statistics that the mean return is −0.205% and the standard deviation is 0.0404. Hence the 95% VaR is:

$$-6.86\% = -0.205\% - 1.645 \times 0.0404$$

If position is \$1 million, then VaR is \$68,543 for parametric (compared with \$65,620 for historical VaR). For the VaR of 1 week, we must do historical VaR separately, but parametric VaR is simply $68,543 \times \sqrt{7} = 181,348$. Under dollar VaR, we have

$$-\$1.75 = -0.067 - 1.645 \times 1.0234$$

using the mean and standard deviation given in the previous section. For \$1 million position (i.e. 36,010 shares), this is \$63,030 (compared with \$67,699 for historical VaR). Certainly, this VaR can be scaled by the square root of the number of days.

Two-stock example

For two stocks (FB and WMT), the return parametric VaR is \$20,854 (compared with the historical VaR of \$22,857) and the dollar parametric VaR is \$20,326 (compared with the historical VaR of \$21,877).

A portfolio of two stocks, say 40%-60%. Historical VaR is fine as long as the history of two stocks exists. But if not (say one has a few missing values) then we cannot do historical VaR (of course we can back-fill these values but that is up to the assumptions of back-filling.

We could do parametric VaR using the portfolio means and standard deviation, defined as follows:

$$\begin{cases} \mu_P = w_1\mu_1 + w_2\mu_2 \\ \sigma_P = \sqrt{w_1^2\sigma_1^2 + 2w_1w_2\rho\sigma_1\sigma_2 + w_2^2\sigma_2^2} \end{cases} \tag{4.2}$$

where μ represents mean and σ represents standard deviation of stock returns. ρ is the correlation coefficient between the two stocks. Note that $\rho\sigma_1\sigma_2$ is the covariance between the two stocks (usually a notation σ_{12} is used). We can conveniently compute the covariance from data using an Excel function `COVAR(array1, array2)` without calculating the correlation coefficient.

Continue the same example in the previous section. The correlation between two series of returns is -0.1399. Hence, we can compute the variance-co-variance matrix to be:

0.0000983	-0.0000560
-0.0000560	0.0016342

Table 4.8: Variance-Covariance Matrix for Returns

The portfolio consists of 10,223 shares of each stock (at $28.76 for FB and $69.06 for WMT) and hence $1 million in total. This represents 70% WMT and 30% FB.

As a result, the portfolio mean and standard deviation (using equation (4.2)) are: -0.0001561 and 0.0131 respectively and the 95% VaR is -0.021710, or $21,710 (compared to $22,857 historical VaR).

For the price change VaR it is best to use number of shares as weights. To see that:

$$\Delta V = n_1 \Delta S_1 + n_2 \Delta S_2$$

Compute the variance of the price change:

$$\begin{aligned}\mathrm{var}[\Delta V] &= n_1^2 \hat{\sigma}_1^2 + 2 n_1 n_2 \hat{\sigma}_{12} + n_2^2 \hat{\sigma}_2^2 \\ &= [n_1 \ n_2] \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} \\ \hat{\sigma}_{12} & \hat{\sigma}_2^2 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}\end{aligned}$$

where $\hat{\sigma}$ presents dollar variances and covariances. The variance-covariance matrix for the price change becomes Table 4.9.

As a result, the portfolio standard deviation is $11,824.63 and the 95% VaR is $19,452 (compared to $21,877), ignoring the mean.

0.4911838	-0.1016057
-0.1016057	1.0472983

Table 4.9: Variance-Covariance Matrix for Price Changes

n stock example

The portfolio is characterized by a set of portfolio weights: $w_1, \cdots, w_N$. The expected return of the portfolio is:

$$\sum_{i=1}^{N} w_i \mu_i = w'_{1\times N} M_{N\times 1} \tag{4.3}$$

and the variance of the portfolio is:

$$\left(w'_{1\times N} \Omega_{N\times N} w_{N\times 1}\right)_{1\times 1} \tag{4.4}$$

where Ω is the var-cov matrix of stocks. Hence the portfolio VaR is 1.645 (95%) or 2.326 (99%) multiplied the square root of the following variance. Note that for the price change VaR weights are shares.

An example of combining stock and option

It is quite common for a portfolio to contain derivative assets. This could be yield enhancements (i.e. using derivatives to increase leverage) or hedging (i.e. using derivatives to cancel exposures). As a result, we must know how a derivative asset relates to its underlying asset.

In derivatives pricing, Black and Scholes, in their seminal work which won them the Nobel prize in 1997, show that delta hedging is a perfect way to reduce the risk of a derivative asset completely. In other words, one can build a portfolio with a derivative asset and its underlying asset and reach 0 risk. However, in a situation of imperfect hedging or yield enhancement, there is still risk and VaR can still be computed.

The parametric VaR can be computed easily. Assume a portfolio of 1 FB stock and 1 at-the-money call option (on the same stock). The information of the stock is: $S = 27.77$, $K = 20$, $r = 0$, $\sigma = 0.4$, $T - t = 0.5$. The option value is computed as:

$$
\begin{aligned}
C &= SN(d_1) - e^{-r(T-t)}KN(d_2) \\
&= 27.77 \times 0.9035 - 1 \times 20 \times 0.8459 \\
&= 8.17
\end{aligned} \tag{4.5}
$$

Now, you need to calculate (dollar) VaR of your portfolio. The change of value of your portfolio can be computed as:

$$
\begin{aligned}
dV &= dS + dC \\
&= dS + \Delta dS \\
&= dS(1 + \Delta)
\end{aligned} \tag{4.6}
$$

where $\Delta = N(d_1)$ is the delta of your call. As a result, the volatility of dV is:

$$
\begin{aligned}
\text{Std.Dev.}(dV) &= \text{Std.Dev.}(dS) \times (1 + \Delta) \\
&= S \times \sigma \times (1 + \Delta) \\
&= 27.77 \times 0.4 \times (1 + 0.9035) \\
&= 21.14
\end{aligned} \tag{4.7}
$$

As a result, the six-month portfolio VaR is $21.14 \times 1.645 = 34.78$. One could scale this parametric VaR to one-day to be $34.68/\sqrt{126} = 3.10$ (half year is 126 business days.) Note that the historical VaR computed previously is \$3.46 so this parametric VaR is smaller. This is one share of stock and one share of option. If we have 100 shares of stock and 20 shares of option, then the VaR becomes $27.77 \times 0.4 \times (100 + 20 \times 0.9035) = 1311.52$.

Note that if we short the call option, then the option contract would serve as a hedge. In this way, we drastically reduce the risk. The standard deviation becomes:

$$
\begin{aligned}
\text{Std.Dev.}(dV) &= \text{Std.Dev.}(dS) \times (1 - \Delta) \\
&= S \times \sigma \times (1 - \Delta) \\
&= 27.77 \times 0.4 \times (1 - 0.9035) \\
&= 1.07
\end{aligned} \tag{4.8}
$$

and the 5% six-month portfolio VaR $1.07 \times 1.645 = 1.76$. Scaling it to one-day (126 days), we obtain 0.16. Now we know that using derivatives can completely remove risk by selecting proper shares of the option to buy (known as the hedge ratio). Assume that we buy 1,000 shares of the stock (which is worth \$27,770 today) and would like to find out how many shares of 20-strike call option to short. Then the price change equation becomes:

$$\begin{aligned} \text{Std.Dev.}(dV) &= S \times \sigma \times (1,000 - n_C\Delta) \\ &= 0 \end{aligned} \tag{4.9}$$

We would like the risk to be 0 and hence we can easily solve for the number of shares n_C to short which is equal to $n_C = \frac{1000}{\Delta} = 1106.78$, or roughly 1,107 shares. In other words, by shorting 1,107 shares of call option, the risk is completely eliminated.

4.2.3 Factor Model Based VaR

Historical VaR is not very useful. When the asset mix of a firm is simple (e.g. only stocks), it is easier to estimate a historical probability distribution. But if the asset mix is not so simple (e.g. stocks and bonds), then a historical probability distribution is not so easy to estimate. Especially if some assets are not liquidly traded, then there is no history of these assets. Then the historical VaR method will fail.

Historical VaR also suffers from a serious criticism. That is, historical VaR is not forward looking. Past distributions may not be accurate representations of the future. Hence, it is important that we forecast what future distribution should look like. Other criticisms of historical VaR include:

- security returns are usually not normal
- assuming a correlation (again, this is normal-based) is inaccurate
- difficult to integrate VaRs of different securities (e.g. bond and stock)
- derivative assets have certain arbitrage-free relationship with its underlying asset
- empirically very unstable

Given the above obvious reasons, VaR models have become parametric. That is, we employ a model and no longer rely on history to compute VaR. The standard VaR model used by the financial industry is a linear factor model. Although the details vary, a VaR model is typically a linear model in widely observed economic variables. As a result, the VaR models used today are connected with P&L attributions.

From P&L attributions, we know that there are a number of explanatory factors, such as FX rates, interest rates, equity indicies, various volatilities, $\cdots$, etc.

To put all assets under one VaR umbrella, we need a flexible and yet manageable model. First of all the risks must be additive. Standard deviations are NOT additive. So we must use something else. We use delta, or more generally, the first order sensitivity with respect to the underlying risk factor.

Principal Component Analysis

A PCA (Principal Component Analysis) is implemented to decide how many significant factors that affect the equity market. Given N stocks, a PCA could extract only a handful (say 3$\sim$5) of factors (decided by the eigenvalues of the PCA). These are known as common factors. One can think of these common factors as macroeconomic variables such as inflation, GDP, exchange rates, etc. However, these factors extracted using PCA do not have those interpretations (in fact, this the main drawback of the PCA methodology, which is the factors are unintepretable.) Despite of this drawback, PCA is a "honest" methodology. In other words, PCA looks at the data of equity returns and recognizes (from how highly or lowly these equity returns are correlated) how many factors (K) are really needed to explain N stocks.[1]

To use PCA, one must assume a linear factor-based return generating model of the following:

$$R_{i,t} = b_{i,1}F_{1,t} + b_{i,2}F_{2,t} + \cdots + b_{i,N}F_{N,t} \tag{4.10}$$

In matrix form (which is important when we use Excel to carry out calculations), we can write this as:

$$\begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1T} \\ R_{21} & \ddots & & \\ \vdots & & \ddots & \\ R_{N1} & R_{N2} & \cdots & R_{NT} \end{bmatrix} = \begin{bmatrix} b_{11} & \cdots & b_{1N} \\ \vdots & & \vdots \\ b_{N1} & \cdots & b_{NN} \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1T} \\ F_{21} & & & \\ \vdots & & & \\ F_{N1} & F_{N2} & \cdots & F_{NT} \end{bmatrix}$$
$$\mathbf{R}_{N\times T} = \mathbf{B}_{N\times N} \times \mathbf{F}_{N\times T} \tag{4.11}$$

PCA takes the return matrix (i.e. $\mathbf{R}$ and returns a set of coefficients (eigenvectors) which are the known as factor loadings or "factor betas". In other words, a PCA outputs an $N \times N$ matrix $\mathbf{B}$. Depending upon the matrix arrangement, sometimes it is convenient to arrange the return matrix as an $T \times N$ matrix:

[1] That is, PCA can decide how many dimensions (K) are needed to expand the equity space which is N dimensions.

$$\begin{bmatrix} R_{11} & R_{21} & \cdots & R_{N1} \\ R_{12} & \ddots & & \\ \vdots & & \ddots & \\ R_{1T} & R_{2T} & \cdots & R_{NT} \end{bmatrix} = \begin{bmatrix} F_{11} & F_{21} & \cdots & F_{N1} \\ F_{12} & & & \\ \vdots & & & \\ F_{1T} & F_{2T} & \cdots & F_{NT} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{N1} \\ \vdots & & \vdots \\ b_{1N} & \cdots & b_{NN} \end{bmatrix} \tag{4.12}$$
$$\mathbf{R}'_{T\times N} = \mathbf{F}'_{T\times N} \times \mathbf{B}'_{N\times N}$$

Hence, depending on what return matrix (either $N\times T$ or $T\times$N) is fed to PCA, the corresponding factor loading matrix (either $\mathbf{B}$ or $\mathbf{B}'$) will be generated. As a result, we can obtain the factor time series by inverting the beta matrix as follows:

$$\begin{gathered} \mathbf{B}^{-1}_{N\times N} \times \mathbf{R}_{N\times T} = \mathbf{F}_{N\times T} \\ \text{or} \\ \mathbf{R}'_{T\times N} \times (\mathbf{B}')^{-1}_{N\times N} = \mathbf{F}'_{T\times N} \end{gathered} \tag{4.13}$$

Note that factor loading matrix is an orthogonal matrix (i.e. inverse equals transpose or $\mathbf{B}' = \mathbf{B^{-1}}$)). Hence, we can write $\mathbf{R}' \times \mathbf{B} = \mathbf{F}'$.

Let us have a simple 2×2 example where only two stocks are used. The returns of the two stocks are given as follows:

2 Stock PCA		
Date	Stock 1	Stock 2
2/9/2006	4.32	4.52
2/10/2006	4.36	4.53
2/13/2006	4.38	4.55
2/14/2006	4.42	4.55
2/15/2006	4.39	4.55
...	...	...

After running PCA, we obtain the factor loading matrix ($\mathbf{B}$):

0.970203	-0.242295
0.242295	0.970203

To solve for the factor values, we can solve for the following equations:

$$0.97 \times F_{11} - 0.24 \times F_{21} = 4.32$$
$$0.24 \times F_{11} + 0.97 \times F_{21} = 4.52$$

and hence $F_{11} = 5.29$ and $F_{21} = 3.34$.

So when we pre-multiply the return matrix ($\mathbf{R}'$) by the factor loading matrix, we should get the factor value matrix ($\mathbf{R}' \times \mathbf{B} = \mathbf{F}'$):

2 Stock PCA		
Date	Factor 1	Factor 2
2/9/2006	5.286448	3.338602
2/10/2006	5.327679	3.338612
2/13/2006	5.351929	3.353171
2/14/2006	5.390737	3.343479
2/15/2006	5.361631	3.350748
...	...	...

This matrix operation is quite easy in Excel. `MMULT(··· ,···)` can perform the above matrix multiplication easily. Note that the right area in Excel must be blocked and then use `Ctrl-Shft-Entr`.

If all N stocks are independent, then PCA will find identical eigenvalues and there is no benefit of dimension reduction (i.e. $K = N$). However, if stocks are highly correlated, then PCA will find fast-decaying eigenvalues, indicating that the first few factors can explain most of the returns (i.e. $K < N$). In this case, we find benefits of using PCA.

Theoretically, this can be explained by the following equation:

$$\begin{aligned} R_{i,t} &= b_{i,1}F_{1,t} + b_{i,2}F_{2,t} + \cdots + b_{i,N}F_{N,t} \\ &= \mu_i + (b_{i,1}f_{1,t} + \cdots + b_{i,K}f_{K,t}) + e_{i,t} \\ \text{or, } dP_{i,t} &= a_i + \sum_{k=1}^{K} b_{i,k}f_{k,t} + e_{i,t} \end{aligned} \tag{4.14}$$

with $\mu_i = \sum_{k=1}^{K} b_{i,k}\mathbb{E}[F_k]$. Note that $f_j = F_j - \mathbb{E}[F_j]$ and hence $\mathbb{E}[f_1] = 0$, $\mathbb{E}[f_2] = 0$ and $\mathbb{E}[e_i] = 0$.

As we can see, only K factors (out of potentially N factors) are chosen and the rest are thrown into the "error term". By making the assumptions that the error term (which consists of unimportant factors) is negligible, we only need part of the factor loading matrix: $\hat{\mathbf{B}}_{N\times K}$ and hence we have the variance-covariance matrix of the stocks being replaced with the variance-covariance of the factors as follows:

$$\Omega_{P(N\times N)} = \hat{\mathbf{B}}_{N\times K}\Omega_{F(K\times K)}\hat{\mathbf{B}}'_{K\times N} \tag{4.15}$$

where Ω_P is the variance-covariance matrix of stocks. These factor values are also used for simulations. Note that only the significant factors will be used for predicting stock returns. So out of N factors, only K actors are significant and used. The rest are treated as errors (idiosyncratic).

Portfolio returns over time are computed as:

$$\sum_{i=1}^{N} w_i R_{i,t} \tag{4.16}$$

so the entire time series of the portfolio returns can be computed as:

$$\underline{w}'_{1\times N}\mathbf{R}_{N\times T} \tag{4.17}$$

Using PCA, the portfolio is:

$$\underline{w}'_{1\times N}\mathbf{R}_{N\times T} = \underline{w}'_{1\times N}\mathbf{B}_{N\times K}\mathbf{F}_{K\times T} + \underline{w}'_{1\times N}\mathbf{e}_{N\times T} \tag{4.18}$$

and hence the portfolio variance:

$$\sigma_P^2 = \left(\underline{w}'_{1\times N}\Omega_{P(N\times N)}\underline{w}_{N\times 1}\right)_{1\times 1} \tag{4.19}$$

can be re-expressed in terms of factors as follows:

$$\left(\underline{w}'_{1\times N}\hat{\mathbf{B}}_{N\times K}\Omega_{F(K\times K)}\hat{\mathbf{B}}'_{K\times N}\underline{w}_{N\times 1}\right)_{1\times 1} \tag{4.20}$$

As a result, the VaR number (at α, say 5%) of the portfolio is the following calculation:

$$\begin{aligned} \text{VaR} &= N^{-1}(\alpha) \times \sqrt{\underline{w}'\Omega_P\underline{w}} \\ &= N^{-1}(\alpha) \times \sqrt{\underline{w}'\hat{\mathbf{B}}\Omega_F\hat{\mathbf{B}}'\underline{w}} \end{aligned} \tag{4.21}$$

This method reduces the computation complexity of an $N \times N$ variance-covariance matrix to a $K \times K$ variance-covariance matrix. Furthermore, the factor model can be estimated easily.

In simulations, factors are assumed normal with the means and variances estimated by the PCA. Error terms are also normal with mean 0 and variances estimated by the PCA. This way, we can construct the simulated return series of the stocks and estimate the VaR that is forward-looking.

Regression Model

PCA has a severe disadvantage that the factors carry no intuitive meaning. Often modelers need to run analyses to investigate what these factors mean. The usual method is correlate factor values with known economic indicators (believing that these factors reflect fundamental economy).

An alternative is to directly use known economic indicators as factors. Then we simply run regressions with these observable economic indicators as independent variables.

$$r_j = a_j + b_{j1} r_{F1} + b_{j2} r_{F2} + \cdots + e_j \tag{4.22}$$

where variables in the regression can be any chosen economic variables. Or one could use the popular Fama-French factors where the common factors are posted on the French website. The Fama-French factors have gained tremendous popularity over the years and have become the standard of empirical asset pricing.

A More Complex Fixed Income Example

In this example, we limit ourselves to only interest rate risk. Let the instantaneous rate follow a three-factor model (where factors are independent of one another):

$$r = y_1 + y_2 + y_3 \tag{4.23}$$

where factors are independent and each factor y_i is assumed to follow either the Vasicek or the CIR model as follows:

$$\begin{cases} dy_k = \alpha_k(\mu_k - y_k)dt + \sigma_k dW_k & \text{Vasicek} \\ dy_k = \alpha_k(\mu_k - y_k)dt + \sigma_k \sqrt{y_k} dW_k & \text{CIR} \end{cases} \tag{4.24}$$

where $k = 1, 2, 3$. Under either model, we have a closed form solution to the discount factor (for each k):

$$\begin{aligned} Y_k(t,T) &= \hat{\mathbb{E}}_t \left[\exp\left\{ -\int_t^T y_k(u)du \right\} \right] \\ &= A_k(t,T) \exp\{-B_k(t,T) y_k(t)\} \end{aligned} \tag{4.25}$$

Under the Vasicek model we have the following closed-form solutions to A_k and B_K:

$$
\begin{aligned}
-\ln A_k(t,T) &= \left(\mu_k - \frac{\sigma_k\lambda_k}{\alpha_k} - \frac{\sigma_k^2}{2\alpha_k^2}\right)(T-t-B_k(t,T)) + \frac{\sigma_k^2 B_k(t,T)^2}{4\alpha_k} \\
B_k(t,T) &= \frac{1-e^{-\alpha_k(T-t)}}{\alpha_k}
\end{aligned}
\tag{4.26}
$$

and under the CIR model we have:

$$
\begin{aligned}
A_k(t,T) &= \left[\frac{2\gamma_k e^{(\alpha_k+\lambda_k+\gamma_k)(T-t)/2}}{(\alpha_k+\lambda_k+\gamma_k)\left(e^{\gamma_k(T-t)}-1\right)+2\gamma_k}\right]^{2\alpha_k\mu_k/\sigma_k^2} \\
B_k(t,T) &= \frac{2\left(e^{\gamma_k(T-t)}-1\right)}{(\alpha_k+\lambda_k+\gamma_k)\left(e^{\gamma_k(T-t)}-1\right)+2\gamma_k} \\
\gamma_k &= \sqrt{(\alpha_k+\lambda_k)^2+2\sigma_k^2}
\end{aligned}
\tag{4.27}
$$

where λ is the market price of risk. The fundamental factors, y_ks, span the entire risk free fixed income universe. Hence, the job now is to estimate these factors, namely we need to estimate α_k, μ_k, σ_k, and λ_k for $k = 1, \cdots, K$. To estimate this, we need to implement the MLE (maximum likelihood estimation). In an exercise at the end of this Chapter, readers can practice these calculations with parameters given (estimated by Chen and Yeh and Chen and Scott).

Now because the factors are assumed independent, the risk-free discount factor can be written as:

$$
\begin{aligned}
P(t,T) &= \hat{\mathbb{E}}_t\left[\exp\left\{-\int_t^T r(u)du\right\}\right] \\
&= Y_1(t,T)Y_2(t,T)Y_3(t,T) \\
&= A_1(t,T)A_2(t,T)A_3(t,T)\exp\{-B_1(t,T)y_1 - B_2(t,T)y_2 - B_3(t,T)y_3\}
\end{aligned}
\tag{4.28}
$$

Then the price change can be easily derived as:

$$
dP(t,T) = (-B_1(t,T)dy_1 - B_2(t,T)dy_2 - B_3(t,T)dy_3)P(t,T) \tag{4.29}
$$

Define an arbitrary coupon bond (symbolized as j where $j = 1, \cdots, J$) as:

$$
\Pi_j = \Pi(c_j, T_{n_j}) = c_j \sum_{i=1}^{n_j} P(t,T_i) + P(t,T_{n_j}) \tag{4.30}
$$

where n_j is the number of coupons paid by bond j and c_j is same coupon at each time T_i.

Then, we can derive the price change of a coupon bond as:

$$
\begin{aligned}
d\Pi_j &= c_j \sum_{i=1}^{n_j} dP(t,T_i) + dP(t,T_n) \\
&= c_j \sum_{i=1}^{n_j} (-B_1(t,T_i)dy_1 - B_2(t,T_i)dy_2 - B_3(t,T_i)dy_3)P(t,T_i) \\
&+ (-B_1(t,T_n)dy_1 - B_2(t,T_n)dy_2 - B_3(t,T_n)dy_3)P(t,T_n) \\
&= a_1 dy_1 + a_2 dy_2 + a_3 dy_3
\end{aligned} \tag{4.31}
$$

where for k = 1, 2, and 3,

$$
a_k = -c_j \sum_{i=1}^{n_j} B_k(t,T_i)P(t,T_i) - B_k(t,T_n)P(t,T_n)
$$

Hence, we can easily compute variances and covariances as:

$$
\begin{aligned}
\mathrm{var}[dP] &= P^2(B_1^2 v_1^2 + B_2^2 v_2^2 + B_3^2 v_3^2) \\
\mathrm{cov}[dP, d\Pi] &= b_1 v_1^2 + b_2 v_2^2 + b_3 v_3^2 \\
\mathrm{var}[d\Pi] &= a_1^2 v_1^2 + a_2^2 v_2^2 + a_3^3 v_3^2
\end{aligned} \tag{4.32}
$$

where (for $k = 1, 2, 3$)

$$
\begin{aligned}
b_k &= a_k(-B_k)P \\
v_k^2 &= \mathrm{var}[dy_k]
\end{aligned} \tag{4.33}
$$

If we invest n_1 units in P and n_2 units in Π,[2] then the 5% VaR is:

$$
\mathrm{VaR}_{5\%} = 1.645 \times \sqrt{[n_1, n_2] \begin{bmatrix} \mathrm{var}[dP] & \mathrm{cov}[dP, d\Pi] \\ \mathrm{cov}[dP, d\Pi] & \mathrm{var}[d\Pi] \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}} \tag{4.34}
$$

To write it in a matrix form for the var-cov matrix:

[2]Note that the total investment amount is $V = n_1 P + n_2 \Pi$ and weights are $w_1 = n_1 P/V$ and $w_2 = n_2 \Pi / V$.

$$\begin{bmatrix} \text{var}[dP] & \text{cov}[dP, d\Pi] \\ \text{cov}[dP, d\Pi] & \text{var}[d\Pi] \end{bmatrix} =$$
$$\begin{bmatrix} -PB_1 & -PB_2 & -PB_3 \\ a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} v_1^2 & 0 & 0 \\ 0 & v_2^2 & 0 \\ 0 & 0 & v_3^2 \end{bmatrix} \begin{bmatrix} -PB_1 & a_1 \\ -PB_2 & a_2 \\ -PB_3 & a_3 \end{bmatrix} \tag{4.35}$$

Hence, similar to the PCA, we can write:

$$\hat{\mathbf{B}} = \begin{bmatrix} -PB_1 & a_1 \\ -PB_2 & a_2 \\ -PB_3 & a_3 \end{bmatrix} \tag{4.36}$$

which plays the role of the factor loading matrix (which is $N \times K$) as in equation (4.21).

We can easily compute any other covariance and fill the variance-covariance matrix. As we can see, once all the parameters are estimated, we complete the variance-covariance matrix.

If we do not want to use a full model (CIR or Vasicek), we can let the fundamental factors to be 11 key Treasury rates, $\xi_{t,T}$. Instead of using the CIR or Vasicek model, we adopt the following yield to maturity formula:

$$P(t,T) = e^{-\xi_{t,T}(T-t)} \tag{4.37}$$

and the coupon bond formula remains the same. The market only observes 11 key rates but we need all the rates for $\underline{T}$. So we need some interpolation method. For the time being, lets assume piece-wise flat. That is $y(t,T) = y(t,T_j)$ for $T_j < T < T_{j+1}$.

Commonly (according to a Goldman Sachs white paper), there are three factors to span the yield curve:

- short rate
- slope
- curvature (convexity)

and hence $K = 3$ seems to be appropriate (even though the factors are unobservable). If we have a stock-fixed income portfolio, then we may need factors for the stocks, and commonly:

- market index (e.g., S&P 500)
- industry indices (e.g., SIC code 9 industries)

PCA can easily incorporate the factors in a augmented matrix. For parametric models, we may consider the Merton-Rabinovitch model (1973-1989) where both stocks and interest rates can be random.

4.3 Marginal and Component VaR

Given that VaR is such a powerful and intuitive risk measure, it is then quite important to break down this number into various divisions of the company in order to see which division of the company is responsible for the most risk born by the company. This is understood as the component VaR.

A three-stock example in the following demonstrates how each component VaR is calculated.

$$\begin{aligned}\sigma_P^2 &= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= \begin{cases} x_1^2\sigma_{11} + x_1x_2\sigma_{12} + x_1x_3\sigma_{13} \to X_1 \\ x_1x_2\sigma_{12} + x_2^2\sigma_{22} + x_2x_3\sigma_{23} \to X_2 \\ x_1x_3\sigma_{13} + x_2x_3\sigma_{23} + x_3^2\sigma_{33} \to X_3 \end{cases}\end{aligned} \tag{4.38}$$

Taking April 1, 2003 till October 6, 2009 weekly prices of JNJ (Johnson & Johnson), JPM (JP Morgan), and KO (Coke Cola), we obtain the following variance-co-variance matrix in Table 4.10.

	JNJ	JPM	KO
JNJ	0.000514	0.000575	0.000372
JPM	0.000575	0.004493	0.000631
KO	0.000372	0.000631	0.000714

Table 4.10: JNJ, JPM, KO

We further assume that the investment amounts on these stocks are \$10,000 (35.71%), \$6,000 (21.42%), and \$12,000 (42.86%). The portfolio standard deviation in percent terms is 0.02685 and in dollar terms \$751.852. As a result, the 5% VaR is \$1,236.69.

To calculate the three component VaRs for JNJ, JPM, and KO, we first note that $x_1 = 10,000, x_2 = 6,000$, and $x_3 = 12,000$. Hence, $X_1 = 130,635, X_2 = 241,674$, and $X_3 = 192,972$, which sum to a total of 565,281.399. As a result, each $X_1 \sim X_3$ accounts for 23.11%, 42.75%, and 34.14% respectively. Using these percentages to allocate the portfolio VaR of $1,236.69, we get the three component VaRs as in Table 4.11 which by construction add up to the portfolio VaR of $1,236.69.

JNJ	285.795
JPM	528.720
KO	422.171

Table 4.11: Component VaRs

Note that from Table 4.10, we can directly obtain individual VaRs to be $372.987, $661.505, and $527.564 for JNJ, JPM, and KO respectively (as in Table 4.12), which totals to $1,562.06 – greater than the portfolio VaR of $1,236.69. The difference is of course known as the diversification effect. Given that individual VaRs do not add up to the portfolio VaR, we cannot see how much each stock is contributed to the portfolio VaR. As a result we need component VaRs.

JNJ	372.987
JPM	661.505
KO	527.564

Table 4.12: Individual VaRs

Another useful concept is marginal VaR. Marginal VaRs measure the risk reduction if one division of the company is eliminated. In our example, it is merely VaRs with an elimination of one stock at a time.

$$
\begin{aligned}
Y_1 &= \sqrt{\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}} - \sqrt{\begin{bmatrix} x_2 & x_3 \end{bmatrix} \begin{bmatrix} \sigma_{22} & \sigma_{23} \\ \sigma_{23} & \sigma_{33} \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix}} \\
Y_2 &= \sqrt{\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}} - \sqrt{\begin{bmatrix} x_1 & x_3 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{13} \\ \sigma_{13} & \sigma_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_3 \end{bmatrix}} \\
Y_3 &= \sqrt{\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}} - \sqrt{\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}
\end{aligned}
\tag{4.39}
$$

and each marginal VaR is $Y_i \times 1.645$ (for 5% VaR). This way, we can see how by eliminating the first stock, how much VaR can be reduced. The marginal VaR for JNJ is \$256.057 (note that the two-stock portfolio of KO and JPM has a VaR of \$980.629). The marginal VaRs for JPM and KO are left for an exercise.

4.4 Decay

It is difficult to choose a proper window for VaR calculations. If the window is too short, then the results will miss important economic cycles. If the window is too long, then the results lose the relevancy as recent history which is more relevant is weighted the same as distant history which is no so relevant.

To overcome this problem, it is quite common for a bank to adopt a set of decaying weights where more recent history is weighted more heavily than more distant history. Also, in this regard, one can take as long historical data as possible.

To do this, we first remind the readers that an unweighted standard deviation formula is:

$$\sigma = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}(r_{t-i} - \bar{r}_t)^2} \tag{4.40}$$

where $\bar{r}_t = \frac{1}{n}\Sigma_{i=1}^{n} r_{t-i}$ is the average. In other words, each observation is weighted by $1/n$. We could weigh each observation differently. One popular method is an exponential decay where more recent data are weighted more heavily than more distant data. Let $\theta < 1$. Adjust each term in the above formula as follows:

$$\theta^{i-1}(r_{t-i} - \bar{r}_t)^2 \tag{4.41}$$

where the most recent term r_{t-1} is weighted 1. Note that,

$$1 + \theta + \cdots + \theta^{n-1} = \frac{1-\theta^n}{1-\theta} \tag{4.42}$$

and hence we have:

$$\sigma_t = \sqrt{\frac{1-\theta}{1-\theta^n}\sum_{i=1}^{n}\theta^{i-1}(r_{t-i} - \bar{r}_t)^2} \tag{4.43}$$

Note that as $n \to \infty, \theta^n \to 0$ and the adjustment simplifies to $\frac{1}{1-\theta}$. Equation (4.43), however, becomes an infinite sum. Luckily, we can work out a recursive equation as follows. We can write equation (4.43) as follows:

$$\begin{aligned} \sigma_t &= \sqrt{(1-\theta)\sum_{i=1}^{\infty}\theta^{i-1}(r_{t-i}-\bar{r}_t)^2} \\ \sigma_{t-1} &= \sqrt{(1-\theta)\sum_{i=1}^{\infty}\theta^{i-1}(r_{t-1-i}-\bar{r}_{t-1})^2} \end{aligned} \tag{4.44}$$

and hence we can then write:

$$\sigma_t^2 = \theta\sigma_{t-1}^2 + (1-\theta)(r_{t-1}-\bar{r}_t)^2 \tag{4.45}$$

assuming that $\bar{r}_{t-1} \sim \bar{r}_t$. This recursive equation is convenient and avoids the infinite sum. Note that this result is similar to EWMA (exponentially weight moving average) described in Chapter 10.

4.5 Exercises

Given the parameter values as follows (estimated by Chen and Yeh (2009) and Chen and Scott (1993)), compute the VaR number in equation (4.34). We use prices of the two bonds to solve for $y_1(t)$ and $y_2(t)$. Then, with the following parameter values we can proceed with the calculations.

	Chen-Scott Estimation				New Estimation			
	factor 1	std.err.	factor 2	std.err.	factor 1	std.err.	factor 2	std.err.
α	1.834100	0.222800	0.005212	0.115600	0.879967	0.001014	0.004423	0.000014
μ	0.051480	0.005321	0.030830	0.683300	0.043822	0.000009	0.029555	0.000097
σ	0.154300	0.005529	0.066890	0.002110	0.097855	0.001429	0.095974	0.000018
λ	-0.125300	0.180600	-0.066500	0.115400	-0.146140	0.000151	-0.178846	0.000361
	likelihood function = 7750.82				likelihood function = 11722.81			
	# of obs. 470				# of obs. 416			

Figure 4.10: CIR parameters: 2-factor model

Chapter 5

Fixed Income Risk Management

5.1 Interest Rates

First of all, we must note that there are a large number of interest rates that all need to be modeled. Usually we start with the least risky rate, which is a collection of the U.S. Treasury rates. The U.S. economy is the worlds largest (its GDP is nearly three times as much as that of the second largest which is China) and hence its government bonds (Treasuries) are regarded as the least risky investments.

5.1.1 US Treasuries

There are plenty of popular interest rate models used by various financial institutions. The most popular one is perhaps the Heath-Jarrow-Morton model published in 1992 in Econometrica.

There are a variety of Treasuries outstanding:

- T bills zero coupon, up to 1 year
- T notes semi-annual coupon, 1 10 years
- T bonds semi-annual coupon, 10 30 years
- CMT (Constant Maturity Treasury)
- TIPS (Treasury Inflation Protected Security)
- STRIPS (Separate Trading of Registered Interest and Principal Securities)

Treasury bills (T bills) are less than one year zero coupon Treasury securities. Three T bills are auctioned once a week on Thursdays 4 week (Monday auction), 13 week, and 26 week T bills. Hence, these are "on-the-run" T bills.

Treasury notes (T notes) usually auctioned every month. The "on-the-run" T notes are currently 2, 5, and 10 year T notes. Only 30-year Treasury bonds (T bonds) are auctioned right now, four times a year, in February, May, August, and November.

TIPS are inflation protected Treasuries and are auctioned for 5 and 10 years only on an irregular basis. TIPS pay coupons that are inflation adjusted where the adjustment is tied to CPI (Consumer Price Index) that is published by the monthly by the Bureau of Labor Statistics of the United States Department of Labor.

CMTs are interpolated (weighted average of Treasury yields) Treasury interest rates published on the fly by the Treasury department. Given that the actual Treasury issues have fixed maturities, and hence cannot provide good benchmarking for other interest rates (e.g. swap rates), the Treasury department compiles interest rates for "constant maturities". Note that CMT rates are "semi-annual par rates", which means, it is the coupon rate of a Treasury issue of the given tenor sold at the face value. Currently, there are 1, 3, 6 month, and 1, 2, 3, 5, 7, 10, 20, 30 year CMTs available.

STRIPS are T-Notes, T-Bonds and TIPS whose interest and principal portions of the security have been separated, or "stripped; these may then be sold separately in the secondary market. The name derives from the notional practice of literally tearing the interest coupons off (paper) securities. The government does not directly issue STRIPS; they are formed by investment banks or brokerage firms, but the government does register STRIPS in its book-entry system. They cannot be bought through TreasuryDirect, but only through a broker.

There are two interest rates controlled by the Federal Reserve Bank.

- FED fund rate
- Discount rate

The Fed funds rate is the interest rate at which banks lend their federal funds at the Federal Reserve to banks, usually overnight. Hence, it is an interbank lending rate. This rate is usually higher than the short term Treasury rates.

The discount rate is the interest rate at which member banks may borrow short term funds directly from a Federal Reserve Bank. The discount rate is one of the two interest rates set by the Fed, the other being the Federal funds rate. The Fed actually controls this rate directly, but this fact does not really help in policy implementation, since banks can also find such funds elsewhere. This rate is lower than the short term Treasury rates.

5.1.2 LIBOR (London Interbank Offer Rate)

LIBOR stands for London InterBank Offer Rate and is an interbank rate between major commercial banks, led by the Bank of England, in London. LIBORs have become the benchmark interest rates in the financial industry (in place of Treasury rates), mainly because these are the funding costs of most banks in the financial industry. There are three popular LIBOR rates set by the major commercial banks: 1 month, 3 month, and 6 month rates. However, there are long term LIBOR derivatives such as Eurodollar futures contracts and interest rate swaps.

- LIBOR (London Interbank Offer Rate)
- Eurodollar futures
- IRS (Interest Rate Swaps)

Given that U.K. has no central bank, unlike the U.S., the Bank of England serves the role of the central bank for the U.K. As a result, LIBOR symbolizes the government rate. However, technically the Bank of England is not the central after all, LIBOR as a result remains a private interest rate.

LIBOR is published by the British Bankers Association (BBA) after 11:00 am (and generally around 11:45 am) each day, London time, and is a filtered average of inter-bank deposit rates offered by designated contributor banks, for maturities ranging from overnight to one year. There are 16 such contributor banks and the reported interest is the mean of the 8 middle values. The shorter rates, i.e. up to 6 months, are usually quite reliable and tend to precisely reflect market conditions. The actual rate at which banks will lend to one another will, however, continue to vary throughout the day.

Floating rate products use LIBORs as benchmarks, mainly because companies that issue floating rate bonds cannot borrow at the Treasury rates. Although the actual situations may change, by and large, LIBORs are in between AAA and AA corporate yields.

5.1.3 Agencies

Agencies refer to three government supported financial institutions that underwrite and guarantee residential mortgages.

corporate bonds of GNMA, FNMA, FHLMC

Government National Mortgage Association (GNMA, pronounced Gennie Mae), Federal National Mortgage Association (FNMA, pronounced Fannie Mae), and Federal Home Loan Mortgage Corporation (FHLMC, pronounced Freddie Mac) are three government agencies that underwrite residential mortgages to the secondary market.

These three agencies were set up after the World War II to help low income and military veterans to purchase homes. Recently the congress, after recognizing the historical mission has been successfully accomplished, announced that they were no longer government agencies but private financial institutions. Yet, investors still believe that there is implicit government backing when these three agencies are in trouble. As a result, they issue corporate bonds with lower yields. An informal rating for these three agencies is AAAA, one rating higher than the highest rating given by rating agencies.

5.1.4 The Model

There are two broad branches of interest rate models, known as the reduced-form (or no-arbitrage) models and structural (or equilibrium) models. Reduced-form models focus on market information and hence take market observables (such as interest rates and volatilities) as inputs. Structural models focus on economic fundamentals and hence take economic drivers as inputs.

The most famous structural models for the interest rates are Vasicek (1977) and Cox-Ingersoll-Ross (1985) which can be written as follows:

$$\begin{array}{ll} dr_t = \alpha(\mu - r_t)dt + \sigma dW & \text{Vasicek} \\ dr_t = \alpha(\mu - r_t)dt + \sigma\sqrt{r_t}dW & \text{Cox - Ingersoll - Ross} \end{array} \tag{5.1}$$

and the risk-free discount factor can be derived as:

$$P_{t,T} = A_{t,T}e^{-r_t B_{t,T}} \tag{5.2}$$

$$\begin{aligned} -\ln A_{t,T} &= \left(\mu - \frac{\sigma\lambda}{\alpha} - \frac{\sigma^2}{2\alpha^2}\right)(T - t - B_{t,T}) + \frac{\sigma^2 B_{t,T}^2}{4\alpha} \\ B_{t,T} &= \frac{1 - e^{-\alpha(T-t)}}{\alpha} \end{aligned} \tag{5.3}$$

under the Vasicek model and

$$
\begin{aligned}
A_{t,T} &= \left[\frac{2\gamma e^{(\alpha+\lambda+\gamma)(T-t)/2}}{(\alpha+\lambda+\gamma)\left(e^{\gamma(T-t)}-1\right)+2\gamma} \right]^{2\alpha\mu/\sigma^2} \\
B_{t,T} &= \frac{2\left(e^{\gamma(T-t)}-1\right)}{(\alpha+\lambda+\gamma)\left(e^{\gamma(T-t)}-1\right)+2\gamma} \\
\gamma &= \sqrt{(\alpha+\lambda)^2+2\sigma^2}
\end{aligned}
\tag{5.4}
$$

under the CIR model where λ is the market price of risk. A simple numerical example is provided below.

For the Vasicek model, the parameters can be reasonably set as:[1]

Inputs	
α	0.2456
μ	0.0648
σ	0.0289
λ	-0.2718
$r(t)$	0.0600

and the output yield curve is given below:

Term Structure					
$T-t$	$P(t,T)$	$y(T-t)$	$T-t$	$P(t,T)$	$y(T-t)$
1	0.9380	0.0641	16	0.2639	0.0833
2	0.8740	0.0673	17	0.2413	0.0836
3	0.8106	0.07	18	0.2207	0.0840
4	0.7491	0.0722	19	0.2017	0.0842
5	0.6905	0.0741	20	0.1844	0.0845
6	0.6352	0.0756	21	0.1686	0.0848
7	0.5835	0.077	22	0.1541	0.085
8	0.5354	0.0781	23	0.1409	0.0852
9	0.4908	0.0791	24	0.1288	0.0854
10	0.4497	0.0799	25	0.1177	0.0856
11	0.4118	0.0807	26	0.1076	0.0857
12	0.3769	0.0813	27	0.0984	0.0859
13	0.3449	0.0819	28	0.0899	0.086
14	0.3155	0.0824	29	0.0822	0.0862
15	0.2886	0.0829	30	0.0751	0.0863

[1] This was estimated by Chen and Yang.

For the CIR model, the parameters can be reasonably set as:[2]

Inputs	
α	0.2456
μ	0.0648
σ	0.14998
λ	-0.129
$r(t)$	0.06

The resulting term structure by the CIR model is given as follows.

Term Structure					
$T-t$	$P(t,T)$	$y(T-t)$	$T-t$	$P(t,T)$	$y(T-t)$
1	0.9379	0.0641	16	0.2641	0.0832
2	0.8738	0.0675	17	0.2417	0.0835
3	0.8099	0.0703	18	0.2212	0.0838
4	0.7481	0.0726	19	0.2025	0.0841
5	0.6891	0.0745	20	0.1853	0.0843
6	0.6336	0.0761	21	0.1696	0.0845
7	0.5818	0.0774	22	0.1552	0.0847
8	0.5337	0.0785	23	0.142	0.0849
9	0.4893	0.0794	24	0.1299	0.085
10	0.4483	0.0802	25	0.1189	0.0852
11	0.4107	0.0809	26	0.1088	0.0853
12	0.3761	0.0815	27	0.0996	0.0854
13	0.3443	0.082	28	0.0911	0.0856
14	0.3152	0.0825	29	0.0834	0.0857
15	0.2886	0.0829	30	0.0763	0.0858

The most widely used reduced-form model is the Heath-Jarrow-Morton model. Here we introduce a "baby-version" of the model first derived by Ho and Lee (1986). The model we use for the "risk-free" rate (represented by the U.S. Treasuries) is the Ho-Lee (HL) model. The HL model is a special case of the popular Heath-Jarrow-Morton (HJM) model and is very easy to implement.

The Ho-Lee model is a "forward rate" model and hence belongs to the HJM family (although the Ho-Lee model was published 6 years prior to the HJM model). However, the original version of the Ho-Lee model is a "forward price" model. It was Phil Dybvig who then extended the Ho-Lee model to continuous time, forward rate model in 1989.

[2]This was estimated by Chen and Yang.

Define the zero-coupon bond price as $P(i, n, j)$ where i is current time, n is maturity time, and j represents state of economy. The Ho-Lee model is a simple formula as follows:

$$\begin{cases} P(i, i+k, j) = \dfrac{P(i-1, i+k, j)}{P(i-1, i, j)} d(k) \\ P(i, i+k, j+1) = \dfrac{P(i-1, i+k, j)}{P(i-1, i, j)} u(k) \end{cases} \tag{5.5}$$

where

$$\begin{aligned} u(k) &= \frac{1}{p + (1-p)\delta^k} \\ d(k) &= \frac{\delta^k}{p + (1-p)\delta^k} \end{aligned}$$

and p is the risk-neutral probability and δ is the "volatility" parameter (that is, δ itself is not volatility but it is directly related to volatility.) As we can see, when $\delta = 1$, then $u(k) = d(k)$ for all k. Then there is no volatility. To maintain $u(k) > d(k) > 0$, it must be that $0 < \delta < 1$. As δ becomes small the volatility becomes large.

Take a four-year yield curve as an example:

Yield Curve		
time to maturity	discount factor	yield to maturity
1	0.9524	0.049979
2	0.8900	0.059998
3	0.8278	0.065021
4	0.7686	0.068009

Given the current term structure of discount factors, we then can compute the forward prices,

$$\begin{aligned} 0.934481 &= \frac{0.8900}{0.9524} \\ 0.869173 &= \frac{0.8278}{0.9524} \\ 0.807014 &= \frac{0.7686}{0.9524} \end{aligned}$$

Note that forward rates are returns of forward prices. For example:

$$
\begin{aligned}
f_{0,1,2} &= \frac{1}{0.934481} - 1 = \frac{(1+5.9998\%)^2}{1+4.9979\%} - 1 = 7.0112\% \\
f_{0,1,3} &= \sqrt{\frac{1}{0.869173}} - 1 = \sqrt{\frac{(1+6.5021\%)^3}{1+4.9979\%}} - 1 = 7.2623\% \\
f_{0,1,4} &= \sqrt[3]{\frac{1}{0.807014}} - 1 = \sqrt[3]{\frac{(1+6.8009\%)^3}{1+4.9979\%}} - 1 = 7.4088\%
\end{aligned}
$$

Different from the equity binomial that has only one pair of u and d. In our example, we set $p = 0.6$ and $\delta = 0.9$ and we have:

Ho-Lee Pert Funcs		
k	$d(k)$	$u(k)$
0	1.000000	1.000000
1	0.937500	1.041667
2	0.876623	1.082251
3	0.817631	1.121579
4	0.760749	1.159501

In the HL model, the next periods up term structure and down term structure are computed by applying the proper u and d on the forward price. For convenience, we introduce the following labeling system. $P(i,n,j)$ represents the discount factor value at current time i, for maturity time n, and in state j. For convenience, we also label the lowest state 0 and 1, 2, as we go up. So for today, we have four discount factors $P(0,1,0), P(0,2,0), P(0,3,0), P(0,4,0)$ which are 0.9524, 0.8900, 0.8278, and 0.7686 respectively. For next year, we apply the forward prices computed above and multiply them by corresponding us and ds. For example,

$$0.9734 = \underbrace{P(1,2,1)}_{\text{up}} = \frac{P(0,2,0)}{P(0,1,0)}u(1) = \frac{0.8900}{0.9524} \times 1.041667$$

$$0.8761 = \underbrace{P(1,2,0)}_{\text{down}} = \frac{P(0,2,0)}{P(0,1,0)}d(1) = \frac{0.8900}{0.9524} \times 0.9375$$

$$0.9407 = \underbrace{P(1,3,1)}_{\text{up}} = \frac{P(0,3,0)}{P(0,1,0)}u(2) = \frac{0.8278}{0.9524} \times 1.082251$$

$$0.7619 = \underbrace{P(1,3,0)}_{\text{down}} = \frac{P(0,3,0)}{P(0,1,0)}d(2) = \frac{0.8278}{0.9524} \times 0.876623$$

$$0.9051 = \underbrace{P(1,4,1)}_{\text{up}} = \frac{P(0,4,0)}{P(0,1,0)}u(3) = \frac{0.8278}{0.9524} \times 1.121579$$

$$0.6598 = \underbrace{P(1,4,0)}_{\text{down}} = \frac{P(0,4,0)}{P(0,1,0)}d(3) = \frac{0.7686}{0.9524} \times 0.817631$$

Put in the table,

Year = 1						
current time	maturity time	state 0	state 1	state 2	state 3	state 4
1	1	1	1			
	2	0.8761	0.9734			
	3	0.7619	0.9407			
	4	0.6598	0.9051			

Now we have two term structures of the next year (time 1). The task continues to time 2. For each term structure in time 1, we shall compute two term structures in time 2 (up and down) by applying the same principle. The up and down term structures of the left are:

$$0.9060 = P(2,3,1) = \frac{P(1,3,0)}{P(1,2,0)}u(1) = \frac{0.7619}{0.8761} \times 1.041667$$

$$0.8154 = P(2,3,0) = \frac{P(1,3,0)}{P(1,2,0)}d(1) = \frac{0.7619}{0.8761} \times 0.9375$$

$$0.8151 = P(2,4,1) = \frac{P(1,4,0)}{P(1,2,0)}u(2) = \frac{0.6598}{0.8761} \times 1.082251$$

$$0.6603 = P(2,4,0) = \frac{P(1,4,0)}{P(1,2,0)}d(2) = \frac{0.6598}{0.8761} \times 0.876623$$

and of the right are:

$$
\begin{aligned}
1.0066 = P(2,3,2) &= \frac{P(1,3,1)}{P(1,2,1)}u(1) = \frac{0.9407}{0.9734} \times 1.041667 \\
0.9060 = P(2,3,1) &= \frac{P(1,3,1)}{P(1,2,1)}d(1) = \frac{0.9407}{0.9734} \times 0.9375 \\
1.0063 = P(2,4,2) &= \frac{P(1,4,1)}{P(1,2,1)}u(2) = \frac{0.9051}{0.9734} \times 1.082251 \\
0.8151 = P(2,4,1) &= \frac{P(1,4,1)}{P(1,2,1)}d(2) = \frac{0.9051}{0.9734} \times 0.876623
\end{aligned}
$$

and put in table,

Year = 2						
current time	maturity time	state 0	state 1	state 2	state 3	state 4
2	2	1	1	1		
	3	0.8154	0.9060	1.0066		
	4	0.6603	0.8151	1.0063		

It can be seen that from the left (applying $u(k)$) we arrive at $P(2,3,1) = 0.9060$ which is the same from the right (applying $d(k)$). This also applies to $P(2,4,1) = 0.8151$. This is known as the re-combination assumption. Note that this assumption must be maintained or the binomial model will explode.

Continuing this process going forward, we can derive the complete the 4-year table for the Ho-Lee model:

Complete Ho-Lee Model						
current time	maturity time	state 0	state 1	state 2	state 3	state 4
0	0	1				
	1	0.9524				
	2	0.89				
	3	0.8278				
	4	0.7686				
1	1	1	1			
	2	0.8761	0.9734			
	3	0.7619	0.9407			
	4	0.6598	0.9051			
2	2	1	1	1		
	3	0.8154	0.906	1.0066		
	4	0.6603	0.8151	1.0063		
3	3	1	1	1	1	
	4	0.7592	0.8435	0.9372	1.0414	
4	4	1	1	1	1	1

This concludes the discrete example.

There have been a huge number of term structure models in the literature. The following table roughly summarized (by no means exhaustive but hopefully representative). There is no clear way to classify interest models. Roughly we use the following matrix:

Term Structure Models		
	Factor models (also called Stru. / Eqm. models)	No-arbitrage models (also called Reduced-form models)
Single-factor	Vasicek (1977) Cox-Ingersoll-Ross (1985) Dothan (1978) Ball-Torous (1983) Duffie-Kan (1996)	Ho-Lee (1986) Black-Derman-Toy (1989) Black-Karasinzinsky (1990) Bloomberg model (flat vol) Heath-Jarrow-Morton (1992)
Multi-factor	Cox-Ingersoll-Ross (1985) Langetieg (1980) Brennan-Schwartz (1978) Longstaff and Schwartz (1992) Duffie-Pan-Singleton (2000)	Hull-White (1990) Heath-Jarrow-Morton (1992) Scott (1997) Chen-Yang (2002)
Others	Quadratic Constantinides (1992) Leippold-Wu (2002)	Random field Goldstein (2000)

The no-arbitrage models are discussed in Chapter **??**, yet only Ho-Lee, Black-Derman-Toy, and Heath-Jarrow-Morton models are covered. So far in this chapter, only Vasicek and CIR models are discussed. In this section, we try to bring in some practical flavor, given that Vasicek and CIR models while provide tremendous insight into term structure theories, are too limited to be used in the real world.

The Dothan model is similar to the Vasicek and CIR models but assumes a lognormally distributed short rate. It has been shown that this model could be unstable on the long end of the yield curve.

Ball and Torous model bonds with a Brownian Bridge process as the BB process is still Gaussian and yet the end of the variable can be pinned at any given value (such as face value of the bond). However, as Cheng (1991) pointed out, this model is not arbitrage-free.

One factor models bring tremendous insight to the term structure theories and yet they are not useful in reality. Empirists have found that one factor models are too limited in explaining the curvature and dynamics and the term structure. Hence, researchers have developed multi-factor models. The first such proposal was by Cox, Ingersoll, and Ross. In their paper where the seminal one factor model was introduced, they proposed the following multi-factor framework: $r = \sum_{i=1}^{n} y_j$ where r is the the short rate and it is a sum of multiple independent factors y_j. However, no explicit solution is given by CIR.

Langetieg, using the CIR suggestion, derived an explicit solution to the Vasicek model. This is straighforward as the sum of Gaussian variables is still Gaussian. This is difficult to do in the CIR model as the sum of non-central chi-square variables (unless they are independent) is not a non-central chi-square variable. The mean reverting processes used in Vasicek and CIR models are AR(1) processes and hence they are not independent.

Brennan and Schwartz explicitly specify a separate long rate to build their two factor model. This approach is better than the CIR suggestion in that these factors are observable (as opposed to unobservable y_j). The model has no closed form solution and needs to be solved numerically (e.g. finite difference method). However, since term rates cannot be modeled directly by separate processes as they may conflict each other, the Brennan-Schwartz model is not arbitrage-free.

The Duffie-Kan model broadened the scope and established the general solution of the affine family of term structure models. The Duffie-Pan-Singleton model extends the Duffie-Kan model to include credit risk. We shall discuss them separately.

Outside of the affine family, the solutions become unmanageable. There are two proposals in the literature. One is the quadratic approach where $r = \sum_{j=1}^{n} y_j^2$. This approach is plausible as the short rate will not be negative and each factor y_j can be normally distributed. Constantinides first proposed this idea and then was followed by other researchers, such as Leippold and Wu. The problem is that the solution is expressed in terms of y_j which are non-observable. Empirists have found that such models can explain the dynamics of the term structure better.

Lastly, Goldstein adopted the Random Field theory in physics to model the term structure. In physics, a field is a physical quantity associated to each point of spacetime. A field may be thought of as extending throughout the whole of space. The Random Field theory has given the term structure model infinite degrees of freedom so it is rich in nature. Yet the model is not practical and has not been tested empirically.

5.2 Forward Expectation

5.2.1 A Simple Concept

A forward expectation is an expected value taken under the "forward" probability measure. Under this probability measure the probability of greater than the mean is exactly half and same as the probability of lower than the mean.

We first take a look at the simple Black-Scholes model. The put-call parity states that $p + S = c + P(t, T)K$. Since the ITM probability of the put is exactly the same as the OTM probability of the call. As a result, for $c = p$, it must be that the in and out of money probabilities for both put and call are equal. From the parity, this means that $S = P(t,T)K$, or $K = S/P(t,T)$ which is the forward price of the stock.

In other words, if we set the strike price as the forward price of the stock then it precisely separates the distribution to be half and half.

In swap contracts, we set the swap rates so that the contract has no value. In other words, the swap rate must be the value which set the rate-rising and rate-falling probabilities equal to be 50%.

The forward probability measure can be derived rigorously just like the risk-neutral probability measure. One easy distinction between the two is that under the risk-neutral measure, the expectation of any future asset value is the future price of the asset today; and under the forward measure, the expectation of any future asset value is the forward price.

We also note that any forward price of an asset is not model-independent but only the compound value of the current asset price:

$$F(t,T) = \frac{S(T)}{P(t,T)} \tag{5.6}$$

The forward measure is to show that the forward price can define a probability measure under which its expected value is also its median.

5.2.2 More Formal Mathematics

The forward measure is mainly used under stochastic interest rates, as it involves discounting. Under deterministic interest rates, discounting stays fixed over time and hence the forward measure does not exist. Under this circumstance, the futures price and the forward price are equal.[3]

$$\begin{aligned} X(t) &= \hat{\mathbb{E}}_t \left[\exp\left(-\int_t^T r(u)du \right) X(T) \right] \\ &= P(t,T)F(t,T) \\ &= P(t,T)\tilde{\mathbb{E}}_t^{(T)}[X(T)] \end{aligned} \tag{5.7}$$

We note that (without giving a proof) forward measure is maturity-dependent (superscript) and hence is not unique; unlike the risk-neutral measure which is unique. Given that:

[3] A formal proof is given by Cox, Ingersoll, and Ross (1981).

$$P(t,T) = \hat{\mathbb{E}}_t \left[\exp\left(-\int_t^T r(u)du \right) \right]$$

it is clear that the change of measure can be performed on (5.7). See Chen (2013).

5.3 IR Risk Management

While VaR can be used in IR risk management, additional risk management measures are introduced due to the specific nature of fixed income assets.

The sources of interest rate risk (very different from stocks) can come from:

- fed policies monetary
- inflation
- economy real rate
- reflect more macro risks and micro risks
- PCA = 3
- liquidity and credit are important

5.3.1 Conventions

One thing very unique about fixed income securities markets is the trading and quoting conventions. Fixed income securities have the most complex trading and quoting conventions than any other securities.

Trading Conventions

Accrued interest is a convention in trading fixed income securities. For reason given before I was born, bonds are quoted by their "clean price" instead of the real price (called "dirty price"). Investors of bonds need to compute the price (dirty price) they have to pay by adding accrued interest to the clean price.

Hence, the dirty price is:

$$\text{Dirty Price} = \text{Clean Price} + \frac{60}{182} \times \frac{\text{cpn}}{2}$$

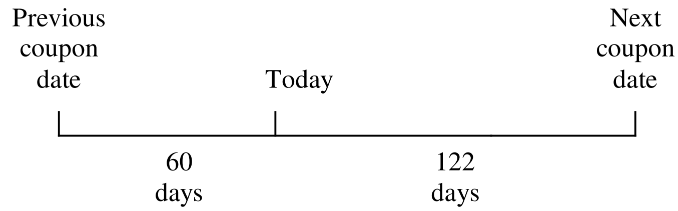


Figure 5.1: Accrued Interest

Tick size. Most fixed income securities are quoted on the thirty second basis. For example, a bond quoted at 100.16 does not mean the bond will be bought and sold at 100 dollars and 16 cents. But rather, the bond will be bought and sold at $100 + 16/32$, which is 100 dollars and 50 cents. One tick in fixed income securities is $1/32$. Stocks used to have ticks too of $1/8$ and $1/16$ but they do not have that anymore.

Rate quotes versus price quotes.

It is important to differentiate rate quotes, that are to determine transaction prices, and rates of return, that represent percentage return of an investment. Rate quotes are subject to day count conventions. For example, T bill quotes are subject to Actual/360 convention. Hence, a quote of 3.24 of a bill 21 days to maturity has a price of:

$$100 - 3.24 \times \frac{21}{360} = 99.81$$

The rate of return of this bill can be computed discretely as:

$$\frac{100 - 99.81}{99.81} \times \frac{365}{21} = 3.285\%$$

Hence 3.24 is not representing rate of return, 3.285 is. Note that 99.81 is a percent quote. It represents the price to pay for acquiring a bond is 99.81% of its face value. If the face value is \$100, it costs \$99.81; if it is $100,000, itcosts$99,810. If it is \$1, then it costs 0.9981. This is known as the discount factor.

Daycount Conventions

Daycount is a very special trading convention in the world of fixed income securities. Other markets have trading conventions (e.g. CBOE does not specify maturity date as a fixed date but the Saturday of the third Friday of the expiration month). The daycount convention specifies how many days should be in a month and in a year. There are 5 popular daycount conventions:

1. 0: 30/360 (corporate fixed)
2. A/A (T notes/bonds)
3. A/360 (corporate floaters, T bills)
4. A/365
5. European 30/360

The code is what is used by the Excel function yearfrac(a,b,c) where a is the beginning date, b is the ending date, and c is the code of the daycount convention . Lets first study 30/360 convention. This convention assumes that there are exactly 30 days in a month for any fraction of a month. For example, from 1/2/2003 to 2/28/2003, both months are not full months. In a normal calendar, January has 29 days and February has 28 days. But under 30/360, there are only 28 days in January (28 = 30 2) and 28 days in February. Hence, the period in years is:

$$0.155556 = \left[\frac{30-2}{30} + \frac{28}{30} \right] \div 12$$

This implies that there will be 1 day in January if we count from 1/29/2003 and 0 day if we count from 1/30/2003:

$$0.08055556 = \left[\frac{30-29}{30} + \frac{28}{30} \right] \div 12$$

Interestingly, since there can be only 30 days in a month, it will be 0 day if we count from 1/30/2003 and 1/31/2003:

$$0.07777778 = \left[\frac{0}{30} + \frac{28}{30} \right] \div 12$$

But on the other hand, if the ending date is 3/1/2003, then we will have a full month for February. Hence, although there is only one day difference between 2/28 and 3/1, the 30/360 daycount treats it as three days apart:

$$0.163888889 = \left[\frac{30-2}{30} + 1 + \frac{1}{30} \right] \div 12$$

5.3.2 Duration (delta) and Convexity (gamma)

Duration and convexity are very close to delta and gamma in option. It is first and second order derivatives of the bond price (or any fixed income security) with respect to a specific interest rate (or a collection of interest rates). Depending on which interest rate(s) chosen, we have different durations and convexities. This lesson introduces various duration and convexity calculations and how to use them. Finally, we shall talk about a very simple idea of immunization.

In the universe of fixed income, participants believe that the price of a fixed income security (bond, swap, ...) is a function of the yield curve. Hence, the change in the price is then a result of the rate change. Taylor's series expansion then provides a nice tool to analyze how the changes of various interest rates affect the price of the fixed income security.

Let P be the price of a fixed income security and y_i be the i-th interest rate on the yield curve that affects the price. Then Taylor's series expansion gives:

$$dP(y_1, y_2, \cdots y_n, t) = \frac{\partial P}{\partial t}dt + \sum\nolimits_{j=1}^{n} \frac{\partial P}{\partial y_j}dy_j + \frac{1}{2}\sum\nolimits_{i=1}^{n}\sum\nolimits_{j=1}^{n} \frac{\partial^2 P}{\partial y_i \partial y_j}dy_i dy_j + o(dt)$$

where $o(dt)$ includes terms that are small and $\lim\limits_{dt\to 0} \frac{o(dt)}{dt} = 0$. The first term is similar to Theta in option and known as roll-down on the yield curve. As time goes by, the life of a fixed income security become shorter. If nothing else in the economy changes (hence, the yield curve stays exactly the same), the yield will become less (in an upward sloping situation). The second term is duration with respect to various interest rates, and the third is convexity with respect to various interest rates.

MaCaulay Duration

MaCaulay duration is to treat the bond as a function of only its own yield. That is, recall from the previous lesson the yield to maturity formula:

$$\begin{aligned} P &= \left[\frac{c/2}{(1+y)^{(T_1-t)}} + \frac{c/2}{(1+y)^{(T_2-t)}} + \cdots + \frac{1+c/2}{(1+y)^{(T_n-t)}}\right] N \\ &= \left[\sum_{j=1}^{n} \frac{c/2}{(1+y)^{(T_j-t)}} + \frac{1}{(1+y)^{(T_n-t)}}\right] N \end{aligned}$$

where c is coupon rate, T_j is the coupon time, y is yield to maturity and N is notional. We must solve for the yield y for the bond. Duration is the first order derivative:

$$\frac{dP}{dy} = \left[\sum_{j=1}^{n} -(T_j - t)\frac{c/2}{(1+y)^{T_j-t-1}} - (T_n - t)\frac{1}{(1+y)^{T_n-t-1}}\right] N$$

MaCaulay duration is “scaled” interest rate sensitivity measure:

$$D_{\text{MaCaulay}} = -\frac{dP}{dy}\frac{1+y}{P} = \frac{1}{P}\left[\sum_{j=1}^{n} (T_j - t)\frac{c/2}{(1+y)^{T_j-t}} + (T_n - t)\frac{1}{(1+y)^{T_n-t}}\right] N$$

Interpretation 1: Note that it is also elasticity:

$$-\frac{dP}{dy}\frac{1+y}{P} = -\frac{{}^{dP}\!/_{P}}{{}^{d(1+y)}\!/_{1+y}} = -\frac{\%\Delta \text{ in } P}{\%\Delta \text{ in } 1+y}$$

Interpretation 2: Note that it is also weighted average of coupon payment times:

$$\begin{aligned}&\frac{1}{P}\left[\sum_{j=1}^{n} (T_j - t)\frac{c/2}{(1+y)^{T_j-t}} + (T_n - t)\frac{1}{(1+y)^{T_n-t}}\right] N \\ &= \sum_{j=1}^{n} (T_j - t) w_j\end{aligned}$$

and $\sum_{j=1}^{n} w_j = 1$

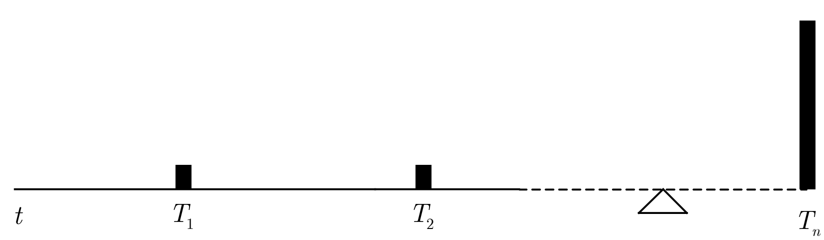


Figure 5.2: Duration as Weighted Average of Coupon Payment Times

Interpretation 3: In physics, MaCaulay has an interpretation of mass center:

Note that in text books and for the exams, we use integer periods to simplify the calculation: annual coupons:

$$P = \left[\frac{c}{(1+y)} + \frac{c}{(1+y)^2} + \cdots + \frac{1+c}{(1+y)^n} \right] N$$

semi-annual coupons:

$$P = \left[\frac{c/2}{(1+y/2)} + \frac{c/2}{(1+y/2)^2} + \cdots + \frac{1+c/2}{(1+y/2)^{2n}} \right] N$$

any arbitrary frequency:

$$P = \left[\frac{c/m}{(1+y/m)} + \frac{c/m}{(1+y/m)^2} + \cdots + \frac{1+c/m}{(1+y/m)^{mn}} \right] N$$

MaCaulay duration:

$$D_{\text{MaCaulay}} = \frac{1}{m} \frac{1}{P} \left[\sum_{j=1}^{mn} j \frac{c/m}{(1+y/m)^j} + mn \frac{1}{(1+y)^{mn}} \right] N$$

Fisher-Weil Duration

The Fisher-Weil duration is similar to the MaCaulay duration. The difference is that Fisher-Weil duration allows the yield curve to be non-flat.

Partial Duration (Key Rate Duration)

Move one rate at a time. These are zero rates. To calculate a key rate duration, we simply bump up (or down) the key rate by 1 basis point and compute the price impact.

Effective Duration

Effective duration measures the price change of the entire yield curve change (parallel shift). The computation is a triangular method. It makes sure that the sum of all key rate durations is equal to the effective duration.

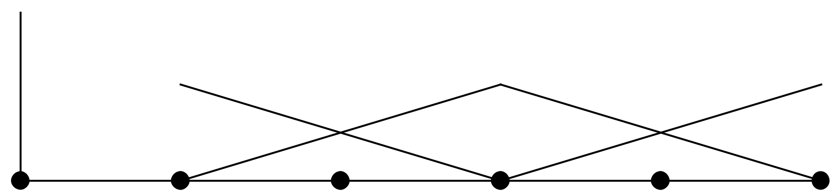


Figure 5.3: Triangular Rule for Key Rate Duration

PV01 (DV01)

DV01 or PV01 simply means price impact of 1 basis point (01) move in "interest rate" whatever that "interest rate" is. Hence it is easier to interpreted it as the key rate. In the case of fixed income securities, this can be interpreted as the key rate duration. (And if move all key rates, then same as effective duration.)

This can be used for other securities as well. In the case of an option this is the Greek letter Rho.

Stochastic Duration

A stochastic duration is calculated when a specific interest rate model is employed to model the yield curve. The sensitivity is not measured with respect to a chosen interest rate from the yield curve but to the risk factor in the theoretical model.

Duration Calculation Using the Ho-Lee Model

See Excel.

5.3.3 IR Swaps

An IRS is (fixed-floating):

$$N \sum_{i=1}^{n} (\ell_i - w_0) P_{0,T_i}$$

where ℓ is the LIBOR rate (usually the index rate in an IRS) and w_0 is today's swap rate.

A deal is made if the following equilibrium is reached:

$$0 = N \sum (\mathbb{E}[\ell_i] - w_0) P_{0,T_i} \tag{5.8}$$

Solving for the swap rate w_0, we get:

$$\begin{aligned} w_0 &= \frac{\sum_{i=1}^n P_{0,T_i} \mathbb{E}[\ell_i]}{\sum_{i=1}^n P_{0,T_i}} \\ &= \frac{\sum_{i=1}^n P_{0,T_i} f_i}{\sum_{i=1}^n P_{0,T_i}} \end{aligned} \tag{5.9}$$

which indicates that the swap rate is a weighted average of future "forward rates". An alternative formula is also often used (less intuitive, but easier to compute):

$$\begin{aligned} w_0 &= \frac{\sum_{i=1}^n P_{0,T_i} f_i}{\sum_{i=1}^n P_{0,T_i}} \\ &= \frac{\sum_{i=1}^n P_{0,T_i}(1 + f_i)}{\sum_{i=1}^n P_{0,T_i}} - 1 \\ &= \frac{\sum_{i=1}^n P_{0,T_i} \frac{P_{0,T_{i-1}}}{P_{0,T_i}}}{\sum_{i=1}^n P_{0,T_i}} - 1 \\ &= \frac{\sum_{i=1}^n P_{0,T_{i-1}} - \sum_{i=1}^n P_{0,T_i}}{\sum_{i=1}^n P_{0,T_i}} \\ &= \frac{1 - P_{0,T_n}}{\sum_{i=1}^n P_{0,T_i}} \end{aligned} \tag{5.10}$$

A Two-Period Example

To get the result of the swap rate being a weighted average of forward rates, we examine the following simple (2pd) example.

2-Period Example		
now	one yr later	two yrs later
long swap: cost 0	$\ell_1 - w_0$	$\ell_2 - w_0$
short 1y fwd: cost 0	$f_1 - \ell_1$	nothing happens
short 2y fwd: cost 0	nothing happens	$f_2 - \ell_2$
short $(f_1 - w_0)$ of $P_{0,1}$	$-(f_1 - w_0)$	nothing happens
short $(f_2 - w_0)$ of $P_{0,2}$	nothing happens	$-(f_2 - w_0)$
$(f_1 - w_0) \times P_{0,1}$ $+(f_2 - w_0) \times P_{0,2}$	0	0

Given that the portfolio generates no cash flow in the future, the value of the portfolio today must also be 0 to avoid arbitrage. As a result,

$$-(f_1 - w_0) \times P_{0,1} - (f_2 - w_0) \times P_{0,2} = 0$$
$$w_0 = \frac{f_1 P_{0,1} + f_2 P_{0,2}}{P_{0,1} + P_{0,2}}$$

An alternative way to form a risk-free portfolio (better way because prices as opposed to rates are used):

2-Period Example		
now	one yr later	two yrs later
long swap: cost 0 short $(1 + \ell_1 - w_0)$ of $P_{0,1}$ long $(1 + w_0)$ of $P_{0,2}$	$\ell_1 - w_0$ $-(1 + \ell_1 - w_0)$ short $(1/P_{1,2})$ of $P_{1,2}$	$\ell_2 - w_0$ nothing happens $(1 + w_0) - (1 + \ell_2)$
$(1 + \ell_1 - w_0) \times P_{0,1} - (1 + w_0) \times P_{0,2}$	0	0

Note that $1 + \ell_1 = \frac{1}{P_{0,1}}$ and $1 + \ell_2 = \frac{1}{P_{1,2}}$. As a result, we have:

$$w_0 = \frac{1 - P_{0,2}}{P_{0,1} + P_{0,2}}$$

The term structure is such that :one-year rate is 5% and two-year rate is 7% (hence the forward rate is 9%) discount factors are:

$$P_{0,1} = \frac{1}{1.05} = 0.9524$$
$$P_{0,2} = \frac{1}{1.07^2} = 0.8734$$

Hence, the swap rate is:

$$w_0 = \frac{P_{0,1}\ell_1 + P_{0,2}f_2}{P_{0,1} + P_{0,2}} = \frac{0.9524 \times 0.05 + 0.8734 \times 0.09}{0.9524 + 0.8734} = \frac{0.1262}{1.8258} = 6.91\%$$

PV01

An IRS is "at the money" at inception. But over time it can be either in or out of the money, as the swap rate moves up or down. The value is equal to:

$$V_t = N \sum (w_t - w_{t-1}) P_{t,T_i}$$

which is a result of doing a reverse swap. Hence the PV01 is:

$$V_t = N \sum (w_t - w_0) P_{t,T_i}$$
$$\frac{\partial V_t}{\partial w_t} = N \sum P_{t,T_i}$$

which is a risk-free annuity.

Take the HL model as an example, the PV01 of a 4-year swap is $0.9524 + 0.8900 + 0.8278 + 0.7686 = 3.4388$. This PV01 (and swap rate and swap value) can be computed for the future as well. For example, the current swap rate is 6.73% (which is equal to $(1 - 0.7686)/3.4388$). The swap rate for the next year is either 14.81% or 3.37% respectively. And hence the swap value is either 18.56% of the notional (in the money) or −9.48% (out of the money).

5.4 FX Risk Management

So far we have only discussed interest rate risk within the domestic economy. Yet many global conglomerates have exposures to other currencies. In an open economy (i.e. multiple nations), we have interest risks under various currencies. For simplicity, here we only discuss foreign exchange (FX hereafter) risk under two nations.

As before, we label domestic discount factors (or risk-free zero coupon bond price) $P(t,T)$. In parallel, we now label foreign discount factors $P^*(t,T)$. Note within each nation, it is a closed economy and hence analyses for $P(t,T)$ can readily apply to $P^*(t,T)$. Finally, the exchange rate is labeled as $e(t)$ which is random.

The following are quick snapshots of the market.

5.4.1 FX Forward and Interest Rate Parity

Interest Rate Parity (IRP) states that a forward exchange rate must bring domestic and foreign discount factors in sync. In other words, a domestic deposit (which earns the domestic interest – $1/P(t,T)$) and a foreign deposit (which earns the foreign interest – $1/P^*(t,T)$) can jointly determine the forward FX rate (which is a forward contract on the FX rate).

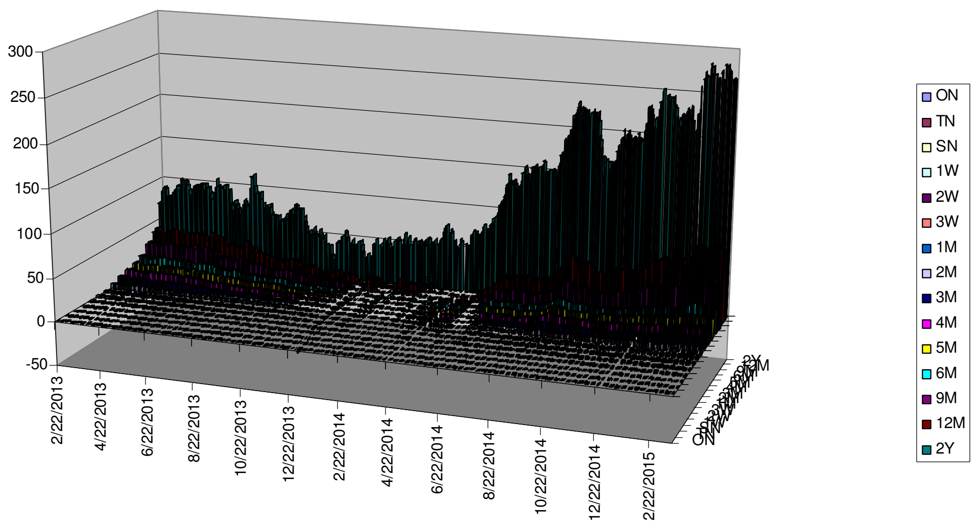


Figure 5.4: 3D Plot of Recent FX Swap Rates

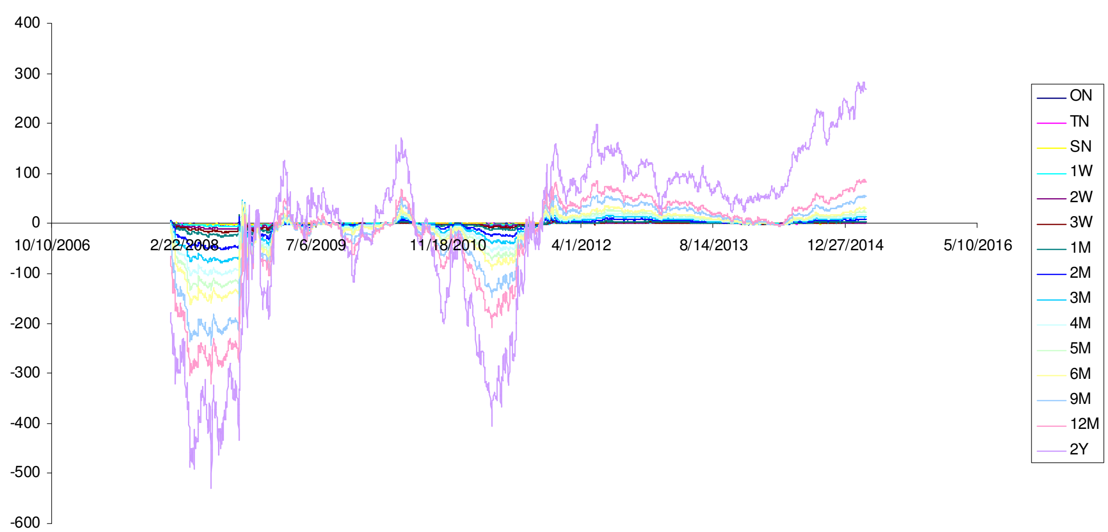


Figure 5.5: 2D Plot of FX Swap Rates since the Lehman Crisis

$$f(t,T) = e(t)\frac{P^*(t,T)}{P(t,T)},[4]$$

FX forward is closely tied with FX swap. Same way how IRS is tied to forward interest rate. Expectations taken are usually forward expectations.

Do not confuse IRP with PPP (Purchasing Power Parity). PPP decides the current FX rate and IRP decides the forward FX rate. But note that usually we adopt a reduced-form approach and directly assume a stochastic process for the FX rate and do not pursue the fundamentals (i.e. PPP) of how FX rates are determined.

5.4.2 FX Swaps

Using FX forward rates, we can not determine FX swap rates. The swap contract must be 0 value at inception. A fixed-fixed FX swap is as follows:

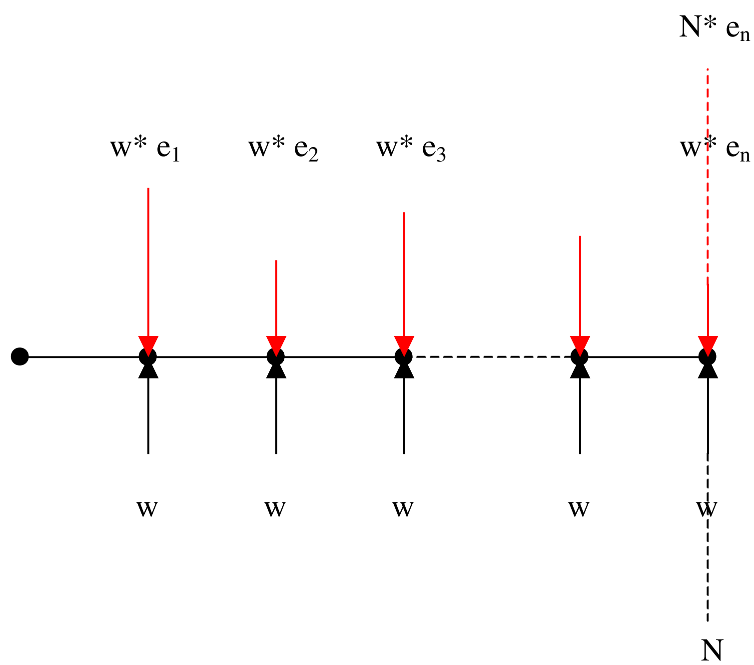


Figure 5.6: Fixed-fixed FX Swap

[4]Under constant interest rates r and r^*, we have $f(t,T) = e(t)e^{(r-r^*)(T-t)}$. This is commonly known as interest rate differential.

Hence, traders must set w and w^* in such a way that the following holds:

$$0 = \sum\nolimits_{i=1}^{n} P(t, T_i)\left(w_t^* N^* \mathbb{E}[e(T_i)] - w_t N\right) + P(t, T_n)\left(N^* \mathbb{E}[e(T_n)] - N\right)$$

where N is notional in domestic currency and N^* is notional in foreign currency, similarly w_t and w_t^* are domestic fixed and foreign fixed rates respectively. Usually notionals are set by $(N^* e(t) - N) = 0$.

It turns out that (known as the forward expectation – see the Appendix for a brief discussion of the forward expectation):

$$\mathbb{E}_t[e(T_i)] = f(t, T_i)$$

which is the forward FX rate. Then:

$$\sum\nolimits_{i=1}^{n} P(t, T_i)\left(w_0^* N^* f(t, T_i) - w_0 N\right) + P(t, T_n)\left(N^* f(t, T_n) - N\right) = 0$$
$$w_0 N = w_0^* N^* \frac{\sum_{i=1}^{n} P(t, T_i) f(t, T_i)}{\sum_{i=1}^{n} P(t, T_i)} + \frac{P(t, T_n)\left(N^* f(t, T_n) - N\right)}{\sum_{i=1}^{n} P(t, T_i)}$$

We can see that if $N^* f(t, T_n) = N$, then the relative fixed rates, i.e. $w_0 N / w_0^* N^*$, are set as in IRS which is the weighted average of forward FX rates.

An Example – Fixed-Fixed FX Swap

A fixed-fixed FX swap is £100 for \$150 as the current exchange rate is 1:1.5 ($e_0 = 1.5$). In each country the interest rate is 5% ($w_0^* = w_0$). Note that

$$N^* e_0 - N = £100 \times 1.5 - \$150 = 0$$

as the initial condition. The exchange of the currencies is 5 for \$7.5 annually.

The actual cash flows (cash flows are fixed but exchange rates are random) are:

$$\sum\nolimits_{i=1}^{n} (w_0^* N^* e_i - w_0 N_0) P_{1,T_i} + (N^* e_n - N_0) P_{1,T_n}$$

One year later the exchange rate changes to 1:2 ($e_1 = 2$). Then a new FX swap will be 5 for \$10. In this case, $N_1 = e_1 N^* = \$200$. In this case, we can do a reverse swap:

$$-\left\{\sum\nolimits_{i=2}^{n}(w_0^* N^* e_i - w_0 N_1)P_{2,T_i} + (N^* e_n - N_1)P_{2,T_n}\right\}$$

which will cancel the previous swap on the pound leg and lead to the following net:

$$\begin{aligned} V_1 &= \sum\nolimits_{i=2}^{n} w_0(N_1 - N_0)P_{2,T_i} + (N_1 - N_0)P_{2,T_n} \\ &= \sum\nolimits_{i=2}^{n} w_0(\bar{e}_1 N^* - \bar{e}_0 N^*)P_{2,T_i} + (\bar{e}_1 N^* - \bar{e}_0 N^*)P_{2,T_n} \\ &= N^* \left\{\sum\nolimits_{i=2}^{n} w_0(\bar{e}_1 - \bar{e}_0)P_{2,T_i} + (\bar{e}_1 - \bar{e}_0)P_{2,T_n}\right\} \end{aligned}$$

init		150	100	1.5					
		dollar leg	pound leg	ex rate					
	0								
	1	-7.5	5		200	100	2		
	2	-7.5	5		10	-5		2.5	0
	3	-7.5	5		10	-5		2.5	0
	4	-7.5	5		10	-5		2.5	0
	5	-7.5	5		10	-5		2.5	0
	6	-7.5	5		10	-5		2.5	0
	7	-7.5	5		10	-5		2.5	0
	8	-7.5	5		10	-5		2.5	0
	9	-7.5	5		10	-5		2.5	0
	10	-157.5	105		210	-105		52.5	0
yield		5%	5%		5%	5%			

Figure 5.7: Same w

Fixed-Floating FX Swap

A fixed-floating FX swap is more popular but more complex as it involves both random exchange rates and random foreign interest rates. Using the same derivation as in fixed-fixed, we have:

init	150	100	1.5					
	dollar leg	pound leg	ex rate					
0								
1	-12	5		150	100	1.5		
2	-12	5		15	-5		3	0
3	-12	5		15	-5		3	0
4	-12	5		15	-5		3	0
5	-12	5		15	-5		3	0
6	-12	5		15	-5		3	0
7	-12	5		15	-5		3	0
8	-12	5		15	-5		3	0
9	-12	5		15	-5		3	0
10	-162	105		165	-105		3	0
yield	8%	5%		10%	5%			

Figure 5.8: Different w and w*

$$\begin{aligned}
&\hat{\mathbb{E}}_t\left[\sum\nolimits_{i=1}^{n}\Lambda(t,T_i)\left(x_i^*N^*e_i-w_0N\right)+\Lambda(t,T_n)\left(N^*e_n-N\right)\right]\\
&=\sum\nolimits_{i=1}^{n}P(t,T_i)\left(N^*\tilde{\mathbb{E}}_t^{(i)}[x_i^*e_i]-w_0N\right)+P(t,T_n)\left(N^*\tilde{\mathbb{E}}_t^{(n)}[e_n]-N\right)=0
\end{aligned}$$

where $\hat{\mathbb{E}}_t$ is risk-neutral expectation and $\tilde{\mathbb{E}}_t$ is forward expectation.

But now we have to deal with the covariance between x_i^* and e_i. This is known as the "quanto" effect in FX analyses. We shall discuss this effect later. Note that if the foreign interest rates (i.e. x^*) and the exchange rate (i.e. e) are not correlated, then the expectation can be separated, and then there is no quanto effect. Hence, quanto is a result of correlation between x^* and e.

To avoid the problem, we keep all foreign cash flows in the foreign country and only discount them once:

$$\begin{aligned}
&e_0\hat{\mathbb{E}}_t^*\left[\sum\nolimits_{i=1}^{n}\Lambda^*(t,T_i)x_i^*N^*+\Lambda^*(t,T_n)N^*\right]-\hat{\mathbb{E}}_t\left[\sum\nolimits_{i=1}^{n}\Lambda(t,T_i)\left(w_0N\right)+\Lambda(t,T_n)\left(N\right)\right]=0\\
&e_0\left[\sum\nolimits_{i=1}^{n}P^*(t,T_i)f_i^*+P^*(t,T_n)\right]N^*-\left[\sum\nolimits_{i=1}^{n}P(t,T_i)w_0+P(t,T_n)\right]N=0\\
&w_0=\frac{1}{\sum_{i=1}^{n}P(t,T_i)}\left\{e_0\sum\nolimits_{i=1}^{n}P^*(t,T_i)f_i^*+\left[e_0P^*(t,T_n)N^*-P(t,T_n)N\right]\right\}
\end{aligned}$$

This way, we by pass the quanto issue.

Floating-Floating FX Swap

Usually there is a fixed spread on one of the sides of the swap so that the swap contract has 0 value. If no spread is specified, then it is known as the differential swap (or diff swap) and in such cases there WILL BE a cash amount exchanged at inception.

$$\begin{aligned}
&\hat{\mathbb{E}}_t\left[\sum\nolimits_{i=1}^{n}\Lambda(t,T_i)\left(x_i^*N^*e_i-(x_i+s)N\right)+\Lambda(t,T_n)\left(N^*e_n-N\right)\right]\\
&=\sum\nolimits_{i=1}^{n}P(t,T_i)\left(N^*\tilde{\mathbb{E}}_t^{(i)}[x_i^*e_i]-(\tilde{\mathbb{E}}_t^{(i)}[x_i]+s)N\right)+P(t,T_n)\left(N^*\tilde{\mathbb{E}}_t^{(n)}[e_n]-N\right)=0
\end{aligned}$$

This is too complicated. So we will not convert till end.

$$e_0\hat{E}^*\left[\sum\nolimits_{i=1}^{n}\Lambda^*(t,T_i)x_i^*N^*+\Lambda^*(t,T_n)N^*\right]-\hat{E}\left[\sum\nolimits_{i=1}^{n}\Lambda(t,T_i)\left((x_i+s)N\right)+\Lambda(t,T_n)\left(N\right)\right]=0$$
$$e_0\left[\sum\nolimits_{i=1}^{n}P^*(t,T_i)f_i^*+P^*(t,T_n)\right]N^*-\left[\sum\nolimits_{i=1}^{n}P(t,T_i)(f_i+s)+P(t,T_n)\right]N=0$$
$$s=\frac{1}{\sum_{i=1}^{n}P(t,T_i)}\left\{e_0\sum\nolimits_{i=1}^{n}P^*(t,T_i)f_i^*-\sum\nolimits_{i=1}^{n}P(t,T_i)f_i+\left[e_0P^*(t,T_n)N^*-P(t,T_n)N\right]\right\}$$

If $e_0P^*(t,T_n)N^*-P(t,T_n)N=0$, then

$$s=\frac{e_0\sum_{i=1}^{n}P^*(t,T_i)f_i^*-\sum_{i=1}^{n}P(t,T_i)f_i}{\sum_{i=1}^{n}P(t,T_i)}$$

which is that the spread is simply the fixed-floating FX swap rate minus the corresponding domestic swap rate.

PV01

FX PV01 is:

$$\frac{\partial V_1}{\partial \bar{e}_1}=N^*\left\{\sum\nolimits_{i=2}^{n}w_0P_{2,T_i}+P_{2,T_n}\right\}$$

which is the price of a domestic bond (fixed rate at w0) with a foreign currency notional.

A FX swap can be coupled with an IR swap. That is, to swap out of the fixed rate w. This way the FX swap will be free from domestic IR risk. In that case, there is an IR PV01 in the combined deal. And that IRS will have a PV01 described above which can be combined with the following PV01:

$$\frac{\partial V_1}{\partial w_0}=N^*\left\{\sum\nolimits_{i=2}^{n}(\bar{e}_1-\bar{e}_0)P_{2,T_i}\right\}$$

Say 100 pounds for \$150 and each year 5 pounds for \$7.5. Now it is 5 pounds for \$8 (wt to 5.3%)

Swap rates are different in two different countries.

See demonstration

5.4.3 Quanto

The quanto effect exists when the exchange rate and the two interest rates are all random and correlated. To see this more clearly, we use the Nikkei option[5] as an example.

Define a "variable rate" call option that pays $\max\{S_T^* - K, 0\}$ in yen. Then in dollars, it is $Y_T \max\{S_T^* - K, 0\}$ at time T. Now we can simply discount it back at the domestic rate r:

$$\begin{aligned} C_t^* &= e^{-r(T-t)}\hat{\mathbb{E}}_t[Y_T \max\{S_T^* - K, 0\}] \\ &= e^{-r(T-t)}e^{(r-r^*)(T-t)}Y_t\hat{\mathbb{E}}_t^{(Y)}[\max\{S_T^* - K, 0\}] \\ &= e^{-r^*(T-t)}Y_t\left[e^{r^*(T-t)}S_t^*N(d_+) - KN(d_-)\right] \\ &= Y_t\left[S_t^*N(d_+) - e^{-r^*(T-t)}KN(d_-)\right] \end{aligned} \tag{5.11}$$

where

$$d_\pm = \frac{\ln S_t^* - \ln K + (r^* \pm 1/2 v^*)(T-t)}{\sqrt{v^*(T-t)}}$$

and hence there is no quanto effect in the option price.

Define a "fixed rate" call that pays $\bar{Y}\max\{S_T^* - K, 0\}$. Then,

$$\begin{aligned} \bar{C}_t &= e^{-r(T-t)}\bar{Y}\hat{\mathbb{E}}_t[\max\{S_T^* - K, 0\}] \\ &= e^{-r^*(T-t)}\bar{Y}\left[e^{(r^*+\rho\sigma_Y v^*)(T-t)}S_t^*N(d_+) - KN(d_-)\right] \end{aligned} \tag{5.12}$$

where

$$d_\pm = \frac{\ln S_t^* - \ln K + (r^* \pm 1/2 v^{*2})(T-t)}{v^*\sqrt{T-t}}$$

and hence the quanto effect exists.

5.4.4 FX Option Formula

Consider the interest rate parity theorem and a simple Black-Scholes type model:

$$\begin{aligned} \frac{de}{e} &= (r - r^*)dt + \sigma dW_e \\ dr &= \kappa(\theta - r)dt + \gamma dW \\ dr^* &= \kappa^*(\theta^* - r^*)dt + \gamma^* dW^* \end{aligned}$$

[5] The Nikkei index put option is ...

and $dW_i dW_j = \rho_{ij} dt$ and $i, j = e, r, *$. The IRP theorem restricts the drift of the exchange rate to be the difference between the two rates. (Note alternatively we can view the foreign interest rate as a "dividend" that takes away the return from investing in the domestic risk-free bond.)

We also modify the Black-Scholes option formula as follows:

$$C_t = P^*_{t,T} e_t N(d_1) - P_{t,T} K N(d_2)$$

where

$$\begin{aligned}
d_1 &= \frac{\ln e_t - \ln K - \ln P_{t,T} + \ln P^*_{t,T} + v^2/2}{\sqrt{v}} \\
d_2 &= d_1 - v\sqrt{T-t} \\
v^2 &= \mathrm{var}[\ln e_T - \ln D_{t,T} + \ln D^*_{t,T}] \\
D_{t,T} &= \exp\left\{ - \int_t^T r_u du \right\} \\
D^*_{t,T} &= \exp\left\{ - \int_t^T r^*_u du \right\}
\end{aligned}$$

5.4.5 FX Basis

An FX basis is the discrepancy between the IRP and actual IR differential. In other words, an FX basis is where the market disagrees with the theory. In theory there will exist arbitrage profits but due to market frictions, such profits are too small to be worth taking.

One key aspect of FX bases is that they are random. This adds an additional risk to consider and hedge/manage.

5.4.6 FX models

First we review the Black-Scholes model. In the case of stocks, we have:

$$\frac{dS}{S} = (\mu - \ell)dt + \sigma dW \tag{5.13}$$

where μ is the expected return of the stock and ℓ is the (continuous) dividend yield.[6] In a general case, d can be regarded as any leakage of return from the underlying

[6] We note that under discrete dollar dividends, the model is quite different.

asset. In the case of FX, it is the foreign risk-free rate; in the case of real estate, it is rent; in the case of commodities, it is convenience yield; and finally in the case of futures, it is risk-free rate itself.

The FX rates are also modeled the same way as (5.13):

$$\frac{dX}{X} = (r_D - r_F)dt + \sigma dW \tag{5.14}$$

where r_D and r_F are domestic and foreign risk-free rates respectively. Compared with equation (5.13), the above equation substitutes r_D for μ and r_F for ℓ. This result is a consequence of what is known as the interest parity theorem. An investor can choose to put his or her money in a domestic risk-free account (making the domestic risk-free return) or a foreign risk-free account (making the foreign risk-free return). Since the investor can only choose one or the other, by choosing one, he or she has to let go of the other. For the domestic point of view, the foreign risk-free return is a "leak".

Another way to look at this is the notion of opportunity cost. By investing domestically, a foreign opportunity is let go and hence it must be considered as a loss (just like dividends) to the "return" on the exchange rate. Such a notion is applied repeatedly in other asset classes (convenience yield in commodities and rent in real estate).

The FX option (e.g. call) hence is evaluated similar to the Black-Scholes formula as in equity as follows:

$$C_t = e^{-r_F(T-t)} X_t N(d_1) - e^{-r_D(T-t)} K N(d_2) \tag{5.15}$$

The binomial model needs to be modified as shown in Figure 5.9.

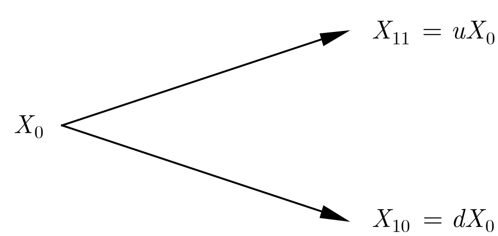


Figure 5.9: Binomial Model for the FX Rate

where the probabilities are modified as:

$$\hat{p} = \frac{e^{(r_D - r_F)\Delta t} - d}{u - d} \tag{5.16}$$

and r_D and r_F are domestic and foreign risk-free rates respectively.

5.5 Total Return Swap

A total return swap (TRS) is usually a floating-floating swap, or diff swap. A typical example is a swap between an equity index (e.g. S&P 500) and an interest rate index (e.g. LIBOR). As in any diff swap, there is a spread on the IR leg.

Another popular TRS can be a real estate index (e.g. NCREIF) for an IR index.

5.6 Residential Mortgage

Residential mortgages (usually refer to Prime mortgages) are guaranteed by government agencies known as GNMA (Ginnie Mae), FNMA (Fannie Mae), and FHLMC (Freddie Mac).[7][8]

Residential mortgages are exposed to prepayment risk. Prepayment is an action taken by mortgage borrowers to pay pre-maturely full or partial amount of the loan. Prepayments affect the incomes of the lenders. Simply speaking, prepayments reduce the loan amounts and hence reduce the interests earned by the lenders. As a result, prepayments represent a risk. However, not all prepayments are bad news. When prepayments occur in a time when interest rates are high, then banks can loan out the early payment amounts at a higher rate and make more interest incomes. On the contrary, when prepayments occur in a time when interest rates are low, banks directly lose interest earnings.

If the reason of a prepayment is non-economical, such as divorce, job change, addition to family, then prepayments can be either good or bad news. However, if a prepayment is due to re-finance of the mortgage, then it is surely bad for the bank. Hence, the models for mortgages are categorized into two big areas: prepayment and refinance.

[7] They are Government National Mortgage Association, Federal National Mortgage Association, and Federal Home Loan Mortgage Corporation respectively.

[8] Note that there are a small number of residential mortgages that are not guaranteed by the three agencies. These mortgages are not included in this section.

5.6.1 Refinance Modeling

Refinancing is a decision directly linked to mortgage rates, which in turn related to Treasury interest rates (most commonly the 10-year rate). As rates drop, borrowers substitute new mortgages that have lower mortgage rates for existing mortgages that have higher mortgage rates. And as borrowers refinance, banks lose future interest revenues. In other words, the values of mortgages that banks make decrease.

As refinancing is a predictable event that happens when rates drop, we can then run an interest rate model to forecast and gauge the impacts of future rate drops. Then the impacts can be factored into today mortgage rate. For example, if rates are expected to fall and refinancing is highly expected, then banks should charge a higher mortgage rate to compensate for the future loss. Figure 5.10 demonstrates how refinance actions are modeled in an interest rate model. As Figure 5.10 demonstrates, the lower part of the model is the refinance area as rates are low.

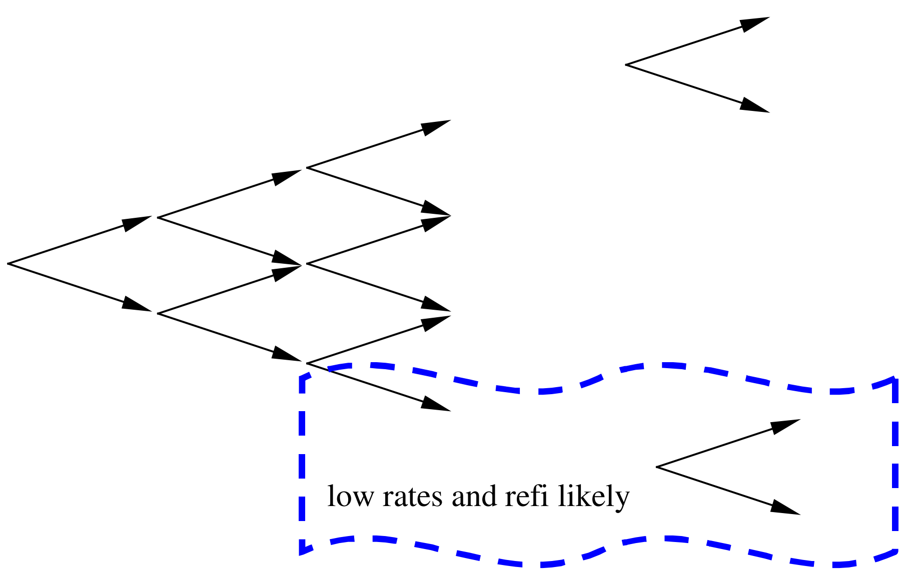


Figure 5.10: An Interest Rate Model

5.6.2 Prepayment Modeling

Prepayment is usually modeled with some kind of "response function" which is graphically similar to an "S" curve. An "S" function is bounded between 0 and 1 and can be modeled via a number of different functional forms such as arc-tangent, logit, and probit. In the mortgage area, the logit function is the most popular choice.[9]

The logit function is:

$$p = \frac{\exp(\alpha + \sum \beta_k x_k)}{1 + \exp(\alpha + \sum \beta_k x_k)} \tag{5.17}$$

where the dependent variable p is the prepayment rate and the explanatory variables are chosen as follows.

- Housing Turnover Rate (When a house is bought and sold, the existing mortgage is usually prepaid completely and a new mortgage is taken (unless the existing mortgage is assumable). As a result, higher turnover rates for a particular housing sector implies higher prepayment rate and hence higher risk for investors. For example, apartments are turned over faster then single homes.)

- Seasonality (Housing markets present seasonality – more people are buying homes in summer than in winter. As mentioned above, housing turnover is directly linked to prepayment. Hence summer prepayment rate is higher than that in winter.)

- Cash-out (This is how much cash is left in the mortgage. As the property appreciates, the difference between house value and mortgage gets larger, and the cash-out effect becomes stronger.)

- Refinance (This is to be discussed separately later. Basically, refinance occurs in low interest rate times.)

- Age (This is the age of the mortgage, not the age of the borrower. The age of a mortgage has a negative impact of prepayment. In other words, the longer a homeowner holds on to the mortgage, the less likely he or she will prepay early.)

- Burnout (In a mortgage-backed security, the percentage of mortgages underlying the security that were not re-financed following a drop in interest rates. Historically, a mortgage that is not re-financed is less likely to re-finance at the next drop in interest rates. As a result, potential investors in mortgage-backed securities often look for a high burnout rate because it reduces prepayment risk, or the risk that investors will be deprived of future interest payments because too many mortgages are prepaid.)

[9] OTS, or Office of Thrift Supervision, used to use an arc-tangent function for prepayment but it was dissolved in 2011.

- Yield Curve (Interest rates have the most impacts on prepayments (or refinance).)
- Equity (As monthly payments are made, the equity of the house increases. The homeowner now can borrow more now as the LTV improves.)
- Credit (As mortgage is paid off gradually (FRM), credit is improved as LTV is improved.)
- Curtailment (As opposed to prepayment that pays off the mortgage in entirety, a curtailment is to pay more than the required monthly payment. Hence, the effect of curtailments is similar to (but milder than) that of prepayment.)

The most common measure of prepayment speed is SMM (single month mortality) rate. SMM is used in interest rate models to evaluate mortgages. SMM is estimated empirically using the above factors. In other words, SMM is the dependent variable (think of a regression) and the above factors are independent (explanatory) variables. The function used is a response function. The usual shape of a response function is an S curve bounded between 0 and 1. Logit or probit functions (both bounded between 0 and 1) are also plausible functions used for estimating SMM. Arc-tangent functions can be used as well.

The other two measures used in practice are:

- PSA (Public Securities Association) Standards
- CPR (constant prepayment rate)

These are measures of the speed of prepayment and not models of prepayment. However, by monitoring these measures one can gain a sense of how the prepayment behavior. Models of prepayment can be found in many fixed income textbooks.

5.6.3 Default Modeling

Since prime residential mortgages are usually underwritten by the Agencies, they are default-protected by the Agencies. This is the credit-enhancement provided by the Agencies so that investors do not need to worry about defaults. This is one of the major functions the Agencies provide in order to facilitate secondary market trading.

While defaults are not a concern in residential mortgages, they are a true concern for commercial mortgages. Since commercial mortgages are not protected by the Agencies, they are subject to default risk.

However, commercial mortgages do not suffer from prepayment risk in that all commercial loans have a "yield maintenance" clause which protects investors by having a guaranteed yield for a certain period (hence yield is "maintained"). In other words, there is a heavy prepayment penalty which is so high that borrowers have no incentive to prepay. The period under protection is called the yield maintenance period (or YM period)

5.6.4 OAS – Option Adjusted Spread

Spread01 or Spread PV01.

5.7 Combined with Other Assets

How do we compute IR risk for a portfolio of stocks and bonds. Stocks are affected by interest rates but there are no clean models (both are "primary assets"). The common practice is to run regression (one could do PCA but then each PC must be investigated to see if it is sensitive to interest rates) of stock on rates.

$$dS = b_0 + \sum_{k=1}^{K} b_k r_k + \cdots + e$$

Then, $dS/dr_k = b_k$ can be used in the portfolio VaR calculation.

A portfolio of a call and a bond. The call option must be evaluated with the Rabinovitch model (1989) which takes into account of random interest rates. Then we have

$$\begin{aligned} dV &= n_C dC + n_2 dP \\ &= n_1 m_C + n_2 m_P \end{aligned}$$

where

$$m_C = C_S dS + {}^1\!/_2 C_{SS} dS^2 + C_t dt + C_r dr + {}^1\!/_2 C_{rr} dr^2 + C_{Sr} dS dr$$

and

$$m_P = P_r dr + {}^1\!/_2 P_{rr} dr^2 + P_t dt$$

Then

$$
\begin{aligned}
\operatorname{var}[dV] &= n_C^2 \operatorname{var}[m_C] + n_P^2 \operatorname{var}[m_P] + 2n_C n_P \operatorname{cov}[m_C, m_P] \\
&= n_C^2 \left\{ C_S^2 \operatorname{var}[dS] + C_r^2 \operatorname{var}[dr] + 2C_S C_r \operatorname{cov}[dS, dr] \right\} \\
&+ n_P^2 P_r^2 \operatorname{var}[dr] \\
&+ 2n_C n_P \left\{ C_S P_r \operatorname{cov}[dS, dr] + C_r P_r \operatorname{var}[dr] \right\}
\end{aligned}
$$

A convertible bond is another good example. A CB is exposed to three important risks – equity (convert), interest rate (coupons and principal), and credit (default). A convertible bond can be expressed as:

$$
B = \begin{cases} \max\{\Pi, \xi S\} & \text{if survives} \\ R & \text{if defaults} \end{cases}
$$

5.8 Appendix

5.8.1 FX Swap Curve and Fixed-Fixed FX Swaps

If we already know the coupons of domestic and foreign bonds, then we can solve for the FX swap rate as follows.

$$
\begin{aligned}
V_0 = 0 &= \sum (w_0^* N^* \bar{e}_0 - w_0 N) P_{0,T_i} + (N^* \bar{e}_0 - N) P_{0,T_n} \\
&= \sum (w_0^* N^* \bar{e}_0 - w_0 N^* e_0) P_{0,T_i} + (N^* \bar{e}_0 - N^* e_0) P_{0,T_n} \\
&= N^* \sum (w_0^* \bar{e}_0 - w_0 e_0) P_{0,T_i} + (\bar{e}_0 - e_0) P_{0,T_n}
\end{aligned}
$$

The job is to solve for the FX swap rate. It is clear that the FX swap rate can be solved the same way as the IR swap rate.

$$
\begin{aligned}
&\sum_{i=1}^{n} w_0 e_0 P_{0,T_i} + e_0 P_{0,T_n} = \bar{e}_0 \left\{ \sum_{i=1}^{n} w_0^* P_{0,T_i} + P_{0,T_n} \right\} \\
&\bar{e}_0 = \frac{\sum_{i=1}^{n} w_0 e_0 P_{0,T_i} + e_0 P_{0,T_n}}{\sum_{i=1}^{n} w_0^* P_{0,T_i} + P_{0,T_n}}
\end{aligned}
$$

where the similarity remains. However, note that this is not to contradict the result in the text where we solve for the relative coupon rates using the FX forward curve.

5.8.2 Commodities and Real Assets

Some times commodities are included in the category of fixed income and so named FICC (Fixed Income, Currencies and Commodities), although the risk characteristics of commodities are quite different from IR and FX.

Commodities are very difficult to price and also their derivatives. Unlike financial assets, commodities (or real assets) are:

- difficult to transact (hence liquidity is very low)
- require large storage cost (including funding cost)
- usually present cycles (including seasonality)

In commodities, two popular terminologies should be paid attention to:

- contango
- backwardation

Contango is defined as the futures price greater than the spot price: $\Phi(t,T) > S(t)$ where T is the settlement date. This is a normal situation as storage cost is high. For a buyer who wants a commodity in the future, he can either buy futures contracts, or he can buy spot and store it, whichever is cheaper. As a result, $\Phi(t,T) = S(t) + C(t,T) > S(t)$.

Backwardation is defined as the futures price smaller than the spot price: $\Phi(t,T) < S(t)$. Under the phenomenon of cost of carry, this is not possible as arbitrageurs of such commodities will just buy futures and sell short the spot to make profits. Hence the only logical explanation of backwardation existing must be that short-selling commodities is prohibitedly costly. Such a cost is termed convenience yield.

Convenience yield happens (i.e. backwardation) when the spot is very rare and hence short-selling is difficult. To short sell, the seller must borrow the spot. When the spot is rare, then the cost of borrowing is consequently high. There are two kinds of rareness. The first is physical, which means the commodity is simply not available. For example, during winter, agriculture products (e.g. corn) are rare, as farmlands cannot produce. Hence to borrow corn to short must pay a higher price. The other kind is the spot is in high demand and its price is skyrocketing. Under this situation, the borrower must pay for the expected growth in price as part of the borrowing cost. For example, gold (or precious metals) is very expensive to borrow during a recession as everyone buys gold to hedge a recession.

As we can see, convenience yield shares the same flavor as seasonality as spot prices of commodities can present cyclical patterns repeatedly. As the two examples earlier, agricultural products present price cycles within a year (seasonality) and precious metals present price cycles along with recessions. Should commodities be absolutely liquid, such cycles cannot exist as one can buy and sell these goods easily at no cost.

By now, we can understand the classical financial models cannot be applied easily to commodities. All financial models assume perfect liquidity that rules out cyclicality and seasonality. However, convenience yield can be regarded as leakage of the spot because it represents the cost of hold the spot. This is similar to dividends of stocks, or foreign interest returns of exchange rates.

Equation (5.13) is then used as the best approximation to model commodities. Eduardo Schwartz (1979) uses equation (5.13) with a mean-reverting convenience yield process:

$$\begin{aligned}\frac{dS}{S} &= (\mu - \ell)dt + \sigma dW_1 \\ d\ell &= \kappa(\theta - \ell)dt + \gamma dW_2\end{aligned} \tag{5.18}$$

where ℓ represents "convenience yield" and $dW_1 dW_2 = \rho dt$. The convenience yield here now can be either positive or negative. When it is substantially negative (larger than μ so $\mu - \ell < 0$), then the futures price will be smaller than the spot price and we have a backwardation. If ℓ is negative, then it is similar to having a contango. Given that ℓ is normally distributed, there is a closed-form solution to the futures price (and futures option):

$$\Phi(t,T) = \hat{\mathbb{E}}[S_T] = \exp\left\{\hat{\mathbb{E}}[\ln S_T] + \frac{1}{2}\hat{\mathbb{V}}[\ln S_T]\right\}$$

where

$$\begin{aligned}\hat{\mathbb{E}}[\ln S_T] &= \ln S_t + (\mu - \tfrac{1}{2}\sigma^2)(T-t) - \left[\ell_t \frac{1}{\kappa}\left(e^{-\kappa(T-t)} - 1\right) + \theta\left(T - t - \frac{1}{\kappa}\left(e^{-\kappa(T-t)} - 1\right)\right)\right] \\ \hat{\mathbb{V}}[\ln S_T] &= \frac{\gamma^2}{2\kappa}\left(e^{2\kappa(T-t)} - 1\right) + \sigma^2(T-t) + \frac{2\rho\gamma\sigma}{\kappa}\left(e^{\kappa(T-t} - 1\right)\end{aligned}$$

Another important pricing question related to commodities (and not so much for financial assets) is the level of inventory. As storing commodities suffers (enjoys) high storage costs (convenience yield), how much to store (level of inventory) is an important decision. William J. Baumol (1952) has a simple model to explain the demand for inventory.[10] Say farmers hold Q bushels of wheat of which Q_1 to be sold at time T_1 and Q_2 to be sold at time T_2. P_1 and P_2 are prices respectively. If $P_1 > \mathrm{PV}[P_2]$, then farmers will sell all Q at T_1.

A marketing cost is assumed as ξQ^2. Also $\mathrm{PV}[P_2] = DP_2$. So the total revenue from both sales is $(P_1Q_1 - \xi Q_1^2) + D(P_2Q_2 - \xi Q_2^2)$. Maximizing the total revenue leads to (substituting $Q - Q_2$ for Q_1 and taking first order derivative with respect to Q_2):

$$-P_1 + 2\xi(Q - Q_2) + D(P_2 - 2\xi Q_2) = 0$$

and the optimal solution for Q_2 is:

$$\hat{Q}_2 = \frac{DP_2 - P_1 + 2\xi Q}{2\xi(1 + D)}$$

which implies a positive inventory (i.e. $Q_2 > 0$) if:

$$DP_2 - P_1 > -2\xi Q$$

that is the price differential must be larger enough to justify an inventory. The larger is the price in the second sale (P_2), the higher is the inventory level (Q_2). This is known as the transaction demand for inventory.

A second theory for inventory is known as the precautionary demand for inventory (S. C. Tsiang, 1969).[11] Let z be amount arrived; I precautionary stock; x a unit loss from the shortage; K demand; and c cost of holding an inventory.[12]

If z is too low to operate at full capacity, then the firm will suffer a shortage cost:

$$(K - I - z)x$$

Expected shortage cost:

$$\int_0^{K-I} (K - I - z)xf(z)dz$$

[10] William J. Baumol, The Quarterly Journal of Economics, Vol. 66, No. 4 (Nov., 1952), pp. 545-556.

[11] S. C. Tsiang, Journal of Political Economy, Vol. 77, No. 1 (Jan. - Feb., 1969).

[12] The following derivation is taken from The Economic Function of Futures MarketsOct 27, 1989 by Jeffrey C. Williams.

Cost of precautionary inventory is Ic so the total cost is:

$$Ic + \int_0^{K-I} (K - I - z) x f(z) dz$$

Optimal level of inventory I^* is:

$$c - \int_0^{K-I^*} x f(z) dz = 0$$

The higher is storage cost c, the less should be the inventory (i.e. I^* is smaller and $K - I^*$ is larger). The higher is the opportunity cost x, the lower is the inventory.

Real estate is another real asset that follows the same model. In investing in real estate, the buyer now needs not to rent and therefore return on the property must be reduced by the amount of rent. Now ℓ represents the percentage rent. Another way to look at this is that the buyer buys the property for purely rental purposes. He or she spends the money and expects to gain returns on the investment. Rental incomes therefore must be a part of the total (cum) return. Given that rents are collected (leaked out) in cash, it must be deducted from the total return.

One thing particular to the real estate market is that properties need to be depreciated. As a result, the minimum return for the property to generate is the depreciate rate. Because of this, the convenience yield now ℓ is rent minus depreciation rate.

Figure 5.11: Triangular Rule for Key Rate Duration

			0.0001										0.0001	
					0.0001								0.0001	
							0.0001						0.0001	
									0.0001				0.0001	
											0.0001		0.0001	
1	0.089432	0.91791	0.089532	0.917825	0.089432	0.91791	0.089432	0.91791	0.089432	0.91791	0.089432	0.91791	0.089532	0.917825
2	0.092061	0.838506	0.092061	0.838506	0.092161	0.838353	0.092061	0.838506	0.092061	0.838506	0.092061	0.838506	0.092161	0.838353
3	0.050936	0.861533	0.050936	0.861533	0.050936	0.861533	0.051036	0.861287	0.050936	0.861533	0.050936	0.861533	0.051036	0.861287
4	0.059417	0.793838	0.059417	0.793838	0.059417	0.793838	0.059417	0.793838	0.059517	0.793539	0.059417	0.793838	0.059517	0.793539
5	0.057017	0.75786	0.057017	0.75786	0.057017	0.75786	0.057017	0.75786	0.057017	0.75786	0.057117	0.757502	0.057117	0.757502
		0.966343		0.966339		0.966335		0.96633		0.966328		0.965966		0.965927
				0.042124		0.076771		0.122943		0.149828		3.763077		4.154744

			0.0001										0.0001	
			0.00005				0.00005						0.0001	
							0.0001						0.0001	
							0.00005				0.00005		0.0001	
											0.0001		0.0001	
1	0.089432	0.91791	0.089532	0.917825	0.089432	0.91791	0.089432	0.91791	0.089432	0.91791	0.089432	0.91791	0.089532	0.917825
2	0.092061	0.838506	0.092111	0.838429	0.092061	0.838506	0.092111	0.838429	0.092061	0.838506	0.092061	0.838506	0.092161	0.838353
3	0.050936	0.861533	0.050936	0.861533	0.050936	0.861533	0.051036	0.861287	0.050936	0.861533	0.050936	0.861533	0.051036	0.861287
4	0.059417	0.793838	0.059417	0.793838	0.059417	0.793838	0.059467	0.793688	0.059417	0.793838	0.059467	0.793688	0.059517	0.793539
5	0.057017	0.75786	0.057017	0.75786	0.057017	0.75786	0.057017	0.75786	0.057017	0.75786	0.057117	0.757502	0.057117	0.757502
		0.966343		0.966335		0.966343		0.966319		0.966343		0.965959		0.965927
				0.080512		0		0.236254		0		3.838		4.154744

Chapter 6

P&L Attribution

6.1 Introduction

The P&L attribution is to explain where the profits and losses come from. Due to the fact that luck plays a critical role in trading, managers need to make sure that their traders make money not due to luck but due to skills (or talents). As a result, P&L attribution has become essential in trading and fund management business.

6.2 Taylor's Series Expansion and Explanatory Risk Factors

Greeks are key to explain relative importance of each risk factor. We define the basic Greeks as follows:

$\Delta = \frac{\partial C}{\partial S}$ = partial derivative of target (e.g. call) with respective to the underlying (e.g. stock)

$V = \frac{\partial C}{\partial \sigma}$ = partial derivative of target (e.g. call) with respective to the volatility

$\mathrm{P} = \frac{\partial C}{\partial r}$ = partial derivative of target (e.g. call) with respective to the the interest rate

$\Theta = \frac{\partial C}{\partial t}$ = partial derivative of target (e.g. call) with respective to time (known as time decay)

$\Gamma = \frac{\partial^2 C}{\partial S^2}$ = partial derivative of target (e.g. call) twice with respective to the underlying

Cross Greeks can be defined, as a few examples, as follows:

$$\Delta_V = V_\Delta = \frac{\partial^2 C}{\partial S \partial \sigma} \quad \text{(Vanna)}$$
$$\Gamma_V = \frac{\partial^2 C}{\partial \sigma \partial \sigma} \quad \text{(Volga)}$$
$$\Delta_{\mathrm{P}} = \mathrm{P}_\Delta = \frac{\partial^2 C}{\partial S \partial r}$$
$$\mathrm{P}_V = V_{\mathrm{P}} = \frac{\partial^2 C}{\partial r \partial \sigma}$$

where the first two are popular in FX.

Take a single product as an example (e.g. call option). The explanatory factors are price, volatility, and time which translate to Delta, Gamma, Vega, and Theta. If the BS model is the correct model, then we know that:

$$\begin{aligned} dC &= \left(\mu S C_S + \sigma^2 S^2 C_{SS} + C_t\right) dt + \sigma S C_S dW \\ &= \underbrace{\left(\mu S \Delta + \sigma^2 S^2 \Gamma + \Theta\right)}_{\text{explanatory factors}} dt + \underbrace{\sigma S \Delta dW}_{\text{unexplained}} \end{aligned} \tag{6.1}$$

and there is no Vega. If the volatility and the risk-free rate are both random, then we must first define the volatility and interest rate processes. Usually we write both of them as mean-reverting square root processes:

$$\begin{aligned} dS &= \mu S dt + \sqrt{V} S dW_1 \\ dV &= \alpha(\beta - V)dt + \sqrt{V}\gamma dW_2 \\ dr &= a(b - r)dt + \sqrt{r} g dW_3 \end{aligned} \tag{6.2}$$

where $dW_i dW_j = \rho_{ij} dt$. Then the P&L attribution becomes:

$$
\begin{aligned}
dC &= C_S dS + C_V dV + C_r dr + C_{SS}(dS)^2 + C_{VV}(dV)^2 + C_{rr}(dr)^2 \\
&\quad + C_{SV}(dS)(dV) + C_{Sr}(dS)(dr) + C_{Vr}(dV)(dr) + C_t \\
&= \begin{bmatrix} \mu S C_S + \alpha(\beta - V)C_V + a(b-r)C_r + C_t + \\ V S^2 C_{SS} + \gamma^2 V C_{VV} + g^2 r C_{rr} + V S \gamma \rho_{12} C_{SV} + r S g \rho_{13} C_{Sr} + \sqrt{rV}\gamma g \rho_{23} C_{Vr} \end{bmatrix} dt \\
&\quad + \sqrt{V} S C_S dW_1 + \sigma \sqrt{V} C_V dW_2 + g\sqrt{r} \mathrm{P} dW_3 \\
&= \underbrace{\begin{bmatrix} \mu S \Delta + \alpha(\beta - V)V + a(b-r)\mathrm{P} + \Theta + \\ V S^2 \Gamma + \gamma^2 V V_V + g^2 r \mathrm{P}_{\mathrm{P}} + V S \gamma \rho_{12} \Delta_V + r S g \rho_{13} \Delta_{\mathrm{P}} + \sqrt{rV}\gamma g \rho_{23} V_{\mathrm{P}} \end{bmatrix}}_{\text{explanatory factors}} dt \\
&\quad + \underbrace{\sqrt{V} S \Delta dW_1 + \gamma \sqrt{V} V dW_2 + g\sqrt{r} \mathrm{P} dW_3}_{\text{unexplained}}
\end{aligned} \tag{6.3}
$$

As we can see, the model can expand to as many random factors as we wish. As shown in the exhibit earlier, the random factors can include FX, multiple key interest rates, credit, prepayment, and any other risk factors. At the end, the equation can be very long and the number of parameters can become quite unimaginable. Some simplifications must be necessary.

The first simplification is as easy as just ignoring higher order (and cross) Greeks. We can assume that higher order Greeks such as Vega-Vega and Delta-Vega effects are small and ignore them. The second simplification is to adopt the VaR methodology and choose several key benchmark indices.

6.3 Pictorial P&L Attribution

Figure 6.1 is taken from http://www.pnlexplained.com/ on 12/15/2009 which describes very well the concept of P&L attribution.

At the bottom layer, we see various first order Greeks such as Delta and Vega, and various higher order Greeks such as Gamma and cross Greeks (cross Gamma and Vega-Gamma).

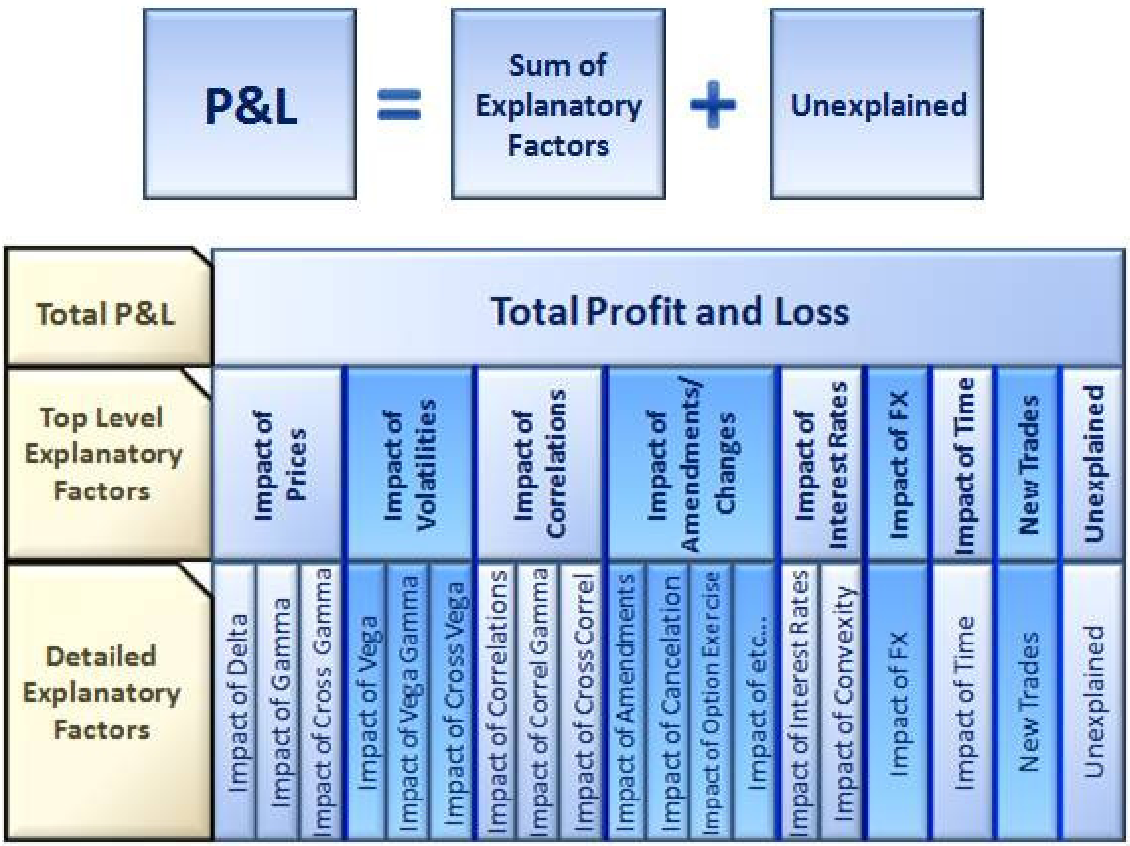


Figure 6.1: P&L Decomposition

Chapter 7

Value Adjustment

7.1 VA – Value Adjustment

As previously discussed, valuation is important in providing reliable risk management. This is particularly so when there are no liquidly observable prices. When prices are not always observable, models are used to estimate those prices. Models make assumptions about the connections between observables in the market place (inputs to the model) to the prices and risk metrics (outputs of the model) of the assets. Complex models are believe to generate more accurate estimates for prices but require a comprehensive set of inputs and often need a long time to compute prices and risk metrics. As a result, these models cannot be used for day-to-day risk management. Simpler models are therefore developed to quickly estimate prices for day-to-day risk management but these estimates can be inaccurate. To balance between the two, a concept of VA (value adjustment) is adopted by the industry.

The idea of a VA is to run simple models every day but periodically check against more comprehensive models. In other words, simple models are "calibrate" to more complex models on a less frequent basis and then run on their own until the next calibration. This constant process of re-calibration has its own become an important methodology which needs to be validated to insure the "calibration" each time is done properly. The complex models themselves are subject to extra scrutiny to insure absolute accuracy in providing proper valuation.

The complex models need to validated thoroughly. This always requires carefully designed econometric methods for parameter estimation (see later), comprehensive back tests and stress tests, sensitivity analyses on parameters, and a full revelation on model risk.

Take a CDO as an example. There are several simple models currently used such as:

- correlation skew
- binomial diversity score

and yet the full model is done over various stochatic input variables via Monte Carlo (with recursive algorithm and Fourier Inversion).

Chapter 8

Parameter Estimation

8.1 Introduction

To use a model, we must first obtain parameter values. There are two ways to estimate parameter values in a model. One is the reduced-form (familiar with the term now?) approach and the other is the econometric approach.

The reduced-form approach is to use current market information to "back out" parameter values. The most famous example is the implied volatility of the option, using the Black-Scholes model. One equation (Black-Scholes option pricing formula) and one unknown (volatility). People use the same method to compute parameter values of other models. For example, we can solve for four parameters of the Vasicek term structure model using four bonds.

The second method is the econometric method. Historical data are collected and the full time series behavior of the model is considered. Hence, theoretically, the econometric method is definitely superior to the reduced-form method. However, given that the parameters are estimated by past prices, hence there is no guarantee that the model can match the current price. This can be a problem for hedging as the hedge ratios will then be incorrect.

8.2 Regression estimation for the Vasicek model

Recall that the Vasicek model is a one-factor mean reverting Gaussian model with the following short rate dynamics

$$dr = \alpha(\mu - r)dt + \sigma dW \tag{8.1}$$

under the real measure and the parameters are defined previously. The mean and variance of r are given as follows:

$$\mathbb{E}[r_s|r_t] = \left(1 - e^{-\alpha(s-t)}\right)\mu + e^{-\alpha(s-t)} r_t \tag{8.2}$$

and

$$\mathbb{V}[r_s|r_t] = \sigma^2 \frac{1 - e^{-2\alpha(s-t)}}{2\alpha} \tag{8.3}$$

for $s \geq t$.

We know that r is normally distributed. Hence, we can write an AR(1) process in discrete time for r:

$$r_t = \left(1 - e^{-\alpha h}\right)\mu + e^{-\alpha h} r_{t-h} + e_t \tag{8.4}$$

where h is time interval and e_t has mean 0 and variance $\frac{1-e^{-2\alpha h}}{2\alpha}\sigma^2$. This essentially an OLS with normally distributed error terms:

$$r_t = a + b r_{t-h} + e_t \tag{8.5}$$

with $a = \left(1 - e^{-\alpha h}\right)\mu$ and $b = e^{-\alpha h}$. We can then run an OLS to obtain the parameters. The regression coefficient, b, can be used to estimate α, and then the constant coefficient, a, can be used for μ. Finally the MSE of the regression model can be used to estimate σ. The t statistics of α and μ can be derived from the t statistics of a and b. With Taylors series expansion, we have:

$$\begin{cases} b \approx 1 - \alpha h \\ a \approx \alpha \mu h \end{cases} \tag{8.6}$$

It then follows that:

$$\begin{cases} \mathrm{var}[b] \approx h^2 \, \mathrm{var}[\alpha] \\ \mathrm{var}[a] \approx (\alpha h)^2 \, \mathrm{var}[\mu] \end{cases} \tag{8.7}$$

and hence the t statistics can be easily calculated. Finally, the asymptotic standard error of the SSE of the regression model is $1/(2\sigma^4)$. Hence, the standard error of the MSE is $1/(2N\sigma^4)$ where N is the sample size.

The problem of the above method is that it fails to estimate the market price of risk. The chosen interest rate series are regarded as instantaneous rate and hence carries no risk premium. Certainly such a substitution, while convenient and often times enough to obtain a major intuition of the interest rate dynamics, is theoretically flawed.

To include the market price of risk into the estimation process, we must realize that any market interest rates are not the instantaneous rate but "term rates" that are usually converted to discount factors $P_{t,T}$ before they can be used for any pricing. In the Vasicek model, this discount factor has a closed-form solution as follows:

$$P_{t,T} = e^{-r_t F_{T-t} - G_{T-t}} \tag{8.8}$$

where

$$\begin{aligned} F_{T-t} &= \frac{1 - e^{-\alpha(T-t)}}{\alpha} \\ G_{T-t} &= \left(\mu - \tfrac{\sigma\lambda}{\alpha} - \tfrac{\sigma^2}{2\alpha^2}\right)(T - t - F_{T-t}) + \frac{\sigma^2 F_{T-t}^2}{4\alpha^2} \end{aligned}$$

where λ is the market price of risk. Inverting the pricing formula gives the instantaneous rate:

$$r_t = -\frac{\ln P_{t,T} + G_{T-t}}{F_{T-t}} \tag{8.9}$$

Note that the above formula must hold regardless of the choice of t. Also note that at the next period, $t+h$, we can write (8.9) as:

$$r_{t+h} = -\frac{\ln P_{t+h,T+h} + G_{T-t}}{F_{T-t}} \tag{8.10}$$

and the functions F and G remain the same as $T + h - (t + h) = T - t$. This result provides tremendous convenience in data collection. We can choose CMT (constant maturity Treasury) rates for the estimation. A 3-month CMT rate have a constant rolling maturity of 3 months. Hence $T - t = 1/4$. We can use daily (i.e. $h = 1/365$) CMT rates for estimation.

Substituting (8.9) and (8.10) back into (8.12), we obtain:

$$\frac{-\ln P_{t+h,T+h}}{T-t} = \left(\frac{1 - e^{-\alpha h}}{T-t}\right)(\mu F_{T-t} + G_{T-t}) + e^{-\alpha h}\frac{-\ln P_{t,T}}{T-t} + \frac{F_{T-t}}{T-t}e_t \tag{8.11}$$

An OLS can be performed to arrive at the following estimates:

$$\begin{aligned}
&\text{Slope} = e^{-\alpha h} \\
&\text{Intercept} = \left(\frac{1-e^{-\alpha h}}{T-t}\right)(\mu F_{T-t} + G_{T-t}) \\
&\text{SSE} = \tfrac{F_{T-t}^2}{(T-t)^2}\tfrac{\sigma^2\left(1-e^{-2\alpha h}\right)}{2\alpha}
\end{aligned} \tag{8.12}$$

where h is data frequency and $T-t$ is the maturity of the bond. Note that (8.12) has three equations but has four unknowns. The market price of risk contained in function G cannot be identified. As a result, we must choose another bond series and run a simultaneous regression as follows:

$$\begin{cases}
\dfrac{-\ln P_{t+h,T+h}}{T-t} = \left(\dfrac{1-e^{-\alpha h}}{T-t}\right)(\mu F_{T-t} + G_{T-t}) + e^{-\alpha h}\dfrac{-\ln P_{t,T}}{T-t} + \dfrac{F_{T-t}}{T-t}e_t \\
\dfrac{-\ln P_{t+h,T'+h}}{T'-t} = \left(\dfrac{1-e^{-\alpha h}}{T'-t}\right)(\mu F_{T'-t} + G_{T'-t}) + e^{-\alpha h}\dfrac{-\ln P_{t,T'}}{T'-t} + \dfrac{F_{T'-t}}{T'-t}e_t
\end{cases} \tag{8.13}$$

where T' represents a different maturity. In the simultaneous regression the regressors in the two separate equations must be restricted to have the same coefficient. This allows us to solve for a unique set of α, μ, and λ. However, there are two SSEs from separate regressions and the choice of σ is arbitrary.

From this exercise, we know that it is impossible to estimate market price of risk without crossectional data. A univariate time series data cannot be used to estimate market price of risk.

Finally, via Taylors series expansion, we can derive the standard errors of the parameters:

$$\begin{aligned}
&\text{var}\left[e^{-\alpha h}\right] \approx \text{var}[1-\alpha h] = h^2\,\text{var}[\alpha] \\
&\text{var}\left[\left(\frac{1-e^{-\alpha h}}{T-t}\right)(\mu F_{T-t} + G_{T-t})\right] \\
&\qquad \approx \text{var}\left[\left(\frac{1-e^{-\alpha h}}{T-t}\right)\left[\mu F_{T-t} + \left(\mu - \tfrac{\sigma^2}{2\alpha^2}\right)(T-t-F_{T-t}) + \frac{\sigma^2 F_{T-t}^2}{4\alpha}\right]\right] \\
&\qquad = \left(\frac{1-e^{-\alpha h}}{T-t}\right)^2 \text{var}\left[\mu(T-t)\right] = \left(1-e^{-\alpha h}\right)^2 \text{var}[\mu]
\end{aligned} \tag{8.14}$$

8.3 The Cox-Ingersoll-Ross Model

The CIR model is a one-factor mean-reverting square-root model with the following short rate dynamics:

$$dr = \alpha(\mu - r)dt + \sigma\sqrt{r}dW \tag{8.15}$$

under the real measure and the parameters are defined previously. This is a complex distribution for r. The transition density for r is:

$$f(r_s|r_t) = ce^{-c(r_s+\xi)}\left(\frac{r_s}{\xi}\right)^{d/2} I_d(2c\sqrt{r_s\xi}) \tag{8.16}$$

where

$$\begin{aligned} c &= \frac{2\alpha}{\sigma^2\left(1 - e^{-\alpha(s-t)}\right)} \\ \xi &= r_t e^{-\alpha(s-t)} \\ d &= \frac{2\alpha\mu}{\sigma^2} - 1 \end{aligned}$$

which after change of variable $x = 2cr$:

$$f(x_s|x_t) = \frac{1}{2}e^{-1/2(x_s+\Lambda)}\left(\frac{x_s}{\Lambda}\right)^{d/2} I_d(\sqrt{\Lambda x_s}) \tag{8.17}$$

is a non-central chi-square distribution with $\Lambda = 2c\xi$ degrees of non-centrality and $\upsilon = 2d + 2 = \frac{4\alpha\mu}{\sigma^2}$ degrees of freedom. We know that the mean and variance of a non-central chi-squared distribution are, respectively, $\upsilon + \Lambda$ and $2(\upsilon + 2\Lambda)$, which lead to (see Chapter **??** for details):

$$\mathbb{E}[r_s|r_t] = \left(1 - e^{-\alpha(s-t)}\right)\mu + e^{-\alpha(s-t)}r_t \tag{8.18}$$

and

$$\mathbb{V}[r_s|r_t] = \frac{\sigma^2}{2\alpha}\left[\left(1 - e^{-\alpha(s-t)}\right)^2\mu + r_t\left(e^{-2\alpha(s-t)} - e^{-\alpha(s-t)}\right)\right] \tag{8.19}$$

It is clear that the mean of r under the CIR model is the same as the mean under the Vasicek model. However, unlike the Vasicek model, the variance of r is no longer independent of r. As a result, the instantaneous rate r under the CIR model is an AR(1) process but is hetroskedestic. A common econometric procedure to account for such a hetroskedesticity is a two-step regression as follows.

- Run OLS of (8.12):

$$
\begin{aligned}
&r_t = a + br_{t-h} + e_t \text{ where} \\
&a = \left(1 - e^{-\alpha(s-t)}\right)\mu \\
&b = e^{-\alpha(s-t)}
\end{aligned}
$$

- The variance of the error term must satisfy:
  $\mathbb{E}[e_t^2] = \frac{\sigma^2}{2\alpha}\left[\left(1 - e^{-\alpha(s-t)}\right)^2 \mu + r_t\left(e^{-2\alpha(s-t)} - e^{-\alpha(s-t)}\right)\right]$

- Run $e_t^2 = a' + b' r_{t-1} + u_t$ where

$$
\begin{aligned}
a' &= \frac{\sigma^2}{2\alpha}\left(1 - e^{-\alpha(s-t)}\right)^2 \mu \\
b' &= \frac{\sigma^2}{2\alpha}\left(e^{-2\alpha(s-t)} - e^{-\alpha(s-t)}\right)
\end{aligned}
$$

- Use the first OLS to solve for α and μ and the use the second OLS to solve for σ.

Since both a' and b' can be used to solve for σ, one can use a more reasonable estimate. Same as in the Vasicek model estimation, this regression is a convenience but it does not contain the market price of risk. To include in the market price of risk, we need to use the closed-form solution of the CIR model for the discount factor $P_{t,T}$. Interested readers can follow the steps as described in the previous sub-section and use two bonds to estimate all four parameters in the CIR model.

8.4 The Ho-Lee Model

Note that from the Ho-Lee model introduced in Chapter **??**, we define the zero-coupon bond price as $P(t, T, j)$ where t is current time, T is maturity time, and j represents the state of economy. The pricing formula in a binomial model is as follows:

$$\begin{cases} P(u,s,j) = \dfrac{P(t,s,j)}{P(t,u,j)} d(s-u) \\ P(u,s,j+1) = \dfrac{P(t,s,j)}{P(t,u,j)} u(s-u) \end{cases} \tag{8.20}$$

where $t < u < s$ and

$$u(\tau) = \frac{1}{p + (1-p)\delta^\tau}$$
$$d(\tau) = \frac{\delta^\tau}{p + (1-p)\delta^\tau}$$

The frequency of the data is $u-t$ and the maturity is τ. Note that $P(t,s,j)/P(t,u,j)$ is a forward price. If the yield curve is flat, then $P(t,s,j)/P(t,u,j)$ becomes $e^{-r(s-u)}$ under state j, which is $P(t, t+s-u, j)$. Note that this bond has the same time to maturity as $P(u,s,j)$ or $P(u,s,j+1)$. As a result, we can write (8.20) as:

$$\begin{cases} P(u, u+\tau, j) = P(t, t+\tau, j) d(\tau) \\ P(u, u+\tau, j+1) = P(t, t+\tau, j) u(\tau) \end{cases} \tag{8.21}$$

Hence, we can collect a series of τ-maturity bond prices (CMT series would be ideal). Write such bond series as $D_t(\tau)$. Then,

$$D_{t+1}(\tau) = \begin{cases} D_t(\tau) \dfrac{\delta^\tau}{p + (1-p)\delta^\tau} & \text{if rate rises} \\ D_t(\tau) \dfrac{1}{p + (1-p)\delta^\tau} & \text{if rate falls} \end{cases} \tag{8.22}$$

where $D_t(\tau)$ represents the zero-coupon bond price at time t for time to maturity τ in data series. We can choose a liquid Treasury series, e.g. 3-month T bills, for the estimation. Equation (8.22) can be written as:

$$\ln D_{t+1}(\tau) - \ln D_t(\tau) = -\ln[p + (1-p)\delta^\tau] + \tau \ln \delta \mathbb{I}_{t+1} + u_{t+1} \tag{8.23}$$

where the indicator function:

$$\mathbb{I}_t = \begin{cases} 1 & \text{if rate rises} \\ 0 & \text{if rate falls} \end{cases}$$

and $u_t \sim N(0,1)$. In a simple regression as follows:

$$y_t = a + bx_t + u_t \tag{8.24}$$

we can set $\delta = e^{b/\tau}$ and $p = \frac{e^{-a}-\delta^{\tau}}{1-\delta^{\tau}}$.

8.5 More sophisticated econometric methods

We can always use more sophisticated econometric methods such as Maximum Likelihood Estimation (MLE) or Generalized Method of Moments (GMM) to estimate parameters. These methods are more robust and accurate. However, this goes beyond the scope of this book. Interested readers are welcomed to study specialized books.

Chapter 9

Simulation

9.1 Introduction

So far we have been discussing how to calculate VaRs from historical data. These VaR numbers are historical and usually do not reflect what the future risk is. As a result, forecasting VaR is crucial in managing market risk. This is where we need simulations.

9.2 Random Numbers

9.2.1 Normal/log Normal

Simulating a normal distribution is the most basic. Although many software packages have such random number generator, it is really easy to do it by oneself.

Inverse normal probability function has an accurate closed form approximation formula and its code is enclosed at the end of this Chapter. In Excel, simply use `NORMSINV(RAND())` to obtain a normal random variable (as `RAND()` generates a uniform random variable).

Log normal random numbers are simply exponential transforms of normal random numbers.

9.2.2 Poisson

Assume the intensity parameter to be λ. Then the Poisson density function (i.e. probability of j events occurring) is:

$$\Pr(X = j) = \frac{(\lambda\Delta t)^j e^{-\lambda\Delta t}}{j!} \tag{9.1}$$

The most popular probability used in finance is $j = 0$ which represents no default. In that case, the probability is $e^{-\lambda\Delta t}$, which is known as the survival probability over the time period Δt.

There are two ways to simulate a Poisson process. One is to simulate directly. Note that a Poisson process is a collection of Bernoulli events. As a result, we can simulate a sequence of 0/1 events (0 for event not happening and 1 for event happening) over the dt period.

This simulation, however, is very time consuming. But it can take care of any flexible dynamics in the state variable. For example, if the intensity parameter is random, then it is only possible to do direct simulation.

An alternative simulation is to simulate exponential time. Note that the probability of no default is identical to the probability that default time τ occurs after the fixed specified time t.

$$\Pr(\tau > t) = e^{-\lambda t} \tag{9.2}$$

Hence, a uniform random number, say u, can be transformed into an exponential random number by $t = -\ln u/\lambda$.

9.2.3 Non-central Chi-square

Non-central Chi-square (nc-χ^2) cumulative distribution function is not closed form, hence an inversion function is not easy to calculate. As a result, simulating nc-χ^2 distribution is not easy. Here is how.

The non-central chi-square cumulative density function, CDF, is (e.g. see Johnson and Kotz):

$$F_{\mathrm{nc}\chi^2}(x) = \sum\nolimits_{j=0}^{\infty} \frac{e^{-\frac{1}{2}\lambda}(\frac{1}{2}\lambda)^j/j!}{2^{\frac{1}{2}\nu+j}\Gamma(\frac{1}{2}\nu + j)} \int_0^x z^{\frac{1}{2}\nu+j-1} e^{-\frac{1}{2}z} dz \tag{9.3}$$

where ν and λ are degrees of freedom and non centrality respectively.

Note that a chi-square density is:

$$f_{\chi^2}(z;k) = \frac{1}{2^{1/2k}\Gamma(1/2k)} z^{1/2k-1} e^{-1/2z} dz \tag{9.4}$$

where k is the degrees of freedom. Also note that the Poisson density is:

$$f_P(x;h) = \sum\nolimits_{j=0}^{\infty} \frac{e^{-h}h^x}{x!} \tag{9.5}$$

where h is the intensity parameter. Hence, we can write the CDF of the non-central, chi-square distribution as:

$$F_{\mathrm{nc}\chi^2}(x) = \sum\nolimits_{j=0}^{\infty} f_P(j; 1/2\lambda) \int_0^x f_{\chi^2}(z; \nu + 2j) dz \tag{9.6}$$

This implies that the CDF of a non-central, chi-square distribution is a CDF of Poisson weighted chi square distribution. Hence, to simulate the non-central, chi-square random variable, we follow the following steps:

- simulate Poisson random variable, x, with an intensity of $1/2\lambda$
- simulate chi-square random variable, y, with the degrees of freedom as $2x + 4\alpha\mu/\sigma^2$
- To simulate the short rate in the CIR model, $r(t + \Delta t)$, we first set $\lambda = 2cr(t)e^{-\alpha\Delta t}$ where $c = 2\alpha/\sigma^2 \left(1 - e^{-\alpha\Delta t}\right)$. Then we note that in the CIR model, the short term interest rate does not follow non-central, chi-square distribution but a scaled non-central, chi-square distribution. The scaled amount is $r(t + \Delta t) = y/(2c)$.

While the above algorithm always works, a simpler algorithm can be employed if the degrees of freedom is greater than 1. When $\nu > 1$, the following recursive result holds (see Johnson and Kotz):

$$Z_{nc\chi^2}(x;\nu,\lambda) = (Z_N + \sqrt{\lambda})^2 + Z_{\chi^2}(x;\nu - 1) \tag{9.7}$$

where $Z_{nc\chi^2}(x;\nu,\lambda$, Z_N, and $Z_{\chi^2}(x;\nu-1)$ represent the random variables from non-central, chi-square, normal, and chi square respectively. In this case, we can simulate the non-central, chi-square variable as follows:

- simulate standard normal

- simulate chi square with $\nu - 1$

- degrees of freedom

- add them together.

9.2.4 Multi-variate Gaussian

To simulate multiple *correlated* random variables, it is important to understand Cholesky decomposition. Cholesky decomposition makes it possible that we first simulate independent random variables and then correlate them. As a result, simulating many correlated random variables is extremely straightforward.

Given the correlation matrix, $\mathbf{R}$, we perform the Cholesky decomposition as follows:

$$\mathbf{C}\mathbf{C}' = \mathbf{R} \tag{9.8}$$

where $\mathbf{C}$ is the Cholesky matrix. In other words, the Cholesky matrix *decomposes* the correlation matrix. This is Cholesky decomposition. Let $\mathbf{X}$ (a vector) follow the diffusion:

$$d\mathbf{X} = \mathbf{M}dt + \Sigma \mathbf{C} dW \tag{9.9}$$

where $\mathbf{M}$ is the drift and Σ is a diagnal matrix of standard deviations. Then,

$$\begin{aligned} \Omega &= (d\mathbf{X})(d\mathbf{X}') \\ &= (\Sigma \mathbf{C} dW)(\Sigma \mathbf{C} dW)' \\ &= \Sigma \mathbf{C}\mathbf{C}'\Sigma' \\ &= \Sigma \mathbf{R} \Sigma' \end{aligned} \tag{9.10}$$

To simulate the joint Gaussian distribution, it is straightforward:

$$\begin{bmatrix} dX_1 \\ dX_2 \\ \vdots \\ dX_n \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} dt + \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{bmatrix}$$

$$\times \begin{bmatrix} x_{1,1} & \cdots & & & & x_{1,n} \\ 0 & x_{2,2} & & & & x_{2,n} \\ \vdots & & \ddots & & & \vdots \\ 0 & 0 & 0 & x_{n-2,n-2} & x_{n-2,n-1} & x_{n-2,n} \\ 0 & 0 & 0 & 0 & x_{n-1,n-1} & x_{n-1,n} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} dW_1 \\ dW_2 \\ \vdots \\ dW_{n-2} \\ dW_{n-1} \\ dW_n \end{bmatrix} \tag{9.11}$$

where

$$x_{j,n} = \rho_{j,n} \text{ for all } j \leqslant n \text{ (note that } \rho_{n,n} = 1)$$
$$x_{i,j} = \frac{\rho_{i,j} - \sum_{k=j+1}^{n} x_{i,k} x_{jk}}{x_{j,j}} \text{ for } i < j$$
$$x_{j,j} = \sqrt{1 - \sum_{k=1}^{n-j} x_{j,j+k}^2} \text{ for } j < n$$

Take a 4-variable example and expand it as:

$$\begin{bmatrix} dX_1 \\ dX_2 \\ dX_3 \\ dX_4 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{bmatrix} dt + \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \sigma_3 & \\ & & & \sigma_4 \end{bmatrix}$$

$$\times \begin{bmatrix} \sqrt{1 - x_{12}^2 - x_{13}^2 - x_{14}^2} & \frac{\rho_{12} - \rho_{14}\rho_{24} - x_{13}x_{23}}{\sqrt{1 - x_{23}^2 - x_{24}^2}} & \frac{\rho_{13} - \rho_{14}\rho_{34}}{\sqrt{1 - \rho_{34}^2}} & \rho_{14} \\ 0 & \sqrt{1 - x_{23}^2 - x_{24}^2} & \frac{\rho_{23} - \rho_{24}\rho_{34}}{\sqrt{1 - \rho_{34}^2}} & \rho_{24} \\ 0 & 0 & \sqrt{1 - \rho_{34}^2} & \rho_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} dW_1 \\ dW_2 \\ dW_3 \\ dW_4 \end{bmatrix} \tag{9.12}$$

The model may not be Gaussian. It could be a different statistical distribution.

9.3 Examples

9.3.1 Black-Scholes

The most straightforward example to practice Monte Carlo is the Black-Scholes model, which we shall demonstrate here. The Black-Scholes model employs the following log normal process for the stock price:

$$dS = rSdt + \sigma SdW \tag{9.13}$$

One way to simulate the stock price is to simulate the following stochastic Euler equation:

$$\Delta S = rS\Delta t + \sigma S\sqrt{\Delta t}\varepsilon \tag{9.14}$$

where $\varepsilon \sim N(0, 1)$. This equation can be written as:

$$S_{t+\Delta t} - S_t = rS_t\Delta t + \sigma S_t\sqrt{\Delta t}\varepsilon \tag{9.15}$$

or

$$S_{t+\Delta t} = S_t(1 + r\Delta t) + \sigma S_t\sqrt{\Delta t}\varepsilon \tag{9.16}$$

where Δt is the small time step (e.g. 1/252 for one daily simulations). Note that subscripts here are time indexes. After a number of steps as the option reaches maturity (e.g. 3 months, or 0.25 year), we can compute the option price as:

$$\max\{S_T - K, 0\}$$

for call and

$$\max\{K - S_T, 0\}$$

for put. The option value today is:

$$\begin{aligned} C_0 &= e^{-rT}\hat{\mathbb{E}}_t[\max\{S_T - K, 0\} \\ &= e^{-rT}\sum\nolimits_{j=1}^{N}\frac{1}{N}\max\{S_{T,j} - K, 0\} \end{aligned} \tag{9.17}$$

The following table is an Excel exercise of 8 weekly Monte Carlo simulation result. The basic information of this example is as follows:

Inputs	
initial stock price (S_0)	100
Δt (week)	1/52
risk-free rate (r)	0.04
volatility (σ)	0.3
strike (K)	98

The simulations of 1 path are done in Excel and shown below:

Outputs 1		
week	rand num	stock price
0		100.0000
1	-0.76519	96.8935
2	-0.70608	94.1218
3	1.41599	99.7388
4	-0.39640	98.1707
5	-0.42778	96.4991
6	-0.24764	95.5792
7	-0.58516	93.3259
8	-0.20290	92.6099

On this path, at the end of the 8th week, the stock price is \$92.61 which is out of the money for the call option. Simulate a different path as follows:

Outputs 2		
week	rand num	stock price
0		100.0000
1	1.59777	106.7241
2	0.24837	107.9089
3	1.45871	114.5405
4	1.47683	121.6660
5	-1.20669	115.6518
6	0.26589	117.0200
7	-0.44651	114.9363
8	0.61087	117.9457

and the stock price at the end of the 8th week is \$117.95 which is in the money of the call option, and the payoff is \$17.95. Simulating a number of paths, say 1000, we obtain 1000 values for the call option. Following (9.17), we take an average of the 1000 option values and then discount at the risk-free rate to arrive at the estimated call option premium.

Such simulations have two problems. First, there is a large number of normal random numbers to generate. Along each path, there need to be n steps and then there are N paths. Hence, there are a total of $n \times N$ random numbers to generate. This is very numerically costly.

An alternative, and more efficient, method is to solve for the stock value at the terminal time – S_T. For this we need the tool from Chapter 1: $\ln S_T = \ln S_t + \left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\sqrt{T-t}\varepsilon$. After obtaining $\ln S_T$, then taking an exponential we obtain the stock price at time T. Then proceed with the option payoff calculation as described and compute the option price. This approach only requires N simulations which is more efficient.

9.3.2 Two stocks

Simulating multiple stocks is very straightforward. We shall use a two-stock example to exemplify.

$$\begin{bmatrix} dS_1 \\ dS_2 \end{bmatrix} = \begin{bmatrix} rS_1 \\ rS_2 \end{bmatrix} dt + \begin{bmatrix} \sigma_1 S_1 & 0 \\ 0 & \sigma_2 S_2 \end{bmatrix} \begin{bmatrix} \sqrt{1-\rho^2} & \rho \\ 0 & 1 \end{bmatrix} \begin{bmatrix} dZ_1 \\ dZ_2 \end{bmatrix} \tag{9.18}$$

The simulation of the second stock follows equations (0-102) through (0-105) as follows:

$$S_{2,t+\Delta t} = S_{2,t}(1 + r\Delta t) + \sigma_2 S_{2,t}\sqrt{\Delta t}\varepsilon_2 \tag{9.19}$$

The first stock can be simulated via the same method, as follows:

$$S_{1,t+\Delta t} = S_{1,t}(1 + r\Delta t) + \sigma_1 S_{1,t}\sqrt{\Delta t}\left(\sqrt{1-\rho^2}\varepsilon_1 + \rho\varepsilon_2\right) \tag{9.20}$$

Use the following parameter values.

Parameter Values		
initial stock price S_0	100	100
volatility σ	0.3	0.3
week Δt	1/52	
risk-free rate r	0.04	
correlation ρ	-0.6	

We simulate 8 weeks of correlated stock prices as follows.

Outputs						
week	rand num	norm r.n.	stock price	rand num	norm r.n.	stock price
0			100			100
1	0.722067	0.588993	102.5273	0.698352	0.519665	100.3363
2	0.061049	-1.54603	96.01172	0.632405	0.33823	105.415
3	0.435153	-0.16327	95.43343	0.771024	0.742223	108.5298
4	0.357303	-0.36568	94.055	0.853696	1.052417	113.4053
5	0.082638	-1.38754	88.698	0.288043	-0.55911	115.3101
6	0.742822	0.652071	91.17241	0.800278	0.842615	116.7556
7	0.823997	0.930706	94.77272	0.470177	-0.07482	113.8423
8	0.157892	-1.00316	90.89039	0.688519	0.491657	118.6433

9.3.3 Stock with random interest rates

We continue to use the Euler approximation for the stock (Black-Scholes):

$$\begin{aligned} dS &= rSdt + \sigma_S S dW_S \\ S_{t+\Delta t} &= S_t + rS_t\Delta t + \sigma S_t\sqrt{\Delta t}\varepsilon_{S,t+\Delta t} \\ &= S_t\left(1 + r\Delta t + \sigma\sqrt{\Delta t}\varepsilon_{S,t+\Delta t}\right) \end{aligned} \tag{9.21}$$

where $\Delta t = \frac{1}{252} = 0.003968254$.

But now the interest rates are random and we assume to follow the Vasicek model:

$$\begin{aligned} dr &= \alpha(\mu - r)dt + \sigma_r dW_r \\ r_{t+\Delta t} &= r_t + \alpha\mu\Delta t - \alpha r_t\Delta t + \sigma_r\sqrt{\Delta t}\varepsilon_{r,t+\Delta t} \\ &= r_t(1 - \alpha\Delta t) + \alpha\mu\Delta t + \sigma_r\sqrt{\Delta t}\varepsilon_{r,t+\Delta t} \end{aligned} \tag{9.22}$$

where the two random processes are correlated with the correlation coefficient ρ : $dW_S dW_r = \rho dt$.

Let (known as the Cholesky decomposition)

$$\begin{bmatrix} dW_S \\ dW_r \end{bmatrix} = \begin{bmatrix} \sqrt{1-\rho^2} & \rho \\ 0 & 1 \end{bmatrix} \begin{bmatrix} dZ_S \\ dZ_r \end{bmatrix} \tag{9.23}$$

Then $dZ_S dZ_r = 0$. Then,

$$
\begin{aligned}
\begin{bmatrix} dS \\ dr \end{bmatrix} &= \begin{bmatrix} rS \\ \alpha(\mu - r) \end{bmatrix} dt + \begin{bmatrix} \sigma_S & 0 \\ 0 & \sigma_r \end{bmatrix} \begin{bmatrix} dW_S \\ dW_r \end{bmatrix} \\
&= \begin{bmatrix} rS \\ \alpha(\mu - r) \end{bmatrix} dt + \begin{bmatrix} \sigma_S S & 0 \\ 0 & \sigma_r \end{bmatrix} \begin{bmatrix} \sqrt{1-\rho^2} & \rho \\ 0 & 1 \end{bmatrix} \begin{bmatrix} dZ_S \\ dZ_r \end{bmatrix}
\end{aligned} \tag{9.24}
$$

Hence,

$$
\begin{aligned}
S_{t+\Delta t} &= S_t + rS_t\Delta t + \sigma_S S_t \sqrt{\Delta t}\varepsilon_{S,t+\Delta t} \\
&= S_t\left(1 + r\Delta t + \sqrt{\Delta t}\sigma_S\left(\sqrt{1-\rho^2}\varepsilon_{S,t+\Delta t} + \rho\varepsilon_{r,t+\Delta t}\right)\right)
\end{aligned} \tag{9.25}
$$

and

$$
r_{t+\Delta t} = r_t(1 - \alpha\Delta t) + \alpha\mu\Delta t + \sigma_r\sqrt{\Delta t}\varepsilon_{r,t+\Delta t} \tag{9.26}
$$

Use the following parameter values.

Parameter Values		
initial value (S_0/r_0)	100	0.02
volatility (σ_S/σ_r)	0.3	0.1
week Δt	1/52	
correlation ρ	-0.6	

We simulate 8 weeks of correlated stock price and interest rates as follows.

Outputs						
week	rand num	norm r.n.	int. rate	rand num	norm r.n.	stock price
0			0.02			100
1	0.26516	-0.62751	0.01168	0.63801	0.35315	99.98257
2	0.35716	-0.36607	0.00715	0.14230	-1.07005	98.30791
3	0.85302	1.04946	0.02234	0.73532	0.62898	101.11923
4	0.68315	0.47651	0.02928	0.76893	0.73533	102.79198
5	0.38716	-0.28673	0.02551	0.88485	1.19960	107.22825
6	0.50613	0.01535	0.02600	0.34620	-0.39560	108.93629
7	0.87028	1.12771	0.04191	0.25562	-0.65691	100.22972
8	0.22975	-0.73966	0.03162	0.63994	0.35831	99.56878

9.4 Homework

1. VaR of one stock: Choose a stock and do the same analysis.

2. VaR of one stock and one option: Do the same analysis with the company of your choice.

9.5 Appendix

Theoretically, the var-cov matrix for the factors should be a diagonal matrix as factors are orthogonal. However, in reality, we do not obtain orthogonal factors from data using PCA. As a result, the following matrix is observed:

$$
\Omega_F = \begin{bmatrix}
\boxed{\begin{matrix} \sigma_{11} & \cdots & \sigma_{1n_1} \\ & \ddots & \\ & & \sigma_{n_1 n_1} \end{matrix}} & 0 & 0 & 0 \\
0 & \boxed{\begin{matrix} \sigma_{n_1+1,n_1+1} & \cdots & \sigma_{n_1+1,n_1+n_2} \\ & \ddots & \\ & & \sigma_{n_1+n_2,n_1+n_2} \end{matrix}} & 0 & 0 \\
0 & & \ddots & 0 \\
0 & 0 & 0 & \boxed{\begin{matrix} \sigma_{n_1+\cdots+n_{n-1},1} & \cdots & \sigma_{n_1+\cdots+n_{n-1},n_n} \\ & \ddots & \\ & & \sigma_{n_1+\cdots+n_n,n_1+\cdots+n_n} \end{matrix}}
\end{bmatrix} \tag{9.27}
$$

where $n_1 + \cdots + n_n = K$ and Ω_F is a $K \times K$ matrix. The above matrix implies that inter-industry correlations are assumed 0. Note that the standard deviations here are for the changes in price and not for return.

Chapter 10

Volatility and Extreme Value Theorem

10.1 Introduction

As we have learned so far, volatility is the key to risk management. The entire market risk (i.e. VaR) is based upon volatility. An inaccurate estimate of the volatility, as a result, can lead to very wrong risk management decisions. Hence, estimating the right levels of volatility is a very important job. In our analyses before, mostly the volatility used is based upon historical estimates, which are not good reflections of future levels of volatility. Consequently, volatility forecast is a key and important task.

The easiest way to obtain an estimate a future volatility is the implied volatility from the options markets. As option prices are bets on future payoffs, their implied volatility estimates reflect what investors' beliefs of future volatility. Unfortunately, we only have a very limited number of options contracts traded in the market place (limited tenors and most tenors are very short term, less than a year). Therefore, methods of forecasting volatility are adopted in risk management.

Another important risk management concept is to recognize the fact that rare events are not rare. In other words, large losses are more than likely than what people think. To over come this problem, probability distributions with fat tails are proposed to replace the popular normal distribution. t distribution, Cauchy distribution, and so on are being proposed. Or alternatively, we can adopt the extreme value theory (EVT) to "fatten" the tails of a normal distribution.

10.2 Implied Volatility and Risk-Neutral Density

The easiest and best forecast of future volatility is the implied volatility from option and/or futures prices.[1]

With enough options and futures contracts we not only can estimate volatility of any future date, we can also imply out the entire distribution of the underlying asset. Options and futures can provide a "volatility surface" which can then be used to imply out the risk-neutral distribution of the underlying asset.

Take the Black-Scholes model as an example. The call option formula is:

$$C = SN(d_1) - e^{-r(T-t)}KN(d_2)$$

where

$$d_1 = \frac{\ln S - \ln K + (r + {}^1\!/_2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$
$$d_2 = d_1 - \sigma\sqrt{T-t}$$

Taking partial derivative of the call price with respect to the strike price (i.e. $\partial C/\partial K$), we arrive at:

$$N(d_2)$$

which is the probability of in the money, or $\Pr[S > K]$ for a given maturity. Taking another partial derivative with respect to the strike, we arrive at the probability density function of the underlying asset:

$$p.d.f. = \frac{\partial N(d_2)}{\partial K}$$

In the Black-Scholes case, this is just a normal density. Yet this method can be used with real options and the resulting $p.d.f.$ is not necessarily normal. To summarize, the risk-neutral density of the stock can be computed as:

$$p.d.f = \frac{\partial^2 C}{\partial K^2}$$

[1]Note that forwards and swaps are model-free, they do not embed volatility information (or any other parameter). Futures and options are model dependent and hence can be used to estimate the volatility which is one of the model parameters.

Sometimes we do not have enough option contracts to derive a granular enough density function and industry often uses smoothing techniques to generate "fake" option prices. The smoothing techniques vary from bank to bank.

10.3 GARCH

GARCH, or Generalized Auto Regressive and Conditional Heteroskedesticity, is a methodology to model time series of volatility. A GARCH(p, q) model specifies the evolution of the volatility as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{10.1}$$

The estimation of GARCH is left to another chapter.

10.4 EWMA

EWMA (exponentially weight moving average) methodology for calculating the volatility estimates.

$$\sigma_{j,t}^2 = (1-\lambda) r_{j,t-1}^2 + \lambda \sigma_{j,t-1}^2 \tag{10.2}$$

where λ is the decay factor (which is usually set between 0.99 and 0.97) and j represents a tenor of choice (e.g. 3-month). EWMA is a special case of a GARCH(1,1) model as follows:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \beta_1 \varepsilon_{t-1}^2 \tag{10.3}$$

where $\alpha_0 = 0$, and $\alpha_1 = 1 - \beta_1$.

When the β's are large and the α's small, the volatility evolves more quickly as more weight is placed on the most recent change in the underlying random variable, and less on the previous volatility levels. On the other hand, when the βs are small and the αs large, the volatility evolves more slowing resulting in a low level of volatility-of-volatility.

RiskMetrics™ estimated [2] the decay (λ) parameter for various regions and

[2] Technical Document. Document date unclear. Morgan Guaranty Trust Company Risk Management Advisory Jacques Longerstaey (1-212) 648-4936 riskmetrics@jpmorgan.com Reuters Ltd International Marketing Martin Spencer (44-171) 542-3260 martin.spencer@reuters.com

instruments as follows (their page 99, Table 5.8: Optimal decay factors based on volatility forecasts):

Decay Parameter Estimates					
Country	Fore Ex	5y Swap	10y Zero	Eq Index	1y MM
Austria	0.945				
Australia	0.980	0.955	0.975	0.975	0.970
Belgium	0.945	0.935	0.935	0.965	0.850
Canada	0.960	0.965	0.960		0.990
Switzerland	0.955	0.835		0.970	0.980
Germany	0.955	0.940	0.960	0.980	0.970
Denmark	0.950	0.905	0.920	0.985	0.850
Spain	0.920	0.925	0.935	0.980	0.945
France	0.955	0.945	0.945	0.985	
Finland	0.995				0.960
Great Britain	0.960	0.950	0.960	0.975	0.990
Hong Kong	0.980				
Ireland	0.990		0.925		
Italy	0.940	0.960	0.935	0.970	0.990
Japan	0.965	0.965	0.950	0.955	0.985
Netherlands	0.960	0.945	0.950	0.975	0.970
Norway	0.975				
New Zealand	0.975	0.980			
Portugal	0.940				0.895
Sweden	0.985		0.980		0.885
Singapore	0.950	0.935			
United States		0.970	0.980	0.980	0.965
ECU		0.950			

To estimate the parameter, one must use GARCH(1,1) with restrictions. We leave all the econometric discussions to another chapter.

10.5 Extreme Value Theory (EVT)

The material in this section is taken from the master thesis of Krenar Avdulaj (2010) at the Charles University in Prague (p.15-19).

The Extreme Value Theory (EVT) is designed to model very large tails. This is know as black swans or rare events. The basic black swan or rare event story is that black swans or rare events are really rare. In other words, tail probabilities (probabilities for very large losses) are not as small as people think.

There are fatter tail distributions, such as t or Cauchy. Yet, a popular method is to use EVT to model the tail probabilities. EVT is based upon the generalized Pareto distribution (cumulative density function) as given below:

$$G_{\xi,\beta,u}(x) = \begin{cases} 1 - \left(1 + \frac{\xi}{\beta}(x-u)\right)^{-1/\xi} & \xi \neq 0 \\ 1 - \exp\left\{-\frac{x-u}{\beta}\right\} & \xi = 0 \end{cases} \tag{10.4}$$

where u can be regarded as a high threshold, $\beta > 0, x \geq 0$ and $x \leq -\beta/\xi$. Furthermore, x can be specified as:

$$\begin{array}{ll} x - u \geq 0 & \xi \geq 0 \\ 0 \leq x - u \leq -\beta/\xi & \xi < 0 \end{array}$$

The conditional distribution of excess losses over a high threshold u is defined to be:

$$F_u(y) = \Pr[X \leqslant y + u | X > u] \tag{10.5}$$

for $0 \leq y \leq v - u$ and v is the end point of the distribution which could be ∞ in many cases. Then this conditional distribution of excess losses can be written in terms of underlying distribution function F as:

$$F_u(y) = \frac{F(y+u) - F(u)}{1 - F(u)} \tag{10.6}$$

From the above expression, it seems that depending on the function shape of $F(u)$, the excess losses distribution can have infinite ways of asymptotic behaviors. But according to a key result in EVT, we have the following Theorem. For a large class of underlying distributions we can find a function $F(u)$ such that (D. V. Gnedenko 1943):

$$\lim_{u \to v} \sup_{0 \leqslant y \leqslant v-u} |F_u(y) - G_{\xi,\beta,u}(x)| = 0 \tag{10.7}$$

where generalized Pareto distribution $G_{\xi,\beta,u}(x)$ is a two parameter distribution function as described in equation (10.4).

10.5.1 Estimating Tails of Distributions

Using equations (10.4), (10.6), and (10.7), we can rewrite distribution function as:

$$F(x) = (1 - F(u))G_{\xi,\beta,u}(x - u) + F(u) \tag{10.8}$$

To estimate the EVT tail probability $G_{\xi,\beta(u)}(x - u)$, we first need to have a proxy for $F(u)$. For this purpose we take the obvious empirical estimator $N - N_u/N$ where N is the total number of simulation points and N_u is the number of points located in the tail.[3]

Hence, we get the tail probability as (adjusted for the EVT):

$$1 - F(x) = \frac{N_u}{N}\left(1 + \hat{\xi}\frac{x - u}{\hat{\beta}}\right)^{-1/\hat{\xi}} \tag{10.9}$$

Note that when $\xi = 0$ (this is when the distribution is degenerated to Gaussian), the tail probability is precisely the sample statistic, which is N_u/N. If $\xi \neq 0$, then the tail probability is higher than N_u/N, basically adding an additional quantity to the tail.

A remark on the left tail estimation

The estimation of tail probability equation apparently gives the right tail. What we need is the left tail estimation. To do that, we simply flip the tail from left to right. Note that losses are all negative, and hence by flipping them we satisfy $x > 0$. Consequently, equation (10.9) applies.

10.5.2 Estimating VaR

For a given probability $\alpha > F(x)$ the VaR estimate is calculated by inverting the tail estimation formula to get

$$V\hat{a}R = u + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{N}{N_u}(1 - \alpha)\right)^{-\hat{\xi}} - 1\right) \tag{10.10}$$

Note that $u = F^{-1}(\alpha)$, i.e. inverse function of F. The rationale behind the VaR equation is quite intuitive. If we adopt normal distribution for our parametric

[3]This is because, of course, the tail probability has a sampling statistic of N_u/N.

VaR then the tail probability $F(u)$ will be too small and u will be too small. As a result, due to the reality of fat tails, we must add value to u. The amount to add is the second term in equation (10.10). For this term to be positive, given both ξ and β are positive, it must be that:

$$\left(\frac{N}{N_u}(1-\alpha)\right)^{-\hat{\xi}} > 1$$

which means that:

$$\frac{N}{N_u} < \frac{1}{1-\alpha} \qquad \text{or} \qquad \alpha < 1 - \frac{N_u}{N}$$

This means that the tail probability of a real distribution must be greater (fatter) than the tail probability of a normal distribution. Otherwise, the tails are THINNER rather than fatter. Now it is clear that the necessary condition for a fatter tail using EVT is that the sampling of a real distribution must be fatter than normal. Then the generalized Pareto distribution helps use its parameters to properly estimate the accurate VaR.[4]

10.5.3 Estimating ES

Expected shortfall is related to VaR by

$$ES_\alpha = VaR_\alpha + E[X - VaR_\alpha | X > VaR_\alpha]$$

It can verified that

$$F_{VaR_\alpha}(y) = G_{\xi,\beta+\xi(VaR_\alpha-u)}(y)$$

By noting the mean of the above distribution is

$$(\beta + \xi(VaR_\alpha - u))/(1-\xi)$$

[4]For example, supposed there there are a total of 100 observations and α is 95%. Then a fat tail suggest that N_u (the tail) should be more than 5 observations.

Chapter 11

Model Risk

11.1 Introduction

Model risk has become one of the major risk management areas post-Lehman crisis. This is because too many positions (especially illiquid ones) have been marked to models that were either manipulated or incorrect. One obvious example is the lack of (il)liquidity discounts.

According to Wikipedia, sources of model risk are (https://en.wikipedia.org/wiki/Model-risk):

- Sources of model risk
- Uncertainty on volatility
- Time inconsistency
- Correlation uncertainty
- Complexity
- Illiquidity and model risk

famous cases in model risk are:

- Natwest – Interest rate options and swaptionsincorrect model specification.
- Bank of Tokyo-Mitsubishi – Interest rate options and swaptions.
- LTCM – lack of stress testing – Crouhy, Galai, and Mark (and over-estimating market appetites – When Genius Failed by Roger Lowenstein)

- Barclays de Zoete Wedd (BZW) – Mispriced currency options.
- National Australia Bank $3 Billion AUD loss on Homeside interest rate model.
- 2007-2012 global financial crisis – Over-reliance on David X. Li's Gaussian copula model misprices the risk of collateralized debt obligations.

quantitative approaches to model risk

- Model averaging vs worst-case approach
- Quantifying model risk exposure
- Position limits and valuation reserves

mitigating model risk

- Theoretical basis
- Implementation
- Testing

examples of model risk mitigation

- Parsimony
- Model risk premium

http://www.bostonfed.org/economic/neer/neer1997/neer697b.pdf

In March of 1997, NatWest Markets, an investment banking subsidiary of National Westminster Bank, announced a loss of £90 million due to mispriced sterling interest rate options. Shortly thereafter, BZW, an investment banking subsidiary of Barclays, sustained a £15 million loss on mispriced currency options and Bank of Tokyo-Mitsubishi announced a loss of $83 million. In April of 1997, Deutsche Morgan Grenfell lost an undisclosed amount. Model errors have been blamed for all these losses.

Part III

Credit Risk

Chapter 12

Introduction

12.1 Sources of Credit Risk

Simply speaking, credit risk is default risk. However, credit risk can be transferred in different ways. That is, there are different transactable instruments that can transfer different parts of credit risk. Typically, there are three types of credit risk:

- bankruptcy
- migration
- spread

Bankruptcy is the ultimate default. The firm no longer operates and liquidates its assets to pay off its debts. Debt holders lose some of their investment value (known as loss given default or severity) and stock holders lose all of their investment value. Credit risk of this sort needs to evaluate the likelihood of bankruptcy (known as probability of default, or PD) and loss given default (LGD).

Usually bankruptcies do not happen suddenly. Companies lose money and ultimately lead to bankruptcy. Rating agencies monitor such transition and provide their assessment of the fundamental financial healthiness of the firm via credit ratings. Hence, rating migrations (especially from high to low) indicate how poorly the firm is operating. As a result, when a firm is downgraded by rating agencies, its bond and stock prices are lowered. Empirical evidence has shown that rating changes have significant impacts on stock and bond prices.

Last but not least is the spread credit risk. Ratings reflect firms' financial health. Yet, rating changes do not occur often. Furthermore, ratings are discrete measures. Stocks and bonds are traded continuously in the marketplace. As a result, we can transfer credit risk much more quickly by trading spreads.

In this chapter, we only discuss the basic math of credit risk modeling. That is, we only evaluate the bankruptcy credit risk. There are models for just migration and spread risks but are not covered here. Readers should be aware that many migration and spread models are ad-hoc and do not relate back to default. Readers should keep in mind that all three types of credit risk result from default because without default risk, there are no migration or spread changes.

12.2 Types of Credit

The credit risk is usually modeled by the Jarrow-Turnbull model published in 1995 by the Journal of Finance.

12.2.1 Corporates

Corporations borrow to finance their investment projects. There is a wide variety of forms of how corporate borrowing ranging from short term borrowings like lines of credit, commercial papers (90 days and 180 days typically), and bank term loans (which themselves take various forms), to medium term corporate notes and bonds, to long term corporate bonds (some can be as long as 100 years!)

There is a wide variety of forms within corporate bonds fixed rate bonds, floating rate bonds (floaters), bonds with sinking funds, bonds with amortizing principals, convertible bonds, callable and puttable bonds, etc. Corporate bonds also vary in terms of covenants, collaterals, and seniorities.

Rating agencies rate corporate bonds by their default likelihoods and recovery values once defaults happen. Hence, ratings provide investors a rough idea of how risky corporate bonds are in a general way. In other words, ratings summarize all the information with a single letter to help investors understand the credit risk of corporate bonds. While ratings are very helpful, due to their simplicity, they are often criticized to be inaccurate and behind market timing.

Despite many rating agencies that provide different rating systems, in general we have 9 rating groups:

- AAA
- AA

- A
- BBB
- BB
- B
- CCC
- CC
- C
- D

where within each rating there could be multiple, usually three, sub-groups (called notches).

Due to different business characteristics (business risk), we often classify companies into industry sectors. The highest level of classification is to divide companies into two groups: financial and industrial. Financial companies have high leverage ratios due the nature of the business, hence cannot be compared with other industrial firms. Industrial companies are further divided into many groups. Different service companies classify industrial companies differently. For example, Compustat, the largest financial data source, classifies the companies as follows:

- Division 0: Agriculture, Forestry, And Fishing
- Division 1: Mining, and Construction
- Division 2,3: Manufacturing
- Division 4: Transportation, Communications, Electric, Gas, And Sanitary Services
- Division 5: Wholesale Trade, Retail Trade
- Division 6: Finance, Insurance, and Real Estate
- Division 7,8: Services
- Division 9: Public Administration

Combining every industry sector and every credit rating, we are able assign each and every firm into a "cohort". For example, 9 credit ratings (from AAA to C) and 9 industry sectors result in 81 cohorts. Within each cohort, companies should be rather homogenous since they belong the same industry and rating. As a result, we can compute 81 cohort yield curves.

12.2.2 Sovereigns ($ denominated)

There are two major types of sovereign bonds traded in the United States (and denominated in dollars):

- Yankee bonds
- Brady bonds

[www.investopedia.com]A Yankee bond is a bond denominated in U.S. dollars and is publicly issued in the United States by foreign banks and corporations. According to the Securities Act of 1933, these bonds must first be registered with the Securities and Exchange Commission (SEC) before they can be sold. Yankee bonds are often issued in tranches and each offering can be as large as $1 billion.

Due to the high level of stringent regulations and standards that must be adhered to, it may take up to 14 weeks (or 3.5 months) for a Yankee bond to be offered to the public. Part of the process involves having debt-rating agencies evaluate the creditworthiness of the Yankee bond's underlying issuer.

Foreign issuers tend to prefer issuing Yankee bonds during times when the U.S. interest rates are low, because this enables the foreign issuer to pay out less money in interest payments.

[wiki]Brady bonds are dollar-denominated bonds, issued mostly by Latin American countries in the 1980s, named after U.S. Treasury Secretary Nicholas Brady.

Brady bonds were created in March 1989 in order to convert bonds issued by mostly Latin American countries into a variety or "menu" of new bonds after many of those countries defaulted on their debt in the 1980's. At that time, the market for sovereign debt was small and illiquid, and the standardization of emerging-market debt facilitated risk-spreading and trading. In exchange for commercial bank loans, the countries issued new bonds for the principal and, in some cases, unpaid interest. Because they were tradable and came with some guarantees, in some cases they were more valuable to the creditors than the original bonds.

The key innovation behind the introduction of Brady Bonds was to allow the commercial banks to exchange their claims on developing countries into tradable instruments, allowing them to get the debt off their balance sheets. This reduced the concentration risk to these banks.

The plan included two rounds. In the first round, creditors bargained with debtors over the terms of these new claims. Loosely interpreted, the options contained different mixes of "exit" and "new money" options. The exit options were designed for creditors who wanted to reduce their exposure to a debtor country.

These options allowed creditors to reduce their exposure to debtor nations, albeit at a discount. The new money options reflected the belief that those creditors who chose not to exit would experience a capital gain from the transaction, since the nominal outstanding debt obligation of the debtor would be reduced, and with it the probability of future default. These options allowed creditors to retain their exposure, but required additional credit extension designed to "tax" the expected capital gains. The principal of many instruments was collateralized, as were "rolling interest guarantees", which guaranteed payment for fixed short periods. The first round negotiations thus involved the determination of the effective magnitude of discount on the exit options together with the amount of new lending called for under the new money options.

In the second round, creditors converted their existing claims into their choice among the "menu" of options agreed upon in the first round. The penalties for creditors failing to comply with the terms of the deal were never made explicit. Nevertheless, compliance was not an important problem under the Brady Plan. Banks wishing to cease their foreign lending activities tended to choose the exit option under the auspices of the deal.

By offering a "menu" of options, the Brady Plan permitted credit restructurings to be tailored to the heterogeneous preferences of creditors. The terms achieved under these deals indicate that debtors used the menu approach to reduce the cost of debt reduction. Furthermore, it reduced the holdout problem where certain shareholders have an incentive to not participate in the restructuring in hopes of getting a better deal.

The principal amount is usually but not always collateralized by specially issued U.S. Treasury 30-year zero-coupon bonds purchased by the debtor country using a combination of International Monetary Fund, World Bank, and the country's own foreign currency reserves. Interest payments on Brady bonds, in some cases, are guaranteed by securities of at least double-A-rated credit quality held with the Federal Reserve Bank of New York.

Countries that participated in the initial round of Brady bond issuance were Argentina, Brazil, Bulgaria, Costa Rica, Dominican Republic, Ecuador, Mexico, Morocco, Nigeria, Philippines, Poland, Uruguay.

12.2.3 Munis

¡wiki¿In the United States, a municipal bond (or muni) is a bond issued by a state, city or other local government, or their agencies. Potential issuers of municipal bonds include cities, counties, redevelopment agencies, school districts, publicly owned airports and seaports, and any other governmental entity (or group of governments) below the state level. Municipal bonds may be general obligations of the issuer or secured by specified revenues. Interest income received by holders of municipal bonds is often exempt from the federal income tax and from the income tax of the state in which they are issued, although municipal bonds issued for certain purposes may not be tax exempt.

Muni bonds are as risky as corporate bonds. Rating agencies rate muni bonds as they rate corporate bonds.

12.2.4 Commercial Mortgages

Commercial mortgages have very low LTVs (loan-to-value ratios) and hence suffer from high default risk. Examples of commercial mortgages are:

- shopping centers
- casinos
- hotels
- rental apartments
- etc.

12.2.5 Retail Credit

The above credit risks are roughly categorized as corporate credits. That is, the credit risk of an entity. In addition, there are also credit risks from individuals. Individuals borrow money just like corporations do and they can default on their loans as well. This is known as retail credit risk. Due to a number of various reasons, we must model retail credit risk different from corporate credit risk. There are a number retail loans that are transacted in the secondary market. We have the following types of retail credit:

- Credit card
- Auto
- Student
- Home equity

Mortgages The most important one of all is residential home mortgages. Home owners borrow from banks to purchase their homes and use their homes as collaterals. These loans are called mortgages. Many mortgage banks lend home owners with the deposits they receive from their depositors. Many others sell their mortgages to the secondary market as mortgage-backed securities.

Regardless if a mortgage is lent directly by deposits or selling to the secondary market, the interest rate charged (called mortgage rate) on the borrower is a function of the borrowers credit history, which is categorized as follows:

- prime
- Alt-A
- sub-prime

These mortgage rates reflect the credit quality of the borrowers and hence vary widely. Prime borrowers are the safest. They must meet many strict criteria such as low LTV (loan to value ratio, typically less than 80%) and low PI (payment to income ratio, usually less than 1/3). Alt-A borrowers are less safe. They may not be high credit-risky but are classified as such due to lack of long credit history or lack of documentation. Sub-prime borrowers are regarded as unsafe or high risk, but they do not necessarily have bad credit history. Some may not have steady job or regular income.

Credit cards Credit card loans are the money owed by card owners by not paying the full amount each month. It is notoriously well known that credit card interest rates are extremely high (like 18% on a per annum basis). Credit card loans, along with other retail lendings are packaged in “asset backed securities” and transacted in the secondary market. A credit card loan can be short or long depending on the borrower’s consumption and financing habit. It can range from a few days to several months.

Auto loans Loans borrowed to purchase automobiles are packaged in asset backed securities as well. However, unlike credit card loans, these auto loans are collateralized (by the vehicles). As a result, they are much less risky than credit card loans. In fact, the asset backed securities backed by auto loans are consider extremely safe and they often receive AAA ratings.

Student loans Student loans are government subsidized loans which are in many cases guaranteed by the government. The student loans that are securitized in the secondary market are performed by SLM Corporation (commonly known as Sallie Mae; originally the Student Loan Marketing Association).

12.3 Economic Default and Liquidity Default

Economic defaults are those follow the structural models (e.g. Merton) where firms lose money over time and deplete the equity (so called "bleed to death"). Liquidity defaults are those where firms are profitable on the accrued basis but are short of cash to pay the immediate expenses (such as a large sum of debts due).[1] This is a sudden death situation which is close to the reduced-form models. We shall see more details in Chapter 17.

12.4 Liquidation Process

A liquidation process is a process of selling a firm's assets under bankruptcy. Lawyers and accountants are involved in determining which debt holder should get what. This is a complex process and will take many years to finish in that all debt contracts must be reviewed including all covenants and clauses of the debt contracts. Selling assets can also itself be a lengthy process and many interested parties (usually competitors) will pick and choose and negotiate prices.

12.5 Recovery (or Severity) and Loss Given Default (LGD)

This is one of the two major building blocks of credit risk. The amount of recovery (or oppositely loss given default) severely and directly impact the credit risk. In an extreme case where recovery is 100%, the credit risk is 0 no matter how high the probability of default (another building block) is. Recovery or LGD is also a major player in portfolio risk analysis (i.e. portfolio loss function which is the foundation of all credit risk metrics.)

12.6 Probability of Default (PD)

This is another building block of credit risk. In many situations PD is more important than LGD, especially when curve construction (because a preferable credit curve is 0 recovery). There are two major approaches in constructing the credit

[1] This happened for Excite@Home where a convertible bond issue was due. The firm was profitable but short of cash. See the Appendix of Chapter 2 for details.

curve – reduced-form approach and structural approach, to be discussed in details in Chapters 13 and 14.

Chapter 13

Reduced-Form Models

13.1 Introduction

There are two general approaches in modeling default – the structural approach and the reduced-from approach. The reduced-form models for default, like any other reduced-form models, take market information as given. Analogously, the structural models, like many others, are built on economic fundamentals.

While details vary, the basic principle of the reduced-form models is that defaults occur according to a Poisson process. In other words, a default event is a Poisson jump event. The representative reduced-form models for default are the Jarrow-Turnbull and Duffie-Singleton models. When a Poisson jump event happens, a firm is in default. Once a firm is in default, it is assumed that it will not become live again. As a result, a default here represent complete bankruptcy. Assets of the company must be liquidated. The usual notion of default such as Chapter 11 (bankruptcy protection) is not a default event by these models.

13.2 Survival Probability

We compute survival probabilities when we model default. The survival probability between now and some future time T for defaults is:

$$Q(t,T) = e^{-\lambda(T-t)} \tag{13.1}$$

where λ is the "intensity" parameter of the Poisson process. This intensity parameter intuitively represents the likelihood of default. When the recovery is 0, then this value is almost identical (exactly identical in continuous time) to the "forward" probability of default.

As we can see, there is extreme similarity between the result of survival probability and risk-free discount, which is $P(t,T) = e^{-r(T-t)}$. In fact, we shall show that they can be combined linearly if the recovery of a risky bond is 0. Due to this similarity, we shall proceed without proof (which is difficult in many situations) to borrow what we have known for the risk-free rate and use it for the intensity. In the insurance literature, the intensity parameter is called the hazard rate.

If the intensity parameter (hazard rate) is non-constant, then we can express the survival probability as:

$$Q(t,T) = \exp\left(-\int_t^T \lambda_u du\right) \tag{13.2}$$

Furthermore, if the hazard rate is random, then we simply compute the risk neutral expectation:

$$Q(t,T) = \hat{\mathbb{E}}_t\left[\exp\left(-\int_t^T \lambda_u du\right)\right] \tag{13.3}$$

Finally, if the interest rate is random, then we must use the forward measure:

$$Q(t,T) = \tilde{\mathbb{E}}_t^{(T)}\left[\exp\left(-\int_t^T \lambda_u du\right)\right] \tag{13.4}$$

Equation 13.2 represent the survival probability till a certain future time (T). As a result, we can have a whole "curve" of survival probabilities, known as the survival probability curve.

Taking the derivative with respect to an arbitrary future time T of equation 13.2, we get:

$$\lambda_T = -\frac{d\ln Q(t,T)}{dT} \tag{13.5}$$

Again, we remind the readers of the similarity between default probabilities and forward rates in the risk-free world.[1]

[1]The risk-free forward rate is $f_{t,T} = -d\ln P(t,T)/dT$.

13.3 Zero Recovery Risky Bond

A zero recovery risky bond has the same analogy to the risk-free bond as the hazard rate to the risk-free rate. First we shall look at the risky discount factor.

13.3.1 Risky Discount Factor

A risky discount factor, similar to the risk-free discount factor, discounts \$1 paid in the future. Assuming the same notation for the risk-free discount factor, $P(t,T)$, that represents the present value of \$1 paid in time T, we can denote the survival probability as $Q(t,T)$. A risky discount factor is the present value of \$1 paid in T only if default does not occur. As a result, the present value of \$1 paid in time T is $P(t,T)Q(t,T)$.

13.3.2 Zero Recovery Risky Bond

A zero recovery risky coupon bond pays a periodic coupon c till maturity T_n. If the bond has no recovery, then its price must be:

$$B(t,\underline{T}) = \sum\nolimits_{j=1}^{n} cP(t,T_j)Q(t,T_j) + P(t,T_n)Q(t,T_n) \tag{13.6}$$

where $\underline{T} =< T_1 \cdots T_n >$.

13.4 Positive Recovery Risky Bond

With recovery, the valuation of the risky bond becomes much more complex. Deciding a recovery value of a defaulted bond is a complex process. When a firm defaults, its assets are under a liquidation process and when it ends, bond holders know what they can recover. This process can sometimes take multiple years to finish. For some bond holders who do not wish to wait, they can sell their bonds to the marketplace (distressed bond market) to gain cash earlier. This is similar to Account Receivable factorization.

When such a market exists, then investors estimate a fair present value for the ultimate recovery. This then represents the fair market value of recovery of the bond. When such a market does not exist, then the recovery value must be estimated. It is common for practitioners to use a historical average. Rating agencies provide historical averages for various categories of bonds. For example, senior unsecured bonds recover on average 35% to 45% and junior unsecured bonds recover on average 15% to 25%.

In the literature, there are two major approaches to model positive recovery. Jarrow and Turnbull (1995) assume recovery to be a fixed percentage of the face value of the bond and Duffie and Singleton (1997) assume the recovery to be a fixed percentage of the market value of the bond immediately prior to default. As we shall show later, the Jarrow-Turnbull model is particularly useful in building the credit curve (i.e. bootstrapping) and the Duffie-Singleton model is useful in integrating with the term structure models.

13.4.1 Recovery of Face Value – The Jarrow-Turnbull Model

When the recovery rate is a fixed amount, we can modify the pricing formula of (13.6) as follows:

$$\begin{aligned} B(t,\underline{T}) = \sum_{j=1}^{n} cP(t,T_j)Q(t,T_j) + P(t,T_n)Q(t,T_n) \\ +R\sum_{j=1}^{n} P(t,T_j)[Q(t,T_{j-1}) - Q(t,T_j)] \end{aligned} \tag{13.7}$$

where c is coupon (or cash flow), $P(t,T)$ is risk-free discount factor between now and time T, $Q(t,T)$ is survival probability between now and time T, and R is the recovery rate that is assumed constant. The last term is added due to recovery. Note that $Q(t,T_{j-1}) - Q(t,T_j)$ is the default probability between T_{j-1} and T_j. Note that in continuous time, this is $-dQ(t,T)$ which is equal to $\pi(t,T)dT$. As a result, the above formula can be written in continuous time as:

$$\begin{aligned} B(t,\underline{T}) = \sum_{j=1}^{n} cP(t,T_j)Q(t,T_j) + P(t,T_n)Q(t,T_n) \\ +R\int_t^{T_n} P(t,u)[-dQ(t,u)] \end{aligned} \tag{13.8}$$

The above equation is not a closed-form solution as it requires integration over the default probability measure. One particularly easy way to keep the closed-form solution is to assume the recovery to be received at a fixed time (and not upon default). Then the above equation can be simplified assuming the recovery is received at T_n:

$$B(t,\underline{T}) = \sum_{j=1}^{n} cP(t,T_j)Q(t,T_j) + P(t,T_n)Q(t,T_n) + RP(t,T_n)[1 - Q(t,T_n)] \tag{13.9}$$

The last term $1 - Q(t, T_n)$ is the cumulative default probability. This is the Jarrow-Turnbull model.

We shall demonstrate numerically how the Jarrow-Turnbull model is used in practice, which is known as "bootstrapping" or "curve cooking". The model has become the industry standard in retrieving survival probability information from market quotes (such as credit default swaps, or CDS). To do that the model needs to be slightly adjusted. We shall discuss this in a separate section later.

13.4.2 Recovery of Market Value – The Duffie-Singleton Model

Another easy way to arrive at a closed-form solution is to let the recovery be proportional to the otherwise undefaulted value. That is, upon default (at default time u), the recovery value is $R_t = \delta Z(t, T)$ where $Z(t, T)$ is the price of a zero coupon risky bond as if it has not defaulted.

Under the Poisson process, for a very small time interval Δt, we can write the bond equation as:

$$Z(t,T) = \frac{Z(t+\Delta t,T)\delta\lambda\Delta t + Z(t+\Delta t,T)(1-\lambda\Delta t)}{1+r\Delta t} \tag{13.10}$$

which then can be simplified to, assuming n periods between now t and maturity T:

$$\begin{aligned} Z(t,T) &= \frac{Z(t+\Delta t,T)(1-\lambda\Delta t(1-\delta))}{1+r\Delta t} \\ &= Z(T,T)\left[\frac{1-\lambda\Delta t(1-\delta)}{1+r\Delta t}\right]^n \\ &\sim Z(T,T)\left[\frac{e^{-\lambda\Delta t(1-\delta)}}{e^{r\Delta t}}\right]^n \\ &\sim Z(T,T)e^{-(r+s)(T-t)} \end{aligned} \tag{13.11}$$

where $s = \lambda(1-\delta)$ can be viewed as a spread over the risk-free rate. This is the Duffie-Singleton model. $Z(T, T)$ is the terminal value of the bond which is usually the face value.

A nice feature of the Duffie-Singleton model is that a coupon bond can then a portfolio of such zeros, as each coupon is treated as a zero bond it recovers market value. That is:

$$B(t,\underline{T}) = \sum\nolimits_{j=1}^{n} cZ(t,T_j) + Z(t,T_n) \tag{13.12}$$

This model is practically appealing in that it reflects the usual industry practice that credit risk is reflected in spreads. This model is also convenient to be combined with existing term structure models. It simply adds a second state variable.

The drawback of the model is that the recovery parameter and the intensity parameter always are inseparable. This adds to difficulty in calibration this model to the market.

Both the Jarrow-Turnbull and the Duffie-Singleton models assume defaults to be unexpected like Poisson events. Different from the Jarrow-Turnbull model that assumes fixed amount recovery (or known as recovery of face value), the Duffie-Singleton model assumes the recovery to be proportional to the market value of the debt (known as recovery of market value) immediately prior to default.

The Jarrow-Turnbull model is suitable for bootstrapping and the Duffie-Singleton model is convenient to combine with term structure models. Following the similar analysis for equations (13.2) $\sim$ (13.4), we can write (13.11) as:

$$Z(t,T) = \hat{\mathbb{E}}_t\left[\exp\left(-\int_t^T (r(u)+s(u))du\right)\right] \tag{13.13}$$

which allows us to directly model spread as another state variable. Note that $s(t) = \lambda(t)(1-\delta)$ according to (13.11) and hence the spread process is similar to the intensity process. If the intensity and the risk-free rate are independent, then the Duffie-Singleton model of (13.12) is similar to (13.6) but with positive recovery.

We can easily conduct a CIR model with the Duffie-Singleton approach. We can have the following joint square-root process:

$$\begin{aligned} dr(t) &= \hat{\alpha}_r(\hat{\mu}_r - r(t))dt + \sigma_r\sqrt{r(t)}d\hat{W}_r(t) \\ ds(t) &= \hat{\alpha}_s(\hat{\mu}_s - s(t))dt + \sigma_s\sqrt{s(t)}d\hat{W}_s(t) \end{aligned} \tag{13.14}$$

where $d\hat{W}_r(t)d\hat{W}_s(t) = 0$. Then (13.13) has a closed-form solution as each expectation in the following equation is a CIR solution.

$$Z(t,T) = \hat{\mathbb{E}}_t\left[\exp\left(-\int_t^T r(u)du\right)\right]\hat{\mathbb{E}}_t\left[\exp\left(-\int_t^T s(u)du\right)\right] \tag{13.15}$$

The following example is to demonstrate how the Duffie-Singleton model can be easily combined with any interest rate model. In the following a simple binomial model for the risk-free rate is given. The probabilities of the up and down branches are $\frac{1}{2}$ and $\frac{1}{2}$.

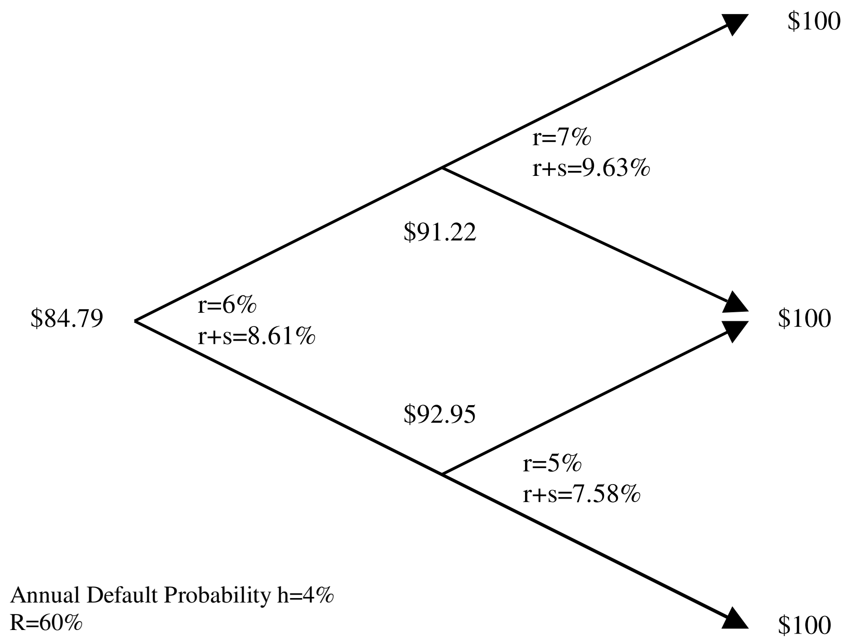


Figure 13.1: Duffie-Singleton Model

From the input information provided, we can compute the bond prices in the diagram, as follows:

$$91.215 = \frac{1}{1.07}[(1-4\%) \times 100 + 4\% \times 40]$$
$$92.95 = \frac{1}{1.05}[(1-4\%) \times 100 + 4\% \times 40]$$
$$84.79 = \frac{1}{1.06}\left[(1-4\%) \times \frac{91.215+92.95}{2} + 4\% \times \frac{91.215+92.95}{2} \times 0.4\right]$$

The spreads of these bonds are computed as follows:

$$\frac{100}{91.215} - 1 - 7\% = 9.63\% - 7\% = 2.63\%$$
$$\frac{100}{92.95} - 1 - 5\% = 7.58\% - 5\% = 2.58\%$$
$$\frac{92.0825}{84.79} - 1 - 6\% = 8.61\% - 6\% = 2.61\%$$

where $92.0825 = \frac{1}{2}(91.215 + 92.95)$. Note that these spreads are not (even though close to) the continuous spread in the Duffie-Singleton model, which is computed as follows:

$$\frac{1}{\text{Default Prob} \times \text{Recovery} + \text{Survival Prob}} - 1$$
$$= \frac{1}{4\% \times 0.4 + 96\%} - 1$$
$$= 2.46\%$$

13.5 Credit Default Swap

A Credit Default Swap, or CDS, is a bilateral contract which allows an investor to buy protection against the risk of default of a specified reference credit. The fee may be paid up front, but more often is paid in a "swapped" form as a regular, accruing cashflow. A CDS is a negotiated contract and there are a number of important features that need to be agreed between the counterparties and clearly defined in the contract documentation.

First and foremost is the definition of the credit event itself. This is obviously closely linked to the choice of the reference credit and will include such events as bankruptcy, insolvency, receivership, restructuring of debt and a material change in the credit spread. This last materiality clause ensures that the triggering event has indeed affected the price of the reference asset. It is generally defined in spread terms since a fall in the price of the reference asset could also be due to an increase in the level of interest rates.

Many CDS contracts define the triggering of a credit event using a reference asset. However, in many cases the importance of the reference asset is secondary as the credit event may also be defined with respect to a class of debt issued by a reference entity. In this case the importance of the reference asset arises solely from its use in the determination of the recovery price used to calculate the payment following the credit event.

The contract must specify what happens if the credit event occurs. Typically, the protection buyer will usually agree to do one of the following:

- Deliver the defaulted security to the protection seller in return for Par in cash. Note that the contract usually specifies a basket of securities which are ranked *pari passu* which may be delivered in place of the reference asset. In effect the protection seller is long a "cheapest to deliver" option.

- Receive Par minus the default price of the reference asset settled in cash. The price of the defaulted asset is typically determined via a dealer poll conducted within a few weeks to months of the credit event, the purpose of the delay being to let the recovery price stabilize.

Some CDS have a different payoff from the standard Par minus recovery price. The main alternative is to have a fixed pre-determined amount which is paid out immediately after the credit event. This is known as a binary default swap. In other cases, where the reference asset is trading at a significant premium or discount to Par, the payoff may be tailored to be the difference between the initial price of the reference asset and the recovery price.

The protection buyer automatically stops paying the premium once the credit event has occurred, and this property has to be factored into the cost of the swap payments. It has the benefit of enabling both parties to close out their positions soon after the credit event and so eliminates the ongoing administrative costs which would otherwise occur.

A CDS can be viewed as a form of insurance with one important advantage – efficiency. Provided the credit event in the default swap documentation is defined clearly, the payment due from the triggering of the credit event will be made quickly. Contrast this with the potentially long and drawn out process of investigation and negotiation which may occur with more traditional insurance.

However it is possible to get a very good idea of the price of the CDS using a simple "static replication" argument. This involves recognizing that buying a CDS on a risky par floating rate asset which only defaults on coupon dates is exactly equivalent to going long a default-free floating rate note and short a risky floating rate note of the same credit quality. If no default occurs, the holder of the position makes a net payment equal to the asset swap spread of the asset on each coupon date until maturity. This spread represents the credit quality of the risky floater at issuance. If default does occur, and we assume that it can only occur on coupon payment dates, the position can be closed out by buying back the defaulted asset in return for the recovery rate and selling the par floater. The net value of the position is equal to the payoff from the default swap. The following table summarizes.

CDS vs. Floater			
Event	Riskless FRN	Risky FRN	CDS
At inception	Pay par	Pay par	0
No default	Receive LIBOR	Receive LIBOR + spread	Pay spread
Upon default	Receive par	Receive recovery	+ par − recovery
Maturity	Receive par	Receive par	0

From the above table, it is clear that the spread of a CDS must equal to the spread of the equivalent risky FRN to avoid arbitrage.

13.6 Restructuring Definitions by ISDA

CDS contracts provide default protection. When a default occurs, CDS buyers stop paying the premium (spread), deliver the defaulted bond (cheapest if possible), and collect full face value as payment. However, default is hard to define. It is extremely rare for a company to file bankruptcy. What is usually happening is that losses happen over the years and reduce the asset value of the company, to a point where the company is at the brink of bankruptcy. Then the management of the company will start looking for alternatives to save the firm. One popular alternative to save the firm is to ask debt holders to change their debt contracts to the company known as debt restructuring. Debt restructuring often means that debt holders convert parts of their debts into equity and participate in the operation of the firm. To protect their own interests, debt holders, especially large ones, will be willing to agree to debt restructuring.

Hence debt restructuring is commonly regarded as a form of default. However, each restructuring can be very different. Some restructurings are major and equivalent to defaults. But some could be minor as precautionary actions to avoid further deterioration of the firm, which are not equivalent to default.

To regulate if a CDS is triggered, ISDA (International Swaps and Derivatives Association) defines various restructuring standards:

- Full restructuring (FR), based on the ISDA 1999 Definition
- Modified restructuring (MR), based on the ISDA 2001 Supplement Definition
- Modified-modified restructuring (MMR), based on the ISDA 2003 Definition,
- No restructuring (NR).
- The definitions are as follows:

  FR Any bond of maturity up to 30 years

  MR $T \leqslant \bar{T} < T + 30$ months

  MMR Allow additional 30 months for the restructured bond.

  For other obligations, same as MR.

CDS contracts traded in different regions follow different ISDA conventions.

13.7 Why Has the CDS Market Developed So Rapidly?

CDS is the most popular credit derivatives contract and has grown rapidly in late 90's and early 00's. The following is a direct quote of Rene Stulz's article on CDS (2010):[2]

Back in the mid-1990s, one of the first credit default swaps provided protection on Exxon by the European Bank for Reconstruction and Development to JP Morgan (Tett, 2009). It took months to negotiate. By 1998, the total size of the credit default swap market was a relatively small $180 billion (Acharya, Engle, Figlewski, Lynch, and Subrahmanyam, 2009). The credit default swap market has grown enormously since then, although there is no definite measure of how much. Based on survey data from the Bank for International Settlements (BIS) at http://www.bis.org/statistics/derstats.htm, the total notional amount of the credit default swap market was $6 trillion in 2004, $57 trillion by June 2008, and $41 trillion by the end of 2008. Credit-default swap contracts that insure default risk of a single firm are called single-name contracts; in contrast, contracts that provide protection against the default of many firms are called multi-name contracts. [3]

In addition to the efficiency in hedging and transferring credit risk, the potential benefits of CDS include:

- A short positioning vehicle that does not require an initial cash outlay
- Access to maturity exposures not available in the cash market
- Access to credit risk not available in the cash market due to a limited supply of the underlying bonds
- The ability to effectively exit credit positions in periods of low liquidity
- Off-balance sheet instruments which offer flexibility in terms of leverage
- To provide important anonymity when shorting an underlying credit

[2] "Credit Default Swaps and the Credit Crisis," Journal of Economic Perspectives, Volume 24, Number 1, Winter 2010, pp. 73-92.

[3] Stulz noted that DTCC statics are a lot smaller.

13.8 Relationship between Default Probabilities and CDS Spreads – Use of the Jarrow-Turnbull Model

There is a simple formula (using the Jarrow-Turnbull model) that relates the CDS spread, the risk-free rate, default/survival probabilities, and the fixed recovery rate. Due to the swap nature, CDS, similar to IRS (interest rate swap), has two legs – floating and fixed. The floating leg of a CDS contract is called the protection leg as it pays only if default occurs. The fixed leg of a CDS contract is called the premium leg because the fixed payments (i.e. spreads) are like insurance premiums. As in a standard swap contract, at inception, the values of the two legs must equal each other. This is how CDS spreads are calculated. Recently affected by the crisis, CDS premiums have been split into an upfront and a spread (which is the index, such as CDX, trading convention). As we shall see later, this extra calculation does not add any complexity to the model. We simply deduct the upfront amount from the protection value of the CDS. For now, we shall proceed with no upfront.

Using the formulation given earlier, the protection and premium values of a CDS are as follows:

$$\begin{aligned} V_{\text{prot}}(t,T) &= (1-R)\sum\nolimits_{i=1}^{n} P(t,T_i)[Q(t,T_{i-1})-Q(t,T_i)] \\ V_{\text{prem}}(t,T) &= s(t,T)\sum\nolimits_{i=1}^{n} P(t,T_i)Q(t,T_i) \end{aligned} \tag{13.16}$$

where $T_n = T$. As a result, the spread (known as par spread) can be computed as:

$$s(t,T) = \frac{(1-R)\sum_{i=1}^{n} P(t,T_i)[Q(t,T_{i-1})-Q(t,T_i)]}{\sum_{i=1}^{n} P(t,T_i)Q(t,T_i)} \tag{13.17}$$

Note that (13.17) (for CDS) and (13.7) (for bond) are very similar. The numerator of (13.17) is similar to the recovery value in (the second line of) (13.7) and the denominator is similar to the (first line of (13.7) coupon value. This should not be surprising as CDS is a natural hedge to the bond. In other words, buy a bond and a CDS is equivalent to buying a default-free bond. If we add the protection value in (13.16) to the coupon bond value in (13.7), recovery disappears and the bond as a result becomes default-free.

Note that (13.17) assumes no accrued interest if default occurs in between coupons. In reality there are accrued interests on both legs and they may not be equal. If default is assumed to happen on cashflow days only, then there is no accrued interest. Note that if there is an upfront, we simply deduct it from the protection value, V_{prot}.

13.9 Back-of-the-envelope Formula

In a one-period model where default is a Bernulli event, as the following picture demonstrates,

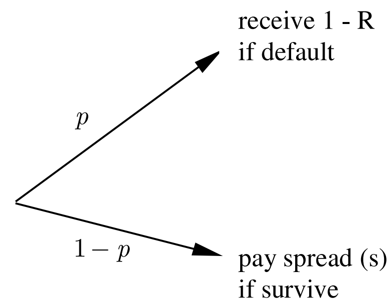


Figure 13.2: One-period Default Diagram

We know that for the CDS to have 0 value it must be true that:

$$p(1 - R) = (1 - p)s \tag{13.18}$$

Note that risk-free discount cancel from both sides. Hence, we arrive at the famous back-of-the-envelope formula for the default probability (by ignoring the term $p \times s$ which is small):

$$p = \frac{s}{1 - R} \tag{13.19}$$

This formula, while simple, provides a powerful intuition of spreads and default probabilities. If the recovery is 0, then spread is (forward) default probability. This is not only true in (13.19) but also true in continuous time. Spreads are not equal to (i.e. smaller) forward default probabilities in that they are compensated by recovery. Note that CDS buyers acquire default protection by paying spreads as an insurance premium. If recovery is high, the protection value is low, and so should be the spread. In an extreme case where the recovery is 100%, the spread should be 0, which is suggested by (13.19).

13.10 Bootstrapping (Curve Cooking)

We need credit curves to price credit derivatives. Credit curves are obtained from liquid cash products such as CDS or corporate bonds. Due to the liquidity concern, CDS is a better choice for curve cooking.

The basic bootstrapping idea of constructing a risky curve is the same as the risk-free curve. We use a pricing formula to back out the parameter(s). In the traditional fixed income world (Treasuries and IRS), we back out spot and forward rates from the market prices of bonds and swap rates. Here, we back out survival probabilities from a series of CDS contracts. As in the world of traditional fixed income, we need a term structure of CDS spreads in order to back out the entire survival probability curve. A popular smoothing technique in LIBOR curves is piece-wise flat.

The CDS market has been standardized over the years to have the following on-the-run maturities: 1, 2, 3, 5, 7, and 10 years to maturities. As in the IRS market, these contracts are on-the-run which are issued periodically. In the early years, only 5-year CDS contracts were issued. A few years ago, the market started to trade 10-year CDS contracts. The other maturities have gradually been introduced to the market but their liquidity is still a concern. Assume for now that we observe market prices of these CDS spreads.

13.11 Poisson Assumption

From (13.2) $\sim$ (13.4), we know that if we assume piece-wise flat intensity values, then the survival probability can be approximated as follows:

$$Q(t, T_n) = \exp\left\{-\sum\nolimits_{i=1}^{n} \lambda_i (T_i - T_{i-1})\right\} \tag{13.20}$$

where $t = T_0$.

13.12 Simple Demonstration (annual frequency)

To make matters simple, we assume CDS spreads are paid annually. There are 6 CDS spreads observed in the market (1, 2, 3, 5, 7, 10). Take Disney as an example. On 12/23/2005, we observe the following spreads (in basis points):

CDS quotes	
term	sprd
1	9
2	13
3	20
5	33
7	47
10	61

Continue to assume fixed 0.4 recovery ratio under MR for Disney. The following binomial chart presents possible cash flows. Lets assume 5% risk-free rate.

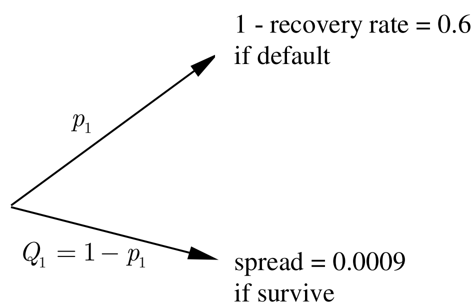


Figure 13.3: First period

CDS is a swap contract so there is no cash changed hands on day 1. Hence, it must be the case that the expected payment (9 basis points) equals the expected compensation (60%). In a single period, since both payment and compensation are discounted, the risk-free does not matter. Note that the survival probability for one year is $Q_1 = 1 - p_1$ (which is also equal to $e^{-\lambda_1}$ if we assume the Poisson process for defaults). Hence, using 5% interest rate, we have:

$$\frac{0.6 \times (1 - Q_1) - 0.0009 \times Q_1}{1.05} = 0$$

which is solved as:

$$0.6 \times (1 - Q_1) = 0.0009 \times Q_1$$

and $Q_1 = 0.9985$. λ_1 can be solved for as $-\ln Q_1$ to be 0.001499, or 14.99 basis points. Hence, the survival and default probabilities are 99.85% and 0.15% respectively.

Now we proceed to bootstrap the second period.

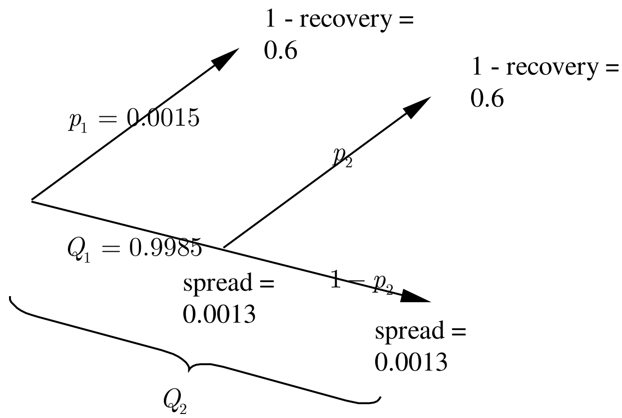


Figure 13.4: Two-period Default Diagram

There are three scenarios. Either Disney defaults in period 1, or default in period 2, or survive in period 2 (note that to survive till period 2, Disney must first survive period 1). We know the first default probability, which is $p_1 = 0.0015$. But we do not know the other two probabilities. Using the same Poisson algorithm, we can compute the second year present value as:

$$\frac{0.6 \times p_2 - 0.0013 \times (1 - p_2)}{1.05}$$

Note that this quantity itself is not 0; but combining it with the first year cash flows is:

$$\frac{0.6 \times 0.0015 + 0.9985 \left(-0.0013 + \frac{0.6 \times p_2 - 0.0013 \times (1 - p_2)}{1.05} \right)}{1.05} = 0$$

Solve for p_2 to get 28.62 basis points. Under the Poisson assumption, $e^{-\lambda_2} = 1 - p_2$ and as a result, $\lambda_2 = -\ln(1 - p_2) = 0.002865$ or 28.65 basis points. The conditional survival probability is $1 - p_2 = 99.7138\%$. The unconditional survival probability, Q_2, equals $(1-p_1)(1-p_2)$ which is 99.5645%. The unconditional default probability is $1 - 99.5645\% = 0.4355\%$.

Similar process applies to all periods as the following figure depicts. Due to the limitation of the space in the table, $Q(t, v)$ is replaced with $Q(\tau)$ where $\tau = v - t$.

The results are given as follows. The first three columns are taken from the

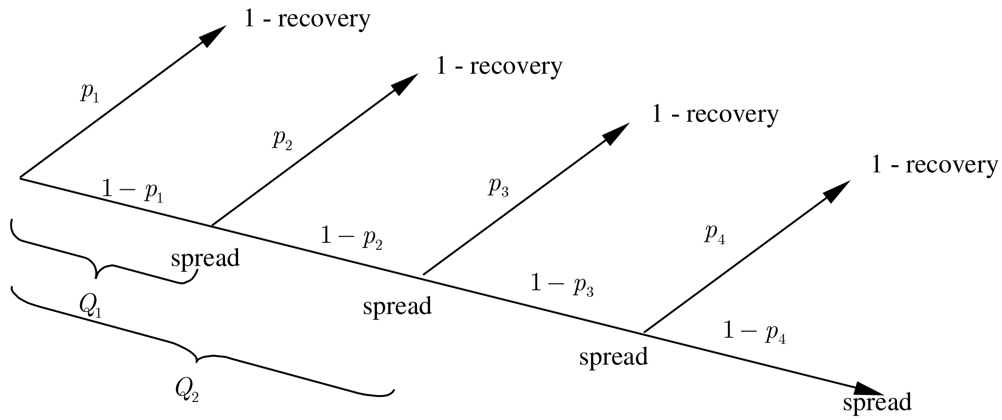


Figure 13.5: Multi-period Default Diagram

market. Columns A and B are the same spread input given earlier. Column C is the risk-free discount factors that are computed using 5% flat in the example.

CDS Bootstrapping					
A	B	C	D	E	F
Term τ	Market Spread	Risk-free $P(\tau)$	Fwd. $\lambda(\tau)$	Surv.Pr. $Q(\tau)$	Def.Pr. $-dQ(\tau)$
1	0.0009	0.9512	0.0015	0.9985	0.0015
2	0.0013	0.9048	0.0029	0.9956	0.0029
3	0.0020	0.8607	0.0059	0.9898	0.0058
4		0.8187	0.0092	0.9808	0.0091
5	0.0033	0.7788	0.0092	0.9718	0.0090
6		0.7408	0.0150	0.9573	0.0144
7	0.0047	0.7047	0.0150	0.9431	0.0142
8		0.6703	0.0176	0.9267	0.0164
9		0.6376	0.0176	0.9106	0.0161
10	0.0061	0.6065	0.0176	0.8948	0.0158

CDS Bootstrapping (cont'ed)					
G	H	I	J	K	L
$P(\tau)Q(\tau)$	$P(\tau)\times$ $[-dQ(\tau)]$	Prem. Leg	Prot. Leg	Model Spread	Cond. Def.Pr.
0.9498	0.0014	0.9498	0.0009	0.0009	0.0015
0.9009	0.0026	1.8507	0.0024	0.0013	0.0029
0.8520	0.0050	2.7027	0.0054	0.0020	0.0058
0.8030	0.0074	3.5056	0.0099	0.0028	0.0092
0.7568	0.0070	4.2625	0.0141	0.0033	0.0092
0.7092	0.0107	4.9717	0.0205	0.0041	0.0148
0.6646	0.0100	5.6363	0.0265	0.0047	0.0148
0.6212	0.0110	6.2575	0.0331	0.0053	0.0174
0.5806	0.0103	6.8381	0.0393	0.0057	0.0174
0.5427	0.0096	7.3808	0.0450	0.0061	0.0174

Column E presents the survival probabilities that are computed sequentially as in equation (13.20):

$$\begin{aligned} Q(t,T_n) &= \exp\left\{-\sum\nolimits_{i=1}^{n}\lambda_i(T_i-T_{i-1})\right\} \\ &= Q(t,T_{n-1})\exp\left\{-\lambda_n(T_n-T_{n-1})\right\} \end{aligned} \tag{13.21}$$

For example, $Q(0,2) = 0.9956 = 0.9985 \times e^{-0.0029\times(2-1)}$. Column F is the *unconditional* default probabilities which is the differences are two consecutive survival probabilities. For example, $0.0029 = 0.9985 - 0.9956$. Column G is known as the risky discount factor (introduecd earlier), or \$1 present value with default risk. These values are needed in order to compute the default swap spread, i.e. the denominator of (13.17). Similarly, column H provides the values for the numerator of (13.17). Columns I and J are accumulations of columns G and H respectively. Column K is the division of column J by column I, which is the spread of CDS. The values of this column must match the market quotes in column B. In fact, we solve for column D so that values in column K are identical to values in column B. Finally, column L contains conditional default probabilities, each of which equals $1-Q(t,T_j)/Q(t,T_{j-1})$. We note that the conditional default probabilities are close to the intensity values (λ), as they should, in that they are exactly equal in continuous time.

The table presented here provides a nice algorithm for further automate the calculations for more complex situations in reality, which we shall demonstrate later. Once all the λ_j values are bootstrapped out, we can then compute any survival probability of any future time, as follows:

$$Q(t, v) = Q(t, T_{n-1}) \exp\{-\lambda_n(v - T_{n-1})\} \tag{13.22}$$

where $T_{n-1} < v < T_n$. For example, the survival probability of 6.25 years is $0.9769 \times e^{-0.0098\times(6.5-5)} = 0.9627$.

13.13 In Reality (quarterly frequency)

In the above example, we assume spreads are paid annually. As a result, the calculation is quite simple. In reality, this is not the case. Spreads are paid by the swap market convention which is quarterly. In this case, default can occur at any quarter. We then need to alter the one period calculation shown above.

Note that within a year (for the first few spreads), all per-quarter default probabilities are equal. This is because we have only one spread (e.g. 0.0009 in year 1) to cover four quarters. The basic formula is still the same. Mainly we solve the following equation for p (note that at each period, the interest rate is 1.25%):

$$\begin{aligned} x_i &= x_0 + 0.6p_1 + (1 - p_1)x_{i-1} \\ x_n &= x_0 \end{aligned} \tag{13.23}$$

where n represents the number of periods that shares the probability. For the first year, $n = 4$ and $x_0 = 0.000225$. We then solve for $p_1 = 0.000375$ or 3.75 basis points and $\lambda_1 = 0.0015$. The full expansion of this equation is shown in the Appendix.

Note that while this equation is solvable by hand if proper re-arrangement of terms is performed, it is much faster if we set up the equation and use the Solver in Excel. This equation can be set up recursively as the discounting and expected values are nested. We can set up an Excel sheet to compute all the results. As a demonstration, we provide the results up the 3 years.

The layout of the table is the same as before. The frequency of the CDS premium payments is now quarterly (see column A). Column B is still market CDS spreads that are available every four quarters. Column C is quarterly risk-free discount factors (at 5%). Columns E $\sim$ H are computed similarly to the previous section, only with quarterly frequency.

Columns, I, J, and K are computed similarly to the previous section but only every year. Note that column K is column J $\div$ column I $\times$ 4 in order to annualize to match annual CDS market quotes in column B.

Readers should complete the table for the full 10 years.

CDS Bootstrapping					
A	B	C	D	E	F
Term	Market Spread	Risk-free $P(\tau)$	Fwd. $\lambda(\tau)$	Surv.Pr. $Q(\tau)$	Fwd.Def.Pr. $-dQ(\tau)$
0.25		0.9876	0.0015	0.9996	0.000375
0.5		0.9753	0.0015	0.9993	0.000375
0.75		0.9632	0.0015	0.9989	0.000375
1	0.0009	0.9512	0.0015	0.9985	0.000374
1.25		0.9394	0.002868	0.9978	0.000716
1.5		0.9277	0.002868	0.9971	0.000715
1.75		0.9162	0.002868	0.9964	0.000715
2	0.0013	0.9048	0.002868	0.9956	0.000714
2.25		0.8936	0.00586	0.9942	0.001457
2.5		0.8825	0.00586	0.9927	0.001455
2.75		0.8715	0.00586	0.9913	0.001453
3	0.002	0.8607	0.00586	0.9898	0.001451

CDS Bootstrapping					
G	H	I	J	K	L
$P(\tau)Q(t)$	$P(\tau)[-dQ(\tau)]$	premium leg	protection leg	computed spread	cond. surv.prob.
0.9872	0.0004				
0.9746	0.0004				0.9996
0.9621	0.0004				0.9996
0.9498	0.0004	3.8737	0.0009	0.0009	0.9996
0.9373	0.0007				0.9993
0.925	0.0007				0.9993
0.9129	0.0007				0.9993
0.9009	0.0006	7.5498	0.0025	0.0013	0.9993
0.8884	0.0013				0.9985
0.8761	0.0013				0.9985
0.8639	0.0013				0.9985
0.8519	0.0012	11.0302	0.0055	0.002	0.9985

Chapter 14

Corporate Finance Approach of Modeling Default

14.1 Merton Model

This approach of modeling default is called structural approach. It does not assume defaults as an unexpected event but rather a firm's failure to pay its debts. The earliest model is Black-Scholes (1973) and Merton (1974) who regard the equity of the firm as a call option. This is because the equity of the firm has a residual claim of the firm's assets. In a simple example where the firm has only one zero coupon debt, the face value of the debt is the strike of the call option. When the asset value at the maturity of the debt is above the face value, the firm is able to pay off its debt and the equity receives the residual value of $A(T) - K$ where A is the asset value of the firm at the maturity date, T, of the debt and K is the face value of the debt. Otherwise (the asset value is smaller than the face value), the equity is worthless.

Figure 14.1 describes the payoffs graphically. The vertical axis is value of the assets of the firm. The horizontal axis is time. The horizontal line represents the face value of the debt, which is flat over time. In the example, a one-year horizon, weekly intervals are used. In Figure 14.1, two samples of the asset value are presented – one good and one bad. The good sample is a profitable sample where the firm does well and the asset value increases (and is higher than the debt face value at time T). The bad sample is where the firm loses money and the asset value ends up lower than the debt face value. In the good sample, the debt is paid in full and the equity receives the residual amount $A(T) - K$. In the bad sample, the firm defaults anfd the debt holder receives a recovery value of $A(T)$ and the equity holder receives nothing. On the very right of the diagram is the log normal distribution that describes the asset

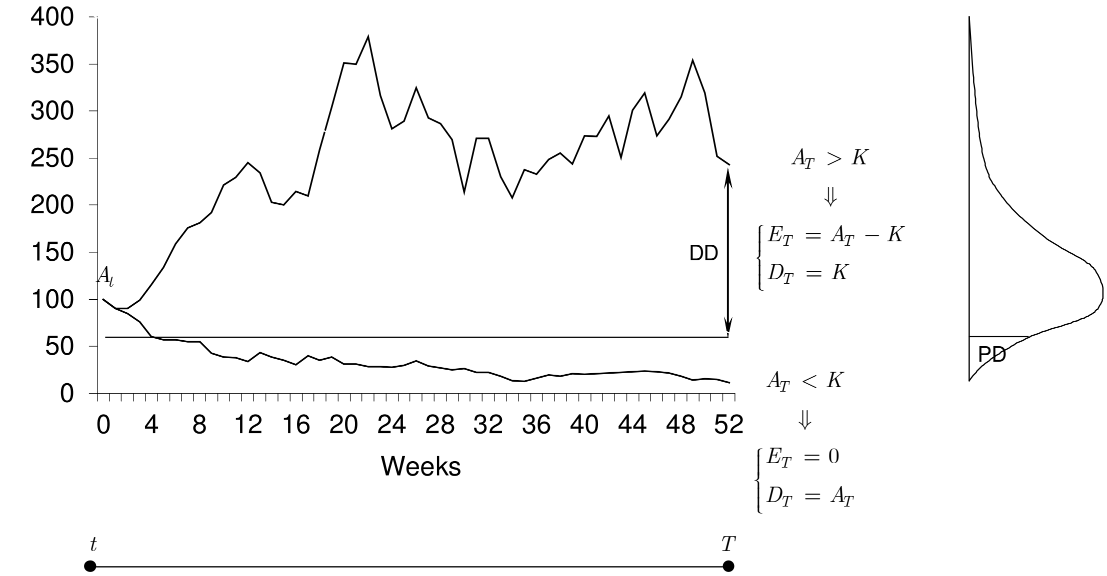


Figure 14.1: Merton Model

price. The area below the default barrier is the probability of default (commonly abbreviated as PD). As a result, the equity value and equity volatility in this simple case can be priced by the Black-Scholes model:

$$\begin{cases} E(t) = A(t)N(d_1) - e^{-r(T-t)}KN(d_2) \\ \sigma_E = \sigma N(d_1)\dfrac{A}{E} \end{cases} \tag{14.1}$$

where σ_E is the equity volatility and:

$$\begin{aligned} d_2 &= \tfrac{\ln A(t) - \ln K + (r - \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \\ d_1 &= d_2 + \sigma\sqrt{T-t} \end{aligned}$$

r is the risk-free rate, and σ is the asset return volatility. The debt value which is the difference between asset and equity is then equal to:

$$\begin{aligned} D(t,T) &= A(t) - E(t) \\ &= \underbrace{A(t)[1 - N(d_1)]}_{\text{Recovery value}} + \underbrace{e^{-r(T-t)}KN(d_2)}_{\text{Survival value}} \end{aligned} \tag{14.2}$$

On the right hand side of the equation, the debt value is decomposed into two parts – surivial value and recovery value. The survival value is the amount received by the debt holders if the firm survives. In this case it is the face value weighted by the survival (in-the-money) probability $N(d_2)$ and discounted at the risk-free rate. The recovery value is the amount received by the debt holders if the firm defaults. In this case it is the asset value weighted by the default probability and discounted at the risk-free rate. But since the asset value is random, we must compute the following expected value: $e^{-r(T-t)}\hat{\mathbb{E}}_t\left[A(T)\mathbb{I}_{A<K}\right]$ which is equal to the recovery value given above.

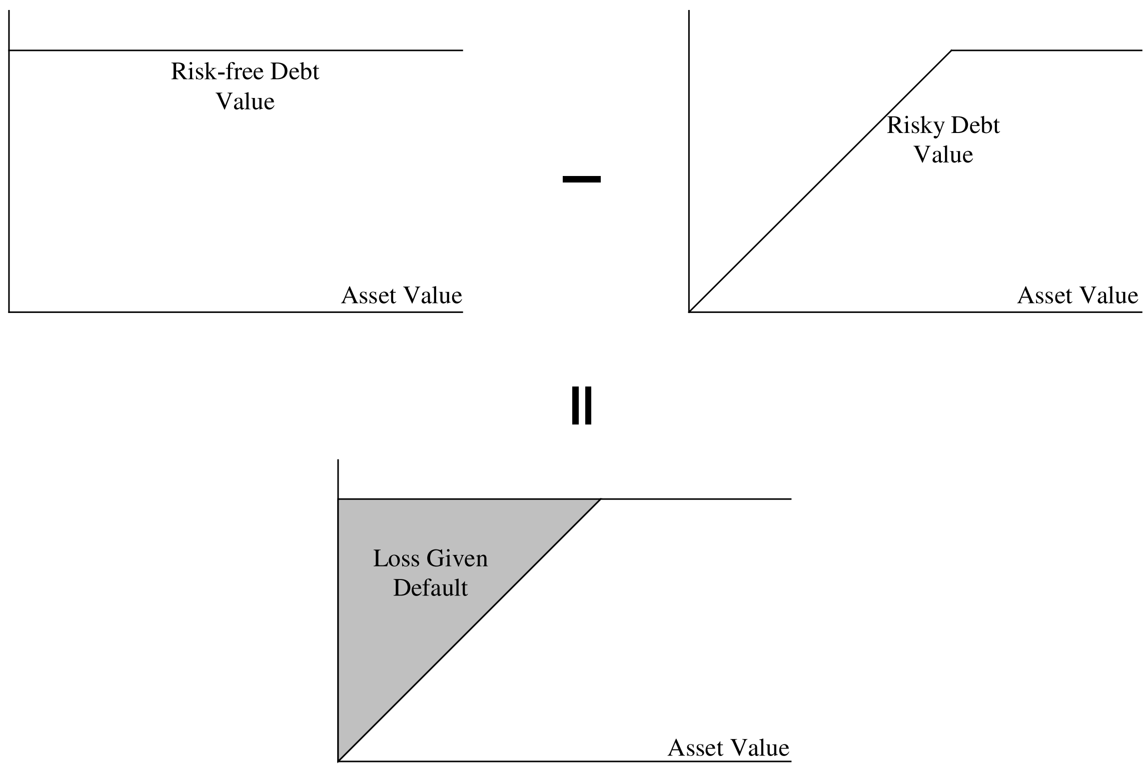


Figure 14.2: Risky Debt

Figure 14.2 describes the equation graphically. The upper-left panel is the risk-free debt whose payoff is fixed regardless of the value of the firm. The upper-right panel is the payoff of the risky debt. When the firm is doing well (i.e. survival), the payment is full (i.e. face value) but when the firm is not doing well (i.e. default), the payment is the recovery value (i.e. asset value). The bottom panel is the difference between the risk-free debt and the risky debt, which gives the result of the loss given default. The convolution of the area and the probabilities is the expected loss (EL).

Note that $D(t,T) < e^{-r(T-t)}K$ which is the risk-free debt value. The (risk-neutral) default probability is $p = 1 - N(d_2)$. The following diagram is usually used to visualize how the Black-Scholes-Merton model works. The plotted are two possible sample paths of the asset value, one drifts upward and the other drifts downward. The asset value starts at 100 with the default barrier at 50. One sample path stays above the default barrier and is a survival path and one crosses the barrier and is a default path. To be explored fully later, the log difference between the asset value and the default barrier is known as the distance to default, or DD.

14.2 KMV Model

In 1989 Stephen Kealhofer, John McQuown and Oldrich Vasicek founded company KMV. They commercialize the Merton (Black-Scholes) model to provide quantitative credit ratings of companies. They coined the term EDF (expected default frequency) using DD. They use equation (14.1) to compute the distance to default (DD) which is d_2, because:

$$d_2 = \frac{\mathbb{E}[\ln A(T)] - \ln K}{\sqrt{\mathbb{V}[\ln A(T)]}} \tag{14.3}$$

Given that $A(T)$ follows a log normal process, this DD makes sense. And naturally the probability of default (PD) is $1 - N(d_2)$, or $N(-d_2)$. The expected recovery (since actual recovery $A(T)$ is random) is $A(t)[1 - N(d_1)]$.

The Merton model allows for only one single debt. To accommodate the multi-debt problem, KMV empirically estimated the one-year equivalent debt value to be:

$$\begin{aligned}\text{One Year Equivalent Debt} &= \text{Short Term Debt} + \frac{\text{Long Term Debt}}{2} \\ &= \frac{\text{Short Term Debt} + \text{Total Debt}}{2}\end{aligned}$$

EDF is similar to PD (i.e. $1 - N(d_2)$, or $N(-d_2)$) (but the translation is a KMV secret) and used to decide a rating for the company. KMV also use the model to predict defaults. See Stein (2002).

Recently, KMV also uses barrier option models (details later) to better capture multiple debts and complex volatility forecast methods (such as GARCH) to improve the historical equity volatility estimates that are not very reliable.

14.3 The Geske Model

Although the Black-Scholes-Merton model reflects the reality that the only reason firms default is due to failure to pay, it is a one period model which is too limited to be used for a regular firm that has multiple debts.

Geske extended the model to include multiple debts. Take a two-period example and assume the firm issues two zero coupon bonds, expiring at T_1 and T_2, with face values K_1 and K_2 respectively. Default at T_1 is defined by Geske (1984) as the firm value less than the face value of the first debt plus the market value of the second debt, that is $V_1 < K_1 + D(T_1, T_2)$ where $D(T_1, T_2)$ is the market value of K_2 at time T_1.

The solution to the equity value and equity volatility can be derived from Geske's compound call option (call on call) model:[1]

$$
\begin{aligned}
E(t) &= A(t)M(h_{1+}, h_{2+}; \rho) - e^{-r(T_2-t)}K_2 M(h_{1-}, h_{2-}; \rho) - e^{-r(T_1-t)}K_1 N(h_{1-}) \\
\sigma_E &= \sigma M(h_{1+}, h_{2+}; \rho)\frac{A}{E}
\end{aligned}
\tag{14.4}
$$

where $N(\cdot)$ is the uni-variate standard normal probability and $M(\cdot, \cdot; \rho)$ is the bi-variate standard normal probability, and:

$$
\begin{aligned}
\rho &= \sqrt{\frac{T_1 - t}{T_2 - t}} \\
h_{j\pm} &= \frac{\ln A(t) - \ln \bar{A}_j + (r \pm \sigma^2/2)(T_j - t)}{\sigma\sqrt{T_j - t}}
\end{aligned}
\tag{14.5}
$$

Note that $\bar{A}_2 = K_2$ is the critical value for the assets to trigger default at T_2 and $\bar{A}_1$ is the solution to $K_1 = E(T_1)$, as follows:

$$
\begin{aligned}
K_1 &= \bar{A}(T_1)N(x_+) - e^{-r(T_2-T_1)}K_2 N(x_-) \\
x_\pm &= \frac{\ln \bar{A}(T_1) - \ln K_2 + (r \pm \sigma^2/2)(T_2 - T_1)}{\sigma\sqrt{T_2 - T_1}}
\end{aligned}
$$

which is the critical value for default at time T_1; and . Recall that $E(T_1)$ is the Black-Scholes value (due to the fact that in the last period, the equity is just a

[1] Note that the equation indicates that the survival probabilities are $N(h_{1-})$ and $M(h_{1-}, h_{2-}; \rho)$ for one year and two years respectively.

simple call option). The two debt values are a bit complex to solve, but the closed-form solutions are given as follows:

$$
\begin{aligned}
D(t,T_1) &= A(t)[1-N(y_{1+})] + e^{-r(T_1-t)}K_1 N(y_{1-}) \\
D(t,T_2) &= A(t)[N(y_{1+}) - M(h_{1+},h_{2+};\rho)] - e^{-r(T_1-t)}K_1[N(y_{1-}) - N(h_{1-})] \\
&\qquad +e^{-r(T_2-t)}K_2 M(h_{1-},h_{2-};\rho)
\end{aligned}
\tag{14.6}
$$

where

$$
y_{j\pm} = \frac{\ln A_{j-1} - \ln K_j + (r \pm \sigma^2/2)(T_j - T_{j-1})}{\sigma\sqrt{T_j - T_{j-1}}}
$$

It is clear that the expected recover value is $A(t)[1-M(h_{1+},h_{2+};\rho)]$ for the entire firm; $A_0[1-N(y_{1+})]$ for debt K_1; and $A(t)[N(y_{1+}) - M(h_{1+},h_{2+};\rho)] - e^{-r(T_1-t)}K_1[N(y_{1-}) - N(h_{1-})]$ for debt K_2.

Note that Geske assumes each time the firm pays off its maturing debt with new equity. As a result, the default condition becomes:

$$
\begin{aligned}
A_1 &< K_1 + D(T_1,T_2) \\
A_1 - D(T_1,T_2) &< K_1 \\
E_1^{(a)} &< K_1
\end{aligned}
\tag{14.7}
$$

which is, the equity value after paying off debt K_1 must be greater than K_1. In other words, the equity value after paying off debt includes the new equity that equals K_1. Hence, this value must be equal to at least K_1. If not, then it indicates that the firm cannot issue new equity, which implies obvious default. Table 14.1 demonstrates the balance sheet before and after paying off K_1.

<table>
<tr><th colspan="5">At time T_1</th></tr>
<tr><th colspan="2">Before K_1 is paid off</th><th></th><th colspan="2">After K_1 is paid off (with equity)</th></tr>
<tr><td>A_1 or $A(T_1)$</td><td>K_1
$D_{1,2}(A_1)$
$E_1^{(b)}$</td><td></td><td>A_1 or $A(T_1)$</td><td>
$D_{1,2}(A_1)$
$E_1^{(a)} = E_1^{(b)} + K_1$</td></tr>
</table>

Table 14.1: Balance Sheet at Time T_1

The default barrier can be calculated $\bar{A}_1 = K_1 + D_{1,2}(\bar{A}_1)$ as $D_{1,2}$ is a function of A_1. Note that $\bar{A}_2 = K_2$ as this is a two period model and the firm liquidates at T_2.

Suppose a company has two zero coupon debts, one and two years to maturity and each has \$100 face value. Also suppose the assets are currently worth \$400 and the debts are together worth \$170 on present value basis. This is graphically represented by the following balance sheet:[2]

Balance Sheet
as of year 0

assets	400	Maturity $t=1$ debt	90
		Maturity $t=2$ debt	80
		Equity	130
total	400	Total	400

note: both debts have face values of \$100

Table 14.2: Balance Sheet at Year 0

First, assume that one year later, the asset grows to \$450 and the firm faces the first debt payment of \$100. The firm at this time should raise equity to pay for the first debt so that the asset value will not have to be decreased. The asset value after paying off the first debt is still \$450. Assume that at this time ($t = 1$), the second debt, now only a year from maturity, has a value of \$90. As a result, the equity should be \$360 (= \$450 − \$90) that includes \$100 new equity and \$260 old equity. The balance sheet becomes:

Now, instead of the assets being worth \$450, suppose that the firm made some bad investment decisions and the asset's value drops to \$150. A bad economy and lower asset value imposes a higher default risk on the second debt so it is priced lower at \$75 due to its higher risk. Hence, the resulting equity value of old equity and of the should be raised equity, or debt due plus net equity, drops to \$75 (150 − \$75 = \$100 − \$25) . The firm, as in the previous case, would like to raise equity to pay off the first debt. But the new equity value needs to be \$100 to retire the debt due which creates a clear contradiction. This means that the new equity owner pays \$100 in cash but in return receives a portion of \$75. No rational investor would invest equity in this firm.

Since the firm cannot raise equity capital to continue its operation, it should not be considered a going concern. There is point where the potential new equity

[2]We assume the risk free rate to be about 10%. Since the company is extremely solvent, both debts are roughly priced at the risk free rate.

Balance Sheet as of year 1 before payment of first debt			
assets	450	Maturity $t = 1$ debt	100
		Maturity $t = 2$ debt	90
		Equity	260
total	450	Total	450

Balance Sheet at year 1 after payment of first debt			
assets	450	maturity $t = 2$ debt	90
		old equity	260
		new equity	100
total	450	Total	450

note: issue new equity to pay for the first debt

Table 14.3: Balance Sheet at Year 1

owner is indifferent and this is the going concern breakeven point for the company. Suppose the (break-even) asset value in one year is falls to \$186.01. At this asset value, the second debt is worth \$86. Consequently, the new equity owner has \$100 and the old equity has \$0.01. And we know that the default point is \$186.[3]

Table 14.4 shows the relationships between the market value of debt (two-year debt at year 1) and market value of equity in previous examples.

We can clearly see that any asset value lower than \$186 will cause default and should require other than a going concern opinion. However, with \$186 of assets, the company can pay the first debt due and continue to operate. One could also consider selling assets to pay off the first debt without raising any new equity. However, this approach to claim dilution would cause the second debt to drop significantly in value as in Table 14.5.

The reason is that the equity immediately has an option value at the cost of the remaining debt. In the above hypothetical tables, we assume \$10 is transferred from debt to equity. At $t = 0$, the debt holders know about this even when there is no information asymmetry. As a result, they will pay less for the debt.

Usually, the company will roll over old debt to new debt instead of issuing

[3]This value is precisely the "implied strike price" in the Geske model. We should notice that \$186.01 $\sim$ \$186 in this example. We leave a minor amount, \$0.01, to old equity holders in order to make this example more reasonable.

Balance Sheet
as of year 1 before payment of first debt

assets	186.01	one-year debt	100
		two-year debt	86
		Equity	0.01
total	186.01	Total	186.01

Balance Sheet
as of year 1 after payment of first debt

assets	186.01	two-year debt	86
		old equity	0.01
		new equity	100
total	186.01	Total	186.01

note: issue new equity to pay for the first debt

Table 14.4: Balance Sheet at Year 1

Balance Sheet
as of year 1 before payment of first debt

Assets	186	one-year debt	100
		two-year debt	86
		Equity	0
total	186	Total	186

Balance Sheet
as of year 1 after payment of first debt

assets	86	two-year debt	76
		old equity	10
total	86	Total	86

note: selling asset to pay for the first debt

Table 14.5: Balance Sheet at Year 1

equity. In the case of extreme solvency, this is not a problem. But in the case of near default, as described above, we have Table 14.6.

Balance Sheet
as of year 1 before payment of first debt

assets	186.01	one-year debt	100
		two-year debt	86
		Equity	0.01
total	186.01	Total	186.01

Balance Sheet
as of year 1 after payment of first debt

assets	186.01	two-year debt	86
		new debt	100
		old equity	0.01
total	186.01	Total	186.01

note: issue new debt to pay for the first debt

Table 14.6: Balance Sheet at Year 1

The principal of the new debt can be extremely high to reflect the very risky situation in order to get a \$100 to retire the first issue. Because the existing debt matures earlier (and hence has a higher seniority) its value should be the same whether there is new equity or debt. The equity will give a different claim whether new equity is raised or new debt is issued. With new equity the original equity will return a small portion after the second debt issue is repaid. With new debt, the original equity will get the entire return if the asset value increases after both debt issues are repaid. In the equilibrium, the original equity value should return to 0.01. It should not matter if the funds come from new equity or new debt at just over break-even point. Either way, the result holds and the old equity holders have a \$0.01 value.

Under the current measurement for going concern status, companies will usually receive a going concern opinion at \$186 and probably at \$150. The company continues to survive and operate. Now at \$150 value, the company is not able to raise capital, but it is certainly able to pay the debt with its assets and leave the second debt with \$50. Under this condition, the junior debt will be worth less than \$50, possibly very little since debt holders do not have the safe covenant to prevent managers/shareholders from selling assets to pay the senior debt. The transferring of wealth from debt owner to equity owner is what we define as the agency problem.

As long as the company spends assets to pay the earlier maturing debt, the later maturing debt holders will be hurt and shareholders will benefit.

We note that at the due date of the first debt, the company faces a decision whether to pay the debt obligation. This is a compound option question in that if the company decides to pay, the company continues to survive much like exercising the compound option to keep the option alive. The company's survival criterion relies upon whether the company can raise new equity capital. In this analysis, the technical condition of staying solvent (paying the coupon) is that the company must use new equity to pay for the coupon. If such new equity conceptually cannot be raised, then the company should go bankrupt. Interestingly, this condition translates into another equivalent condition that the market value of the assets of the company must stay above the market value of the liabilities at the moment of the coupon. This condition is regarded as the no-arbitrage condition and should be the breakeven point in value for receiving a clean going concern audit.

14.4 The Leland-Toft Model

The Leland-Toft model (which is an extension of the Leland model) is another structural model that extended the Black-Scholes-Merton model to include multiple debts. Different from the Geske model, the Leland-Toft model assumes continuous coupons and continuous issuance and redemption of the debts. As a result, the firm has a steady-state debt level that will not change, as the amount of debt issued exactly cancels the amount of debt redeemed.

Note that in the Leland-Toft model, we remove tax shield and dead-weight cost from the model to make it a fair comparison to other models. The model without tax shield and deadweight cost is given as:

$$D = \frac{C}{r} + \left(K - \frac{C}{r}\right)\left(\frac{1 - e^{-rT}}{rT} - I\right) + \left(H - \frac{C}{r}\right)J \tag{14.8}$$

where C is the continuous coupon, K is the face value of debt that expires in time T and:

$$\begin{aligned}
I &= \frac{1}{rT}\left(G - e^{-rT}F\right) \\
J &= \frac{1}{z\sigma_A\sqrt{T}}\left(-(b^{-a+z})N[q_-]q_- + b^{-a-z}N[q_+]q_+\right) \\
F &= N[h_-] + b^{-2a}N[h_+] \\
G &= b^{-a+z}N[q_-] + b^{-a-z}N[q_+]
\end{aligned} \tag{14.9}$$

$$
\begin{aligned}
q_{\pm} &= \frac{-b \pm z\sigma_A^2 T}{\sigma_A \sqrt{T}} \\
h_{\pm} &= \frac{-b \pm a\sigma_A^2 T}{\sigma_A \sqrt{T}} \\
a &= \frac{r - \delta - {}^1\!/_2\sigma_A^2}{\sigma_A^2} \\
b &= \ln\left(\frac{A}{H}\right) \\
z &= \frac{\sqrt{(a\sigma_A^2)^2 + 2r\sigma_A^2}}{\sigma_A^2}
\end{aligned}
\tag{14.10}
$$

$$
\begin{aligned}
H &= \frac{(C/r)[X/(rT) - Y] - XK/(rT)}{1 - Y} \\
X &= 2ae^{-rT} N[a\sigma_A\sqrt{T}] - 2zN[z\sigma_A\sqrt{T} - \frac{2}{\sigma_A\sqrt{T}} n[z\sigma_A\sqrt{T}] + \frac{2e^{-rT}}{\sigma_A\sqrt{T}} n[a\sigma_A\sqrt{T}] + (z - a) \\
Y &= -\left(2z + \frac{2}{z\sigma_A^2 T} N[z\sigma_A\sqrt{T}] - \frac{2}{\sigma_A\sqrt{T}} n[z\sigma_A\sqrt{T}] + (z - a) + \frac{1}{z\sigma_A^2 T}\right)
\end{aligned}
\tag{14.11}
$$

In the model, H is known as the default barrier. In the model, the firm constantly issues and retires the same amount of debt.

14.5 Hybrid Models

The Geske model is difficult to use in reality due to its computational difficulties. Hence, a number of hybrid models emerged to bridge the gap between the true structural model such as Geskes (where the meaning of default is failure to pay) and the reduced form models (that focus on calibration ease). These hybrid models assume a default barrier that is exogenously specified. The default barrier is then calibrated to the market.

A number of hybrid models emerged to bridge the gap between the true structural model such as Geskes (where the meaning of default is failure to pay) and the reduced form models (that focus on calibration ease). These hybrid models assume a default barrier that is exogenously specified. The default barrier is then calibrated to the market.

Once there is an external barrier, the modeling complexity drops substantially.

We can the use the First Passage Time mathematics. There are several barrier option models that are popular. The first one is the flat barrier option model introduced first by Rubinstein and Reiner (1991).

A barrier option is an option whose payoff is activated (knocked in) and terminated (knocked out) when a barrier is reached by the stock price before maturity. Due the barrier can be higher or lower than the current stock price, barrier options are classified as down-and-in and down-and-out barrier options whose barrier lies below the current stock price; and up-and-in, and up-and-out barrier options whose barrier lies above the current stock price.

Down-and-in and up-and-in barrier options become call or put options when the barrier is reached. Similarly, down-and-out and up-and-out barrier options become worthless when the barrier is reached.

The barrier option problem is the first passage time problem. Assume the standard Black-Scholes model that the stock price follows the log normal process. Let the barrier be H and the strike be K. If the barrier is below the current stock price, i.e. $S_t > H$, then the risk neutral probability that the stock price stays above both H and K is either:

$$N\left(\frac{\ln S_t - \ln K + (r - 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right) - \left(\frac{S_t}{H}\right)^{1-\frac{2r}{\sigma^2}} N\left(\frac{2\ln H - \ln S_t - \ln K + (r - 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right) \tag{14.12}$$

if $H < K$ or

$$N\left(\frac{\ln S_t - \ln H + (r - 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right) - \left(\frac{S_t}{H}\right)^{1-\frac{2r}{\sigma^2}} N\left(\frac{\ln H - \ln S_t + (r - 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right) \tag{14.13}$$

if $H > K$. Note that the first part of equation (13-102) is just the probability of $S_T > K$. The second part of equation (13-102) is the joint probability that $S_T > K$ and $S_\tau < H$. As a result, the difference is the joint probability that $S_T > K$ and $S_\tau > H$. When $H > K$, then staying above the barrier automatically guarantees in the money for the option and we simply replace K by H.

As a result, the down-and-out barrier call option model (i.e. $S_t > H$) can be derived as:

$$
\begin{aligned}
C_{\mathrm{do}} &= e^{-r(T-t)}\hat{E}_t[(S_T - K)1_{S_T>K\cap S_\tau>H}] \\
&= e^{-r(T-t)}\hat{E}_t[S_T 1_{S_T>K\cap S_\tau>H}] - e^{-r(T-t)}K\hat{E}_t[1_{S_T>K\cap S_\tau>H}] \\
&= S_t\Pi_+ - e^{-r(T-t)}K\Pi_-
\end{aligned}
\tag{14.14}
$$

where

$$
\Pi_\pm = N\left(\frac{\ln S - \ln K + (r \pm 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right) - \left(\frac{S}{H}\right)^{1-\frac{2r}{\sigma^2}} N\left(\frac{2\ln H - \ln V - \ln K + (r \pm 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right)
\tag{14.15}
$$

if $H < K$ or

$$
\Pi_\pm = N\left(\frac{\ln S_t - \ln H + (r \pm 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right) - \left(\frac{S_t}{H}\right)^{1-\frac{2r}{\sigma^2}} N\left(\frac{\ln H - \ln S_t + (r \pm 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right)
\tag{14.16}
$$

if $H > K$. Note that the two results for Π_- are just probability of first passage time. The results for Π_+ can be obtained via the change of measure.

The down-and-in barrier call option model can be derived by subtracting the down-and-out option model from the Black-Scholes model: $C_{\mathrm{di}} = C_{\mathrm{BS}} - C_{\mathrm{do}}$.

If the barrier is above the current stock price, i.e. $S_t < H$, then the risk neutral probability that the stock price stays above both H and K is either:

$$
\begin{aligned}
&N\left(\frac{\ln S_t - \ln K + (r - 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right) \\
&- \left(\frac{S_t}{H}\right)^{1-\frac{2r}{\sigma^2}} \left\{N\left(\frac{2\ln H - \ln S_t - \ln K + (r - 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right) - N\left(\frac{\ln S_t - \ln K + (r - 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right)\right.
\end{aligned}
\tag{14.17}
$$

if $H < K$ or

$$
\begin{aligned}
&N\left(\frac{\ln S_t - \ln H + (r - 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right) \\
&- \left(\frac{S_t}{H}\right)^{1-\frac{2r}{\sigma^2}} \left\{N\left(\frac{\ln H - \ln S_t + (r - 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right) - N\left(\frac{\ln S_t - \ln H + (r - 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right)\right\}
\end{aligned}
\tag{14.18}
$$

if $H > K$. Note that the first part of equation (13-105) is just the probability of $S_T > K$. The second part of equation (13-105) is the joint probability that $S_T > K$ and $S_\tau < H$. As a result, the difference is the joint probability that $S_T > K$ and $S_\tau > H$. When $H > K$, then staying above the barrier automatically guarantees in the money for the option and we simply replace K by H.

The second one is by Black and Cox (1976) where the barrier is not flat but exponentially decayed. The first passage time (survival probability) by Black and Cox is given as follows:

$$N\left(\frac{\ln V - \ln Ke^{-\gamma(T-\tau)} + (r - a - {}^1\!/_2\sigma^2)(\tau - t)}{\sqrt{\sigma^2(T-t)}}\right) - \left(\frac{V}{Ke^{-\gamma(T-t)}}\right)^{1-\frac{2(r-a-\gamma)}{\sigma^2}} \times N\left(\frac{2\ln Ke^{-\gamma(T-t)} - \ln V - \ln Ke^{-\gamma(T-\tau)} + (r - a - {}^1\!/_2\sigma^2)(\tau - t)}{\sqrt{\sigma^2(T-t)}}\right) \tag{14.19}$$

where K is the barrier amount, γ is the decay rate of the barrier (in other words, the barrier is smaller currently than close to maturity), τ is the default time (i.e., the first passage time to the boundary), and V is the asset value. To put it plainly, if t^* is the first passage time to the boundary, then the probability that $t^* \geq \tau$ is given by the above equation.

The last one is the CreditGrades model (2002) that was developed by Goldman Sachs, JP Morgan, Deutsche Bank, and RiskMetrics. This is essentially a random barrier model. The technical document can be downloaded from creditgrades.com. The main equation is its approximation formula for the survival probability, as follows:

$$Q(t) = N\left(-\frac{A}{2} + \frac{\ln d}{A}\right) - dN\left(-\frac{A}{2} - \frac{\ln d}{A}\right) \tag{14.20}$$

where

$$A = \sqrt{\sigma_S^2 \xi^2 t + \lambda^2}$$
$$d = \frac{1}{1-\xi} e^{\lambda^2}$$
$$\xi = \frac{E}{E + D}$$

and λ is the volatility parameter for the stochastic barrier, E is the equity value and D is the "target" debt value. Hence, it is logical to view ξ as the leverage ratio. With the survival probability function, we can proceed with CDS valuation.

Chapter 15

Credit Portfolio and Credit Correlation

15.1 Introduction

So far we have been assessing single-name credit risk and how to model it. The key risk management factor – correlation has not been discussed. As in market risk management, correlation is the key driving factor to lower the risk. Little or even negative correlation among assets can dramatically reduce the risk of a portfolio. However, correlation in credit is much complex than it is in market risk. In market risk, especially in parametric VaR, correlation is used among normally distributed asset returns, and as a result a simple covariance matrix can describe very well how diversification works and then VaR is easily inferred.

The reason why correlation in market risk so easily used is of course a result of normally distributed asset returns. In market risk, although there are various criticisms on normally distributed asset returns, it is not a bad first order approximation to regard asset returns as normal. Unfortunately, such an approximation is totally unsuitable in credit risk.

First of all, credit losses are only negative. That is, there is no gain in defaults – everyone suffers. Hence, symmetry in normality is out of question. Secondly, both PD (probability of default) and LGD (loss given default) are both random and inter-related (see Introduction). These two variable add extra complexity to modeling the credit risk, unlike in market risk there is only one variable – returns of assets.

Hence, here we first introduce the concept of credit correlation. We start with default correlation. We use a simple and yet powerful credit derivative contract – credit default basket to introduce the concept credit correlation. Briefly, we demonstrate, via default basket, how correlation in credit risk can be misleading. Then, via the example, we show that why industry moves to copula to replace credit correlation. Finally, a full example of CDO (collateralized debt obligation) is introduced to utilize copula in its entirety.

15.2 Basics

The following Venn diagram helps explain the basic setup of portfolio credit risk valuation.

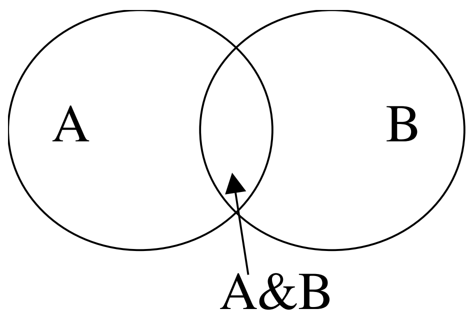


Figure 15.1: Venn Diagram for Default

where A and B represent the two default events for company A and company B respectively. The area in the middle represents joint default.

In order to properly correlate default events, maintain the flexibility of freely specifying individual default probabilities, and be able to uniquely define a joint distribution, we use conditional default probabilities to describe the dependency of two default events (rather than specifying the correlation). In other words, we specify the default probability of the second bond given that the first bond has already defaulted.

For a single period, the default event should follow a Bernoulli distribution where 0 represents no default and 1 represents default. Under the Bernoulli distribution, the specification of the conditional probability (given the marginal probabilities) uniquely defines the joint default probability. The joint default probability can be formally delineated as:

$$p(A \cap B) = \begin{cases} p(A|B)p(B) & \text{or} \\ p(B|A)p(A) \end{cases} \tag{15.1}$$

where A and B represent default events, i.e. a short hand notation for A = 1 and B = 1.

This specification of the joint distribution is desirable in that it is realistic in defining the default relationship between two bonds. If two bonds have different default probabilities (marginals) but they are extremely highly correlated, it means (in our model) that if the less risky one defaults, the riskier one will surely default but not the other way around. Imagine a small auto part supplier whose sole client is a large automobile manufacturer. If the manufacturer defaults, it is highly likely that the small part supplier will eventually default as well. The small part supplier may also default on its own, independent of what happens to the automobile manufacturer.

For example, we let A be default on the part of the manufacturer and B be the default on the part of the small supplier. Each has 10% and 20% default probability for the period, respectively. The part supplier is completely dependent upon the manufacturer, i.e. the conditional default probability of the supplier on the manufacturer is one. Then the joint probability of both firms defaulting is:

$$\begin{aligned} p(A \cap B) &= p(B|A)p(A) \\ &= p(A) \\ &= 10\% \end{aligned} \tag{15.2}$$

Hence, we obtain the following joint distribution for the two companies:

Bivariate Bernulli Distribution			
	0	1	A
0	80%	0%	80%
1	10%	10%	20%
B	90%	10%	100%

The survival of the small supplier depends completely on the large auto manufacturer:

$$p(B|A) = \frac{p(A \cap B)}{p(A)} = 100\% \tag{15.3}$$

The dependency of the manufacturer on the supplier can be calculated to be:

$$p(A|B) = \frac{p(A \cap B)}{p(B)} = \frac{10\%}{20\%} = 50\% \tag{15.4}$$

which is equal to the ratio of the two individual default probabilities. The unconditional correlation of the two companies is 0.6667. The two companies can also have perfectly negative dependency: default of one company implies the survival of the other. In this case, the conditional probability is:

$$\begin{gathered} p(B^C|A) = 1 \\ \text{or} \\ p(B|A) = 0 \end{gathered} \tag{15.5}$$

Using the numbers in the previous example yields the following joint distribution:

Bivariate Bernulli Distribution			
	0	1	A
0	70%	10%	80%
1	20%	0%	20%
B	90%	10%	100%

Notice that, even though there is perfect negative dependency, the unconditional correlation is only -0.1667.

The examples here show that perfect dependency does not translate to perfect correlation. This is because dependency is directional but correlation is not. In the first case, A defaults causes B defaults but not vice versa. Hence the dependency of B on A is 100% but A on B is only 50%. The correlation is in between, 67%. In the second case, A can survive only if B defaults and vice versa. This is perfectly negative dependency but the correlation is only -17%. The unconditional default correlation is calculated as follows:

$$\begin{aligned} \rho(A,B) &\overset{\text{either}}{=} \frac{p(B|A)p(A) - p(A)p(B)}{\sqrt{p(A)(1-p(A))p(B)(1-p(B))}} \\ &\overset{\text{or}}{=} \frac{p(A|B)p(B) - p(A)p(B)}{\sqrt{p(A)(1-p(A))p(B)(1-p(B))}} \end{aligned} \tag{15.6}$$

So it is seen that it is not possible to reach 100% or -100% correlation unless the following conditions are satisfied:

- $p(A) = p(B)$ for 100% correlation
- $p(A) + p(B) = 1$ for -100% correlation

When there is perfect dependency (i.e. $p(B|A) = p(A|B) = 1$), for $\rho(A, B) = 1$, it must be true that:

$$\begin{aligned} 1 &= \frac{p(A)(1 - p(B))}{\sqrt{p(A)(1 - p(A))p(B)(1 - p(B))}} \\ &= \frac{\sqrt{p(A)(1 - p(B))}}{\sqrt{(1 - p(A))p(B)}} \end{aligned} \tag{15.7}$$

in which the only solution is $p(A) = p(B)$. When there is perfect negative dependency (i.e. $p(B^C|A) = 1$, or equivalently $p(B|A) = 0$, and vice versa), for $\rho(A, B) = -1$, it must be true that:

$$\begin{aligned} -1 &= \frac{-p(A)p(B)}{\sqrt{p(A)(1 - p(A))p(B)(1 - p(B))}} \\ 1 &= \frac{\sqrt{p(A)p(B)}}{\sqrt{(1 - p(A))(1 - p(B))}} \end{aligned} \tag{15.8}$$

in which the only solution is $p(A) + p(B) = 1$.

15.3 Default Baskets (First to default)

Default baskets usually contain only a few credit names (individual bonds). The first-to-default basket pays principal and accrued interest minus the recovery value of the first defaulted bond in the basket.

For the sake of easy exposition, we examine a two-asset basket with no counterparty risk. The model can easily be extended to a multi-asset basket, but the value must be solved numerically. Following the previous discussion, the probability of the first-to-default in a two-asset case is $p(A \cup B)$. To find the value of such a probability, we first need to find values of CDS spreads for A and B. Given that this is a single period model, we can use 13.19 with 40% recovery rate. Then, we have the CDS spread for A as $20\% \times (1 - 40\%) = 12\%$ (or 1200 basis points) and for B as $10\% \times (1 - 40\%) = 6\%$ (or 600 basis points). As a result, we can compute the spread for the first-to-default (FTD) as:

$$\begin{aligned} &p(A \cup B) = p(A) + p(B) - p(A \cap B) = 10\% + 20\% - 10\% = 20\% \\ &s(A \cup B) = p(A \cup B) \times (1 - 40\%) = 12\% \end{aligned} \tag{15.9}$$

We can also calculate the second-to-default (STD) as follows:

$$
\begin{aligned}
&p(A)+p(B)-p(A\cup B)=p(A\cap B)=10\% \\
&s(A\cap B)=6\%
\end{aligned}
\tag{15.10}
$$

It is easy to demonstrate that the basket value is a (negatively) linear function of the default correlation. When the default correlation is small (or even negative), the issuers tend to default alternately: this increases the basket risk. When the correlation is large, the issuers tend to default together: this decreases the basket risk. As will be demonstrated below, the value of the default basket will vary between that of a single default swap (when there is perfect dependency) and the sum of the two individual default swaps (when there is zero dependency).

Note that in the event of perfect dependency, i.e. $p(B|A)=1$, then $p(A\cap B)=p(B|A)p(A)=p(A)$ and then $p(A\cup B)=p(B)$. As a natural result, the spread for the FTD equals the spread for B (same can be said if $p(A|B)=1$), which implies that the value of the basket is equal to the value of a single name default swap.

On the other hand, if the dependency is perfectly negative, i.e. $p(B^C|A)=1$ or $p(B|A)=0$, then $p(A\cup B)=p(A)+p(B)$. Then the spread of the FTD equals the sum of the spread for A and the spread for B, which implies that the value of the basket is equal to the sum of the two individual swaps. (Note that in this case the value of the STD is 0, as one of the two must survive.)

Note that if there is no dependency and the two companies were to have default probabilities that sum to exactly 100%, then the joint distribution would degenerate and result in an unconditional correlation of -1.

The results can be extended to multiple assets, though the calculations of the probabilities become multi-dimensional. We can use Monte-Carlo methods to calculate the joint normal probabilities in high dimensions. We can write down the valuation equation as follows:

$$
V=\hat{\mathbb{E}}_t\left[P(t,\min\{u_j\})\mathbb{I}_{\min\{u_j\}<T}(1-R_j(u_j))N_j\right] \tag{15.11}
$$

where $\mathbb{I}$ is the indicator function, u_j is the default time of the j-th bond, R_j is recovery rate of the j-th bond, and N_j is the notional amount of the j-th bond. The basket pays when it experiences the first default, i.e. $\min\{u_j\}$.

Obviously, the above equation has no easy solution when the default events (or default times, u_j) are correlated. Hence, in the following, we only solve the model under indepedent defaults. Under independence, we can write (15.11) as:

$$
V=\int_t^T \sum\nolimits_{i=1}^n P(t,u)\left[-d\Pi_{j=1}^i Q_j(t,u)+d\Pi_{j=0}^{i-1}Q_j(t,u)\right](1-R_i(u)) \tag{15.12}
$$

where $Q_0(t,u) = 1$ and hence $dQ_0(t,u) = 0$. The above formula assumes that the last bond (i.e. bond n) has the highest priority in compensation, i.e. if the last bond jointly defaults with any other bond, the payoff is determined by the last bond. The second to last bond has the next highest priority in a sense that if it jointly defaults with any other bond but the last, the payoff is determined by the second to last bond. This priority prevails recursively to the first bond in the basket.

15.4 Copula and CDO Pricing

15.4.1 Background

A Collateral Debt Obligation, or CDO, is a set of securities known as tranches that are backed by a pool (portfolio) of default-risky fixed income securities (e.g. corporate bonds, loans, default swaps, and asset backed securities). These obligations, or tranches, are then sold to investors. Commonly it is set up as a SPV (special purpose vehicle) as depicted in Figure 15.2.

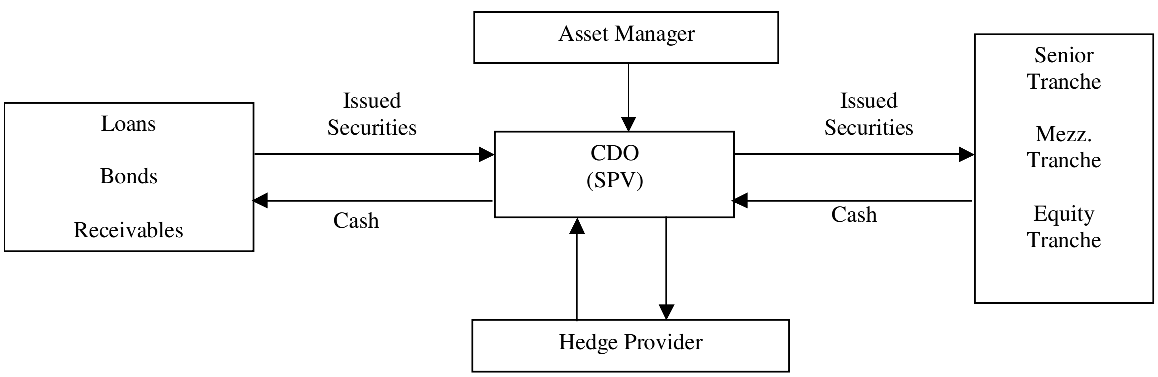


Figure 15.2: CDO Structure

These SPVs are "paper companies" set up by various financial institutions (most likely investment banks) and the assets are also acquired from these financial institutions. During the crisis, it had become apparent that many banks dumped their bad assets to these SPVs and then issued very opaque tranches so that investors would not find out how bad the quality of the assets is in the pool.

While frauds had been committed and lawsuits had been filed (e.g. JP Morgans \$13 billion settlement with the U.S. Justice Department), efforts have been made to resume the securitization market and try to make the market more healthy and transparent.

15.4.2 Basics

A CDO is a collection of bonds (or loans and in that case it is called CLO) that are resold to the secondary market in tranches. The actual tranche structure is highly flexible and takes no definite form but in general, it has a waterfall structure where tranches are paid off in a sequential order. This structure is known as the n-th loss tranche structure. For example, a CDX CDO is a CDO with 125 credit default swaps with US$ 10 million notional. The CDO is sliced up into the following tranches: 0-3%, 3-7%, 7-10%, 10-15%, and 15-30%. Losses over 30% are extremely unlikely (due to collaterals) and hence are not analyzed. The structure is the standard waterfall structure that the first tranche (equity tranche) takes upon loss of default up to 3%, or $300,000, Then the following mezzanine tranche will take over the next 4% loss (from 3 to 7%), or $400,000, and so on.

As we see above, the basic structure of a CDO is similar to that of a CMO (Collateral Mortgage Obligations) in that the waterfall defines the sequence of the payoff to each tranche. The difference lies in tranche payoffs. A typical CMO uses all its revenues (mortgage payments) to pay its senior tranche. After the senior tranche is completely paid, the mezzanine tranche can take over. A CDO structure is to pay every tranche if there is enough revenue (coupons from bonds). Any reduction in revenue due to defaults is taken from the junior tranche, and then the mezzanine tranche, and then the senior tranche.

The primary pricing objective of pricing CDOs is to compute the spreads of the tranches. However, an equally important objective is to study the loss distributions of the tranches. As it will be clear later, even though spreads can be computed for the investors to trade, risk managing these tranches require much more than just the spreads. Due to the extremely exotic loss distributions, without the whole distribution, it is nearly impossible to do a good job in risk management.

A typical number of a CDOs constituents is in hundreds. And all constituents are highly correlated. As a result, to understand the loss distribution of a CDO is difficult. Monte Carlo simulation is the usual technique to accomplish such a goal. However, for 125 bonds in a CDO, there will be $\frac{125\times124}{2} = 7750$ correlations to estimate and simulations themselves take a large amount of time. As a result, a fast approximation method is developed. The use of FI quickly reduces the multi-dimensional problem to uni-dimensional. However, some simplifications must be made. While details are to be seen later, the major intuition is as follows. First, we must assume a factor model (single factor to begin with) through which all bonds are correlated. This limits the flexibility of how assets are correlated and yet it provides consistency (of correlation estimates) over time and a parametric structure that makes FI possible. Secondly, we must assume a barrier structural model that explain how defaults occur.

Tranches are regarded as a series of unsecured debts from senior to junior. The equity tranche of a CDO is like the regular equity of a company, which is a call option. The problem here is to model the underlying asset which is portfolio loss.

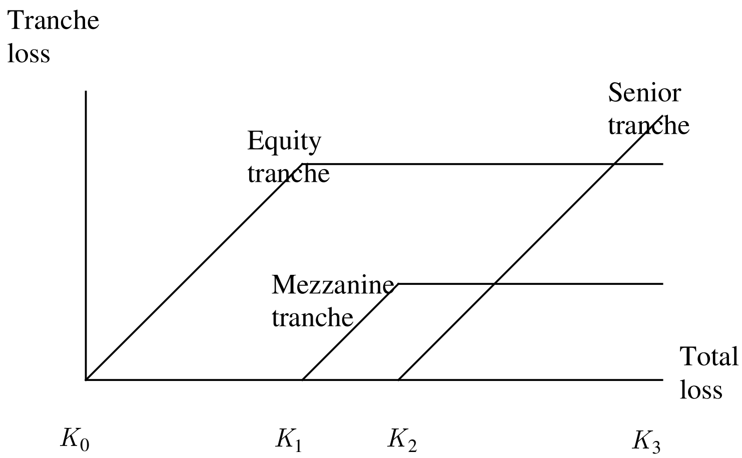


Figure 15.3: CDO Waterfall

In Figure 15.3, $K_0 = 0$ and $K_j - K_{j-1}$ is the size of the jth tranche. For example, K_1 (or $K_1 - K_0$) is the notional of the equity tranche, $K_2 - K_1$ is the notional of the mezzanine tranche, and $K_3 - K_2$ is the notional of the senior tranche.

Note that the total loss of the pool cannot exceed the total size of the pool as long as there is recovery. In a simple case where all bonds are equal size (e.g. $ 100,000 notional for each bond) and equal recovery (e.g. 40%), a pool of 100 bonds has a total size of $10 million and the total loss will not exceed $4 million. One can size up the pool to three tranches as given in the above diagram to be $2 million equity, $5 million mezzanine, and $3 million senior tranches. Identically, this implies $K_1 = 2$ million, $K_2 = 7$ million, and $K_3 = 10$ million. Note that the maximum loss in this example is $6 million (recovery rate is 40% for all bonds) and therefore the senior tranche will never suffer any loss. As a result, the senior tranche is risk-free.

A typical contract of a CDO tranche is a swap. That is, the protection buyer pays a series of spreads and in return receives payments identical to default losses. In other words, a CDO tranche is itself a CDS. However, different from a simple corporate CDS where default can happen only once and the contract stops, a CDO pool will have multiple defaults (up to the total number of bonds in the pool, which is 100 in the above example). Upon each default, a payment is made to a proper tranche investor. According to the loss payoff diagram above, for the first few defaults experienced by the pool, the loss amounts are paid to equity investors up to K_1 ($2 million in the example). Then losses of the next defaults are paid to mezzanine investors up to K_2. The process continues until the either all bonds in the pool default or the CDO contract expires, whichever earlier.

As a result, as long as we can model the total loss, we can easily compute the cumulative tranche loss and in turn the value of the protection leg and the value of the premium leg of each tranche. The valuation of a CDO (i.e. its tranches) is a very technical matter and hence beyond the scope of this book. Interested readers can refer to, for example, Chen's Mathematical Finance for full details.

15.4.3 Factor Copula

In the standard factor copula model used now as the industry standard, we write the following:

$$x_i = \sqrt{\rho}\hat{W}_M + \sqrt{1-\rho}\hat{W}_i \tag{15.13}$$

where x_i is the factor that computes the default probability for firm i and ρ is the base correlation. Both W_M and W_i are normally distributed with 0 mean. Hence x_i is normally distributed. For any given period (say a year), we can write the conditional default probability as follows:

$$\begin{aligned}
\hat{p}_{i|f} &= \widehat{\Pr}\left(x_i < K_i | W_M = f\right) = \widehat{\Pr}\left(\sqrt{\rho}f + \sqrt{1-\rho}W_i < K_i\right) \\
&= \widehat{\Pr}\left(W_i < \tfrac{K_i - \sqrt{\rho}f}{\sqrt{1-\rho}}\right) \\
&= N\left(\tfrac{K_i - \sqrt{\rho}f}{\sqrt{1-\rho}}\right) \\
&= N\left(\tfrac{N^{-1}(\hat{p}_i) - \sqrt{\rho}f}{\sqrt{1-\rho}}\right)
\end{aligned} \tag{15.14}$$

where $\widehat{\Pr}(x_i < K_i) = N(K_i) = \hat{p}_i$ is the unconditional default probability. Naturally,

$$\int_{\Omega(f)} \hat{p}_{i|f}\phi(f)df = \hat{p}_i \tag{15.15}$$

where $\phi(f)$ is the standard normal distribution for the common factor, f. Conditional on the market factor f, firms default independently. Hence, under the conditional probability $\hat{p}_{i|f}$, defaults are independent. There are several methods to model conditional loss distribution. Here we introduce the Vasicek method, the Fourier inversion method, and the recursive algorithm.

The Vasicek method is a special case of the Fourier inversion method and the recursive algorithm. It assumes that the bond notionals and recoveries in the CDO pool are the same. The Fourier inversion method and the recursive algorithm have relative advantages and disadvantages. Recursive algorithm is faster but it works best when the recoveries of the bonds are the same. Fourier inversion method is more general but it is generally slower and it gives continuous loss distribution where the true loss distribution is discrete.

15.4.4 The Vasicek Model

If the exposures (notionals) are identical, then the percentage loss can be found by carrying out the following equation,

$$
\begin{aligned}
\Pr(\ L = i/m\ ) &= \binom{m}{i} \Pr(A_1 < K_1, \cdots, A_i < K_i, A_{i+1} > K_{i+1}, \cdots, A_m > K_m) \\
&= \binom{m}{i} \int_{-\infty}^{\infty} \Pr(A_1 < K_1, \cdots, A_i < K_i, A_{i+1} > K_{i+1}, \cdots, A_m \\
&\quad > K_m | W_M = f) dF(W_M < f)
\end{aligned}
\tag{15.16}
$$

Independence conditional on $W_M = f$ gives:

$$
\begin{aligned}
\Pr(L = i/m) &= \binom{m}{i} \int_{-\infty}^{\infty} \Pr(A_1 < K_1 | W_M = f) \cdots \Pr(A_i < K_i | W_M = f) \\
&\quad \Pr(A_{i+1} > K_{i+1} | W_M = f) \cdots \Pr(A_n > K_n | W_M = f) dF(W_M < f) \\
&= \binom{m}{i} \int_{-\infty}^{\infty} \prod_{j=1}^{i} N\left(\tfrac{N^{-1}(p_j) - \sqrt{\rho} f}{\sqrt{1-\rho}}\right) \prod_{j=i+1}^{m} N\left(-\tfrac{N^{-1}(p_j) - \sqrt{\rho} f}{\sqrt{1-\rho}}\right) \\
&\quad dF(W_M < f)
\end{aligned}
\tag{15.17}
$$

Note that "m choose i" or $(m \quad i)'$, is only symbolic and not a real combination function. Vasicek further simplifies by setting all probabilities equal, i.e. $p_j = p$:

$$
\Pr(L = i/m) = \binom{m}{i} \int_{-\infty}^{\infty} N\left(\tfrac{N^{-1}(p) - \sqrt{\rho} f}{\sqrt{1-\rho}}\right)^{i} N\left(-\tfrac{N^{-1}(p) - \sqrt{\rho} f}{\sqrt{1-\rho}}\right)^{m-i} dF(W_M < f)
\tag{15.18}
$$

Here $(m \quad i)'$ is the combination function. The integral can be implemented either via Riemman sum or Monte Carlo simulations. Cumulative loss is:

$$\Pr(L \leqslant i/m) = \sum\nolimits_{\ell=0}^{i} \Pr(L = \ell/m) \tag{15.19}$$

This can be very computationally expensive if probabilities are not equal, and still quite expensive if probabilities are equal (sum over m-choose-i terms). Vasicek (Limiting Loan Loss Probability Distribution) proposes an asymptotic formula and Schonbucher presents the FFT method. The FFT method is more plausible because it is computationally possible and less limitation than the Vasicek method.

15.4.5 Fourier Inversion and Recursive Algorithm

To evaluate CDOs correctly, two alternative (competing) methods are commonly used – Fourier Inversion and Recursive Algorithm. These methods can accurately calculate tranche values of a CDO. The Fourier Inversion method is more flexible and the Recursive Algorithm method is faster. For the details of the these methods and examples, see for example Chen's Financial Mathematics.

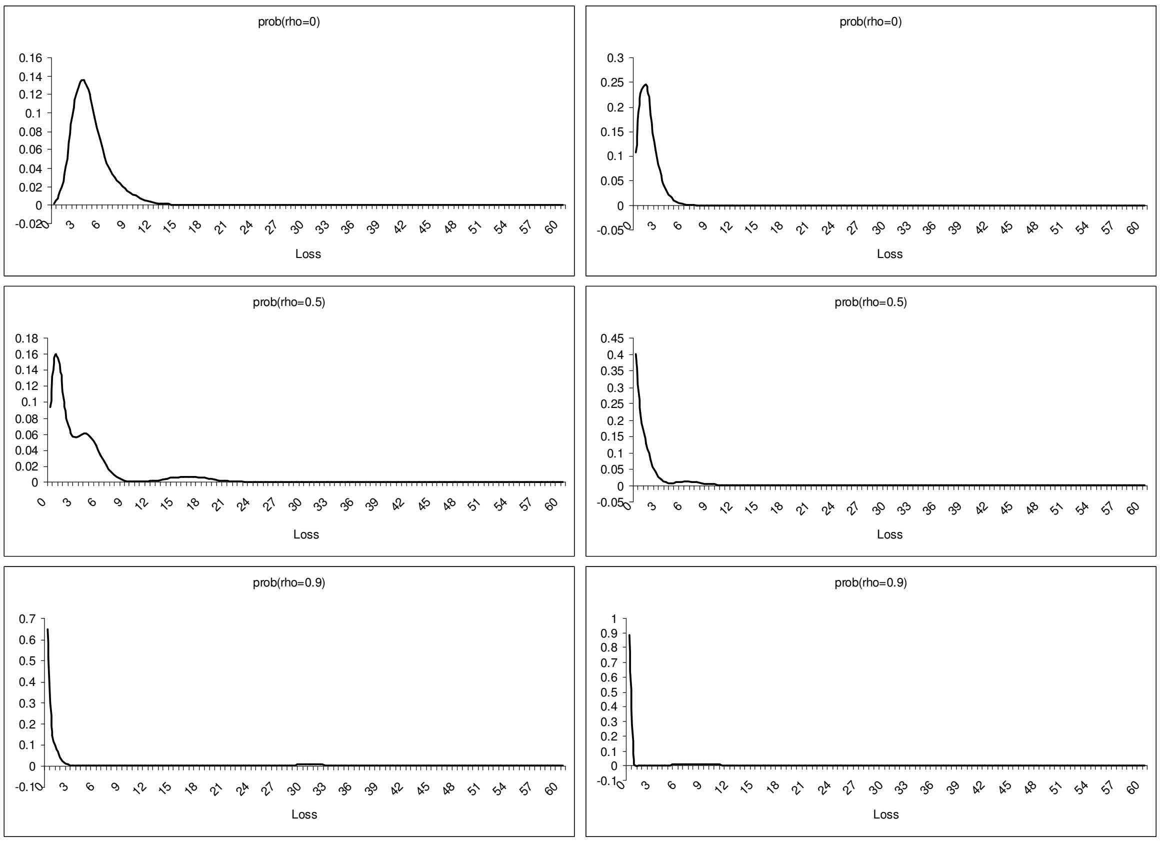
15.4.6 An Example

In the following, we show an example where there are 100 bonds in the pool. This probability is exaggerated so to magnify the behavior of the loss distribution. The recovery rate is 40% for all bonds. Hence, the maximum loss of the pool is 60% of the total pool size.

The shape of the loss distribution is very sensitive to the correlation. In the following several diagrams, we shall see how the loss distribution changes as the correlation increases. On the left is plotted the loss distribution for various correlation numbers when the default probability is 6%. On the right is plotted the same correlation scenarios when the default probability is 2%.

As we can see obviously, the lower is the default probability; the more right-skewed is the loss distribution, which implies the safer is the pool. Also we see that distribution is multi-modal when the correlation is high. The multi-modal phenomenon is more pronounced when the default probability is high. This can be easily seen from the two graphs when the correlation is 0.5. Figure 15.4 depicts the loss functions under various assumptions.

Once we obtain the loss distribution, we can then proceed to price various CDO tranches. Figure 15.5 provides the pricing results under 2% and 6% (for



6% default probability

2% default probability

Figure 15.4: Fourier Inversion Results

better visual) default probability of each name in the CDO with various correlation levels. As we can see the equity tranche is negatively correlated with correlation. That is, the more clustering the defaults, the more risk associated with the equity tranche. This is because high correlation bring the risks of various tranches closer. In an extreme case where one defaults, all default (perfect correlation), then senior tranches suffer losses the same time as junior tranches. On the other hand, when defaults are independent, the equity tranche is wiped out first and hence the risk is high.

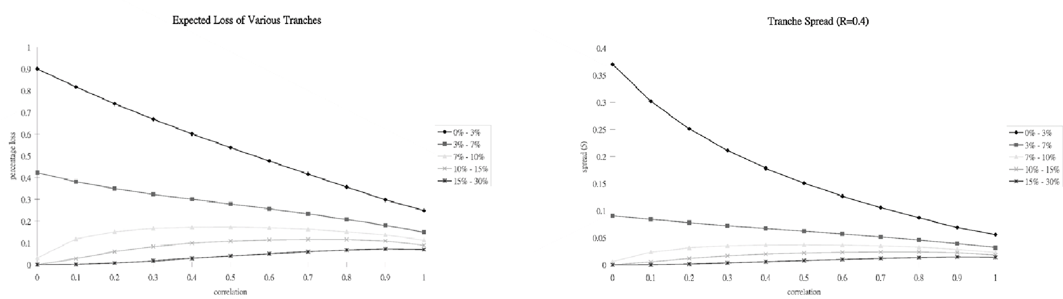


Figure 15.5: Tranche Prices

The expected losses are then translated to spreads because often these tranches are sold in a form of swap. In other words, there is no cash payment at inception. Over time investors receive fixed spreads and yet in the event of default the investor must replenish the loss.

One significant drawback of the copula model is that it is a single period model. It successfully incorporates multiple bonds and generates the loss distribution but it is not able to take into account defaults and losses over time. For this, we need to combine the Merton model with the copula model. Also, in the factor copula model, is the correlation everybody has with the market. However, this is not necessarily the case.

15.5 Monte Carlo Simulations

15.5.1 Default Basket

Default basket contracts (i.e. FTD (first to default) to NTD (nth to default)) have no closed-form solutions. Hence, Monte Carlo simulations are the only method to evaluate these deals. We want to simulate correlated defaults. Steps are as follows:

1. Get CDS spreads (for at least 4 names). Compute PDs (1y) using the back-of-envelope formula (i.e. one period model, equation 13.19 on page 175). Compute survival probabilities (1y), that is Q, for each of the four companies.

2. Compute lambdas from Q by $\lambda = -\ln Q$ (equation 13.1 on page 163).

3. Get a random number, u (i.e. $u = \texttt{RAND}()$), for each of the companies.

4. Compute z (normal random number) by z by inversing the uniform random variable u (i.e. $z = \texttt{NORMSINV}(u)$) for each of the companies.

5. Correlate z's by Cholesky matrix (and get X's) by doing the following calculation (correlation numbers below can be arbitrary):

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} \sqrt{1 - x_{12}^2 - x_{13}^2 - x_{14}^2} & \frac{\rho_{12} - \rho_{14}\rho_{24} - x_{13}x_{23}}{\sqrt{1 - x_{23}^2 - x_{24}^2}} & \frac{\rho_{13} - \rho_{14}\rho_{34}}{\sqrt{1 - \rho_{34}^2}} & \rho_{14} \\ 0 & \sqrt{1 - x_{23}^2 - x_{24}^2} & \frac{\rho_{23} - \rho_{24}\rho_{34}}{\sqrt{1 - \rho_{34}^2}} & \rho_{24} \\ 0 & 0 & \sqrt{1 - \rho_{34}^2} & \rho_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix}$$

6. Compute u^* by $N(x)$ (i.e. $u^* = \texttt{NORMSDIST}(x)$)

7. Compute τ with the formula $\tau = -\ln u^*/\lambda$ for each of the companies. Now these τ's (default time) are correlated.

8. Compute default (0 or 1) by comparing τ with 1

$$I_i = \begin{cases} 0 & \tau_i > 1 \\ 1 & \tau_i < 1 \end{cases}$$

9. FTD contracts pays if there is at least one default, that is: $q_1 = \sum_{i=1}^{4} I_i \geqslant 1$. 2TD are those $q_2 = \sum_{i=1}^{4} I_i \geqslant 2 - q_1$.

10. Repeat 3∼9 N times.

11. Count the number of those that $\Sigma I_i \geqslant 1$ and then divide it by N.

15.5.2 CDO

In the simulation, we simulate a market factor plus idiosyncratic factors, all normals. The simulation for each reference entity is a bivariate normal (or multi-variate normal for multiple factors) and we run this normal for each reference entity through the normal probability function to transform the simulations for each normally distributed x_i into 0-1 random variables, which will be correlated. Then the 0-1 random variables are applied to the survival probability functions to determine default times.

Chapter 16

Risk Management for Credit Risk

16.1 Introduction

While there has been more a long history in VaR, the risk management technology in credit risk management is very limited. People often are confused between market credit risk and default credit risk. Market credit risk refers to spread changes over time. Spreads reflect how default probabilities and expected recoveries are priced. Hence changes in spreads reflect how default probabilities and expected recoveries change over time. As default probabilities and expected recoveries are expectations, changes in spreads hence have nothing to do with actual defaults. As a result, spread changes are similar to any price changes of any financial securities.

Default credit risk is different. It measures losses caused by actual defaults. Banks have been computing "counterparty risk exposures" for a long time. This is usually understood as a call option. That is, if the counterparty defaults, then all its debts to the bank will vanish (assuming no recovery). As a result, the loss is identical to the postive net exposure to the counterparty. If the net exposure is negative (i.e. the bank owes to its counterpary), then there is no risk. Hence, the loss mimics a call option. Banks usually simulate all possible scenarios of net exposures and sum those that have positive net exposure paths.

16.2 Unexpected Loss

Recovery of a debt is a lengthy process and the amount is highly unpredictable. For those distressed debts that are traded in the secondary market, recovery is often assumed to be the fair market price of the debt. But for those who do not have secondary market prices, rough estimates are applied. Rating agencies such as

Moodys and S & P provide rough statistics for senior unsecured and subordinated debts. Often these constant statistics are used as the recoveries for these debts. Under constant recovery, we may compute the expected loss and unexpected loss for a reduced form Bernoulli model:

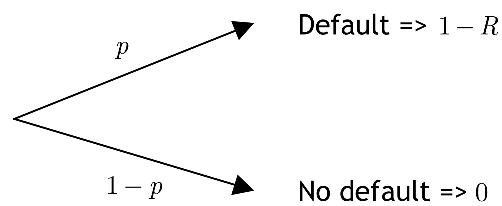


Figure 16.1: Expected Loss and Unexpected Loss

Hence, we compute the expected loss as,

$$\mathrm{EL} = p(1-R) \tag{16.1}$$

and unexpected loss as,

$$\mathrm{UL} = \sqrt{p(1-p)}(1-R) \tag{16.2}$$

Note that the unexpected loss is the standard deviation of loss under the Bernoulli distribution. Also note that the quantity of UL is at the maximum when $p = \frac{1}{2}$. Clearly this is when the uncertainty is the highest. When p is either high (extremely likely to default) or low (extremely unlikely to default), the uncertainty is low. Hence, there is little unexpected loss.

16.3 Term Structure of Credit VaR

ISDA in March 1998 published a manuscript entitled Credit Risk and Regulatory Capital in which it calls for a term structure of Credit VaR. Returns that embed credit risk are highly skewed. The standard nomality assumption for returns cannot apply. As a result, Credit VaR is not scalable by the square root of time. Credit-risky returns are affected by three major sources of risk:

- Market risk – over a short horizon where migration and defaults are not likely to occur,

- Migration risk – for medium horizon where actual defaults are not likely to occur but the perception of default likelihood has changed (binned into the transition matrix, Moodys 2001 one year average default rate for a BBB rating is 15 basis points but the likelihood to be downgraded is 13%)
- Default loss risk – loss due to actual default over the long run.

which is described in the following diagram:

Term Structure of Credit Risk		
Short horizon (< 1/4 year)	Medium horizon (between 1/4 and 1 year)	Long horizon (over 1 year)
Market risk	Migration risk	Default risk

Hence, Credit VaR is expressed as:

$$\begin{aligned} \text{credit VaR} &= \text{market risk weight} \times \text{market risk VaR} \\ &+ \text{migration risk weight} \times \text{migration risk metrics} \\ &+ \text{default loss weight} \times \text{unexpected loss} \\ &+ \text{idiosyncratic risk} \end{aligned} \tag{16.3}$$

As a result, a full term structure of Credit VaR must be considered. Recall in Lesson 1 where a loss distribution is presented, which is now used to compute CVaR. In this diagram, Credit VaR (CVaR) is the highest 5th percentile (or the 1st percentile) of the loss distribution. We notice that the loss distribution is skewed and hence unlike normal distribution, it cannot be scaled. That is, a 5-day CVaR is not $\sqrt{5}$ times of the 1-day CVaR. Hence, a CVaR must be separately calculated for each time horizon.

16.4 CVA – Credit Value Adjustment

CVA, or credit value adjustment, is a method to compute proper value if a trade takes on subtantial counterparty risk. If the trade has subsequent net cash inflows (so called in-the-money) from the counterparty but the counterparty is in default, then there is loss of unrealized gains. Under the requirement of marking to market (as investment banks), such unrealized gains have already been recognized at the time of the trade (so called day-1 P&L). As a result, the loss of the unrealized gains that have been recognized must be credited back and show in the current

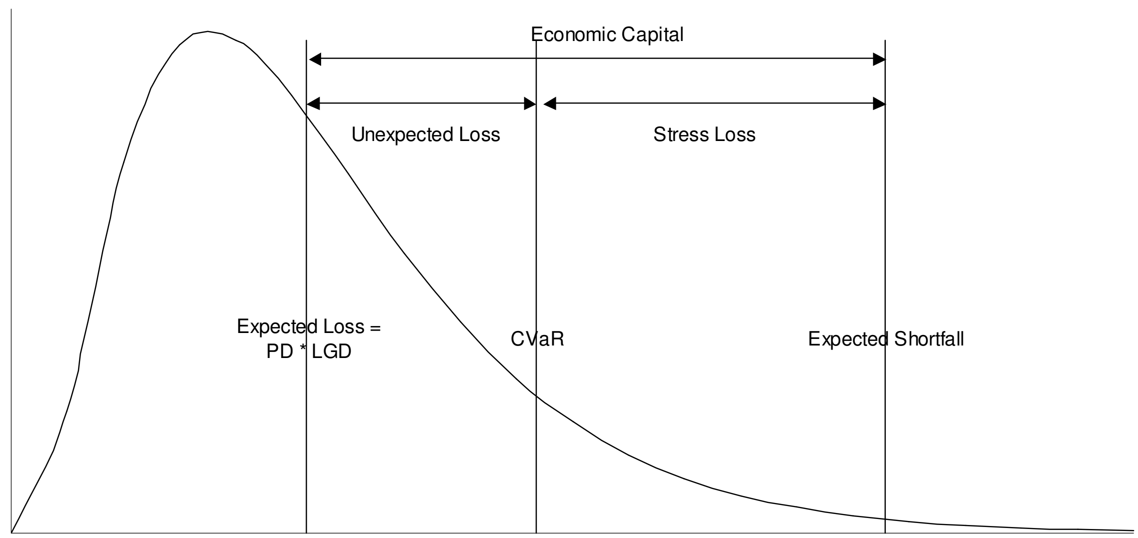


Figure 16.2: Credit Value at Risk

financial statements as a loss. In order not to show lumpy P&L fluctuations due to counterparty defaults, it has been a recent Wall Street trend that trading desks must put up reserves over time to accomondate losses due to counterparty defaults. This is called Credit Value Adjustment, or CVA.

To operationalize such a reserve-charging process, investment banks start to establish a CVA trading desk that sells protection on counterparty defaults. As opposed to burdening each and every trading desk to compute its counterparty default reserve, the CVA trading desk computes the fair price of the protection and sell CDS on each counterparty. CVA traders are then responsible for their own P&L trading CDS on counterparties. This is another successful securitization move on Wall Street that it securitizes default charges. Trading with CVA desk, each and every other trading desk gains protection on its counterparties. As the protection is provided in a form of swap, it is like an installment plan that requires a small payment periodically that is exactly like a reserve-charging system.

In a way, CVA is a mechanism of internal transfer pricing used widely in regular corporations. Yet through securitization, it lets market decide what the correct transfer price should be. This is consistent with the strong belief by Wall Streeters that the free economy, via the Invisible Hand, will set the right equilibrium. While CDS and CDO have been widely studied (See Chapter **??**) and powerful models have been developed, the modeling of CVA so far is still completely dependent on Monte Carlo, as no closed-form solutions exist and the correlation between counterparty default risk and the risk of the deal is hard to capture.

16.4.1 Exposure

The easiest way to understand the nature of counterparty risk is a concept of "exposure". An exposure is a potential loss upon a counterparty's default. Usually this happens when our counterparty owes us money (i.e. our unrealized gain). Note that if we owe our counterparty money, then the default of our counterparty is a good news for us as we do no longer need to pay our debt.

The concept of exposure is often described as a call option, as in Figure 16.3. The dotted line represents the "moneyness of the deal". If the deal is in-the-money (i.e. our counterparty owes us money), then there is a positive exposure dollar-for-dollar. Otherwise, i.e. out-of-the-money, the exposure is 0. The solid line is the PV of such exposure. As we can see, the exposure is a call option and we can apply the Black-Scholes model for it.

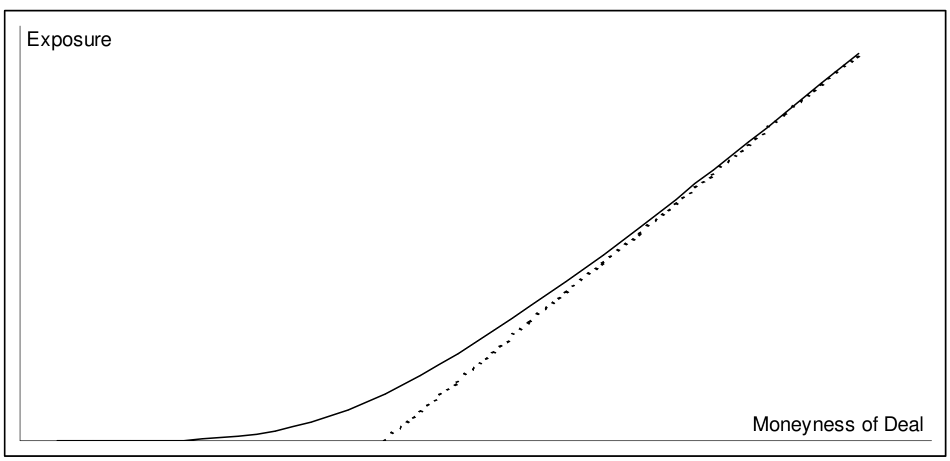


Figure 16.3: Counterparty Credit Exposure

If a deal has multiple cash flows, then the exposure is a collection (sum) of many such call options. Certainly, the simple sum of call options is incorrect as these exposures are inter-dependent. More precisely, later call options are dependent on earlier call options in that if a counterparty defaults at time t then there will be no exposure after t. A rough approximation is the probability-weigh each exposure but such an approximation still ignores inter-dependence. Furthermore, this exposure method ignores the correlation of the counterparty default risk and the underlying risks of the deal.

Lets now turn to CVA calculation. CVA is defined as:

$$\text{Present Value of Expected [exposure} \times \text{PD} \times \text{LGD]}$$

Before we write down the formal formula, lets take a look at a numerical example. Often the typical example is an Interest Rate Swap (IRS). The swap rate of an IRS deal can be set using the yield curve (arbitrate-free valuation). Formally, a swap rate can be calculated as the weighted average of the forward rates (on coupon dates). Lets have a four-year yield curve as follows (which has been used in the demonstration of the Ho-Lee model):

Yield Curve	
term	yield
1	5.0%
2	6.0%
3	6.5%
4	6.8%

The discount factors and forward rates can then be calculated as follows:

Yield Curve				
term	yield	disc. fact.	forward rate	weighted fwd rate
1	5.0%	0.9512	5.0%	1.39%
2	6.0%	0.8869	7.0%	1.81%
3	6.5%	0.8228	7.5%	1.80%
4	6.8%	0.7619	7.7%	1.71%
	PV01	3.4228	swp rate	6.72%

where the numbers are calculated using the formulas as follows.

$$f(0,i,j) = \frac{y(0,j)j - y(0,i)i}{j-i}$$

$$w(0,k) = \frac{\sum_{i=1}^{k} f(0,i-1,i)P(0,i)}{\sum_{i=1}^{k} P(0,i)}$$

$$P(0,i) = \exp\{-y(0,i) \times i\}$$

In two years, the swap deal has two years remaining and hence the market swap rate is determined by the current yield curve up to two years:

Yield Curve				
term	yield	disc. fact.	forward rate	weighted fwd rate
1	6.0%	0.9418	6.0%	3.12%
2	7.0%	0.8694	8.0%	3.84%
3	7.5%	-	-	-
4	7.8%	-	-	-
	PV01	1.8111	swp rate	6.96%

The swap rate has gone up in two years, from 6.72% (four-year swap rate) to 6.96% (two-year swap rate). In other words, the same deal (two-year swap) now investors have to pay a higher fixed rate to exchange for LIBOR. The existing contract is thought of as "in-the-money". Equivalently speaking, the existing contract can be sold for a profit. Imagine that the holder of the existing contract can do a reverse swap to receive 6.96% and pay LIBOR. As a result, the existing investor has no more LIBOR revenue but 6.96% (because LIBOR is netted out). Hence, the investor of the existing swap can pocket the difference between 6.96% and 6.72% for the next two years. Multiplying by the annuity factor (PV01), it is $(6.96\% - 6.72\%) \times 1.8111 = 43.44$ basis points (per dollar notional). If the notional is \$10 million, then the existing swap is \$43,436.34 in the money. When the counterparty defaults at this time, the investor will lose this profit. Hence, this is the exposure of the swap in two years. If the counterparty defaults but is able to pay some of this amount (i.e. recovery), then the lost is not the entire exposure but exposure times $1 - R$.

While this above simple example gives a concrete exposure number, the reality is that we never know what the yield curve is in two years. As a result, we do not know the swap in two years, and hence we do not know the exposure amount in two years. To estimate future yield curves, we then must use Monte Carlo simulations.

To write down the exact valuation formula:

$$
\begin{aligned}
\text{CVA} &= \mathbb{E}\left[\sum\nolimits_{j=1}^{n} \exp\left(-\int_t^{t_j} r_u du\right) X_j \mathbb{I}_{t_{j-1}<\tau<t_j}(1-R_j)\right] \\
&= \sum\nolimits_{j=1}^{n} P_{t,t_j} p_j (1-R_j)\mathbb{E}\left[X_j\right] \\
&= \sum\nolimits_{j=1}^{n} P_{t,t_j} (Q_{t_{j-1}} - Q_{t_j})(1-R_j)\mathbb{E}\left[X_j\right] \\
&= \sum\nolimits_{j=1}^{n} \xi_j \mathbb{E}\left[X_j\right]
\end{aligned}
$$

where p_j is the default probability between time t_{j-1} and t_j, which is equal to $Q_{t_{j-1}} - Q_{t_j}$ (the difference of two survival probabilities), X_j is the exposure lost during t_{j-1} and t_j, R is the recovery (assumed fixed), and $\mathbb{I}$. is the indicator function.

In a single period setting ($n = 1$), the above equation can be simplified as loss rate ξ times expected exposure $\mathbb{E}[X]$ and the loss rate is composed of risk-free discount, default probability, and LGD. Such an interpretation (often times we gain most of the intuition through a single period model) is very easy to understand the complex CVA calculation.

Due to the inter-connections among exposure, default probability, recovery and discounting, CVA can only be computed via Monte Carlo simulations. Figure 16.4 explains the process of the Monte Carlo simulations.

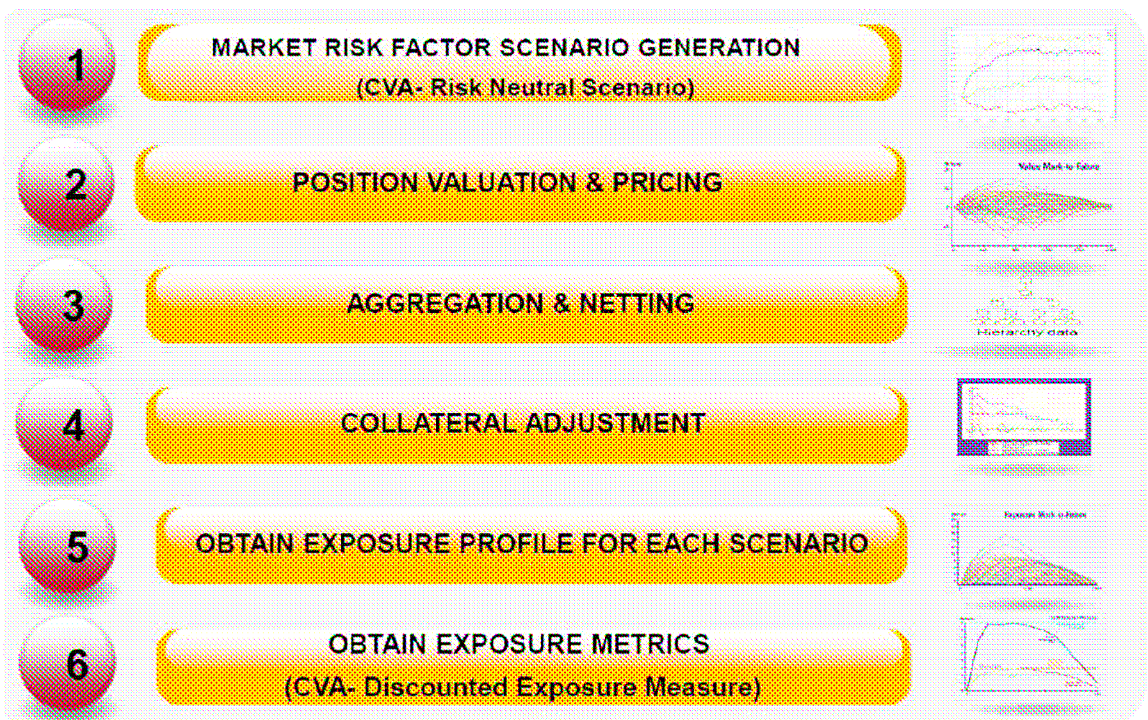


Figure 16.4: Process of Monte Carlo Simulations

In general (a generic deal) the exposure $X(t)$ can be affected by a number of risk factors, $y_i(t)$. Say

$$y_i(t) = a_i + \sum\nolimits_{k=1}^{K} b_{i,k} F_k(t) + e_i(t)$$

where $F_k(t)$ is a risk factor (see discussions on market risk (PCA) for this), $e_i(t)$ is idiosyncratic, and i represents various risk factors. Under this framework, all assets can be priced within a consistent framework. $y_i(t)$ is a state variable used to determine asset prices. If equity, then it is the return (or price change) of the i-th asset as described in Chapter 4. If interest rates, then it could be one of several factors such as: $r(t) = y_m(t)+y_{m+1}(t)+\cdots+y_{m+h}(t)$ assuming h factors that explain the interest rate. Then a bond valuation model can be used to price risk-free bonds (such as Vasicek in Chapter 4) or risky bonds (such as the Jarrow-Turnbull model in Chapter 13).[1] This is step 2.

Step 3 is to aggregate all positions into a portfolio. Apparently the exposure at any time t, $X(t)$, is a result of valuations of all assets owned and owed (netting). Step 4 is related to risky funding, which is to incorporate liquidity risk (to be discussed later). Step 5 is simulations and step 6 is taking the average.

Take the above swap as an example. Say the interest rates are results of a two-factor model (see market risk). Then we evaluate fixed leg and floating leg (evaluation) to get exposure.

16.4.2 CCDS (contingent credit default swap)

Given that CVA is now formally incorporated in valuation, it is proposed that CVA to be securitized. This is CCDS. It is IRS + CVA. That is, the floating leg of a CCDS is the loss due to default of the IRS counterparty. As a result, CCDS is a perfect hedge to CVA.

WWR (wrong way risk) In general, the exposure with a counterparty is not independent of the counterparty's credit quality. Wrong Way Risk is cases where the exposure increases when the credit quality of the counterparty deteriorates – i.e. exposure tends to be high when PDs are high.

There are two types of WWR:

- General WWR: the counterparty's credit quality is for non-specific reasons correlated with macroeconomic factors which also affect the value of the derivatives (e.g. correlation between declining corporate credit quality and high (or low) interest rates causing higher exposures
- Specific WWR: future exposure to a specific counterparty is highly correlated with counterparty's PD. (e.g. a company writing put options on its own stock, derivatives collateralized by own shares)

WWR quantification is still an open challenge other than the self referencing specific WWR due to:

- Difficulty to separate statistical noise from systematic correlation

[1] Jump intensities can also be functions of the factors.

- Challenge of dynamic forward looking adjustment to historical calibration

16.4.3 CVA Hedging

As CVA having become a real cost to trading desks, the hedging of CVA has become an important task as the value of CVA is now part of the P&L of the trading desk. The hedges of the CVA incorporate hedges of the market risk factors driving the exposures and hedges of the credit spreads of the counterparties. As in any hedging, correlation among different risk factors is usually the most important feature evaluating CVA. Empirical evidence has suggested that convexity and cross-convexity (gamma and cross-gamma) plays an important role in evaluating CVA, especially when the changes in spreads and exposures are large.

If the bank marks to market its CVA and the bank does not hedge it, it will experience P&L (and earnings) volatility. More importantly, in a trending or deteriorating credit market environment, the bank could suffer a substantial cumulative loss. The risk management of CVA requires dynamic rebalancing of the hedges. When counterparty exposures and credit spreads of the counterparties are large and volatile, rebalancing requirements can be intense and costly.

In general, changes in the exposure can be hedged by taking positions on the market risk factors that drive the exposure, but the hedging for its own credit spread is more challenging to hedge. The systematic risk component can be hedged. The bank-specific, idiosyncratic risk component is more difficult to hedge.

The ISDA Master Agreement lists two different tools to reduce exposure:

- Collateralization, the right of recourse to some asset of value that can be sold or the value that can be applied in the event of default on the transaction
- Close-out Netting rules, which state that if a default occurs, multiple obligations between two parties are consolidated into a single net obligation

16.4.4 CSA (credit support annex)

Counterparty credit risk can also be mitigated by margining practice through incorporating Credit Support Annex (CSA agreements). A CSA provides credit protection by setting forth the rules governing the mutual posting of collateral.

CSAs are used in documenting collateral arrangements between two parties that trade privately negotiated (over-the-counter) derivative securities. The trade is documented under a standard contract called a master agreement, developed by the International Swaps and Derivatives Association (ISDA). The two parties must sign the ISDA master agreement and execute a credit support annex before they trade derivatives with each other.

In addition to executing the ISDA master agreement and credit support annex, issuers must implement proper resolutions that give authorization to execute any derivative transactions. Each issuer must also obtain an opinion from its respective legal counsel about whether both parties can enter into swap transactions. Issuers must also ensure that such contracts are binding and enforceable, and obtain final credit approval from a bank.

ISDA Credit Support Annex (CSA)

- Permits posting one or several types of collateral with periodic rebalancing and interest paid by the receiving party based on collateral currency and type
- Permits thresholds, minimum transfer amounts, and rounding (to reduce operational costs)
- A new standard form of CSA (SCSA) is in development which will reduce the complexity due to the collateral type switch options embedded in traditional CSA
- Fully collateralized counterparty is counterparty with a perfect CSA all thresholds, minimum transfer amounts, and rounding are zero, with daily rebalancing
- If the market moves against the trade and CSA is in place, massive additional funding will be required immediately to post collateral
- In the absence of CSA, expected future losses may cause immediate crisis of confidence causing creditors to pull funding from the firm

Key exposure / CVA affecting CSA terms include:

- Margin call frequency
- Threshold and Minimal Transfer Amount (MTA)
  - At the end of period (day, week, etc.), if trade MTM value exceeds a threshold, collateral must be posted.
  - Exposure is to the threshold plus the market moves between default and liquidation
  - Symmetry and currency (one-way-in, one-way-out, two way)

- Initial Margin (IM): only affects claims at default and close-out

Clearing houses mitigate risk via netting, collateralization and reassignment of contracts. Clearing houses charge initial and variation margins to make needed cash available in the event of a default. Furthermore,

- If that proves insufficient, coverage comes from backers of clearing house and equity of the firm itself
- Reassignment of contracts prevents losses due to market impact. But in a real crisis, reassignment might fail Another approach to risk mitigation is to include break clause (a.k.a, additional termination events)
- Contract can be close-out at replacement value if the counterpartys rating drops
- Contract can be closed out at market value prior to maturity
- Issues
  - What exactly is replacement value?
  - How is closing out of a swap early any different from entering into the reversing swap?
  - By forcing a closeout upon a rating change, are you decreasing counterparty risk while increasing systemic risk?
  - What if the market is illiquid?

16.4.5 Counterparty Credit Risk (CCR) as Market Risk

If CCR is actively managed and hedged, it is appropriate to treat CCR as part of market risk. CCR can be incorporated in the trading book by adding to the trading book one defaultable exotic virtual trade per counterparty. For each counterparty, the virtual trade is defined according to:

- if the counterparty has not defaulted by time t , the virtual trades value at time t is equal to $-CVA(t)$
- no cash flows occur unless the counterparty defaults
- at the time of the counterpartys default at time τ, the bank pays a single cash flow equal to LGD multiplied by exposure, and the trade terminates

16.4.6 Counterparty Credit Risk (CCR) as Credit Risk

Many banks do not actively manage CCR, but hold this risk to the portfolio maturity. For such banks, joint treatment of market risk and CCR may not be appropriate. Note that the time horizon used for market risk calculations is usually short, which can be justified only if the risk is actively managed. Hence, treating CCR as credit risk (jointly with the banking book) may be more appropriate for such banks.

A primary challenge in treating CCR as credit risk is uncertain nature of counterparty credit exposure. Very often, this challenge is overcome by using loan portfolio models with deterministic loan equivalent exposures that are calculated from counterparty exposure distributions.

16.4.7 Counterparty Credit Risk (CCR) Capital under Basel II & III

CCR capital under Basel II (BCBS, 2006) is treated as credit risk

- For default scenarios, asymptotic single risk factor (ASRF) model is used (see Gordy, 2003).
- Capitalization of no-default scenarios (credit migration risk) is done via calibrating a maturity adjustment (MA) factor to a MTM credit risk model, similar to KMV Portfolio Manager and consistent with the ASRF model (BCBS, 2004).

Basel III (BCBS, 2010) treats CCR in a mixed way:

- Default capital charge treats CCR as credit risk via ASRF framework (similar to Basel II, with minor changes).
- MA factor (same as in Basel II i.e., credit risk treatment) is still used for capitalization against credit migration losses.
- A new CVA capital charge is used to capitalize against CVA losses due to credit spread changes (EE is assumed to be fixed). CVA capital charge treats CCR as stand-alone market risk

16.4.8 CVA Capital Charge and Basel III

Counterparty Credit Risk (CCR) is one of the primary focus points of Basel III (BCBS, 2010). Credit Valuation Adjustment (CVA) is part of the regulatory capital calculations for CCR under Basel III. In addition to default capital charge, banks will be required to calculate a CVA capital charge.

Default charge capitalizes against losses from counterparties defaults. CVA charge capitalizes against losses from CVA increases for surviving counterparties.

Basel III motivation of CVA capital charge.

"Roughly two-thirds of CCR losses were due to CVA losses and only about one third were due to actual defaults. The current framework addresses CCR as a default and credit migration risk, but does not fully account for market value losses short of default."

Basel III CVA Capital Charge

CVA capital charge is calculated for the entire portfolio of OTC derivatives and allowable CVA hedges (securities financing transactions (SFT) are not included). Allowable CVA hedges include single-name and index credit default swaps (CDS) contracts (allowable hedges are removed from market risk calculations).

Two methods are available: advanced (based on simulations) and standardized (based on formula)

- Banks with Internal Models Method (IMM) approval for CCR and approval to use the market risk internal models approach for the specific interest-rate risk of bonds must use advanced method
- All other banks must use the standardized method

16.4.9 Advanced CVA Capital Charge

Advanced CVA charge is based on stand-alone CVA VaR which is a VaR charge resulted from CVA. Usually Monte Carlo simulations are used to compute this quantity. This is similar to the calculation of credit exposures that can only be simulated. In simulations, changes of credit spreads for all counterparties and of credit indexes are simulated for a 10-day horizon using internal models on a stand-alone basis. For each scenario, changes of CVA for each counterparty and those of MTM values of allowable hedges are calculated. For each scenario, changes of counterparty CVAs and those of hedge MTMs are aggregated across counterparties and hedges, resulting in a distribution of changes in portfolio-level hedged CVA. Then 99% VaR is calculated from the simulated distribution of changes in the portfolio-level hedged CVA.

Next, CVA capital charge is obtained from CVA VaR according to market risk rules (BCBS, 2009). The calculation must include Stress VaR, but exclude Incremental Risk Charge (IRC).

Standardized CVA Capital Charge

Standardized CVA charge under Basel III, K, is calculated according to a formula:

$$K = \beta\sqrt{T}N^{-1}(q)$$

where

$$\beta^2 = \left[\frac{1}{2}\sum_{i=1}^{N} w_i \left(X_i M_i - B_i^* M_i^{(h)}\right) - \bar{w}\bar{B}^*\bar{M}^{(h)}\right]^2 + \frac{3}{4}\sum_{i=1}^{N} w_i^2 \left(X_i M_i - B_i^* M_i^{(h)}\right)^2,$$

$T = 1$ year is the time horizon and $q = 99.9\%$ is the confidence level, w_i is standardized "weight" based on credit rating of counterparty i, X_i and M_i are the exposure at default (EAD) and is effective maturity respectively for counterparty i, B_i^* and $M_i^{(h)}$ are the discounted notional and the maturity of single-name CDS hedge respectively on counterparty i, and $\bar{B}^*$ and $\bar{M}^{(h)}$ are the discounted notional and the maturity of index CDS hedge.

16.5 Risky Funding

The recent 2007-8 crisis has revolutionized what financial modeling has long believed in – the law of one price, a result achieved by no-arbitrage. No-arbitrage was the basic principle Black and Scholes used to derive their option formula. No-arbitrage gave birth to risk-neutral pricing that has dominated the quant finance area for half of a decade. No-arbitrage guarantees the law of one price.

We should recall how Black and Scholes derived the pricing equation (i.e. the partical differential quation). This is known as "price by replication". In other words, Black and Scholes taught us that the price of any security is equal to its cost of hedging (or replicating). Later on, this was shown as the Martingale Representation Theorem (a.k.a. self-financing, which is the base of no-arbitrage). However, to achieve this result, one must have the same borrowing and lending rate. Well, this is not a problem under no-arbitrage, since the only rate relavent under no-arbitrage (or risk-neutral pricing) is the risk-free rate.

Until the 2007-8 crisis, the paradigm seemed to work just fine. Everything is calibrated to LIBOR which is regarded as the risk-free rate by the investment community. The crisis changed everything. During the crisis, deals that were safely collaterized were priced very differently (at much higher prices) from those that were not safely collaterized (either under-collaterized or un-callaterized). Buyers did not have faith in their counterparties's ability to fund the transactions (if the deals were under- or un-collaterized) and consequently demanded for lower prices. LIBOR and OIS rates started to diverge and LIBOR was no longer the risk-free rate anymore. This is the issue called Risky Funding.

There has not been any solution yet for this problem. Proposals have been provided but no conclusion has been drawn. This is a still very live and challenging problem.

16.6 Appendix

16.6.1 Poisson Process of Defaults

A Poisson process, named after the French mathematician Simon-Denis Poisson (1781 - 1840), is the stochastic process in which events occur continuously and independently of one another. The Poisson process is a continuous-time process: its discrete-time counterpart is the Bernoulli process.

A homogeneous Poisson process is characterized by a rate parameter λ, also known as intensity, such that the number of events in time interval $(t, t+\tau]$ follows a Poisson distribution with the associated parameter $\lambda\tau$. This relation is given as:

$$\Pr[N(t+\tau) - N(t) = j] = \frac{(\lambda\tau)^j e^{-\lambda\tau}}{j!} \tag{16.4}$$

Hence,

$$\Pr[N(t+\tau) - N(t) = 0] = e^{-\lambda\tau} \tag{16.5}$$

The Poisson process is an ideal way to model unexpected defaults as an event occurs with no prior memory. (16.5) describes the probability of no default between $(t, t+\tau]$. The mean and variance of a Poisson distribution are:

$$\mathbb{E}[j] = \mathbb{V}[j] = \lambda\tau \tag{16.6}$$

Exponential Distribution

If an event occurs with a Poisson distribution, then it can be shown that the time of the event occurring follows an exponential distribution. The exponential distribution is (with parameter λ):

$$f(x) = \lambda e^{-\lambda x} \tag{16.7}$$

Hence, its cumulative density function is:

$$\begin{aligned}\Pr[x < \tau] &= F(\tau) \\ &= \int_0^t f(x)dx \\ &= 1 - e^{-\lambda\tau}\end{aligned} \tag{16.8}$$

or

$$\begin{aligned}\Pr[x > \tau] &= 1 - F(\tau) \\ &= e^{-\lambda\tau}\end{aligned} \tag{16.9}$$

which is the survival probability when t represents the default time. When the time interval is small, we can use Taylor's series expansion to approximate the survival probability as:

$$\Pr[x > \tau] = e^{-\lambda\tau} \approx 1 - \lambda\tau \tag{16.10}$$

or the default probability $\lambda\tau$. The mean and variance of the exponential distribution are:

$$\begin{cases}\mathbb{E}[x] = \dfrac{1}{\lambda} \\ \mathbb{V}[x] = \dfrac{1}{\lambda^2}\end{cases} \tag{16.11}$$

Note that the Poisson process has no memory. Hence, the exponential distribution has the following property:

$$\begin{aligned}\Pr[x > s + t | x > s] &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\ &= e^{-\lambda t} \\ &= \Pr[x > t]\end{aligned} \tag{16.12}$$

In plain words, what this means is that the expected wait time (till the event) has nothing to do with how long one has waited. The event is not expected to happen sooner even though you have already waited for a long time.

We compute survival probabilities when we model default. The survival probability between now and some future time T if we adopt the Poisson process (i.e. default time τ) for defaults is:

$$Q(t, T) = \Pr[\tau > T] = \hat{\mathbb{E}}_t[\mathbb{I}_{\{\tau > T\}}] = e^{-\lambda(T-t)} \tag{16.13}$$

where $\mathbb{I}$ is an indicator function.

16.6.2 Equation 13.23

For the first year, which contains four quarters, $n = 4$ and we have:

$$
\begin{aligned}
x_1 &= x_0 + 0.6p_1 + (1 - p_1)x_0 \\
x_2 &= x_0 + 0.6p_1 + (1 - p_1)x_1 \\
x_3 &= x_0 + 0.6p_1 + (1 - p_1)x_2 \\
x_4 &= x_0 = -0.000225
\end{aligned}
$$

Substitute recursively into x_4 to get:

$$
0 = \frac{0.6p_1 + (1 - p_1)\left[-0.000225 + \frac{0.6p_1 + (1-p_1)\left[-0.000225 + \frac{0.6p_1 + (1-p_1)\left[-0.000225 + \frac{0.6p_1 - 0.000225(1-p_1)}{1.0125}\right]}{1.0125}\right]}{1.0125}\right]}{1.0125}
$$

and solve for $p_1 = 0.000375$ or 3.75 basis points.

Part IV

Liquidity Risk

Chapter 17

Liquidity Quantification

17.1 Introduction

The default of Lehman Brothers Inc on September 15, 2008 marked the unprecedented crisis in the global history. This crisis differed from the previous ones in that:

- it caused the largest bankruptcy in the U.S. history (over $600 billion)
- it caused a global recession
- it recorded the longest recession in the U.S. history
- it started global awareness on regulation

And the trigger of Lehman default is liquidity. While the cause of the crisis, just line any other, is a bubble burst in the real estate market, and frauds committed in making unjust mortgage loans (so called subprime loans), the snowball of one default of Bear Sterns has been unprecedented and liquidity driven. As a result, liquidity risk management has been the focal point of regulation reform in EU (Basel) and the U.S. (Dodd-Frank).

During the last financial crisis (known as the liquidity crisis), a large number of liquidity-squeezed events had occurred. These incredible events had caused Wall Street to consider abandoning its long proud tradition – marking to market. For as long as Wall Street has existed, it has prided itself in respecting the market prices. Wall Street had always regarded market prices to be a reflection of collective wisdom of the entire investment community. The fact that there is a price reflects an equilibrium. Wall Street is proud to be able to take the full advantage of such equilibriums and develop tools that are consistent with market prices. But during the crisis, such a belief had been put under a strong test. Here are some famous examples.

17.1.1 Some Liquidity Squeeze Examples

Lone Starr-Merrill Lynch Deal

On July 28, 2008, a Monday, as the crisis of subprime gradually unfolded, Merrill Lynch, in realizing huge toxic portfolios it owned, would like to sell a particular subprime portfolio that had a face value of $31 billion to a private equity fund owned by Mr. Lone Star. The purchase price was rumored to be 22 cents on the dollar – a 78% discount. On top of that, Merrill Lynch needed to buy back any defaulted asset in the portfolio. In other words, Merrill Lynch was implicitly sold a CDS (for free) to Star, which was worth about 15 cents on the dollar. As a result, Merrill Lynch sold Lone Star a portfolio of financial assets at a discount of 93%!

According to Reuters, the deal had prompted Merrill Lynch to write down all of its toxic assets by at least 75%, an unprecedented price discount seen on Wall Street.

Bear Stearn-JP Morgan Deal

On June 22, 2007, the problems of the two subprime portfolios managed by Bear Stearns Asset Management (BSAM) – Bear Stearns High-Grade Structured Credit Enhanced Leveraged Fund and Bear Stearns High-Grade Structured Credit Fund, started to surface. Its CEO James Cayne initially wanted to keep the two funds separate from the main Bear Stearn bank (as how deals are usually done on Wall Street – so called special purpose vehicle, or SPV). Yet the pressure from the investors and the government forced Bear Stearns to take over. As losses gradually unfolded, Richard A. Marin, a senior executive at Bear Stearns Asset Management responsible for the two hedge funds, was replaced on June 29 by Jeffrey B. Lane, a former Vice Chairman of rival investment bank, Lehman Brothers.

On March 14, 2008, Bear Stearns finally reached to a point where it could no longer operate. Either someone would have to buy it or else it would face liquidation. The Fed called in JP Morgan for the bail out. The price was at $30 per share on March 14, a Friday. After a weekend-long investigation and evaluation of Bear's assets, JP Morgan offered $2 per share. Bear rejected and accused JP Morgan of exploiting the liquidity situation. In fear of a bank run that Bear's bankruptcy could have caused, the Fed yielded to JP Morgan for liquidity concern by agreeing a $30 billion loan to J.P. Morgan. The final deal was $10 per share.

It is a consensus that Bear's default would have been an economic consequence (that is, they lost money in their investments). Yet the bailout amount (from $2 per share to $10 per share) is undoubtedly a consequence of a liquidity squeeze. The liquidity discount here is 80%!

Other Examples

In addition to these two unprecedented liquidity events, there have been numerous examples during the crisis period where asset prices were largely compressed due to urgent needs to unwind those positions (clear evidence of liquidity squeeze).

17.2 Understanding Liquidity and Liquidity Risk

Liquidity standards aim to ensure that a bank is able pay its liabilities on time by holding enough highly liquid assets that could be quickly converted to cash.

Ever since the recent financial crisis, liquidity has been the centerpiece in many regulations. Basel III explicitly calls for liquidity regulations (LCR or Liquidity Coverage Ratio, and NSFR or Net Stable Funding Ratio). Dodd-Frank Section 165 directs the Federal Reserve Board (FRB) to establish heightened liquidity standards for both bank holding companies with over $50 billion in assets and FSOC-designated non-bank systemically important financial institutions (covered companies).

17.3 How to Measure Liquidity

Liquidity is generally understood in the following three areas:

- market microstructure liquidity
  - volume
  - bid-offer spread
  - price movement
- banking liquidity
  - Basel III – LCR and NSFR

  - Dodd-Frank – to be determined

- accounting going concern liquidity (whether or not a firm has enough cash (or liquid assets that can turn into cash quickly) to pay of its expenses)

Accountants for centuries have played an essential in diagnosing the liquidity healthiness of a firm. The so-called "going concern audit" is an audit opinion to tell the stock holders if a firm is likely to survive through the next year (going concern is an annual audit).

In five accounting ratio groups – productivity (e.g. ROE), profitability (e.g. profit margin), market (e.g. PE ratio), efficiency (e.g. turnover ratios), and liquidity (acid ratio), liquidity measures take into account of if the firm has enough cash for the coming year to pay for its short-term liabilities.

This is the most conservative measure of liquidity as only cash (or the extremely marketables) is considered. Other assets that can be liquidated to pay for the liabilities are not considered. However, if each asset can be measured with a liquidity discount, then the firm can remain liquid as long as the total asset value is enough to cover the short-term liabilities. However, doing so requires a model to evaluate each asset with a liquidity discount.

17.4 Liquidity and Liquidity Risk

Liquidity is a static measure of how much should there be enough liquid assets (such as cash) for a firm to survive in a specified period (usually a year). Liquidity risk, on the other hand, refers to future liquidity needs and the likelihood that the firm may not be able to meet the needs. In other words, liquidity risk is about future liquidity shortfalls. The following summarizes the difference between liquidity and liquidity risk:

- liquidity is static and liquidity risk is dynamic (probabilistic)
- liquidity is single period and liquidity risk is multiple periods
- liquidity risk requires models and liquidity does not

The diagram below depicts liquidity shortfalls (a.k.a. liquidity gaps). In the diagram the demand for liquidity on the asset side exceeds the supply of liquidity on the liabilities side. Hence there is a liquidity gap. Liquidity risk is to study future such gaps and compute the impact of these gaps.

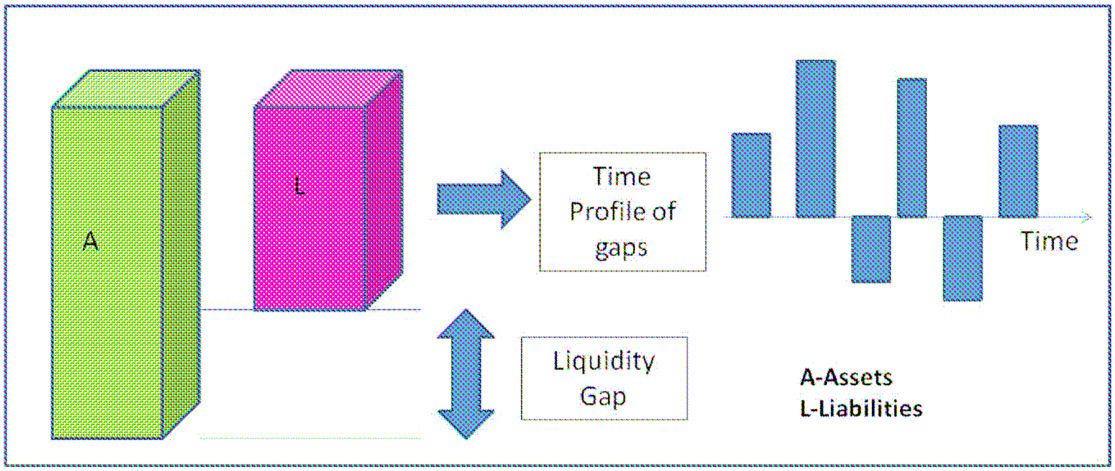


Figure 17.1: Liquidity Risk Gap

17.5 How to Measure Liquidity Risk

17.5.1 Liquidity Discount as a Put Option

Liquidity discount occurs as supply reacts more sensitively to economy than demand does in a non-linear fashion. Liquidity discount reaches the extreme as demand reaches its maximum capacity. Figure 1 depicts this general idea. In Figure 1, we assume that demand and supply of a financial asset jointly determine the equilibrium price of the asset at any given time. The economy is represented by a single state variable (say wealth) symbolized by V. As the economy grows, the supply curve moves to the right and so does the demand curve. To derive the liquidity discount model, we must assume that supply grows faster than demand does in that less elastic demand function is the main reason to cause price discount.

As a result, the growth of economy results in an increased equilibrium quantity and a lower price. In the diagram, we assume that the demand is totally insensitive to economy growth and it has a maximum capacity at Q^*. Clearly, our model requires only demand be less sensitive to supply and such an exaggerated demonstration is just for the purpose of easy exposition. At the maximum capacity of the demand function, the quantity can no longer increase and equilibrium can only draw the price down.

In the diagram, the vertical axis represents price (by S) and the horizontal axis represents quantity (by Q). We let the liquidity-constrained price be S^*. In a usual situation, $S = S^*$. In a liquidity-squeezed situation (represented by the situation

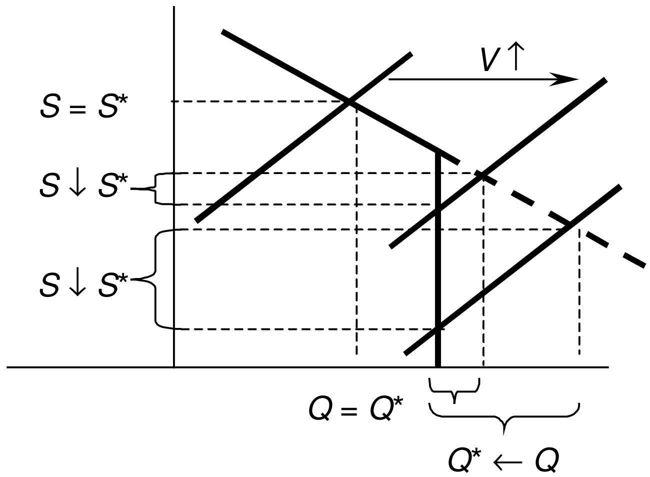


Figure 17.2: Demand-Supply Analysis of Liquidity Discount

where demand reaches its maximum capacity Q^*), $S > S^*$. Liquidity discount in this setting is defined as the impact of the slower demand reaction (than supply) and becomes dramatic as the maximum is approached. Based upon the above diagram, we can draw the following conclusion depicted in Figure 2.

On the left of Figure 17.3, we depict the relationship between the economy (represented by, say wealth V) and perfectly liquidity price S. As shown in Figue 17.2, this is a downward sloping curve. To be shown later, the curvature of S in V must be convex in order to obtain liquidity discount. If the relationship is linear (under which Q^* cannot exist), then there can be no liquidity discount. ,

On the right, we depict the relationship between the perfectly liquidity price S and the liquidity-constrained price S^*. The line $\overline{\text{ABC}}$ is a 45-degree line on which the illiquid price is equal to the liquid price. At point B where the quantity reaches its maximum capacity Q^* the illiquid price starts to decrease rapidly due to problems in liquidity and bends over toward point D. Again, the linear result (by $\overline{\text{BD}}$ line) is just a demonstration. In the next section where we derive the formal model, the graph between points B and D is not linear and has a reflection point.

From the diagram where discount is depicted linearly, it can be seen that liquidity can be explained by a put option. That is,

$$S^* = S - \text{put} \tag{17.1}$$

In the next section, we demonstrate that this is not a simple put due to convexity requirements. But intuitively, the simple put explanation serves the purpose

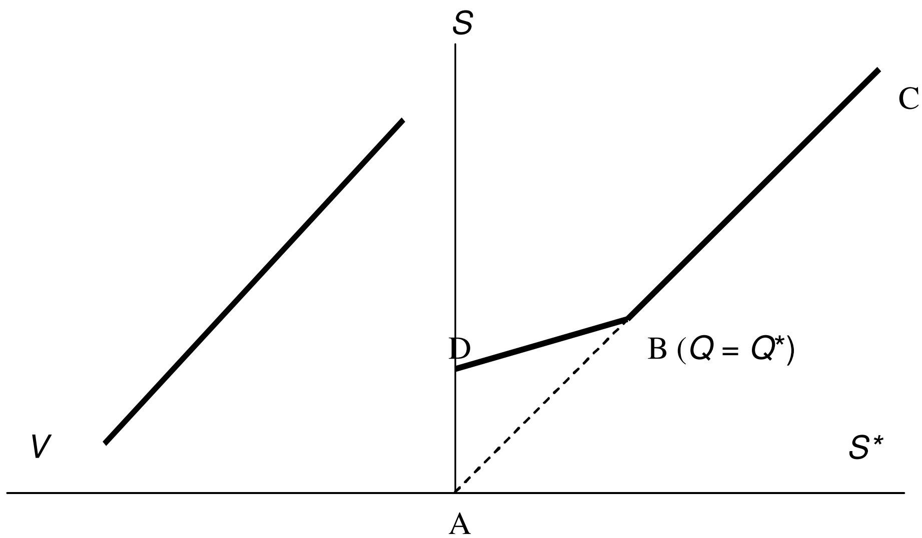


Figure 17.3: Liquidity Discount as a Put

well and can even be used as an approximation, which has already been discussed in the industry. What is offered in this paper is an equilibrium model that is consistent with the Merton model widely used in modeling credit risk. While we do not link our liquidity model yet to credit risk in this paper, it is quite straightforward to do so.

Take the Black-Schole put formula as an example:

$$\text{put} = e^{-rT}K[1 - N(d_2)] - S[1 - N(d_1)] \tag{17.2}$$

where

$d_1 = \dfrac{\ln S - \ln K + (r + 0.5\sigma^2)T}{\sigma\sqrt{T}}$

$d_2 = d_1 - \sigma\sqrt{T}$

Assuming $S = 100, K = 120, r = 0.04, \sigma = 0.3, T = 1$, the discount for liquidity is \$21.88, which means that the illiquid price is $100 - 21.88 = \$78.12$. Note that K is a parameter in the model that decides critically the value of the discount. The larger is K, the more severe is the discount. When K is 0, then there is no discount. As K approaches infinity, the value of discount is approaching \$100 and the illiquid price becomes 0.

In this section, the demonstration is intuitive but not realistic. In reality, the left and the right panels are related. In other words, how the liquidity price reacts to the economy affects the magnitude of liquidity discount. Furthermore, it is not likely to know the exact amount of Q^*. In the next section, we endogenize the relationships amount the economy V, the perfect liquid price S, and the liquidity-constrained price S^*. We let the liquid price be a convex function of the economy and derive the illiquid price directly. In doing so, we arrive an equilibrium without the specification of Q^*.

17.5.2 The Model

In this section, we develop a formal pricing model for illiquidity which is defined as inability to transact. Inability to transact is the exact description of Figure 17.2 in which lack of demand leads to no transactions. In a perfect world where the Black-Scholes model holds, transactions can take place at any time and hence investors can trade securities and rebalance their portfolios continuously. When transactions are not permitted to be continuous, it presents an extra risk born by the buyer and hence the buyer, in return, should ask for compensation and lower the price of the security.

Once continuous trading is not allowed, market is not (dynamically) complete in the Duffie-Huang sense (1985) and the resulting model is not preference-free. As a result, one must adopt a utility function to gauge the magnitude of the risk premium. We first present a model with the quadratic utility so the standard CAPM can be used. It is straightforward to extend the model to a broader class of utility functions. Using a more complex utility function is certainly better in terms of explaining the reality and providing model flexibility but it loses the closed-form CAPM formula.

The price of an arbitrary security at current time t when no trading is allowed until a future time T can be priced by the most fundamental discounted cash flow method:

$$X(t) = e^{-\xi(T-t)}\mathbb{E}_t[X(T)] \tag{17.3}$$

where $X(t)$ is the cash flow of an arbitrary security at time t, $\mathbb{E}_t[\cdot]$ is the conditional expectation taken at time t, and ξ is the (continuously compounded) risk-adjusted return for the security. Such a discounted cash flow method requires further modeling substances in order to be operational.

To build an explicit model for (17.3), we first assume the Black-Scholes/CAPM model where the underlying economy (represented by a single state variable, say wealth) obeys the following log normal process:

$$\frac{dV}{V} = \mu dt + \sigma dW \tag{17.4}$$

where $W(t)$ is the standard Wiener process and μ and σ are (continuous time) mean and standard deviation of the return of the state variable V.

A perfectly liquid price, S, is a contingent claim on the state variable. The assumption of liquidity discount we make in this paper indicates that S must be a monotonic function in wealth V. In a theorem we shall prove later, the function must be convex in order to arrive at liquidity discount. There are a number of ways to construct such an explicit function. For simplicity, we choose a function that imitates a put option (as opposed to an arbitrary polynomial function) for the following reasons. First, the maturity parameter in the put option can be made equal to the liquidity discount horizon. This provides an extreme convenience in modeling liquidity discount. Secondly, there is a closed form solution for the price of a perfectly liquid asset, which allows us to compare with the price of an equal asset but constrained by liquidity. Lastly, the strike price in the put function ideally characterizes the strength of the liquidity squeeze. The higher is the strike price, the stronger is the liquidity discount and the zero strike price ideally represents perfect liquidity. However, the use of put function does suffer from one drawback. It exists a maximum value for the liquid stock price (at the strike level), which could be unrealistic as economy contracts the price of the security could become unboundedly high. Fortunately, this is the situation where the liquidity discount is small and hence the impact would be small.

Given that S is the price of a security that can be continuously traded, it can be easily computed with the Black-Scholes model when its payoff for a fixed time horizon (time to maturity) mimics a put:

$$\begin{aligned} S(t) &= e^{-r(T-t)}\hat{\mathbb{E}}_t[S(T)] \\ &= e^{-r(T-t)}\hat{\mathbb{E}}_t[\max\{K - V(T), 0\}] \\ &= e^{-r(T-t)}KN(-d_-) - V(t)N(-d_+) \end{aligned} \tag{17.5}$$

where K is the strike price that reflects the strength of the liquidity squeeze, $\hat{\mathbb{E}}_t[\cdot]$ is the risk-neutral expectation, and

$$d_{\pm} = \frac{\ln V(t) - \ln K + (r \pm {}^1\!/_2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

Note that the adoption of the put payoff is just a convenience to incorporate convexity. Due to the fact that the liquidity-constrained price does not have a closed-form solution (except for the extreme cases), we implement a binomial model to approximate this Black-Scholes result. For the remainder of the paper, when we refer to the Black-Scholes model, it is actually the binomial approximation.

We also note that although the liquidity discount model has no closed form solution, the binomial implementation is actually better in that the model can then be easily augmented to include credit risk proposed by Chen (2002), and Chen, Fabozzi, Pan, and Sverdlove (2008). In this broader model, interactions between liquidity and credit risks can be studied.

Let S^* be the liquidity-constrained price where trading of the security is not permitted until time T. Equation (17.3), as a result, can be replaced with the following equation:

$$X(t) = e^{-\xi(T-t)}\mathbb{E}_t[f(V(T), \Theta), 0\}] \tag{17.6}$$

where Θ represents other parameters needed for the model, $X(t)$ is either S or S^* depending if the equation is used for an instantaneous period or a longer time period, respectively. Note that in an instantaneous period, the physical expectation as in (17.6) is identical to the risk-neutral expectation (details to be explored later in Theorem 1) and hence (17.6) provides the solution to the liquid price. The expected return, ξ, must follow the Capital Asset Pricing Model (CAPM).

Note that S and S^* are identical securities with just one difference: S^* cannot be transacted (or hedged) until time T. The purpose of the paper is to derive the price difference between S and S^*. We shall prove that $S^* < S$ and this is the model of liquidity discount. As we shall demonstrate in an analysis that such discount can be substantial, even under very reasonable assumptions.

We begin our modeling of liquidity discount with the standard CAPM. Note that the Black-Scholes model is consistent with the standard CAPM as follows:

$$\xi = \mathbb{E}\left[\frac{dX}{X}\right] = rdt + \eta\left[\mathbb{E}\left[\frac{dV}{V}\right] - rdt\right] \tag{17.7}$$

where r is the risk-free rate and

$$\eta = \frac{\partial X}{\partial V}\frac{V}{X} \tag{17.8}$$

is the elasticity of the state contingent claim with respect to the underlying economic state variable (similar to an option on its underlying stock). Note that this result is exact only in continuous time. Black and Scholes (1973) prove that the option price is a CAPM result with the underlying stock as the market. Note that $\eta = \beta$ as shown in the following:

$$\begin{aligned}\beta &= \mathbb{K}\left[\frac{dX}{X}, \frac{dV}{V}\right] \Big/ \mathbb{V}\left[\frac{dV}{V}\right] \\ &= \frac{V}{X}\frac{\mathbb{K}[X_V dV, dV]}{dV} \\ &= \frac{V}{X}X_V \\ &= \eta\end{aligned} \tag{17.9}$$

where $X_V = \partial X/\partial V$. Re-writing (17.7) in discrete time for a small interval h, we have:

$$\frac{\mathbb{E}[X(t+h)]}{X(t)} = (1+rh) + \beta\left[\frac{\mathbb{E}[V(t+h)]}{V(t)} - (1+rh)\right] \tag{17.10}$$

As a result, we can derive a pricing model for the security as:

$$X(t) = \frac{1}{R(t,t+h)}(\mathbb{E}[X(t+h)] - X_V\{\mathbb{E}[V(t+h)] - RV(t)\}) \tag{17.11}$$

where $R(t,T) = 1 + r(T-t) \approx e^{r(T-t)}$. This result provides an alternative proof of the CAPM argument by Black and Scholes. Note that

$$\begin{aligned}X_V &= \frac{\mathbb{K}[X(t+h), V(t+h)]}{\mathbb{V}[V(t+h)]} \\ &= \beta^{\$}\end{aligned} \tag{17.12}$$

is the "dollar beta" (that is, delta is dollar beta). Hence, (17.11) is also an alternative derivation to Jensen's model (1972).

It is important to note that CAPM holds only under one of the two assumptions: quadratic utility for the representative agent in the economy or normality for the returns of the risky assets. Consequently, the above CAPM result holds in the Black-Scholes model without any utility assumption must be due to the fact that option returns and stock returns are both normally distributed. It is clear that the stock return is normally distributed, as equation (17.4) postulates. The option return is normally distributed only under continuous time. This is because in continuous time, the option value is linear in stock and must follow the same distribution of that of the stock and hence its return is normally distributed.

In a discrete time where h is large, the option return is no longer normally distributed and the equation (17.11) can no longer hold without the assumption of quadratic utility that guarantees the validity of the CAPM quadratic utility function. The quadratic utility function is a bad assumption in that it has the wrong sign for the relative risk aversion. However, fortunately, the liquidity discount model we derive in this paper suffers very little from the second order effect of the utility function. All the model requires is that investors are risk adverse.

Under quadratic utility, CAPM holds for all securities. Hence, equation (17.11) holds for an h that is not infinitesimally small as follows:

$$X(t) = \frac{1}{R(t,T)}(\mathbb{E}[X(T)] - \beta^{\$}\{\mathbb{E}[V(T)] - R(t,T)V(t)\}) \tag{17.13}$$

where $R(t,T) = e^{r(T-t)}$ and the dollar beta is computed, for any T, as:

$$\beta^{\$} = \frac{\mathbb{K}[X(T), V(T)]}{\mathbb{V}[V(T)]} \tag{17.14}$$

Equation (17.13) is the main result of our model. It states that under quadratic utility, all assets must follow CAPM in determining their values. What we shall demonstrate is that when liquidity discount is present, the value computed by equation (17.13) is less than the perfectly liquid price computed by the Black-Scholes model.

To derive our liquidity discount model, we shall show first that if the relationship between the economy (represented by the state variable V) and the liquid price (S) is linear, then the liquidity discount is nil. Then we show that liquidity discount can exist only if the relationship between the economy and the liquid price is convex.

Theorem 1 When the payoff is linear, then liquidity discount is nil.

In continuous time, there are only two states in every infinitesimal time step (as described in Duffie and Huang (1985)) and hence no liquidity discount can exist. This indicates that if trading is continuous then at each infinitesimal step the payoff is linear and as a result Theorem 1 holds. In other words, continuous trading breaks up a fixed time horizon into small infinitesimal time steps, each of which is a linear payoff and hence liquidity discount does not exist. In the next section, we shall demonstrate this property in a numerical example and Theorem 1 is explicitly demonstrated.

Theorem 2 in the following proves that if the payoff is not linear and is convex, then the equilibrium price is always less than the linear price which, by Theorem 1, is the risk-neutral price where continuous rebalancing is possible.

Although a general proof with any form of convexity is not available, a proof based upon the binomial model that is consistent with our formulation of (17.4) is provided. In particular, as any three points define the convexity, we use a two period binomial model for the proof. It can be referred that with more points (more periods) in the binomial model, the proof stays valid.

Theorem 2 If the payoff is not linear and convex, then $X^{\mathrm{cvx}}(t) < X^{\mathrm{lnr}}(t)$ where $X^{\mathrm{lnr}}(t)$ is defined in [Theorem 1] and identical to $S(t)$ which is the perfectly liquid price and $X^{\mathrm{cvx}}(t)$ is the same as $S^*(t)$ which is the liquidity-constrained price.

From Theorem 1, we know that $X_0^{\mathrm{lnr}} = S(t)$ as the liquid price. Here, $X_0^{\mathrm{cvx}} = S^*(t)$ represents the illiquid price. Hence, in summary, $S^*(t) < S(t)$ for all values of finite $R(0,2)$ and the theorem is proved.

Note that under linearity (between wealth and liquid price), there exists no liquidity discount, which is the same result as continuous trading. As a consequence, the price under linearity X_0^{lnr} is identical to the price under continuous trading $S(t)$. Similarly, when the relationship between wealth and liquidity price is convex, there exists liquidity discount. And the price X_0^{cvx} represents the price under liquidity squeeze, $S^*(t)$.

17.6 Some Analysis

In this section, we provide a numerical example to demonstrate the enormity of liquidity discount. The main model is equation (17.13). While the analysis in this section is based upon an arbitrarily chosen set of parameter values, the result holds in general. While equation (17.13) is closed-form, (13) needs to be computed numerically as:

$$\beta = \frac{\sum_{j=1}^{n} \begin{pmatrix} n \\ j \end{pmatrix} p^j (1-p)^{(n-j)} \{V_i(T) - \bar{V}(T)\}\{X_j(T) - \bar{X}(T)\}}{\sum_{j=1}^{n} \begin{pmatrix} n \\ j \end{pmatrix} p^j (1-p)^{(n-j)} [V_j(T) - \bar{V}(T)]^2} \tag{17.15}$$

where

$$\bar{V}(T) = \sum_{j=1}^{n} \binom{n}{j} p^j (1-p)^{(n-j)} V_i(T)$$
$$\bar{X}(T) = \sum_{j=1}^{n} \binom{n}{j} p^j (1-p)^{(n-j)} X_i(T)$$

are the means of the economic state variable (wealth) and the state contingent claim respectively. In order to demonstrate convergence to the Black-Scholes model in continuous trading, we adopt the binomial framework of Cox, Ross and Rubinstein (1979) with n periods. Given that the binomial model will converge to the Black-Scholes model as n gets large, we shall use a sufficiently large n to represent the limiting Black-Scholes case. In other words, we shall demonstrate that under continuous trading, there is no discount for illiquid trading.

For the sake of easy exposition, we set up the following base case for the binomial model:

Paremeters	
n	100
$T-t$	1 year
σ	0.5
μ	10%
r	5%
V	\$80

To carry out a numerical example, we need to have an explicit functional form for the relationship between the state variable and the state contingent claim that represents either the liquid price (linear payoff) and illiquid price (convex payoff), in the context of the Cox, Ingersoll, and Ross model (1985). As a convenience, we choose a put payoff for the task. The put payoff is convex and negatively monotonic in the underlying economy, and hence can serve the purpose well. We choose an arbitrary strike of 100 to characterize convexity. We shall note that higher is the strike, higher is the convexity.

We first compute the price of the perfect liquid contingent claim, i.e. $X(t) = S(t)$. Note that the liquid price is such that trading takes place continuously and is represented by the Black-Scholes model. In the binomial model where n = 100 the value of the put (representing the perfectly liquid asset price) is almost identical to the Black-Scholes price of \$25.85. We feel that for the sake of computational time, this is minor enough difference for us to demonstrate the value of liquidity discount. This provides us the comfort that n = 100 is a good enough proxy for continuous trading. In the rest of the paper, we shall use the binomial model with n = 100 as the benchmark to examine the properties of the liquidity discount model.

Next we turn to computing the illiquid price, i.e. $X(t) = S^*(t)$. We start our analysis with n = 1 where we demonstrate that in this case, the liquid price (which is computed by the binomial model based upon the risk-neutral probabilities) is identical to the illiquid price (which is computed by the CAPM based upon the physical probabilities). In other words, when rebalancing is permitted at every node in the binomial model, the illiquid price is identical to the liquid price.

In a one period model (i.e. $n = 1$), it is clear that the up and down movements in the binomial model are $u = e^{\sigma\sqrt{\Delta t}} = e^{0.5\times 1} = 1.6487$ and $d = 1/u = 0.6065$. The state variable lattice and the stock payoff are given below (left and right respectively):

State Variable *V*

	131.90
80	
	48.52

State Contingent Claim Payoff *S*

	0
28.07	
	51.48

The risk-neutral probabilities are $\hat{p} = \frac{\exp(r\Delta t)-d}{u-d} = \frac{\exp(5\%\times 1)-0.6065}{1.6487-0.6065} = 0.4267$ and $1-\hat{p} = 0.5733$. The liquid price is (call): $0.5733 \times 51.48 \div e^{-5\%} = 28.07$, as in the above binomial tree.

The physical probabilities are $p = \frac{\exp(\mu\Delta t)-d}{u-d} = \frac{\exp(10\%\times 1)-0.6065}{1.6487-0.6065} = 0.4785$ and $1-p = 0.5215$. Following (17.13), we arrive at the same exact price of \$28.07 where the expected level of the state variable is \$88.41; the dollar beta $\beta^{\$}$ is 0.6174; the expected value of the illiquid price is \$26.84; and the risk-free discount is 0.9512.

The fact that the price of the illiquid asset equals the price of the liquid asset suggests that the illiquid price is independent of the physical probability p (and also independent of μ). The reason is that in a single period binomial model, the option payoff is linear in the underlying asset and Theorem 1 applies. This result is extremely crucial in our model in that once we approach continuous trading/rebalancing, the binomial model suggests that the option price within a period is linear in the underlying asset and as a result, the illiquid price must equal the liquid price the boundary condition it must satisfy by definition.

As n becomes large, liquidity discount becomes large. As in the base case where n = 100, the discount is substantial. The binomial value is \$25.86 and the liquidity-constrained price, i.e. (17.13), is \$24.47 representing a 5% discount.

To demonstrate the binomial implementation of our model, we let i be the time index and j be the state index. At the end of the binomial lattice $i = n$. At each time i, the state j is labeled as $0 \leqslant j \leqslant i$. The level of the state variable, the

Volatility and u and d				
σ	0.8	0.6	0.4	0.2
u	1.083287	1.061837	1.040811	1.020201
d	0.923116	0.941765	0.960789	0.980199

price of the liquid security, and the price of the illiquid security for the economy are then labeled as V_{ij}, S_{ij}, and S^*_{ij} respectively. The risk-neutral probability and the physical probability are defined as usual and given above.

In the following, we present results when the risk preference represented by the Sharpe ratio ranges from 0 (risk free case) to 1.6. Rebalancing frequency is represented by k. And $k = n - 1$ represents perfect rebalancing or continuous trading (that is, for 100 periods, rebalancing 99 times in between is identical to continuous trading).

In the binomial model, the number of periods n must be divisible by one plus the rebalancing frequency, i.e. $k + 1$, to avoid unnecessary numerical errors. In the base case where $n = 100, k$ can be 0, 1, 3, 4, 9, ..., 99. Take $k = 3$ as an example, rebalancing is allowed at $i = 25$, 50, and 75. At each of these times, equation (17.13) is used to compute the illiquid price at every node at the given time. Specifically, at $i = 75$, equation (17.13) is used to compute values at the nodes that are represented by $j = 0 \sim 75$ (i.e. $S^*_{75,0}$ till $S^*_{75,75}$). Then at $i = 50$, equation (17.13) is again used to compute values where $j = 0 \sim 50$ (i.e. $S^*_{50,0}$ till $S^*_{50,50}$) using the prices from $S^*_{75,0}$ till $S^*_{75,75}$. This process repeats backwards until we reach todays price which is S^*_0.

We compute a number of liquid and illiquid prices under various scenarios. Unless otherwise mentioned, the values of the input variables are taken from the base case. Note that liquidity discount is more severe as investors are more risk averse. To measure the magnitude of risk aversion, we adopt the Sharpe ratio on the underlying state variable, which is excess return scaled by the volatility. We simulate various degrees of Sharpe ratio from 0 (risk free case) to 1.6 with the volatility scenarios from 0.2 to 0.8. At the risk-free rate of 5%, we obtain the required rate of return ($\mu = r + \lambda\sigma$ where λ is Sharpe ratio) from 5% (risk-free case) to 133% ($\lambda = 1.6$).

The results are summarized in Table 17.1. The top panel are binomial parameter values where $u = \exp(\sigma\sqrt{\Delta t})$ and $d = 1/u$ under $n = 100$. The middle panel contains different expected returns) under different volatility scenarios for each Sharpe ratio. The bottom panel presents physical probability values using the binomial formula $p = [\exp(\mu\Delta t) - d]/[u - d]$. Interestingly, each Sharpe ratio corresponds to a physical probability value roughly.

Combine the information of the physical probabilities and other input values,

μ (drift)				
0	0.05	0.05	0.05	0.05
0.4	0.37	0.29	0.21	0.13
0.8	0.69	0.53	0.37	0.21
1.2	1.01	0.77	0.53	0.29
1.6	1.33	1.01	0.69	0.37
p (physical probability)				
0	0.4831	0.4892	0.4963	0.5075
0.4	0.5032	0.5092	0.5163	0.5275
0.8	0.5232	0.5293	0.5363	0.5476
1.2	0.5434	0.5494	0.5564	0.5676
1.6	0.5636	0.5695	0.5765	0.5877

Table 17.1: Shapre Ratio and μ and p

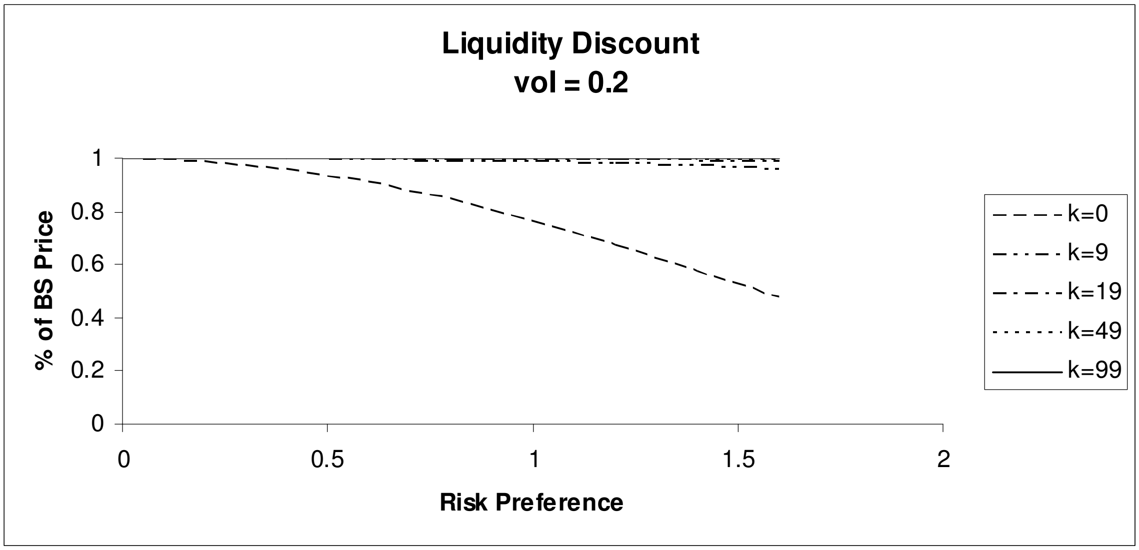
we can then compute liquid and illiquid prices using equation (17.13). Note that when the Sharpe ratio is 0, there is no liquidity discount and the liquid value equals the illiquid value, as proved by Theorem 1. This allows us to examine the magnitude of the liquidity discount as a function of risk preference. Table 2 provides all the liquid (Sharpe ratio is 0) and illiquid prices. The strike price for the result is set at 100 and the state variable is set at 80.

Reported in Table 17.2 are simulated liquid (Sharpe ratio is 0) and illiquid prices (Sharpe ratio is greater than 0). Note that by construction, our model for illiquid prices degenerates to liquid prices as the Sharpe ratio approaches 0. Also as required by the model, when continuous trading is reached ($k = 99$), we obtain liquid prices and risk preference does not matter (bottom panel). Liquidity discount is at maximum when no trading/rebalancing is allowed ($k = 0$).

To visualize the effect, we translate the values in Table 2 from dollar terms to percentage terms. In each case, the liquid price serves as the benchmark (named Black-Scholes value). This is the value consistent with continuous trading. Various comparisons are provided in Figure 17.4 and Figure 4 as follows.

σ (volatility)				
$k=0$	0.8	0.6	0.4	0.2
0	35.4303	29.0305	22.7066	16.9860
0.4	29.7751	25.1990	20.5670	16.2768
0.8	21.9833	19.2094	16.5282	14.3871
1.2	14.1540	12.734	11.6111	11.4428
1.6	7.8894	7.2781	7.0476	8.0259
$k=9$	0.8	0.6	0.4	0.2
0	35.4303	29.0305	22.7066	16.9860
0.4	34.8789	28.6789	22.5240	16.9334
0.8	33.9291	28.0265	22.1478	16.8068
1.2	32.5311	27.0385	21.5570	16.5993
1.6	30.6061	25.6581	20.7160	16.2998
$k=19$	0.8	0.6	0.4	0.2
0	35.4303	29.0305	22.7066	16.9860
0.4	35.1880	28.8759	22.6264	16.9630
0.8	34.7790	28.5946	22.4642	16.9087
1.2	34.1950	28.1806	22.2166	16.8222
1.6	33.4223	27.6247	21.8784	16.7021
$k=49$	0.8	0.6	0.4	0.2
0	35.4303	29.0305	22.7066	16.9860
0.4	35.3702	28.9921	22.6867	16.9803
0.8	35.2700	28.9231	22.6469	16.9670
1.2	35.1297	28.8233	22.5871	16.9461
1.6	34.9489	28.6925	22.5073	16.9176
$k=99$	0.8	0.6	0.4	0.2
0	35.4303	29.0305	22.7066	16.9860
0.4	35.4303	29.0305	22.7066	16.9860
0.8	35.4303	29.0305	22.7066	16.9860
1.2	35.4303	29.0305	22.7066	16.9860
1.6	35.4303	29.0305	22.7066	16.9860

Table 17.2: Liquidity Discount Results



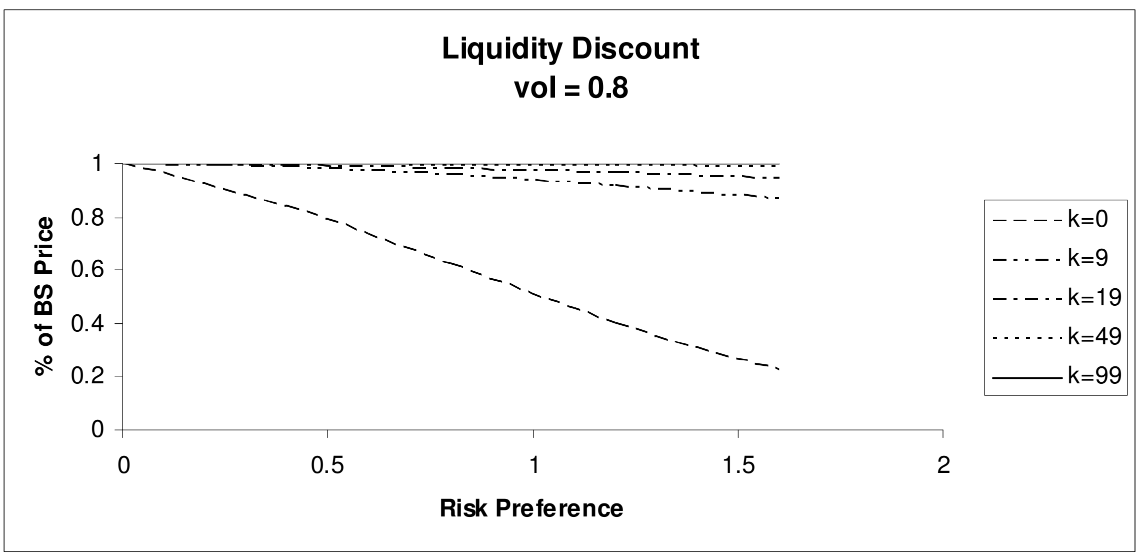


Figure 17.4: A Liquidity Discount Under Various Rebalancing Frequencies

In the upper panel of Figure 17.4, we present the result of liquidity discount under various trading frequencies. The binomial model of $n = 100$ is used as the benchmark and regard as the perfectly liquid price which, in the limiting case, converges to the Black-Scholes model. As a result, in the perfectly liquid case where $n = 100$, the number of rebalancing times is $k = 99$. Under no rebalancing, $k = 0$

and this represents the case of extreme illiquidity where investors hold their securities to maturity. Panel A of Figure 17.4 plots the result using $\sigma = 0.2$. The horizontal axis is the expected rate of return of the stock, used to represent risk preference.

It is clear that as the model allows for continuous rebalancing (represented by $k = 99$ in the case of $n = 100$), the price should be the same as the Black-Scholes price where continuous rebalancing is part of the assumption. In Figure 17.4, we do see that the price ratio (of equilibrium over Black-Scholes) is 1 through out the whole range of risk preference. When $k = 0$, the discount because of illiquidity can be severe. We can see from Figure 17.4 that the discount is as bad as 40% as the Sharpe ratio reaches 1.6 and no rebalancing is permitted.

The lower panel of Figure 17.4 is similar to Panel A with a higher volatility value (0.8). As we can see, the liquidity discount is more severe as the volatility is higher. We recall that during the 2007-8 crisis, the volatility was high. For example, in the case of Lehman (see Chen, Chiddi, Imerman, and Soprazetti (2010)) the volatility in many months of 2008 exceeded 100%.

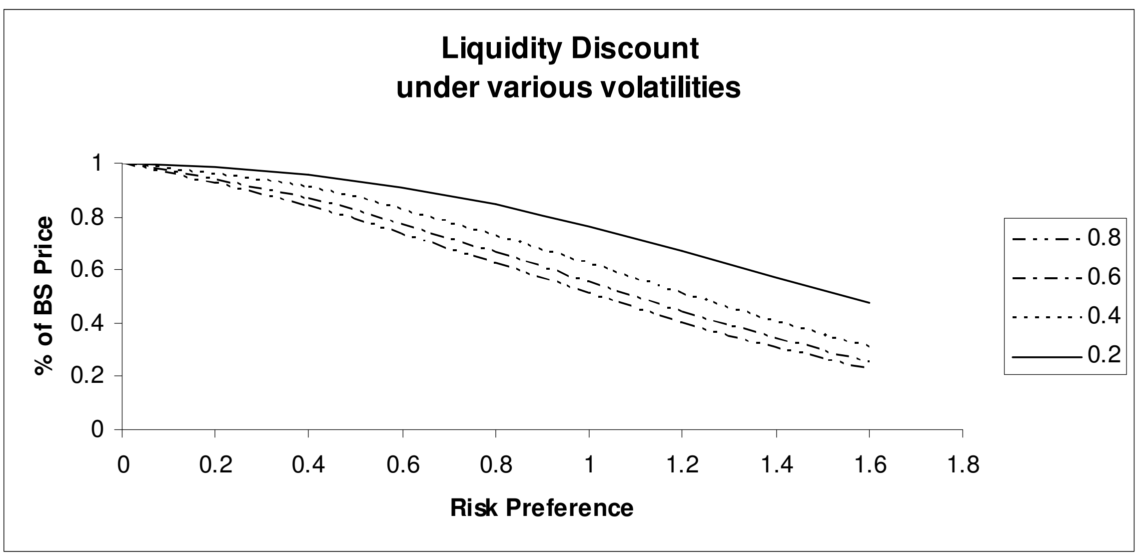


Figure 17.5: A Liquidity Discount Under Various Volatility Levels

Figure 17.5 presents the same result as Figure 17.4 but examines how various volatility levels affect the liquidity discount. Figure 17.5 sets k to be 0 for the maximum amount of liquidity discount. As we can see from the diagram, the deterioration of the asset price is rather fast. As the volatility is higher, the deterioration is faster.

The next necessary step is to test the model against liquid prices. In other

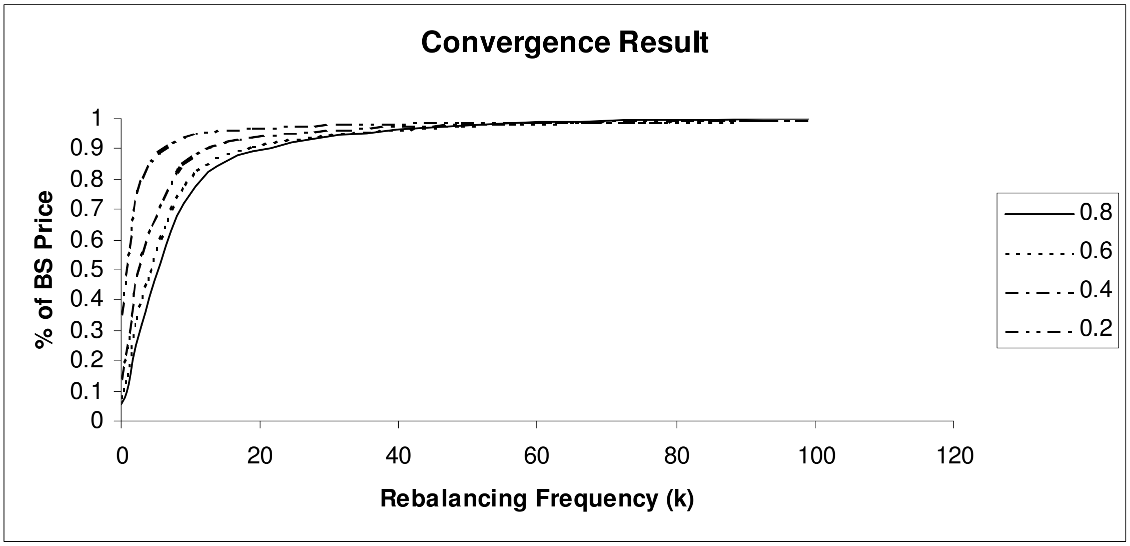


Figure 17.6: A Liquidity Discount Under Various Rebalancing Frequencies

words, when we permit perfect liquidity, i.e. $k = n - 1$, At this situation the equilibrium price must equal to the Black-Scholes (or binomial) price. Figure 17.6 presents the result of convergence under various volatility levels. As we can see, convergence is faster when the volatility is smaller.

In Figure 17.7, we provide the result of our model on the relationship between liquid and illiquid prices. This is the main result of our model which describes the illiquid price (S^*) as a function of the liquid price S. Figure 17.7 is similar to Figure 17.3 but presented with our model. The physical probability is set at 0.6 to exaggerate the result for the visual presentation.

We see in Figure 17.7 that as the liquid price decreases, the illiquid price decreases but at a much faster rate. This is consistent with the description in Figure 17.3 where B-D line bends over to touch the vertical axis.

When the liquid price is high, liquidity discount is small and the two prices are equal to each other. In the numerical example plotted in Figure 17.7, toward the right where the prices are both high, liquidity discount disappears and the curve approaches the 45-degree line asymptotically.

Note that in our model, there is no explicit put option as in Figure 17.3. Our liquidity discount model is derived by limiting trading/rebalancing in a binomial model. Our liquidity discount is computed by assuming a quadratic utility function so that pricing can be achieved via the Capital Asset Pricing Model. Nevertheless,

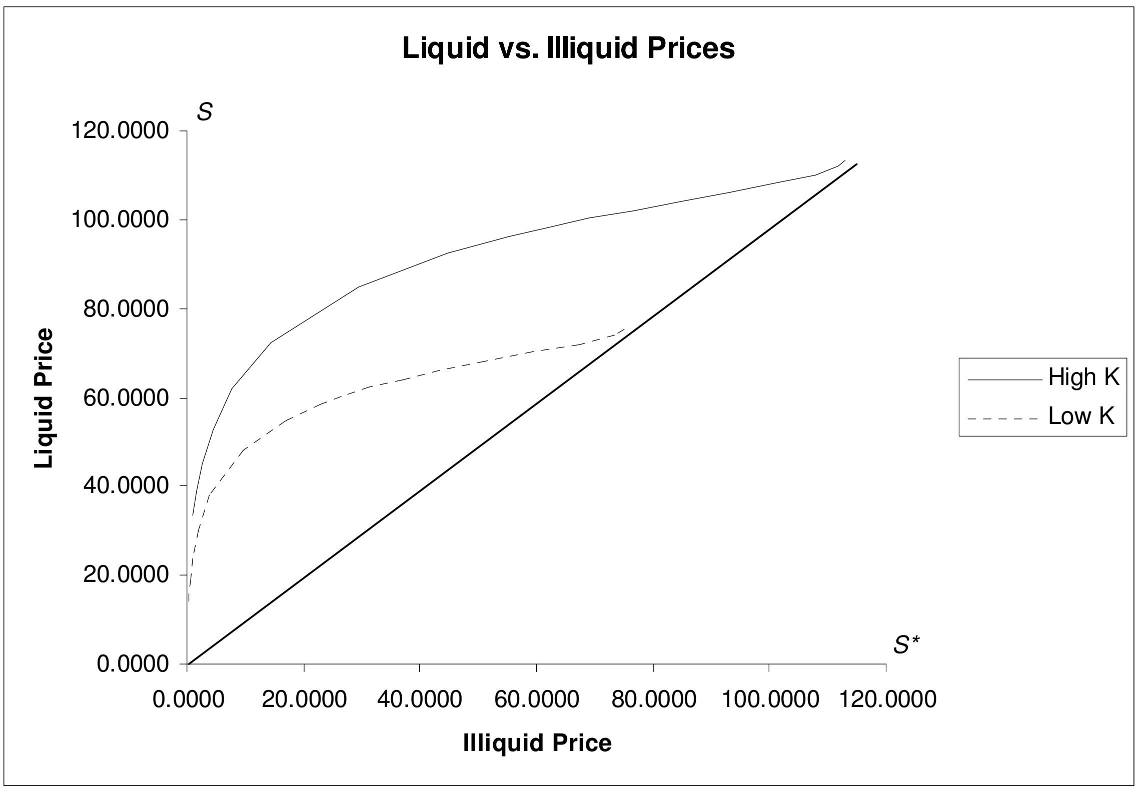


Figure 17.7: Relationship between Liquid and Illiquid Prices

the result of our model describes liquidity discount as a put option.

We also note that the liquidity discount model derived in this paper is closely connected how the relationship between the economy (represented by a single state variable: wealth) and the perfectly liquid price. As a result, the convexity of the relationship determines the severity of the liquidity discount.

17.7 Liquidity Premium

One immediate extension to our model is that we can analyze assets with liquidity premiums. While there are a number of financial assets that suffer from liquidity discount (i.e. lower prices due to limited trading), other assets, not necessarily financial, enjoy liquidity premium. Gold, oil, and even real estate are good examples of assets of such kind.

Symmetrical to the cause of liquidity discount, the cause of liquidity premium is the limited capacity of supply. Parallel to the model of liquidity discount, when an agent is risk adverse and the payoff of a security is concave, limited supply shall cause such an asset to enjoy liquidity premium. The analysis is straightforward as follows.

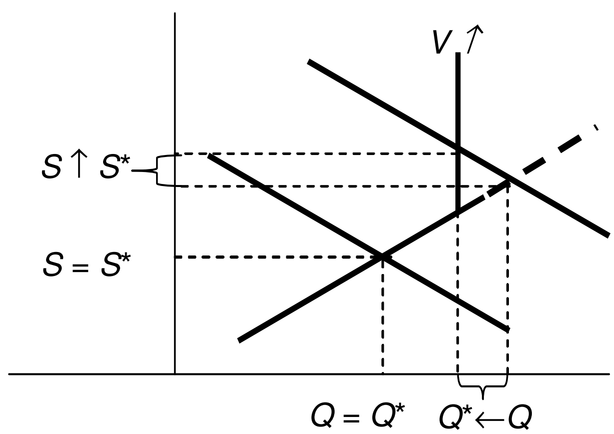


Figure 17.8: Demand-Supply Analysis of Liquidity Premium

Similar to Figure 17.2, Figure 17.8 depicts a situation where demand is more sensitive to economic changes than supply is and supply is bounded by a fixed quantity Q^*. As economy grows, it approaches the maximum capacity and liquidity squeeze (supply-driven) takes place. Contrary to the demand-driven squeeze, now the price under liquidity squeeze is higher with squeeze than without squeeze. In this situation, equilibrium price rises but quantity falls.

A counterpart of Figure 17.3 is shown in Figure 17.9. The left panel of Figure 17.9 describes the relationship between the economy V and perfectly liquid price S. The right panel describes the relationship between liquid price S and illiquid price S^*.

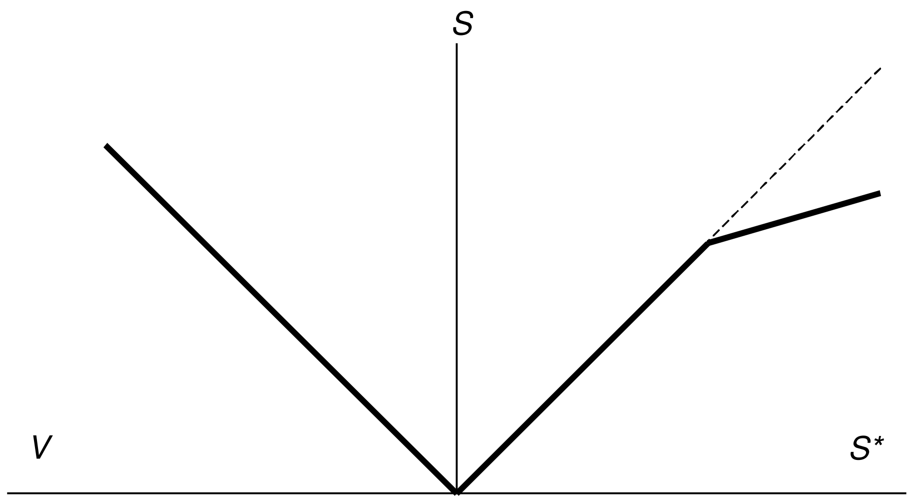


Figure 17.9: Liquidity Premium as a Call

From Figure 17.9, it is clear that liquidity premium can be described as a call option:

$$S^* = S + \text{call} \tag{17.16}$$

As in the liquidity discount case, in our model, there is no need to explicitly model the call option. As long as the liquid price is concave in the state variable of economy, the call-option-like result will be naturally derived.

The following theorem shows that in such a case, the liquidity-squeezed price S^*, is higher than the perfectly liquid price S when the liquid price is a concave function of the economy.

Theorem 3 If the payoff concave, then $S^{\text{lnr}}(t) < S^{\text{cav}}(t)$ where $S^{\text{lnr}}(t)$ is defined in Theorem 1 as $S(t)$ which is the perfectly liquid price and $S^{\text{cav}}(t)$ is the same as $S^*(t)$ which is the liquidity-constrained price.

Proof Repeat the same procedure of the proof of Theorem 2 and the result follows.

□

The model for liquidity premium is a result that represents the case where supply of the asset is less sensitive than demand to the underlying economy. This is useful in explaining prices of several commodities in the current situation such as gold and oil. These commodities are assets with very limited supply. As the demand of such assets grow stronger, prices rise dis-proportionally to the rest of the economy, resulting in liquidity premiums. Our model argues that if the liquid price is linear in the state variable, then the price of the asset will rise, but there is no liquidity premium. If the liquid price is concave in the state variable, then there is liquidity premium. Similar to liquidity discount, such a liquidity premium can be substantial even the fundamental economy does not change materially.

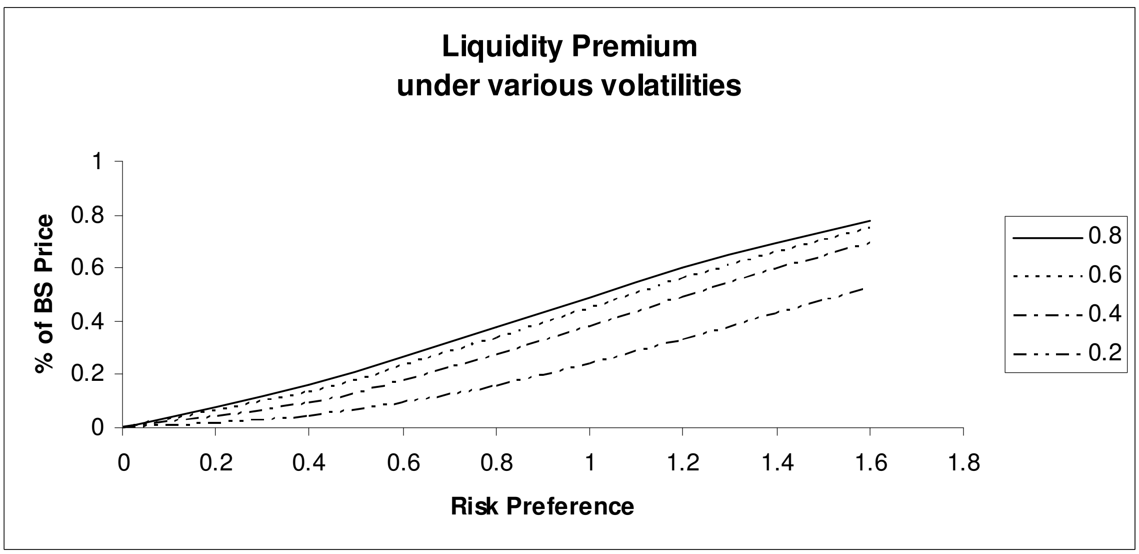


Figure 17.10: A Liquidity Premium Under Various Volatility Levels

A counterpart of Figure 17.5 is plotted in Figure 17.10. As the expected return becomes lower, the impact of liquidity premium is more profound. This diagram is generated with the same inputs as the base case with the relationship between the liquid price and the economic state variable as $S = \min\{V, K\}$.

Remark

There are a number of applications of the liquidity premium model. Any asset that is under supply squeeze will enjoy a liquidity premium. Obvious recent examples include oil, precious metals like gold, silver and platinum, and agriculture products. In a squeeze situation, these commodity prices deviate from their “fundamental” values and highly inflated prices. As in the analysis for the liquidity discount cases, these situations disappear once the liquidity pressure disappears.

Chapter 18

Funding Value Adjustment

18.1 FVA in a Netshell

18.1.1 What is FVA?

Fair value of a derivative portfolio is related to the discounting of the derivative portfolio which depends on

- Counterparty default – CVA
- Own default – DVA
- The collateral posted for this transaction – FVA

FVA is Funding Valuation Adjustment, which is similar concept to CVA but reflects the market value of the cost to fund a derivative instrument. FVA can be positive or negative depending on whether there is a net funding cost or benefit.

Collateral posted to a bank has 2 main benefits:

- Mitigates counterparty credit risk (CVA)
- Reduce funding requirement (FVA)

18.1.2 FVA for Collateralized Trades

For uncollateralized trades, any future positive cash flow is equivalent to investors are purchasing a bond issued by the counterparty, hence its value should simply be given by

$$TV = \max\{Z, 0\}e^{-(r+s_c)(T-t)}$$

For uncollateralized trades, any future negative cash flow is equivalent to investors are issuing a bond to the counterparty, hence its value should simply be given by

$$TV = \min\{Z, 0\}e^{-(r+s_u)(T-t)}$$

When netting is allowed, then

$$\begin{aligned}
&TV \\
&= \max\{Z, 0\}e^{-(r+s_c)(T-t)} + \min\{Z, 0\}e^{-(r+s_u)(T-t)} \\
&= Ze^{-r(T-t)} - \max\{Z, 0\}e^{-r(T-t)}\left(1 - e^{-s_c(T-t)}\right) + \min\{Z, 0\}e^{-r(T-t)}\left(1 - e^{-s_u(T-t)}\right) \\
&= RV - CVA + DVA - FVA + e
\end{aligned}$$

where

$$FVA = Ze^{-r(T-t)}\left(1 - e^{-b(T-t)}\right)$$

and b is cash-synthetic basis (assumed to be same for both counterparty and investor).

In general, FVA can be approximated through:

$$\begin{aligned}
CVA &= \int V_C(t)P_c(t)c_c(t)dt \\
DVA &= \int V_D P_u(t)c_u(t)dt \\
FVA &= \int V_F b(t)dt
\end{aligned}$$

where

$$\begin{aligned}
V_C(t) &= N_0 E\left[\frac{\max\{V_t - C_t, 0\}}{N_t}\right] \\
V_D(t) &= N_0 E\left[\frac{\min\{V_t - C_t, 0\}}{N_t}\right] \\
V_F(t) &= N_0 E\left[\frac{V_t - C_t}{N_t}\right]
\end{aligned}$$

18.1.3 FVA for Collateralized Trades

For collateralized trades, the formula remains the same, but with collateral put into consideration for $V_C(t)$ EEPV, $V_D(t)$ RevEEPV and $V_F(t)$ MEPV:

$$TV = RV - CVA + DVA - FVA + e$$

where

$$FVA = \int V_F(t)b(t)dt - \int V_F^*(t)\alpha(t)dt$$
$$V_F^*(t) = N_0 E\left[\frac{C_t}{N_t}\right]$$

where α is collateral basis (collateral difference between collateral investor posted and collateral counterparty posted)

For fully collateralized trades ($V = C$), the fair value would reduced to

$$TV = RV + \int V_F^*(t)\alpha(t)dt$$

One-way-in CSA

$$TV = RV + DVA + \int V_D(t)b(t)dt + \int V_C(t)\alpha(t)dt > RV$$

One-way-out CSA

$$TV = RV - CVA - \int V_C(t)b(t)dt - \int V_D(t)\alpha(t)dt < RV$$

18.1.4 Conclusion

CVA and counterparty risk is a challenging hybrid. This is probably the most complex instrument we have ever priced!

CVA and counterparty risk is an enormous challenge, it has the most combined and extended modeling challenge

- Compared to Market risk, MTM, and traditional credit risk
- Wrong-way risk

CVA can be evaluated through CCDS

CVA can be mitigated through collateralization, netting, CSA, break clause etc.

CVA capital charge is the main focus of latest Basel regulatory capital requirement

FVA is the latest development of fair value for derivatives to consider funding cost besides credit risk

18.2 Modeling Risky Funding

Risky funding has become a significant problem since the 2007-08 financial crisis. Not only has it posed a challenge on the long-time Wall Street golden rule of "law of one price" but also has brought out an important academic problem in economic costs of capital.

Various funding costs in different banks pose a significant problem in deciding what the true cost of a financial security is. If a bank is asked to incorporate its own funding costs into pricing the securities it buys and sells, then these prices will likely deviate from its competitors' and the market prices cannot be determined without a clearing (market microstructure) mechanism and a general equilibrium theory. In a classical microeconomic theory, in equilibrium, those banks who benefit from superior funding costs will enjoy consumer (buyer) or production (seller) surplus.

The justification of adopting funding costs can come rightfully from the consideration of the bank's costs of capital. Banks need to allocate its cost of equity and cost of debt (two broad types of cost of funding) across its assets in order to measure profitability and if an asset can generate a positive net present value. Without consideration of cost of funding, banks may purchase assets that are not profitable even though the "prices are right".

While these problems are profoundly important, in this article we simply address a small issue which is how various funding costs can be used in pricing. In particular, we extend the Morini–Prampolini model [12] when the funding cost is explicitly considered.[1]

To our knowledge, to date, the literature has assumed exogenous funding costs. This "reduced-form" approach is convenient to derive models that can be easily implemented. More importantly, such an approach is consistent with the current

[1]We note that in an equilibrium framework, funding costs must be endogenously determined and cannot be exogeneously given. Funding costs must be a function of credit and liquidity risks that are in turn results of various capital structures.

curve methodology already widely adopted by the industry and hence adoption of such as approach seems natural.[2]

Morini and Prampolini [12] argue that in order to properly model risky funding, each trade should be decomposed into two legs – a funding leg and a deal leg. We follow the same strategy in this paper. Yet we want to price general derivative contracts. In order to do so, we first evaluate the simple bullet loan assuming a specific close-out convention that is common in the OTC market. We derive a number of closed-form solutions for bullet loans under some simplifying assumptions and semi-closed-form solutions for the general derivative contracts. In doing so, we pave the way to evaluating generic derivative contracts. Furthermore, we propose a general equilibrium framework in which funding costs are incorporated into pricing endogenously.

18.3 Notation and basic layout

In general we consider a derivative contract between two counterparties B and L with maturity T. We let $\mathrm{CF}_X(s)$ denote the cash-flow, possibly stochastic, where $X \in \{B, L\}$ received at time s.

Let r be the spot interest rate, and define the risk-free discount factor D as[3]

$$D(t,s) = \exp\left(-\int_t^s r(u)\mathrm{d}u\right). \tag{18.1}$$

If $t \leq t' \leq t''$, then we denote by $V_X(t;t',t'')$ the risk-free value at time t of all cash-flows between time t' and t'' from X's perspective as agreed in the contract, i.e.

$$V_X(t;t',t'') = \int_{t'}^{t''} \mathbb{E}_t\left(D(t,s)\mathrm{CF}_X(s)\right)\,\mathrm{d}s, \tag{18.2}$$

where $\mathbb{E}_t$ denotes the expectation under the risk-neutral measure conditional on all the information available at time t.

Finally we denote by τ_X the default time of X and we make the following assumptions:

[2]The literature includes, for example, Piterbarg [13], Fries [6], and Burgard and Kjaer [2] and [3].

[3]The spot interest rate can be stochastic and the main results will remain the same.

Following the Jarrow-Turnbull model [9] where the recovery rate R_X is assumed to be fixed[4] (so-called "recovery of face"), we assume that if X defaults, then a fraction R_X of the risk-free value of the contract at time τ_X can be recovered by X's counterparty, i.e. the recovered amount at time τ_X is equal to $-\mathrm{R}_X V_X(\tau_X; \tau_X, T)$, provided $V_X(\tau_X; \tau_X, T)$ is negative. Following the ISDA Close-Out Protocol [8] we assume that if X defaults during the duration of the contract and $V_X(\tau_X; \tau_X, T)$ is positive, then X's counterparty has to pay X the full risk-free value of the contract at time $t = \tau_X$. Combining this with assumption i) we have for instance from B's perspective that:

$$\mathrm{CF}_B(\tau_B) = \mathrm{R}_B V_B(\tau_B; \tau_B, T)^- + V_B(\tau_B; \tau_B, T)^+, \tag{18.3}$$
$$\mathrm{CF}_B(\tau_L) = \mathrm{R}_L V_B(\tau_L; \tau_L, T)^+ + V_B(\tau_L; \tau_L, T)^-, \tag{18.4}$$

where $x^+ = \max\{x, 0\}$ and $x^- = \min\{x, 0\}$. This is in agreement with the standard literature on counterparty credit risk, see for instance Gregory [7].

We define the hazard rate curve $\lambda_X(t)$ of X by

$$\mathbb{P}(\tau_X > t) = e^{-\lambda_X(t)t}, \tag{18.5}$$

where $\mathbb{P}$ denotes the risk-neutral measure. The hazard rate curve of X can be determined from X's CDS prices. For simplicity we also do not consider simultaneous default of B and L.

Also we assume that the funding spread s_X is constant. This assumption can be relaxed to include more complex funding structures. This however makes the valuation of the funding leg more complicated and could potentially lead to an optimization problem.

Finally, we assume that all the funding is done in the debt market. In Figure 18.1, we highlight the funding structure where both B and L fund their transactions through the "market" which follows a different convention than the deal leg between the two counterparties. In the funding leg, neither B nor L takes into account the default risk of the market.

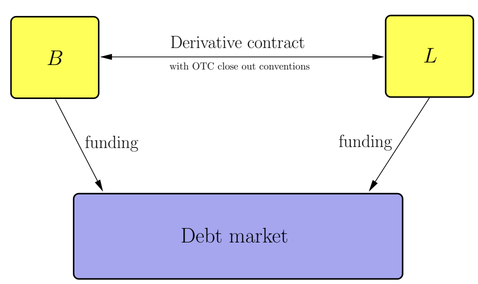


Figure 18.1: Funding in the debt market

[4] This is opposed to the Duffie-Singleton [5] assumption of the recovery to be proportional to the market value.

18.4 Valuation of bullet loans

In this section, we derive results for a bullet loan. While simple, the result contains all the intuition sufficient for understanding risky funding. Following Morini and Prampolini, we valuate the loan with two legs – funding leg and deal leg. However, in contrast, we provide valuation of a bullet loan as opposed to an issuance of a zero-coupon bond.[5]

In the contract L, to be called the Lender hereafter, lends B, to be called the Borrower hereafter, an amount equal to P at time $t = 0$ and B promises to repay L the amount K at time $t = T$, the maturity of the contract. Using the notation from the previous section we have from B's perspective that

$$\mathrm{CF}_B(s) = P\delta(s) - K\delta(s - T), \tag{18.6}$$

where $\delta(s)$ denotes the Dirac delta function. We find that

$$V_B(0; 0, T) = \int_0^T \big(D(0, s)P\delta(s) - D(0, s)K\delta(s - T)\big)\, \mathrm{d}s = P - D(0, T)K, \tag{18.7}$$

as expected.

In the next two subsections we will value the deal leg and funding leg of this loan contract in the spirit of Morini and Prampolini.

18.4.1 The deal leg of a bullet loan

The deal leg of a general derivative contract is by definition the risk-neutral expectation of all possible future cash-flows. Notice that in the case of a loan between two risky counterparties the cash-flows at time $t = 0$ are deterministic and all future cash-flows are stochastic, since they depend on the default times of both counterparties.

We now figure out all the possible future cash-flows in the loan contract between B and L from B's perspective. Clearly B receives an amount equal to P at time $t = 0$ and, when neither party defaults during the duration of the contract, then B pays L an amount equal to K at time $t = T$. The net-present value of these potential future cash-flows is equal to

$$P - \mathbb{E}\Big(D(0, T)K\mathbb{I}_{\{\tau_B > T, \tau_L > T\}}\Big), \tag{18.8}$$

[5]Note that the two are the same when the conventions of the deal and funding legs are the same.

where $\mathbb{I}_A$ represents the indicator of the event A.

It remains to consider the possible cash-flows when one of the parties defaults during the duration of the contract. Let us first consider the case when the Borrower B defaults during the duration of the contract and before L does. In this case, L retrieves a fraction R_B of the risk-free value of the contract at time $t = \tau_B$. The cash-flow at time $t = \tau_B$ from B's perspective is $-\mathrm{R}_B D(\tau_B, T)K$ and the net-present value of this potential future cash-flow is equal to

$$-\mathbb{E}\Big(\mathrm{R}_B D(0,T) K \mathbb{I}_{\{\tau_B<\tau_L, \tau_B \le T\}}\Big). \tag{18.9}$$

Finally we consider the case where L defaults during the duration of the contract and before B does. In this case the contract will be closed out at the risk-free value, i.e. B has to pay L an amount equal to $D(\tau_L, T)K$ at time $t = \tau_L$. The net-present value of this potential future cash-flow is equal to

$$-\mathbb{E}\Big(D(0,T) K \mathbb{I}_{\{\tau_L<\tau_B, \tau_L \le T\}}\Big). \tag{18.10}$$

Summing equations, we find that the value of the deal leg, $V_B^{\mathrm{deal}}(0)$, from B's perspective at time $t = 0$ is

$$\begin{aligned} V_B^{\mathrm{deal}}(0) = P &- \mathbb{E}\Big(D(0,T)K\mathbb{I}_{\{\tau_B>T, \tau_L>T\}}\Big) \\ &- \mathbb{E}\Big(\mathrm{R}_B D(0,T)K\mathbb{I}_{\{\tau_B<\tau_L, \tau_B\le T\}}\Big) - \mathbb{E}\Big(D(0,T)K\mathbb{I}_{\{\tau_L<\tau_B, \tau_L \le T\}}\Big). \end{aligned} \tag{18.11}$$

Note that there are six permutations for the order of T, τ_B, and τ_L, two of which represent the survival of both counterparties (first term), two of which represent the situations where Borrower defaults before maturity and before the Lender does (second term), and two of which represent the situations where Lender defaults before maturity and the Borrower does (last term).

If we denote by $V_L^{\mathrm{deal}}(0)$ the value of the deal leg from L's perspective, then it readily follows from the definition of the deal leg that $V_L^{\mathrm{deal}}(0) = -V_B^{\mathrm{deal}}(0)$ and the value of L's deal leg directly follows from above equation.

18.4.2 The funding leg of a bullet loan

We now discuss the valuation of the funding leg. The value of the funding leg will give us a measure of the funding advantages/disadvantages that B and L have from entering the loan contract. The underlying idea of the valuation of the funding leg is as in Morini and Prampolini [12].

Since the 2007-08 crisis, uncollateralized trades must pay a higher funding cost than those that are collateralized. In order to provide a positive NPV, any trade must be evaluated with its relevant cost of funding, which is a result of the use of capital – debt or equity.

The cost of debt or equity is a result of the profitability (i.e. market risk) and the capital structure (i.e. credit risk) of the firm. We assume, for the moment, that the credit risk of the firm is represented by the credit spread s_δ of its debts in aggregate with δ being the current value of all the debts.

Further let α and ϵ denote the total value of the firm and equity value of the firm. Then, by WACC (weighted average of cost of capital), we have:

$$c_\alpha = \frac{\delta}{\delta+\epsilon} c_\delta + \frac{\epsilon}{\delta+\epsilon} c_\epsilon \tag{18.12}$$

where c_ϵ is cost of equity, $c_\delta = r + s_\delta$ is cost of debt, and c_α is cost of asset. As we can see, the cost of funding (either c_δ or c_ϵ) is closely tied to the credit risk of the firm and must be determined endogenously. In a structural modeling framework, cost of equity c_ϵ and cost of debt c_δ are jointly modeled (as equity is a call option and debt is a covered call (short put)) and both are closely tied to the credit risk of the firm. We note that this cost of funding is also closely tied to the deal leg, via credit risk.

In a real situation, a bank buys a number of assets in wide varieties. The model for each transaction can certainly be quite complex. This is the reason why banks adopt reduced-form approaches to obtain quick solutions.

Equation (18.12) also indicates that if liquidity must be priced into cost of funding, it must go through $\delta, \epsilon, c_\epsilon$, and s_δ. In general, we believe that equity, ϵ, is relatively liquid and hence ϵ and c_ϵ should contain little liquidity impact. In other words, the c_ϵ of an illiquid firm should be the same as that of a liquid firm, after the adjustment of the market risk of course. In the simplest Merton argument [10], investors are indifferent if two stocks yield the same risk-adjusted excess return in a perfectly liquid stock market despite that one firm's assets can be more liquid than the other's.[6] As a result, bond δ and its spread s_δ should carry the weight of the liquidity impact. In an equilibrium setting, equity investors correctly price in extra discount due to liquidity in assets and migrate the impact over to debts.

[6] We can write the equation as $(c_\epsilon^{(i)} - r)/\sigma_\epsilon^{(i)} = (c_\epsilon^{(j)} - r)/\sigma_\epsilon^{(j)}$ where i and j represent two different stocks, σ_ϵ represents the volatility of the stock, and r represent the risk-free rate which is assumed constant for the sake of simplicity.

From equation (18.12), we know that no transaction can be completely funded by debt. This is so because then $s_\delta \to \infty$ as the credit risk of the firm goes to infinity. Unfortunately reduced form approaches assume such an assumption and let $c_\alpha = c_\delta = r + s_\delta$. The implication behind this is that some other deals are funded by the more expensive equity. Hence, we can see that reduced form models used today violate equilibrium.

While a more complete result will be derived (see [1]) where all risks are properly and endogenously evaluated, in this paper, we assume s_L (Lender spread) and s_B (Borrower spread) to be exogenously given and constant.

We begin with the valuation of B's funding leg, and let us assume for the moment that L is risk-free. By receiving P at time $t = 0$, B has a funding advantage since it allows him to reduce his funding by an amount equal to P. In fact, if B would not enter the loan contract with L and receive the amount P at time $t = 0$, he would have to issue a bond with principal $e^{(r+s_B)T}P$ at time $t = 0$. By issuing such a bond, B would receive a premium P at time $t = 0$. If B has not defaulted before maturity, he would have to repay the principal to the market at time $t = T$.

We can now quantify the funding advantage that B has from entering the loan contract by

$$-P + \mathbb{E}\Big(e^{s_B T} P \mathbb{I}_{\{\tau_B > T\}}\Big) + \mathbb{E}\Big(e^{s_B \tau_B} P \mathbb{I}_{\{\tau_B \le T\}}\Big). \tag{18.13}$$

So far we have assumed for simplicity that L is risk-free. However, B can enjoy its funding advantage only if L stays alive, and therefore the valuation of B's funding leg must depend on B's credit exposure to L. If L defaults during the duration of the loan contract and before B does, then B has to close out the loan contract with L and has to repay the risk-free value of the contract. Therefore B has only a funding advantage until time $t = \tau_L$. The cash-flow in the case where L defaults during the duration of the loan contract and before B does is equal to

$$\mathbb{E}\Big(e^{s_B \tau_L} P \mathbb{I}_{\{\tau_L < \tau_B, \tau_L \le T\}}\Big) \tag{18.14}$$

and therefore it follows that the value of B's funding leg is

$$\begin{aligned} V_B^{\text{fund}}(0) = &- P + \mathbb{E}\Big(e^{s_B T} P \mathbb{I}_{\{\tau_B > T, \tau_L > T\}}\Big) + \mathbb{E}\Big(e^{s_B \tau_B} P \mathbb{I}_{\{\tau_B < \tau_L, \tau_B \le T\}}\Big) \\ &+ \mathbb{E}\Big(e^{s_B \tau_L} P \mathbb{I}_{\{\tau_L < \tau_B, \tau_L \le T\}}\Big). \end{aligned} \tag{18.15}$$

Next we consider the value of the funding leg from L's perspective. As before let us first assume that the Borrower B is risk-free. When entering the loan contract with B, L has to raise an amount equal to P by issuing bonds at time $t = 0$ and has to repay an amount $e^{(r+s_L)T}P$ at time $t = T$ if he has not defaulted by that time. Therefore, when we assume that B is risk-free, the value of L's funding leg at time $t = 0$ is

$$V_L^{\text{fund}}(0) = P - \mathbb{E}\Big(e^{s_L T} P \mathbb{I}_{\{\tau_L > T\}}\Big) + \mathbb{E}\Big(e^{s_L \tau_L} P \mathbb{I}_{\{\tau_L \leq T\}}\Big). \tag{18.16}$$

Now we need to include the possibility of B defaulting during the duration of the loan contract. The same arguments as in the valuation of B's funding leg then imply that the value of L's funding leg is

$$\begin{aligned} V_L^{\text{fund}}(0) = & P - \mathbb{E}\Big(e^{s_L T} P \mathbb{I}_{\{\tau_B > T, \tau_L > T\}}\Big) - \mathbb{E}\Big(e^{s_L \tau_L} P \mathbb{I}_{\{\tau_L < \tau_B, \tau_L \leq T\}}\Big) \\ & - \mathbb{E}\Big(e^{s_L \tau_B} P \mathbb{I}_{\{\tau_B < \tau_L, \tau_B \leq \tau_L\}}\Big). \end{aligned} \tag{18.17}$$

In the next two sub-sections, we derive a closed-form solution under the independence assumption of the default times and a semi-closed-form solution under the correlated assumption.

18.4.3 An example when default times are independent

We define the value $V_X(0)$ of the loan contract to X at time $t = 0$ by

$$V_X(0) = V_X^{\text{deal}}(0) + V_X^{\text{fund}}(0). \tag{18.18}$$

For our first numerical example we assume that the default times of B and L are independent, and further we assume that r, λ_B and λ_L, s_B and s_L, and R_B are constant. In this case equation simplifies to

$$V_B^{\text{deal}}(0) = P - e^{-(r+\lambda_B+\lambda_L)T} K - e^{-rT} K \frac{\mathrm{R}_B \lambda_B + \lambda_L}{\lambda_B + \lambda_L} \left(1 - e^{-(\lambda_B+\lambda_L)T}\right). \tag{18.19}$$

In a similar way we can simplify equation and show that

$$V_B^{\text{fund}}(0) = P\left(1 - e^{-(\lambda_B+\lambda_L-s_B)T}\right) \frac{s_B}{\lambda_B + \lambda_L - s_B}. \tag{18.20}$$

Since $V_L^{\text{deal}}(0) = -V_B^{\text{deal}}(0)$ we further find from equation

$$V_L^{\text{deal}}(0) = -P + e^{-(r+\lambda_B+\lambda_L)T} K + e^{-rT} K \frac{\mathrm{R}_B \lambda_B + \lambda_L}{\lambda_B + \lambda_L} \left(1 - e^{-(\lambda_B+\lambda_L)T}\right), \tag{18.21}$$

and finally from equation we conclude that the value of the Lenders funding leg is

$$V_L^{\text{fund}}(0) = -P\left(1 - e^{-(\lambda_B+\lambda_L-s_L)T}\right)\frac{s_L}{\lambda_B+\lambda_L-s_L}. \tag{18.22}$$

Let us now consider an explicit numerical example. We consider a bullet loan with $T = 1$year, $K =$£100.00, and $r = 5\%$. We further assume that $\lambda_B = 5\%$ and $\lambda_L = 3\%$, $s_B = 10\%$ and $s_L = 7\%$, and that $\mathrm{R}_B = 40\%$.

Figure 18.2 plots equation (18.18) for both the Borrower and Lender as functions of the premium P (the amount paid by the Lender to the Borrower at time $t = 0$). Both are linear but with opposite slopes. The contract value for the Borrower V_B is a positive function of premium and that of the Lender function. The two lines cross at $P =$£85.12. This diagram indicates that it is possible for the two counterparties to successfully engage in a trade. In fact the premium can be set in such a way that $V_B(0), V_L(0) \geq 0$.

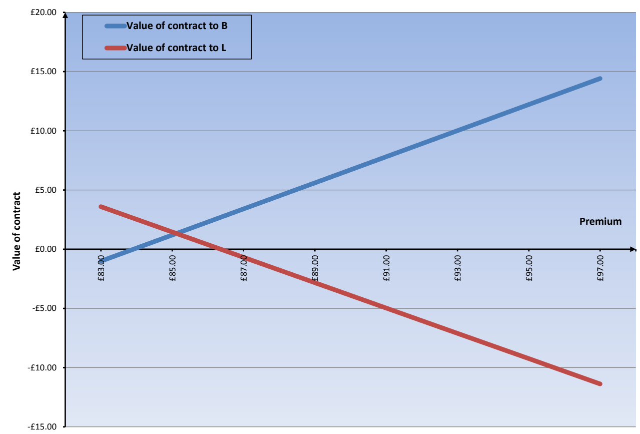


Figure 18.2: Contract value as a function of premium

This analysis indicates that the law of one price that has prevailed since the discovery of the Black-Scholes model must be adjusted. There must be a market clearing condition that decides what the market price (i.e. P) should be.

We notice that in a risk-free world, i.e. in a world without counterparty credit risk and funding costs at the risk-free rate, the breakeven premium is given by $P =$£95.12 and the NPV for both parties is £0. Looking at the figure we can see that pricing under the risk-free assumption leads to severe mispricing. More precisely, in this case the premium paid by L to B at time $t = 0$ is too large and the NPV for L is negative. The premium is not large enough to compensate L for the the risk that B defaults during the duration of the contract and the costs incurred on L from raising money in the market in order to be able to enter the contract.

In a similar way, if we consider a world with counterparty credit risk but with funding costs at the risk-free rate, then we find that the breakeven premium is given by $P =$£92.38. As before we can see that in our model, where funding costs differ from the risk-free rate, this premium leads to a negative NPV for L. It makes economic sense for B and L to enter the contract as long as it has a positive NPV. If the premium is sufficiently large then it makes sense for B to enter the deal, and if the premium is sufficiently small then it makes economic sense for L to enter the deal. We can see from the figure that there is a whole range of values for the premium P which causes the deal to have a positive NPV for B and L.

It is now interesting to think about the breakeven premium in a world in which information is freely available. In this case B will know the hazard and funding curve of L and L will know the hazard and funding curve of B. Both parties will agree to enter the contract at the breakeven premium $P =$£85.12, where the lines in the graph cross. Entering the contract at this premium will cause the contract to have the same positive NPV for both parties.

18.4.4 An example when default times are correlated

Again we assume that r, λ_B and λ_L, s_B and s_L, and R_B are constant. However, we now consider the case where τ_B and τ_L are correlated under the risk-neutral probability measure according to a Gaussian copula with correlation parameter $\rho \in (-1,1)$. More precisely, we assume that

$$\mathbb{P}(\tau_B \leq t, \tau_L \leq s) = C_\rho(\mathbb{P}(\tau_B \leq t), \mathbb{P}(\tau_L \leq s)), \tag{18.23}$$

where C_ρ is the Gaussian copula function with parameter ρ defined by

$$C_\rho(t,s) = \Phi_\rho\left(\Phi^{-1}(t), \Phi^{-1}(s)\right), \tag{18.24}$$

where Φ and Φ_ρ denote the normal distribution and the bivariate normal distribution with parameter ρ.

Using the Gaussian copula model and equation we find that

$$\begin{aligned} V_B^{\mathrm{deal}}(0) &= P - e^{-rT} K C_\rho\left(e^{-\lambda_B T}, e^{-\lambda_L T}\right) \\ &- \lambda_B \lambda_L e^{-rT} K \int_0^T \int_t^\infty c_\rho(1-e^{-\lambda_B s}, 1-e^{-\lambda_L t}) e^{-\lambda_B s} e^{-\lambda_L t} \mathrm{d}s\, \mathrm{d}t \\ &- \lambda_B \lambda_L \mathrm{R}_B e^{-rT} K \int_0^T \int_t^\infty c_\rho(1-e^{-\lambda_B t}, 1-e^{-\lambda_L s}) e^{-\lambda_B t} e^{-\lambda_L s} \mathrm{d}s\, \mathrm{d}t, \end{aligned} \tag{18.25}$$

where c_ρ denotes the density of C_ρ. Similarly we find from equation that

$$
\begin{aligned}
V_B^{\text{fund}}(0) &= -P + e^{s_B T} P C_\rho(e^{-\lambda_B T}, e^{-\lambda_L T}) \\
&\quad + \lambda_B \lambda_L P \int_0^T \int_t^\infty c_\rho(1 - e^{-\lambda_B t}, 1 - e^{-\lambda_L s}) e^{-(\lambda_B - s_B)t} e^{-\lambda_L s} \mathrm{d}s\, \mathrm{d}t \\
&\quad + \lambda_B \lambda_L P \int_0^T \int_t^\infty c_\rho(1 - e^{-\lambda_B s}, 1 - e^{-\lambda_L t}) e^{-\lambda_B s} e^{-(\lambda_L - s_B)t} \mathrm{d}s\, \mathrm{d}t.
\end{aligned}
\tag{18.26}
$$

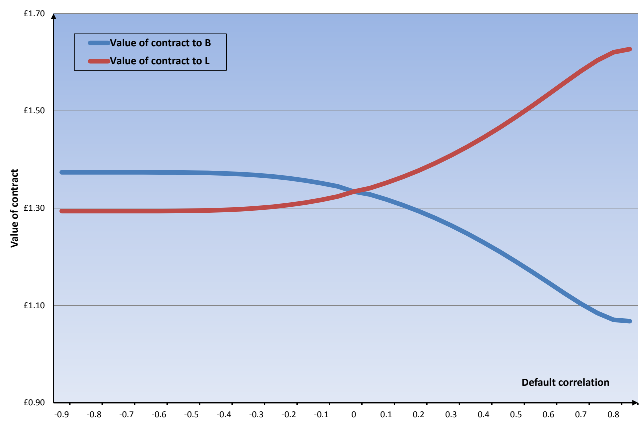


Figure 18.3: Contract value as a function of correlation

The value of L's deal leg and L's funding leg can be found in the same way. Similar to Figure 18.2, we plot the Borrower's and the Lender's contract values each as a function of the default correlation in Figure 18.3. Again we assume that $T = 1$year, $K =$£100, $r = 5\%$, $\lambda_B = 5\%$ and $\lambda_L = 3\%$, and that $s_B = 10\%$ and $s_L = 7\%$. Finally we assume that $\mathrm{R}_B = 40\%$. In the figure, the premium is set at the $P =$£85.12 which is the equilibrium value of the Borrower and the Lender shown in Figure 18.2.

18.5 Valuation of a general derivative contract

In this section we extend the result of the previous section and derive results for general derivative contracts. While we lose closed-form solution, the results remain semi-analytical and computations can be performed fairly efficiently.

18.5.1 Valuation of the deal leg

Recall that the value of the deal leg from X's perspective is, by definition, equal to the present value of all possible future cash-flows agreed in the contract. Therefore the value of the deal leg from X's perspective at time $t = 0$ is given by

$$\begin{aligned} V_X^{\text{deal}}(0) = &\int_0^T \mathbb{E}\Big(D(0,s)\mathrm{CF}_X(s)\mathbb{I}_{\{\tau_B>s,\tau_L>s\}}\Big)\mathrm{d}s \\ &+ \mathbb{E}\Big(D(0,\tau_B)\mathrm{CF}_X(\tau_B)\mathbb{I}_{\{\tau_B<\tau_L,\tau_B\le T\}}\Big) \\ &+ \mathbb{E}\Big(D(0,\tau_L)\mathrm{CF}_X(\tau_L)\mathbb{I}_{\{\tau_L<\tau_B,\tau_L\le T\}}\Big). \end{aligned} \tag{18.27}$$

A straightforward calculation shows that the above equation is equivalent to

$$V_X^{\text{deal}}(0) = V_X(0;0,T) - \mathrm{CVA}_X + \mathrm{DVA}_X, \tag{18.28}$$

where CVA_X is the credit value adjustment and DVA_X is the debit value adjustment from X's perspective. Our valuation of the deal leg therefore agrees with the standard literature on counterparty credit risk, see for instance Gregory [7].

18.5.2 Valuation of the funding leg

The valuation of the funding leg for the general derivative contract follows the same ideas that we used in the valuation of the funding leg for bullet loans in the above section (FundingLegBulletLoan).

After decomposing a general derivative contract into a series of expected future cash-flows we are only left with the valuation of the funding advantage/disadvantage of a forward starting bullet loan with a possibly stochastic premium.

Let us assume that today, at time $t = 0$, we expect to receive a cash-flow $\mathrm{CF}_X(t)$ at time t. Receiving this cash-flow is conditional on neither party defaulting by that time. In particular this cash-flow yields a funding advantage/disadvantage only if neither party has defaulted by time t. Moreover, when neither party has defaulted by time t, then we can understand this cash-flow simply as a bullet loan. Therefore, when we condition on B and L not defaulting by time t, then the funding advantage/disadvantage from receiving $\mathrm{CF}_X(t)$ is equivalent to the funding advantage/disadvantage from a bullet loan initiated at time t and with maturity T. Thus we can understand every expected future cash-flow as a forward starting bullet loan with a possibly stochastic premium. In particular, following the valuation procedure for bullet loan from section FundingLegBulletLoan, we find that the funding advantage/disadvantage from receiving the cash-flow $\mathrm{CF}_X(t)$ at time t has value today equal to

$$
\begin{aligned}
\mathbb{E}\Big(\Big(&e^{s_X(T-t)}D(0,t)\mathrm{CF}_X(t)\mathbb{I}_{\{\tau_B>T,\tau_L>T\}} \\
&- D(0,t)\mathrm{CF}_X(t) + e^{s_X(\tau_B-t)}D(0,t)\mathrm{CF}_X(t)\mathbb{I}_{\{\tau_B<\tau_L,\tau_B\le T\}} \\
&+ e^{s_X(\tau_L-t)}D(0,t)\mathrm{CF}_X(t)\mathbb{I}_{\{\tau_L<\tau_B,\tau_L\le T\}}\Big)\mathbb{I}_{\{\tau_B>t,\tau_L>t\}}\Big).
\end{aligned} \tag{18.29}
$$

Adding all the cash-flows together we therefore find that the value of the funding leg for the general derivative contract is given by

$$
\begin{aligned}
V_X^{\mathrm{fund}}(0) = \int_0^T \mathbb{E}\Big(\Big(&e^{s_X(T-t)}D(0,t)\mathrm{CF}_X(t)\mathbb{I}_{\{\tau_B>T,\tau_L>T\}} \\
&- D(0,t)\mathrm{CF}_X(t) + e^{s_X(\tau_B-t)}D(0,t)\mathrm{CF}_X(t)\mathbb{I}_{\{\tau_B<\tau_L,\tau_B\le T\}} \\
&+ e^{s_X(\tau_L-t)}D(0,t)\mathrm{CF}_X(t)\mathbb{I}_{\{\tau_L<\tau_B,\tau_L\le T\}}\Big)\mathbb{I}_{\{\tau_B>t,\tau_L>t\}}\Big)\,\mathrm{d}t.
\end{aligned} \tag{18.30}
$$

18.6 Liquidity

So far we have not considered liquidity. In Morini and Prampolini [12], a constant liquidity spread is added to discounting exogenously. We recognize that liquidity discount is often combined with credit (at least very high correlation). While it is convenient to include liquidity as an extra discounting factor exogenously, it could lead to severe mispricing. The literature on liquidity has argued that liquidity is an endogenous result of economic conditions.

To model liquidity risk along with credit risk (so that they can be endogenously correlated), a structural approach such as the Merton model (adopted by KMV) [11] or its extensions must be employed. Recently Chen [4] argued that liquidity discount can be viewed as a put option on the traded asset and as a result is directly linked to all the risks the traded asset inherits. In other words, both deal leg and funding leg are affected by the liquidity discount. Equation (18.12) explains why liquidity cannot be modeled separately from the credit risk.

A liquidity default is a situation where the firm has not enough assets to pay for its current cash-flow obligations. This is different from an economic default where the firm's assets become unproductive and ultimately are not enough to cover the firm's debt (not just cash-flow) obligations. Equity investors observe every fundamental factor of the firm and assign the right value to the equity. If a firm suffers from a liquidity squeeze (as in the recent crisis) then its assets are highly discounted so that it faces a possible liquidity-driven default. Then the equity value will properly reflect such an evaluation and then, liquidity risk dominates and is priced (in a form of spread). As a result, using a fixed spread (in particular, using a bond-CDS basis) is not accurate.

While developing a full model that incorporates all risks– market, credit and liquidity is beyond the scope of this paper, we shall lay out the framework of how these risks should interact and why that can severely dictate how a model should be designed. Taking equation (18.12) as an example. In a "normal" situation where securities are priced with perfect liquidity (we ignore, for the sake of exposition, normal/minor liquidity discounts that exist all the time), as described in Section 18.4.2, the equity value ϵ reflects a perfectly liquid value for the assets. The debt value δ which is just the difference of asset value and equity value then reflects only credit risk. In a dramatic situation (such as the recent crisis) where assets are discounted heavily due to liquidity concerns, the equity that is transacted with perfect liquidity, will properly reflect the lowered asset value. In such a situation, the debt value δ must reflect both liquidity and credit risks. As a result, s_δ embeds the liquidity risk.

Continue with equation (18.12) which is an accounting identity, we argue that under the liquidity squeeze, the same equation holds but with different values of debt and equity:

$$c^*_{\alpha^*} = \frac{\delta^*}{\delta^* + \epsilon^*} c^*_{\delta^*} + \frac{\epsilon^*}{\delta^* + \epsilon^*} c^*_{\epsilon^*} \tag{18.31}$$

where the asterisk superscripts represent values under the liquidity squeeze. Chen [4] demonstrates how to link α with α^* via a put option analogy. As $\alpha^* < \alpha$, $c^*_\alpha > c_\alpha$ and the difference is a liquidity spread $\ell_\alpha = c^*_\alpha - c_\alpha$. Similarly, $c^*_\delta = c_\delta + \ell_\delta$ and $c^*_\epsilon = c_\epsilon + \ell_\epsilon$. Recall that $c_\delta = r + s_\delta$.

For the sake of argument, we simplify the problem by making the assumption that $\ell_\alpha = \ell_\delta = \ell_\epsilon$ and furthermore $\delta^*/(\delta^* + \epsilon^*) = \delta/(\delta + \epsilon)$. Then equation (18.31) differs from equation (18.12) by the liquidity spread. As a result, the current industry practice of using the CDS-bond basis as a proxy for liquidity risk is an acceptable approximation, in that c_δ is represented by CDS and c^*_δ is represented by bond which is justified by equation (18.31).

Funding costs should reflect credit risk and liquidity risk. Note that liquidity risk can be driven by the general economy as well as market-specific (or product-specific) factors, and as a natural result market microstructure mechanism can be introduced and general equilibrium prices (market prices) can be identified.

However, while the idea of liquidity risk valuation is intuitive and straightforward, details need to be worked out. As argued in Chen [4], each case is different and hence it requires further research to bring about how liquidity risk can be explicitly incorporated in the various derivatives contracts.

18.7 Summary and Future Research

In this paper, we study risky funding. We price general derivative contracts in the context of risky funding. We first evaluate the simple bullet loan assuming a specific close-out convention where we arrive at a closed-form solution. We then derive semi-closed-form solutions for the general derivative contracts. Lastly, we propose a general equilibrium framework in which funding costs are incorporated into pricing endogenously.

Following Morini and Prampolini [12], we continue to assume an exogenous funding discount rate. Note that with funding costs explicitly considered in pricing, each bank will generate different valuation results. As mentioned earlier, this violates the law of one price. Market microstructure mechanisms and equilibrium theories must be introduced in order to arrive at market prices.

The confusion arises in the literature as whether or not risky funding should be part of evaluation in that risky funding often is used in conjunction with liquidity risk which is part of deal evaluation. While risky funding indeed should incorporate liquidity risk, such liquidity risk should be separated from the liquidity risk used in calculating prices of assets.

To reconcile different views in the literature and avoid the confusion of risky funding and evaluation, a fundamental pricing theory needs to be developed where liquidity risk, costs of capital, and valuation can be simultaneously assessed within a consistent framework, which is beyond the scope of this paper and requires further research.

18.8 Collateral Management

To discuss collateral management, we must first define rehypothecation. The practice by banks and brokers of using, for their own purposes, assets that have been posted as collateral by their clients. Clients who permit rehypothecation of their collateral may be compensated either through a lower cost of borrowing or a rebate on fees.

The recent crisis has changed how OTC trading is conducted. Prior to the crisis, most of the trades were naked, or uncollateralized. This means that the two parties that trade with each other (counterparties) trust each other. If one party defaults, then the other party must suffer from the consequences (i.e. LGD). As trades and trade parties became convoluted, the defaults of counterparties have become magnified. And that was the cause of the crisis. Basically the defaults of Bear Sterns and Lehman had spiraled. All banks had been affected at the end.

A collateralized trade is different. The out-of-money counterparty must post assets that equal (or at least a big percentage of) the value of the trade. For example, a 5-year, $10 million IRS is transacted between JP Morgan and Goldman Sachs. At inception, there is no value of this IRS (known as par). A fixed swap rate is signed between JP Morgan and Goldman Sachs (say 5%). In the next five years, JP Morgan will pay Goldman Sachs 1.25% of the notional (or $125,000) every quarter and Goldman Sachs will pay JP Morgan the LIBOR rate. Afterwards, the swap rate goes up or down. If it goes up, then JP Morgan makes profits out of this swap so JP Morgan is in-the-money and if Goldman Sachs defaults now, then JP Morgan would lose the money it makes. However, if Goldman Sachs posts a collateral that equals the value of the trade, then JP Morgan can liquidate the collateral and its profit is retained. Similarly, if the swap rate goes down, then JP Morgan would have to post a collateral for the value of the trade so if JP Morgan defaults, Goldman Sachs will not suffer any loss.

After the crisis, most trades are collateralized and the size of the collateralized assets grows. These assets that are used for collaterals, if no default, still belong to the banks that post them. Any interests or dividends still belong to the banks that post them. These assets are generally be managed by a third-party bank known as a custodian bank. This is similar to any stock investor who wants to do margin trades or short sales. Stocks in his or her account will be used as collaterals for such trades. These stocks are already with the broker under his or her account so it is very convenient. It is well-known that these stock brokers lend these securities out to make additional profits. Such an action is legally allowed as long as these brokers operate under the limits by the regulation.

Similarly, these custodian banks can do the same thing. They can loan these collaterals out via their securities lending business to make additional profits. Often these profits are shared with the banks who post them so it is a win-win strategy. As the amount the collaterals grows substantially after the crisis, the profits from rehypothecation become substantial.

Different from stocks held by stock brokers as collaterals, these assets are not liquid. As a result, lending them out for profits is tricky. That is, they cannot be easily bought back when a delivery due to default becomes necessary. As a result, managing these illiquid assets is difficult. The illiquid nature of these assets reduces the amount that can be rehypothencated. As a result, banks seek aggressively how they can evaluate and manage these assets more effectively.

There are two major tasks in managing collaterals:

- maintaining a good inventory of assets
- accurately liquidity pricing all assets

Note that billions of dollars of assets covering a huge variety are used for collaterals. These assets must be properly and accurately priced and categorized. This is as important as a valuation task as an IT task.

18.9 Appendix

18.9.1 Proof of Theorem 1

Note that if payoff is linear, $X(T) = aV(T) + b$, then the following results hold:

$$\begin{array}{ll} (i) & \mathbb{E}[X(T)] = a\mathbb{E}[V(T)] + b \\ (ii) & \mathbb{V}[X(T)] = a^2\mathbb{V}[V(T)] \\ (iii) & \mathbb{K}[V(T), X(T)] = a\mathbb{V}[V(T)] \\ (iv) & \beta^{\$} = a \end{array}$$

where the expectation, variance, and covariance operations are taken under the physical measure. As a result, following (17.13), we have:

$$X(t) = \frac{1}{R(t,T)} \{\mathbb{E}[X(T)] - a(\mathbb{E}[V(T)] - R(t,T)V(t)\} \tag{18.32}$$

where $R(t,T) = e^{r(T-t)}$ and

$$\mathbb{E}[X(T)] - R(t,T)X(t) = a\mathbb{E}[V(T)] - aR(t,T)V(t) \tag{18.33}$$

which gives rise to the following result:

$$X(t) = \frac{1}{R(t,T)}\{aR(t,T)V(t) + b\} \tag{18.34}$$

Note that this is exactly the result of risk neutral pricing (expected value with a "$\hat{\mathbb{E}}$"), i.e.,

$$\begin{aligned} X(t) &= \frac{1}{R(t,T)}\hat{\mathbb{E}}[aV(T) + b] \\ &= aV(t) + b\frac{1}{R(t,T)} \end{aligned} \tag{18.35}$$

This is also the result described by the Martingale Representation Theorem which indicates that any contingent claim under continuous trading can be replicated by the underlying asset and the risk-free asset. The result shown in (18.35) states that equation (17.13) computes also the liquid price, i.e. $X(t) = S(t)$ if the relationship between economy and the asset price is linear. □

18.9.2 Proof of Theorem 2

We use a two-period binomial model to demonstrate the proof. We let V_{ij} represent the level of the state variable at time period i and state j respectively. In a two-period binomial model, the three state variable values at time 2 are V_{20}, V_{21}, and V_{22} representing low, medium, and high prices respectively. Without loss of generality, we also let $V_{21} = V_0$. Similarly, we also let X_{ij} represent the state contingent claim price at time period i and state j respectively where X_{20}, X_{21}, and X_{22} represent low, medium, and high prices respectively.

We let the convex payoff differ from the linear one by slightly altering the middle one as follows: $X_{21}^{\text{cvx}} = X_{21}^{\text{lnr}} - \varepsilon$ where ε is an arbitrary small positive amount to create the convexity of X in V. The real probability per period is p which is time invariant. Consistent with the notation above, we symbolize compounding at the risk-free rate in two periods as $R(0,2)$. Our goal is to prove that:

$$\begin{aligned} X_0^{\text{cvx}} &= \frac{1}{R(0,2)}\left\{\mathbb{E}[X_{2\cdot}^{\text{cvx}}] - \beta^{\$}(\mathbb{E}[V_{2\cdot}] - R(0,2)V_0)\right\} \\ &< \frac{1}{R(0,2)}\left\{\mathbb{E}[X_{2\cdot}^{\text{lnr}}] - \beta^{\$}(\mathbb{E}[V_{2\cdot}] - R(0,2)V_0)\right\} \\ &= X_0^{\text{lnr}} \end{aligned} \tag{18.36}$$

where the symbol "2 in the subscript represents the three states in period 2, and cvx or lnr in the superscript represent convex or linear payoffs respectively. In the binomial model, the three physical probabilities are p, $2p(1-p)$, and $(1-p)^2$ for high, medium, and low states respectively. Hence, we have:

$$\begin{array}{ll} (i) & \mathbb{E}[V_{2\cdot}X_{2\cdot}^{\text{cvx}}] = \mathbb{E}[V_{2\cdot}X_{2\cdot}^{\text{lnr}}] - \varepsilon V_{21}2p(1-p) \\ (ii) & \mathbb{E}[X_{2\cdot}^{\text{cvx}}] = \mathbb{E}[X_{2\cdot}^{\text{lnr}}] - 2\varepsilon p(1-p) \\ (iii) & \mathbb{E}[V_{2\cdot}]\mathbb{E}[X_{2\cdot}^{\text{cvx}}] = \mathbb{E}[V_{2\cdot}]\mathbb{E}[X_{2\cdot}^{\text{lnr}}] - 2\mathbb{E}[V_{2\cdot}]\varepsilon p(1-p) \\ (iv) & \text{cov}[V_{2\cdot}, X_{2\cdot}^{\text{cvx}}] = \text{cov}[V_{2\cdot}, X_{2\cdot}^{\text{lnr}}] + 2\varepsilon p(1-p)(\mathbb{E}[V_{2\cdot}] - V_{21}) \end{array}$$

and then the dollar beta under the convex function can be derived as:

$$\beta^{\text{cvx}} = \beta^{\text{lnr}} + \frac{\varepsilon 2p(1-p)(\mathbb{E}[V_{2\cdot}] - V_{21})}{\text{var}[V_{2\cdot}]} \tag{18.37}$$

As a result,

$$\begin{aligned}&\beta^{\mathrm{cvx}}\{\mathbb{E}[V_{2\cdot}] - RV_0\} = \\ &\beta^{\mathrm{lnr}}\{\mathbb{E}[V_{2\cdot}] - R(0,2)V_0\} + \frac{2p(1-p)\varepsilon(\mathbb{E}[V_{2\cdot}] - V_{21})(\mathbb{E}[V_{2\cdot}] - R(0,2)V_0)}{\mathrm{var}[V_{2\cdot}]}\end{aligned} \tag{18.38}$$

and

$$\begin{aligned} X_0^{\mathrm{cvx}} &= \mathbb{E}[X_{2\cdot}^{\mathrm{cvx}}] - \beta^{\mathrm{cvx}}\{E[V_{2\cdot}] - R(0,2)V_0\} \\ &= \mathbb{E}[X_{2\cdot}^{\mathrm{lnr}}] - 2p(1-p)\varepsilon - \beta^{\mathrm{lnr}}\{\mathbb{E}[V_{2\cdot}] - R(0,2)V_0\} \\ &\qquad\qquad - \frac{2p(1-p)\varepsilon(\mathbb{E}[V_{2\cdot}] - V_{21})(\mathbb{E}[V_{2\cdot}] - R(0,2)S_0)}{\mathrm{var}[V_{2\cdot}]} \\ &= X_0^{\mathrm{lnr}} - 2p(1-p)\varepsilon\left\{1 + \frac{(\mathbb{E}[V_{2\cdot}] - V_{21})(\mathbb{E}[V_{2\cdot}] - R(0,2)V_0)}{\mathrm{var}[V_{2\cdot}]}\right\} \\ &< X_0^{\mathrm{lnr}} \end{aligned} \tag{18.39}$$

This is because, clearly,

$$\mathbb{E}[V_{2\cdot}] > V_{21} = V_0$$

(the expected payoff should be larger than todays value) and

$$\mathbb{E}[V_{2\cdot}] > R(0,2)V_0$$

(the expected return should be more than the risk-free rate) to avoid arbitrage. □

Bibliography

[1] T. Behrndt, R.-R. Chen, in preparation, 2012.

[2] C. Burgard and M. Kjaer, *In the Balance*, Risk, October 2011.

[3] C. Burgard and M. Kjaer, *PDE Representations of Options with Bilateral Counterparty Risk and Funding Costs*, Journal of Credit Risk, Volume 7, Number 3, 2011.

[4] R.-R. Chen, *Valuing Liquidity Discount*, Journal of Fixed Income, forthcoming, 2011.

[5] D. Duffie and K. Singleton, *Modeling Term Structures of Defaultable Bonds*, Review of Financial Studies, 12, 687–720, 1999.

[6] C. P. Fries, *Discounting Revisited: Valuation Under Funding, Counterparty Risk and Collateralization*, http://ssrn.com/abstract=1609587, August 2010.

[7] J. Gregory, *Being two faced over counterparty credit risk*, Risk, February 2009.

[8] International Swaps and Derivatives Association, Inc., *ISDA Close-out Amount Protocol*, 27 February 2009.

[9] R. A. Jarrow and A. M. Turnbull, *Pricing Derivatives on Financial Securities Subject to Credit Risk*, The Journal of Finance, Vol 50, 53–85, 1995.

[10] R. Merton, *Option Pricing when Underlying Stock Returns are Discontinuous*, Journal of Financial Economics, 3, 125–144, 1976.

[11] R. Merton, *On the Pricing of Corporate Debt: The Risk Structure of Interest Rates*, Journal of Finance, 29, 449–470, 1974.

[12] M. Morini and A. Prampolini, *Risky funding: a unified framework for counterparty and liquidity risk*, Risk, March 2011.

[13] V. Piterbarg, *Funding beyond discounting: collateral agreements and derivative pricing*, Risk, February 2010.

Chapter 19

A Story about the Financial Crisis – A Case Study of Lehman Brothers

19.1 Introduction

The 2008 financial crisis shocked the whole world. It all began with Lehman default. The default of Lehman Brothers Inc. on September 15, 2008 (Monday) shook the whole world in many ways. Not only was it the largest bankruptcy case in the United States, it started an enormous chain effect around the world and started a bank run unseen before. A new type of systemic risk was witnessed.

To provide a little background of how connected Lehman default is to the entire crisis – known as systemic risk, European Central Bank in 2000 defines three types of systemic risk:

- bank run
- contagion
- failure in interbank systems

However, the recent 2008 crisis defines a new systemic risk in our financial systems. Allen and Carletti (2013) view this new systemic risk as banking crises due to asset price falls. They further define such a problem as "mispricing due to inefficient liquidity provision and limits to arbitrage." Shin (2009) explicitly characterizes this liquidity-driven crisis as a new type of bank run. He contends that illiquidity, together with excess leverage and credit risk, ultimately affects nearly every financial institution. In summary, Lehman default has caused a bank-run type of phenomenon unseen before. And for the first time in the banking history, Lehman default brings out awareness of a new type of systemic risk that is driven by (lack of) liquidity. This liquidity driven systemic risk has been the focus of the recent regulation. Basel III explicitly highlights the need of liquidity regulation via two liquidity ratios – LCR and NSFR, which we will discuss in details in the next part. In this chapter, we take materials from Chen, Chidambaran, Imerman, and Soprazetti (2014) and introduce the risk Lehman faced back in 2008.

19.2 Richard (Dick) Fuld

Fuld began his career at Lehman Brothers in 1969 and had stayed with Lehman till its bankruptcy in 2008. Lehman was his only employer. In 1969, the leader of the company, Robert Lehman died and for the first time it was succeeded by a non-family member, Pete Peterson. Lehman had gone through many transitions. Most notably was the merger with Shearson, an American Express-owned securities company in 1984 and then with Hutton in 1988. It was the time of the Shearson-Lehman-Hutton. In 1994, American Express spun off Lehman in an initial public offering as Lehman Brothers Holding Inc. Fuld was elected as the CEO then and stayed at the position till its bankruptcy.

Throughout his term as the CEO, Fuld has experienced the Asian financial crisis. The company was near bankruptcy.

19.3 Lehman Time Line

As the documentary movie "Too Big To Fail" shows, the default of Lehman was more or less a government decision to reflect the regret of Bear Sterns bailout. The government realized the damage of the Bear Sterns bailout (see Chen, Chidambaran, Imerman, and Soprazetti (2014)) and decided to reverse the trend. The government felt that the market should learn the consequence of extra risk taking. According to the movie, the government then decided to let Lehman default but to save the economy by putting a sound policy around the next bank – Merrill Lynch. The government arranged for Bank of America to bail out Merill Lynch.

A detailed Lehman timeline in 2008 is provided in the Appendix. We can see that when the bailout of Bear Sterns is underway, Lehman obtained a $2 billion credit line from a consortium of 40 banks in March and $4 billion in capital in April to improve its liquidity situation. Because of this, the price of Lehman stock surged by almost 50%. After the bailout of Bear Sterns, Lehman decided not to raise anymore capital, although there was a loss for the second quarter. In the entire months of July and August, Lehman did not continue to strengthen its financial situation. By the time Lehman realized the seriousness of the situation and its CEO Richard (Dick) Fuld announced a spin-off of its real estate assets in September, it was too late. In Figure 19.1, we can see clearly how stock price reacts to actions taken by Lehman.

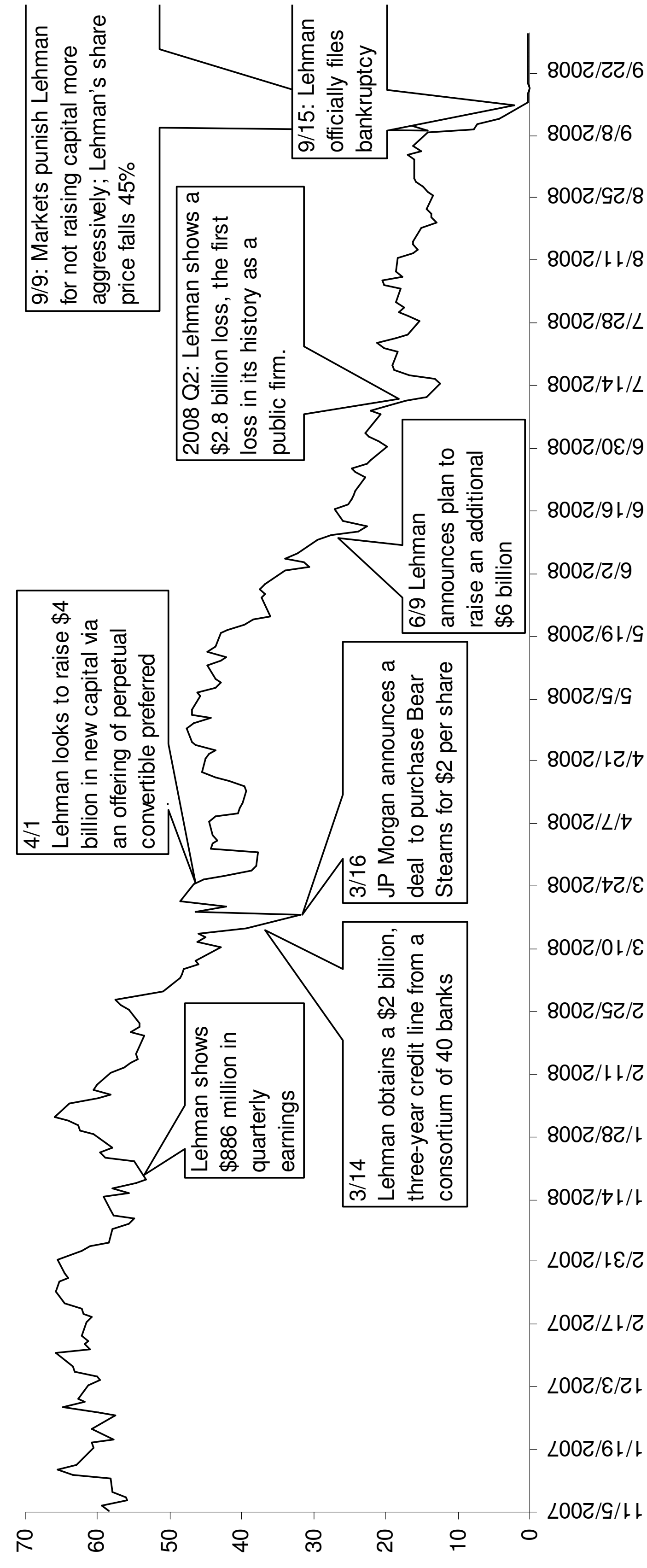


Figure 19.1: Lehman Timeline and Stock Prices

19.4 Lehman Default Probability

Chen, Chidambaran, Imerman, and Soprazetti (2014), CCIS for short from now on, take a novel approach to estimate the default probability of Lehman. They use the Geske model (an extension of Merton) to estimate the default probabilities of various tenors of Lehman over time.

Table 3: Lehman Brothers Debt Term Structure ($ Million)

Years to Maturity	December 2007	January 2008	February 2008	March 2008	April 2008	May 2008	June 2008	July 2008	August 2008	September 2008	Mean	Median	Standard Deviation
1	20,628	19,172	16,893	16,271	14,389	11,453	10,695	10,316	8,597	8,069	13,648	12,921	4,453
2	21,202	22,138	27,527	27,678	27,839	28,173	28,315	28,306	28,308	28,308	26,779	28,006	2,717
3	15,478	15,792	16,030	16,193	16,333	16,549	16,677	16,695	16,695	16,695	16,313	16,441	433
4	17,508	17,577	17,662	17,812	17,817	17,859	20,021	20,029	20,030	20,030	18,634	17,838	1,203
5	17,408	17,385	17,473	17,587	17,570	17,599	17,605	17,608	17,610	17,610	17,545	17,593	88
6	11,532	15,560	15,817	16,078	16,152	16,208	16,353	16,702	16,788	16,798	15,798	16,180	1,553
7	9,896	9,815	9,833	9,916	9,871	9,880	9,911	9,984	9,991	9,993	9,909	9,903.5	63
8	10,665	10,685	10,675	10,721	10,673	10,687	10,688	10,696	10,696	10,696	10,688	10,687	15
9	5,886	5,893	5,971	5,971	5,970	5,975	5,975	5,978	5,978	5,978	5,958	5,973	36
10	9,003	8,999	8,999	9,022	9,020	9,020	9,066	9,066	9,066	9,066	9,033	9,021	29
11	548	716	736	867	3,411	4,427	4,471	4,471	4,480	4,480	2,861	3,919	1,874
12	1,778	1,778	1,778	1,778	1,778	1,778	1,778	1,778	1,778	1,778	1,778	1,778	0
13	465	493	510	514	545	545	561	577	577	577	536.4	545	39
14	269	305	348	348	348	348	348	348	348	348	335.8	348	27
15	1,281	1,233	1,233	1,173	1,152	1,142	1,133	1,128	1,128	1,128	1,173	1,147	55
16	287	441	669	884	950	971	1,016	1,017	1,017	1,017	826	960	268
17	29	29	29	29	29	29	29	29	29	29	29	29	0
18	13	13	13	13	13	13	13	13	13	13	13	13	0
19	427	427	427	427	427	427	427	427	427	427	427	427	0
20+	21,769	21,820	22,635	22,404	22,469	24,311	24,311	24,309	24,309	24,309	23,264	23,472	1,133
Total Public Debt	166,072	170,271	175,258	175,686	176,756	177,394	179,393	179,477	177,865	177,349			

Table 4: Lehman Brothers Collateralized Transactions

This table presents data related to Lehman Brothers' collateralized transactions including repurchase agreements and reverse repurchase agreements, loaned securities and borrowed securities (cash collateral borrowed and lent, respectively), as well as the amount and sources of collateral pledged. Amounts are in millions of dollars. Data are from Lehman Brothers' 10-K and 10-Q filings and the accompanying footnotes.

	2007:Q4	2008:Q1	2008:Q2
Repo	$181,732	$197,128	$127,846
Reverse Repo	$162,635	$210,166	$169,684
Net	$19,097	($13038)	($41838)
Loaned Securities	$55,420.0	$54,847.0	$53,307
Securities Borrowed	$124,842.0	$158,515.0	$138,599.0
Net	($69,422.0)	($103,668.0)	($85,292)
Own Collateral Pledged	$150,000	$155,000	$123,031
Collateral Permitted to Re-pledge	$798,000	$929,000	$518,000
Collateral Actually Re-pledged	$725,000	$852,000	$427,000
Percentage	90.85%	91.71%	82.43%

19.5 Lehman Liquidity Problems

While subprime portfolios were the main cause for the fall of Bear Sterns and Lehman Brothers, the bankruptcy of Lehman was clearly a liquidity event. Note that the 1998 Q1 profit was positive for Lehman. Even Q2 was a loss, the amount was only $2.8 billion; while the market capitalization of Lehman were $9.1 and $8.4 billion in July and August respectively. Hence economic default could have been inevitable for Lehman ultimately, the September default had been no doubt triggered by lack of liquidity.

The press has reported that Lehman was still hopeful on Thursday (September 11) that it could survive. Yet a $3 billion margin call by JP Morgan removed that hope and Lehman had to finally file bankruptcy.

19.6 Appendix

19.6.1 Lehman Timeline

2007 to January 2008: Lehman scales back its mortgage business, cutting thousands of mortgage-related jobs and closing mortgage origination units.

2007 Q4: Lehman shows $886 million in quarterly earnings (at compared to third quarter) and reported earnings of $4.192 billion for fiscal year 2007 (a 5% increase from the previous fiscal year).

January 29, 2008: Lehman announces an increase in dividends and plans to repurchase up to 100 million shares of common stock.

2008 Q1: Lehman increases holding of Alt-A mortgages despite the prevailing troubles in the real estate market.

March 14, 2008: Lehman obtains a $2 billion, three-year credit line from a consortium of 40 banks, including JPMorgan Chase and Citigroup. On the same day, the Federal Reserve and JPMorgan Chase begin to put together a deal to bail out Bear Stearns.

March 16, 2008: JP Morgan announces a deal to purchase Bear Stearns for $2 per share.

March 18, 2008: Lehman shares surged up almost 50% after the Federal Reserve gives investment banks access to the discount window.

April 1, 2008: Lehman looks to raise $4 billion in new capital via an offering of perpetual convertible preferred stock.

Figure 8: Lehman Brothers Default Probability Term Structure

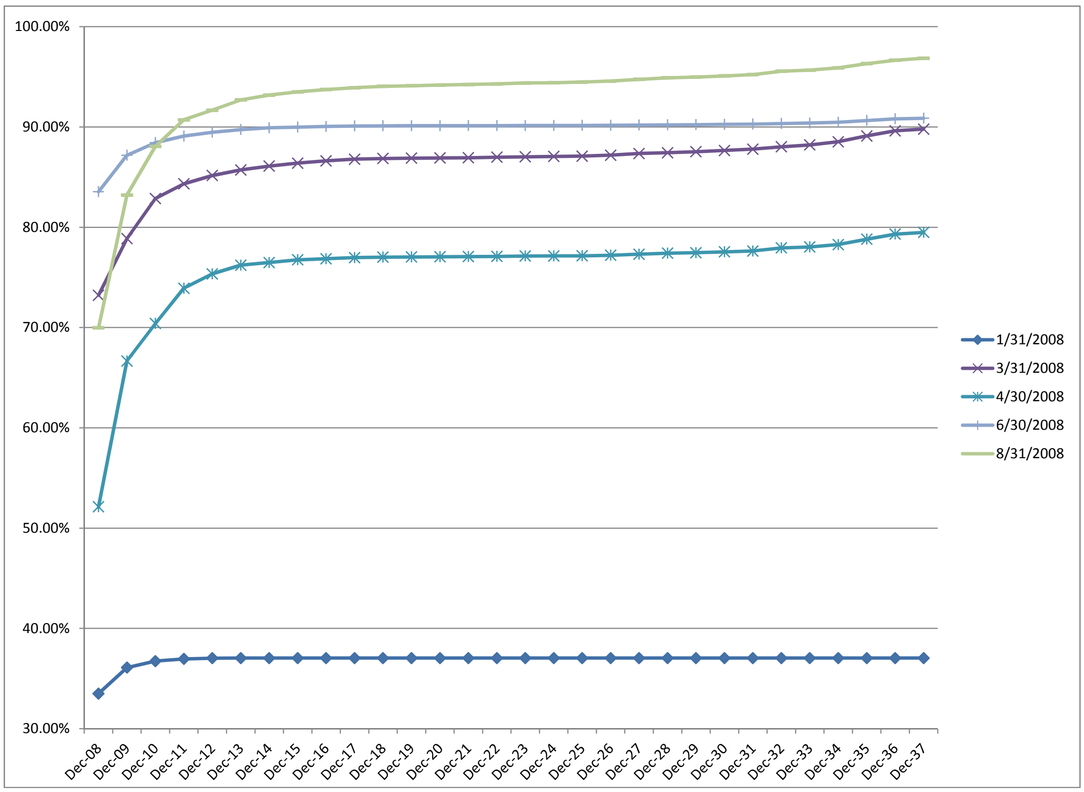


Figure 8 shows the cumulative default probabilities, or the term structure of default probabilities, for Lehman Brothers as of January 2008, March 2008, April 2008, June 2008, and August 2008. Each curve depicts the probability, at that date, of Lehman Brothers defaulting between then and the end every year from 2008 to 2032.

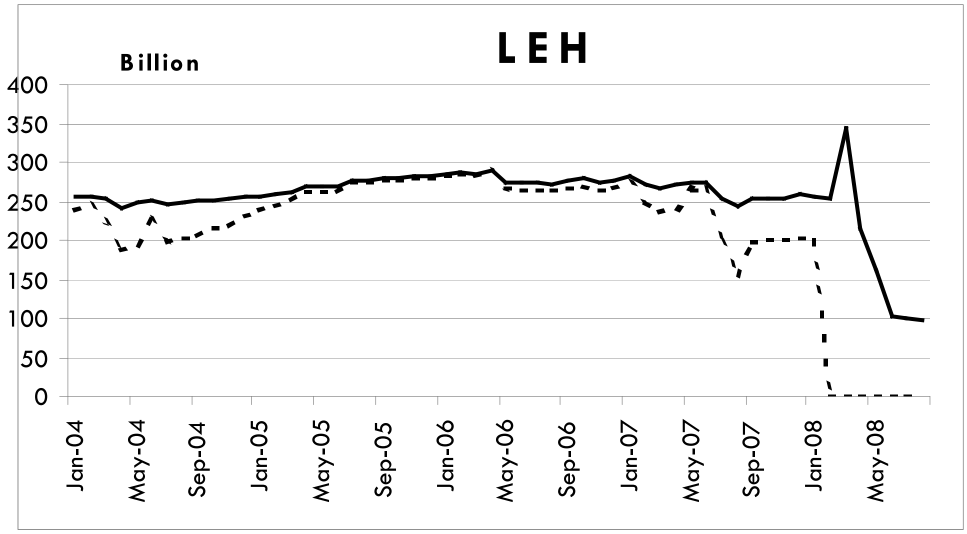


Figure 19.2: Lehman Asset Values (Liquid and Illiquid)

2008 Q2: Lehman shows a \$2.8 billion loss, the first loss in its history as a public firm. It admits the losses came not only from mortgage-related positions but also from hedges against those positions.

June 9, 2008: Lehman announces plan to raise an additional \$6 billion in new capital (\$4 billion in common stock, \$2 billion in mandatory convertible preferred stock).

July 7 to July 11, 2008: Lehman shares plunge more than 30% for the week amid rumors that the firm's assets have not been priced appropriately to reflect the true value.

September 9, 2008: Markets punish Lehman for not raising capital more aggressively; Lehman's share price falls 45% to \$7.79 on fears that the firm's capital levels are insufficient to support exposure to deteriorating real estate investments.

September 10, 2008: Lehman CEO Dick Fuld reveals plans to spin off real estate assets and sell a portion of the asset management division, insisting that the firm is solvent enough to survive.

September 11, 2008: Talks of a Lehman takeover permeate the markets as Lehman shares fall further, closing at \$4.22.

September 12, 2008: Lehman approaches several potential buyers, including Bank of America and Barclays.

September 15, 2008: Lehman officially files bankruptcy after Treasury Secretary

Paulson refuses to back any takeover; Shares close at $0.21.

September 16, 2008: Lehman is dropped from the S&P 500 Index.

September 18, 2008: Lehman shares close at $0.052 in over-the-counter trading as effects of the biggest bankruptcy in history ripple through the financial markets.

September 22, 2008: Lehman's U.S. operations reopen for business under Barclays Capital after approval for the acquisition was granted by the federal bankruptcy court presiding over the liquidation.

Part V

Others

Chapter 20

Operational Risk Management

20.1 Introduction

Events such as the September 11 terrorist attacks, rogue trading losses at Socit Gnrale, Barings, AIB and National Australia Bank serve to highlight the fact that the scope of risk management extends beyond merely market and credit risk.

The list of risks (and, more importantly, the scale of these risks) faced by banks today includes fraud, system failures, terrorism and employee compensation claims. These types of risk are generally classified under the term 'operational risk'.

The identification and measurement of operational risk is a real and live issue for modern-day banks, particularly since the decision by the Basel Committee on Banking Supervision (BCBS) to introduce a capital charge for this risk as part of the new capital adequacy framework (Basel II).

20.2 Basel II event type categories

The following lists the official Basel II defined event types with some examples for each category:

- Internal Fraud - misappropriation of assets, tax evasion, intentional mismarking of positions, bribery

- External Fraud- theft of information, hacking damage, third-party theft and forgery

- Employment Practices and Workplace Safety - discrimination, workers compensation, employee health and safety
- Clients, Products, & Business Practice- market manipulation, antitrust, improper trade, product defects, fiduciary breaches, account churning
- Damage to Physical Assets - natural disasters, terrorism, vandalism
- Business Disruption & Systems Failures - utility disruptions, software failures, hardware failures
- Execution, Delivery, & Process Management - data entry errors, accounting errors, failed mandatory reporting, negligent loss of client assets. The P&L attribution is to explain where the profits and losses come from. Due to the fact that luck plays a critical role in trading, managers need to make sure that their traders make money not due to luck but due to skills (or talents). As a result, P&L attribution has become essential in trading and fund management business.

20.3 Methods of Operational Risk Management

(The content of this section is taken from a BIS document bcbsca07.pdf entitled "Consultative Document: Operational Risk")

Basel II and various Supervisory bodies of the countries have prescribed various soundness standards for Operational Risk Management for Banks and similar Financial Institutions. To complement these standards, Basel II has given guidance to 3 broad methods of Capital calculation for Operational Risk

- Basic Indicator Approach - based on annual revenue of the Financial Institution
- Standardized Approach - based on annual revenue of each of the broad business lines of the Financial Institution
- Internal Measurement Approaches - based on the internally developed risk measurement framework of the bank adhering to the standards prescribed (methods include IMA, LDA, Scenario-based, Scorecard etc.)

The Operational Risk Management framework should include identification, measurement, monitoring, reporting, control and mitigation frameworks for Operational Risk.

20.3.1 Basic Indicator Approach

The most basic approach allocates operational risk capital using a single indicator as a proxy for an institution's overall operational risk exposure. Gross income is proposed as the indicator, with each bank holding capital for operational risk equal to the amount of a fixed percentage, α, multiplied by its individual amount of gross income. The Basic Indicator Approach is easy to implement and universally applicable across banks to arrive at a charge for operational risk. Its simplicity, however, comes at the price of only limited responsiveness to firm-specific needs and characteristics. While the Basic Indicator Approach might be suitable for smaller banks with a simple range of business activities, the Committee expects internationally active banks and banks with significant operational risk to use a more sophisticated approach within the overall framework.

The calibration of this approach is on a similar basis to that outlined in Annex 3 for the Standardized Approach. The current provisional estimate is that be set at around 30% of gross income. This figure needs to be treated with caution as it is calibrated on a limited amount of data. Also, it is based on the same proportion of capital (20%) for operational risk as the Standardized Approach and may need to be reviewed in the light of wider calibration.

For instance, in order to provide an incentive to move towards more sophisticated approaches, it may be desirable to set at a higher level, although alternative means of generating such an incentive are also available, for instance under Pillar 2 or by making the Standardized Approach the entry point for internationally active banks. It is also worth noting that a sample of internationally active banks has formed the basis of this calibration. As it is anticipated that the Basic Indicator Approach will mainly be used by smaller, domestic banks, a wider sample base may be more appropriate.

20.3.2 Standardized Approach

The Standardized Approach represents a further refinement along the evolutionary spectrum of approaches for operational risk capital. This approach differs from the Basic Indicator Approach in that a bank's activities are divided into a number of standardized business units and business lines. Thus, the Standardized Approach is better able to reflect the differing risk profiles across banks as reflected by their broad business activities. However, like the Basic Indicator Approach, the capital charge would continue to be standardized by the supervisor.

The proposed business units and business lines of the Standardized Approach mirror those developed by an industry initiative to collect internal loss data in a consistent manner. Working with the industry, regulators will specify in greater detail which business lines and activities correspond to the categories of this framework, enabling each bank to map its structure into the regulatory framework. Annex 2 presents such a mapping. This mapping exercise is yet to be finalized and further work, in consultation with the industry, will be needed to ensure that businesses are slotted into the appropriate broad categories to avoid distortions and the potential for arbitrage.

Within each business line, regulators have specified a broad indicator that is intended to reflect the size or volume of a bank's activity in this area. The indicator is intended to serve as a rough proxy for the amount of operational risk within each of these business lines. The table below presents the business units, business lines and size/volume indicators of the Standardized Approach.

Business Units	Business Lines	Indicator
Investment Banking	Corporate Finance	Gross Income
	Sales and Trading	Gross Income
Banking	Retail Banking	Annual Average Assets
	Commercial Banking	Annual Average Assets
	Payment and Settlement	Annual Settlement Throughout
Others	Retail Brokerage	Gross Income
	Asset Management	Total Funds under Management

Table 20.1: Standardized Approach

Within each business line, the capital charge is calculated by multiplying a bank's broad financial indicator by a .beta. factor. The beta factor serves as a rough proxy for the relationship between the industry's operational risk loss experience for a given business line and the broad financial indicator representing the banks. activity in that business line, calibrated to a desired supervisory soundness standard. For example, for the Retail Brokerage business line, the regulatory capital charge would be calculated as follows:

$$\text{K(Retail Brokerage)} = \beta\text{(Retail Brokerage)} \times \text{(Gross Income)}^1$$

where K(Retail Brokerage) is the capital requirement for the retail brokerage business line, β(Retail Brokerage) is the capital factor to be applied to the retail brokerage business line (each business line has a different beta factor), and Gross Income

[1] An alternative to Gross Income may be VaR.

is the indicator for this business line.

The total capital charge is calculated as the simple summation of the capital charges across each of the business lines. Annex 3 outlines a possible calibration mechanism based on existing data and 20% of current minimum regulatory capital.

The primary motivation for the Standardized Approach is that most banks are in the early stages of developing firm-wide data on internal loss by business lines and risk types. In addition, the industry has not yet been able to show a causal relationship between risk indicators and loss experience. As a result, banks that have not developed internal loss data by the time of the implementation period of the revised New Basel Capital Accord and/ or do not meet the criteria for the Internal Measurement Approach will require a simpler approach to calculate their regulatory capital charge. In addition, certain institutions may not choose to make the investment to collect internal loss data for all of their business lines, particularly those that present less material operational risk to the institution. A final important feature of the Standardized Approach is that it provides a basis for moving, on a business line by business line basis, towards the more sophisticated approaches and as such will help encourage the development of better risk management within banks.

20.3.3 Internal Measurement Approach

Methodology

The Internal Measurement Approach provides discretion to individual banks on the use of internal loss data, while the method to calculate the required capital is uniformly set by supervisors. In implementing this approach, supervisors would impose quantitative and qualitative standards to ensure the integrity of the measurement approach, data quality, and the adequacy of the internal control environment. The Committee believes that, as the Internal Measurement Approach will give banks incentives to collect internal loss data step by step, this approach is positioned as a critical step along the evolutionary path that leads banks to the most sophisticated approaches. However, the Committee also recognizes that the industry is still in a stage of developing data necessary to implement this approach.

Currently, there is not sufficient data at the industry level or in a sufficient range of individual institutions to calibrate the capital charge under this approach. The Committee is laying out, in some detail, the elements of this part of the approach and the key issues that need to be resolved (discussed below). In particular, in order for this approach to be acceptable, the Committee will have to be satisfied that a critical mass of institutions have been able individually and at an industry level to assemble adequate data over a number of years to make the approach workable.

Structure of Internal Measurement Approach

Under the Internal Measurement Approach, a capital charge for the operational risk of a bank would be determined using the following procedures.

- A banks activities are categorized into a number of business lines, and a broad set of operational loss types is defined and applied across business lines.
- Within each business line/loss type combination, the supervisor specifies an exposure indicator (EI) which is a proxy for the size (or amount of risk) of each business lines operational risk exposure.
- In addition to the exposure indicator, for each business line/loss type combination, banks measure, based on their internal loss data, a parameter representing the probability of loss event (PE) as well as a parameter representing the loss given that event (LGE). The product of EI*PE*LGE is used to calculate the Expected Loss (EL) for each business line/loss type combination.
- The supervisor supplies a factor (the gamma term) for each business line/loss type combination, which translates the expected loss (EL) into a capital charge. The overall capital charge for a particular bank is the simple sum of all the resulting products. This can be expressed in the following formula:

  Required capital = Σi Σj [γ (i,j) * EI(i,j) * PE(i,j) * LGE(i,j)] (where i is the business line and j is the risk type.)
- To facilitate the process of supervisory validation, banks supply their supervisors with the individual components of the expected loss calculation (i.e. EI, PE, LGE) instead of just the product EL. Based on this information, supervisors calculate EL and then adjust for unexpected loss through the gamma term to achieve the desired soundness standard.

Business lines and loss types

The Committee proposes that the business lines will be the same as those used in the Standardized Approach. It is also proposed that operational risk in each business line then be divided into a number of non-overlapping and comprehensive loss types based on the industry's best current understanding of loss events. By having multiple loss types, the scheme can better address differing characteristics of loss events, while the number of loss types should be limited to a reasonable number to maintain the simplicity of the scheme. The Committee's provisional proposal on the grid for business lines, loss types and exposure indicators, which has reflected considerable discussion with the industry, is shown in Annex 4. Whilst further work will be needed to specify the indicators for each risk type per business line, the Committee has more confidence that the business lines and loss types are those which will form the basis of the new operational risk framework. The Committee believes that there should be continuity between approaches, and that the indicators under the Standardised Approach and Internal Measurement Approach should, where possible, be similar. The Committee therefore welcomes comment on the choice of indicators under both approaches, including whether a combination of indicators might be used per business line in the Standardised Approach, and if so, what these might be. The Committee also welcomes comment on the proposed loss categories.

Parameters

The exposure indicator (EI) represents a proxy for the size of a particular business lines operational risk exposure. The Committee proposes to standardise EIs for business lines and loss types, while each bank would supply its own EI data. Supervisory prescribed EIs would allow for better comparability and consistency across banks, facilitate supervisory validation, and enhance transparency.

Probability of loss event (PE) represents the probability of occurrence of loss events, and Loss given event (LGE) represents the proportion of transaction or exposure that would be expensed as loss, given that event. PE could be expressed either in "number" or "value" term, as far as the definitions of EI, PE and LGE are consistent with each other. For instance, PE could be expressed as .the number of loss events / the number of transactions. and LGE parameters can be defined as .the average of (loss amount / transaction amount).. While it is proposed that the definitions of PE and LGE are determined and fixed by the Committee, these parameters are calculated and supplied by individual banks (subject to Committee guidance to ensure the integrity of the approach). A bank would use its own historical loss and exposure data, perhaps in combination with appropriate industry pooled data and public external data sources, so that PE and LGE would reflect each banks own risk profile.

Risk weight and gamma (scaling factor)

The product of EI*PE*LGE produces an Expected Loss (EL) for each business line/risk type. The term γ represents a constant that is used to transform EL into risk or a capital charge, which is defined as the maximum amount of loss per a holding period within a certain confidence interval. The scale of γ will be determined and fixed by supervisors for each business line/loss type. In determining the specific figure of γ that will be applied across banks, the Committee plans to develop an industry wide operational loss distribution in consultation with the industry, and use the ratio of EL to a high percentile of the loss distribution (e.g. 99%).

Correlations

Current industry practice and data availability do not permit the empirical measurement of correlations across business lines and risk types. The Committee is therefore proposing a simple summation of the capital charges across business line/loss type cells. However, in calibrating the gamma factors, the Committee will seek to ensure that there is a systematic reduction in capital required by the Internal Measurement Approach compared to the Standardized Approach, for an average portfolio of activity.

Further evolution

While the Committee believes that the definitions of business lines/loss types and parameters should be standardized at least in an early stage, the Committee also recognizes such standardization may limit banks. ability to use the operational risk measures that they believe most accurately represent their own operational risk (although banks could map their internal approaches into regulatory standards). As banks and supervisors gain more experience with the Internal Measurement Approach and as more data is collected, the Committee will examine the possibility of allowing banks greater flexibility to use their own business lines and loss types.

Key issues

The committee should pay close attention to the following key issues:

1. In order to use a bank's internal loss data in regulatory capital calculation, harmonisation of what constitutes an operational risk loss event is a prerequisite for a consistent approach.

2. In order to calibrate the capital calculation, an industry wide distribution will be used.

3. The historical loss observation may not always fully capture a bank's true risk profile.

4. As noted previously, a regulatory specified gamma term γ, which is determined based on an industry wide loss distribution, will be used across banks to transform a set of parameters, such as EI, PE and LGE, into a capital charge for each business line and risk type.

5. More work is needed to determine if there is a stable relationship between EL and UL and what the role of external data (to include severity) should be in assessing this relationship.

Loss Distribution Approach (LDA)

A more advanced version of an .internal methodology. is the Loss Distribution Approach. Under the LDA, a bank, using its internal data, estimates two probability distribution functions for each business line (and risk type); one on single event impact and the other on event frequency for the next (one) year. Based on the two estimated distributions, the bank then computes the probability distribution function of the cumulative operational loss. The capital charge is based on the simple sum of the VaR for each business line (and risk type). The approach adopted by the bank would be subject to supervisory criteria regarding the assumptions used. At this stage the Committee does not anticipate that such an approach would be available for regulatory capital purposes when the New Basel Capital Accord is introduced. However, this does not preclude the use of such an approach in the future and the Committee encourages the industry to engage in a dialogue to develop a suitable validation process for this type of approach. The LDA is discussed further in Annex 6.

Chapter 21

Types of Capital

21.1 Introduction

21.2 Regulatory Capital (wiki)

Capital requirement (also known as Regulatory capital or Capital adequacy) is the amount of capital a bank or other financial institution has to hold as required by its financial regulator. This is in the context of fractional reserve banking and is usually expressed as a capital adequacy ratio of liquid assets that must be held compared to the amount of money that is lent out. These requirements are put into place to ensure that these institutions are not participating or holding investments that increase the risk of default and that they have enough capital to sustain operating losses while still honoring withdrawals.

In the Basel II accord bank capital has been divided into two tiers [6] , each with some subdivisions.

21.2.1 Tier 1 capital

Tier 1 capital, the more important of the two, consists largely of shareholders' equity and disclosed reserves. This is the amount paid up to originally purchase the stock (or shares) of the Bank (not the amount those shares are currently trading for on the stock exchange), retained profits subtracting accumulated losses, and other qualifiable Tier 1 capital securities (see below). In simple terms, if the original stockholders contributed $100 to buy their stock and the Bank has made $10 in retained earnings each year since, paid out no dividends, had no other forms of capital and made no losses, after 10 years the Bank's tier one capital would be $200. Shareholders equity and retained earnings are now commonly referred to as Core Tier 1 capital, whereas Tier 1 is core Tier 1 together with other qualifying Tier 1 capital securities.

Owned funds stand for paid up equity capital, preference shares which are compulsorily convertible into equity, free reserves, balance in share premium account and capital reserves representing surplus arising out of sale proceeds of asset, excluding reserves created by revaluation of asset, as reduced by accumulated loss balance, book value of intangible assets and deferred revenue expenditure, if any.

21.2.2 Tier 2 (supplementary) capital

Tier 2 capital, or supplementary capital, comprises undisclosed reserves, revaluation reserves, general provisions, hybrid instruments and subordinated term debt.

Undisclosed Reserves

Undisclosed reserves are not common, but are accepted by some regulators where a Bank has made a profit but this has not appeared in normal retained profits or in general reserves. Most of the regulators do not allow this type of reserve because it does not reflect a true and fair picture of the results.

Revaluation reserves

A revaluation reserve is a reserve created when a company has an asset revalued and an increase in value is brought to account. A simple example may be where a bank owns the land and building of its headquarters and bought them for $100 a century ago. A current revaluation is very likely to show a large increase in value. The increase would be added to a revaluation reserve.

General provisions

A general provision is created when a company is aware that a loss may have occurred but is not certain of the exact nature of that loss. Under pre-IFRS accounting standards, general provisions were commonly created to provide for losses that were expected in the future. As these did not represent incurred losses, regulators tended to allow them to be counted as capital.

Hybrid debt capital instruments

They consist of instruments which combine certain characteristics of equity as well as debt. They can be included in supplementary capital if they are able to support losses on an on-going basis without triggering liquidation. Sometimes, it includes instruments which are initially issued with interest obligation (e.g. Debentures) but the same can later be converted into capital.

Subordinated-term debt

Subordinated debt is classed as Lower Tier 2 debt, usually has a maturity of a minimum of 10 years and ranks senior to Tier 1 debt, but subordinate to senior debt. To ensure that the amount of capital outstanding doesn't fall sharply once a Lower Tier 2 issue matures and, for example, not be replaced, the regulator demands that the amount that is qualifiable as Tier 2 capital amortises (i.e. reduces) on a straight line basis from maturity minus 5 years (e.g. a 1bn issue would only count as worth 800m in capital 4years before maturity). The remainder qualifies as senior issuance. For this reason many Lower Tier 2 instruments were issued as 10yr non-call 5 year issues (i.e. final maturity after 10yrs but callable after 5yrs). If not called, issue has a large step - similar to Tier 1 - thereby making the call more likely.

21.2.3 Common capital ratios

Common capital ratios are:

- Tier 1 capital ratio = Tier 1 capital / Risk-adjusted assets $\geq$ 6%
- Total capital (Tier 1 and Tier 2) ratio = Total capital (Tier 1 + Tier 2) / Risk-adjusted assets $\geq$ 10%
- Leverage ratio = Tier 1 capital / Average total consolidated assets $\geq$ 5%
- Common stockholders equity ratio = Common stockholders equity / Balance sheet assets

21.2.4 Capital adequacy ratio

Capital Adequacy Ratio (CAR), also called Capital to Risk (Weighted) Assets Ratio (CRAR), is a ratio of a bank's capital to its risk. National regulators track a bank's

CAR to ensure that it can absorb a reasonable amount of loss and complies with statutory Capital requirements.

Capital adequacy ratios (CARs) are a measure of the amount of a bank's core capital expressed as a percentage of its risk-weighted asset.

Capital adequacy ratio is defined as:

$$\text{CAR} = \frac{\text{Tier 1 capital} + \text{Tier 2 capital}}{\text{Risk weighted assets}}$$

- TIER 1 CAPITAL = (paid up capital + statutory reserves + disclosed free reserves) - (equity investments in subsidiary + intangible assets + current & b/f losses)
- TIER 2 CAPITAL = A) Undisclosed Reserves + B) General Loss reserves + C) hybrid debt capital instruments and subordinated debts

where Risk can either be weighted assets (a) or the respective national regulator's minimum total capital requirement. If using risk weighted assets,

$$\text{CAR} = \frac{T_1 + T_2}{a} \geq 10\%$$

The percent threshold varies from bank to bank (10% in this case, a common requirement for regulators conforming to the Basel Accords) is set by the national banking regulator of different countries.

Two types of capital are measured: tier one capital (T_1 above), which can absorb losses without a bank being required to cease trading, and tier two capital (T_2 above), which can absorb losses in the event of a winding-up and so provides a lesser degree of protection to depositors.

Risk weights

Since different types of assets have different risk profiles, CAR primarily adjusts for assets that are less risky by allowing banks to discount lower-risk assets. The specifics of CAR calculation vary from country to country, but general approaches tend to be similar for countries that apply the Basel Accords. In the most basic application, government debt is allowed a 0% risk weighting - that is, they are subtracted from total assets for purposes of calculating the CAR. Risk weighted assets - Fund Based : Risk weighted assets mean fund based assets such as cash, loans, investments and other assets. Degrees of credit risk expressed as percentage weights have been assigned by RBI to each such assets.

Non-funded (Off-Balance sheet) Items : The credit risk exposure attached to off-balance sheet items has to be first calculated by multiplying the face amount of each of the off-balance sheet items by the Credit Conversion Factor. This will then have to be again multiplied by the relevant weightage.

Local regulations establish that cash and government bonds have a 0% risk weighting, and residential mortgage loans have a 50% risk weighting. All other types of assets (loans to customers) have a 100% risk weighting.

Bank "A" has assets totaling 100 units, consisting of:

- Cash: 10 units
- Government bonds: 15 units
- Mortgage loans: 20 units
- Other loans: 50 units
- Other assets: 5 units

Bank "A" has debt of 95 units, all of which are deposits. By definition, equity is equal to assets minus debt, or 5 units.

Bank A's risk-weighted assets are calculated as follows:

- Cash 10 * 0% = 0
- Government securities 15 * 0% = 0
- Mortgage loans 20 * 50% = 10
- Other loans 50 * 100% = 50
- Other assets 5 * 100% = 5
- Total risk
- Weighted assets 65
- Equity 5
- CAR (Equity/RWA) 7.69%

Even though Bank "A" would appear to have a debt-to-equity ratio of 95:5, or equity-to-assets of only 5%, its CAR is substantially higher. It is considered less risky because some of its assets are less risky than others.

21.3 Economic Capital (Wiki)

In finance, mainly for financial services firms, economic capital is the amount of risk capital, assessed on a realistic basis, which a firm requires to cover the risks that it is running or collecting as a going concern, such as market risk, credit risk, and operational risk. It is the amount of money which is needed to secure survival in a worst case scenario. Firms and financial services regulators should then aim to hold risk capital of an amount equal at least to economic capital.

Typically, economic capital is calculated by determining the amount of capital that the firm needs to ensure that its realistic balance sheet stays solvent over a certain time period with a pre-specified probability. Therefore, economic capital is often calculated as value at risk. The balance sheet, in this case, would be prepared showing market value (rather than book value) of assets and liabilities.

The first accounts of economic capital date back to the ancient Phoenicians, who took rudimentary tallies of frequency and severity of illnesses among rural farmers to gain an intuition of expected losses in productivity. These calculations were advanced by correlations to predictions of climate change, political outbreak, and birth rate change.

The concept of economic capital differs from regulatory capital in the sense that regulatory capital is the mandatory capital the regulators require to be maintained while economic capital is the best estimate of required capital that financial institutions use internally to manage their own risk and to allocate the cost of maintaining regulatory capital among different units within the organization.

21.3.1 E&Y Model

Economic Capital (EC) is calculated to measure the amount of capital that an insurer should hold to withstand all risks it undertakes on the economic basis. Capital adequacy is the core use of EC for most insurers providing a measure of capital that truly captures the risk of the insurers own portfolio. Besides, EC is frequently featured as an important component of an insurers risk appetite framework to facilitate risk measurement and monitoring processes.

Moreover, EC plays an important role in performance management, risk-based decision making, risk-based pricing and business strategic planning. In short, EC serves as a key tool for insurance companies to adjust their risk-taking behaviors and to improve capital efficiency within the defined risk and capital framework.

In recent years, EC has drawn increasing attention from the Chinese and other Asia-Pacific insurers. The emergence of the new financial reporting standards, regulatory capital requirements and business value measurements continue to drive the principle-based and market-consistent framework. We are experiencing a structural evolution in the global insurance industry and are moving towards the Solvency II and IFRS 4 Phase II regimes. Within the context of the global regulatory environment changes, China Insurance Regulatory Commission (CIRC), has published a set of guidelines on risk management and governance for life and health insurance companies in 2010, requiring insurers to set up an EC framework as the risk management tool before 2014. Recently, Monetary Authority of Singapore (MAS) has published consultation papers on insurer Enterprise Risk Management (ERM) and investment activities requirements as well. MAS encouraged insurance companies in Singapore to compute EC in their Own Risk and Solvency Assessment (ORSA) report.

The proposed ERM and investment activities requirements would be effective from 1 January 2014 (subject to approval by the Board). Meanwhile, at the other end of Eurasia, the credit and sovereign crises continue to ferment and drive up global market volatility. In such a market, the capability to manage an insurance business in a risk-sensitive EC framework becomes necessary, particularly in Asia.

Economic capital refers to the required capital under an economic accounting framework, where assets and liabilities are measured using a market-consistent approach. EC is commonly defined as the amount of capital required to withstand a maximum loss under the market-consistent basis, over a one-year time horizon, with a certain confidence level (e.g. 99.5In other words, EC is calculated based on economic principles and linked to the companys own risk profile. From a risk measurement point of view, EC can be interpreted as a one-year Value-at-Risk (VaR) of the companys Market-consistent Value of Surplus (MVS).

The starting point for EC is an economic, market consistent approach to valuing both assets and liabilities. In an economic framework, there is a clear distinction between the role of capital and the role of Market-consistent Value of Liability (MVL). Capital is used to buffer risks during a given, defined time horizon, while provisions cover the expected liability. In other words, expected liability is covered by MVL, whereas any deviations from the expected liability are covered by capital. For most Asian insurers, economic capital is disjointed with its actual operations, as the assets are not marked to market and liabilities are still based on the Net Level Premium reserve approach. As we are moving towards the market-consistent environment, such disjoints will eventually be eliminated. The MVL is the value at which the liability could be transferred to a willing, rational, diversified counterparty in an arms length transaction under normal business conditions. The MVL consists of:

- The Best Estimate Liability (BEL): this is the expected present value of future

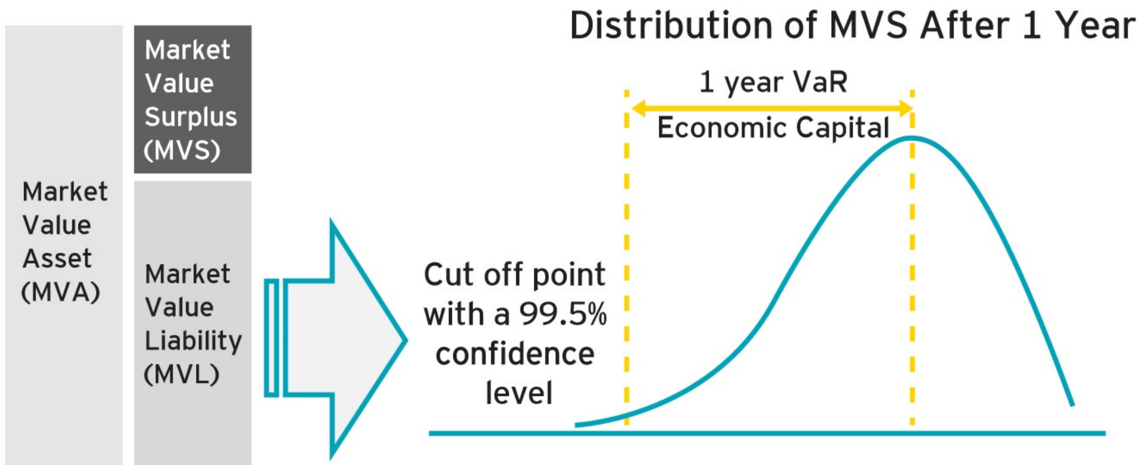


Figure 21.1: MVS

liability cash flows, and

- The Risk Margin (RM) this is an additional, explicit cost for non-hedgeable risks

There are three approaches: Confidence Level; Conditional Tail Expectation; and Cost of Capital.

Approach 1: Confidence Level

Under this approach, the MVL is considered to be the highest estimate liability within a specified confidence interval. This view of MVL presumes that the liability holder is reasonably confident that the costs due to liability are less than the payments to be received for holding the liability (under certain confidence level) Under this view, the RM is actually implied by the MVL because RM is the difference between MVL and BEL. To calculate RM under the confidence level approach, the following steps are required: Calculate best estimate liability stochastically Plot a probability distribution of the best estimate liability Calculate the risk margin based on the best estimate liability at the desired percentiles on the distribution In practice, the confidence level approach is usually simplified to the Provision of Adverse Deviation (PAD) approach. PAD will be added to the insurance and operation assumptions when calculating MVL, while BEL is calculated under best estimate assumptions. In this case, the RM is measured as the difference between the BEL with and without the PAD. When implementing such an approach, the PAD should

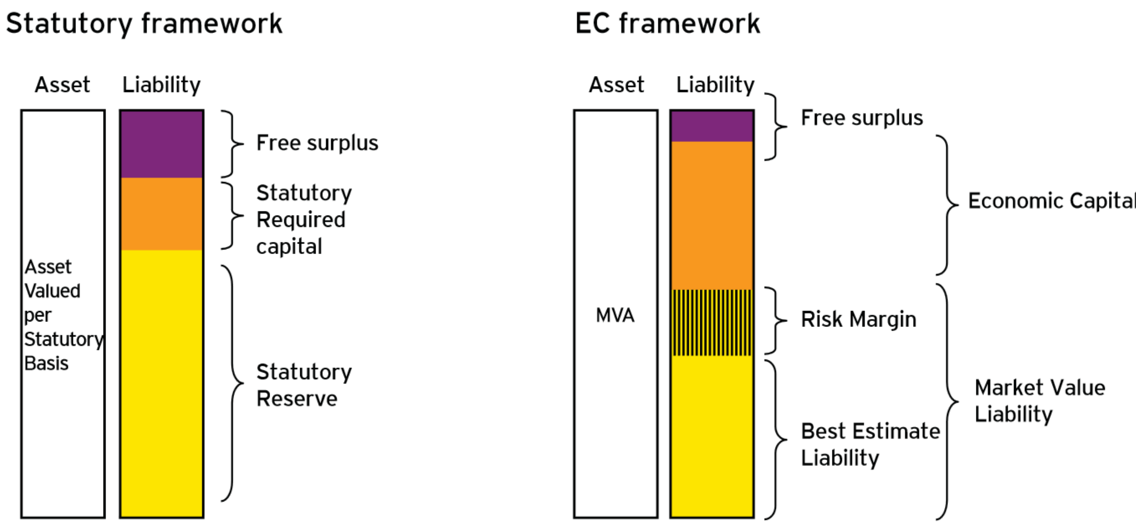


Figure 21.2: Framework

be calibrated to the required confidence level for each assumption. The calculation of this approach is straightforward and easy to explain. However, it has the drawbacks that the required confidence interval is subjective and the associated PAD parameters are difficult to calibrate. Also, the approach requires insurance shock calibration and inter-risk correlation.

Approach 2: Conditional Tail Expectation (CTE)

This approach is similar to the confidence level approach, except the risk measure is replaced by Conditional Tail Expectation (CTE). The RM is thereby the difference between the CTE (e.g. at 60 percentile) and the mean of the BEL distribution.

Approach 3: Cost-of-capital (CoC)

The cost-of-capital approach sets RM equal to the present value of the required risk premiums for each period, where the risk premiums are assumed to be proportional to the amount of capital required to support the liability. In order to calculate the RM, the following steps are required.

The CoC approach requires the projection of capital requirements across the whole projection period. The capital requirement is the amount required to be able to withstand a maximum loss over a one year period, which is consistent with the EC calculation horizon. It is less subjective than the confidence level approach. The

CoC approach is also the CRO Forum recommended approach. The main drawback of the CoC approach is that it is computationally intensive. Furthermore, there is no consensus on what cost-of-capital rate to use.

To calculate EC, the distribution of future (one year horizon) MVS under real world scenarios (outer scenarios) is required. Each MVS in the distribution will need to be calculated based on the simulated risk neutral scenarios (inner scenarios) if contractual options and guarantees exist. EC therefore requires stochastic on stochastic (SOS) calculations.

There is a widespread concern in the industry regarding the intensive computation of SOS. For example, assume there are (1) 1000 outer scenarios; (2) 1000 inner scenarios; (3) 1000 liability model points; and (4) 100 CPUs (running 2 model points per second per CPU). Under such a situation, it would take 58 days to complete the calculation. Below are three common approaches used to avoid the intensive SOS calculation.

Replicating Portfolio

The Replicating Portfolio approach matches the cash flows (magnitude and timing) of a standard asset portfolio (which may include derivatives) with the asset and liability cash flows of the insurer. The MVS under the outer scenarios can be determined by the market value of the replicating portfolio (normally closed form formula is available). The valuation of liabilities under inner scenarios on the stochastic basis is therefore not required when quantifying EC at each time period.

Curve Fit Model

The Curve fit model finds a polynomial to fit the asset and liability changes corresponding to the variation of risk factors, where the variation of risk factors are input variables of the polynomial. The regression method is used to determine the coefficients of the polynomial based on results of calibration points. The MVS under the outer scenarios can therefore be determined by the polynomial with changes in the risk factors in the scenario. Under this approach, the calculation under risk neutral scenarios is only required for a limited number of calibration points

Stress Test

Under the stress test approach, future MVS is estimated within two outer scenarios: (1) best estimate scenario and (2) stress test scenario. EC is then determined based on the difference between the MVS under these two scenarios. Usually, the stress test method is used to derive EC of the individual risk factor and a correlation matrix is used to aggregate the individual risk factor EC to derive the overall EC.

21.4 Risk Adjusted Return On Capital (RAROC)

Risk-adjusted return on capital (RAROC) is a risk-based profitability measurement framework for analysing risk-adjusted financial performance and providing a consistent view of profitability across businesses. The concept was developed by Bankers Trust and principal designer Dan Borge in the late 1970s. Note, however, that more and more return on risk adjusted capital (RORAC) is used as a measure, whereby the risk adjustment of Capital is based on the capital adequacy guidelines as outlined by the Basel Committee, currently Basel III.

21.4.1 Basic formula

Two alternative and yet identical formulas:

- RAROC = (Expected Return)/(Economic Capital) or
- RAROC = (Expected Return)/(Value at risk)

Broadly speaking, in business enterprises, risk is traded off against benefit. RAROC is defined as the ratio of risk adjusted return to economic capital. The economic capital is the amount of money which is needed to secure the survival in a worst case scenario, it is a buffer against expected shocks in market values. Economic capital is a function of market risk, credit risk, and operational risk, and is often calculated by VaR. This use of capital based on risk improves the capital allocation across different functional areas of banks, insurance companies, or any business in which capital is placed at risk for an expected return above the risk-free rate.

RAROC system allocates capital for two basic reasons:

- Risk management
- Performance evaluation

For risk management purposes, the main goal of allocating capital to individual business units is to determine the bank's optimal capital structure – that is economic capital allocation is closely correlated with individual business risk. As a performance evaluation tool, it allows banks to assign capital to business units based on the economic value added of each unit.

21.5 David Chow

The United States passed the Gramm-Leach-Billy Act in 1999 (a.k.a the Financial Services Modernization Act of 1999) to repeal the 1933 GlassSteagall Act that limited commercial bank securities activities and affiliations between commercial banks and securities firms. This article deals with that limited meaning of the GlassSteagall Act. A separate article describes the entire Banking Act of 1933. With the passage of the Gramm-Leach-Bliley Act, commercial banks, investment banks, securities firms, and insurance companies were allowed to consolidate. The legislation was signed into law by President Bill Clinton. Citi Bank is a typical recent example.

Large international banks started the one-stop service in 1990. The rise of these large international banks was a consequence of globalization in the banking industry. These banks offer integrated services in the following areas: (1) mergers and acquisitions, (2) private equity funds, (3) brokerage, (4) consulting, and (5) risk management. These large international banks expand via acquiring specialized financial institutions like brokerage firms, small banks, insurance companies, etc. These acquisitions can be multi-national. The major benefit is two-fold: (1) economical scale and (2) diversification. The former is reflected in securities trading and asset management. Theatrically, managing $10 billion and managing $100 million require similar resources. Yet $10 billion investment can generate 100 times of the returns. The latter is obvious that diversification can be most effective with the inclusion of the entire investment universe.

21.5.1 Gap Analysis

EVA (economic value added) and SVA (shareholder value added)

EVA=NOPAT-(Invested Capital x WACC)

Chapter 22

Stress Testing, DFAST, CCAR, and CVaR

22.1 Introduction

While stress testing has been a long- and well-discussed topic since Basel II, it only becomes mandatory and part of the regulation by many governments after the recent crisis. Now, all BHCs (bank holding companies) must perform stress tests on a regularly basis. Those that fail must provide remedies timely or penalties will be applied.

22.2 Stress Testing

A stress test, as the name suggests, is to simulate an extreme (bad) economic environment and see if a bank (BHC, or bank holding company) can survive under the extreme condition. Usually, there are three ways to conduct such a test – historical, parametric, and

22.2.1 Historical Stress Testing

A historical stress test is to look back in (a very long) history and find historically the worst cases and use those scenarios as the stress tests. For example, for a stock, a stress test can be its 10 worst returns in the past 10 years. These 10 worst returns are the stress scenarios used today.

22.2.2 Parametric Stress Testing

A parametric stress test is to use a model to calibrate to the target portfolio and then stress the parameters of the model. For example, for a fixed income portfolio, we can fit the value with an Heath-Jarrow-Morton model for the term structure and Jarrow-Turnbull model for the credit spreads. Then we stress the yield curve, the spreads, and the correlations between them.

Scenarios

- A 100-basis-point parallel shift (up or down) in a yield curve.
- Increasing or decreasing all the implied volatilities by 20% of current values.
- Increasing or decreasing an equity index by 10%.
- Increasing or decreasing the exchange rate for a major currency by 6%.
- Increasing or decreasing the exchange rate for a minor currency by 20%.

Regression

Historical data and linear.

Option-theoretic

Forward-looking and non-linear.

Liquidity discount model is one example.

22.2.3 Conditional Stress Testing

Let peripheral variables be functions of core variables

22.2.4 Reverse Stress Testing

While stress testing addresses the concern of extreme events and their impacts on a bank's assets/portfolios. How it is implemented is often subject to question. Basically, models that are used to generate stressed losses are the same models that are used to measure non-catastrophic risks. The behaviors of these models are well understood and as a result the stressed losses are highly predicted. [A simple analogy is basically that if a daily VaR is $20 million then a stressed loss, after shocking the scenarios by 5 times, will be simply $100 million. So what is there to learn and what is the purpose of doing such a stress test?!]

Hence a result-driven analysis is proposed. Instead of shocking the economic variables and see how the impacts are, regulators ask the following question: under what scenarios the bank is in default?

As a result, banks need to find out where the vulnerabilities lie. This is a much more efficient stress test than shocking the economic variables. Once the vulnerabilities are found, banks can then focus on those vulnerable areas and enhance risk management in those areas.

22.2.5 A Simple Demonstration

The simplest stress test can be done in the following manner. We run regression (linear) of a target portfolio on all chosen economic variables. Once the regression coefficients are estimated, we shock the coefficients by the defined scenarios.

$$\underline{y} = \underline{a} + X\beta$$

where y is a time series of portfolio returns (% or $), X contains time series of economic variables (hence a matrix), and β is the coefficients that need to be shocked by defined stressed scenarios.

Of course, such a model is too simplistic. Furthermore, the regression coefficients may not be economically or even statistically significant. So what is the meaning of such a model!

Hence, more advanced methodologies must be used. One such methodology is combine pricing models and regressions. From the pricing models we know clearly and confidently how much input variables impact asset values. Then we try to find using regressions to investigate how these input variables are affected by macro economic variables.

This two-step method is more desirable in that not only is it able to measure more accurately the relationships between asset values and economic variables, but it can incorporate nonlinearities of the relationships because pricing models are usually non-linear.

For example, if we run option values (meaning returns) against VIX (one of the chosen economic variables in CCAR), we may not be able to find significant coefficient. But if we run the implied volatility against VIX and then use the shocked VIX coefficient to compute the implied volatility and then compute the option value (return), we can much better measure the impact of shocked VIX on option value, as the implied volatility and VIX must be highly related and implied volatility can capture the true of the option well. That is:

$$\hat{\sigma} = a + bv$$
$$c = SN(d_1) - P(t,T)KN(d_2)$$

where $\hat{\sigma}$ is implied volatility and v is VIX. The second line is the Black-Scholes model for the option.

22.3 Comprehensive Capital Analysis and Review (CCAR)

CCAR is an exercise includes a supervisory stress test to evaluate whether firms would have sufficient capital in times of severe economic and financial stress to continue to lend to households and businesses. The Federal Reserve estimated revenue and losses under the stress scenario based on detailed data provided by the firms and verified by supervisors. The CCAR draws on the expertise of hundreds of staff throughout the Federal Reserve System, including supervisors, economists, markets analysts, and others.

Scenarios

- Five measures of economic activity and prices:
- Real and nominal Gross Domestic Product (GDP),
- unemployment rate of the civilian non-institutional population aged 16 and over,
- nominal disposable personal income, and the Consumer Price Index (CPI);

Four aggregate measures of asset prices or financial conditions:

- CoreLogic National House Price Index,
- National Council for Real Estate Investment Fiduciaries Commercial Real Estate Price Index
- Dow Jones Total Stock Market Index,
- Chicago Board Options Exchange Market Volatility Index;

Four measures of interest rates:

- rate on the three-month Treasury bill,
- yield on the 10-year Treasury bond,
- yield on a 10-year BBB corporate security,
- the interest rate associated with a conforming, conventional, fixed-rate, 30-year mortgage.

For the international variables,

- Three variables in four countries/country blocks.
- percent change in real GDP,
- Percent change in the Consumer Price Index or local equivalent,
- U.S./foreign currency exchange rate.

The four countries/country blocks included:

- Euro area

  The euro area is defined as the 17 European Union member states that have adopted the euro as their common currency
- United Kingdom,
- developing Asia

  The developing Asia is defined as the aggregate of China, India, Hong Kong, and Taiwan
- Japan

22.4 Dodd-Frank Act Stress Testing (DFAST)

Dodd-Frank regulation asks banks (BHCs) to conduct CCAR to examine required capital. CCAR is based upon results from the stress tests.

Supervisory scenarios

- base line

- adverse
- severely adverse

28 variables (16 on economy and GDP growth, inflation, and FX of 4 countries)

Models and methodology -31 participating BHCs -Project revenues, expenses, etc. -How b/s, rwa, n/i, of BHC are affected by economy changes -Loss = PD * LGD * EAD

More than a dozen individual models

- Wholesale loans
- commercial and industrial (C&I) loans, commercial real estate (CRE) loans
- Retail loans
- residential mortgages, credit cards, student loans, auto loans, small business loans, and other consumer lending.
- Subcategories

22.5 Credit VaR

A CVaR in an intuitive way is the necessary capital for the firm to survive its most harsh possible economic conditions. As a result, many regard CVaR as an output of CCAR.

Index