ABSTRACT

LI, LE. Numerical Studies for CVA with DWR and Portfolio Optimization with Mixed Normal Distribution. (Under the direction of Dr. Tao Pang and Dr. Wei Chen.)

Credit value adjustment (CVA) is an adjustment added to the fair value of an over-the-counter trade due to the counterparty risk. When the exposure to the counterparty changes in the same direction as the counterparty default risk the so-called *wrong-way-risk* (WWR) must be taken into account. On the other hand, if these two quantities change in the opposite direction, *right-way-risk* (RWR) takes place. These two sides of effects are also called *directional-way risk* (DWR). Calculating CVA with DWR has been a computationally challenging task especially because it has to be done frequently. In this thesis, we start with the fact that the ratio of CVA with DWR to CVA under the independent exposure and default assumption depends on the means and standard deviations of exposure and default probability and their linear correlation. The CVA DWR ratio is then decomposed into two factors, a robust correlation and a profile multiplier with further economic insight into the CVA DWR ratio. The distribution free approach in this paper entails an efficient algorithm of curve based CVA DWR calculation. A numerical study illustrates the algorithm and its benefits when CVA with WWR is priced. A detailed discussion about Hull and White model is made. Some analytical results are derived. We further show the CVA DWR multiplier decomposition bridges different existing approaches that are used to calculate CVA with DWR. Thus the decomposition provide insights of DWR and better explain some phenomenon when DWR is in present. Portfolio optimization with mixed normal distribution is presented. It's shown that mixed normal distribution can better describe stock index returns than normal distribution. Under a mixed normal assumption Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) can be easily computed, which greatly reduce the computational efforts of the optimization problem. A numerical example with 5 assets is presented to show the computational efficiency.

Numerical Studies for CVA with DWR and Portfolio Optimization with Mixed Normal Distribution

by
Le Li

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Operations Research

Raleigh, North Carolina

2015

APPROVED BY:

Dr. Negash Medhin	Dr. Min Kang
Dr. Tao Pang Co-chair of Advisory Committee	Dr. Wei Chen Co-chair of Advisory Committee

DEDICATION

To my father, Enlai Li, my mother, Ye Wang and my wife, Yijing Cheng for their endless love.

BIOGRAPHY

Le Li was born in Beijing, capital of China. He received his bachelor degree in Statistics from Fudan University in 2007. After that he came to North Carolina State University and received master degree in Financial Maths in 2009. In 2010 he joined the Ph.D. program of Operations Research. He has been working in SAS Institute as a Graduate Industry Trainee for more than 3 years. Meanwhile, he passed all 3 level exams of the Chartered Financial Analyst (CFA). Now he is a senior associate research statistician developer with SAS Institute.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Dr. Tao Pang, for his guidance, encouragement and support throughout my Master and PhD studies since I came to NC State University.

I would also like to extend my appreciation to my co-chair, Dr. Wei Chen. His industry knowledge not only provides valuable opinions to my research but also guides me well when I work in SAS Institute.

I don't know how to thank Dr. Negash Medhin enough. He is always doing all he can to help students achieve their goals. Without his help I could not have continued my PhD study.

Dr. Min Kang's lectures on probability theory and stochastic differential equation are so amazing. She shows nothing but the beauty of maths. I don't have the words to thank her for that.

Special thanks to Dr. Jeff Scroggs for all the efforts he has devoted to the Financial Maths program. The knowledge and skills I am equipped with proves to be valuable to my work with SAS Institute.

Many thanks to Dr. Zhilin Li for being my Graduate School Representative. It is very nice of him to spare his valuable time with me with such a short notice.

To all the friends who offered me a hand, I appreciate it more than you will ever know.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

Over-the-counter derivatives dealers need to adjust the value of a portfolio of derivatives transactions with counterparties since there will be possible losses simply due to defaults by the counterparties even if the trade itself doesn't incur significant profit or loss. This has become a standard practice nowadays. The adjustment amount is known as Credit Value Adjustment or CVA. Prior to the credit crisis in 2007, CVA was considered negligible. According to the Basel Committee, during the financial crisis roughly two-thirds of the credit crisis risk losses were due to CVA losses and only one-third were due to actual defaults[2]. Consequently increasing attention has been paid to counterparty risk and CVA charges ever since. *Wrong-way-risk* (WWR) and right-way-risk (RWR) are introduced. WWR and RWR together are usually called *directional-way-risk* (DWR).

Portfolio optimization has been an interesting topic for several decades. Lots of alternative distributions have been proposed to substitute normal distribution in order to better fit the financial data. Different risk measures have been proposed to address different concerns of risk managers. We focus on Value-at-Risk (VaR) and Conditional-Value-at-Risk (CVaR) optimization with mixed normal distribution.

1.1 Background and Literature Review

A derivatives dealer can have thousands counterparties with over millions derivative transactions in total. For each counterparty a CVA is calculated based on the total net exposure to the counterparty. Each CVA depends on the net value of the portfolio to the counterparty. The computational complication of CVA is significant.

In parallel to CVA, there is a measure called DVA, short for debt value adjustment. It is the credit adjustment value to the dealer's counterparty due to possible defaults by the dealer. DVA is not typically an interesting valuation and capital exercises because for one thing the dealer typically does not realize DVA without the actual default, and for the other the dealer can even book profit and loss against its own credit change. Therefore, our focus will be CVA in this paper.

CVA includes two types of exposures, the potential movements in the counterparty credit spreads and the underlying market variables that affect the risk free instrument values. According to Basel III published in December 2010 by the Basel Committee on Banking Supervision, dealers should only identify the CVA risk resulted from the changes in the counterparty credit quality. Some practitioners and researchers argue that when the market variable is hedged, the hedging positions will increase but not decrease the required capital[24]. The hedging transactions will introduce new counterparties. As a result, it is not possible to hedge CVA perfectly.

In practice CVA calculations often assume the independence between the counterparty's probability of default and the dealer's exposure to this counterparty. However, this assumption usually doesn't hold in reality, which means they can move together. In the case they move in the same direction, the so-called *wrong-way risk* (WWR) materializes. The *right-way risk* (RWR) takes place when the two factors move in the opposite directions. These two sides of effects are also called *directional-way risk* (DWR). WWR can be harmful. Examples like AIG in 2008 have underlined the importance of WWR and brought growing interest in modeling CVA with WWR. RWR tends to reduce the price of CVA.

In response to WWR, Basel III levied a multiplier α to the exposure in the CVA capital charge. By default, α is set to be 1.4. Financial institutions can use a lower α if enough justification is provided but it can not be less than 1.2. However estimates of α by banks are reported to range from 1.07 to 1.10 [26]. Although there is no incentive in BASEL III that rewards RWR, we will discuss both effects in this thesis.

According to Jon Gregory [28], there are mainly two types of approaches to model CVA with DWR, correlation approach and parametric approach. The former one includes Pykhtin and Rosen [36], Rosen and Saunders [39], Brigo and Alfonsi [9], Brigo and Pallavicini [10], Skoglund, Vestal and Chen [40] and Pang, Chen and Li [34]. The latter one is also called stochastic hazard rate approach and was proposed in Hull and White [24]. Studies of Hull and White model are also given in Ghamami and Goldberg [21] and Ruiz, Boca and Pachòn [38].

Pykhtin and Rosen [36] show that with a normal distribution of exposures and a Gaussian copula, analytical expressions can be obtained for CVA even in the case of WWR. Rosen and Saunders [39] propose a robust method, which effectively leverages existing 'pre-computed' exposures into a joint market and credit risk portfolio model. Their approach to WWR is through an artificial copula, the so called *ordered-scenario copula* which governs the relation between an exposure and a counterparty creditworthiness indicator. The method highly depends on the chosen copula and distribution assumption of the creditworthiness indicator and the time to default. However, the choice of the proper copula and the dependence level is also mostly subjective. The authors give no criteria on how to choose a copula or how to select the dependence level.

Brigo and Alfonsi [9], Brigo and Pallavicini [10] and Skoglund, Vestal and Chen [40] correlate the exposure and default probability with two joint stochastic processes. Exposures and default probabilities are then generated from the underlying interest rates and default intensities respectively.

Hull and White [24] defines hazard rate as a deterministic function of the exposure and relate it to default. The dependence of the default probability and the exposure is given between the hazard rate function and the exposure. This approach is tractable but the deterministic function reduces the modeling flexibility. Their model is subject to subjective judgment about the amount of right-way or wrong-way risk the counterparty brings [24]. It requires the users to have a good knowledge of the counterparty's business.

Both types of approaches have advantages and disadvantages. Our goal is not to draw conclusions about which approach is better. We try to propose a semi-parametric model that can calculate CVA with DWR in a more efficient way and better interpret both approaches.

Portfolio optimization has been an interesting topic for several decades. Early work can be traced back to Markowitz [30]. A mean-variance framework was built. The goal is to find the optimal asset allocation that minimizes the variance of a portfolio, meanwhile certain minimum required return level should be exceeded. Portfolio variance was the risk measure.

Due to the leptokurtosis of financial data [18], normal distribution may not describe the historical returns very well. Some other distributions were proposed to fit financial data, like generalized student's t distribution [41] and stable distribution [32] [8] [33]. Instead of those distributions, we use a mixed normal model, which will improve computational efficiency.

Besides the distribution selection, variance is not adequate to describe the risk of a portfolio and some new risk measures were introduced. A popular one is Value-at-Risk (VaR). Hull and White proposed a way to calculate VaR when the returns are not normally distributed [25]. VaR can describe tail risk but can not tell a portfolio manager on average how much the loss is going to be. Conditional Value-at-Risk (CVaR) was used to address this. CVaR is defined to be the expected loss given the loss is beyond VaR. As is shown in Rockafellar and Uryasev [37] CVaR optimization can be quite convenient. Both VaR and CVaR are used as the risk measure in this thesis. With a numerical example, we show the portfolio optimization problems can be solved fast under this framework.

1.2 Contribution of This Research

In this thesis, we investigate existing models that are used to calculate CVA with DWR. Advantages and disadvantages of each model are also discussed. There are several common drawbacks that all these models have. The information embedded in the simulations on a single day can not be reused; all of them try to calculate one CVA value without showing how a confidence interval can be constructed; no efforts have been done to bridge different models; none of the existing models provides a good way to gain insights of CVA with DWR.

We will address these drawbacks one by one. Firstly, a decomposition of CVA DWR multiplier is proposed. Based on this decomposition, a semi-parametric method is proposed. Then an efficient curve fitting algorithm is developed. Guidelines are provided so that the information embedded in the simulations on a single day can be reused with certain conditions satisfied. Re-usability criteria are listed to outline these conditions

A single CVA value is not as a helpful as a confidence interval of CVA to a risk manager. Our work allows a risk manager to construct one without adding much computational burden. Details are shown how this can be done.

By performing different approaches with the same set of exposures, we illustrate how our robust correlation and profile multiplier can be used to interpret different sets of parameters that are used in different types approaches. Thus a bridge is built. With this bridge, attempts of interpreting CVA with DWR are made. We discuss extensively how the correlation parameter, namely *b*, in Hull and White model impacts robust correlation and profile multiplier. Insights of CVA with DWR are also provided and proves to be useful explaining phenomenon observed in other researchers' work.

We show that mixed normal distribution can better describe stock index return data than normal distribution in the sense that mixed normal distribution can better model skewness and heavy tail. It's also shown that VaR and CVaR optimization with mixed normal distribution has great computational advantage.

1.3 Outline of the Thesis

The rest of this thesis is organized as follows. In Chapter 2, preliminary knowledge used in our research is introduced. We briefly review interest rate models and credit risk models. We present, in details, how we simulate interest rates and price interest rate swaps. Pricing methodology of plain vanilla interest rate swaps is given in Appendix A.

In Chapter 3, CVA and DWR are introduced. The standard simulation approach for CVA is set up. We then propose our model to calculate CVA with DWR. We decompose the CVA DWR multiplier into two factors, namely robust correlation and profile multiplier.

We present correlation approach and focus on WWR in Chapter 4. With a series of numerical studies, an efficient curve fitting algorithm is developed. We discuss the use of our model to build a confidence interval for CVA. Some findings about correlation approach are also discussed.

Parametric approach is shown in Chapter 5. An analytical expression of robust correlation and profile multiplier is given. We use a numerical example to show how the analytical results and our model can be applied to gain insights of CVA with DWR.

Portfolio optimization with mixed normal distribution is discussed in Chapter 6. We show that VaR and CVaR of a mixed normal distribution can be found with little computational efforts. Thus the work can be useful in the stress testing framework.

Chapter 2

Preliminary

Now let's introduce some preliminaries that will be used in our research. This chapter has two parts: interest rate models and credit risk models. Pricing of plain vanilla interest rate swaps is given in Appendix A.

2.1 Interest Rate Models

The risk free interest rate is one of the most important factors for pricing the financial instruments and hence has been studied by a lot of researchers and practitioners. Short rate models include those by Merton [31], Vasicek [43], Cox, Ingersoll, and Ross [15](CIR), Ho and Lee [22], Hull and White [23], Black, Derman and Toy [6] and Black and Karasinski [7](BK). Comparison and implementation of affine models has been given by Chan, Karolyi, Longstaff and Sanders [11] and Paseka, Koulis and Thavaneswaran [35].

Since our focus is CVA and we propose a semi-parameter approach, we use the widely known and popular CIR and Vasicek models in our numerical studies. The stochastic process of CIR model is

$$dr_t = \kappa_r(\theta_r - r_t)dt + \sigma_r\sqrt{r_t}dW_t \tag{2.1}$$

and for Vasicek model we have

$$dr_t = \kappa_r(\theta_r - r_t)dt + \sigma_r dW_t \tag{2.2}$$

where W_t is a standard Brownian Motion and κ_r, θ_r and σ_r correspond to the mean-reverting speed, the mean-reverting level and the volatility respectively.

For CIR model, some efforts were made to avoid negative interest rates brought by the discretization. See Deelstra [16], Diop [17] and Alfonsi [1]. We take absolute value of the interest rate which is proposed in Diop [17]. Euler discretization is applied to both models.

Plots of interest rate paths with both CIR and Vasicek models are shown below. The parameters are chosen to be $\kappa_r = 0.1$, $\theta_r = 0.05$ and $\sigma_r = 0.06$ for both models. This set of parameters is also used in Skoglund, Vestal and Chen [40].

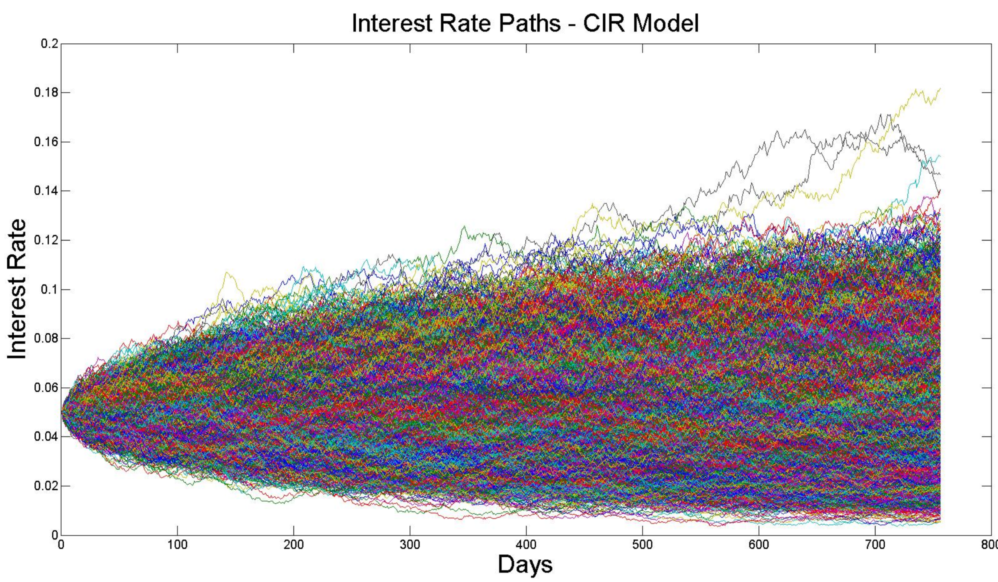


Figure. 2.1 *Simulated interest rate paths - CIR model*

2.2 Credit Risk Modeling

On the other hand, various methods to model the default probability have been proposed and studied. Structural approach proposed by Black and Scholes [5] and Merton [31] is also known as the value-of-the-firm approach. It models the behavior of the total value of the firm's assets. Default is triggered when the total value of the firm's assets fall below a preset barrier - such as its debt. Black and Cox [4] extends Merton's model in a way that allows premature default. Fouque et al. [19] introduces a stochastic volatility model and Fouque et al. [20] is an extension to a multi-name case.

Jarrow, Lando and Yu [27] discusses the relationship of the default intensities under physical and risk neutral measures. A CIR model is used to model the intensity and a 'drift change in the intensity' is derived. They provide conditions under which the empirical and martingale default intensities are equivalent. Berndt et al. [3] opts to use a Black-Karasinski model to model the intensity process. They assume the logarithm $X_t = ln(h_t)$ of the default intensity h_t has the

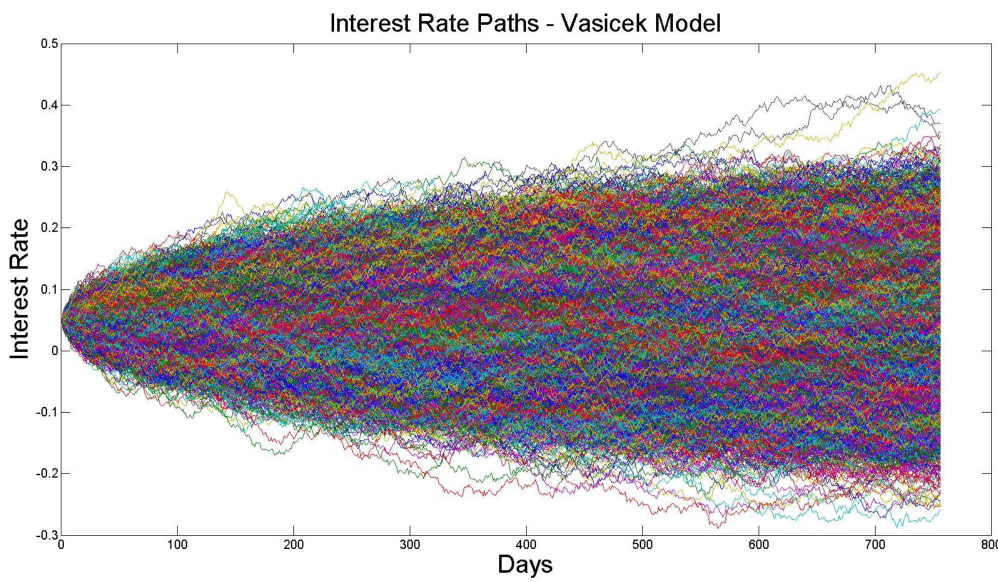


Figure. 2.2 *Simulated interest rate paths - Vasicek model*

following SDE:

$$dX_t = \kappa_X(\theta_X - X_t)dt + \sigma_X dZ_t \tag{2.3}$$

where Z_t is a standard Brownian motion, and κ_X, θ_X and σ_X are unknown constants. These parameters are estimated from the monthly Moody's KMV EDF observations.

We use the recovery rate R of 0.4 proposed by Varma and Cantor [42] and choose κ_X and σ_X estimated in Berndt et al.[3] for the Oil and Gas sector and set the long term mean so that $0.6\exp(-\theta_X) = 0.02$. Hence the parameters are chosen to be $\kappa_X = 0.470$, $\theta_X = 3.401$ and $\sigma_X = 1.223$. We further assume that certain conditions in Jarrow, Lando and Yu[27] are met such that the intensity process is under the risk-neutral measure. Once we get all the default intensities, the default probabilities will be calculated with the following equations.

The relationship between default intensity h_t and survival probability $S(t)$ is given by

$$S(t) = \exp\left(-\int_0^t h_u du\right) \tag{2.4}$$

A discretization of Eq. 2.4 based on equally spaced time period is

$$S(t) = \exp\left(-\Delta t \sum_{t_i \le t} h_{t_i}\right) \tag{2.5}$$

The default probability $q(t_i)$'s in Eq. 3.5 follows

$$q(t_i) = S(t_{i-1}) - S(t_i) \tag{2.6}$$

Chapter 3

A Semi-Parametric Method

In this chapter, we propose an approach that does not depend on the distribution of exposures and the default probabilities.

3.1 CVA and DWR

Throughout our paper, we consider unilateral CVA, which ignores the bank's own default risk. The general formula for CVA is

$$CVA = (1-R)\mathbf{E}\left[\mathbf{1}_{\{\tau \leq T\}} V^{+}(\tau)\right], \tag{3.1}$$

where R is the constant recovery rate; τ is the default time; T is the time to expiration; $\mathbf{1}_{\{\tau \leq T\}}$ is the default indicator function that is 1 if the default time $\tau \leq T$ and 0 otherwise; $V^{+}(t)$ is the risk-neutral discounted exposure at time t.

The exposure $V^{+}(t)$ is subject to netting and collateral but not recovery. Assume the discounted portfolio value and required collateral are $V(t)$ and $C(t)$ respectively then the exposure is

$$V^{+}(t) = \max\left(V(t) - C(t), 0\right). \tag{3.2}$$

Eq. 3.1 can be expressed as

$$CVA = -(1-R)\left[\int_0^T \mathbf{E}\left[V^+(t)\right] dS(t)\right], \tag{3.3}$$

where $S(t)$ is the survival function of the counterparty's default time τ.

We take discretization of the integral into $(t_i)_{i=0}^k$ time steps where $t_0 = 0$ and $t_k = T$. Eq. 3.3 can be written as

$$\begin{aligned} CVA &= (1-R)\sum_{i=1}^k \mathbf{E}\left[V^+(t_i)[S(t_{i-1}) - S(t_i)]\right] \\ &= (1-R)\sum_{i=1}^k \mathbf{E}\left[V^+(t_i)q(t_i)\right], \end{aligned} \tag{3.4}$$

where $q(t_i)$ is the probability of default within $(t_{i-1}, t_i]$.

In a discretization approximation, the representative exposure $\overline{V}^+(t_i)$ is usually used instead of $V^+(t_i)$. The formula used in BASEL III follows

$$\overline{V}^+(t_i) = \frac{V^+(t_{i-1}) + V^+(t_i)}{2}.$$

One can also use a right- or left-point rule, or use the exposures at the middle points

$$\overline{V}^+(t_i) = V^+\left(\frac{t_{i-1}+t_i}{2}\right).$$

Thus Eq. 3.4 can be written as

$$CVA = (1-R)\sum_{i=1}^k \mathbf{E}\left[\overline{V}^+(t_i)q(t_i)\right]. \tag{3.5}$$

Wrong-way risk (WWR) occurs when the exposure $\overline{V}^+(t_i)$ tends to grow while the default probability $q(t_i)$ becomes larger. A good example of the presence of WWR is the situation in which many credit default protection buyers had to deal with the falling giant American

International Group (AIG) during the end of last decade financial crisis. AIG Financial Products (AIGFP) was a division of AIG. It sells a huge amount of credit default swap known as CDS against various types of debts including collateralized debt obligation or CDO. Prior to the crisis, the default of individual bond issuers and investors were considered independently, i.e. counterparty A's default will not affect the likelihood of counterparty B's default. However, this was no longer true during the crisis due to the systematic risk. Many credit instruments that AIGFP sold CDS against suddenly incur high default risk. AIGFP thus had a huge liability from its CDS business which increases the default risk of AIG itself. In this situation, the CDS buyers faced increasing exposure $\overline{V}^{+}(t_i)$ from the CDS contracts because of the credit risk of the underlying debt as well as the increasing default probability of the CDS counterparty i.e. AIG, $q(t_i)$ simultaneously. One approach to capture WWR is to use the positive correlation between stochastic processes of the exposure and the default probability. In other words, WWR is presented when there is a positive correlation between $\overline{V}^{+}(t_i)$ and $q(t_i)$. When there is WWR, the expected credit loss due to counterparty risk amplifies. Therefore, WWR is not negligible.

3.2 CVA DWR Multiplier Decomposition

We assume that the distributions of $\overline{V}^{+}(t_i)$'s are unknown but have finite mean and volatility, which depend on time t_i and can vary across time. Likewise, the distributions of default probability $q(t_i)$'s is also unknown. The corresponding time varying mean and volatility should be finite as well.

Denote the means and variances of the exposures and default probabilities as follows

$$\begin{aligned}\mathbf{E}\left[\overline{V}^{+}(t_i)\right] = \mu_V(t_i), \quad \mathbf{Var}\left(\overline{V}^{+}(t_i)\right) = \sigma_V^2(t_i),\\ \mathbf{E}\left[q(t_i)\right] = \mu_q(t_i), \quad \mathbf{Var}\left(q(t_i)\right) = \sigma_q^2(t_i).\end{aligned}$$

We now consider the correlation coefficient of $\overline{V}^{+}(t_i)$ and $q(t_i)$ directly, namely $\rho(t_i)$. For each time node, we have

$$\rho(t_i) = \frac{\mathbf{E}\left[\overline{V}^{+}(t_i)q(t_i)\right] - \mu_V(t_i)\mu_q(t_i)}{\sigma_V(t_i)\sigma_q(t_i)}. \tag{3.6}$$

With the independence assumption and Eq. 3.5, we can get

$$CVA_{IND} = (1-R)\sum_{i=1}^{k}\mu_V(t_i)\mu_q(t_i). \tag{3.7}$$

Then, according to Eq. 3.5, it is the CVA when exposures and default probabilities are independent. In general, if we do not know whether the exposure and the default are independent, by virtue of Eq. 3.5 and Eq. 3.6, the general formula for CVA is

$$\begin{aligned} CVA &= (1-R)\sum_{i=1}^{k}\mathbf{E}\left[V(t_i)q(t_i)\right] \\ &= (1-R)\sum_{i=1}^{k}\frac{\mathbf{E}\left[V(t_i)q(t_i)\right]-\mu_V(t_i)\mu_q(t_i)+\mu_V(t_i)\mu_q(t_i)}{\sigma_V(t_i)\sigma_q(t_i)}\sigma_V(t_i)\sigma_q(t_i) \\ &= (1-R)\left[\sum_{i=1}^{k}\rho(t_i)\sigma_V(t_i)\sigma_q(t_i)+\sum_{i=1}^{k}\mu_V(t_i)\mu_q(t_i)\right]. \end{aligned} \tag{3.8}$$

We define the ratio of CVA to CVA_{IND} as

$$CVA_{ratio} \equiv \frac{CVA}{CVA_{IND}}. \tag{3.9}$$

Then, by virtue of Eq. 3.7 and Eq. 3.8, we can get

$$CVA_{ratio} = 1+\frac{\sum_{i=1}^{k}\rho(t_i)\sigma_V(t_i)\sigma_q(t_i)}{\sum_{i=1}^{k}\mu_V(t_i)\mu_q(t_i)}, \tag{3.10}$$

where $\rho(t_i)$ is the correlation between $\overline{V}^{+}(t_i)$ and $q(t_i)$ given by Eq. 3.6.

Define

$$\bar{\rho} \equiv \frac{\sum_{i=1}^{k}\rho(t_i)\sigma_V(t_i)\sigma_q(t_i)}{\sum_{i=1}^{k}\sigma_V(t_i)\sigma_q(t_i)}, \tag{3.11}$$

and

$$C_p \equiv \frac{\sum_{i=1}^{k}\sigma_V(t_i)\sigma_q(t_i)}{\sum_{i=1}^{k}\mu_V(t_i)\mu_q(t_i)}. \tag{3.12}$$

We call $\bar{\rho}$ *robust correlation* and C_p *profile multiplier*. It is easy to see that $-1 \le \bar{\rho} \le 1$. C_p describes the profiles of exposure and default probability.

Now the Eq. 3.10 can be written as

$$CVA_{ratio} = 1 + \bar{\rho} C_p. \tag{3.13}$$

and CVA can be expressed as

$$CVA = (1 + \bar{\rho} C_p) CVA_{IND}. \tag{3.14}$$

If the robust correlation is not sensitive to small market changes, then we only need to estimate it once in a small sub-period. With this stability assumption the computational effort can be significantly reduced. We can simply simulate the exposures and default probabilities independently and then derive CVA with WWR in a more efficient way.

It's clear that the ratio depends on two factors $\bar{\rho}$ and C_p. If either of them is sensitive to the market changes, so will be the CVA ratio. Since one can derive C_p easily when calculating CVA_{IND} with full simulation, more attention will be paid to the sensitivity of robust correlation.

First of all, C_p can be considered as a *quasi* coefficient of variation of the defaultable value, that is, the product of default free value and default probability assuming zero recovery rate. It can be a by-product from the CVA_{IND} calculation. In the defined CVA ratio, the robust correlation $\bar{\rho}$ is simply a coefficient of the effect from the *quasi* coefficient of variation C_p. Intuitively, a portfolio with large C_p tends to inhibit high WWR risk. However, the actual impact from C_p can range from none (when $\bar{\rho} = 0$) to full (when $\bar{\rho} = 1$). Our numerical study shown in the next chapter is aiming to find out the effect of C_p and $\bar{\rho}$ on CVA ratio. We observe the factor of $\bar{\rho} C_p$ can be as large as 3. The CVA ratio also matches the spirit of the Basel α, which is currently set to 1.4 in Basel III.

Chapter 4

Efficient CVA Curve

In this chapter, we will perform a series of numerical analyses in the context of vanilla interest rate swaps and develop an algorithm that calculates CVA with DWR efficiently.

4.1 Simulation Models

Although the model does not require any distribution assumption, for the sake of this numeric study we will apply a few industry standard risk free interest rate and default hazard rate models.

The risk free interest rate is one of the most important factors for asset pricing and hence has been studied by a lot of researchers and practitioners. We use short rate models proposed by Cox, Ingersoll, and Ross [15](CIR) and Vasicek [43] in our numerical study. The stochastic process of CIR model is

$$dr_t = \kappa_r(\theta_r - r_t)dt + \sigma_r\sqrt{r_t}dW_t, \tag{4.1}$$

and for Vasicek model we have

$$dr_t = \kappa_r(\theta_r - r_t)dt + \sigma_r dW_t, \tag{4.2}$$

where W_t is a standard Brownian Motion, κ_r, θ_r and σ_r the corresponding mean-reverting speed, mean-reverting level and volatility respectively.

We simulate 3-year pay fixed interest rate swap exposures with quarterly payments by modeling the interest rate $r(t_i)$ with both models. We will study and compare the effect of WWR for both models.

Euler discretization is applied to both models. For CIR model, some efforts were made to avoid negative interest rates brought by the discretization. See Deelstra [16], Diop [17] and Alfonsi [1]. We take absolute value of the interest rate when it is negative, as proposed in Diop [17].

The parameters are chosen to be $\kappa_r = 0.1$, $\theta_r = 0.05$ and $\sigma_r = 0.06$ and we assume at time 0 the interest rate term structure is flat and equal to θ_r. The fixed rate is also set to be θ_r so the swap is at the money. The simulated future exposures have time dependent mean $\mu_V(t_i)$ and volatility $\sigma_V(t_i)$. This set of parameters is also used in Skoglund, Vestal and Chen [40]. Later we will perform some sensitivity analysis when κ_r, θ_r and σ_r change.

Various default probability models have been proposed and studied. Jarrow, Lando and Yu [27] discusses the relationship of the default intensities under physical and risk neutral measures. A CIR model is used to model the intensity and a 'drift change in the intensity' is derived. They provide conditions under which the empirical and martingale default intensities are equivalent. Berndt et al. [3] opt to use a Black-Karasinski [7] model for the intensity process. They assume the logarithm $X_t = \log h_t$ of the default intensity h_t has the following SDE:

$$dX_t = \kappa_X(\theta_X - X_t)dt + \sigma_X dZ_t. \tag{4.3}$$

where Z_t is a standard Brownian motion, and κ_X, θ_X and σ_X are unknown constants. These parameters can be estimated from the monthly Moody's KMV (Kealhofer, McQuown and Vasicek) EDF (Expected Default Frequency) observations.

We use the recovery rate R of 0.4 proposed by Varma and Cantor [42] and choose κ_X and σ_X estimated in Berndt et al.[3] for the Oil and Gas sector and set the long term mean so that $0.6\exp(-\theta_X) = 0.02$. Hence the parameters are chosen to be $\kappa_X = 0.470$, $\theta_X = 3.401$ and $\sigma_X = 1.223$. We further assume that certain conditions in Jarrow, Lando and Yu[27] are met such that the intensity process is under the risk-neutral measure. Once we get all the default intensities, the default probabilities will be calculated with the following equations.

The relationship between default intensity h_t and survival probability $S(t)$ is given by

$$S(t) = \exp\left(-\int_0^t h_u du\right). \tag{4.4}$$

A discretization of Eq. 4.4 based on equally spaced time periods is

$$S(t) = \exp\left(-\Delta t \sum_{t_i \le t} h_{t_i}\right). \tag{4.5}$$

The default probabilities $q(t_i)$'s in Eq. 3.5 follow

$$q(t_i) = S(t_{i-1}) - S(t_i). \tag{4.6}$$

The default probabilities may have time dependent mean $\mu_q(t_i)$ and volatility $\sigma_q(t_i)$. As stated in the previous section, our approach only requires the values of the means and variances of the exposures and default probabilities. It doesn't rely on the assumption of distributions. It's totally safe to use any existing models to generate the default probabilities.

4.2 Wrong Way Risk and Correlation

To correlate the exposures and default probabilities, we use correlated W_t and Z_t to generate the underlying processes and denote

$$\rho_{WZ} \equiv \mathbf{E}\left[dW_t \cdot dZ_t\right]/dt$$

as the correlation coefficient of W_t and Z_t. ρ_{WZ} reflects the correlation between the exposure and the default probability.

ρ_{WZ} can take values in $[-1, 1]$ to reflect the different level of correlations between the exposures and the default probabilities. Given a level of the underlying correlation ρ_{WZ}, the robust correlation given in Eq. 3.11 can be calculated by simulations. In particular, for a given value of ρ_{WZ}, we generate N paths of the interest rate process r_t described by Eq. 4.1 or Eq. 4.2 and N paths of the default density process $h_t = \log X_t$ where X_t is given by Eq. 4.3. Then we can generate the exposures and default probabilities at each time point t_i. At each time point t_i, we calculate the means and variances of the exposures and the default probabilities, $\mu_V(t_i), \sigma_V(t_i), \mu_q(t_i), \sigma_q(t_i)$. Then the robust correlation $\bar{\rho}$ can be calculate as the following:

$$\bar{\rho} = \frac{\sum_{i=1}^{k} \rho(t_i)\sigma_V(t_i)\sigma_q(t_i)}{\sum_{i=1}^{k} \sigma_V(t_i)\sigma_q(t_i)}.$$

To calculate CVA using the full Monte Carlo simulation described above, we need to regenerate all paths each time when we consider a new level of ρ_{WZ}. This is very time consuming if one only cares about the CVA while ρ_{WZ} changes in a small range around the current value. On the other hand, in the formula (3.14), once we generate the exposures and the default probabilities for the independent case, we are ready to calculate the profile multiplier C_p, and the CVA corresponding to different level of ρ_{WZ} can be obtained by changing the robust correlation $\bar{\rho}$ without rerunning the full simulations. To guarantee the consistence, we need to establish the relationship between ρ_{WZ} and the robust correlation $\bar{\rho}$. This will be done in section 4.4. But first, we must determine N, the number of paths we need to simulate, to achieve the desired accuracy.

4.3 Full Simulation and Convergence

We derive exposures and default probabilities in the way introduced in the previous sections. Then we apply the full simulation approach given in equation Eq. 3.4 to calculate CVA. N paths of the interest rate r_t and the default density h_t are generated for a given level of ρ_{WZ}. For each path $j, j = 1, 2, \cdots, N$, we can calculate the exposure $V_j^+(t_i)$ and the default probability $q_j(t_i)$ at each time t_i. On the other hand, according to Eq. 3.4, the CVA is given by

$$CVA = (1-R)\sum_{i=1}^{k} \mathbf{E}[V^+(t_i)q(t_i)] = (1-R)\mathbf{E}\left[\sum_{i=1}^{k} V^+(t_i)q(t_i)\right]. \tag{4.7}$$

We define CVA_j as

$$CVA_j = (1-R)\sum_{i=1}^{k} V_j^+(t_i)q_j(t_i). \tag{4.8}$$

The sample mean of the simulated CVA is:

$$\overline{CVA}_N = \frac{1}{N}\sum_{j=1}^{N} CVA_j.$$

Apparently, by the Law of Large Numbers and Eq. 3.4, we know that

$$CVA = \mathbf{E}[\overline{CVA}_N]. \tag{4.9}$$

The sample standard deviation s_N is given by

$$s_N^2 \equiv \frac{1}{N-1}\sum_{j=1}^{N} \left(CVA_j - \overline{CVA}_N\right)^2.$$

Then the standard error e_N of the sample mean $\overline{CVA}_N$ can be estimated by

$$e_N^2 = \frac{s_N^2}{N}.$$

Since CVA may be sensitive to the input parameters, the standard error e_N of the sample mean $\overline{CVA}_N$ is not a good measure for convergence. Hence, for convergence study, we consider the coefficient of variation the sample mean $\overline{CVA}_N$:

$$\varepsilon_N = \frac{e_N}{\overline{CVA}_N} = \frac{s_N}{\sqrt{N}\,\overline{CVA}_N}. \tag{4.10}$$

N is set to be 10,000, 30,000, 50,000, 70,000 and 100,000 and we compute ε_N for both CIR and Vasicek models with all ρ_{WZ} values. For illustration purpose, Fig. 4.1 shows ε_N when we use CIR model and set ρ_{WZ} to be 0.5. We use a threshold of 0.01 or 1% to determine the number of paths needed. From Fig. 4.1, we can see that when the total number of paths is 100,000 the CVA curve is good enough ($\varepsilon_N < 0.01$) to be used as a benchmark. We also conduct convergence studies for different values of ρ_{WZ}, such as $0, 0.05.0.1, 0.15, \cdots, 1$, and the results are the same. For ρ_{WZ} taking negative values, the results are the same, too. Since we focus on the WWR, we mainly consider $\rho_{WZ} \in [0,1]$ in this paper. Next, we will use $N = 100,000$ and call it the full simulation if we use the procedure above to obtain CVA.

On the other hand, once we have generated the paths and obtained the exposures $V_j^+(t_i)$ and the default probability $q_j(t_i)$ at each time t_i, we can estimate CVA_{IND}, too. Recall that

$$CVA_{IND} = (1-R)\sum_{i=1}^{k} \mathbf{E}[V^+(t_i)]\mathbf{E}[q(t_i)]. \tag{4.11}$$

Then we can estimate CVA_{IND} by the following estimator:

$$\overline{CVA^0}_N = (1-R)\sum_{i=1}^{k}\left(\frac{1}{N}\sum_{j=1}^{N} V_j^+(t_i)\right)\left(\frac{1}{N}\sum_{j=1}^{N} q_j(t_i)\right). \tag{4.12}$$

Now we can derive the CVA_{ratio} by Eq. 3.9 and we denote it as CVA^S_{ratio}:

$$CVA^S_{ratio} \equiv \frac{\overline{CVA}_N}{\overline{CVA^0}_N}. \tag{4.13}$$

4.4 Robust Correlation and Efficient Curve Fitting

As we have seen in the last section, CVA_{IND} is relatively easy to obtain, but if the level of dependence ρ_{WZ} has changed, we have to redo the whole simulation process to get the CVA. Here we propose a more efficient way to derive CVA when there is WWR.

Since CVA_{IND} is already obtained, our goal is to obtain the CVA_{ratio}. First, by virtue of the equation for the profile multiplier C_p given by Eq. 3.12, we can obtain C_p after one full simulation and we do not need to re-do the simulation even the ρ_{WZ} changes. Secondly, according to Eq. 3.13, we only need to estimate $\bar{\rho}$ to obtain the CVA_{ratio}. Therefore, if we can identify stable mapping between ρ_{WZ} and $\bar{\rho}$, we will not need to redo the full simulation when ρ_{WZ} changes. We denote the CVA obtained by this formula as CVA^F_{ratio}.

Fig. 4.2 gives the relation between ρ_{WZ} and $\bar{\rho}$. As we can see, an exponential function may give us a very good fit. So we will propose an exponential function to map ρ_{WZ} to $\bar{\rho}$:

$$\bar{\rho} = a[\exp(b\rho_{WZ}) - 1] \tag{4.14}$$

where a and b are fitted parameters and $\hat{\rho}$ denotes the estimate of $\bar{\rho}$. So there are only 2 parameters to estimate.To fit the curve, we divide the interval $[0,1]$ into K small subintervals and define

$$\delta_\rho = \frac{1}{K}, \quad \rho^i_{WZ} = i\delta_\rho, \quad i = 1,2,\cdots,K.$$

For each given ρ^i_{WZ}, we run the simulations to generate the exposures and default probabilities to calculate $\bar{\rho}^i$ using formula (3.11). Then we fit the curve with the $\bar{\rho}_i$'s by estimating the values of a,b by least square method.

We use two measures to test the accuracy of our approach. First we consider the relative error of the ratio at each level of ρ_{WZ}, which is defined as

$$\varepsilon_{\rho_{WZ}} \equiv \left| \frac{CVA^F_{ratio} - CVA^S_{ratio}}{CVA^S_{ratio}} \right|.$$

It's easy to see that this relative error is also the relative error of CVA. We check the maximum $\varepsilon_{\rho_{WZ}}$ across all ρ_{WZ} levels to see if our approach is accurate enough. The other measure we use is the mean squared error (*MSE*) which is defined as

$$MSE \equiv \frac{1}{K}\sum_{i=1}^{K}\left[CVA^F_{ratio}(\rho^i_{WZ}) - CVA^S_{ratio}(\rho^i_{WZ})\right]^2.$$

As discussed in the previous section we calculate CVA_{IND} with 100,000 paths and derive $\mu_V(t_i)$, $\mu_q(t_i)$, $\sigma_V(t_i)$ and $\sigma_q(t_i)$. First we calculate the robust correlations for $K = 20$, i.e., at the points where ρ^i_{WZ} equals 0.05, 0.1, ..., 0.95, 1. At each point we generate 10,000, 30,000, 50,000 or 70,000 paths to test if we need more paths to fit a robust correlation curve that can be used in our approach. From Tables 4.1 and 4.2, we can see that 10,000 paths is good enough.

To reduce the computational burden, we further reduce the value of K from 20 to 10 and 5, i.e. ρ_{WZ} equals 0.1,0.2, ..., 1.0 or 0.1,0.3, 0.5, 0.7 and 0.9, respectively. The corresponding results are given in Tables 4.3 and 4.4. For comparison purpose, in Fig. 4.3 , the CVA ratio curves derived using the full simulation and using our approach are plotted. Obviously increasing the fitting points of ρ_{WZ} may not give us more benefit. So in this case, we conclude that 5 points of ρ_{WZ} between 0 and 1 is adequate to fit a robust correlation curve. Thus the total paths we generate are 150,000. But one can get the CVA with any ρ_{WZ} between 0 and 1.

Throughout the rest of this paper, we fit the robust correlation curve with 5 points of ρ_{WZ} and 10,000 paths at each ρ_{WZ}. In the next section, we will test the robustness of the approach and perform some sensitivity analysis and extreme scenarios analysis.

4.5 Sensitivity Analysis and Extreme Scenarios

To test the robustness of our method, we first consider small perturbations of the parameters and investigate the range of change for the robust correlation $\bar{\rho}$. Here we perturb all 3 interest process parameters, κ_r, θ_r and σ_r, by $\pm 10\%$ of their current level. We denote the current level as 'MMM'. A -10% level is denoted as 'L' and +10% as 'H'. Since we have 3 input parameters, there are 3^3 possible levels to examine. We want to see if it is possible to fit the robust correlation curve once for the 'MMM' level and reuse it for all other levels.With the fitted curve for level 'MMM', we calculate adjusted R-squares for all 27 levels to check the accuracy of fit. The statistics and measures of accuracy are in Table 4.5. From this table, we can see that the adjusted R^2 statistics are very close to 1, which means the fitted curve can represent all other levels very well. The accuracy measures indicate the fitted robust correlation can be used to calculate CVA and capture WWR with relative error below 4% in magnitude. Therefore, instead of regenerating $\rho(t_i)$ and calculating $\bar{\rho}$ for all levels, we can just use $\hat{\rho}$ fitted from level 'MMM' to represent all other levels. We can replace $\bar{\rho}$ with $\hat{\rho}$ and plug it to our formula in Eq. 3.13, CVA ratio curves for all 27 levels can be calculated easily without losing much accuracy. Such analysis can also be performed against parameters for X_t or h_t process. However, a company's credit quality information is usually updated quarterly or monthly and we assume that the X_t or h_t process is estimated less frequently unless there is a credit event. A credit event can trigger unwinding of transactions, which will definitely impact the CVA calculation. One should always re-fit the robust correlation in case of a credit event.

The previous experiments may not reveal which input parameter has greater impact on the CVA ratio. Also we need to know how sensitive $\bar{\rho}$ and C_p are with respect to each input parameter. We start from level 'MMM'. Every time we only perturb one input parameter of the interest rate process within a certain range. The range we use for κ_r, β_r and σ_r are $\kappa_r \in [0.05, 0.5], \beta_r \in [0.04, 0.4]$ and $\sigma_r \in [0.04, 0.4]$ respectively. These ranges are all equally spaced. Table 4.6 shows the sensitivity of robust $\bar{\rho}$ when ρ_{WZ} is 1. Fig. 4.4 and 4.5 show the value of C_p for all different parameter levels.

Table 4.6 contains the value of $\bar{\rho}$ when ρ_{WZ} is 1 where $\bar{\rho}$ differs most. We can see that except in the case a CIR model is assumed and σ_r changes, the maximum difference of $\bar{\rho}$ in percentage is below 6% and we don't have to re-fit our curve. This means for partial perturbation, the robust correlation can be reused without losing too much accuracy unless the volatility of CIR model changes a lot.

From Fig. 4.4 and 4.5 we observe similar trends for both CIR and Vasicek models. C_p decreases as κ_r or β_r increases. A larger κ_r draws the interest rate to its mean reverting level faster and this reduces the volatility of the exposure and the numerator of C_p given in Eq. 3.12. A larger β_r leads to a higher mean reverting level and increases the mean of exposure and the denominator of C_p. A larger σ_r means more volatile and increases the volatility of the exposure and the numerator of C_p. We view the jump of C_p in Fig. 4.5a as simulation error. The sensitivity analysis of C_p tells us that portfolios with different C_p should not be treated the same when we consider WWR. Holding everything else the same, a more volatile portfolio has more risk to be amplified and it needs a bigger adjustment factor. Note here that people don't have to worry about this when they use our model since C_p is derived every time when CVA_{IND} is calculated via full simulation. This will not add any computational effort.

In the real world the long term mean reverting level is usually stabler than the other two inputs, so we further examine two more extreme scenarios. Notice that for a given level of β_r, the interest rate process is less volatile when κ_r is large and σ_r is small. Our 'low volatile' scenario uses $\kappa_r = 0.5$, $\beta_r = 0.1$ and $\sigma_r = 0.04$ and 'high volatile' scenario takes $\kappa_r = 0.05$, $\beta_r = 0.1$ and $\sigma_r = 0.4$. We fit robust correlation curves under both scenarios and use our formula to derive the CVA ratio curve and check its accuracy. Fitted parameters, accuracy measures, C_p and maximum CVA ratio are given in Table 4.7.

As a result from the extreme scenario analysis the CVA ratio can vary from 2.7434 to 4.1460. Both ρ_{WZ} and C_p change a lot. See Table 4.7. Therefore in the situation where dramatic change of the underlying interest rate process, a reusable robust correlation curve is not guaranteed and thus one should re-fit the curve for the subsequent analysis.

4.6 Confidence Interval for CVA and CVA Ratio

The correlation between the exposures and the default probabilities ρ_{WZ}, plays a key role in the CVA calculation. However, it is not directly observable from the market and the risk manager has to estimate its value. The correlation may change from time to time, and the risk manager may just have a range of the correlation with certain confidence level. On the other hand, to estimate the CVA with a different correlation level ρ_{WZ}, the risk manager needs to run a full simulation. So to get a confidence interval for CVA, a risk manager may need to repeat the full simulation many times to get the distribution of the CVA to obtain the confidence level. On the other hand, in our approach, from the Eq. 3.14, we can see that, if the risk manager has a confidence interval about $\bar{\rho}$, she can get the confidence internal for CVA immediately. If she has a confidence interval for ρ_{WZ} instead of the robust correlation $\bar{\rho}$, by virtue of the monotone property of the mapping from ρ_{WZ} to $\bar{\rho}$, we can use the Eq. 3.14 and the values of $\bar{\rho}$ corresponding to the two ending points of the confidence interval for ρ_{WZ} to obtain the confidence interval for the CVA.

With our proposed method, we only need to run the simulations with computational efforts of $150,000$ paths in this particular study case to generate the whole CVA profile for $\rho_{WZ} \in [0,1]$. Not only it is easy to obtain the confidence interval, but it is also convenient to obtain the CVA for any correlation level very quickly. For a different level of ρ_{WZ}, we can use the formula (4.14) to obtain $\hat{\rho}$ then plug it into Eq. 3.13 to obtain the new CVA ratio and then use formula (3.14) to obtain the CVA under the new correlation level. No need to run any simulations for the new levels of ρ_{WZ}. In this way, a distribution of CVA ratio or CVA can be derived and a confidence interval is constructed. If we need 100 CVA values to construct the distribution of CVA or to construct a confidence interval, we at least reduce the computational burden by 98.5%.

Finally, although ρ_{WZ} must be specified to perform the simulations to describe the WWR, it is more nature to consider the robust correlation $\bar{\rho}$ instead of ρ_{WZ}. From our definition of $\bar{\rho}$, we can see that it describes the average correlation coefficient of the exposures and the default probabilities. So if a risk manager has a view on the correlation, it should be described by $\bar{\rho}$ instead of ρ_{WZ}. Further, it makes more sense to build a confidence level of CVA based on a confidence level of $\bar{\rho}$ instead of a confidence level of ρ_{WZ}.

4.7 Implementation Guide

Based our approach and numerical studies, we propose a step-by-step guide for CVA calculation. Consider a risk manager who needs to calculate CVA with WWR. We assume that the risk manager already have adapted her interest rate model and the default probability model, and she need to calculate the CVA under different correlation levels. In other words, we assume that the risk manager needs to calculate CVA under different level of ρ_{WZ} for $\rho_{WZ} \in [0,1]$. The step-by-step guide for CVA calculation is as follows:

Step 1. Run a full simulation (e.g. 100,000 paths) and use Eq. 4.11 to get CVA_{IND} and use Eq. 3.12 to calculate C_p;

Step 2. Choose a grid of 5 values of ρ^i_{WZ}, e.g. $0.2, 0.4, 0.6, 0.8, 1.0$ and for each of them calculate $\bar{\rho}_i$ using Eq. 3.11 with 10,000 paths;

Step 3. Fit a robust correlation curve given by Eq. 4.14;

Step 4. For any given value of $\rho_{WZ} \in [0,1]$, calculate $\bar{\rho}$ for ρ_{WZ} using the fitted curve.

Step 5. Apply CVA_{IND}, C_p and $\bar{\rho}$ into equation Eq. 3.14 to get CVA and its confidence interval;

Step 6. In the next period, the risk manager should have an updated view on $\bar{\rho}$ or ρ_{WZ}. If the robust criteria (to be given below) is satisfied, go to Step 4 to calculate a new CVA. Otherwise a full simulation has to be run, that is, return to step 1.

From the results shown in Section 4.5, the model can be used without a full simulation if one of the following robust criteria is satisfied:

i. The input parameters deviate from current level by no more than 10%;

ii. In the case of Vasicek model, if the parameters change in the following range: $\kappa_r \in [0.05, 0.5], \beta_r \in [0.04, 0.4]$ and $\sigma_r \in [0.04, 0.4]$.

iii. In the case of CIR model, if the parameters change in the following range: $\kappa_r \in [0.05, 0.5], \beta_r \in [0.04, 0.4]$ and σ_r doesn't change more than 10%.

From this step-by-step algorithm, we can see that if a risk manager obtains a confidence interval for the correlation ρ_{WZ}, she just needs to follow Step 4-6 for the maximum value and the minimum values of ρ_{WZ} to build the confidence interval for CVA. However this algorithm is not a recipe for any situations without caution. When extreme market scenarios or credit events with the counterparty occur, a full simulation needs to be performed to reestablish the foundation of the algorithm.

Table 4.1 Fitted parameters and accuracy with 20 ρ_{WZ} points and different number of paths - CIR model

Number of paths at each ρ_{WZ}	a	b	Adjusted R^2	MSE	Maximum $\varepsilon_{\rho_{WZ}}$
10,000	1.0272	0.6267	0.9999	0.0002	1.63%
30,000	1.0223	0.6273	0.9998	0.0001	1.69%
50,000	1.0085	0.6339	0.9999	0.0001	1.67%
70,000	1.0287	0.6245	0.9999	0.0001	1.70%

Table 4.2 Fitted parameters and accuracy with 20 ρ_{WZ} points and different number of paths - Vasicek model

Number of paths at each ρ_{WZ}	a	b	Adjusted R^2	MSE	Maximum $\varepsilon_{\rho_{WZ}}$
10,000	1.5091	0.4579	0.9999	0.0001	1.14%
30,000	1.5319	0.4524	0.9998	0.0001	1.21%
50,000	1.4717	0.4678	0.9999	0.0001	1.11%
70,000	1.4401	0.4762	0.9999	0.0001	1.05%

Table 4.3 Fitted parameters and accuracy with 10,000 paths and different number of ρ_{WZ} points - CIR model

Number of ρ_{WZ} points	a	b	Adjusted R^2	MSE	Maximum $\varepsilon_{\rho_{WZ}}$
5	1.0195	0.6321	0.9998	0.0002	1.44%
10	1.1040	0.5931	0.9998	0.0002	1.50%
20	1.0272	0.6267	0.9999	0.0002	1.63%

Table 4.4 Fitted parameters and accuracy with 10,000 paths and different number of ρ_{WZ} points - Vasicek model

Number of ρ_{WZ} points	a	b	Adjusted R^2	MSE	Maximum $\varepsilon_{\rho_{WZ}}$
5	1.3791	0.4934	0.9999	0.0001	1.16%
10	1.4522	0.4721	0.9999	0.0001	1.25%
20	1.5091	0.4579	0.9999	0.0001	1.14%

Table 4.5 Goodness of fit statistics and accuracy using the robust correlation curve fitted at level 'MMM'

Level	Adjusted R^2		MSE		Maximum $\varepsilon_{\rho_{WZ}}$	
	CIR	Vasicek	CIR	Vasicek	CIR	Vasicek
LLL	0.9998	0.9999	0.0003	0.0003	1.52%	1.72%
LLM	0.9996	0.9999	0.0004	0.0005	1.69%	2.01%
LLH	0.9996	0.9999	0.0001	0.0002	1.37%	1.64%
LML	0.9997	0.9999	0.0003	0.0009	1.95%	3.24%
LMM	0.9996	0.9999	0.0006	0.0001	2.14%	1.20%
LMH	0.9996	0.9999	0.0002	0.0003	1.43%	1.91%
LHL	0.9998	0.9999	0.0001	0.0001	1.41%	1.16%
LHM	0.9998	0.9999	0.0004	0.0005	2.07%	2.10%
LHH	0.9997	0.9999	0.0009	0.0013	2.03%	2.75%
MLL	0.9998	0.9999	0.0002	0.0004	1.51%	1.79%
MLM	0.9997	0.9999	0.0009	0.0004	2.43%	1.84%
MLH	0.9997	0.9999	0.0002	0.0001	1.89%	2.00%
MML	0.9999	0.9999	0.0001	0.0001	1.45%	1.34%
MMM	0.9998	0.9999	0.0002	0.0001	1.44%	1.16%
MMH	0.9997	0.9999	0.0003	0.0002	1.78%	1.59%
MHL	0.9999	0.9999	0.0006	0.0001	2.65%	2.30%
MHM	0.9998	0.9999	0.0001	0.0001	1.11%	1.09%
MHH	0.9997	0.9999	0.0003	0.0002	1.96%	1.58%
HLL	0.9999	0.9999	0.0002	0.0002	1.62%	1.61%
HLM	0.9998	0.9999	0.0003	0.0001	1.71%	1.38%
HLH	0.9998	0.9999	0.0002	0.0001	1.21%	1.19%
HML	0.9999	0.9999	0.0003	0.0003	1.47%	1.50%
HMM	0.9999	0.9999	0.0002	0.0001	1.41%	1.41%
HMH	0.9998	0.9999	0.0001	0.0001	1.14%	1.08%
HHL	0.9999	0.9999	0.0003	0.0001	2.21%	1.49%
HHM	0.9999	0.9999	0.0003	0.0002	1.77%	1.65%
HHH	0.9999	0.9999	0.0003	0.0002	2.14%	1.74%

Table 4.6 $\bar{\rho}$ ranges with perturbed input parameters

Input parameter	Parameter range	CIR			Vasicek		
		Min $\bar{\rho}$	Max $\bar{\rho}$	% change	Min $\bar{\rho}$	Max $\bar{\rho}$	% change
κ_r	0.05-0.5	0.8815	0.9337	5.92%	0.8690	0.9189	5.74%
β_r	0.04-0.4	0.8623	0.8896	3.17%	0.8630	0.8773	1.66%
σ_r	0.04-0.4	0.7808	0.8914	14.16%	0.8366	0.8800	5.19%

Table 4.7 Fitted parameters, accuracy measures, C_p and CVA ratio - extreme scenarios

Model	Scenario	a	b	Adj. R^2	MSE	Max. $\varepsilon_{\rho_{WZ}}$	C_p	Max. CVA ratio
CIR	Low Vol	1.1454	0.5980	0.9993	0.0003	2.00%	1.8595	2.7434
	High Vol	0.4109	1.0452	0.9982	0.0008	2.35%	4.1520	4.1460
Vasicek	Low Vol	1.4592	0.4921	0.9997	0.0004	2.47%	1.6303	2.5122
	High Vol	2.2641	0.3078	1.0000	0.0002	2.12%	1.8312	2.4943

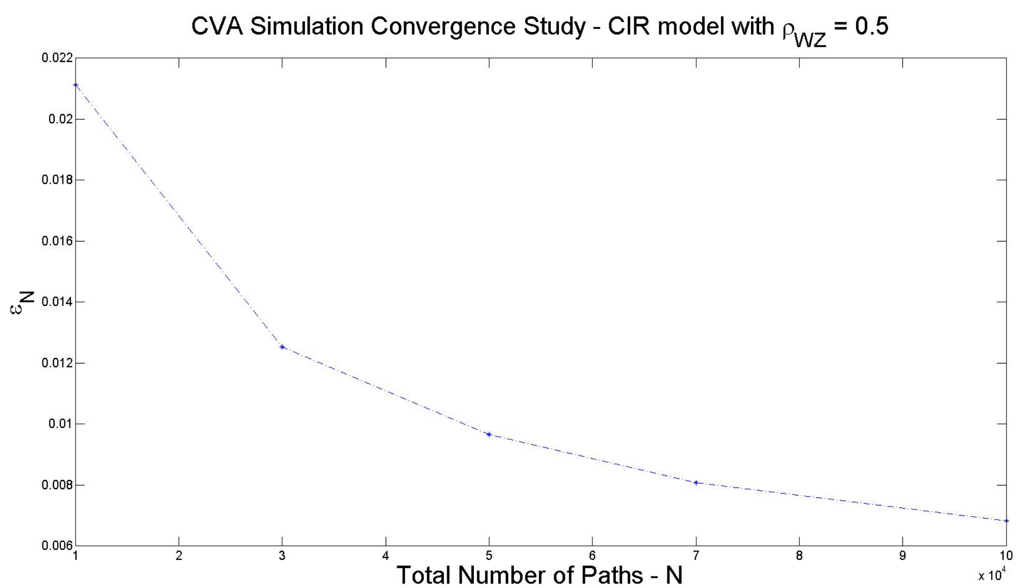
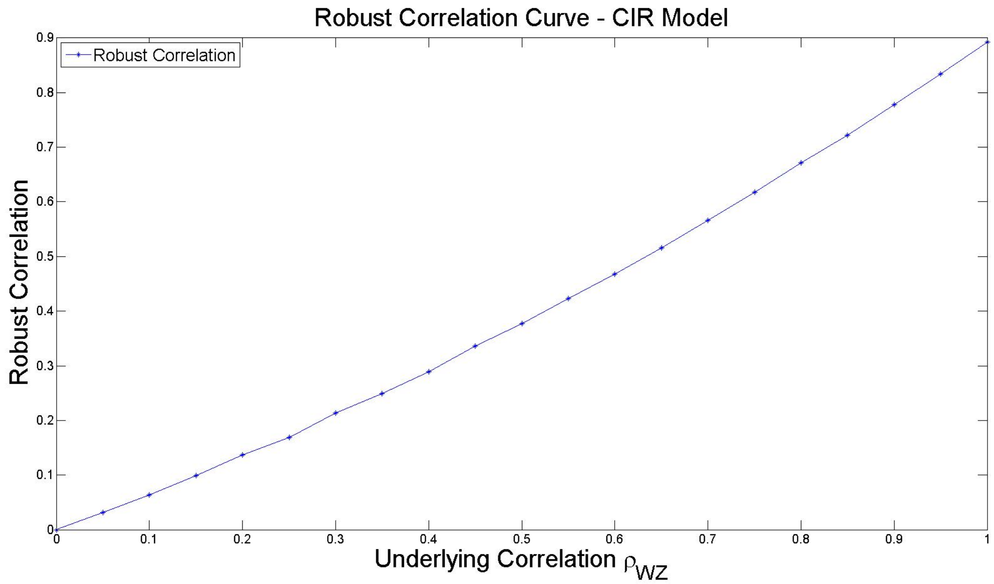
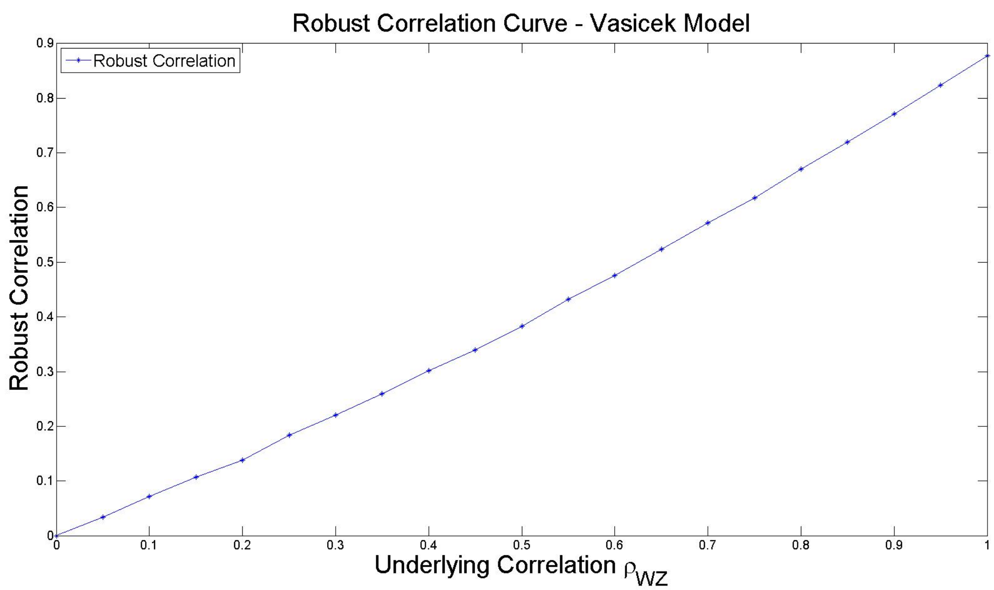


Figure. 4.1 *CVA simulation convergence study*

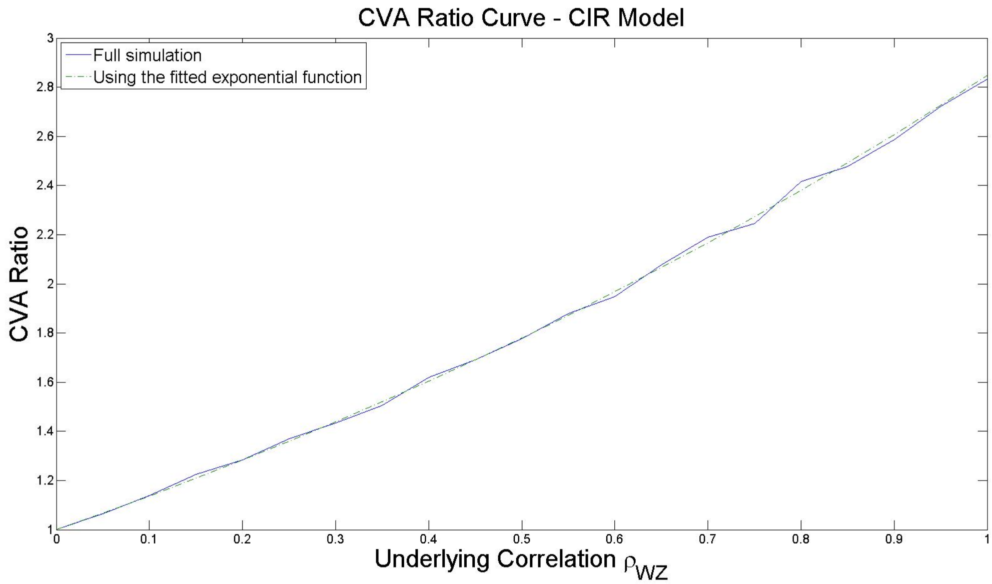


(a) *Robust correlation curve - CIR model*

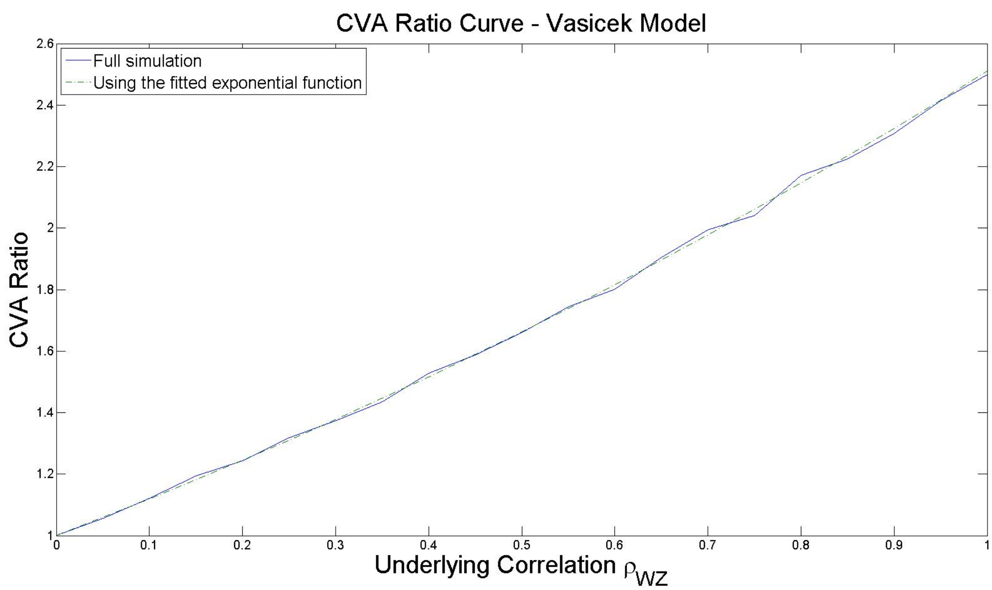


(b) *Robust correlation curve - Vasicek model*

Figure. 4.2 *Robust correlation curve*

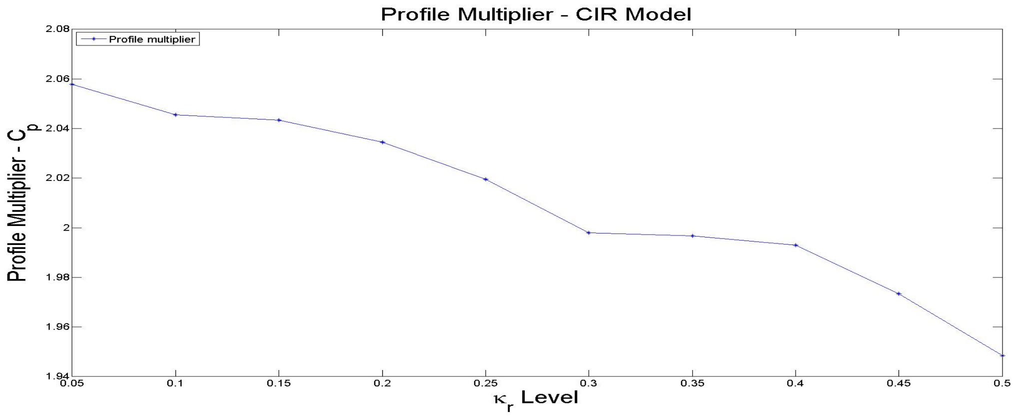


(a) *CVA ratio curve - CIR model*

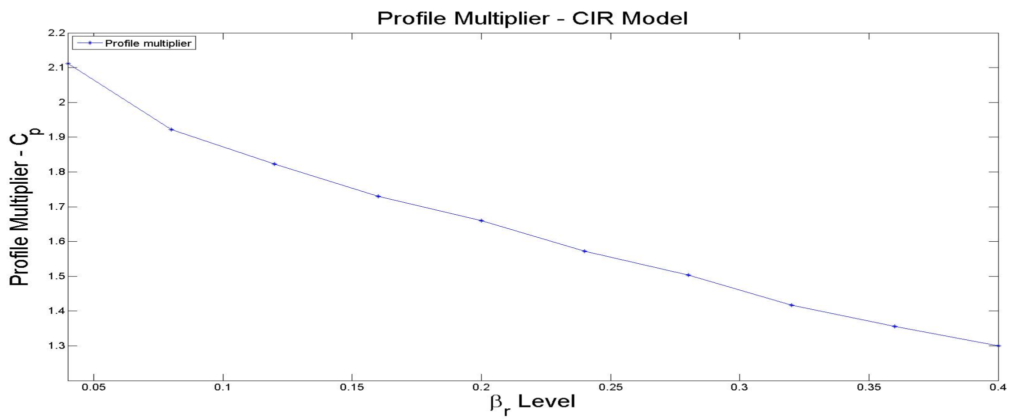


(b) *CVA ratio curve - Vasicek model*

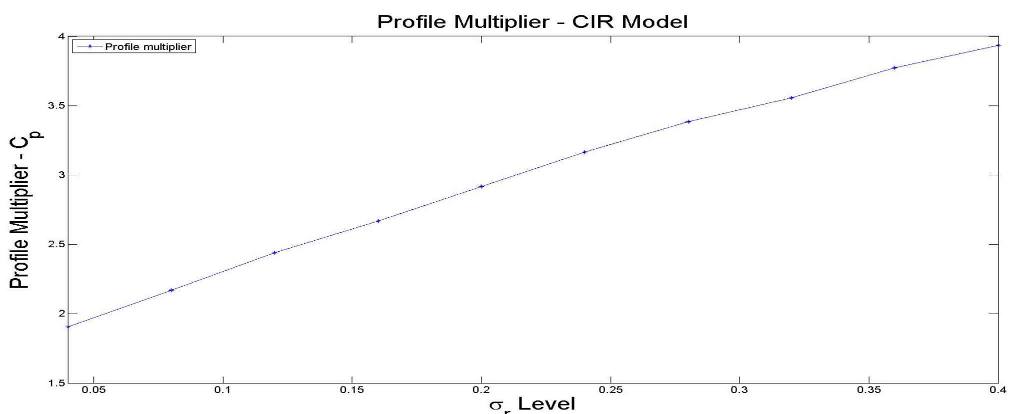
Figure. 4.3 *CVA ratio curve*



(a) *Profile multiplier with perturbed κ_r*

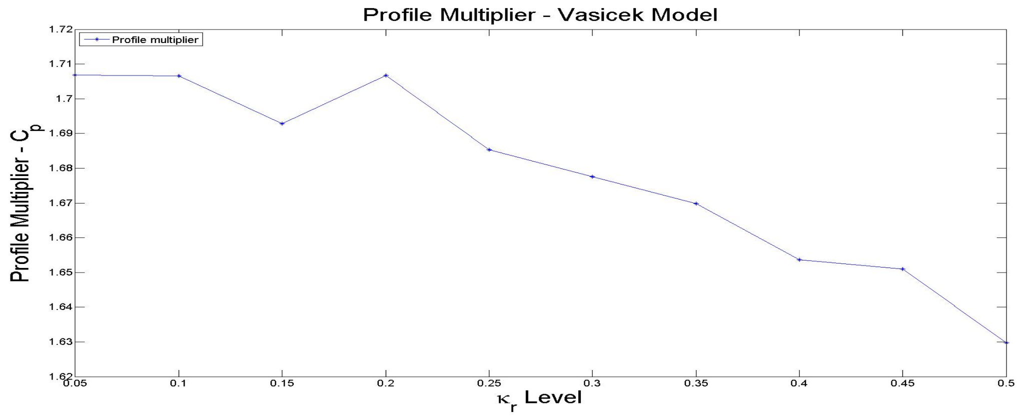


(b) *Profile multiplier with perturbed β_r*

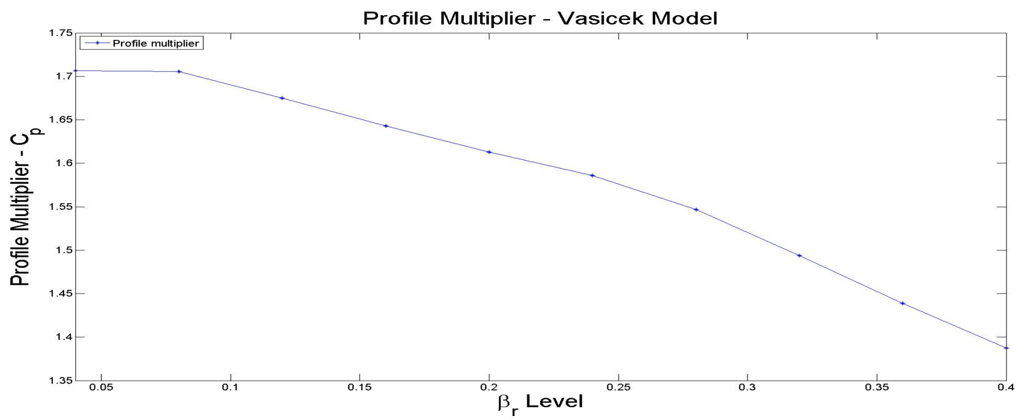


(c) *Profile multiplier with perturbed σ_r*

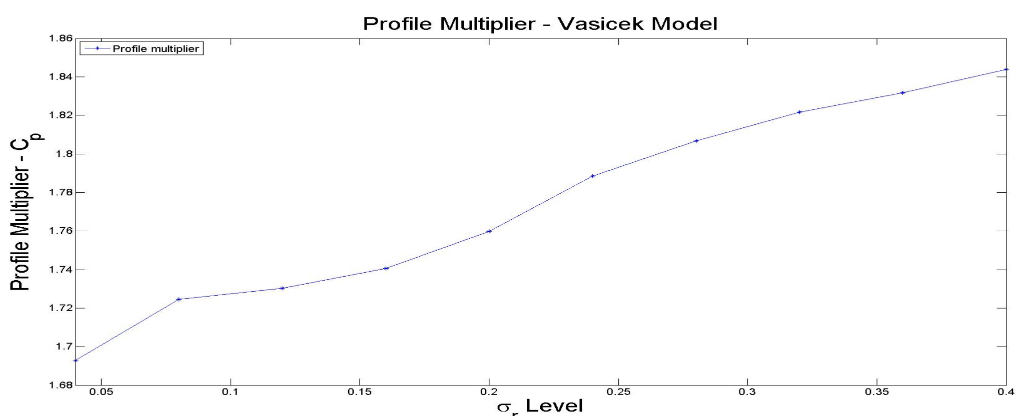
Figure. 4.4 *Results from perturbed input parameters - CIR model*



(a) *Profile multiplier with perturbed* κ_r

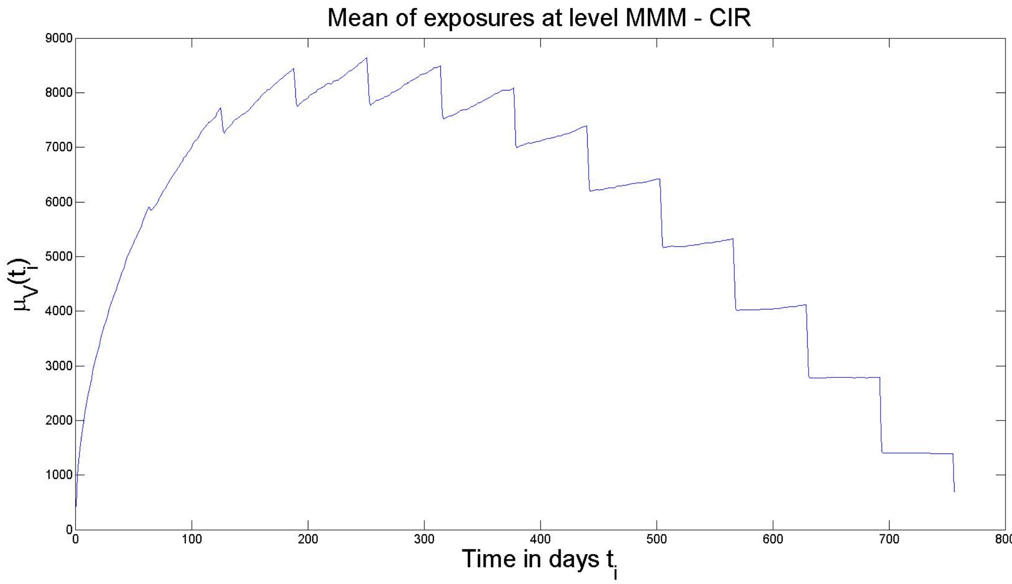


(b) *Profile multiplier with perturbed* β_r

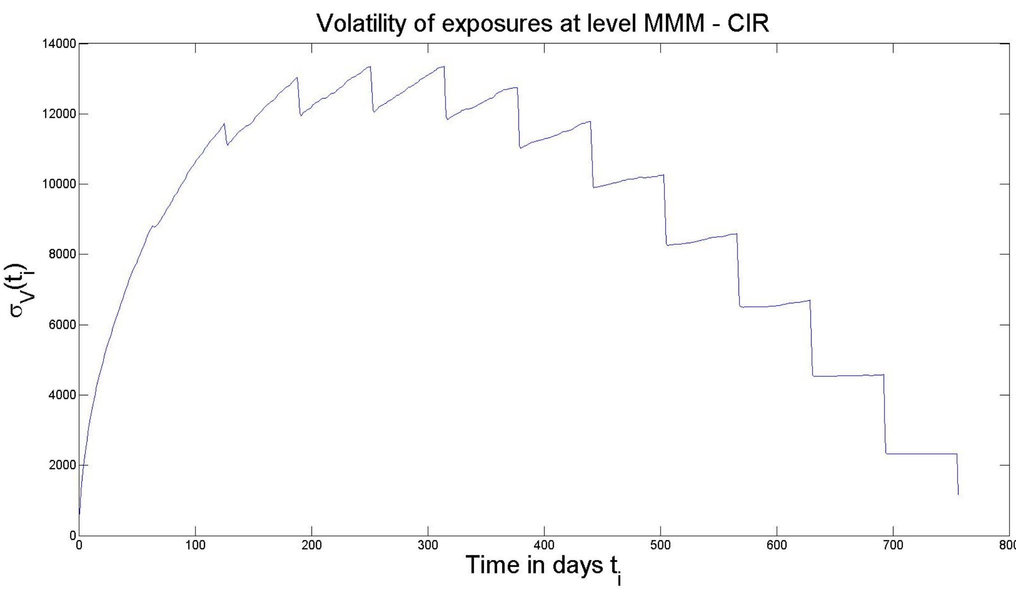


(c) *Profile multiplier with perturbed* σ_r

Figure. 4.5 *Results from perturbed input parameters - Vasicek model*

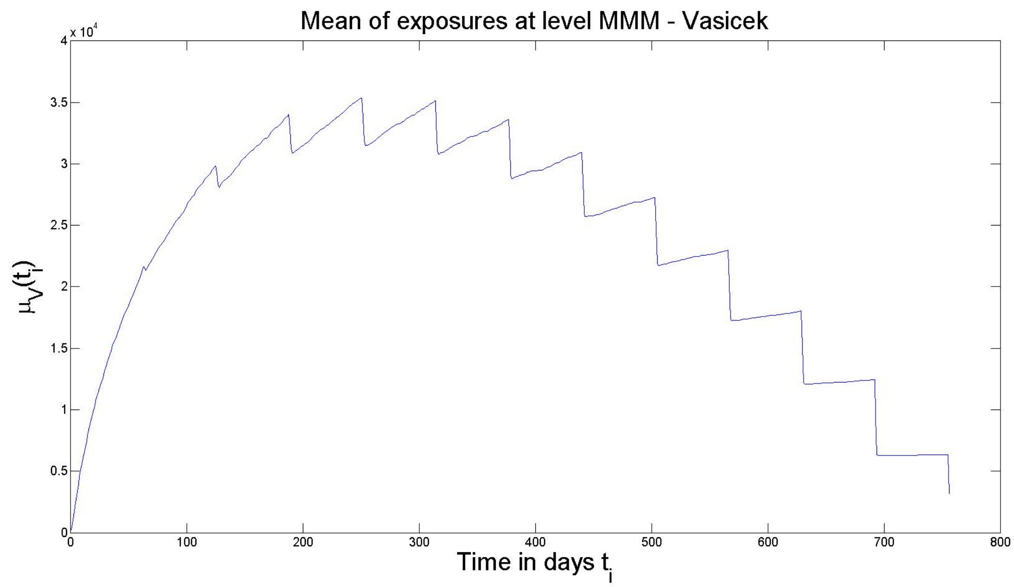


(a)

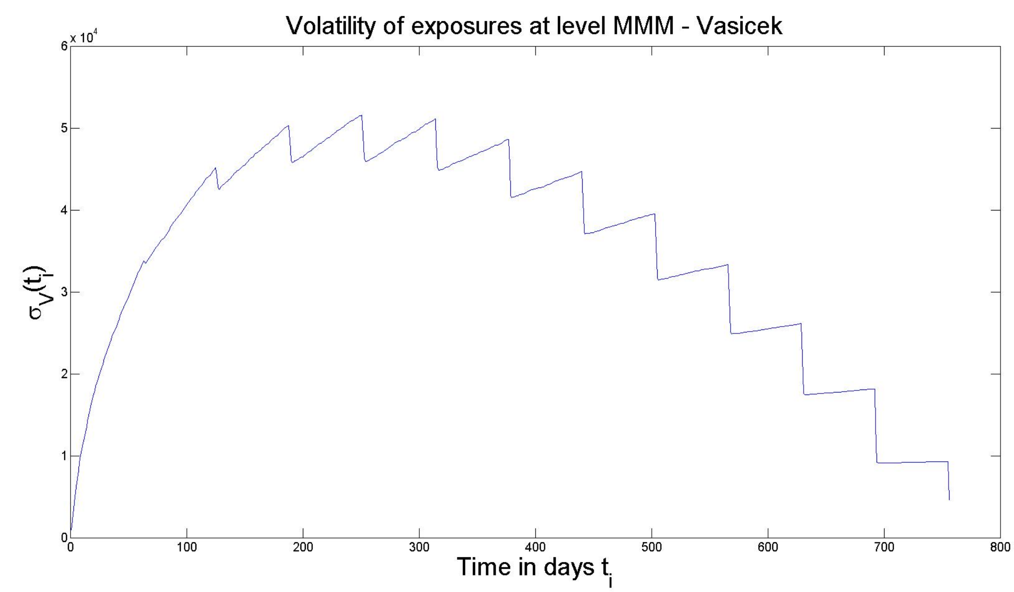


(b)

Figure. 4.6 *Mean and volatility of exposure at level MMM - CIR*

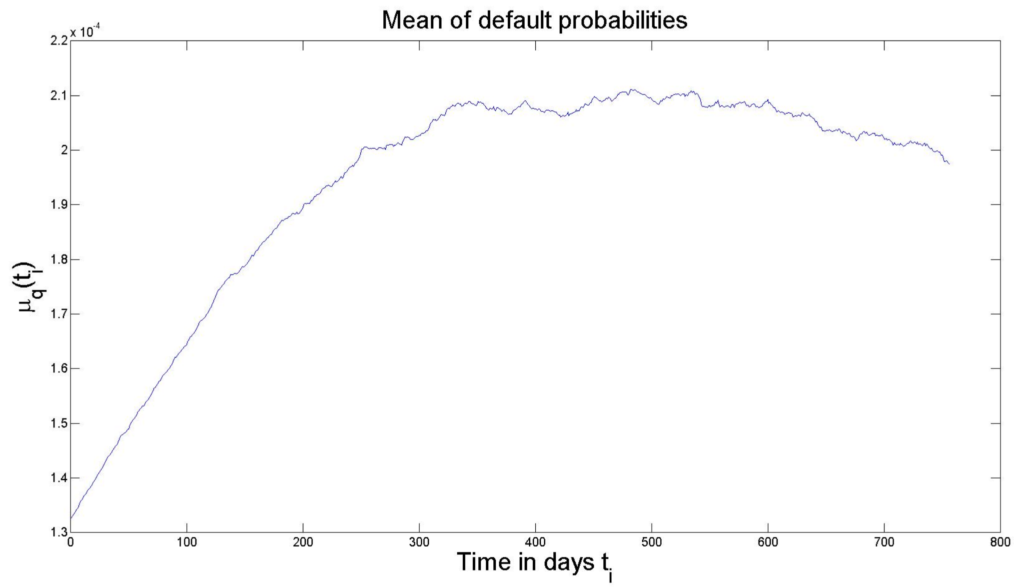


(a)

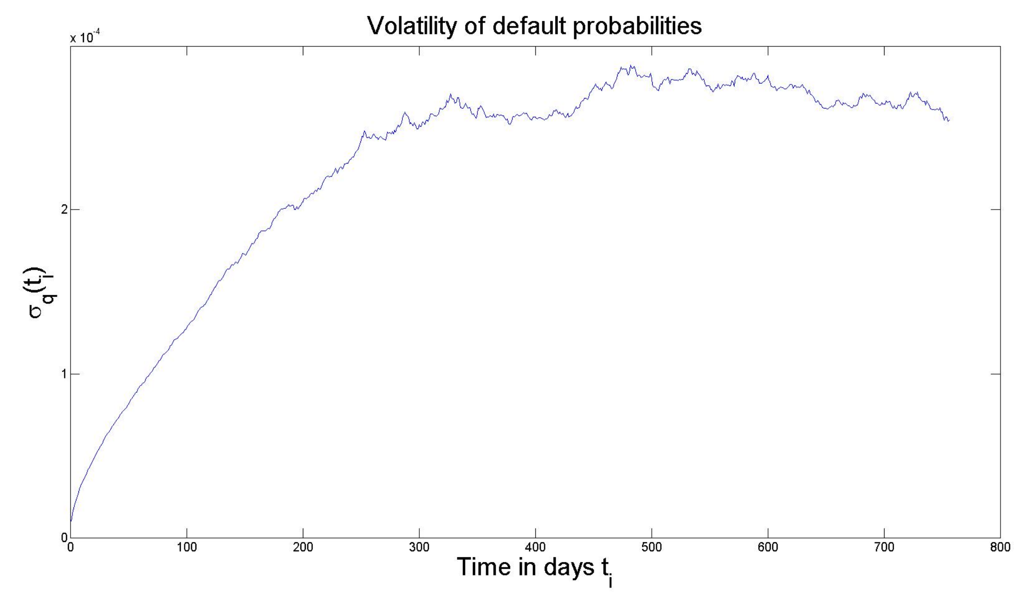


(b)

Figure. 4.7 *Mean and volatility of exposure at level MMM - Vasicek*

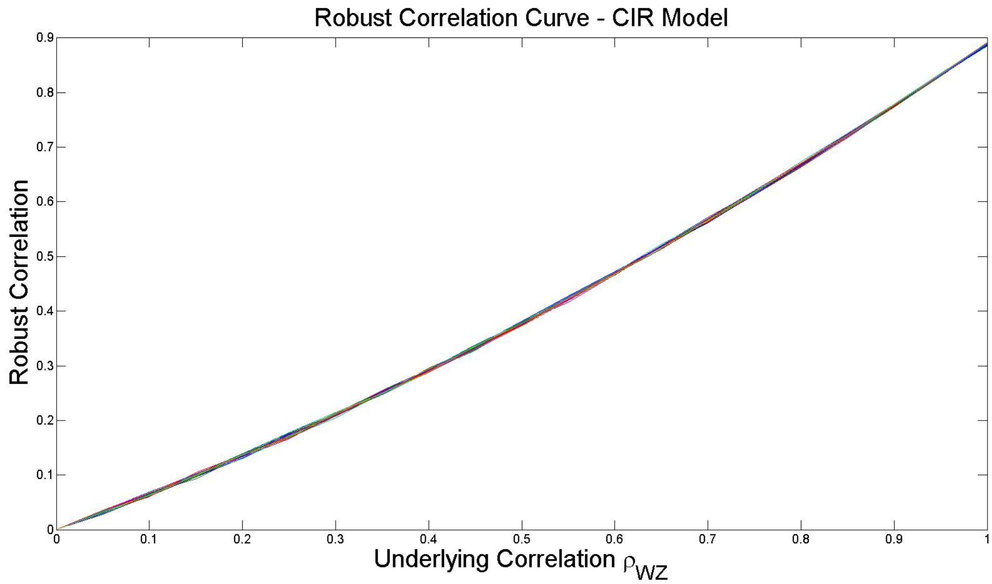


(a)

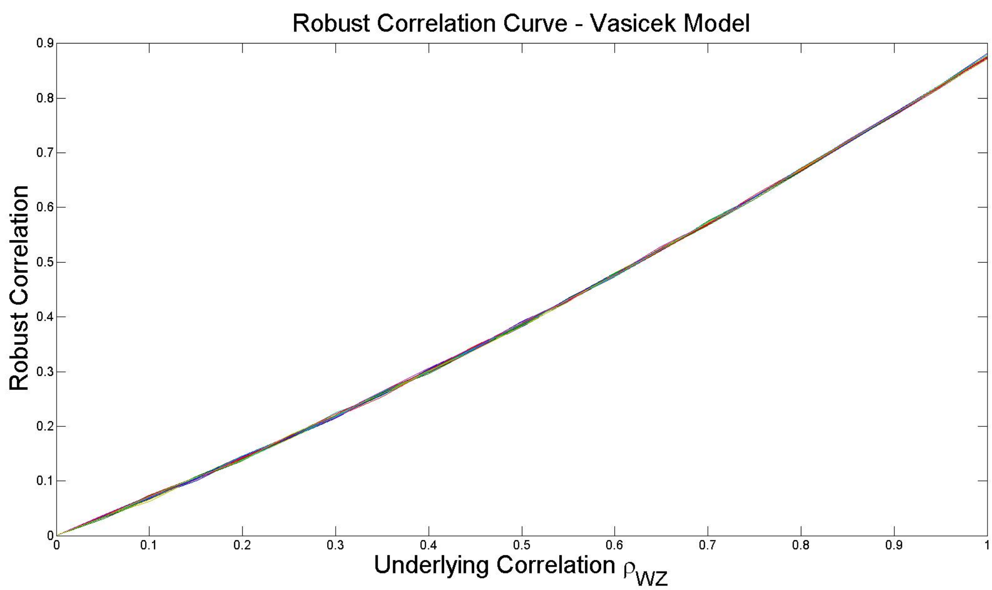


(b)

Figure. 4.8 *Mean and volatility of default probability*



(a) *All 27 levels of robust correlation* $\bar{\rho}$



(b) *All 27 levels of robust correlation* $\bar{\rho}$

Figure. 4.9 *Robust correlation*

Chapter 5

CVA DWR Multiplier Decomposition as A Bridge

5.1 The Hull and White Model

In the case of WWR, when the dealer's exposure is high the default probability of a counterparty is also high. RWR takes the opposite side. The usual approach models DWR by calculating conditional exposure. While Hull and White have proposed a different approach and do the reverse. They change the calculation of $q(t)$ so that the evolution of $q(t)$ is related to that of $V(t)$, where $V(t)$ can be the value of portfolio at time t.

Hull and White define the conditional default probability by linking the hazard rate, $h(t)$, to the underlying future value of the portfolio $V(t)$. Conditional on no earlier default, the probability of default in any small period Δt is $h\Delta t$. Viewing today as time 0, $\exp(-ht)$ is the probability of no default occurs before time t. If the hazard rate varies as a deterministic function of time then this no default probability is $\exp\left(-\int_0^t h(u)du\right)$.

Hazard rates can change stochastically and are not directly observable from the market. But credit spreads are observable and can be used as a good approximation. In their approach, the following relationship must be satisfied

$$\mathbf{E}\left[\exp\left(-\int_0^t h(u)du\right)\right] = \exp[-s(t)t], \tag{5.1}$$

where $s(t)$ is the counterparty credit spread with maturity t; 0 recovery rate is assumed.

In a Monte Carlo simulation, assuming the discretization is equally spaced, Eq. 5.1 is

$$\mathbf{E}\left[\exp(-\sum_{i=1}^{j} h_i \Delta t)\right] = \exp(-s_j t_j), \tag{5.2}$$

where h_i and s_j are $h(t_i)$ and $s(t_j)$ respectively.

Hull and White assume

$$h_i = g(V(t_i)) = \exp(a_i + bV(t_i)), \tag{5.3}$$

where b is a constant that measures the amount of DWR, a_i is a function of t_i that should be calibrated with Eq. 5.2. The detailed procedure of this calibration is given in Hull and White. Our discussion in the next section will be based on their model.

5.2 Analytical Results

In this section, we make some assumptions and derive some analytical results to have some insights of Hull and White model.

We assume the following assumptions hold:

1. There are K time periods;
2. $t_j = j\Delta t$ for $j = 1, 2, \ldots, K$ and $\Delta t = \frac{1}{252}$;
3. The profit and loss at t_j is denoted as X_j;
4. X_j follows normal distribution with mean μ_j and variance σ_j^2 and X_i and X_j are independent for $i \neq j$;

5. Suppose the initial value of the portfolio is a positive constant V_0;

6. The observed credit spread with maturity t_j is s_j;

7. t_0, X_0 and s_0 are all 0;

8. Recovery rate and discount rate are 0;

9. The exposure at time t_j is V_j and $V_j = V_0 + \sum_{i=1}^{j} X_i$;

10. For all j, the probability of V_j falls below 0 is negligible.

We define the hazard rate at time t_j as h_j. Following Hull and White framework, h_j is a function of V_j and

$$h_j = g(V_j) = \exp(a_j + bV_j), \tag{5.4}$$

where b is a predetermined constant and a_i is time dependent and should be calibrated so that

$$\mathbf{E}\left[\exp(-\sum_{i=1}^{j} h_i \Delta t)\right] = \exp(-s_j t_j)$$

We approximate the left-hand side of the above equation with

$$\mathbf{E}\left[\exp(-\sum_{i=1}^{j} h_i \Delta t)\right] \approx 1 - \mathbf{E}\left[\Delta t \sum_{i=1}^{j} h_i\right] \tag{5.5}$$

Then the market implied probability of default between time t_{j-1} and t_j, denoted as C_j, is

$$\begin{aligned} C_j &\equiv \exp(-s_{j-1}t_{j-1}) - \exp(-s_j t_j) \\ &= \mathbf{E}\left[\exp(-\sum_{i=1}^{j-1} h_i \Delta t)\right] - \mathbf{E}\left[\exp(-\sum_{i=1}^{j} h_i \Delta t)\right] \\ &\approx \mathbf{E}\left[\Delta t \sum_{i=1}^{j} h_i\right] - \mathbf{E}\left[\Delta t \sum_{i=1}^{j-1} h_i\right] \\ &= \mathbf{E}\left[h_j \Delta t\right] \end{aligned} \tag{5.6}$$

Hence C_j's can be expressed in terms of the expectation of hazard rate times Δt and we have

$$\mathbf{E}[h_j \Delta t] = C_j. \tag{5.7}$$

In Appendix B, we use some numerical examples to show the error of this approximation is acceptable.

From the definition of h_j,

$$\mathbf{E}[h_j \Delta t] = \exp(a_j - a_{j-1}) \mathbf{E}[h_{j-1} \Delta t \exp(bX_j)]. \tag{5.8}$$

From Eq. 5.8 one can see

$$\exp(a_j - a_{j-1}) \mathbf{E}[\exp(bX_j)] = \frac{C_j}{C_{j-1}}. \tag{5.9}$$

CVA with independent exposure and default probability can be expressed as

$$\begin{aligned}
CVA_{IND} &= \mathbf{E}\left[\sum_{j=1}^{K} V_j C_j\right] \\
&= \mathbf{E}\left[\sum_{j=1}^{K} C_j (V_0 + \sum_{i=1}^{j} X_i)\right] \\
&= \sum_{j=1}^{K} C_j \left(V_0 + \sum_{i=1}^{j} \mathbf{E} X_i\right) \\
&= \sum_{j=1}^{K} C_j \left(V_0 + \sum_{i=1}^{j} \mu_i\right).
\end{aligned} \tag{5.10}$$

CVA with DWR is

$$\begin{aligned}
CVA_{DWR} &= \sum_{j=1}^{K} \mathbf{E}\left[V_j h_j \Delta t\right] \\
&= \sum_{j=1}^{K} \mathbf{E}\left[\left(V_0 + \sum_{i=1}^{j} X_i\right) h_j \Delta t\right].
\end{aligned} \tag{5.11}$$

We show

$$\mathbf{E}\left[\left(V_0+\sum_{i=1}^{j}X_i\right)h_j\Delta t\right]=C_j\left[V_0+\sum_{i=1}^{j}(\mu_i+b\sigma_i^2)\right] \tag{5.12}$$

by induction.

Since V_0 is a constant and by Eq. 5.7, it is enough to show

$$\mathbf{E}\left[\sum_{i=1}^{j}X_ih_j\Delta t\right]=C_j\sum_{i=1}^{j}(\mu_i+b\sigma_i^2) \tag{5.13}$$

Stein's Lemma:

Suppose X follows N(μ,σ^2). Further assume g is a function for which both $\mathbf{E}\left[g(X)(X-\mu)\right]$ and $\mathbf{E}\left[g'(X)\right]$ exist. Then

$$\begin{aligned}\mathbf{E}[g(X)(X-\mu)]&=\sigma^2\mathbf{E}[g'(X)]\\ \mathbf{E}[g(X)X]&=\sigma^2\mathbf{E}[g'(X)]+\mu\mathbf{E}[g(X)].\end{aligned} \tag{5.14}$$

Following our assumptions and the definition of function g given by Eq. 5.4, we know the conditions of Stein's Lemma are satisfied.

First let's check when $j=1$

$$\begin{aligned}\mathbf{E}\left[X_1h_1\Delta t\right]&=\mathbf{E}\left[X_1\exp(a_1+bX_1)\Delta t\right]\\&=(\mu_1+b\sigma_1^2)\mathbf{E}[h_1\Delta t]\\&=(\mu_1+b\sigma_1^2)C_1.\end{aligned}$$

Eq. 5.13 holds. Next, let's assume Eq. 5.13 holds for $j=n-1$ and that is

$$\mathbf{E}\left[\sum_{i=1}^{n-1}X_ih_{n-1}\Delta t\right]=C_{n-1}\sum_{i=1}^{n-1}(\mu_i+b\sigma_i^2) \tag{5.15}$$

Then we move on to $j = n$

$$
\begin{aligned}
\mathbf{E}\left[\sum_{i=1}^{n} X_i h_n \Delta t\right] &= \mathbf{E}\left[(\sum_{i=1}^{n-1} X_i + X_n)\exp(a_n - a_{n-1})\exp(bX_n)h_{n-1}\Delta t\right] \\
&= \exp(a_n - a_{n-1})\mathbf{E}[\exp(bX_n)]\mathbf{E}\left[\sum_{i=1}^{n-1} X_i h_{n-1}\Delta t\right] \\
&+ \exp(a_n - a_{n-1})\mathbf{E}[X_n \exp(bX_n)]\mathbf{E}[h_{n-1}\Delta t] \\
&= \exp(a_n - a_{n-1})\mathbf{E}[\exp(bX_n)]\mathbf{E}\left[\sum_{i=1}^{n-1} X_i h_{n-1}\Delta t\right] \\
&+ \exp(a_n - a_{n-1})\mathbf{E}[\exp(bX_n)](\mu_n + b\sigma_n^2)\mathbf{E}[h_{n-1}\Delta t]
\end{aligned}
$$

Taking Eq. 5.7, Eq. 5.9 and Eq. 5.15 to the right-hand side of the above equation we have

$$
\begin{aligned}
\mathbf{E}\left[\sum_{i=1}^{n} X_i h_n \Delta t\right] &= C_n \sum_{i=1}^{n-1}(\mu_i + b\sigma_i^2) \\
&+ (\mu_n + b\sigma_n^2)\mathbf{E}[h_n \Delta t] \\
&= C_n \sum_{i=1}^{n-1}(\mu_i + b\sigma_i^2) \\
&+ C_n(\mu_n + b\sigma_n^2) \\
&= C_n \sum_{i=1}^{n}(\mu_i + b\sigma_i^2).
\end{aligned}
$$

Hence Eq. 5.13 holds for all j.

Thus CVA with DWR given by Eq. 5.11 can be expressed as

$$
CVA_{DWR} = \sum_{j=1}^{K} C_j \left[V_0 + \sum_{i=1}^{j}(\mu_i + b\sigma_i^2)\right] \tag{5.16}
$$

So the CVA ratio is given by

$$CVA_{ratio} \equiv \frac{CVA_{DWR}}{CVA_{IND}} = 1 + b \frac{\sum_{j=1}^{K} \left[C_j \sum_{i=1}^{j} \sigma_i^2 \right]}{\sum_{j=1}^{K} \left[C_j \left(V_0 + \sum_{i=1}^{j} \mu_i \right) \right]}. \tag{5.17}$$

The above equation shows that the CVA ratio is a function of b, moments of exposures and the observed credit spreads. It is irrelevant with a_j's. This result is not surprising since all a_j's are calibrated such that the expectation of default probabilities match those implied by the observed market credit spreads. Thus the information carried by a_j's should be embedded in b, moments of exposures and the observed market credit spreads. In other words a_j's should be functions of those factors.

Denote the means and variances of the exposure and default probabilities as follows

$$\begin{aligned} \mathbf{E}\left[V_j\right] &= \mu_V(t_j), \quad \mathbf{Var}\left(V_j\right) = \sigma_V^2(t_j), \\ \mathbf{E}\left[h_j \Delta t\right] &= \mu_{PD}(t_j), \quad \mathbf{Var}\left(h_j \Delta t\right) = \sigma_{PD}^2(t_j). \end{aligned}$$

With our assumptions, the followings hold

$$\begin{aligned} \mu_V(t_j) &= V_0 + \sum_{i=1}^{j} \mu_i, \\ \sigma_V^2(t_j) &= \sum_{i=1}^{j} \sigma_i^2, \\ \mu_{PD}(t_j) &= C_j, \\ \sigma_{PD}^2(t_j) &= C_j^2[\exp(b^2 \sum_{i=1}^{j} \sigma_i^2) - 1]. \end{aligned}$$

The first two equations above are relatively straight forward. The third one is a calibration condition in Hull and White approach. The derivation of the fourth one follows

$$\begin{aligned}\mathbf{E}\left[h_j\Delta t\right] &= \mathbf{E}\left[\exp(a_j)\exp[b(V_0+\sum_{i=1}^{j}X_i)]\Delta t\right] \\ &= \exp(a_j)\exp(bV_0)\exp\left(b\sum_{i=1}^{j}\mu_i+\frac{b^2\sum_{i=1}^{j}\sigma_i^2}{2}\right)\Delta t \\ &= C_j.\end{aligned} \tag{5.18}$$

Thus we have

$$\exp(a_j)\Delta t = C_j\exp(-bV_0)\left[\exp\left(b\sum_{i=1}^{j}\mu_i+\frac{b^2\sum_{i=1}^{j}\sigma_i^2}{2}\right)\right]^{-1}. \tag{5.19}$$

Next, let's derive the second moment of $h_j\Delta t$

$$\begin{aligned}\mathbf{E}\left[(h_j\Delta t)^2\right] &= (\Delta t)^2\mathbf{E}\left[\exp(2a_j)\exp[2b(V_0+\sum_{i=1}^{j}X_i)]\right] \\ &= \exp(2a_j)(\Delta t)^2\exp(2bV_0)\exp\left(2b\sum_{i=1}^{j}\mu_i+2b^2\sum_{i=1}^{j}\sigma_i^2\right).\end{aligned} \tag{5.20}$$

Taking Eq. 5.19 into Eq. 5.20

$$\mathbf{E}\left[(h_j\Delta t)^2\right] = C_j^2\exp(b^2\sum_{i=1}^{j}\sigma_i^2). \tag{5.21}$$

Hence

$$\sigma_{PD}^2(t_j) = \mathbf{E}\left[(h_j\Delta t)^2\right]-\left(\mathbf{E}\left[h_j\Delta t\right]\right)^2 = C_j^2\left[\exp(b^2\sum_{i=1}^{j}\sigma_i^2)-1\right].$$

We now consider the correlation coefficient of V_j and h_j directly, namely $\rho(t_j)$. For each time node, we have

$$\rho(t_j) = \frac{Cov(V_j,h_j\Delta t)}{\sigma_V(t_j)\sigma_{PD}(t_j)} \tag{5.22}$$

By Stein's Lemma,

$$\begin{aligned}
Cov(V_j, h_j \Delta t) &= \mathbf{E}\left[g'(V_j)\Delta t\right] \sum_{i=1}^{j} \sigma_i^2 \\
&= b\mathbf{E}\left[h_j \Delta t\right] \sum_{i=1}^{j} \sigma_i^2 \\
&= C_j b \sigma_i^2.
\end{aligned}$$

Hence

$$\begin{aligned}
\rho(t_j) &= \frac{C_j b \sum_{i=1}^{j} \sigma_i^2}{C_j \sqrt{\exp(b^2 \sum_{i=1}^{j} \sigma_i^2) - 1} \sqrt{\sum_{i=1}^{j} \sigma_i^2}} \\
&= \frac{b \sqrt{\sum_{i=1}^{j} \sigma_i^2}}{\sqrt{\exp(b^2 \sum_{i=1}^{j} \sigma_i^2) - 1}}.
\end{aligned} \tag{5.23}$$

In order to claim the magnitude of $\rho(t_j)$ decreases as b increases in magnitude, it is enough to show this is true when $b > 0$ due to the symmetry of the function above. We first set

$$x = b \sqrt{\sum_{j=1}^{K} \sigma_j^2} > 0.$$

Eq. 5.23 becomes

$$\begin{aligned}
\rho(t_j) &= \frac{x}{\sqrt{\exp(x^2) - 1}} \\
\rho^2(t_j) &= \frac{x^2}{\exp(x^2) - 1}
\end{aligned}$$

One can see that both the numerator and denominator of the equation above are 0 when $x = 0$ and the first order derivative of the numerator is less than that of the denominator when $x > 0$. So $\rho(t_j)$ decreases in magnitude as b increases in magnitude when $b > 0$.

Next we want to show $\rho(t_j)$ takes value in the interval of $(-1, 1)$. Using the expression of

$\rho^2(t_j)$ given above and denote $u = x^2$

$$\lim_{u \to 0} \frac{u}{\exp(u) - 1} \stackrel{\text{H}}{=} \lim_{u \to 0} \frac{1}{\exp(u)}$$
$$= 1.$$

So when $b > 0$, $x > 0$ and the limit of $\rho(t_j)$ is 1. When $b < 0$, x approaches 0 from left and the limit of $\rho(t_j)$ is -1.

In the case when defaults are independent with exposures or $b = 0$. The hazard rates are constant over all simulation paths for all time nodes t_j's. So they have 0 variance and covariance with exposure. $\rho(t_j)$ should also be 0.

Let's see how the expression given by Eq. 5.17 can be interpreted in terms of CVA DWR multiplier decomposition given in Pang, Chen and Li [34]. They discussed the use of CVA multiplier decomposition to gain insights for WWR. Actually the formulas also hold for RWR. The followings are defined

$$CVA_{ratio} = 1 + \frac{\sum_{j=1}^{K} \rho(j_i)\sigma_V(t_j)\sigma_{PD}(t_j)}{\sum_{j=1}^{K} \mu_V(t_j)\mu_{PD}(t_j)}, \tag{5.24}$$

$$\bar{\rho} \equiv \frac{\sum_{j=1}^{K} \rho(t_j)\sigma_V(t_j)\sigma_{PD}(t_j)}{\sum_{j=1}^{K} \sigma_V(t_j)\sigma_{PD}(t_j)}, \tag{5.25}$$

and

$$C_p \equiv \frac{\sum_{j=1}^{K} \sigma_V(t_j)\sigma_{PD}(t_j)}{\sum_{j=1}^{K} \mu_V(t_j)\mu_{PD}(t_j)}. \tag{5.26}$$

$\bar{\rho}$ and C_p are called *robust correlation* and *profile multiplier* respectively. It is easy to see that $0 \le \bar{\rho} < 1$. C_p describes the profiles of exposure and default probability.

Now the Eq. 5.17 can be written as

$$CVA_{ratio} = 1 + \bar{\rho} C_p. \tag{5.27}$$

and CVA can be expressed as

$$CVA = (1 + \bar{\rho} C_p) CVA_{IND}. \tag{5.28}$$

We claim that for given $\sigma_V(t_j)$ and $\sigma_{PD}(t_j)$, the magnitude of $\bar{\rho}$ decreases as b increases in magnitude. It's enough to show this hold for $\rho(t_j)$ since $\bar{\rho}$ is a weighted average of $\rho(t_j)$ with all positive weights.

Taking $\mu_V(t_j)$, $\mu_{PD}(t_j)$, $\sigma_V(t_j)$ and $\sigma_{PD}(t_j)$ into Eq. 5.26

$$C_p = \frac{\sum_{j=1}^{K} \left[C_j \sqrt{\sum_{i=1}^{j} \sigma_i^2} \sqrt{\exp(b^2 \sum_{i=1}^{j} \sigma_i^2) - 1} \right]}{\sum_{j=1}^{K} \left[C_j \left(\sum_{i=1}^{j} \mu_i + V_0 \right) \right]}.$$

So C_p increases of as b increases in magnitude.

5.3 Discussion of Our Results

With CVA DWR multiplier decomposition, we can see that the parameter b in Hull and White approach plays two roles. It amplifies the profile multiplier meanwhile reduces robust correlation. The increase of profile multiplier dominates the decrease of robust correlation. Hence the combined effect is increasing in b.

If we compare parametric approach with correlation approach, we can see that parametric approach tends to consider the combined effect as a whole. While the latter one can better describe robust correlation and profile multiplier separately since profile multiplier is only a function of the mean and volatility of the exposures and default probabilities and robust correlation is more sensitive to the correlation of the underlying risk drivers.

In almost all the uncollateralized cases, the direction of DWR is not hard to decide. It depends on the nature of the trade and the position a bank takes. If a risk manager has a good estimate of the change of default probability relative to the change of the portfolio exposure but not the level of correlation, a parametric approach is more appropriate.

If the estimates of underlying risk drivers are believed to be of high accuracy, then a correlation approach is better. Even when the risk manager only has a confidence interval of the underlying or robust correlation, a corresponding confidence interval of CVA can be built.

From both of these two types of models, we conclude that both the robust correlation and the profile multiplier play key roles in DWR. Aiming only on either of them will not be adequate. For Hull and White approach we want to emphasis the importance of profile multiplier. If the portfolio exposure tends to be more volatile or market CDS spread volatility of a counterparty tends to increase, we should be more careful about DWR even though the correlation is still unchanged.

With our results in this paper, one can better explain a phenomenon shown with Figure 9 in Ruiz, Boca and Pachòn [38]. That is CVA price decreases as the underlying volatility increases in the presence of RWR. Given the denominator of profile multiplier unchanged, a bigger numerator would increase the effect of RWR and reduce the CVA price. So CVA price is decreasing in the volatility of the underlying (oil price). For CVA Vega, the same amount of change of underlying volatility has relatively smaller impact on profile multiplier when the numerator is bigger.

5.4 Numerical Study

A series of numerical analyses in the context of vanilla interest rate swaps will be performed in this section.

5.4.1 Simulation Models

Neither Hull and White approach nor CVA DWR multiplier decomposition requires any distribution assumption. For illustration purposes we will use the same set of exposures that are used for 'MMM' level in Chapter 4. DWR dependency is modeled in the spirit of Hull and White approach. Parameter b ranges from -0.4 to 0.4.

5.4.2 Numerical Results

From Fig. 5.1 and 5.3, we can see that the profile multiplier increases as the magnitude of b increases. This is true for both RWR and WWR. This result also holds no matter we use CIR or Vasicek model. In the presence of RWR, the magnitude of profile multiplier ranges from 0 to a number slightly over 1. On the other hand, if WWR is in place this number could be almost 50 for CIR model and 100 for Vasicek model. This means WWR can be more amplified given the same correlation magnitude. The huge difference between RWR and WWR is not surprising since the exposure is truncated below.

An interesting relation we observe is the absolute value of robust correlation decreases as b increases in magnitude. That's to say a stronger codependent parameter lends to weaker correlation. We consider this a property of Hull and White approach. The volatility of default probabilities increases faster than the covariance of exposures and default probabilities.

Finally, the combined effect or CVA ratio is increasing in b. This is also expected for both RWR and WWR. It's quite intuitive to see if a trade faces WWR and the default probability of a counterparty is more volatile, then the potential credit loss is bigger.

On the other hand, if there is RWR it seems that a trade may benefit from a more volatile counterparty. This can be counter-intuitive at the first glance. Suppose a portfolio manager can make deals with several counterparties to accomplish her goal. Given the same level of expected default probability, a counterparty with less volatile default probability is considered to have better quality. Then it's not making sense to choose the ones with worse quality. This is true only when counterparties are viewed standalone. However, this will ignore the RWR effect on CVA value. A negative co-dependence can indeed reduce the potential credit loss. This is just like portfolio optimization. It's not always a bad idea to include a more risky asset due to its negative correlation to current assets in the portfolio.

5.5 Conclusion and Further Discussion

CVA DWR multiplier decomposition can build a bridge linking different approaches. With our analytical and numerical results one can see the following:

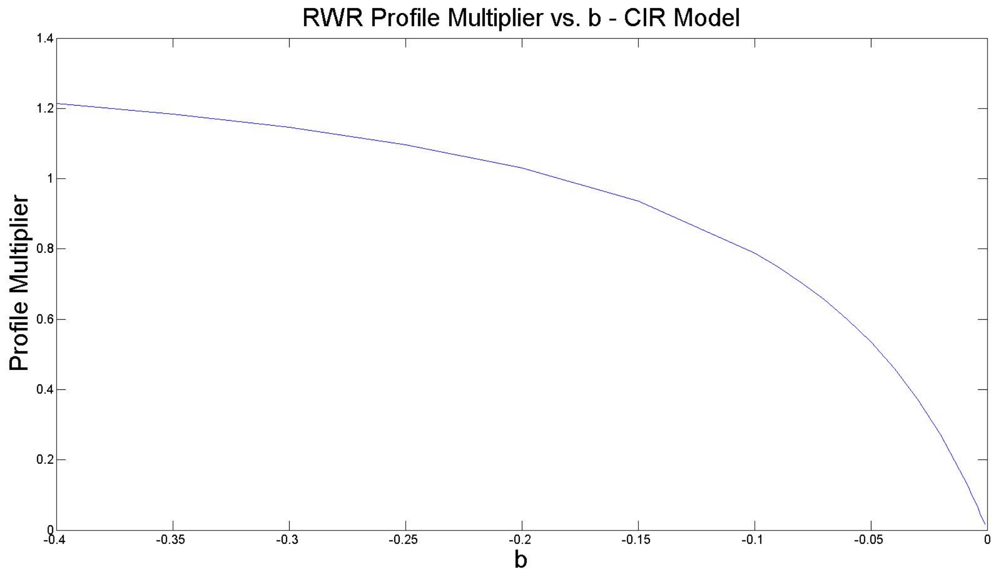
- The parameter b in Hull and White model can impact both robust correlation and profile multiplier given by Pang, Chen and Li [34] at the same time;
- With the same level of b in magnitude, b tends to have much more amplification effect on the WWR side than the RWR;
- Better insights of RWR are achieved by examining how underlying volatility affects the profile multiplier;
- The CVA ratio sensitivity relative to b is steeper when b is small in magnitude.

The parameter b in Hull and White is calibrated as the change of CDS spread relative to the change of portfolio exposures. It's intuitive to see the profile multiplier is closely related to the level of b and has such monotonicity shown in Fig. 5.1 and 5.3.

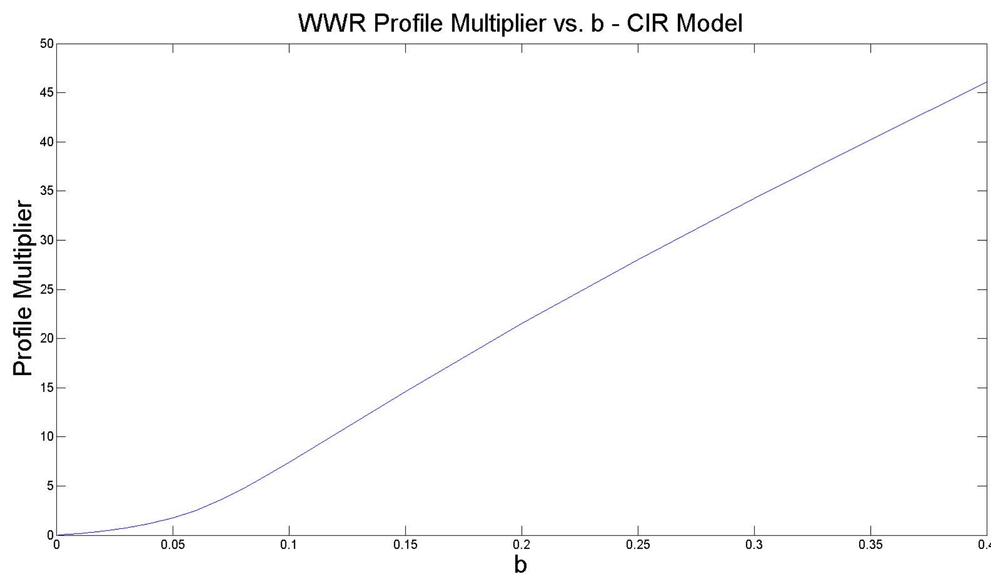
When the amplification effect on the WWR side is compared to the one on the RWR side, it's clear that the profile multiplier is a lot bigger when there is WWR. This is a very nice property from the standpoint of risk management. The alert is hard to neglect. Even when the linear correlation of the exposure and default probability is small, the whole ratio can be large enough to have risk managers to keep a close look at what's happening. In some cases, market tends to get more volatile much earlier than one notices the change in the correlation between credit risk driver and underlying risk factor.

In the case one tries to gain insights of CVA with RWR, the CVA DWR decomposition offers a big hand. It's counterintuitive to see the CVA price decreases as b increases in magnitude when there is RWR. However, from the decomposition one can easily see why this happens. Holding the mean of the underlying as a constant and bumping up its volatility will increase the profile multiplier. Since profile multiplier has more impact on the CVA ratio and robust correlation is negative when there is RWR, the CVA price is for sure to be smaller.

Another important finding is that when b is close to 0, one should be more careful. The CVA ratio is quite sensitive to the estimates of b when b lies closely around 0. This is quite like the behavior of options that are at the money. Delta is quite sensitive to the price change of the underlying when option is at the money. A bad estimate of b around 0 can be misleading. So when Hull and White model is used and b is small in magnitude, a risk manager may want to try different values of b and see how sensitive the CVA price is relative to b.

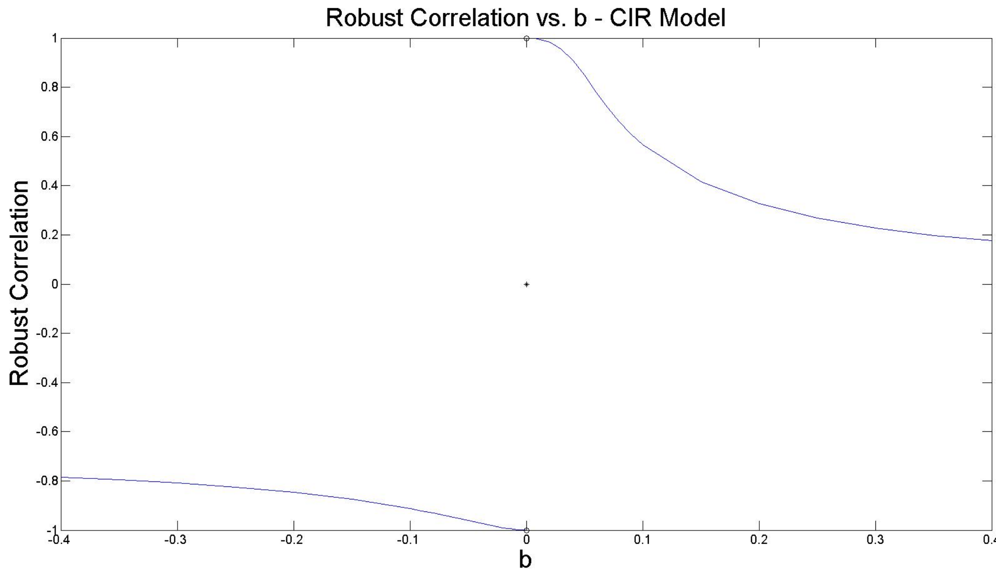


(a) *RWR profile multiplier - CIR model*

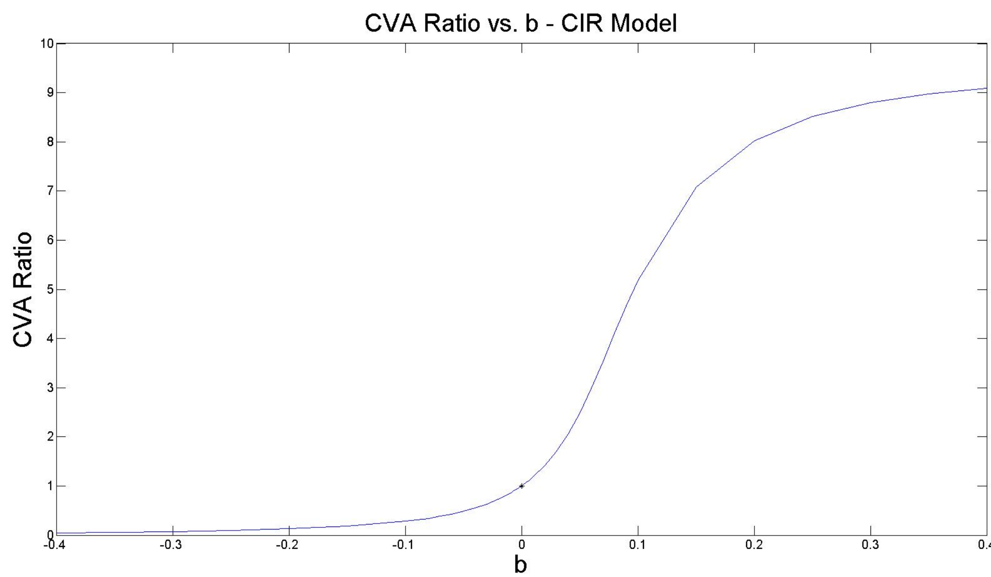


(b) *WWR profile multiplier - CIR model*

Figure. 5.1 *Profile multiplier - CIR model*

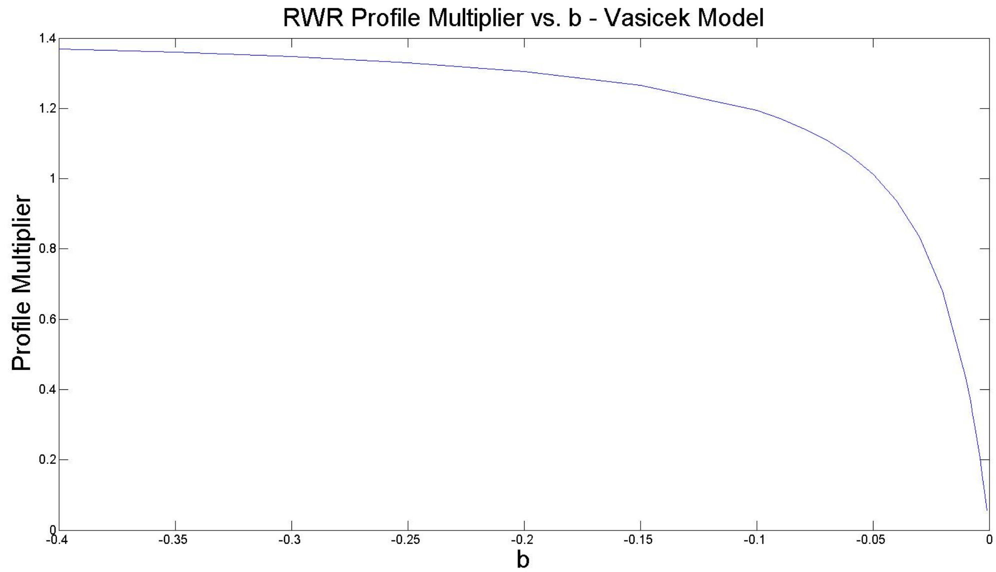


(a) *Robust correlation - CIR model*

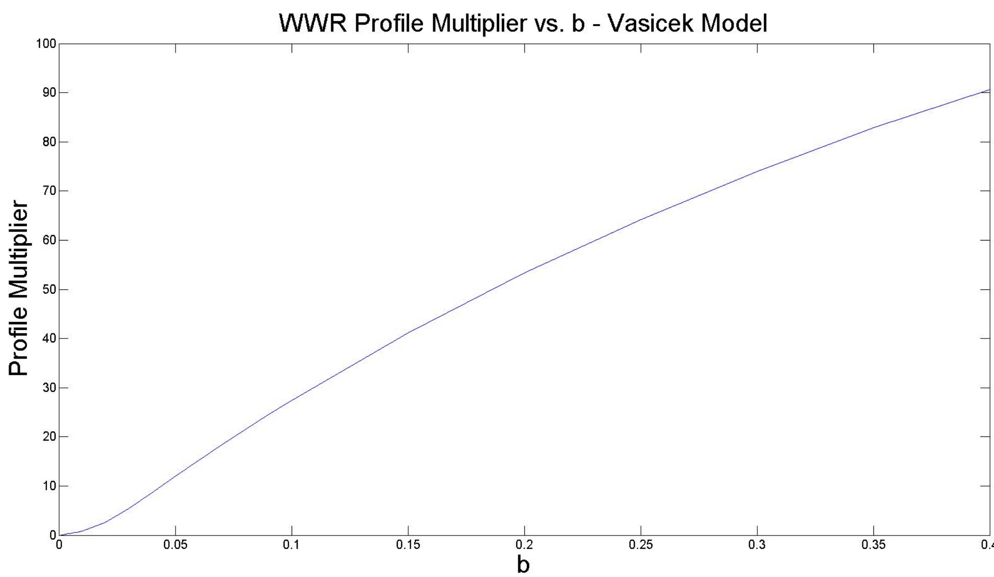


(b) *CVA ratio - CIR model*

Figure. 5.2 $\bar{\rho}$ *and CVA ratio - CIR model*

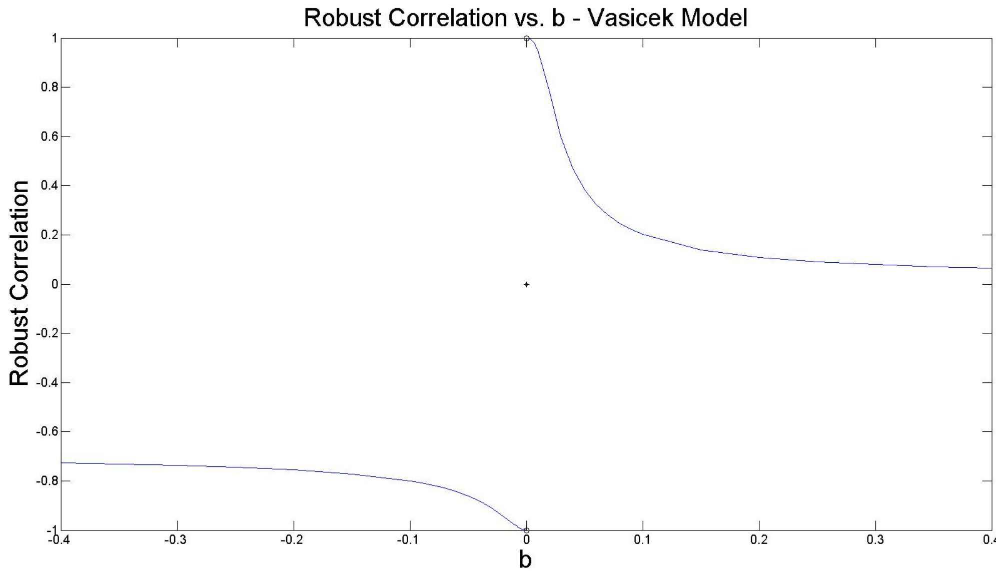


(a) *RWR profile multiplier - Vasicek model*

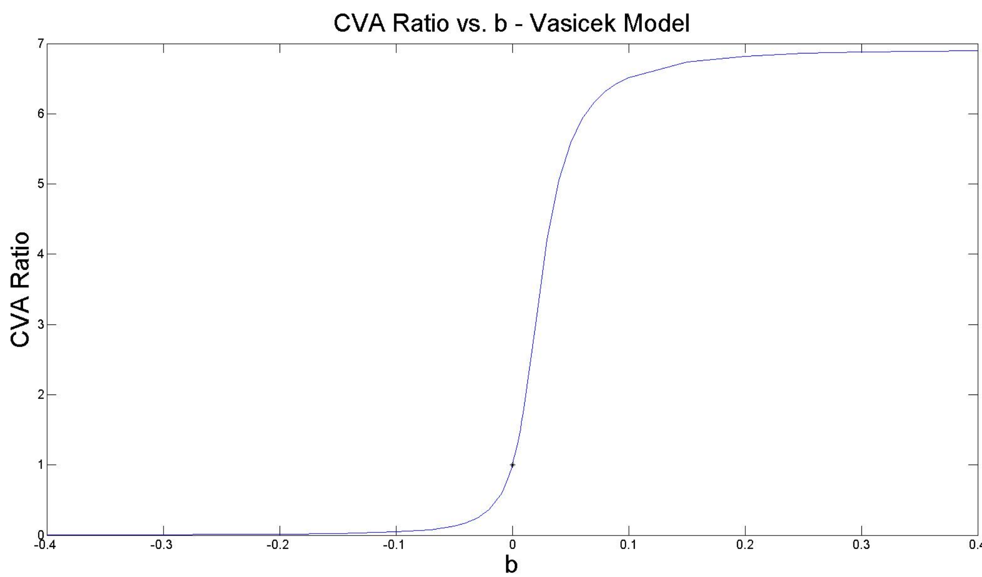


(b) *WWR profile multiplier - Vasicek model*

Figure. 5.3 *Profile multiplier - Vasicek model*



(a) *Robust correlation - Vasicek model*



(b) *CVA ratio - Vasicek model*

Figure. 5.4 $\bar{\rho}$ *and CVA Ratio - Vasicek model*

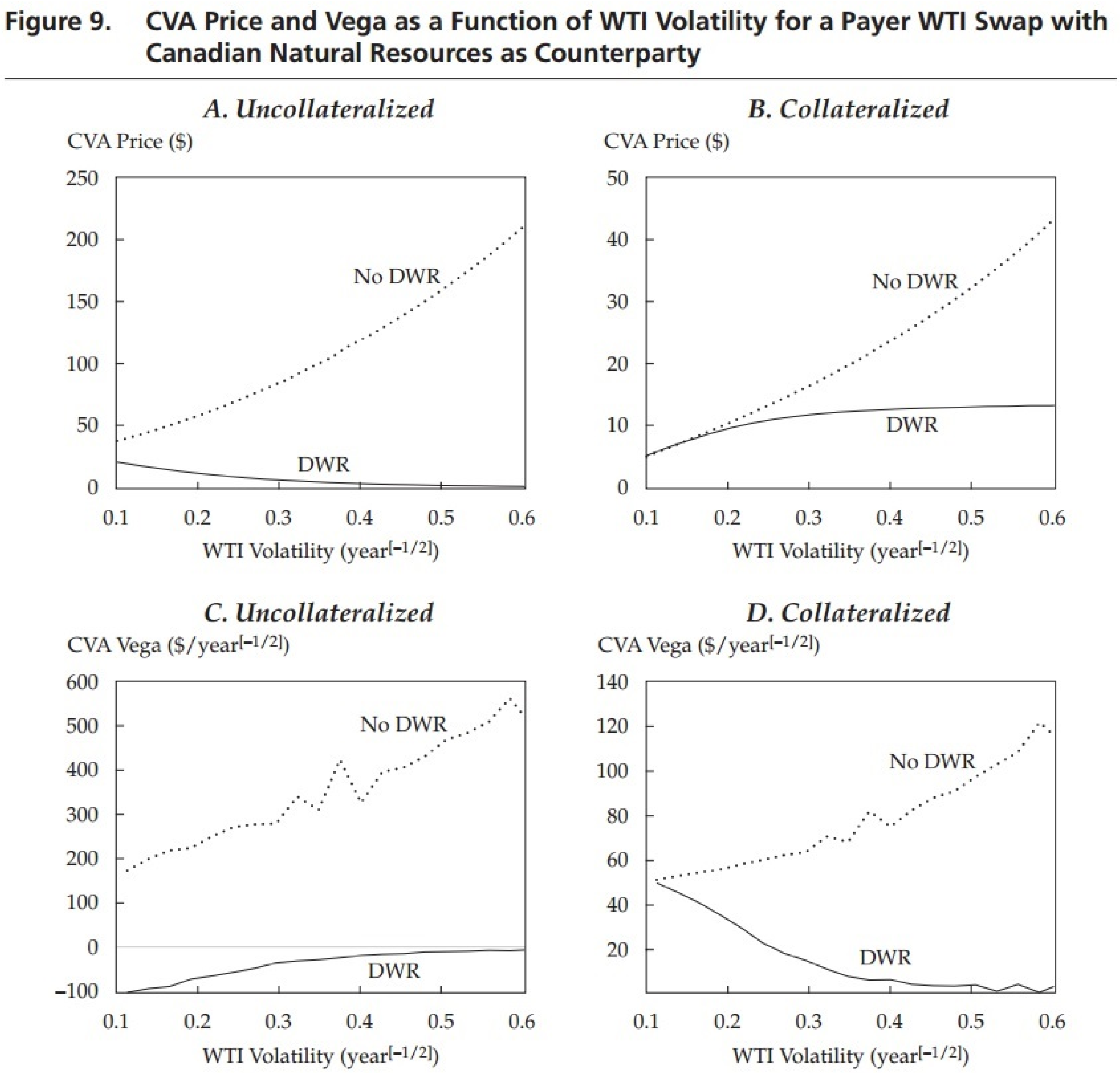


Figure. 5.5 *Snapshot of Figure 9 from Ruiz, Boca and Pachòn [38]*

Chapter 6

VaR and CVaR Optimization with Mixed Normal Returns

Portfolio optimization has been an interesting topic for several decades. Lots of alternative distributions have been proposed to substitute normal distribution in order to better fit the financial data, such as student's t distribution [41] and stable distribution [32] [8] [33]. At the very beginning the optimization used the portfolio return variance as the objective function [30]. Later on different risk measures have been proposed, like Value-at-Risk (VaR) and Conditional-Value-at-Risk (CVaR). They are used to address different concerns of portfolio managers. In this thesis, we focus on VaR and CVaR optimization with mixed normal distribution.

6.1 Definitions

Definition 6.1.0.1. We denote R_p as the rate of return of a portfolio. R_p is a continuous random variable that has cumulative distribution function $F_{R_p}(r_p)$. VaR can be expressed as a threshold

$$VaR_\alpha(R_p) \equiv \max\{r_p \in \mathbb{R} : Pr(R_p \leq r_p) \leq \alpha\}. \tag{6.1}$$

Definition 6.1.0.2. CVaR has the following expression

$$CVaR_\alpha(R_p) \equiv \mathbf{E}\left[R_p | R_p \leq VaR_\alpha(R_p)\right] \tag{6.2}$$

6.2 Data Selection

Our investment universe contains 5 different broad indices across the world. They are Financial Times Stock Exchange 100 (FTSE100), Hang Seng Index (HSI), NASDAQ Composite Index (IXIC), NIKKEI 225 (NI225) and Standard and Poors 500 (GSPC). All of them are widely used as benchmarks by portfolio managers. Since portfolio management may not benefit from rebalancing too often [14], end of month adjust close prices are quoted.

For all the 5 assets, historical monthly log returns are calculated as

$$R_p(t) = ln\left(P_t\right) - ln\left(P_{t-1}\right). \tag{6.3}$$

Descriptive statistics of monthly log returns are shown below.

Table 6.1 Mean of monthly log returns

Asset	FTSE100	HSI	IXIC	NI225	GSPC
μ	0.00365	0.00603	0.00742	-0.00030	0.00595

Table 6.2 Standard deviation of monthly log returns

Asset	FTSE100	HSI	IXIC	NI225	GSPC
σ	0.04559	0.07928	0.06586	0.06192	0.04439

Table 6.3 Return correlation

Asset	FTSE100	HSI	IXIC	NI225	GSPC
FTSE100	1	0.6562	0.6718	0.4662	0.8048
HSI	-	1	0.5924	0.4177	0.6355
IXIC	-	-	1	0.4931	0.9485
NI225	-	-	-	1	0.5157
GSPC	-	-	-	-	1

6.3 Mixed Normal Distribution

A mixed normal distribution with k components has the following CDF and probability density function (PDF):

$$F(x) = \sum_{i=1}^{k} \pi_k F_k(x)$$

and

$$f(x) = \sum_{i=1}^{k} \pi_k f_k(x),$$

where $F_k(x)$ and $f_k(x)$ are CDF and PDF of the k-th component which follows normal distribution with parameters μ_k and σ_k.

First, we standardize each returns and draw quantile-quantile (Q-Q) plots. These plots provide insights of how well the market data fit a normal distribution.

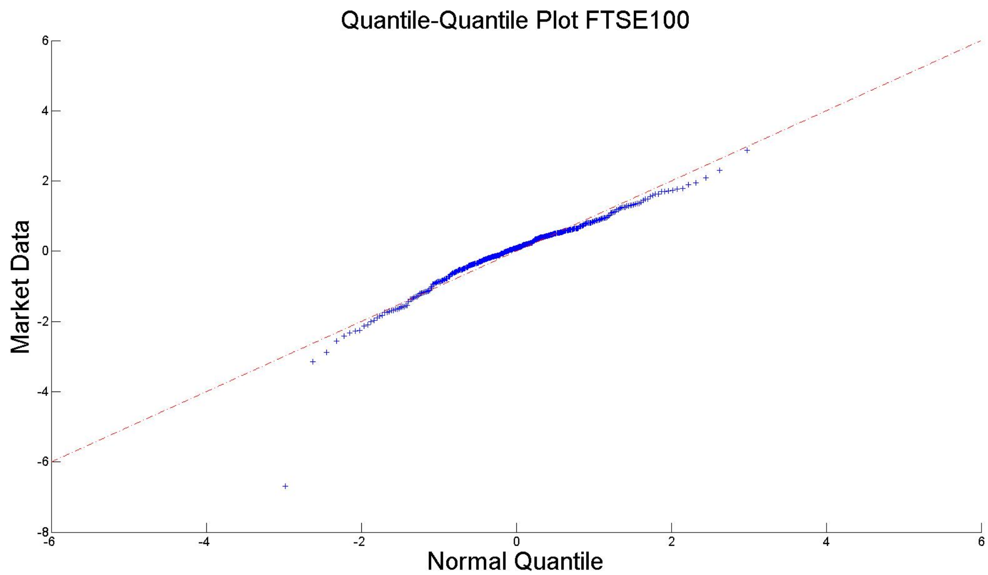


Figure. 6.1 *Q-Q plot for FTSE100 - normal distribution*

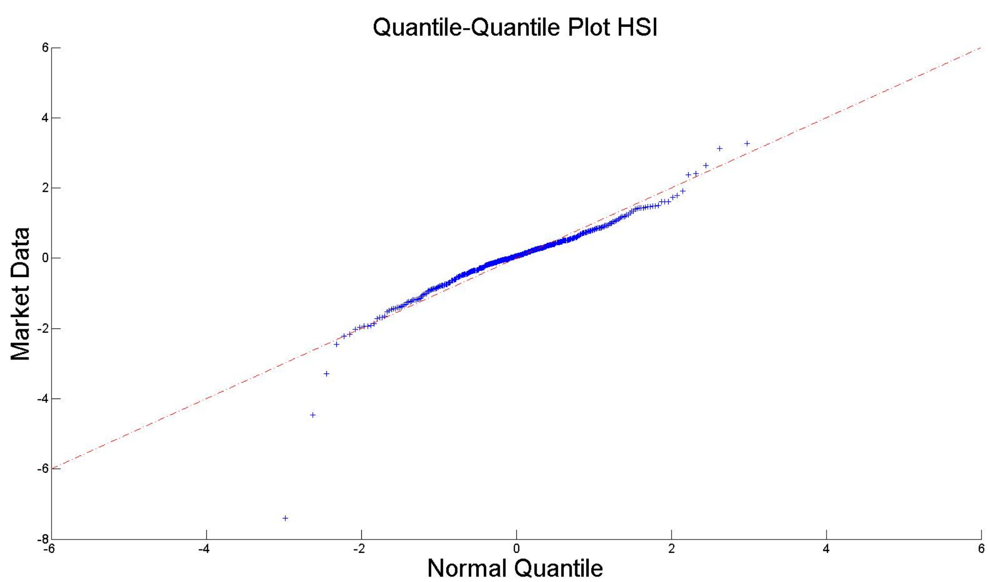


Figure. 6.2 *Q-Q plot for HSI - normal distribution*

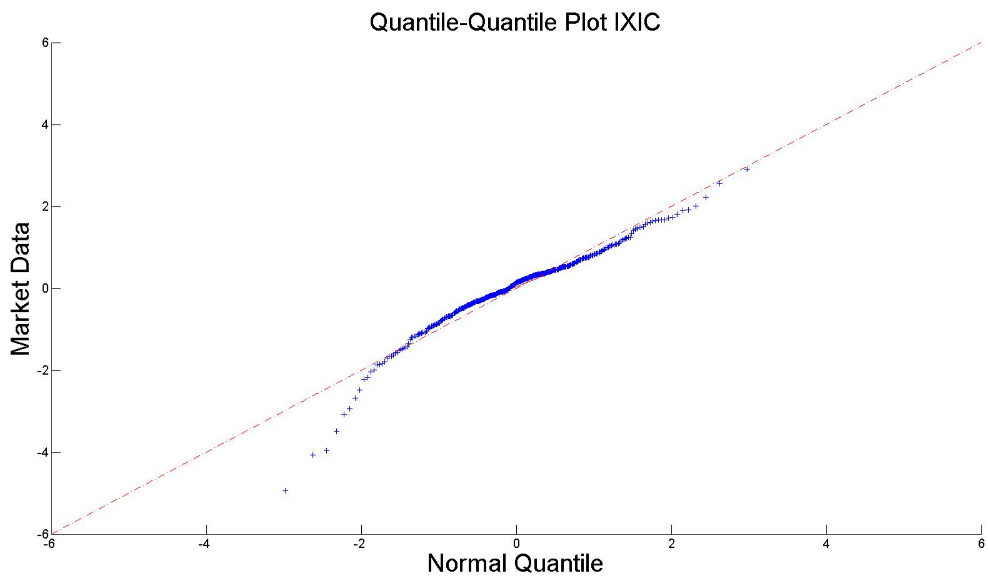


Figure. 6.3 *Q-Q plot for IXIC - normal distribution*

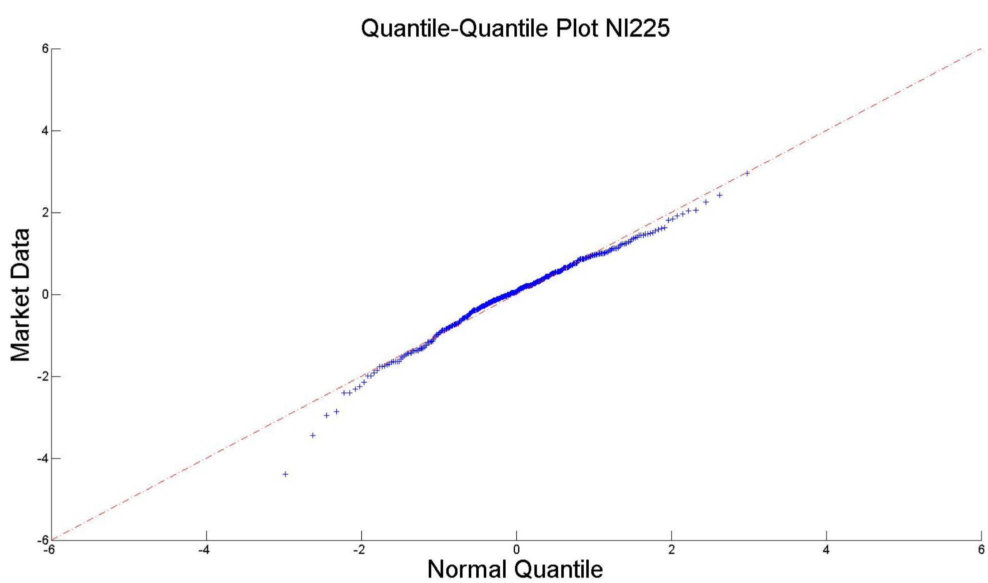


Figure. 6.4 *Q-Q plot for NI225 - normal distribution*

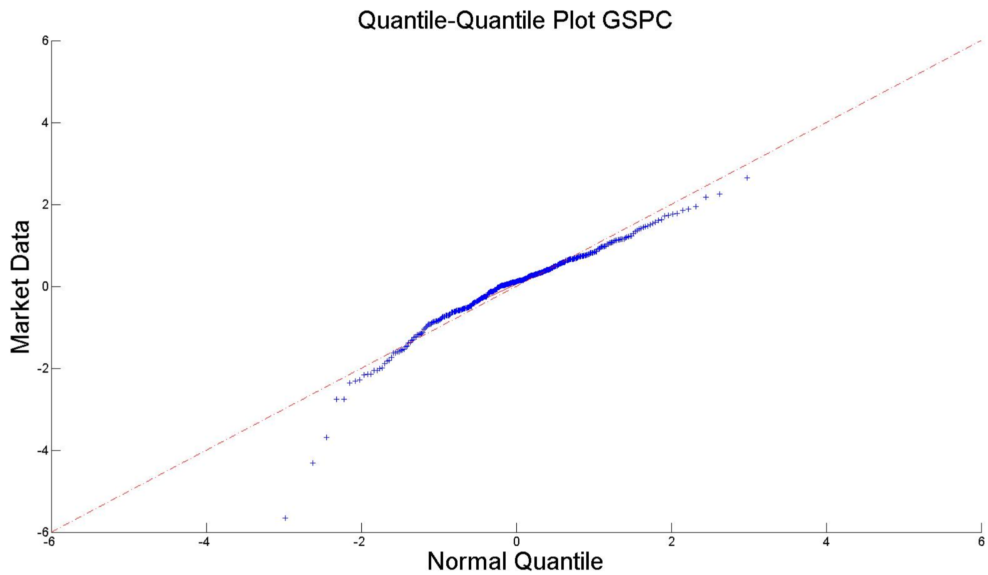


Figure. 6.5 *Q-Q plot for GSPC - normal distribution*

From the shape of the Q-Q plots, one can see that the returns of all 5 assets are negative or left skewed. This suggests we should find alternative distributions to fit the market data. In our work, we choose to fit two component mixed normal distributions. Again Q-Q plots are drawn to check the fit.

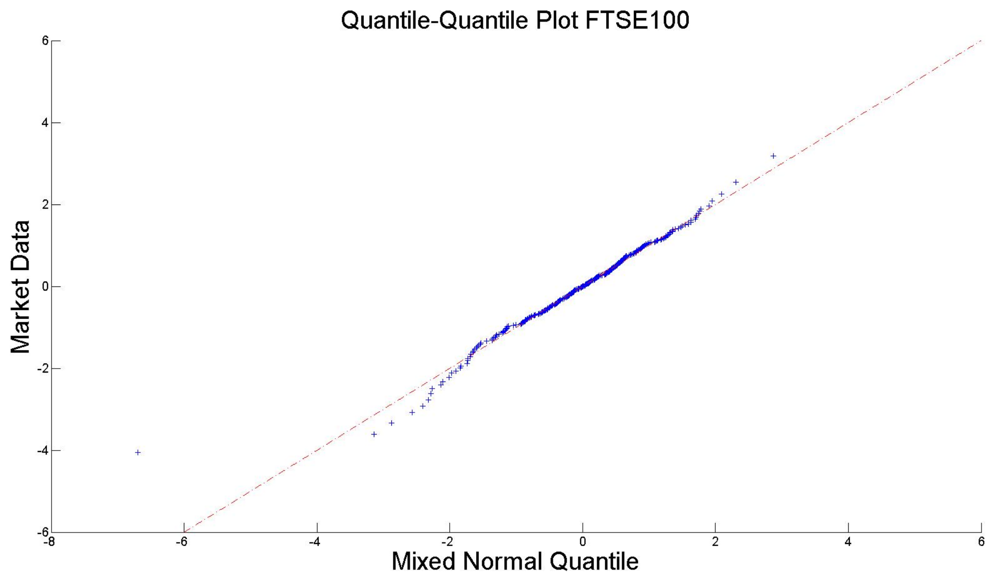


Figure. 6.6 *Q-Q plot for FTSE100 - mixed normal distribution*

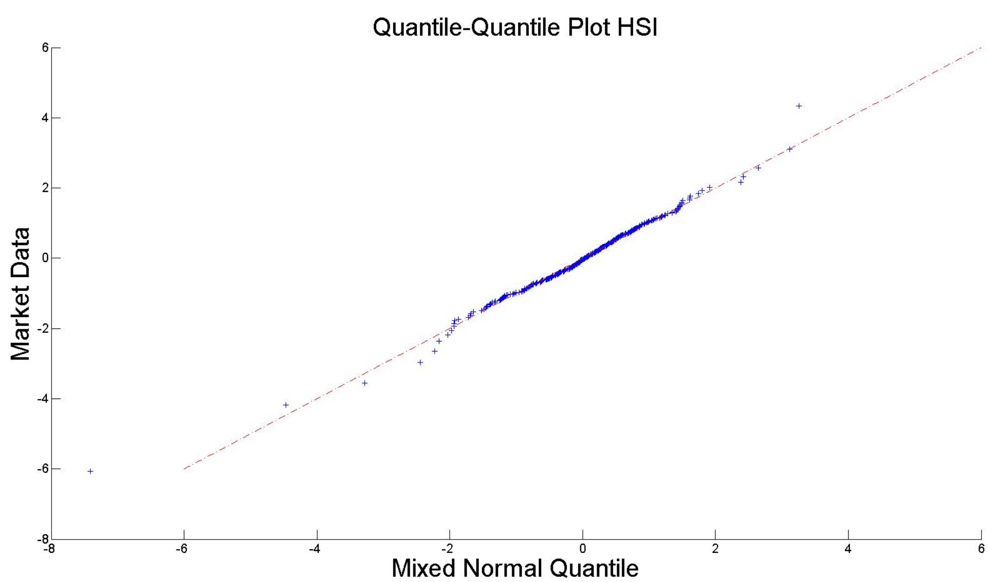


Figure. 6.7 *Q-Q plot for HSI - mixed normal distribution*

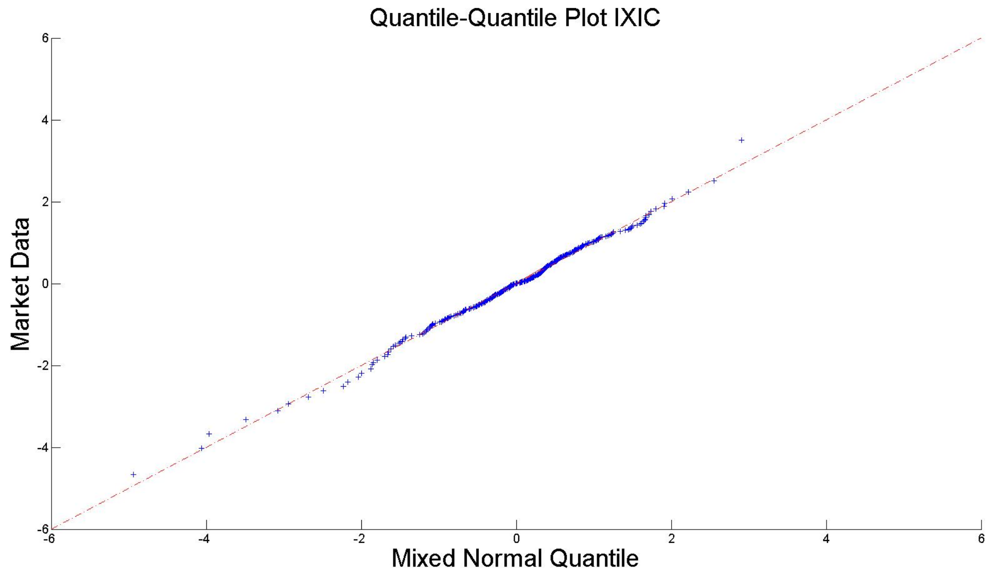


Figure. 6.8 *Q-Q plot for IXIC - mixed normal distribution*

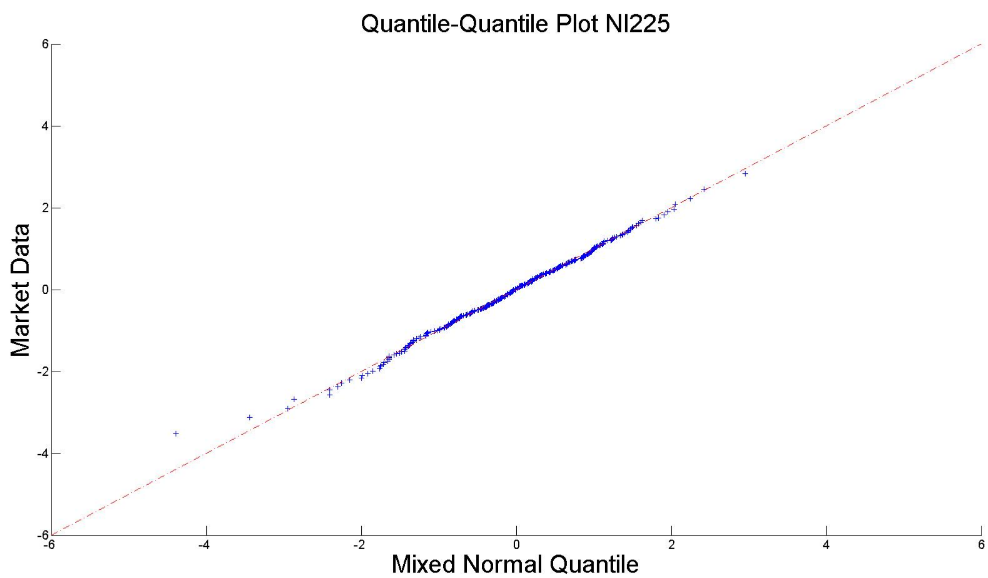


Figure. 6.9 *Q-Q plot for NI225 - mixed normal distribution*

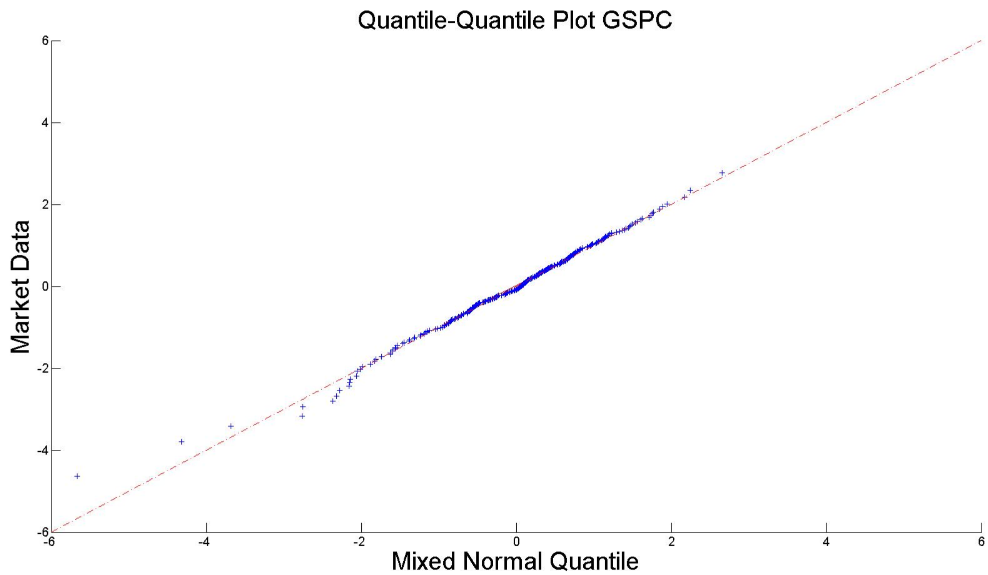


Figure. 6.10 *Q-Q plot for GSPC - mixed normal distribution*

Almost all the market data lie on the 45 degree line, which means a two component mixed normal distribution can better fit the data than a normal distribution. And this holds for all 5 assets we selected.

6.4 Problem Setup

Assume that the goal of a professional investment manager is to earn a fixed target return. Meanwhile the risk of the portfolio needs to be minimized. Candidate risk measures are variance, VaR and CVaR. If the portfolio value is normally distributed then these three measures are equivalent. However, this is not true for other distributions like mixed normal distribution.

We assume the market has two different states: normal and stressed. In each state, all the assets follow a joint normal distribution. The distribution parameters in each state are different. Let's consider one period T without re-balancing.

We define the followings:

1. N: The total number of individual assets in the manager's universe - scalar.

2. x_P: Weights of each individual assets in the manager's portfolio with N elements, x_P^1, x_P^2, $\cdots x_P^N$ - N by 1 vector.

3. θ_k: Expected return of assets in state k, for $k = 1,2$ denoting normal and stressed markets respectively - N by 1 vector.

4. Σ_k: Variance-Covariance matrix of assets in state k, for $k = 1,2$ denoting normal and stressed markets respectively - N by N matrix.

5. π_k: The probability of the market is in state k at the end of the period T; $\sum_{k=1}^{2} \pi_k = 1$.

6. W: The target return.

The portfolio return in state k follows $N(x_P'\theta_k, x_P'\Sigma_k x_P)$ and the return at time T should follow a mixed normal distribution. $MN(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \pi_1, \pi_2)$ is used to denote the distribution, where $\mu_1 = x_P'\theta_1$, $\mu_2 = x_P'\theta_2$, $\sigma_1 = x_P'\Sigma_1 x_P$ and $\sigma_2 = x_P'\Sigma_2 x_P$.

The characteristic function of this mixed normal distribution can be expressed as

$$\Psi_{R_p}(t) = \mathbf{E}\left[e^{itR_p}\right] = \sum_{k=1}^{2} \pi_k \exp\left(it\mu_k - \frac{1}{2}\sigma_k^2 t^2\right) \tag{6.4}$$

where i is the imaginary unit i.e. $i = \sqrt{-1}$.

It's not hard to see that

$$\mathbf{E}[R_p] = \sum_{k=1}^{2} \pi_i \mu_k. \tag{6.5}$$

Taking second derivative of $\Psi_{R_p}(t)$ with respect to t we have

$$\frac{d^2\Psi_{R_p}(t)}{dt^2} = \sum_{k=1}^{2} \pi_k \left[\left(i\mu_k - \sigma_k^2 t\right)^2 - \sigma_k^2\right] \Psi_{R_p}(t). \tag{6.6}$$

By setting $t = 0$ and multiplying i^2, we get the second moment of R

$$\mathbf{E}\left[R_p^2\right] = \sum_{k=1}^{2} \pi_k \left(\sigma_k^2 + \mu_k^2\right). \tag{6.7}$$

Then the mean variance optimization problem can be written as:

$$\begin{aligned}
\min : & \sum_{k=1}^{2} \pi_k \left(\sigma_k^2 + \mu_k^2\right) - \left(\sum_{k=1}^{2} \pi_k \mu_k\right)^2 \\
& s.t. \\
& \sum_{i=1}^{N} x_P^i = 1 \\
& \sum_{k=1}^{2} \pi_k \mu_k \geq W.
\end{aligned}$$

Note here σ_k^2 and μ_k are nothing but functions of x_P assuming all the other parameters are given. Hence the mean-variance optimization problem can be expressed as a pretty simple programming problem with linear constraints and quadratic objective function and can be solved very fast without any simulation at all.

The VaR and CVaR optimization problems should have the same constraints as the mean variance optimization. But they have different objective functions, which are

$$\min : -VaR_\alpha(R_p; \mu, \sigma, \pi)$$

and

$$\min : -CVaR_\alpha(R_p; \mu, \sigma, \pi).$$

Both types of optimization problems can be solved numerically. In the next subsection, we show how that VaR can be found quickly with a bisection search. When VaR is given, CVaR calculation is straight forward.

6.5 Bisection Search For VaR of Mixed Normal Distribution

An example of a mixed normal distribution with two components is shown here. The first and second component have cumulative distribution function $F_1(r)$ and $F_2(r)$ and weights π_1 and π_2. Then the cumulative function of R is

$$F_{R_p}(r_p) = \pi_1 F_1(r_p) + \pi_2 F_2(r_p).$$

First, let's find the lower α quantile for both F_1 and F_2, namely c_1 and c_2 respectively.

$$c_1 = \max\{r_p \in \mathbb{R} : F_1(r_p) \leq \alpha\}$$

$$c_2 = \max\{r_p \in \mathbb{R} : F_2(r_p) \leq \alpha\}$$

Without loss of generality, let's say $c_1 \leq c_2$.

Next, we use bisection search to find level α VaR for the mixed normal distribution. The solution must lie in the interval with end points c_1 and c_2. It's not hard to see the followings hold

$$F_1(c_1) = \alpha$$

$$F_2(c_1) \leq \alpha$$

and

$$F_R(c_1) = \pi_1 F_1(c_1) + \pi_2 F_2(c_1) \leq \alpha;$$

$$F_1(c_2) \geq \alpha$$

$$F_2(c_2) = \alpha$$

and

$$F_{R_p}(c_2) = \pi_1 F_1(c_2) + \pi_2 F_2(c_2) \geq \alpha.$$

Moreover, cumulative distribution function F_{R_p} is non-decreasing, so the solution must lie in the interval $[c_1, c_2]$.

6.6 CVaR Calculation

For a random variable X that follows $N(\mu,\sigma^2)$, one can see that

$$\begin{aligned}
\mathbf{E}[X|X\le c] &= \frac{1}{Pr(X\le c)\sigma\sqrt{2\pi}}\int_{-\infty}^{c} x\exp^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx \\
&= \frac{1}{\Phi(\frac{c-\mu}{\sigma})\sqrt{2\pi}}\int_{-\infty}^{\frac{c-\mu}{\sigma}} (\sigma y+\mu)\exp^{-\frac{y^2}{2}}\,dy \\
&= \frac{1}{\Phi\left(\frac{c-\mu}{\sigma}\right)}\left[\Phi\left(\frac{c-\mu}{\sigma}\right)\mu-\sigma\phi\left(\frac{c-\mu}{\sigma}\right)\right] \\
&= \frac{1}{\alpha}\left[\Phi\left(\frac{c-\mu}{\sigma}\right)\mu-\sigma\phi\left(\frac{c-\mu}{\sigma}\right)\right]
\end{aligned} \tag{6.8}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are cumulative distribution function and probability density function of standard normal distribution.

In the case where return R follows a mixed normal distribution with two components, CVaR can be expressed as

$$\begin{aligned}
\mathbf{E}[R_p|R_p\le VaR_\alpha] &= \frac{1}{\alpha}\sum_{k=1}^{2}\pi_k\left[\Phi\left(\frac{VaR_\alpha-\mu_k}{\sigma_k}\right)\mu_k-\sigma_k\phi\left(\frac{VaR_\alpha-\mu_k}{\sigma_k}\right)\right] \\
&= \sum_{k=1}^{2}\frac{\pi_k\Phi\left(\frac{VaR_\alpha-\mu_k}{\sigma_k}\right)}{\alpha}\left[\mu_k-\frac{\sigma_k\phi\left(\frac{VaR_\alpha-\mu_k}{\sigma_k}\right)}{\Phi\left(\frac{VaR_\alpha-\mu_k}{\sigma_k}\right)}\right]
\end{aligned} \tag{6.9}$$

Note here VaR_α is the α level VaR of the mixed normal distribution. As is shown in the previous subsection, VaR can be found with a bisection search given.

6.7 Stressed States

Stressed states can be defined based on a portfolio manager's view. Due to the efficiency of solving a VaR or CVaR optimization with mixed normal distribution, more than one stressed states can be defined and easily added to the optimization problem. One can also calibrate a mixed normal distribution with historical data [13] [44] [29] and add stressed states.

The stressed states we use is defined in the following way.

1. In a stressed state, we assume the mean of returns is 90% of what it is in the normal state;
2. The returns are more volatile and we multiply the volatility by 1.1;
3. Denote Q as the correlation matrix for normal state, we assume in the stressed state the correlation matrix is $Q^{\frac{1}{2}}$.

This kind of setup is also used in Chen and Skoglund [12]. In the next section, we show some numerical results and discuss them.

6.8 Numerical Results and Discussions

The following plots show the VaR and CVaR optimization results. There are two main advantages of this optimization. First, it's very flexible. More components can be added and the components weight can be adjusted based on a portfolio manager's view.

Second, unlike stable distribution [33], PDF and CDF don't have to be calculated numerically. Moreover since one can get the explicit expression of CVaR, the optimization problem can be solved much faster. In this numerical example, 38 different target return values were chosen over the interval of [0.0036, 0.0073] equally spaced. MATLAB was used to solve the optimization problem and the code was run on a quattro core i-7 processor @ 2.80GHz with 8 gigabyte memory. It only took 1.4738 and 1.6538 seconds to find the optimal solution to minimize VaR and CVaR respectively.

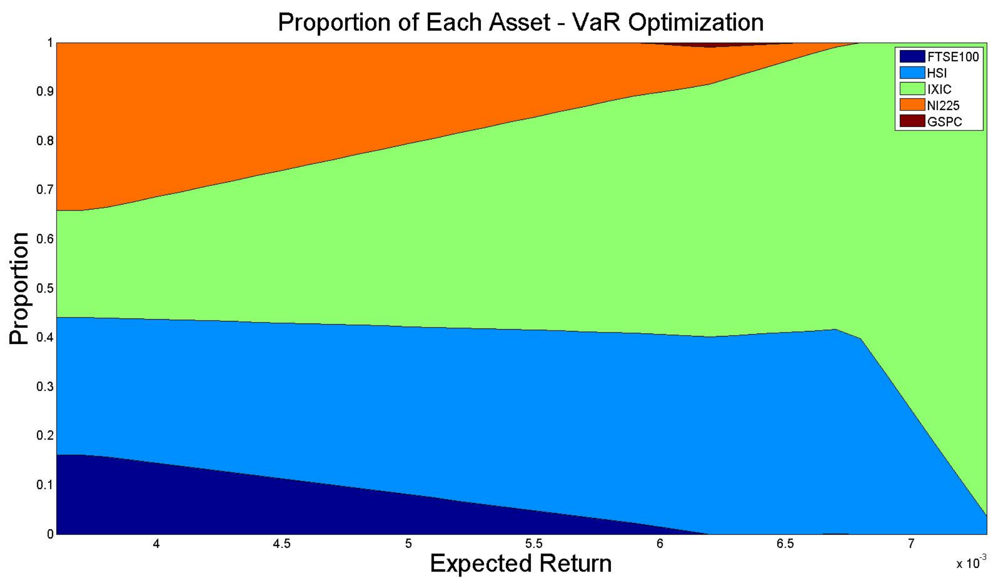


Figure. 6.11 *VaR asset allocation*

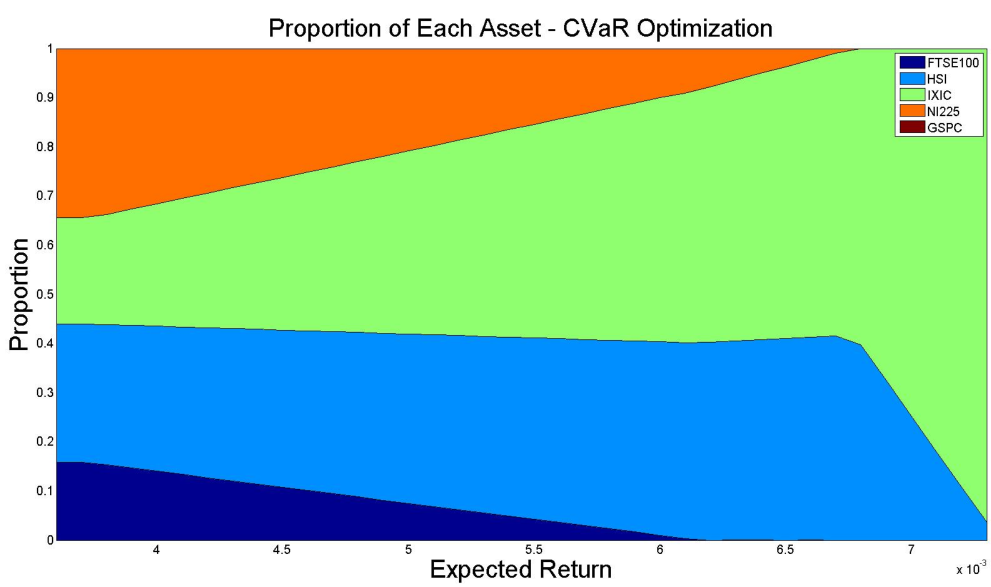


Figure. 6.12 *CVaR asset allocation*

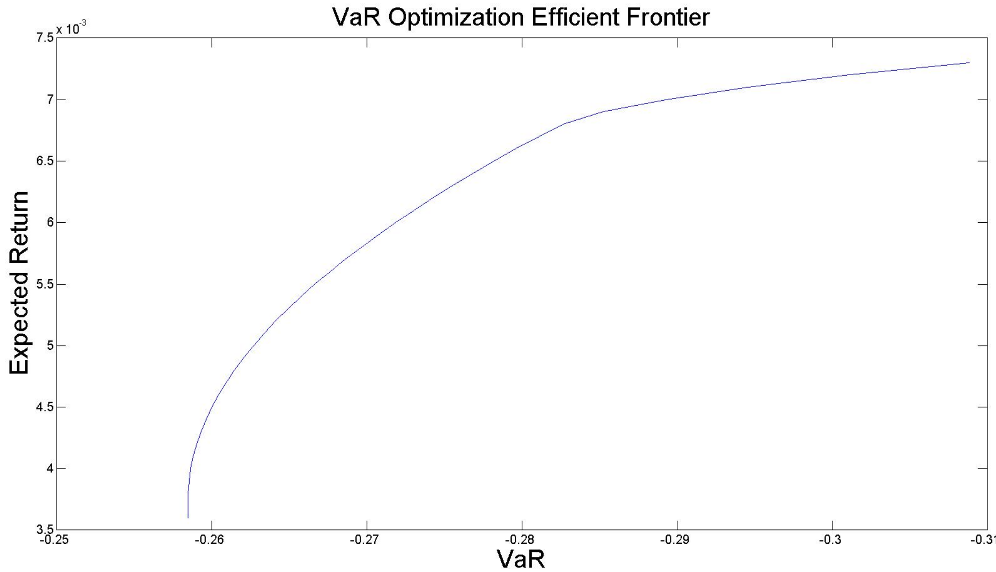


Figure. 6.13 *VaR efficient frontier*

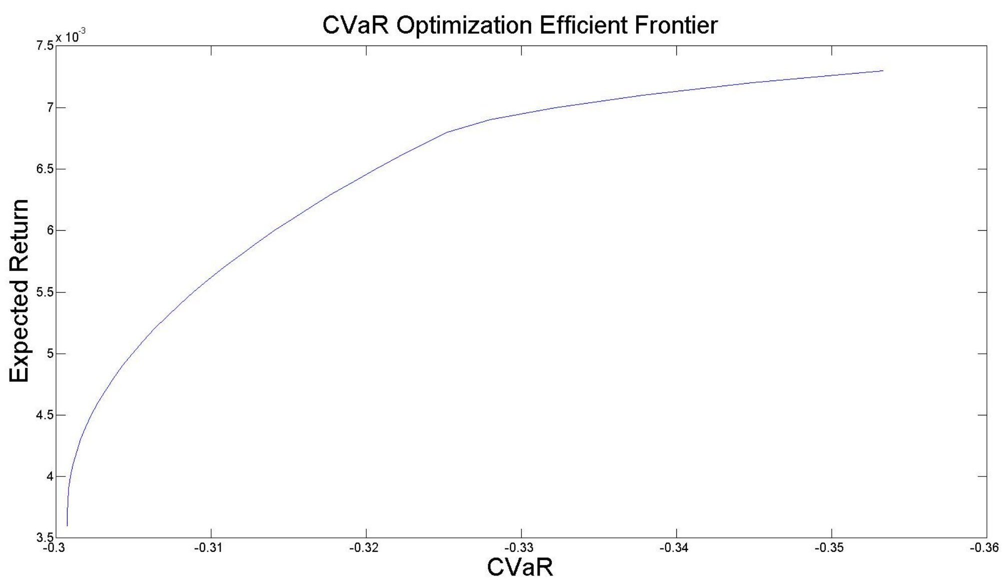


Figure. 6.14 *CVaR efficient frontier*

BIBLIOGRAPHY

[1] Alfonsi A. (2005), ‘On the Discretization Schemes for the CIR (and Bessel Squared) Processes’, *Monte Carlo Methods and Applications mcma*, Vol. 11, No. 4, pp. 355-384.

[2] Basel Committee on Banking Supervision (BCBS) (2011), ‘Capital treatment for bilateral counterparty credit risk finalised by the Basel Committee’, available at: http://www.bis.org/press/p110601.htm (accessed 16th November, 2014).

[3] Berndt A., Douglas R., Duffie D., Ferguson M. and Schranz D. (2005), ‘Measuring Default Risk Premia from Default Swap Rates and EDFs’, Bank for International Settlements Working Paper No. 173.

[4] Black F. and Cox J.C. (1976), ‘Valuing Corporate Securities: Some Effects of Bond Indenture Provisions’, *Journal of Finance*, Vol. 31, No. 2, pp. 351-367.

[5] Black F. and Scholes M. (1973), ‘The Pricing of Options and Corporate Liabilities’, *Journal of Political Economy*, Vol. 81, No. 3, pp. 637-654.

[6] Black F., Derman E. and Toy W. (1990), ‘A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options’, *Financial Analysts Journal*, Vol. 46, No. 1, pp. 33-39.

[7] Black F. and Karasinski P. (1991), ‘Bond and Option Pricing when Short Rates are Lognormal’, *Financial Analysts Journal*, Vol. 47, No. 4, pp. 52-59.

[8] Borak S., Härdle W. and Weron R. (2005), ‘Stable Distributions’, SFB 649 Discussion Paper 2005-08 . Humboldt University, Berlin.

[9] Brigo D. and Alfonsi A. (2005), 'Credit Default Swaps Calibration and Derivatives Pricing with the SSRD Stochastic Intensity Model', *Finance and Stochastic*, Vol. 9, No. 1, pp. 29-42.

[10] Brigo D. and Pallavicini A. (2008), 'Counterparty Risk and Contingent CDS Under Correlation', *Risk Magazine*, Vol. 21, No. 2, pp. 84-88.

[11] Chan K.C., Karolyi G.A., Longstaff F.A. and Sanders A.B. (1992), 'An Empirical Comparison of Alternative Models of the Short-Term Interest Rate', *The Journal of Finance*, Vol. 48, No. 3, pp. 1209-1227.

[12] Chen W. and Skoglund J. (2013), 'An Integrated Stress Testing Framework via Markov Switching Simulation', *Journal of Risk Model Validation*, Vol. 7, No. 2, pp. 3-27.

[13] Cohen A.C. (1967), 'Estimation in Mixtures of Two Normal Distributions', *Technometrics*, Vol. 9, No. 1, pp. 15-28.

[14] Colleen M.J., Francis M.K. Jr. and Yan Z. (2010), 'Best Practices for Portfolio Rebalancing', Vanguard Research, available at http://www.vanguard.com/pdf/icrpr.pdf.

[15] Cox J.C., Ingersoll J.E. and Ross S.A. (1985), 'A Theory of the Term Structure of Interest Rates', *Econometrica* Vol. 53, No. 2, pp. 385-407.

[16] Deelstra G. and Delbaen F. (1998), 'Convergence of Discretized Stochastic (Interest Rate) Processes with Stochastic Drift Term', *Appl. Stochastic Models Data Anal.*, Vol. 14, No.1, pp. 77-84.

[17] Diop A. (2003), 'Euler Scheme for SDEs with Non-Lipschitz Diffusion Coefficient: Strong Convergence', Ph.D Thesis, INRIA, available at http://hal.inria.fr/docs/00/05/42/25/PDF/RR-5637-V2.pdf.

[18] Fama E.F. (1965), 'The Behavior of Stock Prices', *The Journal of Business*, Vol. 38, No. 1, pp. 34-105.

[19] Fouque J.-P., Sircar R. and Solna K. (2006), 'Stochastic Volatility Effects on Defaultable Bonds', *Applied Mathematical Finance*, Vol. 13, No. 3, pp. 215-244.

[20] Fouque J.-P., Wignall B.C. and Zhou X. (2008), 'Modeling Correated Defaults: First Passage Model Under Stochastic Volatility', *Journal of Computational Finance*, Vol. 11, No. 3, pp. 43-78.

[21] Ghamami S. and Goldberg L.R. (2014), 'Stochastic Intensity Models of Wrong Way Risk: Wrong Way CVA Need Not Exceed Independent CVA', *The Journal of Derivatives*, 2014, Vol. 21 Issue 3, pp. 24-35.

[22] Ho T.S.Y. and Lee S.B. (1986), 'Term Structure Movements and Pricing Interest Rate Contigent Claims', *Journal of Finance*, Vol. 41, No. 5, pp. 1011-1029.

[23] Hull J. and White A. (1990), 'Pricing Interest-Rate Derivative Securities', *The Review of Financial Studies*, Vol. 3, No. 4, pp. 573-592.

[24] Hull J. and White A. (2012), 'CVA and Wrong Way Risk', *Financial Analysts Journal*, Vol. 68, No. 5, pp. 58-69.

[25] Hull J. and White A. (1998), 'Value at Risk When Daily Changes in Market Variables are not Normal Distributed', *Journal of Derivaties*, Vol. 5, pp. 9-19.

[26] International Swap and Derivatives Association (ISDA) (2003), ‘Counterparty Risk Treatment of OTC Derivatives and Securities Financing Transactions’, available at: http://www.isda.org/c_and_a/pdf/counterpartyrisk.pdf (accessed 16th November, 2014).

[27] Jarrow R., Lando D. and Yu F. (2005), ‘Default Risk and Diversi?cation: Theory and Applications’, *Mathematical Finance* Vol. 15, NO. 1, pp. 1-26.

[28] Gregory J. (2012), Counterparty Credit Risk and Credit Value Adjustment: A Continuing Challenge for Global Financial Markets. 4th ed. Chichester, UK: John Wiley and Sons. pp. 380.

[29] Kim H.-M. and Genton M.G. (2011), ‘Characteristic Functions of Scale Mixtures of Multivariate Skew-Normal Distributions’, *Journal of Multivariate Analysis*, 102(7), pp. 1105-1117.

[30] Markowitz, H.M. (1952), ‘Portfolio Selection’, *Journal of Finance*, Vol. 7, No. 1, pp. 77-91.

[31] Merton R.C. (1973), ‘Theory of Rational Option Pricing’, *Bell Journal of Economics and Management Science* , Vol. 4, No. 1, pp. 141-183.

[32] Nolan J.P. (2003), ‘Modeling Financial Distributions with Stable Distributions’, *Handbooks in Finance*, Vol. 1, Chap. 3, pp. 105-130.

[33] Pang T. and Yang S. (2015), ‘Copulas and GARCH Models for Market Risks of Relative Returns’, *Journal of Finance and Management Research*, Vol. 1, No. 1, pp. 19-38

[34] Pang T., Chen W. and Li L. (2015), ‘CVA Wrong Way Risk Decomposition and Efficient CVA Curve’, *Journal of Risk Management in Financial Institutions*, 2015, Vol. 8 Issue 4

[35] Paseka A., Koulis T. and Thavaneswaran A. (2012), 'Interest Rate Models', *Journal of Mathematical Finance*, Vol. 2, No. 2, pp. 141-158.

[36] Pykhtin M. and Rosen D. (2010), 'Pricing Counterparty Risk at the Trade Level and CVA Allocations', *Journal of Credit Risk*, Vol. 6, No. 4, pp. 3-38.

[37] Rockafellar, R.T. and Uryasev, S. (2000), 'Optimization of Conditional Value-at-Risk', *Journal of Risk*, Vol. 7, No. 1, pp. 77-91.

[38] Ruiz I., Boca P.D. and Pachòn R. (2015), 'Optimal Right- and Wrong-Way Risk from a Practitioner Standpoint', *Financial Analysts Journal*, 2015, Vol. 71, No. 2, pp. 47-60.

[39] Rosen D. and Saunders D. (2012), 'CVA The Wrong Way', *Journal of Risk Management in Financial Institutions*, Vol. 5, No. 3, pp. 152-272.

[40] Skoglund J., Vestal D. and Chen W. (2013), 'Credit Valuation Adjustment Tail Risk and the Impact of Wrong Way Trades', *Journal of Risk Management in Financial Institutions*, 2013, Vol. 6 Issue 3, pp. 280-301

[41] Theodossiou P. (1998), ' Financial Data and the Skewed Generalized T Distribution', *Management Science*, 44(12-part-1), pp.1650-1661.

[42] Varma P. and Cantor R. (2005), 'Determinants of Recovery Rates on Defaulted Bonds and Loans for North American Corporate Issuers: 1983-2003', *Journal of Fixed Income*, Vol. 14, No. 4, pp. 29-44.

[43] Vasicek O. (1977), 'An Equilibrium Characterisation of the Term Structure', *Journal of Financial Economics*. Vol. 5, No. 2, pp. 177-188.

[44] Xu D. and Wirjanto T. S. (2010), ‘An Empirical Characteristic Function Approach to VaR Under a Mixture-of-Normal Distribution with Time Varying Volatility’, *Journal of Derivatives*, Vol. 18, No. 1, pp. 39-58.

APPENDICES

Appendix A

Pricing of Vanilla Interest Rate Swap

Without loss of generality, we assume the notional principal amount is 1. Each fixed rate payment B_i, for $i = 0,1,\ldots, N-1$, has the present value

$$PV(B_i) = K\tau_i D(0,t_{i+1}) \tag{A.1}$$

where N is the total number of payments; K is the fixed rate determined at $t = 0$; τ_i is the time between the t_i and t_{i+1} as a fraction of a year; $D(0,t_{i+1})$ is the price of zero coupon bond with maturity t_{i+1}; t_i's are repricing dates.

Each floating rate payment A_i, for $i = 0,1,\ldots, N-1$, has the present value

$$PV(A_i) = E[R_i\tau_i D(0,t_{i+1})] \tag{A.2}$$

where R_i is the spot rate at time t_i with maturity t_{i+1}.

Assuming that there is no arbitrage, we should set

$$R_i = \frac{D(0,t_i)/D(0,t_{i+1}) - 1}{\tau_i} \tag{A.3}$$

The present value of the fixed rate leg is

$$\sum_{i=0}^{N-1} PV(B_i) = \sum_{i=0}^{N-1} K\tau_i D(0, t_{i+1}) \tag{A.4}$$

The present value of the floating rate leg is

$$\sum_{i=0}^{N-1} PV(A_i) = \sum_{i=0}^{N-1} R_i\tau_i D(0, t_{i+1}) \tag{A.5}$$

The equilibrium swap rate X_t is defined such that the time t net present value of the swap is 0, given that a new zero curve is available at time t. That is to say standing at time t

$$\begin{aligned}\sum_{t_i \geq t} PV(B_i) &= \sum_{t_i \geq t} X_t\tau_i D(t, t_{i+1}) \\ &= \sum_{t_i \geq t} R_i\tau_i D(t, t_{i+1}) = \sum_{t_i \geq t} PV(A_i)\end{aligned}$$

Solve for X_t

$$X_t = \frac{\sum_{t_i \geq t} R_i\tau_i D(t, t_{i+1})}{\sum_{t_i \geq t} \tau_i D(t, t_{i+1})} \tag{A.6}$$

We assume $K = X_0$ so that the initial value of the swap is 0.

To calculate CVA, we need to simulate the daily risk-free mark-to-market value of the vanilla swap. Let's take a look at how the net present value of a fixed rate payer swap evolves as time t changes. Because our focus is CVA and not the value of the swap itself, we further assume that all τ_i's are equal. Given the repricing frequency is m times per year, $\tau_i = \frac{1}{m}$.

There are two cases: t is a repricing date or between two repricing dates. For the first case, we have

$$NPV(t) = -\frac{1}{m}\sum_{t_i \geq t} X_0 D(t, t_{i+1}) + \frac{1}{m}\sum_{t_i \geq t} R_i D(t, t_{i+1}) \tag{A.7}$$

Apply Eq. A.6, we have

$$NPV(t) = \frac{X_t - X_0}{m}\sum_{t_i \geq t} D(t, t_{i+1}) \tag{A.8}$$

Now the new set of repricing rates should be derived from the new zero curve, we can write Eq. A.3 as

$$R_i = m[D(t,t_i)/D(t,t_{i+1}) - 1] \tag{A.9}$$

Consider the numerator of Eq. A.6 and express R_i with Eq. A.9

$$\begin{aligned}\sum_{t_i \geq t} R_i \tau_i D(t,t_{i+1}) &= \sum_{t_i \geq t} [D(t,t_i) - D(t,t_{i+1})] \\ &= D(t,t_i) - D(t,t_{n+1}) \\ &= D(t,t) - D(t,t_{n+1})\end{aligned} \tag{A.10}$$

Since $D(t,t) = 1$, we can rewrite X_t as

$$X_t = m\frac{1 - D(t,t_{n+1})}{\sum_{t_i \geq t} D(t,t_{i+1})} \tag{A.11}$$

If t is between two pricing dates, the first floating rate R_{base} is determined on the previous repricing date. Hence X_t should be modified as

$$X_t = m\frac{(\frac{R_{base}}{m} + 1)D(t,t_i) - D(t,t_{n+1})}{\sum_{t_i \geq t} D(t,t_{i+1})} \tag{A.12}$$

$$NPV(t) = \begin{cases} 1 - D(t,t_{n+1}) - PV(B) & t \text{ is repdate} \\ (\frac{R_{base}}{m} + 1)D(t,t_i) - D(t,t_{n+1}) - PV(B) & t \text{ is not repdate} \end{cases} \tag{A.13}$$

where $PV(B) = \frac{X_0}{m}\sum D(t,t_{i+1})$.

Appendix B

Error of Appoximation

We set the initial value of the portfolio as \$10,000. On each day there is a profit or loss that follows a normal distribution with 0 mean and \$100 standard deviation. The time horizon we use is one year or 252 trading days. We choose b to be 0.4. 100,000 paths are used in this simulation. The following figures show that the first order approximation we use in Chapter 5 can be safely assumed.

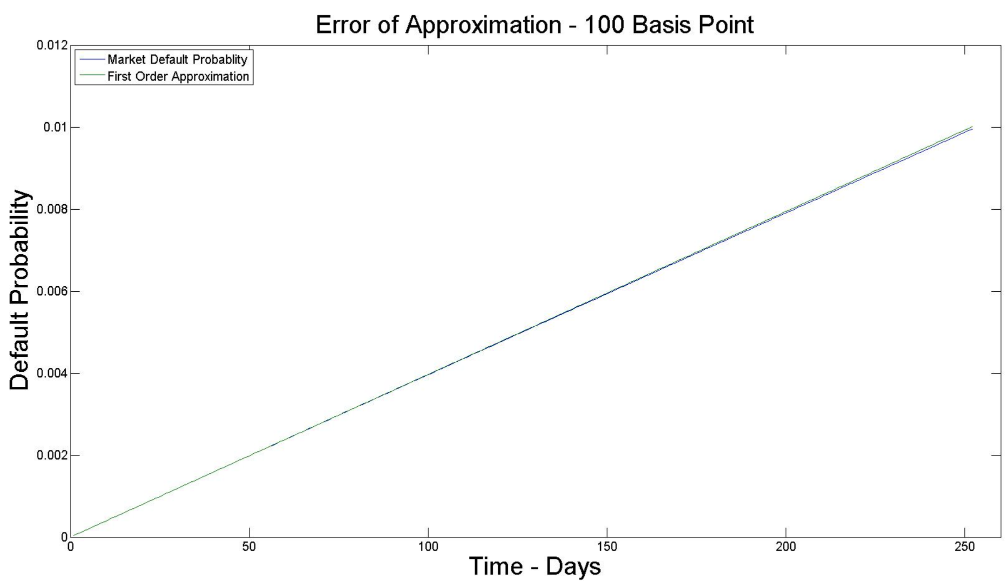


Figure. B.1 *Default Probability - Market CDS Spread 100 Basis Point*

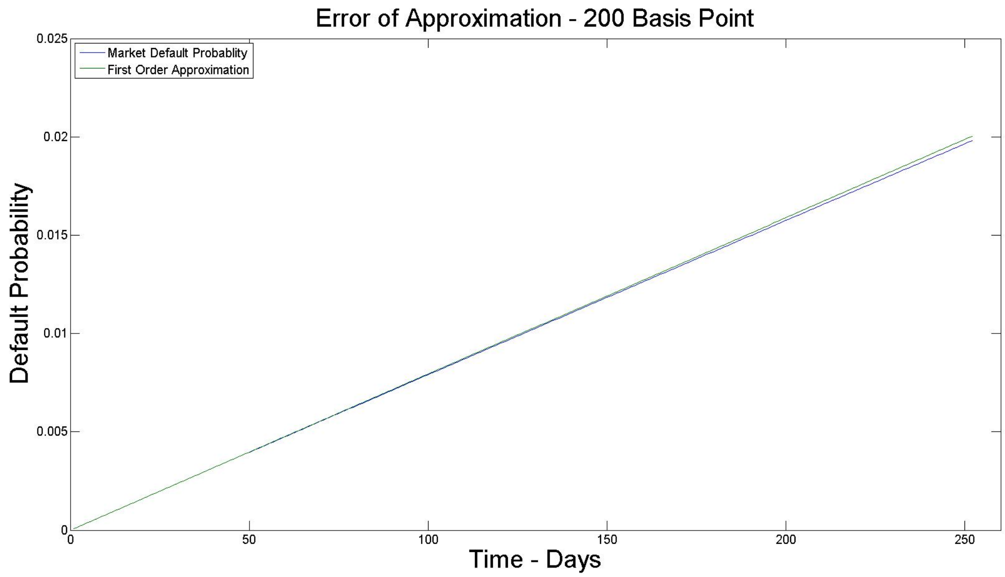


Figure. B.2 *Default Probability - Market CDS Spread 200 Basis Point*

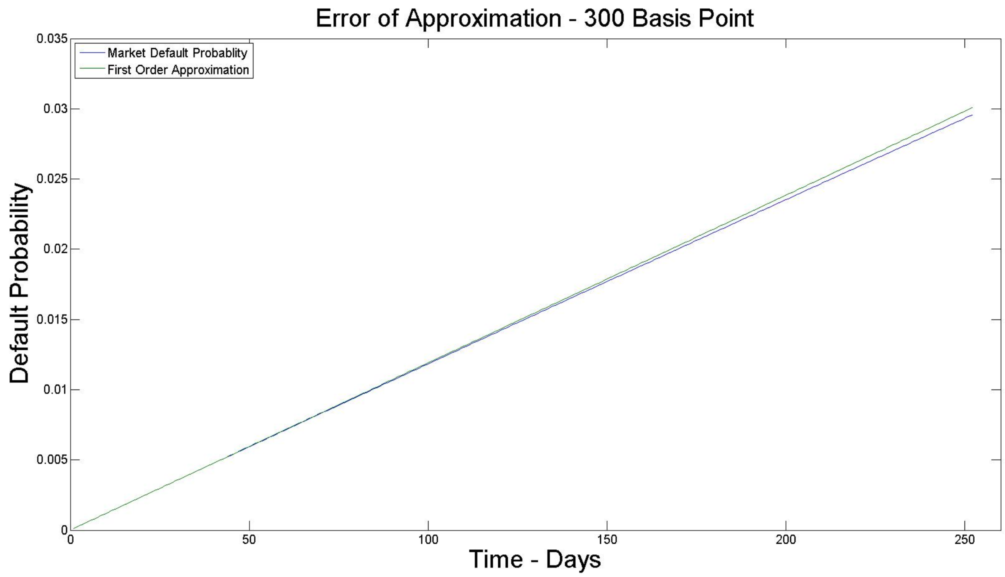


Figure. B.3 *Default Probability - Market CDS Spread 300 Basis Point*

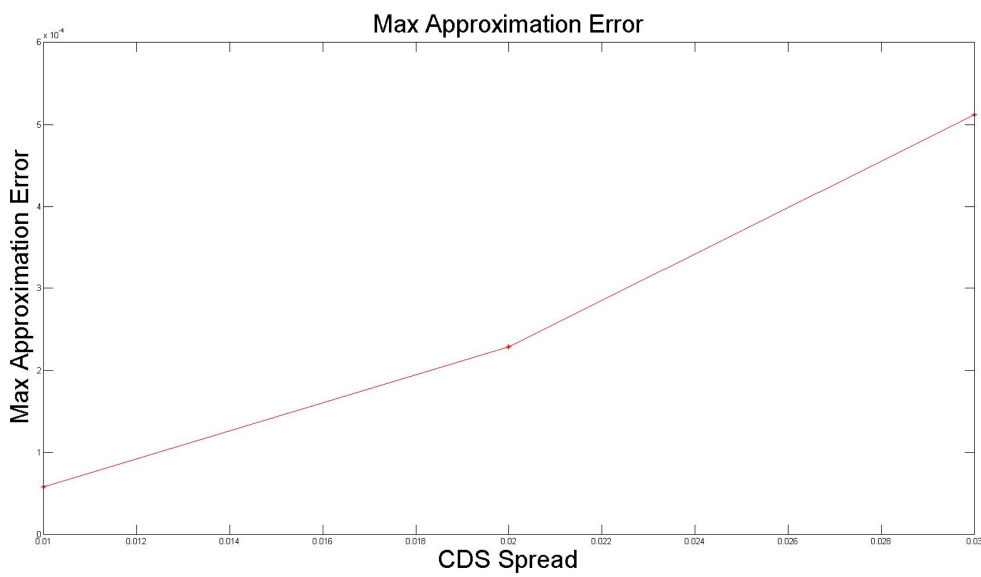


Figure. B.4 *Max Error*

From the first four figures above, we can see that this approximation has small errors that can be

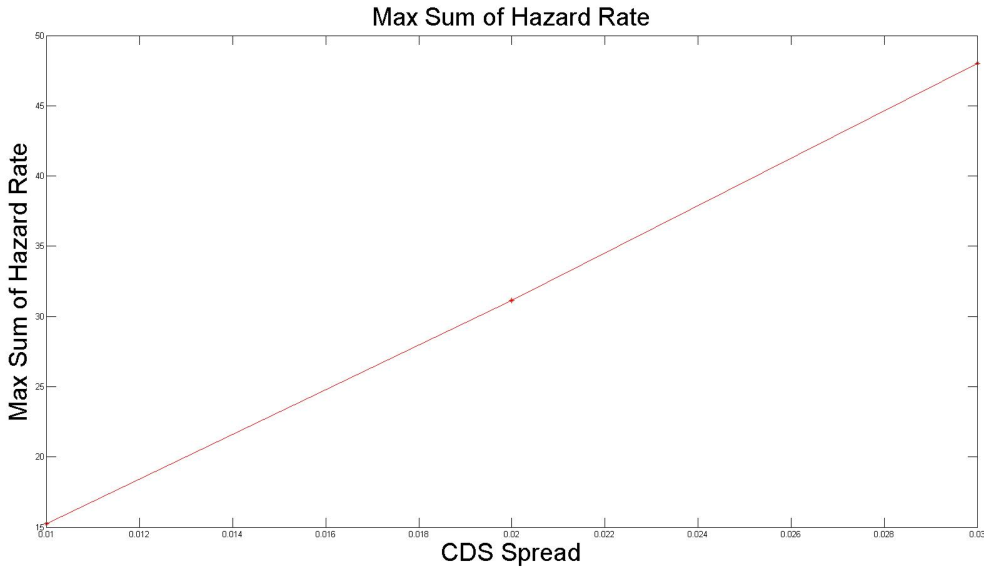


Figure. B.5 *Max Summation of Hazard Rates*

tolerated. Fig. B.5 shows how big the summation of hazard rates can be. The last figure on the next page is a screen-shot of market quoted CDS spread. For one year time horizon, 300 basis point can cover all levels of ratings except Caa/CCC. For three year time horizon, a rating better than Baa3/BBB- can be safely used.

Reuters Corporate Spreads for Industrials*

Rating	1 yr	2 yr	3 yr	5 yr
Aaa/AAA	5	10	15	20
Aa1/AA+	10	15	20	30
Aa2/AA	15	25	30	35
Aa3/AA-	20	30	35	45
A1/A+	25	35	40	50
A2/A	35	44	55	60
A3/A-	45	59	68	75
Baa1/BBB+	55	65	80	90
Baa2/BBB	60	75	100	105
Baa3/BBB-	75	90	110	115
Ba1/BB+	115	125	140	170
Ba2/BB	140	180	210	205
Ba3/BB-	165	200	230	235
B1/B+	190	215	250	250
B2/B	215	220	260	300
B3/B-	265	310	350	400
Caa/CCC	1125	1225	1250	1200

Figure. B.6 *Example of Market Quoted CDS Spread*