PRICING OF COLLATERALIZED DEBT OBLIGATIONS AND CREDIT DEFAULT SWAPS USING MONTE CARLO SIMULATION

by

MARK NEIER

B.S., Kansas State University, 2009

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Industrial and Manufacturing Systems Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2009

Approved by:

Major Professor
Dr. Chih-Hang Wu

Abstract

The recent economic crisis has been partially blamed on the decline in the housing market. This decline in the housing market resulted in an estimated 87% decline in value of collateralized debt obligations (CDOs) between 2007 and 2008. This drastic decline in home values was sudden and unanticipated, thus it was incomprehensible for many investors how this would affect CDOs. This shows that while analytical techniques can be used to price CDOs, these techniques cannot be used to demonstrate the behavior of CDOs under radically different economic circumstances. To better understand the behavior of CDOs under different economic circumstances, numerical techniques such as Monte Carlo simulation can be used instead of analytical techniques to price CDOs. Andersen et al (2005) proposed a method for calculating the probability of defaults that could then be used in the Monte Carlo simulation to price the collateralized debt obligation.

The research proposed by Andersen et al (2005) demonstrates the process of calculating correlated probability of defaults for a group of obligors. This calculation is based on the correlations between the obligors using copulas. Using this probability of default, the price of a collateralized debt obligation can be evaluated using Monte Carlo simulation. Monte Carlo simulation provides a more simple yet effective approach compared to analytical pricing techniques. Simulation also allows investors to have a better understanding of the behaviors of CDOs compared to analytical pricing techniques. By analyzing the various behaviors under uncertainty, it can be observed how a downturn in the economy could affect CDOs. This thesis extends on the use of copulas to simulate the correlation between obligors. Copulas allow for the creation of one joint distribution using a set of independent distributions thus allowing for an efficient way of modeling the correlation between obligors.

The research contained within this thesis demonstrates how Monte Carlo simulation can be used to effectively price collateralized debt obligations. It also shows how the use of copulas can be used to accurately characterize the correlation between obligor defaults for pricing collateralized debt obligations. Numerical examples for both the obligor defaults and the price of collateralized debt obligations are presented to demonstrate the results using Monte Carlo simulation.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my major advisor Dr. John Wu for his support while completing this research. Dr. Wu's knowledge and advice were invaluable with regards to this thesis. Along with Dr. Wu, I would like to thank the students, faculty, and staff of the Kansas State University Industrial and Manufacturing Systems Engineering Department for their support over the last five and a half years. The friendships and memories created during this time will stay with me for the rest of my life.

Dedication

I would like to dedicate this thesis to my parents (Bob and Evelyn Neier), my fiancé Kinley, as well as the rest of my friends and family. Without their unconditional love and support this would not have been possible. I am forever grateful for all they have done for me and helping build me into the person I am today.

Chapter 1 Introduction

1.1 Introduction to Collateralized Debt Obligations and Credit Default Swaps

In the 2002 Annual Report for Berkshire Hathaway, Warren Buffett was quoted as saying "derivatives are financial weapons of mass destruction, carrying dangers that, while now latent, are potentially lethal (Berkshire, 2003)." Now those weapons of mass destruction are no longer latent and have been viewed as the culprits to the current financial crisis which has been compared to an economic Pearl Harbor by Warren Buffett in his 2008 interview with Charlie Rose (Rose & Vega, 2008). To accurately calculate the value of these financial derivatives a new field of study has emerged. This field known as Financial Engineering or Quantitative Finance is the application of engineering and mathematics in the aiding of investment decisions based on the expected worth of derivatives as well as risk control.

Recently the focus of Financial Engineering has revolved around stock options, Credit Default Swaps (CDSs), and Collateralized Debt Obligations (CDOs). Stock options come in various different forms with the two most common being American and European options, also known as "vanilla options." In general there are three main parts to an option. First, there is the expiration date, which is the date at which point the option can no longer be exercised. Second, there is the strike price, which is the amount for which underlying asset will be exchanged. Lastly, there is the type of option, which can be either a put or a call. A put option gives the purchaser the right to sell the underlying asset at the pre-agreed strike price whereas the call option gives the purchaser the right to buy the underlying asset at the pre-agreed strike price.

Collateralized debt obligations and credit default swaps are both credit derivatives. A credit derivative covers a group of financial instruments, whose value is based on the credit risk of the underlying asset (Das, 2005). Primarily, credit derivatives are used to efficiently repackage and transfer risk. Credit derivatives are usually classified into two main formats: funded and unfunded. Funded credit derivatives involve the seller of protection making an initial payment to cover potential credit events, such as defaults. CDOs are an example of a funded credit derivative. An unfunded credit derivative involves both parties making payments based on the underlying contract. A CDS is an example of an unfunded credit derivative (O'Kane, 2001). Figure 1.1 below shows the hierarchy of credit derivatives. The research focus of this thesis is on collateralized debt obligations. The fundamental mechanisms of these two credit derivatives are explained in the next two subsections.

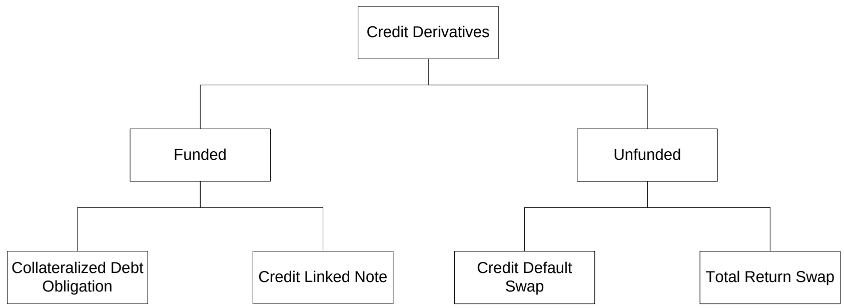

Figure 1.1 Credit Derivatives Hierarchy

adapted from Das 2005

1.1.1 Collateralized Debt Obligations (CDOs)

A CDO is a type of an asset-backed security comprised of a portfolio of assets. CDOs are frequently divided into one to three different pieces called tranches with the three main ones being the junior equity tranche, the mezzanine tranche, and the senior tranche each bearing a certain amount of exposure or risk. Each tranche can be owned by one too many investors. Each tranche i, has a notional principal, P_i, or the original value of the tranche, and an upper and lower default loss level denoted as L_b and L_a , respectively, where $L_b > L_a$. The seller of protection earns a return, also referred to as a spread, based on the notional principal. This is why the CDO is considered a funded credit derivative; the seller of protection makes an initial payment to cover any potential defaults as mentioned in Section 1.1 above. In the beginning when the cumulative losses on the portfolio, L, are less than L_a the spread is earned on the entire P_i. After time when $L_b > L > L_a$, the spread is earned on $P_i(L_b - L)/(L_b - L_a)$, where $P_i(L_b - L)/(L_b - L_a)$ is the remaining value of the tranche. A spread is no longer earned once the losses exceed L_b (Glasserman & Suchintabandid, 2006).

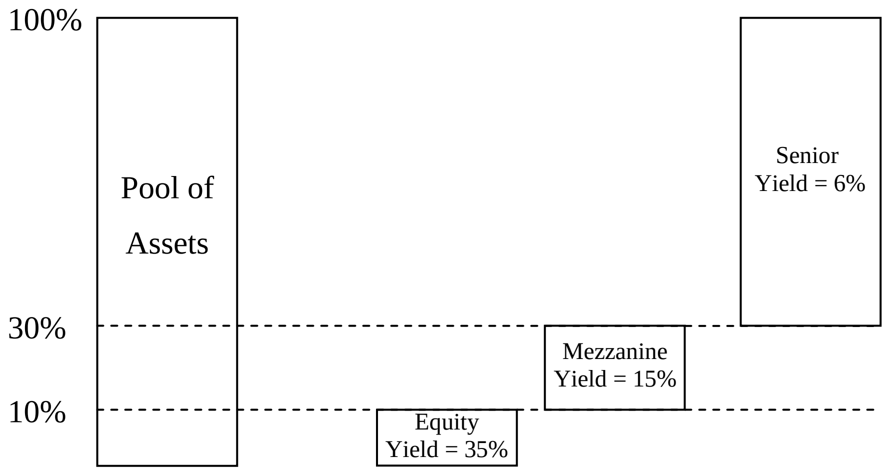

Figure 1.2 CDO Structure

adapted from Glasserman and Suchintabandid 2006.

Illustrated in Figure 1.2 is a CDO with three different tranches. Assuming that the pool of assets was worth $1 million, the notional value of the equity tranche, P_1, would be $100,000 with a L_a=0% and L_b=10%. Similarly, the notional value of the mezzanine tranche, P_2, would be $200,000 with L_a=10% and L_b=30%. Finally, the notional value for senior tranche, P_3, would be $700,000 with L_a=30% and L_b=100%.

Every period, the seller of protection receives a coupon payment, an amount of money based on the remaining notional value. When coupon payments are made, they are paid out in reverse order losses are accumulated. This means the senior tranche is paid first, the mezzanine tranche is paid second, and the equity tranche is paid last, assuming there are still funds available. Since the equity tranche holder is paid last and incurs losses due to defaults first, it receives the highest yield because it contains the highest risk. In Figure 1.2 each tranche is shown to have a yield, y_i. This yield is the amount of interest each tranche will receive in a given period. In this example, the equity tranche will receive an interest payment worth v_1 = 35% of the remaining notional value of the tranche, P_1. Thus in the beginning each tranche will receive coupon payments relative to their perspective yields for the entirety of the value of the tranche, but as defaults occur the value of the tranches will decrease, thus the coupon payments will decrease. As an example, in the beginning when no defaults have occurred, L=0, thus all tranches will receive coupon payments equal to y_iP_i. For example the equity tranche buyers will receive coupon payment of (35%)(100,000) or $35,000 each period, such as a month or a year.

Once a default occurs, say the losses are L=5%, the coupon payments thereafter to the equity tranche will become (35%)($100,000)((10%–5%)/(10%–0%)) or $17,500. Note, that since L=5% is less than L_a for both the mezzanine and senior tranche, both of these tranches will continue to receive coupon payments on each of the respective P_i. If more backed assets are defaulted over time, say the losses are now L=15%, then the equity tranche would be wiped out and receive no more coupon payments since $L > L_b$ and the mezzanine tranche would start bearing the losses and receive discounted coupon payments in the amount of $22,500, or (15%)($200,000) * ((30%–15%)/(30%–10%)).

CDOs have received a great deal of attention and criticisms lately due to the downturn in financial industries. According to the Securities Industry and Financial Markets Association the global value of the CDO market peaked in 2006 at an estimated value of $520 billion and fell to an estimated value of $61 billion in 2008 (Securities, 2009). This can be seen below in Table 1.1.

Table 1.1 Global CDO Value (Securities, 2009)

Year	USD Billions
2004	$157.4
2005	$271.8
2006	$520.6
2007	$481.6
2008	$61.1

This shows that with the economic downturn, the global value of CDOs has decreased drastically since 2006. Part of this is due to the mark to market difference in the notional value of the underlying assets brought on by the drastically decreases in home prices. Even though the value has dropped off lately, credit derivatives are still the fundamental financial instruments for lenders and investors to hedge risks or seek protection. Since there are still significant benefits to the economy using these credit derivatives, they should be evaluated to better understand their behavior.

CDOs were developed around 1987 but did not really take hold until the 1990s, and have since become one of the fastest growing areas in the market. Originally, CDOs were designed for investors to repackage high yield bonds, which lacked liquidity, thus making them hard to sell on the secondary market. These CDOs were also developed because the requirements established by the National Association of Insurance Commissioners' (NAIC) on the reserve required to hold equities, which made these bonds expensive to hold. This was because the requirements established by NAIC required the bank to hold a percentage of capital in relationship to the credit quality. For this reason, the riskier tranches were either sold or transferred to the insurance companies holding companies, also referred to as a special purpose vehicle. Since the holding companies were not subject to the requirements set up by the NAIC, the banks could minimize the capital requirements by holding on to the tranches with the highest credit quality (Das, 2005). Other reasons for structuring CDOs include:

- The transfer of risk
- The creation of tailor-made positions in credit risks
- The funding benefits
- The capital relief due to regulations

To better optimize the return on capital, banks can take either long or short with their positions. Here, going short involves selling the credit risk and thus buying protection. When a bank goes long, they are buying credit risk and are thus selling protection. This works to eliminate the traditional buy and hold strategy and thus improves the return to risk ratio of the credit portfolio (Das, 2005).

To hedge against the risk of default a couple of different strategies can be used. One common approach to hedging the risk to the investor is the standard "bull-bear" approach. This approach is used when the investor believes the default rates will stay relatively low. In this approach the investor sells protection in the equity tranche while buying protection in the senior tranche. Another technique for hedging against defaults is by using individual credit default swaps which can be an expensive strategy (Rajan, McDermott, & Roy, 2007).

CDOs are structured for several difference reasons from the bank's perspective. First, it reduces the amount of capital needed to cover credit portfolios. Second, it allows the bank to remove assets from its balance sheet, thus enhancing the return on equity. Third, for smaller banks, it can help to raise funding. Fourth, it helps to transfer the risk to investors, to help reduce risk.

There are several key drivers of CDOs from the investor's point of view as well. First, the structure of CDOs allows investors to take part in a diversified portfolio easily. Second, it allows investment managers to manage more assets while still being leveraged. Investors can also receive higher returns on CDOs relative to similar credit risk. CDOs also provide investors with structured exposure. An example of this is that in the market highly rated assets, such as those rated AAA, are not very common, but a CDO can create a tranche with a AAA rating even though the underlying assets aren't all AAA. Before the downturn in the economy there was a very low level of defaults, thus many assets were given higher ratings than they deserved due to the history of the market (Weitzner, 2009). Lastly, there remains a certain level of liquidity in the CDO market since they can be sold on the secondary market.

Transactions of CDOs can be issuer driven, investor driven, arbitrage driven, or some combination of the three. Typically, if a CDO is issuer driven, it is to either access funding for the issuer or to transfer some of its risk to an investor. Investor driven CDOs are typically driven by the need to provide tailor made structured exposure to credit risk. CDOs are arbitrage driven when a difference between the market price of the assets and the worth of the assets in a structured form, exists.

The following figure, Figure 1.3, represents a single tranche CDO. The CDO has three main groups involved: the obligors, the special purpose vehicle, and the investor. The special purpose vehicle (SPV) is a stand-alone entity, frequently owned by the bank issuing the underlying CDO, that is used to limit the liability to the issuers. This has a couple of advantages. First, it doesn't require the issuer to hold as much capital to back the assets. Second, if the portfolio were to under-perform it would not affect the issuer. The obligor receives capital from the SPV. In this illustrated example the SPV provides 25 obligors with a $150,000 mortgage each. In return each obligor will make a monthly payment to the SPV in the amount of $997.95. The SPV takes the assets and packages them into a CDO to sell to an investor for a given amount. In this case the SPV will sell all 25 mortgages to the investor for an amount relative to the expected number of defaults. In return, the investor will receive a coupon payment from the issuer of $997.95, less 3% in fees, for each obligor who has not defaulted at this point. Note, the 3% in fees is paid to the issuer in this illustrated case. When an obligor defaults, the cash flows decrease accordingly. In this illustrated case once the first obligor defaults, the SPV receives only 24 coupons and thus pays the investor for only 24 coupons.

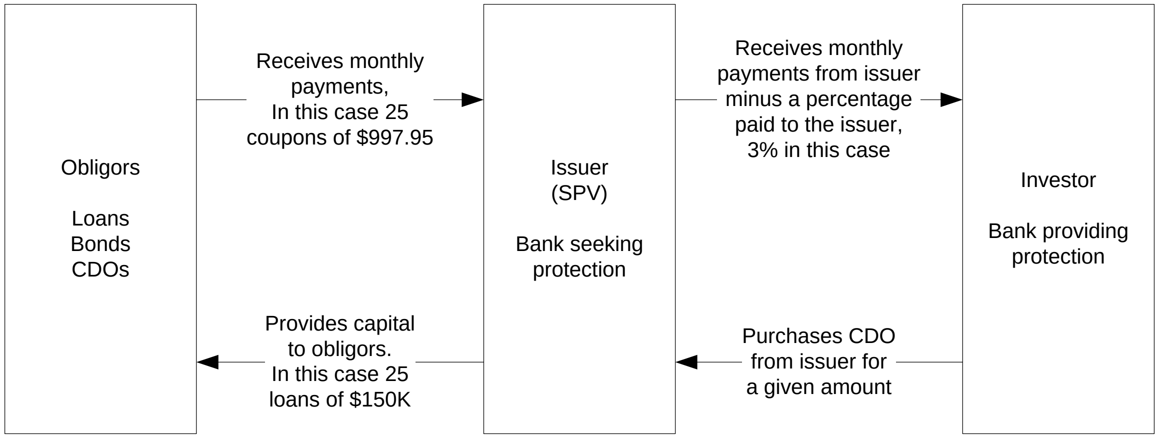

Figure 1.3 How a CDO Operates

CDOs can be structured with various underlying assets from mortgages to other CDOs. The most common types are:

- Loans: Collateralized Loan Obligations
- Bonds: Collateralized Bond Obligations
- Mortgages: Mortgage-backed Securities
- Credit Default Swaps: Synthetic CDO
- CDO tranches: CDO of CDOs or CDO-squared structures

1.1.2 Credit Default Swaps (CDSs)

Credit default swaps (CDSs) are one of the most important credit derivatives and currently make up the majority of the trading volume in the credit derivatives market (Das, 2005). It is estimated that this credit derivative is worth $20 trillion globally (Barrett, 2006). A credit default swap is a structured investment that has a predetermined payout that will pay off if a credit event occurs. The structure of the CDS can be seen below in Figure 1.4.

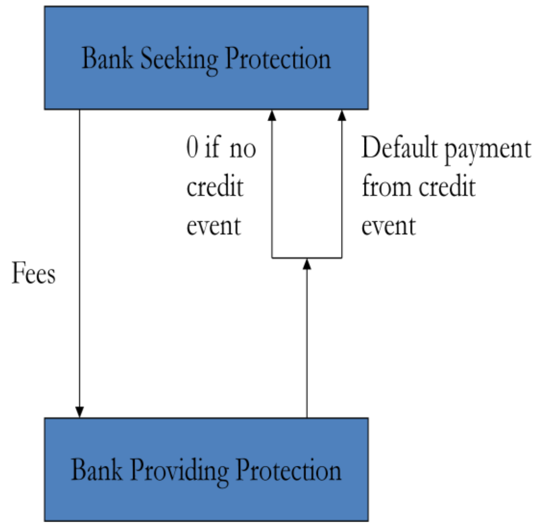

Figure 1.4 Structure of a CDS

adapted from Das 2005

First, the bank providing protection receives fees from the bank seeking protection for a specific entity. In return, if a credit event occurs, the bank providing protection will make a predetermined payment to the bank seeking protection. If a credit event does not occur, then the bank providing protection does not make a payment to the bank seeking protection. In essence, the bank seeking protection is buying insurance on an entity from the bank providing protection. A credit event in most cases is a default. The payment received by the bank seeking protection is usually one of four types: fixed payment, physical settlement, recovery value, or cash settlement. A fixed payment is when the buyer of protection receives a predetermined amount, an estimate of loss, from the provider of protection. The physical settlement is when the protection provider receives an asset from the protection seeker while the protection provider purchases the defaulted security at a given price. The recovery value is when the bank providing protection pays the bank seeking protection the full face value of the swap. The protection buyer will also pay the protection provider for any amounts recovered after the default event. A cash settlement is when the protection seeker receives an amount relative to the change in price of the asset between when the CDS was issued and when the credit event occurred (Das, 2006). In essence, a CDS is the same as a one obligor CDO.

Since CDOs and CDSs can have complicated structures and there is no closed form solution for pricing or evaluating them for most of the structures, a more simplistic method should be used to evaluate these credit derivatives. One method to calculate their worth uses

Monte Carlo simulation. Monte Carlo simulation allows for a better understanding of the behavior of these derivatives.

1.1.3 Monte Carlo Simulation

Simulation works by randomly generating scenarios based on given probabilities to generate possible outcomes. These outcomes can then be analyzed to determine such things as risk and likely results. Boyle (1977), was the first to apply Monte Carlo simulation to the financial markets where he used simulation to value options. Monte Carlo methods for valuing derivatives involves simulating the paths that the security can take as the price evolves based on time, interest rate, or other parameters (Glasserman, 2004). The main issue with using Monte Carlo methods is the fact that it can be rather slow since a large number of paths must be simulated to accurately predict what will happen (Glasserman, 2004). This is because the error rate is equal to $1/\sqrt{c}$ where c is the number of generated paths. To reduce the error rate, more replications must be performed thus taking longer. Monte Carlo simulation is popular because it is easy to implement and is fairly accurate when using sufficient replications.

1.2 Research Motivations

Lately, more people have become skeptical of CDOs because of their involvement in the economic downturn. One of the reasons for the skepticism of CDOs is due to the lack of knowledge about them. Allan Greenspan was recently quoted on a 2009 CNBC documentary "House of Cards," that even he does not understand the complex mathematical formulas used to price CDOs (Weitzner, 2009). This is why an easier technique needs to be developed to better understand the CDO behavior and nature.

Copulas are also needed to represent the correlation between the different obligors. The correlation between obligors is determined based on economic factors. These economic factors include such things as market indexes, economic events such as defaults, presidential elections, natural disasters, geographical location, etc.

The research effort within this thesis extends on Andersen’s probability of default model and uses Monte Carlo simulation to calculate the expected number of defaults, expected losses to a portfolio, and the value of a CDO. Monte Carlo simulation can be used to evaluate CDOs as well as better represent their behavior.

1.3 Research Objectives and Contributions

The purpose of this research study is to test the effectiveness of Monte Carlo simulation in the pricing of collateralized debt obligations. The study will show the simplicity of using Monte Carlo simulation and relate the CDO pricing to the following input parameters: (1) initial asset price, (2) coupon payment value, (3) life of the asset, (4) risk-free interest rate, (5) value at default, and (6) correlation between obligors. The dependent variable will be generally defines as the price of the tranches, and the control variables, interest rate and time period, will be statistically controls in the study. This will be done using MATLAB r2007b to calculate the probability of default, and simulation software, Rockwell Software Arena 12.0 to simulate the defaults to the CDO to show the flexibility in using simulation. The main contribution of this thesis is the exploration of the use of Monte Carlo simulation in conjunction with copulas to represent the correlation between defaults to accurately and straightforwardly price CDOs in a method that allows for the observation of the behavior of the CDO over time.

1.4 Thesis Overview

The remainder of this thesis is presented in the following way. Chapter 2 contains a literature review of Collateralized Debt Obligations and Credit Default Swaps and current methods to determine their prices. Then Chapter 3 discusses the calculations of the probability of default. This includes discussions on both loading matrices and correlation matrices. Chapter 4 details the Monte Carlo simulation used to calculate the value of a CDO while Chapter 5 describes the results of this research. Finally, Chapter 6 provides the conclusion of the thesis.

Chapter 2 Literature Review

When pricing credit derivatives it is important to know when an asset defaults because of the financial losses due to default. In pricing credit derivatives there are two main techniques for determining defaults: reduced form models and structural models. A reduced form model focuses on modeling the probability of default based on market data compared to a structural model which models a default based on if the value of the firm, based on debt and equity, drops below a certain barrier. The two models also differ in what type of data (e.g., available credit/debit information) is available to the modeler. Using structural models implies that the default time is predictable because complete knowledge of all of the firm's assets and liabilities is known while the use of reduced form models infers that the firms default time is unattainable because the only information available is that which is available to the market which is incomplete in reference to all of the current conditions of a firm. The main difference between the two models as mentioned before is what information is available to the modeler. Most of the debates between the two modeling approaches have focused on which model is better at forecasting performance such as the expected number of defaults, instead of if the model should be based off of data observed by the market or not. Data observed by the market should be used in pricing credit risks such as CDOs, thus the reduced form model should be used because complete knowledge of a firms assets and liabilities is very rarely known, which is required to use structural models. Even though the structural model does not seem to be the best approach based on the information available to the modeler, it is still the most popular technique (Jarrow, 2004). Because of this, existing literatures in both the structural and reduced form model will be reviewed and discussed in this chapter.

This chapter is a literature review of some of the existing works on copulas, Monte Carlo simulation, structural models, and reduced form models. Because copulas are used by structural and reduced form models and is used in Monte Carlo simulations, Section 2.1 first discusses the existing literature on the use of copulas to model the correlation between obligors which can be used with either model. This use of copulas is expanded on and used to calculate the probability of default of obligors in Chapter 3. With the information of copulas, Section 2.2 discusses the use of Monte Carlo simulation in pricing credit derivatives, which again can be employed by either model. Section 2.3 introduces pricing credit derivatives with reduced form models while Section 2.4 introduces structural models for pricing credit derivatives.

2.1 Copula Distribution

Being able to accurately represent the correlation between obligors is one of the key aspects of analyzing a credit derivative. More recently this correlation between obligors has been modeled using copula distributions. A copula is a function that obtains a joint distribution with a certain dependence structure using a combination of univariate distributions and can be used with either the structural model or the reduced form model. The theorem that stands as the foundation of copulas is Sklar's theorem, which states that there exists a copula function that relates a given joint multivariate distribution and relevant marginal distributions:

Sklar's Theorem

Let F_{XY} be a joint distribution with marginals F_X and F_Y. Then there exists a function $C:[0,1]^2 \rightarrow [0, 1]$ such that

$$F_{XY}(x,y) = C(F_X(x),F_Y(y))$$

If X and Y are continuous, then C is unique; otherwise, C is uniquely determined on the (range of X)(range of Y).*

Conversely, if C is a copula and F_X and F_Y are distribution functions, then the function F_{XY} is a joint distribution with margins F_X and F_Y.

Though this theorem is for a bi-variate setting it can also be extended to the case of *n* marginal distributions.

Copula's work in the same way as using the CDF of a distribution to generate a random variate. To simulate a univariate distribution, a sample from a uniform $U(0,1)$ distribution can be inverted to obtain the random normal variate obtained from the PDF. For example, using a sample from a uniform $U(0,1)$ distribution, the random normal variate can be obtained by inverting the CDF of the normal distribution using the sample from a uniform $U(0,1)$. This process of generating the random normal variate can be seen below in Figure 2.1. For this example the red line is the uniform $U(0,1)$ sample and the light blue line indicates the random normal variate. Here, the random uniform value is about 0.8 and the random normal variate is about 0.8 as well. In practice, any distribution can be used, however the normal distribution is used here.

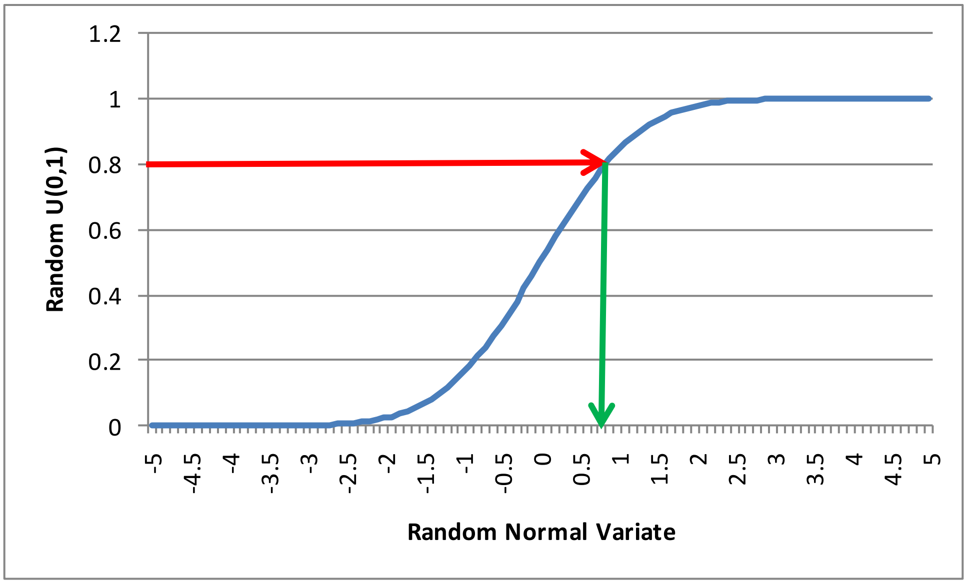

Figure 2.1 Simulation of Univariate Distribution

Copula's are different in the fact that they take this method and convert it from a single distribution to two or more distributions. They also differ in the fact that the CDF on the vertical axis and PDF on the horizontal axis are probabilistically linked and thus are not necessarily linked in a straight-line manner. Here x is the random normal variate and y is the random $U(0,1)$. It should be noted that in such cases as the multivariate normal distribution, the two values, x and y are not linearly related. This is primarily due to the fact the copula is multidimensional (Dorey & Joubert, 2007).

A more formal definition of a Copula is as follows:

C is a Copula if $C:[0,1]^2 \rightarrow [0,1]$ *and*

(i) $C(0,u) = C(v,0) = 0$ for every u, v in [0,1]

(ii) $C(1,u) = C(u,1) = u$ for every u, v in [0,1]

(iii) $C(u_2,v_2) - C(u_1,v_2) - C(u_2,v_2) + C(u_1,v_1) \geq 0$ *for all* $v_1 \leq v_2$, $u_1 \leq u_2$. For every u_1, u_2, v_1, v_2 in [0,1]

Some of the more frequently used copulas in the financial industry are the Gaussian Copula, the Archimedean Copula, and the Student-*t* Copula.

2.1.1 Gaussian Copula

The general definition of the *n*-variate Gaussian copula is as follows:

$$C_{m,\Gamma}(u_1, \ldots, u_m) = N_m[N^{-1}[u_1], \ldots, N^{-1}[u_m]; \Gamma].$$

Here, $N[\bullet]$ is the standard normal function, $N^{-1}[\bullet]$ is the inverse of the standard normal function, and $N[\ldots;\Gamma]$ is the multivariate Gaussian distribution function with correlation matrix

$$\Gamma = (\rho_{ij})_{1\le i,j\le m},$$

and mean of zero. Here, as in most cases, m refers to the number of assets or obligors in the reference portfolio. Through the use of Cholesky decomposition, a positive definite matrix can be created. This decomposition will result in an n x n matrix, A, where $\Gamma=AA^T$. The matrix of A will be a lower triangular matrix with the values representing the covariance matrix of the n variates. This in turn with give a Gaussian copula of the form

$$\mu + AZ \sim N(\mu, \Gamma),$$

where $Z = [Z_1, \ldots, Z_n]$ are normally distributed N(0,1) and are independent. One of the major advantages of the Gaussian is the fact that only linear correlations between assets or obligors need to be calculated (Bluhm & Overbeck, 2007).

2.1.2 Student-t Copula

The second most widely used copula after the Gaussian is the Student-*t* copula. The Student-*t* copula can be defined as follows:

$$C_{m,\Gamma,d}(u_1,\ldots,u_m) = \Theta_{m,\Gamma,d}\left[\Theta_d^{-1}[u_1],\ldots,\Theta_d^{-1}[u_m]\right].$$

Here, $\Theta_{m,\Gamma,d}$ is the multivariate t-distribution with d degrees of freedom and linear correlation matrix Γ, Θ_d is a t-distribution with d degrees of freedom and Θ_d^{-1} is the inverse of the t-distribution with d degrees of freedom.

2.1.3 Marginals

Copulas can be used in combination with a marginal to obtain different forms of correlation. In practice the marginal is the distribution of the individual pricing or defaults whereas the copula is used for the joint probability. The four most common combinations of copulas and marginals are the Gaussian copula with Gaussian marginals, the Gaussian copula with Student-*t* marginals, the Student-*t* copula with Gaussian marginals, and the Student-*t* copula with Student-*t* marginals. Another type of copula and marginal that is not as widely used is the Archimedian otherwise known as the Clayton copula and marginal.

The copula can thus be used to calculate the probabilities of defaults among different obligors. The copula can be used to show the correlation between obligors which can then be used to calculate the correlated probabilities of defaults to be used in the Monte Carlo simulation. This will be discussed further in Chapter 3.

2.2 Monte Carlo Simulation

The research contained within this thesis uses Monte Carlo simulation to price CDOs instead of complex mathematical models. Like Copulas, Monte Carlo simulation can be used for both structural and reduced form models. In the credit derivatives market, Monte Carlo simulation is most commonly known for pricing options. Monte Carlo simulation was first used in 1977 by Boyle in the financial markets to price options. Monte Carlo methods for valuing derivatives involves simulating the paths that the security can take as the price evolves based on time, interest rate, or other parameters (Glasserman et al, 2006). The main issue with using Monte Carlo methods is the fact that it can be rather slow since a large number of paths must be simulated to accurately predict what will happen (Glasserman et al, 2006). This is because the error rate is equal to $1/\sqrt{c}$ where c is the number of generated paths. To reduce the error rate, more replications must be performed thus taking longer. In general the running time for Monte Carlo simulation grows at a rate of $O(MN)$ where N is the number of paths and M is the number of replications.

CDOs are complex credit derivatives that along with the involved underlying assets default behavior make Monte Carlo simulation an attractive choice to be able to accurately capture the CDOs behavior over time. Morokoff (2003) demonstrated how single step approximations and multi-step simulations can produce significantly different results. Monte Carlo simulation allows for the migration of the asset over time, which is significantly more realistic compared to traditional mathematical modeling techniques (Morokoff, 2003).

2.3 Reduced Form Models

A reduced form model focuses on modeling the likelihood of default compared to a structural model which models a default based on if the value of a firm drops below a certain barrier. The main advantage of reduced form models over structural models is their ability to be calibrated to different types of credit structures. Reduced form models are used when the firms default time is inaccessible. These models assume that the information on the set of obligors is less detailed than that of structural models much like that of the observed market. Based on this, reduced form models are better suited for pricing risk.

Glasserman and Suchintabandid (2006) proposed a reduced form model for pricing CDOs through a series of independent-obligor models. This approximation used the multifactor Normal Copula model to capture the default correlation between obligors. The approach they used involved calculating independent defaults using power series expansions to scale the underlying correlations. More specifically, the obligors' correlation matrix is scaled while the tranche price is expanded as a power series with each term in the power series expansion being expressed as a weighted finite sum of independent-obligor prices. In this paper they proposed approximations for weak and strong correlation.

Approximation for weak correlation

We calculate $E\ (L-y)^+$, the expected loss at a certain coupon date with attachment point *y*. To calculate the tranche price $E\ (L-y)^+$ of a dependent obligor model, they use an infinite sum of independent obligor tranche prices. The correlation matrix of $X_1,..., X_M$ is *C* such that:

$$C_t = \begin{bmatrix} 1 & t\rho_{12} & t\rho_{13} & \cdots & t\rho_{1M} \\ t\rho_{21} & 1 & t\rho_{23} & & t\rho_{2M} \\ t\rho_{31} & t\rho_{32} & 1 & & t\rho_{3M} \\ \vdots & & & \ddots & \\ t\rho_{M1} & t\rho_{M2} & t\rho_{M3} & & 1 \end{bmatrix}, \tag{2.1}$$

where $t \in [0,1]$. Since this is the case, *t* is set equal to 0, thus a total independent correlation matrix is constructed. Since the values are independent, the tranche price can be calculated as follows:

$$E_t(L-y)^+ = \delta_0 + \delta_1 t + \delta_2 \frac{t^2}{2!} + ...$$

where $\sum_J w_J \tilde{E}_J (L-y)^+$ approaches δ_n as *s*, the perturbation parameter, goes to 0. Here, $w_J = \frac{\lambda_J}{(2s^2)^n}$. To do this, the exact value of $\tilde{E}_J (L-y)^+$ is computed using the recursive algorithm presented by Andersen et al.

Approximation for strong correlation

For the case of strong correlation, t is no longer set equal to 0, but corresponds to a reference correlation matrix R such that: $C_t = (1-t)R + tC$, $t \in [0,1]$. Where R contains a single factor structure such that for the $(i,j)th$ element $R = \begin{cases} 1, & i = j \\ \gamma_i \gamma_j, & i \neq j \end{cases}$

$$E_t(l-y)^+ \approx \Delta_0 + \Delta_1 t + \Delta_2 \frac{t^2}{2!} + \ldots + \Delta_n \frac{t^n}{n!}$$

$$X \sim N(0, C_t)$$

X can be decomposed into $X_i = \gamma_i Z + \sqrt{1-\gamma_i^2}\,\tilde{X}_i$ where $Z \sim N(0,1)$ and $\tilde{X} = (\tilde{X}_1, \ldots, \tilde{X}_M)$ is a multivariate normal vector.

$$C_t = \begin{bmatrix} 1 & t\sigma_{12} & t\sigma_{13} & \cdots & t\sigma_{1M} \\ t\sigma_{21} & 1 & t\sigma_{23} & & t\sigma_{2M} \\ t\sigma_{31} & t\sigma_{32} & 1 & & t\sigma_{3M} \\ \vdots & & & \ddots & \\ t\sigma_{M1} & t\sigma_{M2} & t\sigma_{M3} & & 1 \end{bmatrix},$$

where $\sigma_{ij} = (\rho_{ij} - \gamma_i \gamma_j) / \sqrt{(1-\gamma_i^2)(1-\gamma_j^2)}$.

Similar to the weak correlation $E_t\left[(l-y)^+ \mid Z\right]$ can be expanded to:

$$E_t\left[(l-y)^+ \mid Z = z\right] = \delta_0(z) + \delta_1(z)t + \delta_2(z)\frac{t^2}{2!} \ldots$$

$$\Delta_k = E\left[\delta_k(Z)\right] = \int_{-\infty}^{\infty} \delta_k(z) \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz.$$

This integration to calculate $\Delta_0, \ldots, \Delta_n$ can be used to approximate $E_t(l-y)^+$ (Glasserman & Suchintabandid, 2006). As we can see, the calculations to price a CDO using this model can be very complex.

2.4 Structural Models

As mentioned in Section 2.3, structural models use a default triggered by the value of the firm dropping below a certain barrier. Using structural models implies that there is complete knowledge of the obligors data set. This assumption implies that the default time for an obligor is predictable.

Structural models were first introduced by Black and Scholes (1973) and Merton (1974). Fischer Black and Myron Scholes (1973) presented the Black-Scholes formula originally to price options in “The Pricing of Options and Corporate Liabilities”. This formula incorporated a random walk, which demonstrated how a stock or derivative can only change in small amounts over time. This, however, only covered stocks and could potentially be used to cover other types of assets as well.

Merton (1974) presented a model that stated that an obligor would default if at the time of maturity the obligor’s debt was larger than its assets. The model by Merton is considered the first model of defaults as well as one of the first structural models, because a structural model determines the time of default based on the evolution of a firms’ debt and equity (Elizalde, 2005). Merton’s model used a diffusion process to calculated the obligor’s asset value, V_t:

$$dV_t = rV_t dt + \sigma_V V_t dW_t,$$

where r is the risk-free rate, σ_V is the volatility and W_t is a Brownian motion at time t. With debt, D, and maturity, T, Merton’s model examined whether $V_T < D$. If this was the case then the obligor would default. This assumes that the obligor can only default at time T. For a case such as a European option this would be fine since the option can only be exercised upon maturity but for the case of a CDO the obligor can default at any time prior to time T. Since the obligor can only default upon maturity, Merton’s model drastically underestimates the total probability of default.

Black and Cox (1976) extended on the Black-Scholes formula and Merton’s model to create a structural model, which could be extended to corporate liabilities. This structural model showed an obligor defaults once the value of its assets reaches a certain barrier assuming volatility and the risk-free rate are constant over time. Black and Cox (1976) used a non constant barrier for determining if an obligor defaulted. The barrier they presented was a time dependent barrier given by:

$$Pe^{-\gamma(T-t)},$$

where P is the constant default threshold and γ is a given rate of change over time. One interesting case is when $\gamma = r$, where r is the risk-free rate. In this case the barrier would then be equal to the face value of the default threshold discounted at the risk-free rate. This structural model was an improvement on Merton's model because it allowed for the obligor to default at any point prior to maturity; which is important in pricing CDOs. However, this only would cover one obligor.

Zhou (2001) proposed a more realistic structural model for calculating the probability of default over time instead of only at maturity which extended on the Black and Cox (1976) structural model. Zhou (2001) also extended the calculations for the probability of default to include the correlation between 2 obligors. However, the main drawback to the work by Zhou is that it only covered the correlation between two obligors and couldn't be used to calculate the correlation between a large numbers of obligors.

In Li (2000) a couple of standard structural models were introduced for the study of default correlation. In this the introduction of a random variable, time-until-default, was created as well as the definition of default correlation between two assets as the correlation between their survival times. The introduction of using copulas in credit studies was discussed as well as how to calibrate the factor in a normal copulas function. Li (2000) is credited with first using copulas in the financial industry to demonstrate the dependence between obligors. This use of copulas was brought from the biostatistics and actuarial sciences area where it was already being used to demonstrate correlation between objects. Li (2000) mentioned that if everything was independent it would be easy to handle any situation in the portfolio. However, this is unrealistic because in a bear market defaults are higher and in a bull market defaults are lower. This demonstrates the existence of macroeconomic influences that obligors share. The addition of copulas to the financial markets was important because it helped introduce the correlation structure between obligors to the portfolio. To do this however, a specific joint distribution of survival times must be determined with given marginal distributions.

Hull, Predescu, and White (2005) demonstrated the use of a structural model in comparison to a reduced form model using the Gaussian copula. Their model was an extension to Zhou's model and addressed the problem with Zhou's work, that Zhou's technique could only be used to calculate the probability of default between 2 obligors. They mentioned the advantage of the structural model is it provides a method for simultaneously modeling defaults and changes in credit rating. However, this structural model did not show a good fit to market prices of CDO tranches. The problem occurs when the correlation parameter is chosen to mirror the price of the equity tranche, the mezzanine tranche became overpriced while the senior tranche became under priced.

Chapter 3 Probability of Default

3.1 Introduction

This chapter introduces the techniques to determine the correlated probabilities of defaults for a set of obligors in an asset pool of collateralized debt obligations. Specifically, this chapter presents a technique for calculating the probability of default between assets while incorporating correlation. To accurately price a CDO, the relationships among a pool of asset obligors needs to be carefully evaluated. These relationships can be explained through the use of a correlation matrix. In practice it is difficult to determine the correlation between two individual obligors but the correlation between the obligors and economic factors can be estimated in a more straight forward manner. The correlations between obligors and factors can be explained through a loading matrix and it can be used to estimate the correlation matrix among a set of obligors.

Section 3.2 discusses the loading matrix and how it is calculated. The loading matrix is introduced first since it is easier to estimate in practice and is then used to calculate the correlation matrix. The correlation matrix will be explained in Section 3.3. Section 3.4 presents a technique for calculating the default time vector, while Section 3.5 discusses how the default time vector can be converted to a monthly probability of default.

3.2 Loading Matrix

The loading matrix, A, can be estimated through the use of factor analysis. The factor analysis works by testing to see if the set of p individual predictor variables, X, can be explained using a set of m different factors denoted as F_i, $i = 1, \ldots, m$. This factor analysis is not intended to predict individual scores for X variables using other X variables but rather to explain patterns in the structure of the data (Warner, 2008). The predictor variables, X, could be measures such as credit rating: AAA, AA, A, BBB, etc. The vector of factors, F, could represent economic indices or events. Economic indices could include such things at the Dow Jones Industrial Average (DJIA), PHLX Semiconductor Index (SOX) the S&P 500(SPX), the Dow Jones Transportation Average (DJTA), etc. Events could include such things as presidential elections, legislation of new bills, oil hitting a certain price, etc. Other factors could include geographical locations, seasons, and so forth.

The use of a path model as seen in Figure 3.1 can be used to visually explain the relationship between the variables and the factors. In this thesis, the factors are denoted F_1 through F_m and the measured variables are denoted $Z(X_i)$. In this research the measured variables, $Z(X_i)$, would indicate the default probabilities of a rated security X_i. The correlations between a measured variable $Z(X_i)$ to factor F_j, is denoted as a_{ij}.

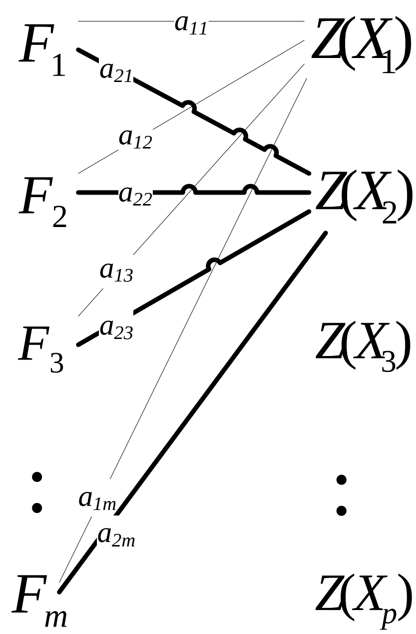

Figure 3.1 Path Model for Loading Matrix

adapted from Warner 2008

The paths model is then used to construct the loading matrix A as seen in Equation 3.1. For example, the path a_{23} is the correlation between the variable $Z(X_2)$ and the factor F_3.

We can then arrange a_{ij} into a matrix format called the loading matrix as follows,

$$A = \begin{array}{c} \\ Z(X_1) \\ Z(X_2) \\ \vdots \\ Z(X_p) \end{array} \begin{array}{c} \begin{array}{cccc} F_1 & F_2 & \cdots & F_m \end{array} \\ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{bmatrix} \end{array} \tag{3.1}$$

The loading matrix, A, is then used to calculate the correlations among the various measured variables $Z(X_i)$ (e.g., the probability of default for mortgage X_i). This is done by taking each path between any two given $Z(X_i)$ variables and multiplying the two path coefficients together as seen in the path model Figure 3.1. An example for the correlation coefficient between variables $Z(X_1)$ and $Z(X_2)$ can be seen below in Equation 3.2.

$$r_{12} = (a_{11}a_{21}) + (a_{12}a_{22}) + (a_{13}a_{23}) + \cdots + (a_{1p}a_{2p}) \tag{3.2}$$

Here, r_{12} is the correlation between variables $Z(X_1)$ and $Z(X_2)$. The correlation based on the path from $Z(X_1)$ to $Z(X_2)$ through factor F_1 is equal to $a_{11}a_{21}$. Likewise, the correlation based on the path from $Z(X_1)$ to $Z(X_2)$ through factor F_2 is equal to $a_{12}a_{22}$. This calculation is repeated for all factors F_j to account for all possible paths (Warner, 2008).

3.3 Correlation Matrix

Once the correlation between all the variables $Z(X_i)$ have been calculated, the correlation matrix can be constructed. The correlation matrix takes on the form in Equation (3.3).

$$R = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & 1 \end{bmatrix} \tag{3.3}$$

The correlations, thus, represent the relationship between the different measured variables instead of the relationship between the variables and factors like the loading matrix. All the diagonal elements are equal to 1. It should be noted that if the set of p variables are uncorrelated, then all of the off diagonal elements of the matrix R are equal to 0.

The elements of the correlation matrix can be either positive or negative, thus they are considered either positively or negatively correlated respectively. If two elements are positively correlated, that means as one element increases the other element increases likewise, if one element decreases, the other element decreases as well. For example, if two houses are located on the same block, there is a high likelihood that these two houses are positively correlated. This is because since they are both in the same neighborhood, if the price of one increases, the value of the other home increases as well. Similarly, if one home mortgage defaults, another home mortgage in the similar location could increase its probability for defaulting. When two elements are negatively correlated they inversely change compared to each other. An example of two variables that are negatively correlated could be home values verse default rates. As home values decrease, the default increases as seen in the recent economy.

The computation of the correlation can be done by taking p factors and N participants and rate the importance of each factor on the participant. Then the correlation can then be defined using Pearson correlation coefficients (Warner, 2008).

In practice, the loading matrix is used more often because it can represent the relationship between the asset and its related factors (market indices, industrial indices, and others). The coefficients of the paths are usually estimated, and can be used to reconstruct the correlation between variables. This is easier to estimate because the relationship between a factor and an asset is far less complex than the correlation between two assets which can have several different factors influencing it. When the correlation matrix is used in practice it is calculated using factor analysis or regression. It is usually estimated using the loading matrix, since the loading matrix is easier to estimate than the correlation matrix.

3.4 Default Time Vector

Andersen, Sidenius, and Basu (2003) presented an algorithm for the pricing of CDOs. Their algorithm works by calculating the random default time vector, evaluate a payout function, and compute the average loss to the portfolio. In this chapter, the algorithm to calculate the random default time vector is discussed, while Chapter 4 examines the simulation used for the payout function, and Chapter 5 discusses the outputs of the simulation which are used to compute the average loss to the portfolio.

The first step of the algorithm is to generate an $M+N$ dimensional sample of uncorrelated Gaussian numbers called $U^T = [X^T \quad e^T]$. The first M of these numbers is used to represent X, while the remaining N numbers are used to represent e by the following equation:

$$Y = cX + e, \tag{3.4}$$

where Y is an N x 1 standard Gaussian vector with correlation matrix Σ, and c is an $N \times M$ loading matrix, e is an $N \times 1$ matrix denoting the residual value, and X is the risk factor such as bond rating. In this research, the estimated correlation between assets is estimated to be 0.25. Next, the correlated sample Y is generated using the equation:

$$Y = AU, \tag{3.5}$$

where $U = [X^T e^T]^T$ and $A = [c \ \sqrt{F}]$. In this thesis, F is initially set to be an identity matrix, and F and c can be iteratively calculated as follows:

1. Calculate $c^i = E^i \sqrt{\Lambda_M^i}$ using a principal components analysis decomposition of $\Sigma - F^i$, where E is a matrix of normalized eigenvectors of $\Sigma - F^i$ and Λ_M^i is a diagonal matrix which contains the M largest eigenvalues of $\Sigma - F^i$.
2. Calculate F^{i+1} where $diag\ F = 1 - diag\ cc^T$
3. Repeat steps 1 and 2 until $tr\left(F^{i+1} - F^i\right)\left(F^{i+1} - F^i\right)^T \rightarrow 0$

Once the correlated sample Y has been calculated, generate a single chi-square value, g, with v degrees of freedom. Then calculate the correlated Student-t vector, Z, using the following equation:

$$Z = Y\sqrt{v/g}. \tag{3.6}$$

Using this Z, the random default time vector can be calculated as:

$$\tau_k = p_k^{-1}\left(t_{1,v}(Z_i)\right), k = 1, \ldots, N, \tag{3.7}$$

where k is the number of assets. These steps have been diagramed into a flowchart that can be seen in Figure 3.2:

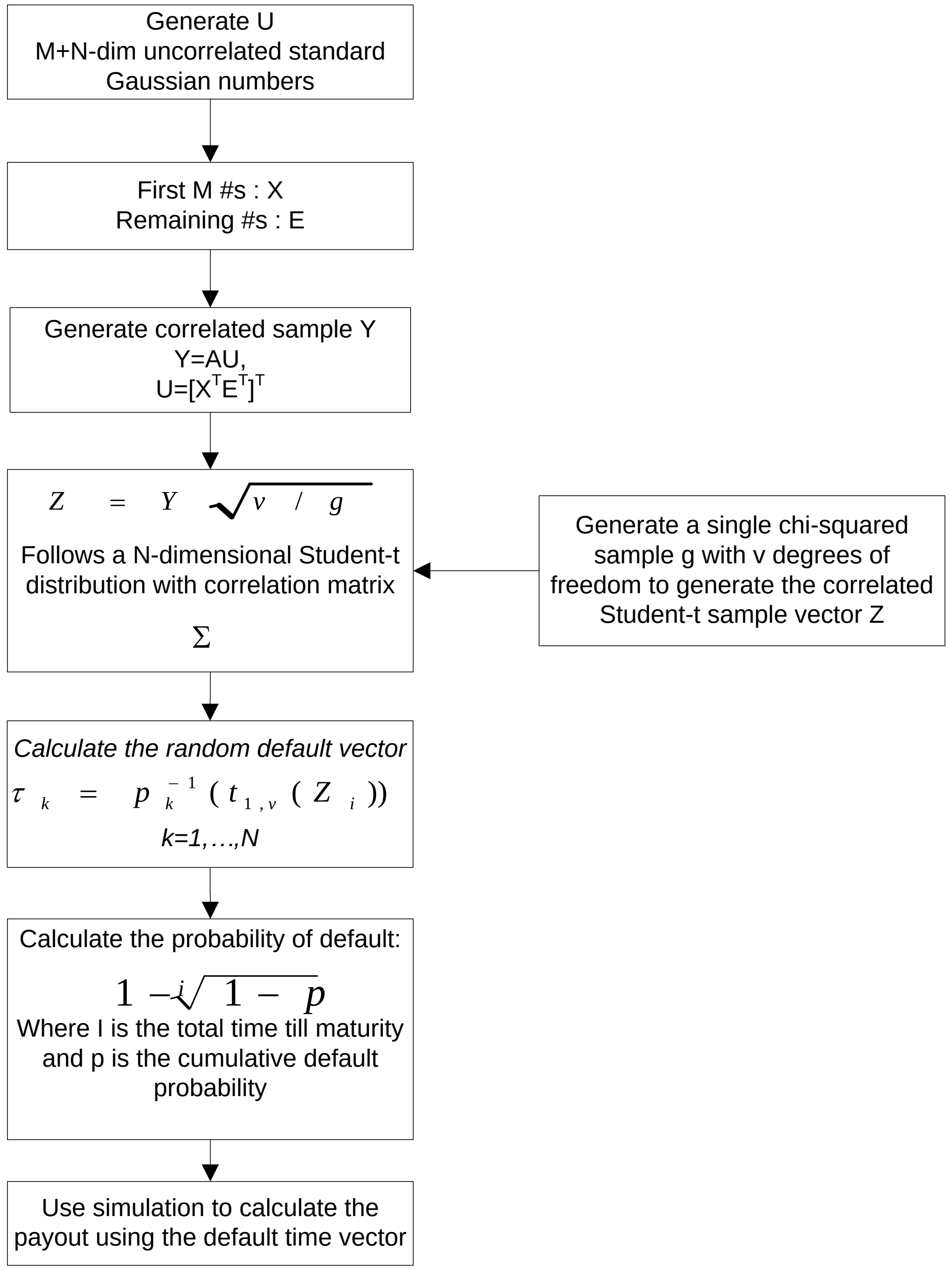

Figure 3.2 Default Time Vector Calculation Flowchart

3.5 Converting Overall Probability of Default to Different Period Lengths

These cumulative probabilities must be converted to monthly probability of defaults, which is geometric in nature. A cumulative default probability, *s*, can be expressed as the summation of the each periods probability of default. The default in the first time period is *p*. The default in the second time period is $(1-p)p$ because there is a $(1-p)$ probability that the default does not occur in the first period and probability *p* that the default occurs in the second period. The probability of default for the third time period is $(1-p)^2p$ because there is a $(1-p)$ probability that the default does not occur in the first period, a $(1-p)$ that the default does not occur in the second period, and a probability of p that the default will occur in the third period. This is then continued for all N periods, 360 in this case. The summation of all the periods can be seen below in Equation 3.8:

$$s = p + (1-p)p + (1-p)^2 p + ... + (1-p)^{N-1} p \tag{3.8}$$

For the simulation, the probability of default of each period is needed, which can be calculated using the cumulative probability of default. To do this, the original equation can be multiplied by $(1-p)$ and subtracted from the original equation at which point p can be solved in terms of s.

$$s = p + (1-p)p + (1-p)^2 p + ... + (1-p)^{N-1} p \tag{3.9}$$

$$s(1-p) = (1-p)p + (1-p)^2 p + (1-p)^3 p ... + (1-p)^N p \tag{3.10}$$

Now subtract Equation (3.9) from Equation (3.10),

$$\begin{aligned} ps &= p - (1-p)^N p \\ \Rightarrow s &= 1-(1-p)^N \\ \Rightarrow 1-s &= (1-p)^N \\ \Rightarrow 1-p &= \sqrt[N]{1-s}\,. \end{aligned} \tag{3.11}$$

Thus, the periodic independent default probability for each month is:

$$p = 1 - \sqrt[N]{1-s}\,. \tag{3.12}$$

These calculations assume that the probability, p, is constant for each time period.

Table 3.1 has the cumulative default probability and the monthly default probability used for the simulation of the portfolio of 25 assets.

Table 3.1 Probability of Default

Entity	Cumulative Default Probability	Monthly Default Probability
1	0.2218	0.0007
2	0.5180	0.0020
3	0.1468	0.0004
4	0.3923	0.0014
5	0.3836	0.0013
6	0.3983	0.0014
7	0.3051	0.0010
8	0.7298	0.0036
9	0.0446	0.0001
10	0.5332	0.0021
11	0.6754	0.0031
12	0.6271	0.0027
13	0.5802	0.0024
14	0.4107	0.0015
15	0.6108	0.0026
16	0.5774	0.0024
17	0.3225	0.0011
18	0.2888	0.0009
19	0.3112	0.0010
20	0.0800	0.0002
21	0.3287	0.0011
22	0.4350	0.0016
23	0.4972	0.0019
24	0.8120	0.0046
25	0.2960	0.0010

This means that for the first asset, there is a 0.07% chance that the default occurs in the current month and a 22.18% chance that the asset will default at some point during the 360 months. Similarly, the second asset has a 0.2% chance of defaulting in the current month and a 51.8% chance of defaulting at some point during the 360 months.

The historic one year average probability of defaults for each rated class of assets was published in a 2005 study by Standard & Poor's. These historic average probabilities were calculated using data between 1981 and 2004 and are summarized below in Table 3.2 (Standard & Poor's, 2005).

Table 3.2 One Year Probability of Default

S&P Rating	Probability of Default
AAA	0.00%
AA	0.01%
A	0.04%
BBB	0.29%
BB	1.28%
B	6.24%
CCC	32.35%

This table means that given a portfolio of 10,000 BBB rated bonds, 29 obligors are expected to default at some point within one yea. This expected number of defaults can be extended out to multiple years using the following equation derived from Equations 3.9 and 3.10 from above to calculate the cumulative probability:

$$s = 1 - (1 - p)^N, \qquad (3.13)$$

where s is the cumulative probability, p is the yearly probability, and N is the number of periods. For example, for a portfolio of 10,000 BBB rated bonds the cumulative probability of default for 30 years would be about 8.34%. This means the expected number of obligors to default over the 30 year period would be about 834.

Chapter 4 Monte Carlo Simulation of CDOs

4.1 Introduction

This chapter extends on the method for pricing collateralized debt obligations. Specifically, this chapter takes the probability of default derived in the previous chapter and uses it in simulating a CDO setting to determine the price of a CDO. Section 4.2 discusses the Monte Carlo simulation used to determine the amount of losses due to default. In Section 4.3, the example outputs of the simulation setups are explained to demonstrate how the CDO prices can be calculated using the simulation results.

In Chapter 3, a technique for calculating the probability of default is evaluated. This technique integrated the correlation between obligors necessary for accurately representing a CDO. A Monte Carlo simulation can then be developed using these probabilities of defaults which are developed in Section 4.2 that can be used to calculate the price for a CDO. Also using these probabilities of defaults, the size of the equity tranche can be developed in Section 4.3.

4.2 Monte Carlo Simulation

To illustrate how to evaluate the price of a CDO, a portfolio of 25 mortgages is simulated. To simplify the illustration, each asset is a 30 year \$150,000 mortgage with a 7% interest rate paid monthly. Thus, the monthly payment, A, is calculated using the following formula:

$$A = P\left[\frac{i(1+i)^N}{(1+i)^N-1}\right], \tag{4.1}$$

where P is the original amount of the mortgage, i is the monthly interest rate, and N is the length of the mortgage. In this example, P is \$150,000, i is 0.583%, and N is 360 months. Given this information, the monthly payment, A, is calculated to be \$997.95. It is also assumed that if an obligor defaults, the asset can be sold for 70% of the original value in the future.

The correlation among the 25 assets is assumed to be 0.25. Based on this assumption, the correlated default probabilities for each asset over the 360 periods can be calculated using the procedure proposed by Andersen, Sidenius, and Basu (2003) as seen in Chapter 3. In addition, the independent default probabilities for each month can be obtained using Equation 3.12 from

Chapter 3. Figure 4.1 illustrates the CDO pricing simulation logics using a flowchart presentation and the logics are explained in detail using an example.

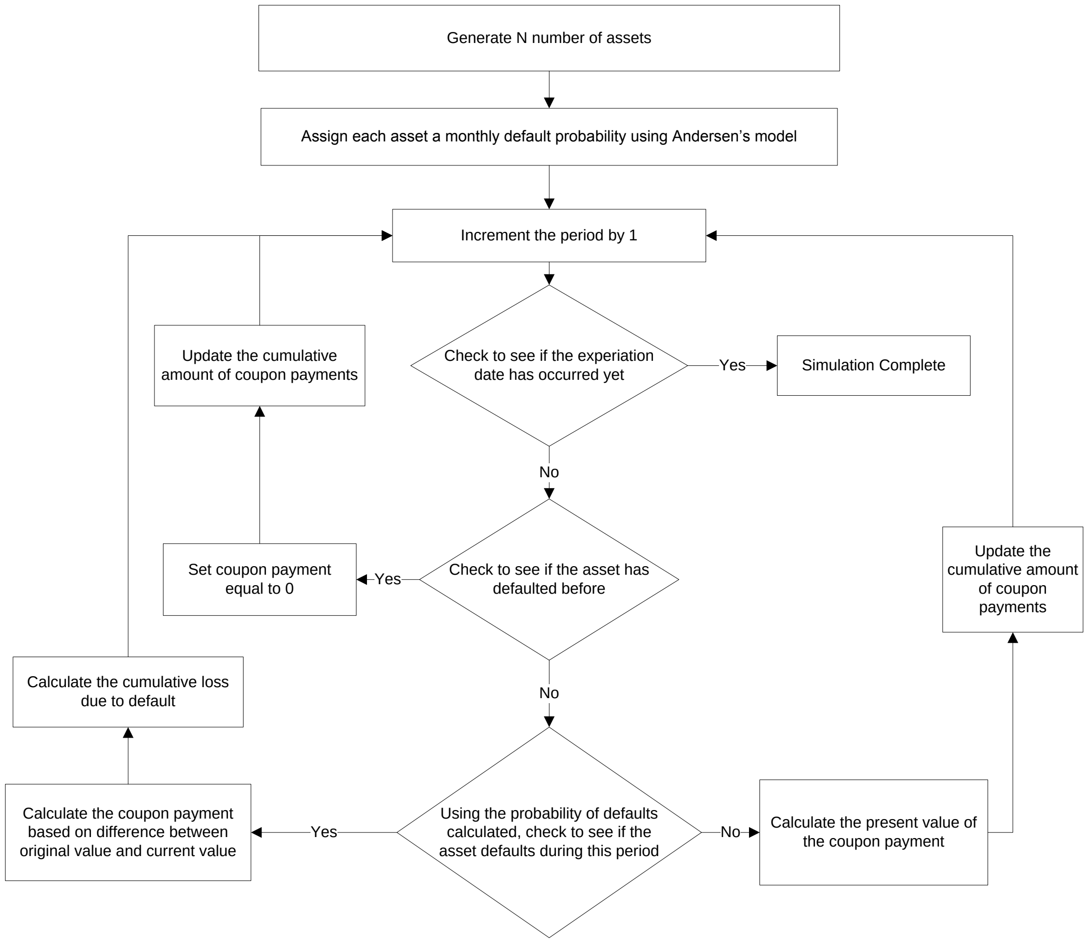

Figure 4.1 Simulation Flow Diagram

To simulate the value of a CDO, 25 assets are created as individual entities at the beginning of the simulation. Each entity contains 4 attributes:

- Asset number
- Period
- Coupon payment amount
- Probability of default

The asset number is used to differentiate between the different obligors and it ranges between 1 and 25 in this example. The period is used to reflect the current month of the mortgage in the simulation. The coupon payment is the discounted value of the monthly mortgage payment based on the period number. Lastly, the probability of default is the probability of defaulting calculated in Chapter 3 Equation 3.12.

The simulation model as a whole uses 3 global variables:

- Cumulative coupon payments
- Cumulative losses to CDO
- Number of defaulted obligors

The cumulative coupon payments variable records a running total of all the coupon payments received from the obligors. The cumulative loss to the CDO is a running total of all the losses associated with defaults for keeping track of the current notional during the simulation. The number of defaulted obligors is the total number of obligors that have defaulted up to the current time period.

To begin the simulation, each entity representing an asset is assigned an asset number and a corresponding probability of default. This probability of default is read into the simulation from a file that is created using the techniques in Chapter 3 proposed by Andersen et al (2003). The period, which starts at 0 is then incremented by 1 for each iteration and is tested to see if the expiration date, 360 in this case, has occurred.

If the expiration date has not been reached, the simulation checks to see if the asset has defaulted in a previous period. If the asset has not yet defaulted, the simulation checks to see if the asset defaults in the current month using the monthly default probability attribute. If the asset does default in the current month, then the loss due to default can be calculated by taking the original value of the mortgage, $150,000, and subtracting the discounted $105,000 from it. This represents a recovery of 70% of the original value in the future and will be discounted back at an annualized rate of 4%. The formula used to calculate the present value given the future value is:

$$P = \frac{F}{(1+i)^N}, \tag{4.2}$$

where F is the future value, i is the interest rate, N is the number of time periods, and P is the present value. The value is then used to update the cumulative loss given default by taking the previous cumulative loss given default value and adding the losses in this period given default. The asset is then flagged as having defaulted and updates the total number of assets to default by taking the previous number of defaulted obligors and incrementing it by 1. After these steps the following outputs are generated for this asset:

- the current period
- the asset number
- the amount of the current coupon payment
- the cumulative amount of coupon payments
- the cumulative amount of losses to the CDO
- the number of assets that have defaulted

These outputs are used later in Section 4.3 to show the cash flows over the 360 periods and to calculate the yield for the equity tranche. Once the simulation writes out the aforementioned information, the asset is then sent through the loop again updating the period, checking to see if the expiration date has occurred yet, etc.

If an asset did not default in the current month, the coupon payment is calculated to account for discounting using Equation 4.2 and updates the cumulative amount of coupon payments received by taking the previous cumulative coupon payments and adding the current coupon payment. The simulation then writes out the information as listed above and goes through the cycle again. This cycle is then repeated until all 360 months have been simulated.

If an asset has defaulted in a previous period, the asset is assigned a coupon payment of $0, increments the cumulative amount of coupon payments by taking the previous total amount of coupon payments made and adding $0 and then the simulation writes out the aforementioned information.

Prior to running the simulation, the expected number of defaults can be calculated by summing the cumulative probability of defaults in Table 3.1 from Chapter 3. For 25 assets, the expected number of defaults is 10.53. Based on 30 replications, the average number of defaults was 9.9 with a standard deviation of 2.77.

4.3 Simulation Outputs

Table 4.1 shows the example outputs from the first period of the first replication of the simulation. The first column is the period, which is 1 for all assets to begin with, and will increment by 1 for each period. The second column is the asset number which there is 25 of them total. The third column is the coupon payment for the current period. As long as an asset has not defaulted this value should be constant for each value and will decrease over time due to discounting. Note that even though this is the first time period, the cash flow of $997.95 needs to be discounted back for one period to its present value of $994.63. The next column takes the previous cumulative coupon payment and adds the current coupon payment to it. Thus, this value increases with time. The fifth column is the cumulative loss and keeps a running total of the sum of the losses due to default. The next column is the portfolio value which will start at $3,750,000 based on the $150,000 initial value for the 25 assets. Each time an asset defaults, this value will decrease by $150,000 with no discounting. The seventh column keeps a running total of the number of assets that have defaulted. The last column shows the monthly probability of default for each asset. Note the monthly probability of default is not constant and in subsequent periods each asset continues to have the same probability of default.

Table 4.1 Example Outputs for 1st Period

Period	Asset Number	Coupon Payment	Cumulative Coupon Payment	Cumulative Loss	Portfolio Value	Number of Assets Defaulted	Monthly Probability of Default
1	1	$994.63	$994.63	$0.00	$3,750,000.00	0	0.07%
1	2	$994.63	$1,989.27	$0.00	$3,750,000.00	0	0.20%
1	3	$994.63	$2,983.90	$0.00	$3,750,000.00	0	0.04%
1	4	$994.63	$3,978.54	$0.00	$3,750,000.00	0	0.14%
1	5	$994.63	$4,973.17	$0.00	$3,750,000.00	0	0.13%
1	6	$994.63	$5,967.81	$0.00	$3,750,000.00	0	0.14%
1	7	$994.63	$6,962.44	$0.00	$3,750,000.00	0	0.10%
1	8	$994.63	$7,957.08	$0.00	$3,750,000.00	0	0.36%
1	9	$994.63	$8,951.71	$0.00	$3,750,000.00	0	0.01%
1	10	$994.63	$9,946.35	$0.00	$3,750,000.00	0	0.21%
1	11	$994.63	$10,940.98	$0.00	$3,750,000.00	0	0.31%
1	12	$994.63	$11,935.61	$0.00	$3,750,000.00	0	0.27%
1	13	$994.63	$12,930.25	$0.00	$3,750,000.00	0	0.24%
1	14	$994.63	$13,924.88	$0.00	$3,750,000.00	0	0.15%
1	15	$994.63	$14,919.52	$0.00	$3,750,000.00	0	0.26%
1	16	$994.63	$15,914.15	$0.00	$3,750,000.00	0	0.24%
1	17	$994.63	$16,908.79	$0.00	$3,750,000.00	0	0.11%
1	18	$994.63	$17,903.42	$0.00	$3,750,000.00	0	0.09%
1	19	$994.63	$18,898.06	$0.00	$3,750,000.00	0	0.10%
1	20	$994.63	$19,892.69	$0.00	$3,750,000.00	0	0.02%
1	21	$994.63	$20,887.33	$0.00	$3,750,000.00	0	0.11%
1	22	$994.63	$21,881.96	$0.00	$3,750,000.00	0	0.16%
1	23	$994.63	$22,876.59	$0.00	$3,750,000.00	0	0.19%
1	24	$994.63	$23,871.23	$0.00	$3,750,000.00	0	0.46%
1	25	$994.63	$24,865.86	$0.00	$3,750,000.00	0	0.10%

The next table, Table 4.2, shows the first 20 periods for one asset, asset number 13. Since this example is just one asset, the asset number and the monthly probability of default are the same for all time periods. This asset was chosen because it was the first to default. It can be seen that in the 11th time period the default occurs. Prior to this, the cumulative loss had been \$0 as well as the number of assets that had defaulted. In time period 11 however, the coupon payment is –\$48,774 because of the loss due to default. This is calculated by taking \$150,000 and discounting \$105,000 for 11 periods and subtracting it from \$150,000. The total portfolio value decreases by \$150,000 to \$3,600,000 and the cumulative loss is set to –\$48,774 to account for the losses. The number of assets to default is also set to 1 at this point. This also shows that since the asset has defaulted there will not be any more coupon payments made and thus the coupon payment is set to \$0 for all remaining periods following the default.

Table 4.2 Example Outputs for first 20 Periods of Asset 13

Period	Asset #	Coupon Payment	Cumulative Coupon	Cumulative Loss	Portfolio Value	# of Assets Defaulted	Monthly Probability of
1	13	\$994.63	\$12,930.25	\$0.00	\$3,750,000.00	0	0.24%
2	13	\$991.33	\$37,753.16	\$0.00	\$3,750,000.00	0	0.24%
3	13	\$988.04	\$62,493.59	\$0.00	\$3,750,000.00	0	0.24%
4	13	\$984.75	\$87,151.84	\$0.00	\$3,750,000.00	0	0.24%
5	13	\$981.48	\$111,728.16	\$0.00	\$3,750,000.00	0	0.24%
6	13	\$978.22	\$136,222.83	\$0.00	\$3,750,000.00	0	0.24%
7	13	\$974.97	\$160,636.13	\$0.00	\$3,750,000.00	0	0.24%
8	13	\$971.73	\$184,968.32	\$0.00	\$3,750,000.00	0	0.24%
9	13	\$968.50	\$209,219.67	\$0.00	\$3,750,000.00	0	0.24%
10	13	\$965.29	\$233,390.45	\$0.00	\$3,750,000.00	0	0.24%
11	13	-\$48,774.10	\$256,518.85	-\$48,774.10	\$3,600,000.00	1	0.24%
12	13	\$0.00	\$279,570.41	-\$48,774.10	\$3,600,000.00	1	0.24%
13	13	\$0.00	\$302,545.39	-\$48,774.10	\$3,600,000.00	1	0.24%
14	13	\$0.00	\$325,444.04	-\$48,774.10	\$3,600,000.00	1	0.24%
15	13	\$0.00	\$348,266.62	-\$48,774.10	\$3,600,000.00	1	0.24%
16	13	\$0.00	\$371,013.37	-\$48,774.10	\$3,600,000.00	1	0.24%
17	13	\$0.00	\$393,684.55	-\$48,774.10	\$3,600,000.00	1	0.24%
18	13	\$0.00	\$416,280.41	-\$48,774.10	\$3,600,000.00	1	0.24%
19	13	\$0.00	\$438,801.20	-\$48,774.10	\$3,600,000.00	1	0.24%
20	13	\$0.00	\$461,247.18	-\$48,774.10	\$3,600,000.00	1	0.24%

A cash flow plot in Figure 4.2 provides the visual of the total amount of losses and coupon payments for each period. This shows the 9 defaults in red with the first being about \$48,774 in period 11 as mentioned above. This shows how the inflow of payments decreases over time due to discounting and dips slightly every time there is a default.

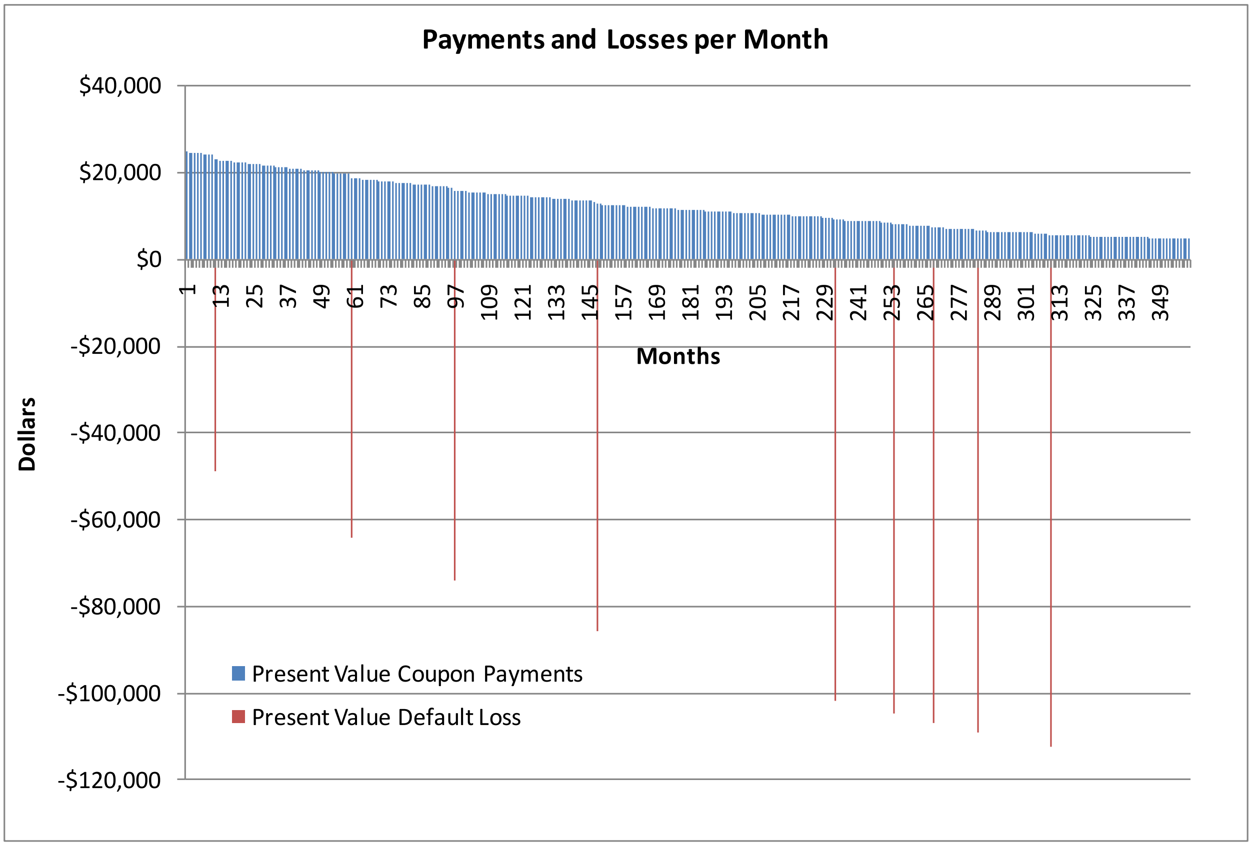

Figure 4.2 Portfolio Cash Flow Revenue and Losses per Month

To price the CDO, the size of the equity tranche was first determined. It was assumed that the notional value of the equity tranche would be equal to the expected number of defaults. In this example, the expected number of defaults according to Table 3.1 in Chapter 3 is 10.53 which is rounded up to 11. The notional value is then $1,650,000, (11 defaults * $150,000). The yield for the equity tranche was calculated so the total yield from the equity tranche would be at least equal to the expected losses to the CDO. This is necessary otherwise nobody would invest in the equity tranche because there would be no financial benefit associated. To calculate the yield that should be paid to the equity tranche holder, the following equation is used:

$$y \sum_{N=1}^{360} p_N - L = 0, \tag{4.3}$$

where y is the monthly yield, p_N is the remaining principal in month N, and L is the cumulative loss due to default. To calculate this, the cumulative loss and the cumulative remaining principal for each replication were written out by the simulation. Both of these values included discounting using Equation 4.2. By dividing the absolute value of the cumulative loss by the cumulative notional value the monthly yield was determined. These values can be seen in Table 4.3 below.

Table 4.3 Monthly and Yearly Yield

Replication	Cumulative Loss	Cumulative Notional	Monthly Yield	Yearly Yield
1	-$807,861.96	$234,719,014.66	0.34%	4.13%
2	-$1,130,107.41	$132,891,292.96	0.85%	10.20%
3	-$725,857.40	$199,457,052.84	0.36%	4.37%
4	-$1,229,170.46	$97,193,308.76	1.26%	15.18%
5	-$1,155,783.29	$138,173,227.60	0.84%	10.04%
6	-$465,597.53	$291,222,053.79	0.16%	1.92%
7	-$847,005.65	$251,550,800.70	0.34%	4.04%
8	-$771,655.98	$219,150,444.88	0.35%	4.23%
9	-$1,332,338.97	$207,457,114.66	0.64%	7.71%
10	-$799,674.73	$231,198,504.11	0.35%	4.15%
11	-$1,024,894.93	$177,270,952.26	0.58%	6.94%
12	-$805,566.04	$233,731,767.84	0.34%	4.14%
13	-$854,684.96	$203,933,718.23	0.42%	5.03%
14	-$691,252.63	$235,496,188.26	0.29%	3.52%
15	-$1,080,059.77	$199,097,203.58	0.54%	6.51%
16	-$814,941.19	$186,843,896.09	0.44%	5.23%
17	-$883,684.79	$165,484,460.41	0.53%	6.41%
18	-$774,096.62	$169,280,732.12	0.46%	5.49%
19	-$1,007,642.97	$218,786,476.01	0.46%	5.53%
20	-$833,580.02	$245,777,781.43	0.34%	4.07%
21	-$1,044,999.88	$94,655,375.70	1.10%	13.25%
22	-$535,739.73	$219,544,829.40	0.24%	2.93%
23	-$914,625.15	$129,598,086.08	0.71%	8.47%
24	-$562,824.00	$231,191,063.89	0.24%	2.92%
25	-$320,443.53	$279,725,022.49	0.11%	1.37%
26	-$621,585.84	$205,539,468.72	0.30%	3.63%
27	-$659,184.87	$170,787,864.14	0.39%	4.63%
28	-$591,547.03	$192,622,779.52	0.31%	3.69%
29	-$1,046,384.74	$235,445,437.42	0.44%	5.33%
30	-$580,421.26	$238,757,885.95	0.24%	2.92%
Average			0.47%	5.60%

By taking the average of the 30 replications, the annual yield was determined. In this scenario the yield was calculated to be 5.6% which means that the SPV must provide a yield rate of 5.6% or more on the equity tranche for the owner of the equity tranche to break even.

An example illustrating the losses to the equity tranche can be seen below in Figure 4.3. In this example there are 30 portfolios simulated. Each simulated CDO has an equity tranche detachment point of $12,900,000 shown as a horizontal red line in the graph. This is chosen because the expected number of defaults for 200 assets is 85.219. Thus $12,900,000 is determined by rounding up the expected number of defaults to 86 and multiplying by the $150,000 for each obligor.

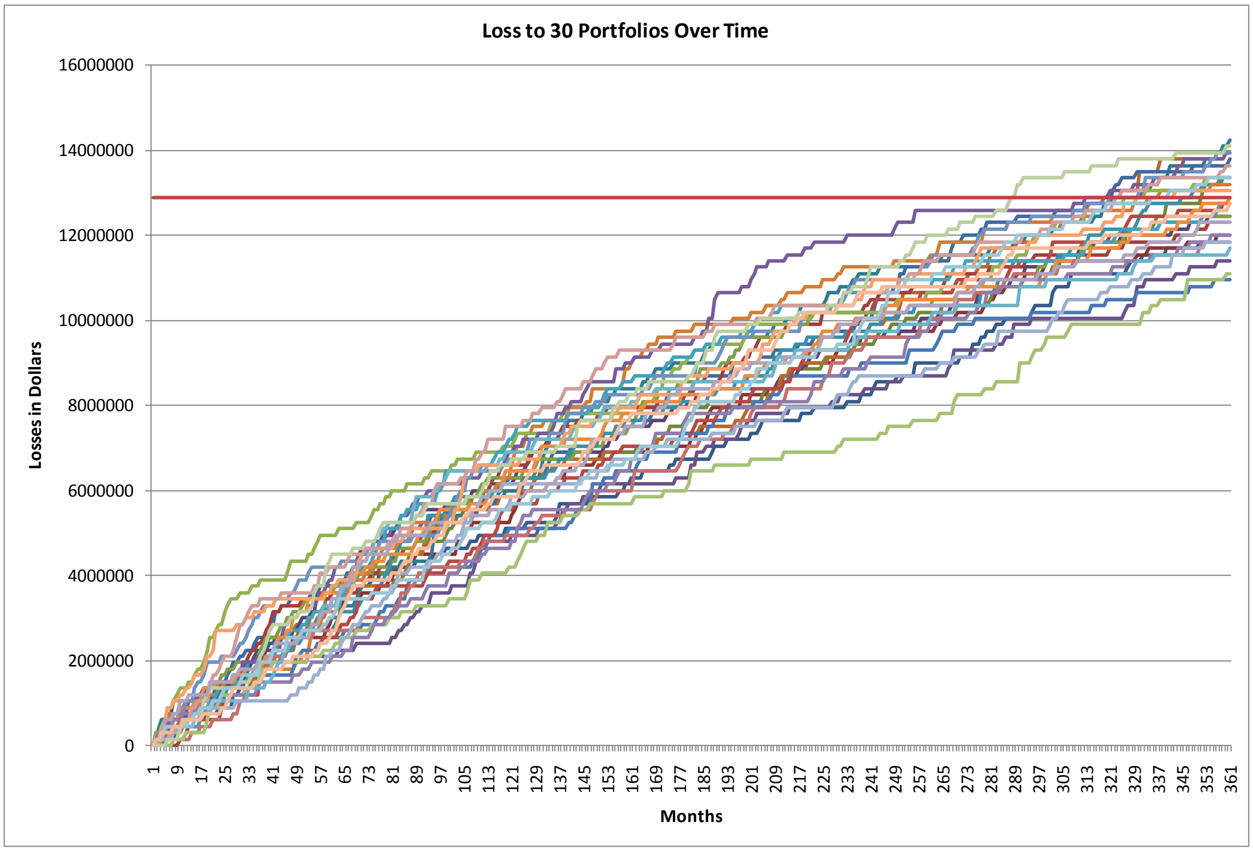

Figure 4.3 Cumulated Losses for 200 Assets over 30 year period- 30 individual simulations

This figure illustrates there are 13 simulations with losses above the equity tranche detachment point. This is used to visualize how a detachment point of an equity tranche is determined using the expected number of defaults. In this illustration, about half of the portfolios should be on each side of the detachment point.

Chapter 5 Simulation Results

5.1 Introduction

The previous two chapters discussed the techniques that are used in this research to calculate the probabilities of default and price the collateralized debt obligations yields via Monte Carlo simulation. This chapter examines the computational results of the Monte Carlo simulation used to price CDOs in two main aspects: precision and accuracy. Specifically, this chapter demonstrates high precision based on the expected number of defaults following a linear trend as the size (i.e., number of obligors) increases, while high accuracy is characterized by constant standard deviation between different sizes. Section 5.2 shows the results for 30 replications of CDOs with 25, 50, 100, and 200 obligors with each replication using the same probability of defaults. This section demonstrates how the Monte Carlo Simulation can produce high precision results across various numbers of obligors. Section 5.3 will show the results for 30 different CDOs with 25, 50, 100, and 200 obligors with each CDO having a different set of probability of defaults. This section demonstrates how the Monte Carlo Simulation produces high accuracy results across various CDOs and sizes. Finally, Section 5.4 concludes with the findings from these computational results.

5.2 Test of Precision and Repeatability

The objective of this section is to prove that the simulation is highly precise based on a constant standard deviation for all sets of obligors. This section evaluates the outputs for CDOs with 25, 50, 100, and 200 obligors when each replication uses the same probability of default. Since the probability of default is the same for each of the 30 replications it is expected that the data should be fairly consistent thus having a low standard deviation. For each CDO size the mean and standard deviation for the number of defaults is calculated based on 30 replications. These values are then used to perform a one-sample *t*-test to determine if the expected mean (μ) and the sample mean $\bar{x}$are the same. The null hypothesis is that the expected number of defaults is equal to the sample mean number of defaults from the simulation:

Null Hypothesis: $\bar{x} = \mu$

Alternative Hypothesis: $\bar{x} \neq \mu$

Equation 5.1 is use to calculate the *t* value based on the one-sample *t*-test.

$$t = \frac{\bar{x}-\mu}{\hat{\sigma}/\sqrt{n}}, \tag{5.1}$$

where $\bar{x}$ is the sample mean, μ is the expected mean, $\hat{\sigma}$ is the sample standard deviation, and n is the number of replications. Then using the t distribution with $n-1$ degrees of freedom, the p value is determined. The calculated p value represents the probability of the obtaining a similar test statistic at least as great as the observed, while assuming the means are equal. If this p value is greater than 0.05 we can fail to reject the null hypothesis that the sample mean, the average number of defaults from the simulation, and the expected mean, the calculated expected number of defaults, are equal. This means that any scenario that has less than a 5% probability of happening is considered extraordinary and would thus reject the null hypothesis.

The size of the four test cases were chosen to provide enough obligors to generate relative data but not too many obligors which would require additional and unnecessary running time to compute the price of the CDO. A common size of a CDO in practice can range from about 50-200 obligors.

5.2.1 25 Obligors

To calculate the expected number of defaults the cumulative probabilities of defaults for each of the 25 obligors is summed together as described in Section 3.5. For the illustrated example with 25 obligors in Table 5.1 the expected number of defaults is 10.53. Table 5.1 contains the cumulative probabilities of defaults as well as the monthly probabilities of defaults. For the illustrated example of 25 obligors the 1st obligor has a 22.18% probability of defaulting at some point during the 360 periods while it has a 0.07% probability of defaulting in a given month. Table 5.1 can be seen at the end of this chapter.

Based on the expected number of defaults for 25 obligors, 10.53, the expected loss to the notional for the equity tranche was determined to be $1,650,000 by rounding up the number of expected defaults to 11 and multiplying by $150,000, the value of each mortgage. This means over 30 years, the expected loss of the total starting value of the CDO will be $1,650,000 due to defaults. On the other hand, using the Monte Carlo simulation, the average number of defaults for 30 replications was 9.9 with a standard deviation of 2.77 with a 95% confidence interval of (8.865, 10.935). Using Equation 5.1 the p value of the one sample t-test was calculated to be 22.3%. Since this p value is more than 5%, we do not have sufficient evidence to reject the null hypothesis that the means are equal. Using Equation 4.3, the estimated monthly yield is calculated to be about 5.6%.

5.2.2 50 Obligors

To calculate the expected number of defaults the cumulative probability for each of the 50 obligors is summed together as described in Section 3.5. For the illustrated example with 50 obligors in Table 5.2 the expected number of defaults was 13.05. Table 5.2 contains the cumulative probabilities of defaults as well as the monthly probabilities of defaults. For the illustrated example of 50 obligors the 1st obligor has a 84.95% probability of defaulting at some point during the 360 periods while it has a 0.52% probability of defaulting in a given month. Table 5.2 can be seen at the end of this chapter.

Based on the expected number of defaults for 50 obligors, 13.05, the expected loss to the notional for the equity tranche was determined to be \$1,950,000 by rounding the number of expected defaults to 13 and multiplying by \$150,000, the value of each mortgage. This means it is expected that over 30 years, \$1,950,000 of the total starting value of the CDO will be lost due to default. Using the Monte Carlo simulation, the average number of defaults for 30 replications was 12.13 with a standard deviation of 3.19 with a 95% confidence interval of (10.941, 13.325). Using Equation 5.1 the *p* value was calculated to be 12.7%. Since this *p* value is more than 5% we fail to reject the null hypothesis that the means are equal. Using Equation 4.3 the estimated monthly yield is calculated to be about 6.14%.

5.2.3 100 Obligors

To calculate the expected number of defaults the cumulative probability for each of the 100 obligors is summed together as described in Section 3.5. For the illustrated example with 100 obligors in Table 5.3 and 5.4 the expected number of defaults was 27.97. Table 5.3 and 5.4 contains the cumulative probabilities of defaults as well as the monthly probabilities of defaults. For the illustrated example of 100 obligors the 1st obligor has a 80.90% probability of defaulting at some point during the 360 periods while it has a 0.46% probability of defaulting in a given month. Table 5.3 and 5.4 can be seen at the end of this chapter.

Based on the expected number of defaults for 100 obligors, 27.97, the expected loss to the notional for the equity tranche was determined to be \$4,200,000 by rounding up the number of expected defaults to 28 and multiplying by \$150,000, the value of each mortgage. This means it is expected that over 30 years, \$4,200,000 of the total starting value of the CDO will be lost due to default. Using the Monte Carlo simulation, the average number of defaults for 30 replications was 27.37 with a standard deviation of 3.88 and a 95% confidence interval of (25.917, 28.816). Using Equation 5.1 the *p* value was calculated to be 40.2%. Since this *p* value is more than 5% we fail to reject the null hypothesis that the means are equal. Using Equation 4.3, the estimated monthly yield is calculated to be about 6.02%.

5.2.4 200 Obligors

To calculate the expected number of defaults the cumulative probability for each of the 200 obligors is summed together as described in Section 3.5. For the illustrated example with 200 obligors in Table 5.5, 5.6, 5.7, 5.8 the expected number of defaults was 85.21. Table 5.5, 5.6, 5.7, 5.8 contains the cumulative probabilities of defaults as well as the monthly probabilities of defaults. For the illustrated example of 200 obligors the 1st obligor has a 98.49% probability of defaulting at some point during the 360 periods while it has a 1.16% probability of defaulting in a given month. Table 5.5, 5.6, 5.7, 5.8 can be seen at the end of this chapter.

Based on the expected number of defaults for 200 obligors, 85.21, the expected loss to the notional for the equity tranche was determined to be \$12,900,000 by rounding up the number of expected defaults to 86 and multiplying by \$150,000, the value of each mortgage. This means it is expected that over 30 years, \$12,900,000 of the total starting value of the CDO will be lost due to default. Using the Monte Carlo simulation, the average number of defaults for 30 replications was 85.03 with a standard deviation of 6.2 and a 95% confidence interval of (82.72, 87.35). Using Equation 5.1 the *p* value was calculated to be 87.7%. Since this *p* value is more than 5% we fail to reject the null hypothesis that the means are equal. Using Equation 4.3, the estimated monthly yield is calculated to be about 6.13%.

These four example cases can all be considered consistent because they all fail to reject the null hypothesis. Failing to reject the null hypothesis means that in all four cases there is not sufficient evidence to say that the expected number of defaults and the estimated number of defaults from the simulation are not statistically equal. For the illustrated example of 100 obligors the estimated number of defaults calculated using the simulation is accurate within ± 1.45 defaults while the illustrate example of 200 obligors is accurate within ± 2.32 defualts. This shows that even for large numbers of obligors the outputs are still rather precise. The running time that is observed for 30 replications ranged from 17 seconds for the illustrated example of 25 obligors to 1 minute 27 seconds for the illustrated case of 200 obligors using a computer with a Pentium 4 processor, 3.00 GHz, and 1.00 GB of RAM.

5.3 Test of Accuracy

The objective of this section is to verify the simulation is accurate. To create an accurate simulation the goal is to simulate multiple CDOs, each CDO using a different set of probabilities of defaults. This section evaluates the outputs for CDOs with 25, 50, 100, and 200 obligors when each replication uses a different probability of default. For this, probabilities of defaults for 30 different CDOs are randomly generated using the techniques discussed in Sections 3.4 and 3.5. Each of these 30 CDOs are then simulated with 30 replications. Using the different probabilities of defaults however, should give a more accurate expected number of defaults since it will cover more than just one circumstance.

5.3.1 25 Obligors

Table 5.9, at the end of this chapter, shows the expected and average number of defaults for 25 obligors. The first column is the CDO number. Each one of these CDOs has its own set of probabilities of defaults and is simulated for 30 replications. Thus the second column, the average number of defaults is for the 30 replications each using the same probabilities of defaults. For each CDO, the sample mean is compared against the expected mean to check and see if the means are equal. Here, a one sample *t*-test is again performed using Equation 5.1 which tested the null hypothesis that the sample mean, average number of defaults according to the simulation, is equal to the expected number of defaults. The last column shows the *p* value for each replication. If the *p* value is larger than 5%, we fail to reject the null hypothesis. For the case of 25 obligors 3 replications rejected the null hypothesis meaning the sample mean and the expected mean for the number of defaults was not equal.

A paired *t*-test is also performed to test whether the expected number of defaults to the estimated number of defaults using simulation. The paired *t*-test is used in this case because each CDO has a pair of data points: the expected number of defaults and the estimated number of defaults. The paired *t*-test creates a 95% confidence interval using the difference between the simulated data and the expected data. If the range of the confidence interval covers 0, then the means of the two samples are statistically equal. For the illustrated case of 25 obligors the 95%

confidence interval of the difference in means was (–0.2577, 0.0241). Because the confidence interval covers 0, we fail to reject the null hypothesis the means are equal. It should be noted that the confidence interval is heavier on the negative side thus meaning the estimated data is consistently lower than the expected number of defaults. The expected number of defaults was 8.92 with an expected standard deviation of 3.23. The observed average number of defaults was 8.83 with a standard deviation of 3.29 over the 30 replications.

The yield for each replication number was calculated using Equation 4.3 and is summarized at the end of this chapter in Table 5.10. For this calculation the equity tranche was set at $1,350,000, (9 defaults x $150,000), based on the expected number of defaults as mentioned above, 8.92.

Figure 5.1 below shows the relationship between the expected number of defaults and the calculated yield. The green trend line shows a power trend with an R^2 of 0.9536. The red trend line is an exponential with an R^2 of 0.9805. Both trend lines accurately represent the data, however the exponential trend line fits the data a little bit better due to the higher R^2 value.

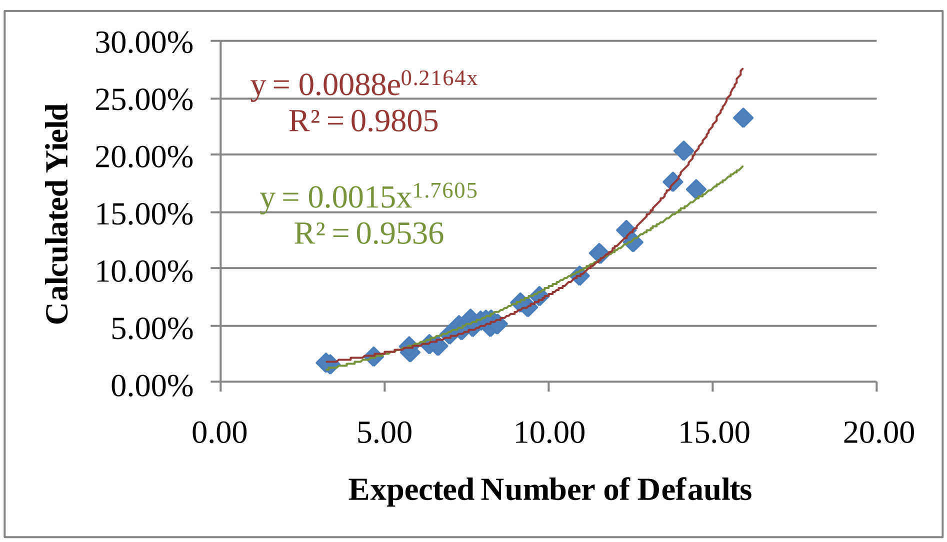

Figure 5.1 Expected Number of Defaults vs. Calculated Yield for 25 Obligors

5.3.2 50 Obligors

Table 5.11, at the end of this chapter, shows the expected and average number of defaults for 25 obligors. The first column is the replication number. Each one of these replications has its own set of probabilities of defaults and is run for 30 replications. Thus the second column, the average number of defaults is for the 30 replications each using the same probabilities of defaults. For each set of replications, the sample mean is compared against the expected mean to check and see if the means are equal. Here, a one-sample *t*-test is performed using Equation 5.1 which tests the null hypothesis that the sample mean, average number of defaults according to the simulation, is equal to the expected number of defaults. The last column shows the *p* value for each replication. If the *p* value is larger than 5%, we fail to reject the null hypothesis. For the case of 50 obligors, 4 replications rejected the null hypothesis meaning the sample mean and the expected mean for the number of defaults was not equal.

For the illustrated case of 50 obligors the 95% confidence interval of the difference in means was (–0.442, –0.007). Because the confidence interval does not cover 0, we reject the null hypothesis that the means are equal. It should be noted that the confidence interval is heavier on the negative side thus meaning the estimated data is consistently lower than the expected number of defaults. The expected number of defaults was 18.45 with an expected standard deviation of 6.85. The observed average number of defaults was 18.23 with a standard deviation of 6.93 over the 30 replications.

The yield for each replication number is calculated using Equation 4.3 and is summarized at the end of this chapter in Table 5.12. For this calculation the equity tranche was set at \$2,850,000, (19 defaults x \$150,000), based on the expected number of defaults as mentioned above, 18.45.

Figure 5.2 below shows the relationship between the expected number of defaults and the calculated yield. The green trend line shows a power trend with an R^2 of 0.9742. The red trend line is an exponential with an R^2 of 0.9894. Both trend lines accurately represent the data, however the exponential trend line fits the data a little bit better due to the larger R^2 value.

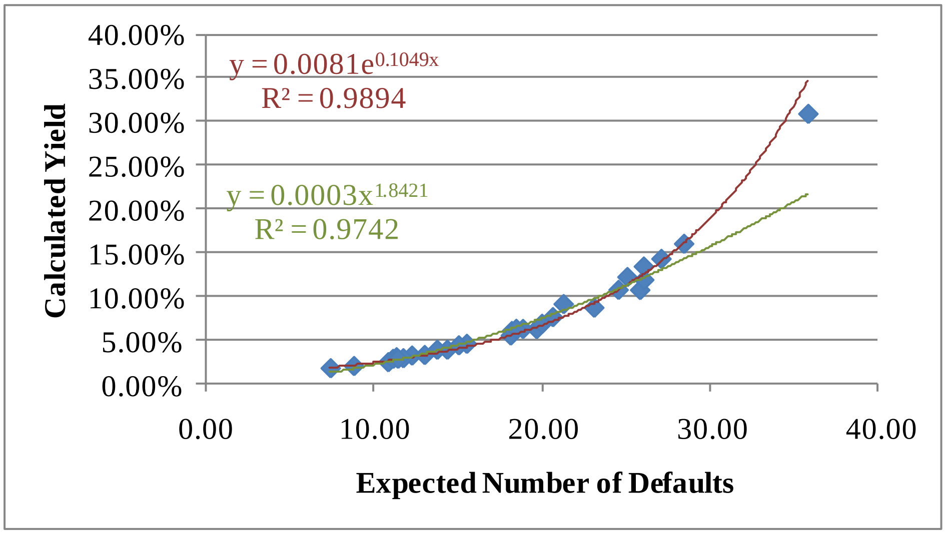

Figure 5.2 Expected Number of Defaults vs. Calculated Yield for 50 Obligors

5.3.3 100 Obligors

Table 5.13, at the end of this chapter, shows the expected and average number of defaults for 100 obligors. The first column is the replication number. Each one of these replications has its own set of probabilities of defaults and was run for 30 replications. Thus the second column, the average number of defaults is for the 30 replications each using the same probabilities of defaults. For each set of replications, the sample mean was compared against the expected mean to check and see if the means were equal. Here, a one-sample *t*-test was performed using Equation 5.1 which tested the null hypothesis the sample mean, average number of defaults according to the simulation, is equal to the expected number of defaults. The last column shows the *p* value for each replication. If the *p* value is larger than 5%, we can fail to reject the null hypothesis. For the case of 100 obligors 1 replication rejected the null hypothesis meaning the sample mean and the expected mean for the number of defaults was not equal.

For the illustrated case of 100 obligors the 95% confidence interval of the difference in means was (–0.508, 0.003). Because the confidence interval covers 0, we fail to reject the null hypothesis that the means are equal. It should be noted the confidence interval is heavier on the negative side thus meaning that the estimated data is consistently lower than the expected number of defaults. The expected number of defaults was 37.33 with an expected standard deviation of 12.80. The observed average number of defaults was 37.08 with a standard deviation of 12.98 over the 30 replications.

The yield for each replication number was calculated using Equation 4.3 and is summarized below in Table 5.14. For this calculation the equity tranche was set at $5,550,000, (37 defaults x $150,000), based on the expected number of defaults as mentioned above, 37.33.

Figure 5.3 below shows the relationship between the expected number of defaults and the calculated yield. The green trend line shows a power trend with an R^2 of 0.9885. The red trend line is an exponential with an R^2 of 0.9918. Both trend lines accurately represent the data, however the exponential trend line fits the data a little bit better due to the higher R^2 value.

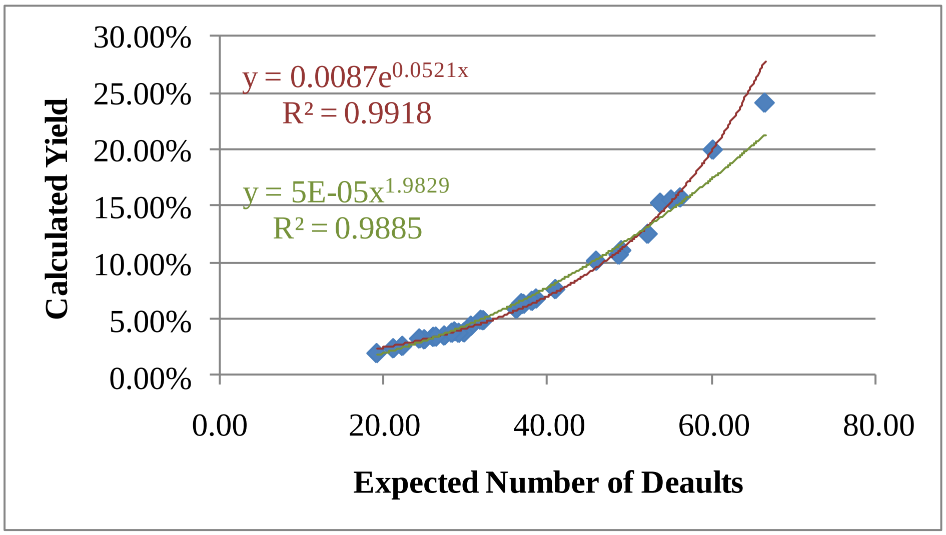

Figure 5.3 Expected Number of Defaults vs. Calculated Yield for 100 Obligors

5.3.4 200 Obligors

Table 5.15, at the end of this chapter, shows the expected and average number of defaults for 200 obligors. The first column is the replication number. Each one of these replications has its own set of probabilities of defaults and was run for 30 replications. Thus the second column, the average number of defaults is for the 30 replications each using the same probabilities of defaults. For each set of replications, the sample mean was compared against the expected mean to check and see if the means were equal. Here, a one-sample t-test was performed using Equation 5.1 which tested the null hypothesis that the sample mean, average number of defaults according to the simulation, was equal to the expected number of defaults. The last column shows the p value for each replication. If the p value is larger than 5%, we can fail to reject the null hypothesis. For the case of 200 obligors 2 replications rejected the null hypothesis meaning that the sample mean and the expected mean for the number of defaults was not equal.

For the illustrated case of 200 obligors the 95% confidence interval of the difference in means was (–0.425, 0.385). Because the confidence interval covers 0, we fail to reject the null hypothesis that the means are equal. The expected number of defaults was 67.24 with an expected standard deviation of 22.44. The observed average number of defaults was 67.22 with a standard deviation of 22.47 over the 30 replications.

The yield for each replication number was calculated using Equation 4.3 and is summarized below in Table 5.16. For this calculation the equity tranche was set at $10,050,000, (67 defaults x $150,000), based on the expected number of defaults as mentioned above, 67.24.

Figure 5.4 below shows the relationship between the expected number of defaults and the calculated yield. The green trend line shows a power trend with an R^2 of 0.9759. The red trend line is an exponential with an R^2 of 0.9849. Both trend lines accurately represent the data, however the exponential trend line fits the data a little bit better due to the higher R^2 value.

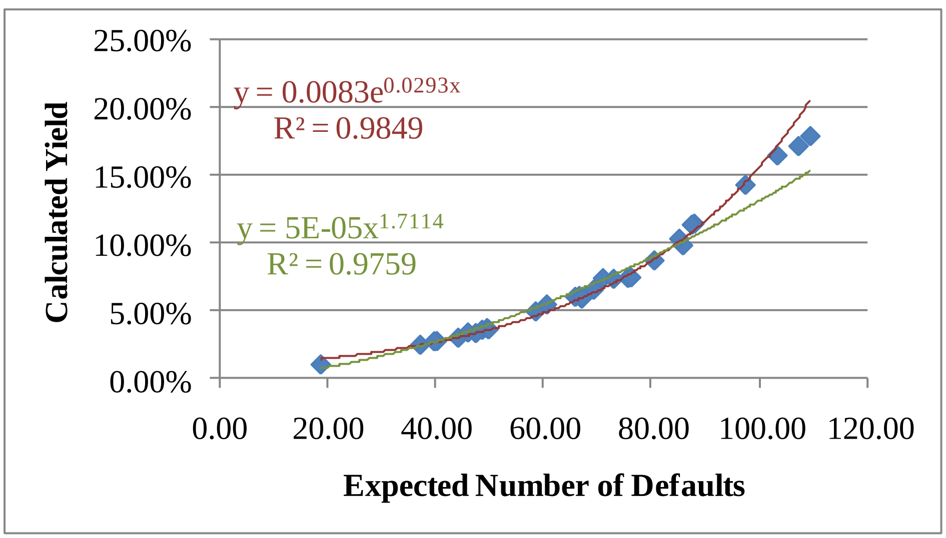

Figure 5.4 Expected Number of Defaults vs. Calculated Yield for 200 Obligors

These four cases all show that both the expected number of defaults and the standard deviation follow a positive correlation with the number of obligors as can be seen below in Table 5.17.

Table 5.17 Average, Standard Deviation, and 95% CI

Number of Obligors	Average Number of Defaults	Expected Number of Defaults	Standard Deviation	95% Confidence Interval
25	8.83	8.92	3.29	(5.656, 11.960)
50	18.23	18.45	6.93	(11.704, 24.749)
100	37.08	37.33	12.98	(23.813, 50.352)
200	67.22	67.24	22.47	(43.166, 91.274)

The standard deviation for the test case when the probabilities of defaults are the same for each replication is expected to be lower than the test case when the probabilities for defaults are different. This is because there is standard deviation due to the simulation and standard deviation due to the different probabilities of default for each CDO for the illustrated example using different probability of defaults. Since the first test case uses the same probabilities of defaults there is no additional standard deviation other than that from the simulation.

These four illustrated examples all show that the relationship between the calculated yield and the expected number of loses is exponential. This is important to note because if the calculated yield is exponential, then as the number of defaults increases the losses to the portfolio increase exponentially. This exponential trend was also seen as these CDO markets imploded in 2008.

5.4 Conclusion

The data for CDOs with the same probabilities of defaults for each replication show that the average numbers of defaults are very precise based on the low standard deviation. The low standard deviation is because there is no difference in the probabilities of defaults between each replication and thus the standard deviation is only due to the simulation. However, using just one set of probability of defaults only gives one situation, thus it does not cover multiple scenarios as shown in Section 5.3. This is evident when plotting the expected numbers of defaults for both cases, same and different probabilities of defaults. This shows that using different probabilities of defaults results in data that follows a linear trend much more closely than data from the same probabilities of defaults and is thus more accurate. Figure 5.5 below shows the linear trends for both cases where the green line shows the trend line when using the same probabilities of defaults while the red line shows the trend line when using different probabilities of defaults.

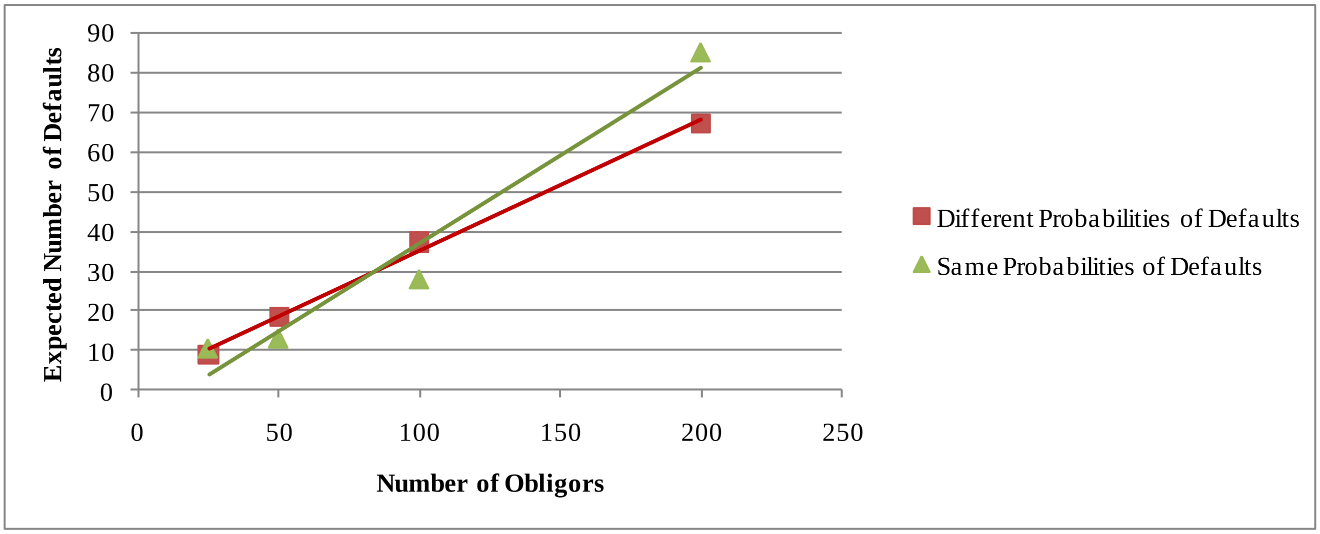

Figure 5.5 Expected Number of Defaults for Same and Different Probabilities of Defaults

Table 5.2 Expected Number of Defaults 25 Obligors

Entity	Cumulative Default Probability	Monthly Default Probability
1	22.18%	0.07%
2	51.80%	0.20%
3	14.68%	0.04%
4	39.23%	0.14%
5	38.36%	0.13%
6	39.83%	0.14%
7	30.51%	0.10%
8	72.98%	0.36%
9	4.46%	0.01%
10	53.32%	0.21%
11	67.54%	0.31%
12	62.71%	0.27%
13	58.02%	0.24%
14	41.07%	0.15%
15	61.08%	0.26%
16	57.74%	0.24%
17	32.25%	0.11%
18	28.88%	0.09%
19	31.12%	0.10%
20	8.00%	0.02%
21	32.87%	0.11%
22	43.50%	0.16%
23	49.72%	0.19%
24	81.20%	0.46%
25	29.60%	0.10%
Total	10.53	

Table 5.3 Expected Number of Defaults 50 Obligors

Entity	Cumulative Default Probability	Monthly Default Probability		Entity	Cumulative Default Probability	Monthly Default Probability
1	84.95%	0.52%		26	28.83%	0.09%
2	7.17%	0.02%		27	21.51%	0.07%
3	12.94%	0.04%		28	1.73%	0.00%
4	13.20%	0.04%		29	26.96%	0.09%
5	27.95%	0.09%		30	74.19%	0.38%
6	9.98%	0.03%		31	51.60%	0.20%
7	10.16%	0.03%		32	21.07%	0.07%
8	1.00%	0.00%		33	15.09%	0.05%
9	12.04%	0.04%		34	38.51%	0.13%
10	38.77%	0.14%		35	20.92%	0.07%
11	25.88%	0.08%		36	17.72%	0.05%
12	26.31%	0.08%		37	32.59%	0.11%
13	27.68%	0.09%		38	70.38%	0.34%
14	36.00%	0.12%		39	0.76%	0.00%
15	18.30%	0.06%		40	49.39%	0.19%
16	14.29%	0.04%		41	9.56%	0.03%
17	46.87%	0.18%		42	61.07%	0.26%
18	20.06%	0.06%		43	0.92%	0.00%
19	37.88%	0.13%		44	24.84%	0.08%
20	10.23%	0.03%		45	32.67%	0.11%
21	26.89%	0.09%		46	19.98%	0.06%
22	54.64%	0.22%		47	25.38%	0.08%
23	11.23%	0.03%		48	4.50%	0.01%
24	35.10%	0.12%		49	8.09%	0.02%
25	24.53%	0.08%		50	12.41%	0.04%
26	0.2883	0.0009		Total	13.05	

Table 5.4 Expected Number of Defaults Obligors 1-50

Entity	Cumulative Default Probability	Monthly Default Probability	Entity	Cumulative Default Probability	Monthly Default Probability
1	80.90%	0.46%	26	11.20%	0.03%
2	27.80%	0.09%	27	86.27%	0.55%
3	0.58%	0.00%	28	52.19%	0.20%
4	11.53%	0.03%	29	25.24%	0.08%
5	8.12%	0.02%	30	15.03%	0.05%
6	26.53%	0.09%	31	25.84%	0.08%
7	16.81%	0.05%	32	52.73%	0.21%
8	15.35%	0.05%	33	94.26%	0.79%
9	6.97%	0.02%	34	34.70%	0.12%
10	21.15%	0.07%	35	56.78%	0.23%
11	1.56%	0.00%	36	49.09%	0.19%
12	46.28%	0.17%	37	43.88%	0.16%
13	2.05%	0.01%	38	62.41%	0.27%
14	7.98%	0.02%	39	72.11%	0.35%
15	18.80%	0.06%	40	24.56%	0.08%
16	31.70%	0.11%	41	14.52%	0.04%
17	20.79%	0.06%	42	30.91%	0.10%
18	85.98%	0.54%	43	62.21%	0.27%
19	83.17%	0.49%	44	70.56%	0.34%
20	46.87%	0.18%	45	9.89%	0.03%
21	26.14%	0.08%	46	58.80%	0.25%
22	6.05%	0.02%	47	18.09%	0.06%
23	11.39%	0.03%	48	14.02%	0.04%
24	1.71%	0.00%	49	43.33%	0.16%
25	13.17%	0.04%	50	59.61%	0.25%

Table 5.5 Expected Number of Defaults Obligors 51-100

Entity	Cumulative Default Probability	Monthly Default Probability		Entity	Cumulative Default Probability	Monthly Default Probability
51	9.42%	0.03%		76	45.70%	0.17%
52	1.38%	0.00%		77	28.05%	0.09%
53	21.58%	0.07%		78	20.07%	0.06%
54	43.81%	0.16%		79	0.08%	0.00%
55	25.86%	0.08%		80	24.03%	0.08%
56	82.00%	0.48%		81	5.29%	0.02%
57	14.14%	0.04%		82	15.68%	0.05%
58	53.95%	0.22%		83	14.21%	0.04%
59	6.33%	0.02%		84	2.95%	0.01%
60	2.35%	0.01%		85	44.93%	0.17%
61	48.56%	0.18%		86	31.39%	0.10%
62	1.76%	0.00%		87	13.69%	0.04%
63	2.16%	0.01%		88	57.97%	0.24%
64	8.85%	0.03%		89	82.02%	0.48%
65	1.41%	0.00%		90	3.29%	0.01%
66	44.04%	0.16%		91	22.33%	0.07%
67	15.40%	0.05%		92	3.27%	0.01%
68	2.20%	0.01%		93	52.09%	0.20%
69	5.16%	0.01%		94	41.63%	0.15%
70	4.24%	0.01%		95	10.34%	0.03%
71	20.94%	0.07%		96	9.60%	0.03%
72	0.14%	0.00%		97	6.72%	0.02%
73	7.51%	0.02%		98	59.35%	0.25%
74	1.76%	0.00%		99	16.10%	0.05%
75	34.81%	0.12%		100	18.70%	0.06%
				Total	27.97	

Table 5.6 Expected Number of Defaults Obligors 1-50

Entity	Cumulative Default Probability	Monthly Default Probability	Entity	Cumulative Default Probability	Monthly Default Probability
1	98.49%	1.16%	26	66.97%	0.31%
2	50.08%	0.19%	27	33.37%	0.11%
3	5.72%	0.02%	28	15.06%	0.05%
4	48.46%	0.18%	29	95.18%	0.84%
5	5.59%	0.02%	30	42.22%	0.15%
6	44.76%	0.16%	31	12.62%	0.04%
7	3.49%	0.01%	32	62.91%	0.28%
8	65.84%	0.30%	33	90.29%	0.65%
9	78.92%	0.43%	34	61.39%	0.26%
10	88.11%	0.59%	35	68.26%	0.32%
11	84.78%	0.52%	36	43.15%	0.16%
12	16.34%	0.05%	37	33.66%	0.11%
13	55.84%	0.23%	38	70.91%	0.34%
14	72.19%	0.35%	39	32.43%	0.11%
15	10.23%	0.03%	40	42.04%	0.15%
16	46.76%	0.17%	41	22.17%	0.07%
17	67.34%	0.31%	42	18.04%	0.06%
18	7.21%	0.02%	43	29.98%	0.10%
19	12.33%	0.04%	44	8.52%	0.02%
20	67.85%	0.31%	45	48.53%	0.18%
21	39.05%	0.14%	46	43.04%	0.16%
22	58.87%	0.25%	47	37.79%	0.13%
23	84.35%	0.51%	48	1.14%	0.00%
24	57.30%	0.24%	49	22.61%	0.07%
25	15.69%	0.05%	50	48.55%	0.18%

Table 5.7 Expected Number of Defaults Obligors 51-100

Entity	Cumulative Default Probability	Monthly Default Probability	Entity	Cumulative Default Probability	Monthly Default Probability
51	53.08%	0.21%	76	55.42%	0.22%
52	45.91%	0.17%	77	24.53%	0.08%
53	37.93%	0.13%	78	40.37%	0.14%
54	4.13%	0.01%	79	28.45%	0.09%
55	51.96%	0.20%	80	74.24%	0.38%
56	34.61%	0.12%	81	39.94%	0.14%
57	57.42%	0.24%	82	69.18%	0.33%
58	61.58%	0.27%	83	65.25%	0.29%
59	36.19%	0.12%	84	31.20%	0.10%
60	51.42%	0.20%	85	35.22%	0.12%
61	27.84%	0.09%	86	43.99%	0.16%
62	28.24%	0.09%	87	15.42%	0.05%
63	26.37%	0.09%	88	30.96%	0.10%
64	12.32%	0.04%	89	36.07%	0.12%
65	47.96%	0.18%	90	11.07%	0.03%
66	81.33%	0.47%	91	73.01%	0.36%
67	20.76%	0.06%	92	9.34%	0.03%
68	22.87%	0.07%	93	17.44%	0.05%
69	65.97%	0.30%	94	8.78%	0.03%
70	40.72%	0.15%	95	76.81%	0.41%
71	87.92%	0.59%	96	18.17%	0.06%
72	3.44%	0.01%	97	3.28%	0.01%
73	10.90%	0.03%	98	69.02%	0.33%
74	16.21%	0.05%	99	86.00%	0.54%
75	95.86%	0.88%	100	64.20%	0.28%

Table 5.8 Expected Number of Defaults Obligors 101-150

Entity	Cumulative Default Probability	Monthly Default Probability	Entity	Cumulative Default Probability	Monthly Default Probability
101	44.39%	0.16%	126	66.68%	0.30%
102	14.15%	0.04%	127	70.16%	0.34%
103	16.47%	0.05%	128	14.77%	0.04%
104	81.19%	0.46%	129	59.18%	0.25%
105	79.16%	0.43%	130	75.87%	0.39%
106	19.99%	0.06%	131	10.59%	0.03%
107	68.79%	0.32%	132	52.81%	0.21%
108	47.34%	0.18%	133	37.95%	0.13%
109	26.65%	0.09%	134	7.98%	0.02%
110	14.72%	0.04%	135	29.76%	0.10%
111	49.97%	0.19%	136	33.60%	0.11%
112	2.74%	0.01%	137	10.08%	0.03%
113	4.36%	0.01%	138	2.10%	0.01%
114	94.46%	0.80%	139	13.15%	0.04%
115	36.69%	0.13%	140	39.07%	0.14%
116	98.29%	1.12%	141	3.29%	0.01%
117	30.80%	0.10%	142	82.02%	0.48%
118	45.69%	0.17%	143	20.19%	0.06%
119	72.57%	0.36%	144	1.68%	0.00%
120	32.98%	0.11%	145	43.36%	0.16%
121	22.31%	0.07%	146	86.91%	0.56%
122	85.80%	0.54%	147	41.21%	0.15%
123	13.93%	0.04%	148	45.27%	0.17%
124	68.84%	0.32%	149	38.30%	0.13%
125	22.61%	0.07%	150	45.70%	0.17%

Table 5.9 Expected Number of Defaults Obligors 151-200

Entity	Cumulative Default Probability	Monthly Default Probability		Entity	Cumulative Default Probability	Monthly Default Probability
151	71.27%	0.35%		176	0.48%	0.00%
152	46.03%	0.17%		177	49.92%	0.19%
153	67.08%	0.31%		178	69.45%	0.33%
154	58.27%	0.24%		179	38.98%	0.14%
155	44.14%	0.16%		180	70.51%	0.34%
156	49.72%	0.19%		181	67.74%	0.31%
157	11.78%	0.03%		182	83.03%	0.49%
158	12.49%	0.04%		183	81.24%	0.46%
159	38.80%	0.14%		184	73.66%	0.37%
160	52.52%	0.21%		185	3.75%	0.01%
161	13.44%	0.04%		186	11.65%	0.03%
162	68.99%	0.32%		187	64.66%	0.29%
163	43.31%	0.16%		188	11.93%	0.04%
164	98.62%	1.18%		189	20.02%	0.06%
165	54.66%	0.22%		190	64.71%	0.29%
166	6.99%	0.02%		191	40.01%	0.14%
167	53.63%	0.21%		192	76.77%	0.40%
168	57.95%	0.24%		193	6.59%	0.02%
169	22.61%	0.07%		194	57.21%	0.24%
170	48.01%	0.18%		195	3.90%	0.01%
171	20.25%	0.06%		196	5.52%	0.02%
172	40.03%	0.14%		197	50.23%	0.19%
173	54.97%	0.22%		198	34.29%	0.12%
174	80.22%	0.45%		199	9.36%	0.03%
175	67.62%	0.31%		200	29.19%	0.10%
				Total	85.21	

Table 5.10 Expected Default, Average Default, and P value for 25 Obligors

CDO Number	Sample Mean	Sample Standard Deviation	Expected Mean	P Value
1	7.74	2.68	7.63	0.818
2	13.61	2.68	14.12	0.313
3	3.35	1.70	3.36	0.989
4	8.17	2.17	8.25	0.828
5	10.80	2.17	10.94	0.718
6	9.00	1.70	9.37	0.242
7	7.80	1.81	8.23	0.201
8	14.43	1.87	14.49	0.861
9	8.27	1.55	7.93	0.249
10	9.13	1.85	9.15	0.972
11	7.63	2.13	8.45	0.044
12	6.03	2.09	6.38	0.378
13	5.13	1.98	5.79	0.079
14	8.00	2.44	8.10	0.832
15	4.63	1.52	4.68	0.860
16	14.00	2.03	13.79	0.567
17	11.80	2.04	11.54	0.494
18	7.63	2.33	8.44	0.069
19	11.63	2.25	11.57	0.886
20	5.83	1.70	5.76	0.816
21	3.67	1.67	3.23	0.163
22	5.93	1.44	6.64	0.012
23	9.83	2.04	9.72	0.770
24	7.00	1.62	6.99	0.974
25	11.77	1.81	12.57	0.021
26	12.43	1.81	12.37	0.847
27	7.57	1.65	7.69	0.675
28	7.60	2.11	7.28	0.407
29	7.30	2.15	7.35	0.892
30	16.50	2.39	15.93	0.199

Table 5.11 Yield for 30 CDOs with 25 Obligors Each

CDO Number	Expected Number of Defaults	Yield
1	7.63	5.59%
2	14.12	20.44%
3	3.36	1.57%
4	8.25	5.51%
5	10.94	9.40%
6	9.37	6.62%
7	8.23	4.88%
8	14.49	17.04%
9	7.93	5.42%
10	9.15	7.03%
11	8.45	5.15%
12	6.38	3.35%
13	5.79	2.63%
14	8.10	5.51%
15	4.68	2.25%
16	13.79	17.69%
17	11.54	11.40%
18	8.44	5.09%
19	11.57	11.32%
20	5.76	3.15%
21	3.23	1.71%
22	6.64	3.20%
23	9.72	7.59%
24	6.99	4.21%
25	12.57	12.36%
26	12.37	13.43%
27	7.69	4.87%
28	7.28	5.03%
29	7.35	4.58%
30	15.93	23.35%

Table 5.12 Expected Default, Average Default, and P Value for 50 Obligors

CDO Number	Sample Mean	Sample Standard Deviation	Expected Mean	P Value
1	7.37	2.25	7.42	0.907
2	9.80	2.51	10.84	0.031
3	18.30	3.24	18.18	0.847
4	11.10	2.52	11.13	0.956
5	11.73	2.03	11.33	0.290
6	14.07	1.82	14.35	0.401
7	12.23	2.40	12.25	0.966
8	11.33	2.44	11.74	0.373
9	11.07	2.68	11.42	0.474
10	25.57	2.45	25.05	0.258
11	21.80	3.20	23.08	0.037
12	24.67	2.70	25.81	0.028
13	18.73	2.24	19.66	0.031
14	22.33	3.38	21.26	0.092
15	27.33	2.86	27.07	0.619
16	15.43	2.73	15.51	0.880
17	23.70	3.32	24.52	0.185
18	17.20	3.24	18.12	0.131
19	12.37	2.75	13.01	0.213
20	8.30	2.45	8.80	0.269
21	18.27	2.59	18.85	0.230
22	13.93	2.92	13.75	0.734
23	19.27	4.03	19.98	0.341
24	20.70	2.84	20.62	0.885
25	18.57	3.17	18.45	0.840
26	28.33	3.09	28.42	0.878
27	25.27	2.57	26.05	0.106
28	36.10	3.02	35.80	0.596
29	15.10	1.79	15.03	0.838
30	26.83	2.28	26.03	0.062

Table 5.13 Yield for 30 CDOs with 50 Obligors Each

CDO Number	Expected Number of Defaults	Yield
1	7.42	1.56%
2	10.84	2.26%
3	18.18	5.85%
4	11.13	2.64%
5	11.33	2.83%
6	14.35	3.70%
7	12.25	3.02%
8	11.74	2.72%
9	11.42	2.65%
10	25.05	12.07%
11	23.08	8.53%
12	25.81	10.57%
13	19.66	6.07%
14	21.26	8.95%
15	27.07	14.16%
16	15.51	4.37%
17	24.52	10.59%
18	18.12	5.33%
19	13.01	3.09%
20	8.80	1.82%
21	18.85	6.09%
22	13.75	3.69%
23	19.98	6.68%
24	20.62	7.45%
25	18.45	6.10%
26	28.42	15.90%
27	26.05	11.74%
28	35.80	30.86%
29	15.03	4.20%
30	26.03	13.28%

Table 5.14 Expected Default, Average Default, and P Value for 100 Obligors

CDO Number	Sample Mean	Sample Standard Deviation	Expected Mean	P Value
1	55.97	5.47	56.25	0.779
2	36.47	3.92	37.15	0.345
3	30.77	3.94	30.74	0.970
4	36.30	3.35	36.28	0.976
5	38.63	4.63	38.68	0.958
6	41.30	3.72	41.04	0.701
7	51.23	4.17	52.30	0.172
8	26.63	3.38	27.49	0.176
9	26.20	3.27	26.14	0.926
10	53.97	3.53	53.83	0.828
11	26.10	3.48	26.46	0.579
12	32.23	4.31	32.24	0.996
13	38.43	3.58	38.17	0.693
14	46.33	4.05	46.01	0.661
15	24.93	3.77	25.05	0.867
16	28.23	3.48	29.90	0.014
17	21.70	3.96	22.37	0.362
18	59.80	2.61	60.25	0.353
19	28.33	3.35	28.72	0.528
20	66.40	3.86	66.57	0.809
21	48.80	3.33	49.08	0.645
22	20.50	2.93	21.26	0.166
23	56.33	4.37	55.16	0.151
24	27.70	4.79	29.25	0.087
25	17.77	2.71	19.24	0.006
26	37.53	4.22	36.89	0.413
27	27.80	3.75	28.40	0.391
28	25.27	4.42	24.44	0.315
29	48.10	3.39	48.77	0.288
30	32.70	3.35	31.93	0.220

Table 5.15 Yield for 30 CDOs with 100 Obligors Each

Replication Number	Expected Number of Defaults	Yield
1	56.25	15.76%
2	37.15	6.35%
3	30.74	4.42%
4	36.28	5.93%
5	38.68	6.82%
6	41.04	7.66%
7	52.30	12.55%
8	27.49	3.56%
9	26.14	3.44%
10	53.83	15.29%
11	26.46	3.46%
12	32.24	4.91%
13	38.17	6.62%
14	46.01	10.16%
15	25.05	3.22%
16	29.90	3.84%
17	22.37	2.64%
18	60.25	19.98%
19	28.72	3.92%
20	66.57	24.13%
21	49.08	11.09%
22	21.26	2.43%
23	55.16	15.57%
24	29.25	3.79%
25	19.24	2.00%
26	36.89	6.41%
27	28.40	3.80%
28	24.44	3.29%
29	48.77	10.70%
30	31.93	4.94%

Table 5.16 Expected Default, Average Default, and P Value for 200 Obligors

CDO Number	Sample Mean	Sample Standard Deviation	Expected Mean	P Value
1	49.73	6.85	49.59	0.911
2	69.10	5.84	69.28	0.864
3	66.47	4.70	66.63	0.849
4	97.63	5.54	97.34	0.771
5	65.40	5.60	65.88	0.639
6	83.93	6.63	85.80	0.135
7	48.70	5.15	48.69	0.992
8	73.53	5.56	75.64	0.047
9	88.87	5.54	87.89	0.343
10	110.17	5.61	109.32	0.417
11	65.83	5.63	67.09	0.231
12	49.17	5.26	49.87	0.470
13	37.80	5.18	37.26	0.569
14	104.07	5.62	103.23	0.423
15	86.13	6.32	85.12	0.386
16	74.40	5.09	76.20	0.062
17	72.30	5.14	71.02	0.182
18	18.33	3.80	18.87	0.443
19	85.80	4.64	87.45	0.061
20	43.50	5.11	44.26	0.419
21	79.33	6.17	80.50	0.309
22	46.77	5.14	47.51	0.433
23	59.00	4.98	58.58	0.646
24	66.90	5.16	67.18	0.767
25	107.73	6.12	107.14	0.597
26	62.70	4.51	60.64	0.018
27	74.53	5.72	72.98	0.147
28	40.83	5.39	40.32	0.607
29	47.20	4.61	46.07	0.190
30	40.73	4.63	39.84	0.300

Table 5.17 Yield for 30 CDOs with 200 Obligors Each

Replication Number	Expected Number of Defaults	Yield
1	49.59	3.72%
2	69.28	6.55%
3	66.63	6.10%
4	97.34	14.29%
5	65.88	6.03%
6	85.80	9.83%
7	48.69	3.60%
8	75.64	7.43%
9	87.89	11.43%
10	109.32	17.89%
11	67.09	5.90%
12	49.87	3.64%
13	37.26	2.49%
14	103.23	16.47%
15	85.12	10.32%
16	76.20	7.47%
17	71.02	7.41%
18	18.87	1.04%
19	87.45	11.35%
20	44.26	3.02%
21	80.50	8.72%
22	47.51	3.36%
23	58.58	4.96%
24	67.18	6.03%
25	107.14	17.15%
26	60.64	5.47%
27	72.98	7.36%
28	40.32	2.77%
29	46.07	3.40%
30	39.84	2.75%

Chapter 6 Conclusions

The recession that has lingered over the last few years has been blamed partially on collateralized debt obligations. One of the reasons for the skepticism of CDOs is due to the lack of knowledge about them. To reduce the skepticism, an easier technique to price CDOs needs to be developed, one that could better represent the nature and behavior of CDOs. The research effort within this thesis extended on the probability of default model described by Andersen et al and then used Monte Carlo simulation to calculate the expected number of defaults, expected losses to a portfolio, and the value of a CDO. This thesis has shown that Monte Carlo simulation along with the implementation of a copula distribution can be used to better represent the nature and behavior of CDOs.

The purpose of this research study was to test the effectiveness of Monte Carlo simulation in the pricing of collateralized debt obligations and credit default swaps. The study showed the simplicity of using Monte Carlo simulation and related the CDO price to the following input parameters: (1) initial asset price, (2) coupon payment value, (3) life of the asset, (4) risk-free interest rate, (5) value at default, and (6) correlation between obligors. The dependent variable was the price of the tranches while the control variables were the interest rate and time period.

The remaining sections of Chapter 6 are as follows. Section 6.1 discusses the conclusions that can be made on the use of copulas in pricing CDOs. Section 6.2 examines the conclusions that are made on the use of Monte Carlo simulation with regards to accuracy and precision. Finally, Section 6.3 discusses future work related to this thesis.

6.1 Conclusion on Copulas

This thesis has shown that copulas can be used as an effective way of including the correlation between obligors when calculating their probabilities of defaults. Chapter 3 demonstrated how the set of correlated probabilities of defaults could be generated using a loading matrix to represent the correlation between obligors and economic factors. This calculation was developed using the process explained by Andersen et al. (2003). The loading matrix used within this thesis assumed the correlation between each obligor and the factor was constant and there was only one factor necessary to determine the correlation between obligors.

6.2 Conclusions on Monte Carlo Simulation

This thesis has demonstrated how Monte Carlo simulation can be used to price collateralized debt obligations. Monte Carlo simulation provides a flexible approach to pricing CDOs that is easy to understand and allows for the observation of the behavior. In this thesis, two different experiments were setup to verify the precision and the accuracy in pricing the CDOs: same probabilities of defaults for each replication and then different probabilities of defaults were tested for each replication. In both scenarios CDOs were priced with 25, 50, 100, and 200 obligors.

In Section 5.2 CDOs were priced using the same probabilities of defaults for each replication. To check for precision, a one-sample *t*-test was conducted for each CDO size: 25, 50, 100, and 200 obligors. This test was used to determine if the estimated number of defaults from the simulation was equal to the expected number of defaults calculated using the cumulative default probabilities of each obligor. All four tests were consistent in lacking sufficient evidence to say that the expected number of defaults and the estimated number of defaults from the Monte Carlo simulations were significantly different. For each illustrated example, a 95% confidence interval was created to demonstrate the precision of the simulation. For the illustrated example of 100 obligors the estimated number of defaults calculated using the simulation was accurate within ± 1.45 defaults while the illustrate example of 200 obligors was accurate within ± 2.32 defaults. This shows that even for large numbers of obligors the outputs are still rather precise. Along with the ease of implementation, the computational effort for Monte Carlo simulation was also shown to be reasonable. The running time that was observed for 30 replications ranged from 17 seconds for the illustrated example of 25 obligors to 87 seconds for the illustrated case of 200 obligors.

In Section 5.3 30 CDOs were priced with each CDO using different probabilities of defaults. This was done as an attempt to observe the accuracy of the simulation. To test for accuracy a paired *t*-test was performed on each of the four illustrated examples. The paired *t*-test was used to compare the estimated number of defaults using simulation to the expected number of defaults for the 30 CDOs generated. Three of the four examples lacked evidence to conclude that the estimated values and the expected values were not the same. However, all of the examples showed that the estimated number of defaults were statistically lower than the expected number of defaults. This was concluded based on the 95% confidence interval on the difference in means being consistently heavier on the negative side of the interval. This illustrated example also showed that the average number of defaults as well as the standard deviation were both positively correlated to the number of obligors and followed a linear trend.

Finally, Figures 5.1 through 5.4 were used to plot the calculated yield verse the expected number of defaults. It was interesting to note that all of these graphs showed an exponential or high-order power trend, meaning that as the expected number of defaults increased, the calculated yield increased exponentially. This was very interesting to note because it helped to explain the drastic market downturns and the mortgage market meltdown behavior in the current economy. This shows that as the number of defaults increases the losses to the portfolio also increase exponentially.

6.3 Future Work

During this research a few questions came up that warrant further research. Five areas involving the use of Monte Carlo simulation with copulas are thought to warrant additional research. These include the use of a memory reduction technique, using multiple factors when calculating the probabilities of defaults using copulas, the application of non constant probabilities of defaults, using a variance reduction technique in the Monte Carl simulation, and modifying the input parameters to accommodate underlying assets with separate input parameters (e.g. initial value, coupon amounts, recovery value, and probability of default). For simplicity in analyzing the Monte Carlo simulation, the input parameters were kept constant in this thesis.

The first concept that could benefit from further research is the addition of a memory reduction technique to the Monte Carlo simulation. For small simulations, ones with a limited number of obligors and few replications (as an example 25 obligors and 30 replications), the Monte Carlo simulation ran relatively fast. However for instances with 200 obligors and 900 replications the simulation ran rather slow, almost 45 minutes in this case. The running time is estimated to be $O(NPM)$ where N is the number of obligors, P is the number of periods, and M is the number of replications. The number of replications has a large impact on the running time, however a sufficient number of replications must still be run to reach an acceptable standard deviation. By not forcing each obligor to loop through this system for each period, the running time could be decreased, but this would come at the expense of no longer being able to observe the behavior of the CDO over time.

In this research the number of factors used to calculate the probabilities of defaults was equal to one for simplicity in calculations. In reality this is not true and should be further researched. It should be evaluated to better understand what types of affects on running time and accuracy the different number of factors causes. While researching this it also might be beneficial to understand how many factors are truly necessary to accurately represent the data.

For simplicity, this thesis only used constant probabilities of defaults. This means that given an obligor the same probability of default would be used for every month. In reality, the monthly probabilities of defaults change over time due to various economic circumstances. To more accurately simulate CDOs with real world conditions, non constant probabilities of defaults should be used in pricing CDOs.

A variance reduction technique to be used in the Monte Carlo simulation should also be evaluated. A variance reduction technique could be used to reduce the number of replications necessary to accurately understand the behavior of a CDO. One variance reduction technique that could be used would be antithetic pairing. Antithetic pairing reduces the variance in Monte Carlo simulation by using negative dependence between pairs replication pairs.

The last area for future research is the enhancement to the Monte Carlo simulation to accommodate different input parameters for the various assets. For simplicity, in this thesis the original asset value, recovery value, coupon value, and correlation between other obligors was kept constant. In practice these values would not be constant for all obligors. To better price a CDO, allowing for various input parameters would give a better representation of real world circumstances.

References

Andersen, L., & Sidinius, J. (2005). Extensions to the Gaussian Copula: Random recovery and random factor loadings. *Journal of Credit Risk*, *1*(1), 29-70.

Andersen, L., Sidinius, J., & Basu, S. (2003). All your hedges in one basket. *RISK*, 67-72.

Barrett, R., & Ewan, J. (2006). BBA Credit Derivatives Report 2006.

Berkshire Hathaway Inc. (2003). *2002 Annual Report.* Retrieved Sept. 18, 2009, Available at http://www.berkshirehathaway.com

Black, F., & Cox, J. (1976). Valuing corporate securities: some effects of bond indenture provision. *The Journal of Finance, 31*(2), 351-367.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, *81*(3), 637-654.

Bluhm, C., & Overbeck, L. (2007). *Structured Credit Portfolio Analysis, Baskets & CDOs*. Boca Raton: Chapman & Hall.

Boyle, P. (1977). Options: A Monte Carlo approach. *Journal of Financial Economics*. 4, 323-338.

Das, S. (2005). *Credit Derivatives CDOs & Structured Credit Products.* Singapore: Wiley Finance.

Dorey, M., & Joubert, P. (2007). Modeling Copulas: An overview. *The Staple Inn Actuarial Society*.

Elizalde, A. (2005). Credit Risk Models II: Structural Models. Available at www.abelelizalde.com.

Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering*. New York: Springer-Verlag.

Glasserman, P., & Suchintabandid, S., (2006). Correlation expansions for CDO pricing. *Journal of Banking & Finance*, *31*, 1375-1398.

Hull, J., Predescu, M., & White, A. (2005) The valuation of correlation-dependent credit derivatives using a structural model. Working Paper, Joseph L. Rotman School of Management, University of Toronto, available at http://www.defaultrisk.com/pp_crdrv_68.htm.

Jarrow, R. (2004). Structural versus reduced form models: A new information based perspective. *Journal of Investment Management*, *2*(2), 1-10.

Li, D.X. (2000, March). On default correlation: A Copula approach. *The Journal of Fixed Income*, 43-54.

Merton, R. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, *29*(2), 449-470.

Morokoff, W. (2003). Simulation methods for risk analysis of collateralized debt obligations. *Winter Simulation Conference.*

O'Kane, D. (2001). *Credit Derivatives Explained*. Lehman Brothers International: Europe.

Rajan, A., McDermott, G., & Roy, R. (2007). *The Structured Credit Handbook*. Hoboken, NJ: John Wiley & Sons, Inc.

Rose, C. (Executive Producer & Host), & Vega, Y. (Executive Producer). (2008, October 1). An exclusive conversation with Warren Buffett [Television series episode]. *Charlie Rose*. New York: WNET.

Securities Industry and Financial Markets Association (2009-01-15). "Global CDO Market Issuance Data". Press release. http://www.sifma.org/research/pdf/CDO_Data2008-Q4.pdf. Retrieved 2009-10-05.

Standard & Poor's; Annual Global Corporate Default Study: Corporate Defaults Poised to Rise in 2005; S&P Global Fixed Income Research, January (2005).

Warner, R. (2008). *Applied Statistics: From Bivariate Through Multivariate Techniques.* Thousand Oaks, CA: Sage Publications, Inc.

Weitzner, M. (Executive Producer). (2009, September 1). *House of Cards* [Documentary]. New York: CNBC.

Zhou, C. (2001). An analysis of default correlations and multiple defaults. *The Review of Financial Studies*, *14*(2), 555.