

2010

Correlation of defaults in complex portfolios using copula techniques

Adam Lodygowski

Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses



Part of the [Applied Mathematics Commons](#)

Recommended Citation

Lodygowski, Adam, "Correlation of defaults in complex portfolios using copula techniques" (2010). *LSU Master's Theses*. 617.
https://digitalcommons.lsu.edu/gradschool_theses/617

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

CORRELATION OF DEFAULTS IN COMPLEX PORTFOLIOS USING COPULA TECHNIQUES

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science

in

The Department of Mathematics

by

Adam Lodygowski

B.S. in Civil Engineering., Poznan University of Technology, 2003

M.S. in Civil Engineering, Poznan University of Technology/University of Hannover, 2005

May, 2010

Acknowledgments

This work, standing as it does at the interface of Mathematics and Finance required me to stretch beyond my previous training and preparation in the field of Civil Engineering and would not have been possible without the help and support of the people I now wish to acknowledge.

First, I would like to warmly thank my advisor Professor Ambar N. Sengupta for his guidance and support. Without him this work would never have been accomplished. To my readers, Professors Eric Hillebrand and Padmanabhan Sundar, whom I would like to thank for their patience in reading this thesis and their many pertinent comments.

Of my peers, I would particularly like to mention Stan Dziobiak and Ben Szozda. I thank Stan for his friendship and help in understanding various mathematical problems. I thank Ben for his help while puzzling on the complicated Stochastic Analysis problems I was facing during this journey.

I would also like to take this opportunity to acknowledge my Civil Engineering advisor, Boyd Professor Dr. George Z. Voyiadjis for his spiritual and financial support. Without him my stay here at LSU would not be possible. I wish to acknowledge the Advanced Computational Solid Mechanics Laboratory (CSM Lab) friends, in particular Danial, for their encouragement and always free will to help.

Further on, I would like to thank my parents, my sisters and my family for their love and support. This was an extraordinary challenge to pursue two graduate degrees at the same time and thanks to their warm words and constant encouragement it was a lot easier to overcome the everyday difficulties and obstacles.

Finally, I thank Agnieszka who gives special sense and meaning to all my efforts. Thank you for constant, unwavering support that helped me through the harder times, for your patience and, above all, for your endless love...

Table of Contents

| | |
|---|----|
| Acknowledgments | ii |
| Abstract | iv |
| 1 Introduction | 1 |
| 2 Theoretical Background | 5 |
| 2.1 Basic Definitions for Copulas | 5 |
| 2.2 Definition of Copulas | 5 |
| 2.3 Sklar's Theorem | 6 |
| 2.4 Copulas and Random Variables | 8 |
| 2.5 Kendall's Tau | 12 |
| 3 Gaussian Model: Simulations and Observations | 14 |
| 3.1 Basic Concepts | 14 |
| 3.2 Model Definition | 20 |
| 3.2.1 Model for a Single Portfolio | 21 |
| 3.2.2 Model for a Single Portfolio with Two Subportfolios | 22 |
| 3.3 Numerical Algorithm | 25 |
| 3.4 Validation of the Implemented Code | 26 |
| 3.5 Obtained Results | 38 |
| 4 Observations and Final Remarks | 48 |
| Bibliography | 51 |
| Appendix- Numerical Algorithm | 52 |
| Vita | 56 |

Abstract

This work, dealing with the correlation between subportfolios in more complex portfolios, begins with a brief survey of the necessary theoretical background. The basic statistical and probabilistic concepts are reviewed. The notion of copulas is introduced along with the fundamental theorem of Sklar. After this background a numerical procedure and code are developed for correlated defaults in multiple correlated portfolio. Further on, interesting results regarding the impact of changes in correlation on the portfolio performance are investigated in the simulations. The most valuable observations regarding the expected default ratios of two subportfolios considered jointly are presented and explained with particular care. These observations are compared with theoretical results. The sensitivity of the tranche losses to correlation parameters is examined carefully. The work is concluded with a brief summary of the most significant observations and their possible impact on portfolio performance.

Chapter 1

Introduction

Copula functions have become the most significant new tool to handle, in a flexible way, co-movement between default risk, risk factors, and other relevant variables important in portfolio performance and risk modeling. While the tool is borrowed from statistical theory, it has been gathering more and more popularity both among academics and practitioners in the field of finance, principally because of the great need to hedge against volatility and erratic behavior of financial markets. One reason for more than simply Black and Scholes formula [4] is the overwhelming evidence of non-normality of the probability distribution of financial asset returns. Some expressions of non-normality has been described using terms such as the "smile effect", which traders now commonly use to formulate strategies, and the "fat-tails" problem, which is a major topic of debate among risk managers and regulators. The result is that nowadays no one address any financial or statistical problem connected to financial markets without taking care of the issue of departures from normality.

People in the field have begun to realize that abandoning the normality assumption for multidimensional problems was a much more involved issue. The multidimensional extension of the techniques devised at the univariate level has also grown all the more as a necessity in market practice. On the one hand, the massive use of derivatives in asset management, in particular from hedge funds, has made a non-normality of returns an investment tool, rather than a mere statistical problem: using non-linear derivatives any hedge fund can design an appropriate probability distribution of those exposures to such markets and risk factors. On the other hand, the need to reach ef-

fective diversification has led to new investment products, bound to exploit the credit risk features of the assets. It is particularly for the evaluation of these new products, such as securitized assets and basket credit derivatives (n^{th} to default options) that the need to account for co-movement among non-normally distributed variables has become an unavoidable task.

Copula functions have been first applied to the solution of these problems, and have been later applied to the multidimensional non-normality problem. In fact, the use of copula functions enables the task of specifying the marginal distributions to be decoupled from the dependence structure of variables. This allows us to exploit univariate techniques at the first step and then to use non-parametric dependence measures at the second step. This avoids the flaws of linear correlation that have, by now, become well known.

The phrase *copula* was first used in 1959 by Sklar, but the traces of copula theory can be found in Hoeffding's work already during the 1940's. This theory languished for three decades in obscurity in theoretical statistics before re-emerging as an important analytical tool in the global scene of financial economics, with particular usefulness in modeling the dependence structure between sets of random variables. Prior to the very recent spread of copula theory and applications in the financial world, the only models available to represent this dependence structure were the classical multivariate models, such as the widely used Gaussian multivariate model. These models entailed rigid assumptions on the marginal and joint behaviors of the variables, and were almost useless for modeling the dependence between real financial data.

Copula theory provides a method of modeling the dependence structure between sets of observations without becoming inextricably tangled in these assumptions. Simply expressed, copulas separate the marginal behavior of variables from the dependence

structure through the use of distribution functions. As the empirical marginal distribution functions can be used instead of their explicit analogues, it is not even necessary to know the exact distribution of the variables being modeled.

Copulas provide big versatility, and can be used as an analytical tool in a broad range of financial situations such as risk estimation, credit modeling, pricing derivatives, and portfolio management, to name but a few. Although much of the vast literature dedicated to copula applications lies in the financial sector, applications of copula theory are not confined to the financial world. Any situation involving more than one random variable can be modeled and analyzed using copula theory. As is usually the case with statistics, the more variables that are present in the model, the more complicated and time-consuming the analysis becomes.

In this work we analyze the sensitivity to correlation factors in security defaults in complex portfolios. We focus only on a portfolio built from two subportfolios; however, the developed numerical code in the MATLAB programming language extends easily to the more complex problem of portfolios including more than two subportfolios.

The first part of this work briefly presents the main assumptions and theoretical background necessary to understand the copula method. Further on, the fundamental theorem underpinning all copula-based analysis, known as Sklar's Theorem, is stated and proved. This theorem shows the copula function as a method of modeling the dependence structure between sets of observations by linking the joint cumulative distribution function (cdf) of an n -dimensional random vector to the marginal distribution functions of the components.

The major contributions of this work are presented in Chapter 3. A numerical algorithm as well as specifically implemented MATLAB code is developed to simulate random defaults in a complex portfolio with a specified Gaussian copula correlation

structure. This approach was adopted from [5] and further slightly modified to account for more than just one subportfolio.

The second part of this work presents the numerical approach and simulation the influence of the correlation of defaults in subprime securitization. The numerical algorithm is presented and explained followed by series of simulations with different correlation factors and for different standard tranches.

The thesis ends with the summary of all innovatory observations and ideas arising from the performed numerical simulations. Interesting trends are described along with their impact important to the modeling market behavior. The correlation factor and the ways how it can affect the portfolio performance is discussed in detail and plotted on the numerous graphs.

Chapter 2

Theoretical Background

2.1 Basic Definitions for Copulas

Copulas are the functions that join or couple multivariate distribution functions to their one-dimensional marginal distributions. Gaining a precise understanding of this process is the main objective of this Chapter. Consider a pair of random variables X and Y , with distribution functions

$$F(x) = P[X \leq x] \quad \text{and} \quad G(y) = P[Y \leq y]$$

respectively, and joint distribution,

$$H(x, y) = P[X \leq x, Y \leq y].$$

For any $(x, y) \in [0, 1]^2$, we associate $F(x)$, $G(y)$, and $H(x, y)$. In particular, each pair (x, y) would lead to a point $(F(x), G(y))$ in the unit square $[0, 1]^2$, and in turn, this ordered pair corresponds to a number $H(x, y)$ in $[0, 1]$. This correspondence, which associates the value $H(x, y)$ of the joint distribution function to the ordered pair $(F(x), G(y))$ is, in essence, called a *copula*, since F and G might not be one-to-one.

2.2 Definition of Copulas

The definition of the copula, followed by the description of its characteristics and properties is given in this Section. We note that with the use of the information stated in Section 2.1 copulas are a class of grounded 2-increasing functions with marginals, with domain I^2 . Here, and always, $I = [0, 1]$.

Definition 2.1. A copula is a multivariate distribution with uniform marginals on I .

Thus a *two-dimensional* copula C is a function from $I \times I$ which

$$C(u, 0) = 0 = C(0, v)$$

$$C(u, 1) = u \text{ and } C(1, v) = v;$$

For all $u, v \in I$, and suitable continuity condition holds.

For each u_1, u_2, v_1, v_2 in I with $u_1 \leq u_2$ and $v_1 \leq v_2$,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.$$

It is also convenient to examine the two-dimensional copula, C on the unit square $I \times I$, as outside the unit square, the values of $C(u, v)$ can be easily determined.

In particular, we have $C(u, v) = 0$ if either $u < 0$ or $v < 0$.

Also, $C(u, v) = u$ if $v > 1$ and similarly, $C(u, v) = v$ if $u > 1$.

Theorem 2.2. *Let C be a copula. Then for every u_1, u_2, v_1, v_2 in $\text{Dom}(C)$,*

$$|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|.$$

Hence C is uniformly continuous on its domain.

The proof of above theorem can be found in [2]

2.3 Sklar's Theorem

We need two preparatory definitions before we turn to the theorem.

Definition 2.3. $D : A_1 \times A_2 \rightarrow \bar{\mathbb{R}}$ is called **2-increasing** if for every rectangle $[v_1, v_2] \times [z_1, z_2]$ whose vertices lie in $A_1 \times A_2$, such that $v_1 \leq v_2, z_1 \leq z_2$

$$D(v_2, z_2) - D(v_2, z_1) - D(v_1, z_2) + D(v_1, z_1) \geq 0$$

Definition 2.4. A two-dimensional subcopula \mathcal{C} is a real function defined on $A \times B$, where A and B are non-empty subsets of $I = [0, 1]$, containing both 0 and 1:

$$\mathcal{C} : A \times B \rightarrow \bar{\mathbb{R}}$$

1. grounded ($\mathcal{C}(v, 0) = \mathcal{C}(0, z) = 0$)
2. such that

$$\mathcal{C}(v, 1) = v, \quad \mathcal{C}(1, z) = z$$

for every (v, z) of $A \times B$

3. 2-increasing

This completes the necessary definitions and we can move to the statement and proof of important Sklar's Theorem.

Theorem 2.5. Sklar's Theorem. *Let F and G be one-dimensional distribution functions.*

1. *If \mathcal{C} is any subcopula whose domain contains $\text{Ran}(F) \times \text{Ran}(G)$, then*

$$(x, y) \mapsto \mathcal{C}(F(x), G(y))$$

is a joint distribution function with margins $F(x), G(y)$;

2. Conversely, if $H(x, y)$ is a joint distribution function with margins $F(x), G(x)$, there exists a unique subcopula \mathcal{C} , with domain $\text{Ran}(F) \times \text{Ran}(G)$, such that

$$H(x, y) = \mathcal{C}(F(x), G(y))$$

If $F(x), G(y)$ are continuous, the subcopula is a copula; if not, there exists a copula C such that

$$C(v, z) = \mathcal{C}(v, z)$$

for every $(v, z) \in \text{Ran}(F) \times \text{Ran}(G)$.

The proof of this theorem can be found in [1].

In effect, the copula can be thought of as a function that ‘couples’ the joint distribution function to its univariate marginals. Furthermore, it may be noted that $H(x, y) = C(F(x), G(y))$ can be inverted to express the copula in terms of a joint distribution function and the inverse of two marginals, if the inverses exist. However, if any of marginal distributions is not strictly increasing, then it does not possess the property of inverse. Sklar’s theorem overcomes that obstacles.

2.4 Copulas and Random Variables

It is important to extend the introduced concept of copula to random variables. First, it is necessary to define the notion of a random variable.

Definition 2.6. A random variable X is a function on the probability space (Ω, \mathfrak{F}) such that

$$[X \leq x] \stackrel{def}{=} \{\omega \in \Omega : X(\omega) \leq x\}$$

is an event, i.e. it is measurable for every $x \in \mathbb{R}$

From now on X, Y will denote random variables, and x, y will denote the values of these random variables. We will denote by F_X the distribution function of a random variable X :

$$F_X(x) = P[X \leq x] \quad \text{for all } x \in \bar{\mathbb{R}}.$$

This function is monotone non-decreasing, right continuous, $F_X(-\infty) = 0$ and $F_X(+\infty) = 1$. Sometimes we use simply F .

Theorem 2.7. *Let X and Y be random variables with distribution functions F and G , respectively, and joint distribution function H . Then there exists a copula C such that*

$$H(x, y) = C(F(x), G(y)).$$

If F and G are continuous, then C is unique.

The copula C from the above theorem of random variables X and Y will be referred from this point as C_{XY} .

Particularly in nonparametric statistics, copulas are useful in obtaining the joint distribution functions of two or more variables. In this work we will concentrate on copulas of two variables; however, the problem statement as well as the implemented simulations can be easily extended. The reason for the increase in the usage of copulas is because copulas used in determining strictly monotone transformations of random variables, are usually either invariant or easy to determine as they change in predictable ways.

The following theorem shows the invariance exhibited by copulas. However, first

we need to recall that if the distribution function of a given random variable X is continuous, and if α is a strictly monotone continuous function, whose domain contains $\text{Ran}(X)$, then the distribution function of the random variable $\alpha(X)$ is also continuous. The strictly increasing transformation is addressed first below.

Theorem 2.8. *Let X and Y be random variables with copula C_{XY} . Suppose α and β are strictly increasing continuous functions on $\text{Ran}(X)$ and $\text{Ran}(Y)$ respectively. Then we obtain,*

$$C_{\alpha(X)\beta(Y)} = C_{XY}$$

Thus, C_{XY} is invariant under strictly increasing transformation of both, X and Y .

Proof. Suppose F_1, G_1, F_2 and G_2 are the distribution functions of $X, Y, \alpha(X)$ and $\beta(Y)$ respectively. Both α and β are strictly increasing on R_X and R_Y . Therefore:

$$F_2(x) = P[\alpha(X) \leq x] = P[X \leq \alpha^{-1}(x)] = F_1(\alpha^{-1}(x)).$$

Similarly,

$$G_2(y) = P[\beta(Y) \leq y] = P[Y \leq \beta^{-1}(y)] = G_1(\beta^{-1}(y)).$$

It follows that for any x, y in $\bar{\mathbb{R}}$

$$\begin{aligned} C_{\alpha(X)\beta(Y)}(F_2(x), G_2(y)) &= P[\alpha(X) \leq x, \beta(Y) \leq y] \\ &= P[X \leq \alpha^{-1}(x), Y \leq \beta^{-1}(y)] \\ &= C_{XY}(F_1(\alpha^{-1}(x)), G_1(\beta^{-1}(y))) \\ &= C_{XY}(F_2(x), G_2(x)) \end{aligned}$$

It follows that $C_{\alpha(X)\beta(Y)} = C_{XY}$. □

It is easy to find the same relationship between $C_{\alpha(X)\beta(Y)}$ and C_{XY} when one of either α and β is strictly decreasing as shown below in the next theorem.

Theorem 2.9. *Let X and Y be continuous random variables with copula C_{XY} . Let α and β be strictly monotone on $\text{Ran}(X)$ and $\text{Ran}(Y)$ respectively. It follows that if α is strictly increasing and β is strictly decreasing, then*

$$C_{\alpha(X)\beta(Y)}(u, v) = u - C_{XY}(u, 1 - v)$$

Proof. Suppose F_1, G_1, F_2 and G_2 are the distribution functions of $X, Y, \alpha(X)$ and $\beta(Y)$ respectively. Suppose α is strictly increasing on $\text{Ran}(X)$ and β is strictly decreasing on $\text{Ran}(Y)$. Therefore:

$$F_2(x) = P[\alpha(X) \leq x] = P[X \leq \alpha^{-1}(x)] = F_1(\alpha^{-1}(x)).$$

However,

$$\begin{aligned} G_2(y) &= P[\beta(Y) \leq y] = P[Y \geq \beta^{-1}(y)] \\ &= 1 - P[Y \leq \beta^{-1}(y)] \\ &= 1 - G_1(\beta^{-1}(y)). \end{aligned}$$

It follows that for any x, y in $\bar{\mathbb{R}}$

$$\begin{aligned}
C_{\alpha(X)\beta(Y)}(F_2(x), G_2(y)) &= P[\alpha(X) \leq x, \beta(Y) \leq y] \\
&= P[X \leq \alpha^{-1}(x), Y \geq \beta^{-1}(y)] \\
&= P[X \leq \alpha^{-1}(x)] - P[X \leq \alpha^{-1}(x), Y < \beta^{-1}(y)] \\
&= F_1(\alpha^{-1}(x)) - C_{XY}(F_1(\alpha^{-1}(x)), G_1\beta^{-1}(y)) \\
&= F_2(x) - C_{XY}(F_2(x), 1 - G_2(y))
\end{aligned}$$

Thus, $C_{\alpha(X)\beta(Y)}(u, v) = u - C_{XY}(u, 1 - v)$, for all $u, v \in [0, 1]$.

□

2.5 Kendall's Tau

Kendall's Tau is a useful measure of association between two random variables. We define Kendall's Tau using the notion of concordance, which is presented below.

Definition 2.10. Let (x_i, y_i) and (x_j, y_j) denote two observations from a vector (X, Y) of random variables. We say that (x_i, y_i) and (x_j, y_j) are concordant if

$$\begin{aligned}
&x_i < x_j \text{ and } y_i < y_j \\
&\text{or } x_i > x_j \text{ and } y_i > y_j
\end{aligned}$$

Similarly, one can define the notion of discordance.

Definition 2.11. Let (x_i, y_i) and (x_j, y_j) denote two observations from a vector (X, Y) of random variables. We say that (x_i, y_i) and (x_j, y_j) are discordant if

$$x_i < x_j \text{ and } y_i > y_j$$

or $x_i > x_j \text{ and } y_i < y_j$

One can extend the notions of concordance and discordance to a correlation measure, Kendall's Tau. The importance of using Kendall's Tau is central to the theory of copulas as it is invariant under monotone transformations of the underlying distributions X and Y . The definition of the sample version of Kendall's Tau is provided below.

Definition 2.12. Let $x_1, y_1, x_2, y_2, \dots, x_n, y_n$ denote values of random variables X, Y . There are $\binom{n}{2}$ distinct pairs x_i, y_i and x_j, y_j of observations in the sample, and each pair is either concordant or discordant. If we let c denote the number of concordant pairs and d denote the number of discordant pairs, then Kendall's Tau for the sample is defined as

$$\tau_s = \frac{c - d}{c + d} = \frac{c - d}{\binom{n}{2}}$$

Chapter 3

Gaussian Model: Simulations and Observations

3.1 Basic Concepts

The statement of the theoretical model utilized in this work requires some brief additional concepts. The full understanding of the model depends on the correct definition of those concepts.

The first obstacle is to obtain a variable with a particular distribution, such as the Gaussian distribution. It can be achieved by introducing the proxy, which is the theoretical representation of the market variable. Despite the fact that in this research we do not work with real market data, the model presented below along with the simulations fully are capable of working on real life data.

We consider the return on a security represented by a proxy variable $X(t)$ given by

$$X(t) = aM(t) + bZ_1(t), \quad (3.1)$$

where a and b are the coefficients of the components $M(t)$ and $Z_1(t)$, respectively. In a similar manner we consider another return on different security represented by proxy variable $Y(t)$ such as

$$Y(t) = cM(t) + dZ_2(t), \quad (3.2)$$

where c and d are again the coefficients of components $M(t)$ and $Z_2(t)$, respectively. It is important to notice that component $M(t)$ is common for both $X(t)$ and $Y(t)$ while

the random variable $Z_1(t)$ affects only X and the random variable $Z_2(t)$ affects only Y . It was shown [3] that the serial correlation in a common risk factor propagates to the default rates. In order to study this observation analytically, a time series model is used for the common risk factor in the Gaussian copula approach. Additionally, $M(t)$, $Z_1(t)$ and $Z_2(t)$ are all independent and identically distributed $N(0, 1)$. Hence, $M(t)$ can be seen as a global variable factor whereas Z_1 Z_2 are idiosyncratic factors. Therefore, we will work at a fixed time and will drop “t”.

The following properties can be seen as consequence of the above definition

$$E[X] = E[Y] = 0, \text{ and}$$

$$\text{Var}(X) = \text{Var}(Y) = 1.$$

The variance of a given distribution X is given by

$$\text{Var}(X) = E[X^2] - (E[X])^2$$

Thus, in one case

$$E[X^2] = 1$$

which is equivalent to,

$$a^2 E[M^2] + b^2 E[Z_1^2] + 2abE[MZ_1] = 1$$

Notice, that $E[MZ_1] = 0$, as $E[M]E[Z_1] = 0$. Moreover, $E[M^2] = E[Z_1^2] = 1$. Hence, (3.1) yields

$$a^2 + b^2 = 1, \tag{3.3}$$

and analogously (3.2) yields

$$c^2 + d^2 = 1. \tag{3.4}$$

We will assume $a, b, c, d > 0$. Thus, from (3.1) with use of (3.3) one obtains

$$X = aM + \sqrt{1 - a^2}Z_1, \tag{3.5}$$

and again analogously from (3.2) and (3.4), we have

$$Y = cM + \sqrt{1 - c^2}Z_2, \tag{3.6}$$

Additionally, the correlation between X and Y denoted by ρ_{XY} can be defined by

$$\rho_{XY} = \frac{E[XY] - E[X]E[Y]}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}$$

Noting that $E[X] = E[Y] = 0$ and $\text{Var}(X) = \text{Var}(Y) = 1$, yields

$$\rho_{XY} = E[XY].$$

From (3.5) and (3.6), it follows that

$$\begin{aligned}\rho_{XY} &= E[(aM + \sqrt{1-a^2}Z_1)(cM + \sqrt{1-c^2}Z_2)]. \\ &= E[acM^2 + c\sqrt{1-a^2}MZ_1 + a\sqrt{1-c^2}MZ_2 + \sqrt{1-a^2}\sqrt{1-c^2}Z_1Z_2] \quad (3.7)\end{aligned}$$

Observe, that in (3.7) only first term of the expectation, namely acM^2 is non zero.

Hence, from (3.7)

$$\rho_{XY} = acE[M^2],$$

and since

$$E[M^2] = 1,$$

we obtain

$$\rho_{XY} = E[XY] = ac.$$

Additionally,

$$\begin{aligned}\rho_{XM} &= E[XM] = E[(aM + \sqrt{1-a^2}Z_1)M] \\ &= E[aM + \sqrt{1-a^2}Z_1M] \\ &= E[aM^2]\end{aligned}$$

Thus, $E[XM] = a$.

Similarly, $E[YM] = c$.

The above equations express the correlation of the underlying returns on X and Y to the risk factor, M . Furthermore, by definition, the pair (X, Y) is given by

$$(aM(t) + \sqrt{1 - a^2}Z_1(t), cM(t) + \sqrt{1 - c^2}Z_2(t))$$

with both $aM(t) + \sqrt{1 - a^2}Z_1(t)$ and $cM(t) + \sqrt{1 - c^2}Z_2(t)$ defined as normally distributed functions with mean 0 and variance 1 *i.e.* $N(0, 1)$.

We will now discuss the use of the Gaussian copula as our choice of copulas to capture the joint distribution function of the defaults on two subportfolios of one major portfolio. Furthermore, for more subportfolios the multivariate Gaussian copula can be used.

Consider the distribution of the returns of a security X , with marginal distribution denoted by $F_X(x)$, where

$$F_X(x) = P[X \leq x]$$

and $x \in X$. Let $U = F_X(x)$ and let $\tilde{X} = \Phi^{-1}(U)$, where Φ is the standard Gaussian distribution function. Then we have the distribution function of \tilde{X} described as follows,

$$\begin{aligned} F_{\tilde{X}}(x) &= P[\tilde{X} \leq x] = P[\Phi^{-1}(U) \leq x] \\ &= P[\Phi^{-1}(F_X(x)) \leq x] \\ &= P[F_X(x) \leq \Phi(x)] \\ &= P[x \leq F_X^{-1}(\Phi(x))] \\ &= F_X[F_X^{-1}(\Phi(x))] \\ &= \Phi(x). \end{aligned}$$

Hence, \tilde{X} is standard Gaussian. We have assumed that F_X is continuous and strictly increasing.

The transformations used above can be applied to obtain the results in Sklar's Theorem in terms of a Gaussian copula as shown below. Let the marginal distribution functions of the return on securities X and Y , be F_X and G_Y respectively. Let $U = F_X(X)$ and $V = G_Y(Y)$, be the uniformized marginal distribution of X and Y and let H be the observed joint distribution between X and Y .

$$C(F_X(\Phi^{-1}(u)), G_Y(\Phi^{-1}(u))) = C(u, v) = H(x, y)$$

In essence, this gives us the choice of determining the joint distribution function of X and Y through either the 'uniformized' marginal distribution or the 'gaussianized' marginal distribution.

Theoretically, Sklar's theorem in this aspect provides a desirable result by linking the joint distribution of two variables with the copula of the marginal distributions of these two variables. In practice however, it is more convenient to express Sklar's Theorem in terms of the density functions of the copula and joint distribution respectively. This is especially useful in scenarios where the range of either the joint distribution that is to be determined. This notion of using the copula density function to express the joint density function of two random variables is particularly useful in our analysis, as the density function of the Gaussian copula can be easily determined.

Consider the bivariate standard Gaussian distribution as shown below,

$$H(x, y) = \int_{-\infty}^{\Phi^{-1}(x)} \int_{-\infty}^{\Phi^{-1}(y)} \exp \left[- \left(\frac{s^2 - 2\rho st + t^2}{2(1 - \rho^2)} \right) \right] \frac{dsdt}{2\pi\sqrt{1 - \rho^2}}$$

Taking the partial derivative of $H(x, y)$ with respect to x, y we obtain

$$\begin{aligned} &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left[- \left(\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)} \right) + \frac{x^2}{2} + \frac{y^2}{2} \right] \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left[- \left(\frac{x^2 - 2\rho xy + y^2 - x^2(1-\rho^2) - y^2(1-\rho^2)}{2(1-\rho^2)} \right) \right] \end{aligned}$$

$$\text{Thus, } h(x, y) = \frac{1}{\sqrt{1-\rho^2}} \exp \left[\frac{-1}{2(1-\rho^2)} (\rho^2 x^2 - 2\rho xy + \rho^2 y^2) \right] \quad (3.8)$$

Recall that from Sklar's Theorem, we have

$$\begin{aligned} h(x, y) &= c(\Phi^{-1}(u), \Phi^{-1}(v)), \text{ yielding} \\ c(\Phi^{-1}(u), \Phi^{-1}(v)) &= \frac{1}{\sqrt{1-\rho^2}} \exp \left[\frac{-1}{2(1-\rho^2)} (\rho^2 x^2 - 2\rho xy + \rho^2 y^2) \right] \end{aligned} \quad (3.9)$$

Here $h(x, y)$ is the density function of the bivariate Gaussian distribution of random variables X and Y . Notice, that, through Sklar's Theorem, the above expression also describes the density function for the Gaussian copula, of the marginal distribution functions are 'gaussianized'.

3.2 Model Definition

The credit industry standard copula model was introduced by Li [5] and is called default time Gaussian copula. This model is applicable to almost all types of CDO, MBS, and almost all other credit derivatives that are derived from the multiple assets with credit risk. The main idea of this model is that each credit asset has a default time, or survival time, after which the mortgage defaults. Instead of modeling the correlation between mortgages, Li proposed a copula approach to capture the joint distribution of

defaults times. A copula in this case takes the marginal distribution of default times and returns their joint distribution.

3.2.1 Model for a Single Portfolio

Consider a portfolio of credits. Let us examine the overall state of the portfolio and its relation to component names. As a simple model, presented already by [10] assume that there is a ‘global factor’ Z that describes the overall state of the economy relevant to this portfolio. Each name i in the portfolio has default behavior described through a combination of Z and an idiosyncratic factor ε which describes the part of its behavior specific to i , independent of the behavior of the overall portfolio. In more detail, the model assumes that the default behavior of i is governed by

$$X_i = \beta Z + \sqrt{1 - \beta^2} \varepsilon_i \tag{3.10}$$

where Z and ε are independent random factors, normalized to have mean 0 and variance 1, and β is a weight whose significance we will see shortly. Default occurs if the value of X_i falls below a threshold level. Our assumption of the specific form of X_i means that the default behavior of each name is correlated to the global factor Z in exactly the same way, measured by the quantity

$$\beta = \text{Corr}(X_i, Z) \tag{3.11}$$

There is no loss of generality in assuming that $\beta \geq 0$. Hence, the portfolio correlation is the correlation between any pair of names:

$$\rho = \text{Corr}(X_i, X_j) = \beta^2 \quad (3.12)$$

The factor X_i controls the default behavior of name i in the portfolio: if X_i is below a threshold level X^* then the name i defaults:

$$X_i \leq X^* \quad \text{corresponds to default of } i \quad (3.13)$$

We see that X_i is a kind of proxy for the asset value of the firm i . More information regarding setting X^* threshold for dynamic copula approach can be found in [12].

3.2.2 Model for a Single Portfolio with Two Subportfolios

In this approach we adopt the Li's default copula, however we do not consider the time, neither for state variables X_i nor for common risk factor Z_i . On the other hand, one considers two different subportfolios, containing i securities each, and govern by independent correlation factors ρ_1 and ρ_2 . The random variables X_{1i} and X_{2i} represent the value of the particular security i , and are compared to the threshold value X^* . The initial risk factors Z_1 and Z_2 are correlated with the common factor ρ . We assume that the distribution of X_i is the same across all securities for particular subportfolio, where j stays for the number of subportfolios (in our case $j = 1, 2$)

$$F_j(s) = \mathbb{P}[X_{j,i} < s], \forall i = 1, 2, \dots, N \text{ and } \forall j = 1, 2. \quad (3.14)$$

Moreover, we assume that F_j is continuous and strictly increasing. Given this information, for each subportfolio separately the Gaussian copula approach provides a way to obtain the joint distribution of the $X_{j,i}$ across i .

Briefly recall, that copula is a joint distribution function

$$C(u_1, u_2, \dots, u_N) = \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2, \dots, U_N \leq u_N), \quad (3.15)$$

where u_1, u_2, \dots, u_N are N uniformly distributed random variables that may be correlated. It can be easily verified that the function

$$C[F_1(x_1), F_2(x_2), \dots, F_N(x_N)] = G(x_1, x_2, \dots, x_N) \quad (3.16)$$

is a multivariate distribution function with marginal distribution functions $F_1(x_1), F_2(x_2), \dots, F_N(x_N)$. It was already shown in previous Chapter, that Sklar showed that for an arbitrary multivariate distribution function $G(x_1, x_2, \dots, x_N)$ with continuous marginal distribution functions $F_1(x_1), F_2(x_2), \dots, F_N(x_N)$, there exist a unique C such that Equation 3.16 holds. Since we consider the case of only two separate subportfolios, there is a C_j for each subportfolio j such that

$$C_j[F_j(x_{j,1}), F_j(x_{j,2}), \dots, F_j(x_{j,N})] = G_j(x_{v,1}, x_{v,2}, \dots, x_{v,N}) \quad \forall j = 1, 2. \quad (3.17)$$

We consider a portfolio of credits as in previous section with the distinction that, the existing portfolio is subdivided into two separate subportfolios. We assume that there is a global risk factor Z that describes the overall state of the economy relevant to this portfolio. Furthermore, we divide this portfolio into two subportfolios j . Each name i in the subportfolio has default behavior described through a combination of Z and an *idiosyncratic* factor ε_i which describes the part of its behavior specific to i ,

independent of the behavior of the overall portfolio. In more detail, the default behavior of i is governed by

$$X_{j,i} = \beta_j Z_j + \sqrt{1 - \beta_j^2} \varepsilon_{j,i} \quad (3.18)$$

where Z_j and ε_i are independent random factor, normalized to have mean 0 and variance 1, and β is a weight whose significance we will see shortly. Default occurs if the value of X_i for particular subportfolio falls below a threshold level X^* .

We also assume that

$$\begin{aligned} Z_1 &= \sqrt{\rho} Z + \sqrt{1 - \rho} \varepsilon_1 \\ Z_2 &= \sqrt{\rho} Z + \sqrt{1 - \rho} \varepsilon_2. \end{aligned}$$

where $Z, \varepsilon_1, \varepsilon_2$ are independent, identically distributed standard Gaussian and $\rho \in (0, 1)$ is a correlation term. The risk factor $X_{j,i}$ controls the default behavior of name i in the subportfolio j : if $X_{j,i}$ is below a threshold level X^* then name i defaults

$$X_{j,i} \leq X^* \quad (3.19)$$

In this sence, X_i is a kind of proxy for the asset value of the firm i .

An ideal subportfolio in which all the names are correlated in the same way with the global factor Z_j , as above, is said to be homogeneous in correlation.

The value of ρ in the Gaussian copula which gives the correct expected loss for a tranche $0 - x\%$ (X^*) is the *base correlation* for that tranche. Generally, tranches of the type $0 - x\%$ do not exist in the market, except for the equity tranche. Thus, market data for the standard traded tranches must be bootstrapped to yield the correlation for ‘theoretical’ tranches such as $0 - 7\%$, and then for all intermediate tranches. For

example, to obtain the correlation for the 0 – 7%, the expected loss (implied from the tranche spread) for the 0 – 3% tranche should be combined with the expected loss for the 3 – 7% tranche, and the correlation for the full 0 – 7% tranche backed out from this combined expected loss.

3.3 Numerical Algorithm

This section describes in detail the implementation of an approach together with the necessary equations, theoretically discussed in the previous chapters. We consider two subportfolios. In the following steps the numerical algorithm is presented and described. The MATLAB environment and programming language is used to conduct the implementation and simulations.

1. Simulate standard Gaussian independent variable using the variable generating option: Z ; ε_{init1} ; ε_{init2} ; $\varepsilon_{1,1}, \varepsilon_{1,2}, \dots, \varepsilon_{1,N}$; $\varepsilon_{2,1}, \varepsilon_{2,2}, \dots, \varepsilon_{2,N}$ where N is the number of securities in each subportfolio, Z is the initial random number and $\varepsilon_{j,N}$ is the idiosyncratic factor necessary for further calculation of risk factor Z_i , and state variable $X_{j,i}$, respectively.
2. Compute the risk factor Z_j based on common risk factor simulated in part 1, idiosyncratic factor ε_j and common correlation factor ρ for all securities in each subportfolio ($\forall j = 1, 2$)

$$Z_1 = \sqrt{\rho}Z + \sqrt{1 - \rho}\varepsilon_1$$

$$Z_2 = \sqrt{\rho}Z + \sqrt{1 - \rho}\varepsilon_2.$$

Note: a good check at this point of the simulations is the fact that the input correlation factor $\rho = \text{Corr}(Z_1, Z_2)$ for a big sample of Monte Carlo simulations,

i.e. repeating all steps of the numerical algorithm consecutively many times (at least 1000).

3. Calculate the state variables $X_{j,i}$ for each subportfolio using the information from previous steps

$$X_{1,i} = \sqrt{\rho_1}Z_1 + \sqrt{1 - \rho_1}\varepsilon_1$$

$$X_{2,i} = \sqrt{\rho_2}Z_2 + \sqrt{1 - \rho_2}\varepsilon_2.$$

4. Define the threshold value X^* and calculate the total losses of each subportfolio:

$$L_1 = \sum 1_{[X_{1,i} \leq X^*]}$$

$$L_2 = \sum 1_{[X_{2,i} \leq X^*]}.$$

5. Corresponding to the standard tranches calculate the minimum of the defaults ratio per subportfolio and x for all $x \in [0, 1]$

$$L_1(x) = \min \left\{ \frac{L_1}{N}, x \right\}$$

$$L_2(x) = \min \left\{ \frac{L_2}{N}, x \right\}.$$

6. Using obtain data from step 5 compute the correlation between different default ratios for both, different correlation factors and different tranches.

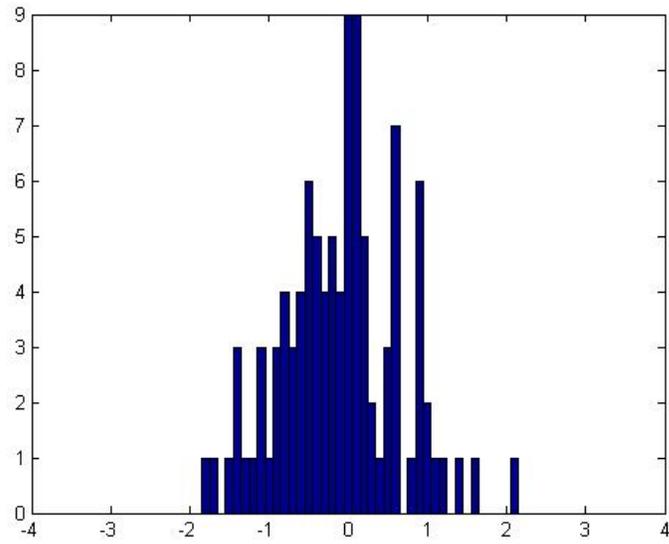
3.4 Validation of the Implemented Code

The numerical simulation does not only experience the common numerical errors due to approximation techniques but it is also possible that seemingly correct code

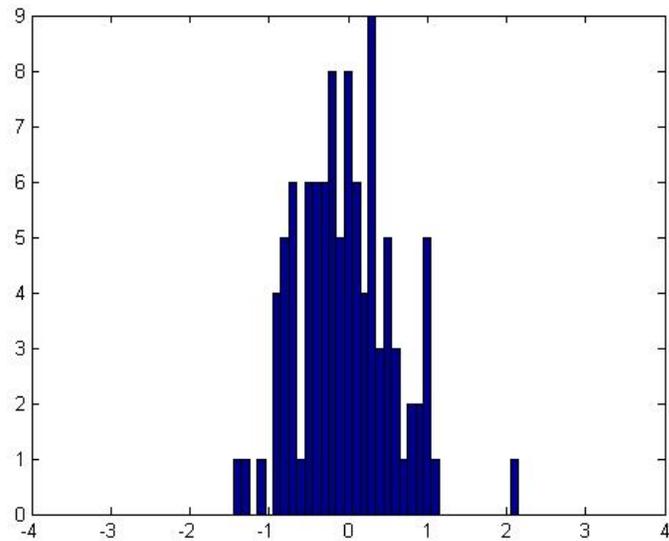
without any syntax mistakes might produce results not consistent with the theoretical approach. Therefore, it is crucial to validate the developed theory and make sure that during each step of the simulation the program is conducting the calculations exactly according to the desired theoretical approach. In this Section we present several *checkpoints* which provides strong confidence in the codes accuracy.

The first important check is the fact that for the significant amount of Monte Carlo simulations the correlation between risk factors for each subportfolio must be close to the given initial correlation factor ρ . That means that sufficiently large Monte Carlo simulations $\rho = \text{Corr}(Z_{i,1}, Z_{i,2})$, where i denotes the number of Monte Carlo simulations. It was shown in numerous conducted simulations that the occurring error is less than 5% i.e. for input value $\rho = 0.5$ and $N = 1000$ Monte Carlo simulations the correlation value $\text{Corr}(Z_{i,1}, Z_{i,2}) \approx 0.46 - 0.54$. Increasing the number of simulation to $N = 10000$, the range of obtained correlations narrows significantly to $\text{Corr}(Z_{i,1}, Z_{i,2}) \approx 0.49 - 0.51$. This provides the first important check showing the correctness of the developed code.

The second crucial validation is to check if the obtained state variable $X_{j,i} = \sqrt{\rho_j}Z_j + \sqrt{1 - \rho_j}\varepsilon_j$ is normal Gaussian distribution with the mean 0 and variance 1, i.e. $N(0, 1)$. It is easy to see that, since Z and ε are standard Gaussian, the outcome should be the same. Again a great number of simulations showed the consistency with the theoretical approach. Conducting over 1000 simulations proved that each time the variable $X_{j,i}$ is standard Gaussian. As an example two histograms for randomly selected simulations are presented in Figure 3.1.



(a)



(b)

FIGURE 3.1. Randomly selected histograms to present the standard normality of X .

Another important fact is considered in aspect of only one simulation. The previously given definition of the default ratio is given as a minimum value from the loss

rate and tranche x , $\forall x \in [0, 1]$ such as

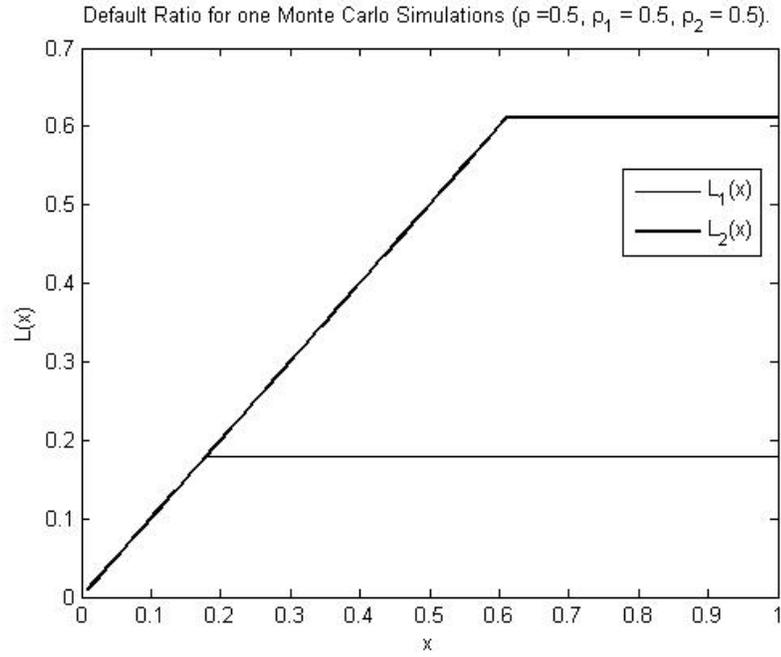
$$L_1(x) = \min \left\{ \frac{L_1}{N}, x \right\}$$

$$L_2(x) = \min \left\{ \frac{L_2}{N}, x \right\}..$$

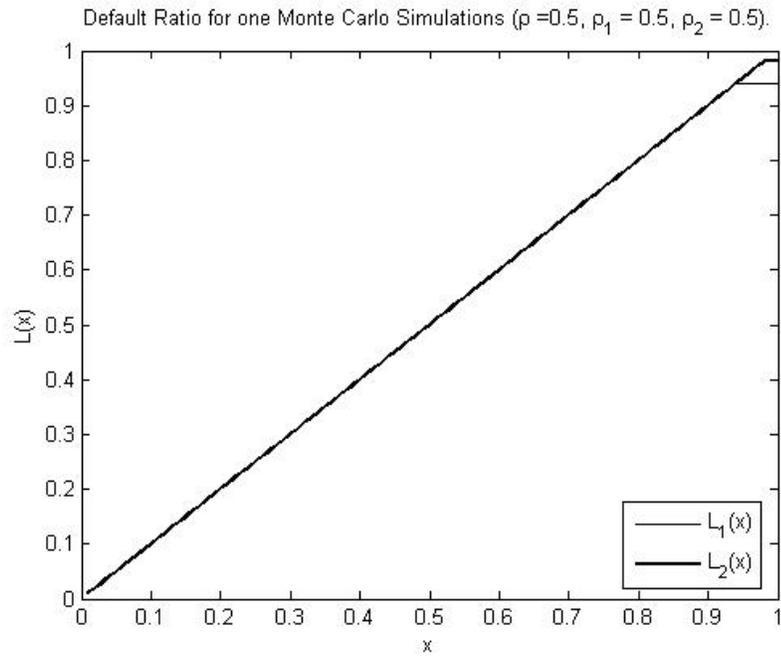
It can be clearly seen that the resulting graph should be linearly increasing for the initial values of default ratio with the increase of x . This will continue until the value of x exceeds the $\frac{L_j}{N}$ loss rate. At that point the graph is going to break and become constant taking the values of loss rate until the end of the range of x . This desired behavior for four randomly selected single simulations can be clearly seen in the following Figures 3.2 and 3.3. The bi-linear function clearly representing the default ratio behavior again confirmed the theoretical approach and added additional drop of certainty for the developed code. The randomly selected two examples for the single simulations with correlations $\rho = 0.5$, $\rho_1 = 0.5$ and $\rho_2 = 0.5$ are presented in Figure 3.2 and 3.3.

Furthermore, it is simple to show, due to the normality of state variable X , the correlation factor $\rho = 0.5$, and sufficiently large M (number of Monte Carlo simulations) the probability of default converges to 0.5, as $x \rightarrow 1$, $E[L_i(x)] = 0.5$, i.e. the expected value of L_1 and L_2 is equal to half of the securities in each subportfolio. This behavior is also captured by the numerical simulations presented in Figure 3.4.

Confirming another step of the algorithm's agreement with existing theory we can move to slightly more advanced relations. Expanding the above check it is interesting to investigate the change of expected value of $L(x)$ with respect to the change of the correlation for each subportfolio, i.e. ρ_1 and ρ_2 . For this case let us state the following Theorem, which was formulated by [6]:

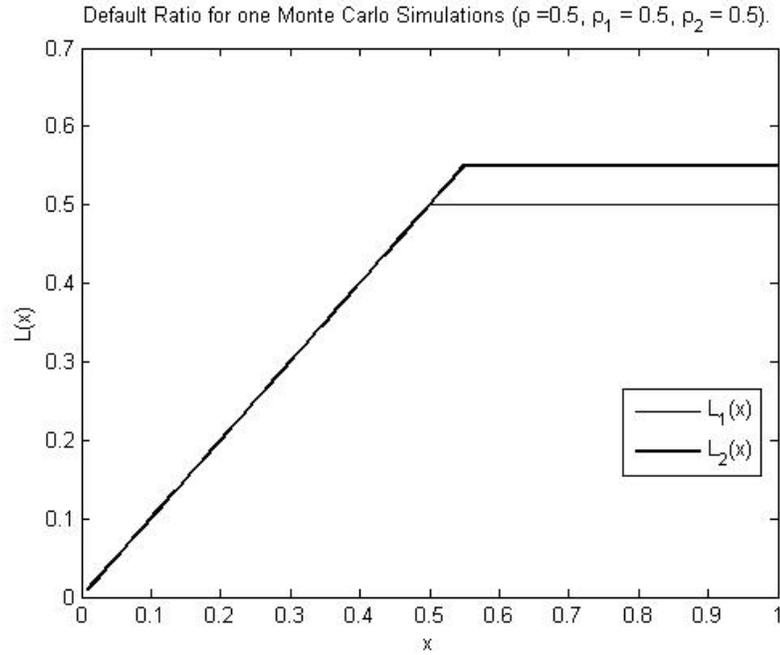


(a)

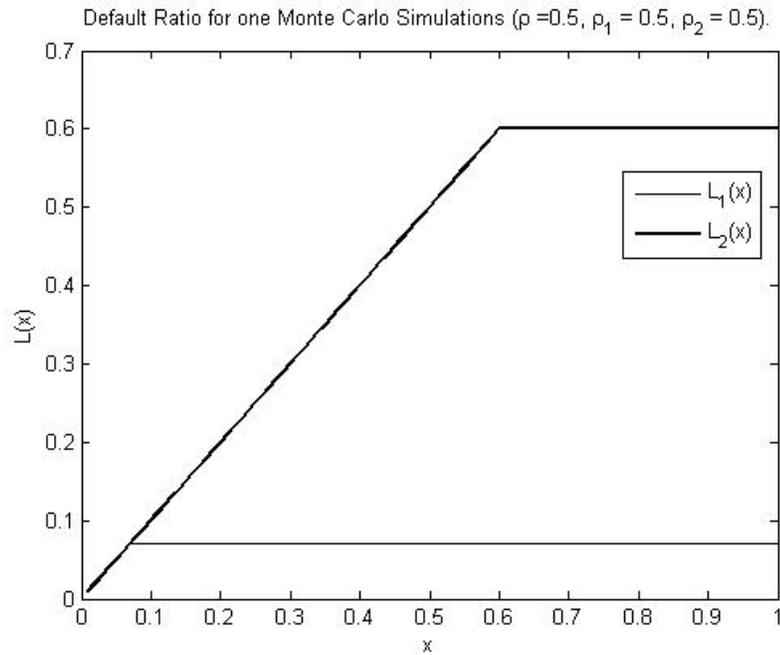


(b)

FIGURE 3.2. Two randomly selected bi-linear functions representing the desired behavior for a single simulation of Default Ratio, with correlations $\rho = 0.5$, $\rho_1 = 0.5$ and $\rho_2 = 0.5$.

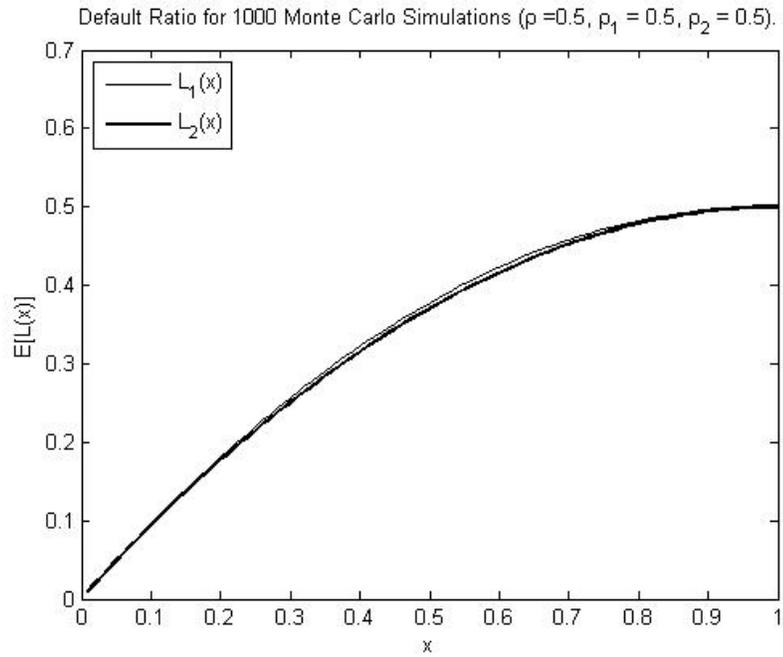


(a)

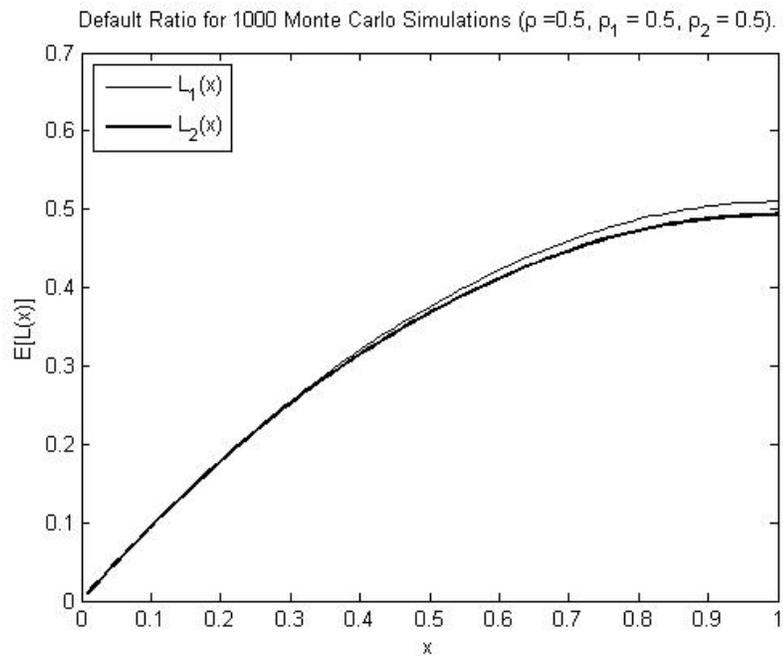


(b)

FIGURE 3.3. Additional two randomly selected bi-linear functions representing the desired behavior for a single simulation of Default Ratio, with correlations $\rho = 0.5$, $\rho_1 = 0.5$ and $\rho_2 = 0.5$, to show the consistency of the numerical simulations.



(a)



(b)

FIGURE 3.4. The default ratio for tranche $x \in [0, 1]$ showing that the expected value of $L(x)$ for sufficiently large M converges to 0.5, as $x \rightarrow 1$.

Theorem 3.1. *Assume that $Z, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_N$ are independent standard Gaussian variables, with $N > 1$, ale let*

$$X_i = \sqrt{\rho}Z + \sqrt{1 - \rho}\varepsilon_i \quad \text{for } i \in 1, \dots, N \quad (3.20)$$

where $\rho \in (0, 1)$. Let $X^* \in \mathbb{R}$. Let v be the random variable which counts the number of X_j with value $< X^*$

$$v = \sum_{i=1}^N 1_{[X \leq X^*]} \quad (3.21)$$

and, for $k \in 1, \dots, N$,

$$v_k = \min\{v, k\} \quad (3.22)$$

$$v_k^s = v - \min\{v, k\} \quad (3.23)$$

Then the expected value of v has no dependence on ρ :

$$\frac{dE[v]}{d\rho} = 0.$$

Moreover,

$$\frac{dE[v_k]}{d\rho} < 0, \quad \text{and} \quad \frac{dE[v_k^s]}{d\rho} > 0. \quad (3.24)$$

for $1 \leq k < N$.

The above theorem shows, that the increase of the correlation factor ρ causes the decrease of expected value of $E[L(x)]$. High portfolio correlation increases the probability of few defaults and so makes the equity tranche safer. The result of above stated

Theorem (for proof see [7] p. 35-38) is now compared with the conducted simulations. In the following Figures the set of simulations is conducted for constant 1000 Monte Carlo Simulations and correlation factor $\rho = 0.5$. Two other parameters, that is the correlations for both subportfolios ρ_1 and ρ_2 governing the change of the expected value $E[L(x)]$ alternate according to the following pattern: in Figure 3.5 $\rho_1 = 0.1$ and $\rho_2 = 0.9$, in Figure 3.6 $\rho_1 = 0.3$ and $\rho_2 = 0.7$, in Figure 3.7 $\rho_1 = 0.5$ and $\rho_2 = 0.5$ (the same scenario as in Figure 3.4 where probability of default for $x \rightarrow 1$ is 0.5), in Figure 3.8 $\rho_1 = 0.7$ and $\rho_2 = 0.3$, and finally in Figure 3.9 $\rho_1 = 0.9$ and $\rho_2 = 0.1$.

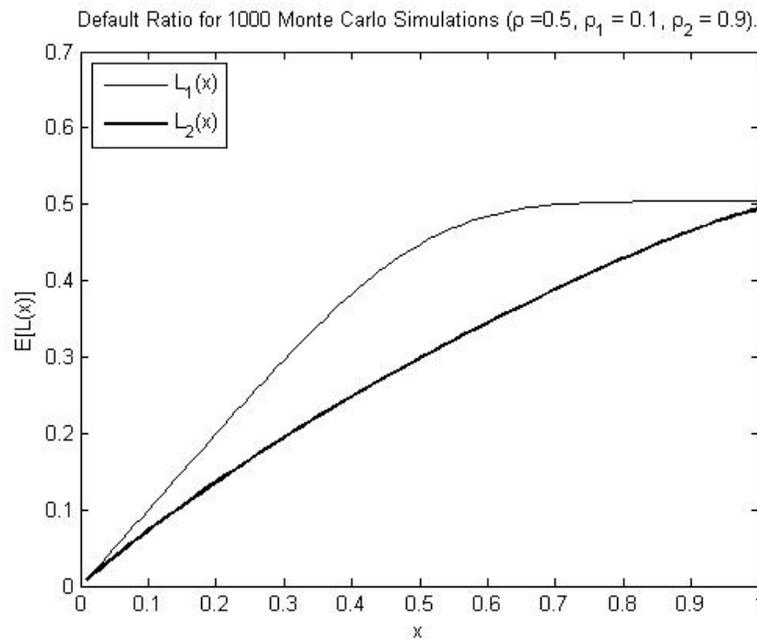


FIGURE 3.5. The default ratio for different correlation factors $\rho_1 = 0.1$ and $\rho_2 = 0.9$ presenting the accuracy of numerical implementation in aspect of Theorem 3.1.

Following Figures 3.5, 3.6, 3.7, 3.8, 3.9 it is clearly visible that the increase of the correlation factor cause the decrease of the expected value of $L(x)$. Presented five figures are randomly selected examples confirming the accuracy of the implemented numerical code with respect to Theorem 3.1. This result gives us additional strong proof of accuracy of the simulations.

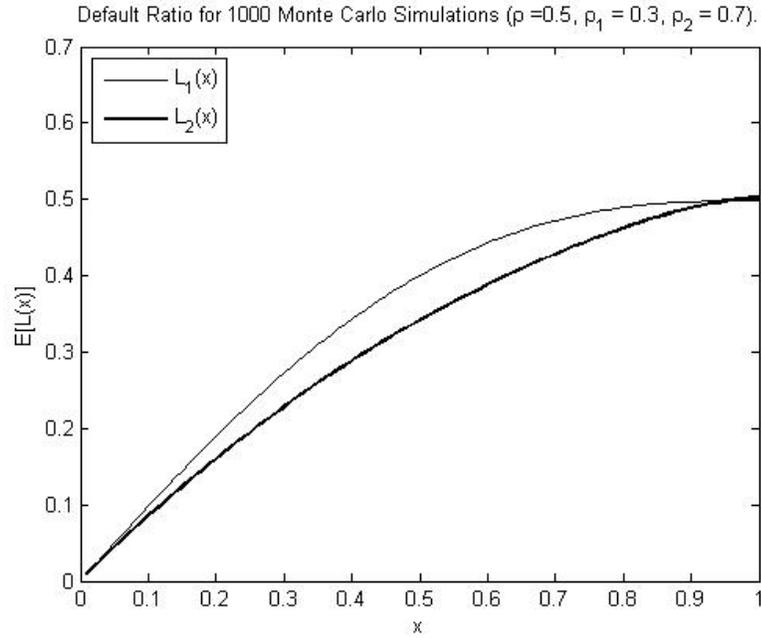


FIGURE 3.6. The default ratio for different correlation factors $\rho_1 = 0.3$ and $\rho_2 = 0.7$ presenting the accuracy of numerical implementation in aspect of Theorem 3.1.

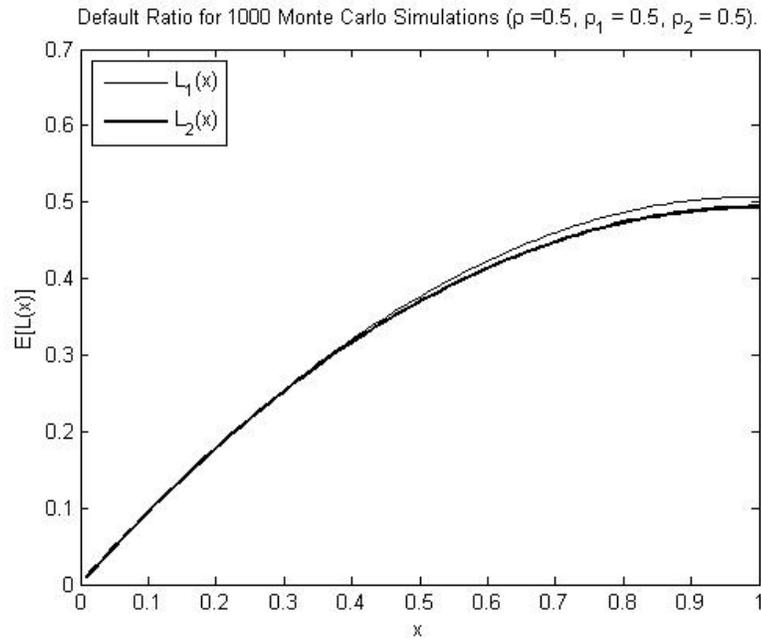


FIGURE 3.7. The default ratio for different correlation factors $\rho_1 = 0.5$ and $\rho_2 = 0.5$ presenting the accuracy of numerical implementation in aspect of Theorem 3.1.

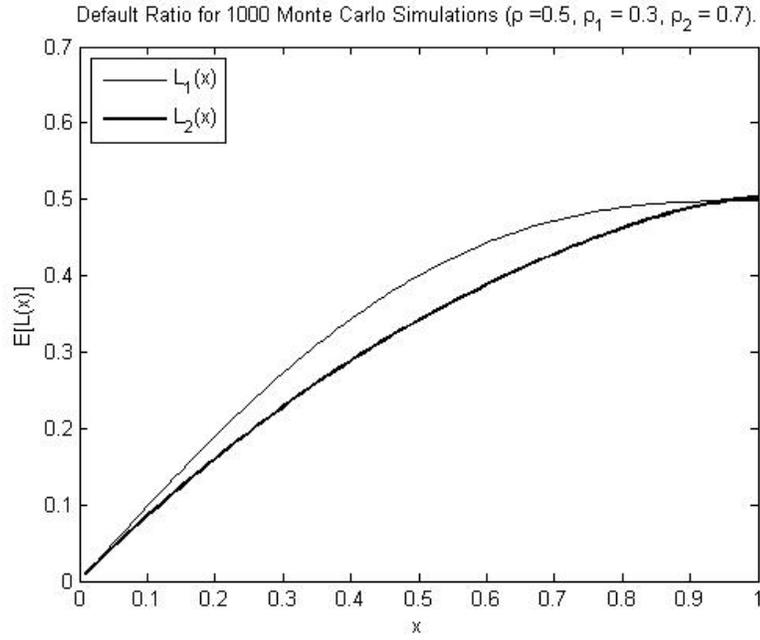


FIGURE 3.8. The default ratio for different correlation factors $\rho_1 = 0.7$ and $\rho_2 = 0.3$ presenting the accuracy of numerical implementation in aspect of Theorem 3.1.

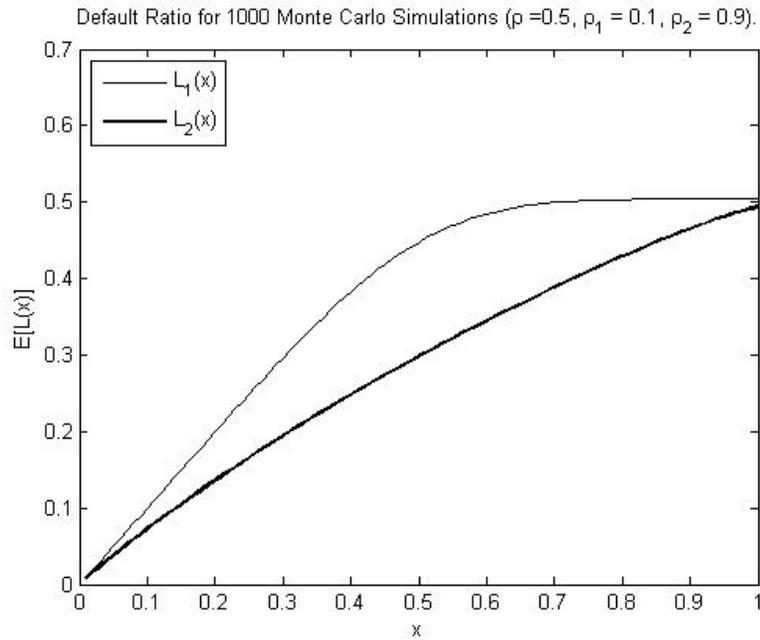


FIGURE 3.9. The default ratio for different correlation factors $\rho_1 = 0.9$ and $\rho_2 = 0.1$ presenting the accuracy of numerical implementation in aspect of Theorem 3.1.

The last aspect considered in this Section refers to the different default thresholds X^* . Obviously, the higher the default threshold the higher the probability of default, since more securities will fail to meet increased standards i.e. X^* . This logical reasoning is represented by the Figure 3.10 and hence, accurately captured by the simulations. The graphs presented in Figure 3.10 are generated according to the following scenario: each simulation conducted for $M = 1000$, $\rho = \rho_1 = \rho_2 = 0.5$ and alternating threshold value X^* . The lowest graph in Figure 3.10 represents the simulation for $X^* = 0$, which is a well know by now. In each graph with the increased probability default and hence, with the increased expectation of default $E[L(x)]$ the threshold value increases 0.25 giving the threshold values of $X^* = 0.25$, $X^* = 0.5$, $X^* = 0.75$ and $X^* = 1$. It can be clearly seen that also in this case the comparison of theory with the numerical approach gives very satisfactory results.

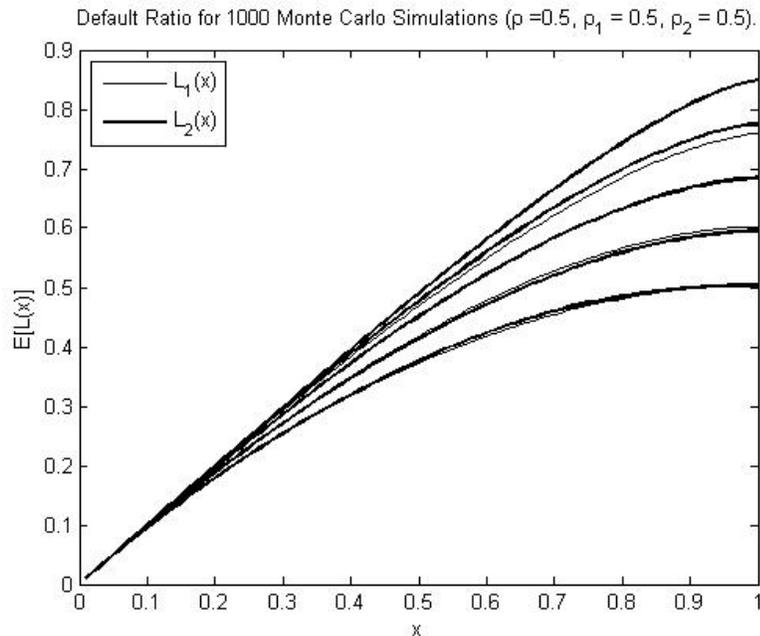


FIGURE 3.10. The default ratios for different thresholds X^* . The lowest graph represents the threshold of $X^* = 0$, and is increased by 0.25 consecutively generating additional graphs with the higher probability defaults. The graph with the highest probability default is for $X^* = 1$.

After presenting all of these checks, it can be stated with high dose of certainty, that the developed numerical code coincides with the presented theoretical approach. In the following Section, the application of developed and checked code resulting in interesting observations are going to be presented and discussed.

3.5 Obtained Results

In this section we concentrate mainly on the results concerning the changes of the correlation factors. It was already stated that in this computational model three different correlation factors were introduced: ρ as a global correlation factor utilized to compute the initial risk factor, ρ_1 capturing the correlation in the first subportfolio and ρ_2 capturing the correlation in the second subportfolio. Normally, the correlation factors are stated as a matrix which connects different securities with each other. However, for the sake of simplicity, this model has only three correlation factors to show the particular relationships in a clear and simplistic manner.

Let us begin our discussion with considering only one single simulation. The simple case of just one common correlation factor $\rho = \rho_1 = \rho_2 = 1$ is presented first. For the model presented in this work it is obvious that number of losses for the first subportfolio are exactly equal to number of losses for the second one, i.e. $L_1 = L_2$. This means that both subportfolios are fully correlated and their graphical representation is the same.

$$Z_1 = \sqrt{\rho}Z + \sqrt{1 - \rho}\varepsilon_1 = Z$$

$$Z_2 = \sqrt{\rho}Z + \sqrt{1 - \rho}\varepsilon_2 = Z$$

and hence

$$X_{1,i} = \sqrt{\rho_1}Z_1 + \sqrt{1 - \rho_1}\varepsilon_1 = Z$$

$$X_{2,i} = \sqrt{\rho_2}Z_2 + \sqrt{1 - \rho_2}\varepsilon_2 = Z$$

Since $X_{1,i} = X_{2,i}$, it results that $L_1 = L_2$ and hence $L_1(x) = L_2(x)$ which proves, that both graphs must be the same, i.e. $E[L_1(x)] = E[L_2(x)]$.

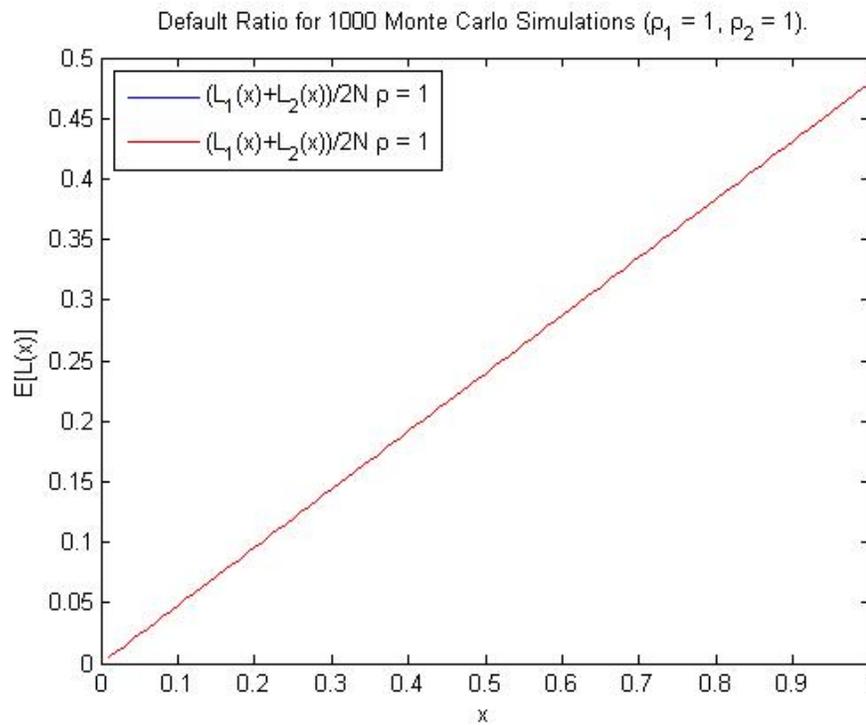


FIGURE 3.11. The same graph representing the expected values of losses $E[L(x)]$ for both subportfolios due to full correlation, i.e. $\rho = \rho_1 = \rho_2 = 1$.

One can now investigate the influence of the initial correlation factor ρ on the default rate considering that both, ρ_1 and ρ_2 are not equal to 1. If $\rho = 1$ than L_1 and L_2 are strongly correlated but not necessarily equal to each other. The smaller the ρ , the smaller the correlation between L_1 and L_2 until the point $\rho = 0$ where entire correlation vanishes and both these variables become independent. This theoretically

observed property can be clearly seen by the graphs generated by simulations and presented in the following Figures. Figure 3.12 shows the relatively small difference between L_1 and L_2 due to high correlation factor ρ whereas Figure 3.13 presents a very particular case of this correlation, when both L_1 and L_2 are equal. On the two following Figures 3.14 and 3.15 the correlation factor is reduced to 0 and hence, number of losses in each subportfolio are entirely independent, adopting the values in random, not correlated, manner. The distances between the plotted functions $L_1(x)$ and $L_2(x)$ is random, hence the probability that both plots are going to be close to each other is by far smaller than the same probability in case of correlated default ratios.

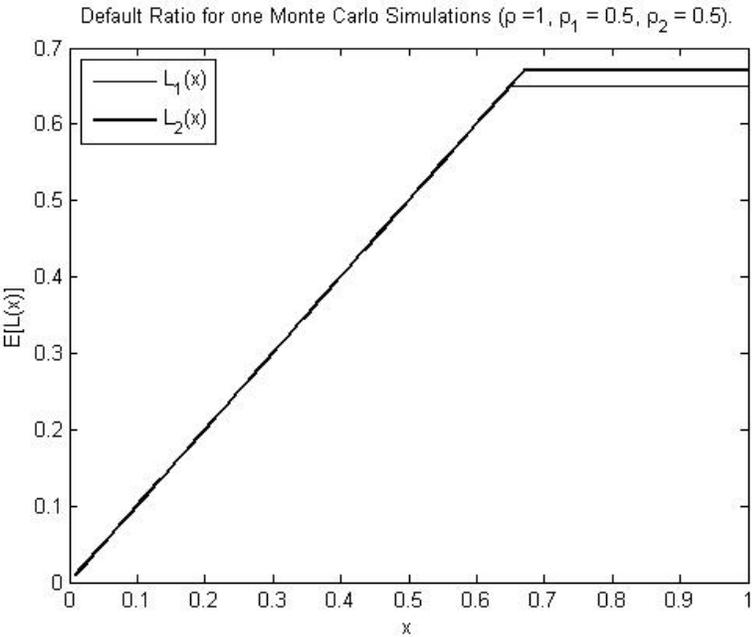


FIGURE 3.12. Strongly correlated default ratios $\rho = 1$ resulting in the small distance between $L_1(x)$ and $L_2(x)$.

After separate consideration of both subportfolios it is crucial to investigate the impact of the correlation factors on joint default ratio. Let us define joint default ratio

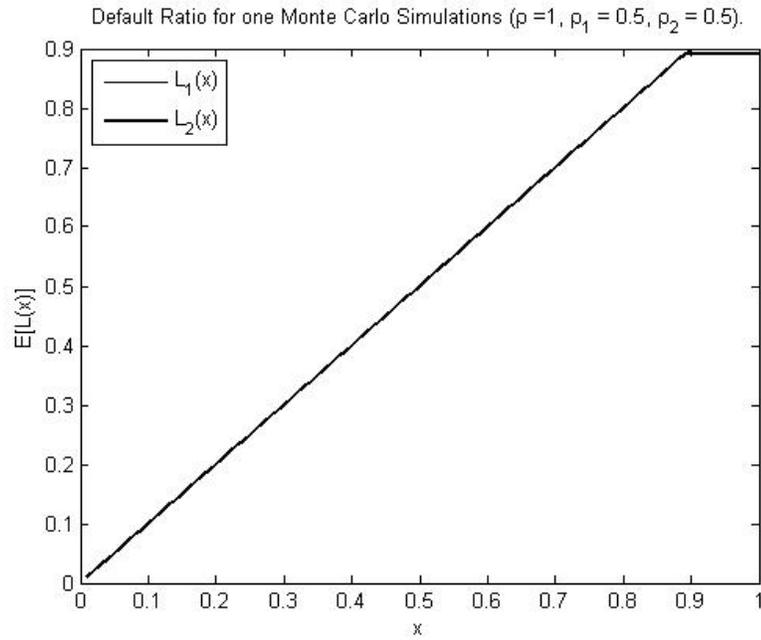


FIGURE 3.13. Strongly correlated default ratios $\rho = 1$ resulting in particular case of the same default ratios for both subportfolios $L_1(x) = L_2(x)$.

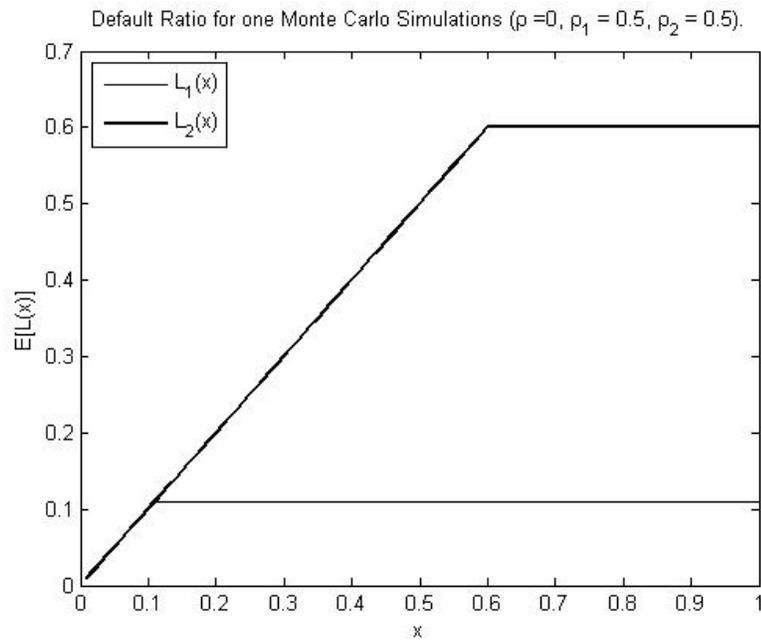


FIGURE 3.14. The independent default ratios $\rho = 0$ resulting in the random distance between L_1 and L_2 .

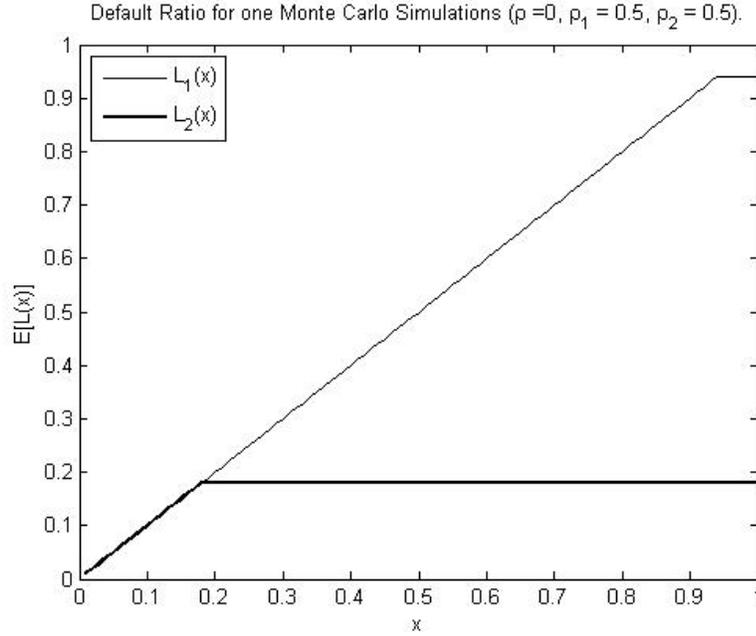


FIGURE 3.15. Another example of the independent default ratios $\rho = 0$ resulting in the random distance between L_1 and L_2 .

as

$$L_{12}(x) = \min \left\{ \frac{L_1 + L_2}{2N}, x \right\} \quad \forall x \in [0, 1] \quad (3.25)$$

It is further shown that $L_{12}(x)$ is sensitive to the joint behavior of L_1 and L_2 , and is affected by the correlation factor ρ . The strong dependency of ρ on L_1 and L_2 was already presented. This implies, that the joint default ratio $L_{12}(x)$ is also affected. Finding this particular pattern and behavior allows one to estimate the default ratio in much more precise way, assuming that the correlation is known.

Let us once more consider 1000 Monte Carlo simulations. The outcome of such simulation is the expected value of $L_1(x)$, $L_2(x)$ and now also $L_{12}(x)$. However, it is difficult to find particular pattern based only on one complete simulation. In case of generating random variables, where the previous computations are never the same as the current ones and different from the future ones as well, it is necessary to provide

sufficient amount of complete simulations to find a repeating pattern, if, naturally, such a pattern exists. Here, the starting point was 50 complete simulations, meaning 50 complete simulations of 1000 Monte Carlo simulations. Each complete simulation results in a graph of L_1 , L_2 and most desirably $L_{12}(x)$.

The first simulation attempt of 50 sets of 1000 Monte Carlo simulations is conducted for two portfolios (with subportfolios correlations $\rho_1 = \rho_2 = 0.5$ in both cases) with the correlation factors $\rho = 0.9$ for the first one and $\rho = 0.1$ for the second one. The major concern are the joint behavior $L_{12}(x)$ for each portfolio which are presented in Figure 3.16.

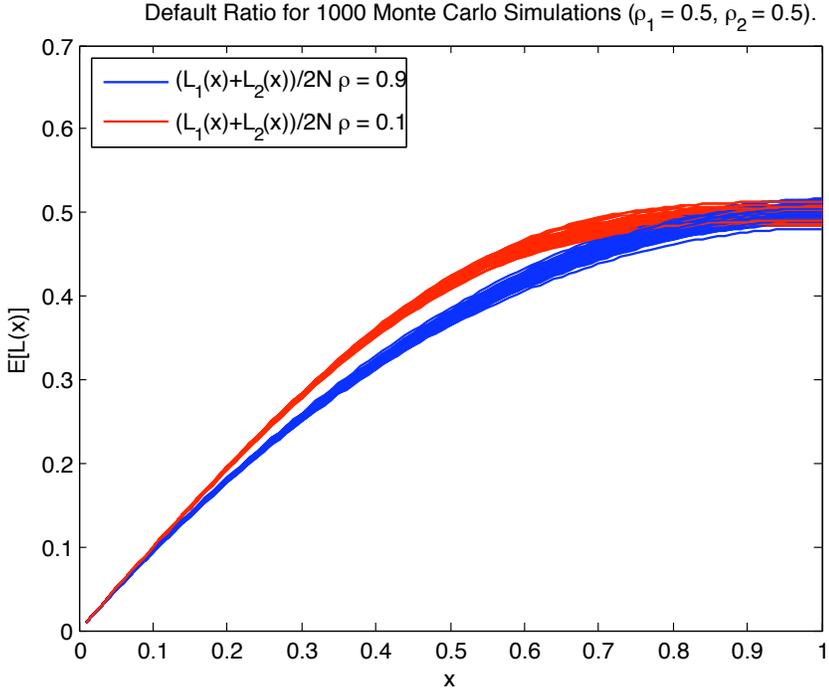


FIGURE 3.16. Joint default ratio for two portfolios and different correlation factors ρ .

With 50 complete simulations it is already possible to notice arising pattern. It is clearly visible that for $x \in [0.2, 0.8]$ the joint behavior of portfolios with different correlations differ. It can be noticed that for this particular range of x the portfolio with the higher correlation factor $\rho = 0.9$ has got smaller joint default ration ($E[L_{12}(x)]$)

(blue in Figure 3.16) than the portfolio with the lower default ration (red in figure 3.16). This arising pattern must, however, be compared to the simulations with different correlation factors in order to confirm or rebut this observation. In order to do so, the two following sets of correlations are performed for the same correlation parameter inside each subportfolio ($\rho_1 = \rho_2 = 0.5$) however, for the different correlation factors ρ . Figure 3.17 presents the simulations for $\rho = 0.5$ for both portfolios. These results confirm the pattern suggested above and visible in Figure 3.16. The same correlation factor for both portfolios provides the same results for the range $x \in [0.2, 0.8]$. This shows that the joint distribution is sensitive to the correlation factor changes.

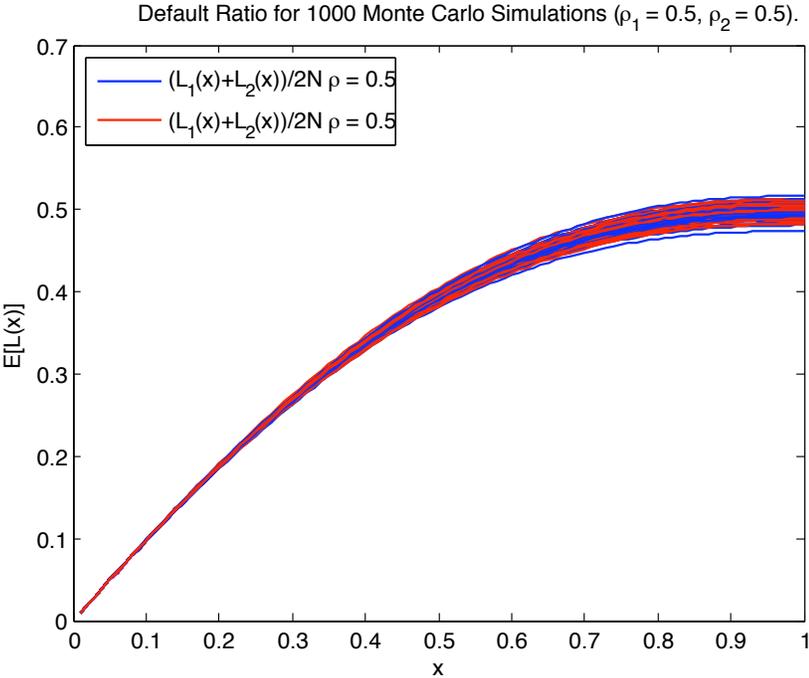


FIGURE 3.17. Joint default ratio for two portfolios and the same correlation factors ρ .

The last stage of completing the accuracy of the pattern already introduced in Figures 3.16 and 3.17 is conducting the simulation with the correlation factors $\rho = 0.1$ and $\rho = 0.9$, respectively. Figure 3.18 presents the obtained results that finally

concludes the initial observation. In this case the red curves with the higher correlation ($\rho = 0.9$) have smaller default ratio than the blue curves with the lower correlation factor ($\rho = 0.1$). Due numerous simulations for the wide range of correlation factors the monotonicity can be observed and clearly stated. The higher the correlation factor ρ the lower the joint expected default ratio of the portfolio.

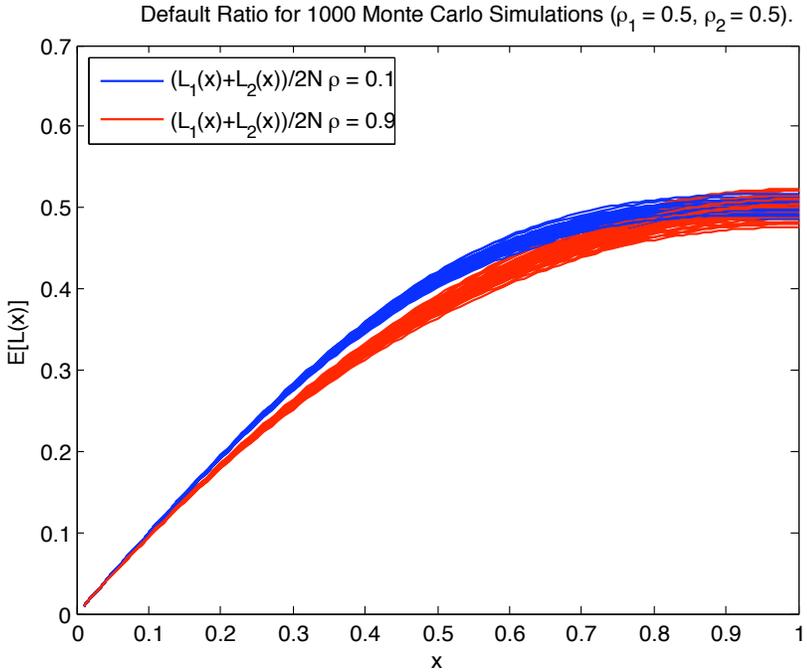


FIGURE 3.18. Joint default ratio for two portfolios with the opposite correlation factors ρ than in Figure 3.16 finally confirming the observed pattern.

Two following Figures 3.19 and 3.20 show, that regardless of the subportfolio correlation factor changes ρ_1 and ρ_2 the observation presented above still holds.

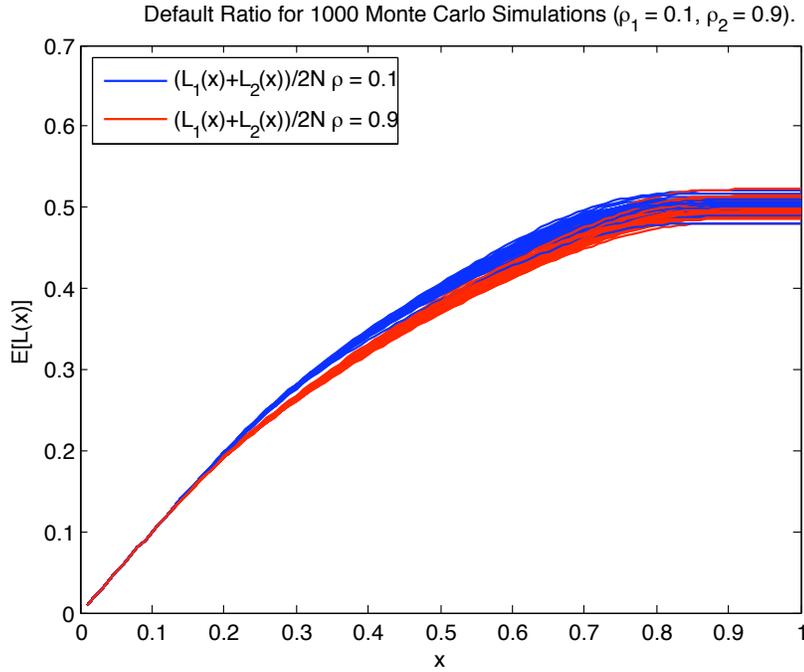


FIGURE 3.19. Joint default ratio for two portfolios with changed ρ_1 and ρ_2 proving that the proposed observation holds.

The issue under consideration is the correlation between the $L_1(x)$ and $L_2(x)$ depending on the correlation factor ρ . This correlation is given by the following formula

$$\text{Corr}(L_1(x), L_2(x)) = \frac{\sum_i (L_{i,1} - \bar{L}_1(x))(L_{i,1} - \bar{L}_2(x))}{\sqrt{\left(\sum_i (L_{i,1}(x) - \bar{L}_1(x))^2\right)\left(\sum_i (L_{i,2}(x) - \bar{L}_2(x))^2\right)}}$$

From the outcome of the numerous simulations it is clearly shown that the correlation coefficient ρ coincides with the correlation obtained from the above equation of the default ratios $\text{corr}(L_1(x), L_2(x))$. The maximum difference is less than 5%. This fact strongly indicates that knowing the correlation factor in the model presented here it is relatively easy to establish the correlation between losses in two subportfolios and moreover it is also proved that the higher the correlation the lower the expected losses in the joint default ratio.

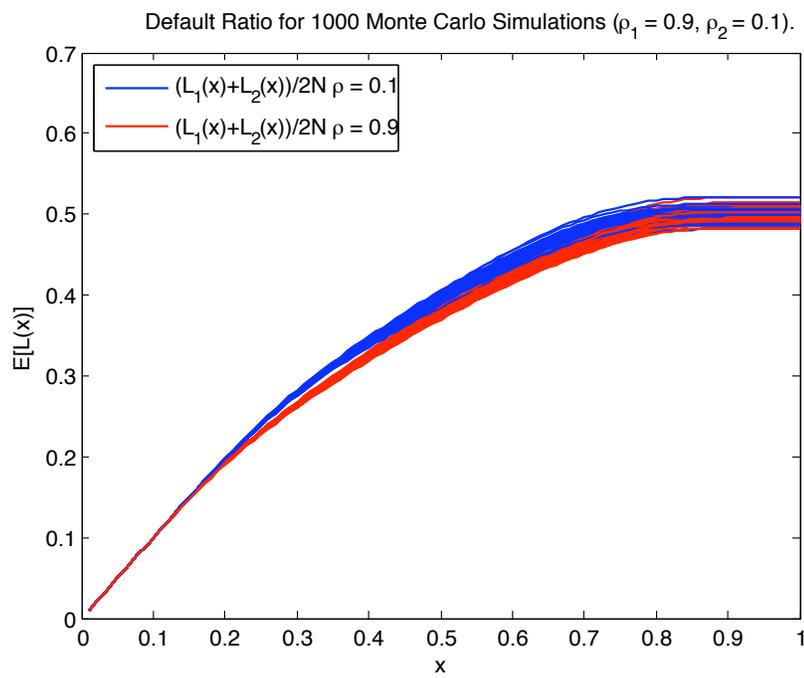


FIGURE 3.20. Last example of the joint default ratio for two portfolios with changed ρ_1 and ρ_2 finally proving the accuracy of made observation.

Chapter 4

Observations and Final Remarks

This work as it stands utilized and modified the initial idea of Li [5]. The first part gives a detailed insight into the theoretical aspects necessary to understand the proposed theoretical model. The main theorem governing the behavior of the copulas is stated showing the importance as well as simplicity of the presented approach (for proof refer to [1]).

After this theoretical introduction, the definition of the proposed model is stated. The first case includes the description of a simple portfolio without any subdivisions. Later one extends the simple approach and expands it to the case where a single portfolio is internally divided into two different subportfolios.

Next, the numerical algorithm is described and the main points are stated. The significant part of this work involves the numerical simulation. Developing the code in accord with the theoretical approach is of a crucial importance leading to obtaining the expected results. Hence, the introduction of the code is followed by its verification, thanks to which proper observation are made.

It was also shown that the numerically generated variable X according to the presented model follows a normal Gaussian distribution, which confirms the theoretical assumptions. For a constant correlation coefficient, the number of losses and explained following the theory presented in this work. It can be clearly seen that the generated graphs are in accord with the theory in this case as well.

The numerical approach comparison with the theory concentrates on larger amount of Monte Carlo simulations M . Taking $\rho = 0.5$ and sufficiently large M , one can derive that the expected value of the default ratios $L_1(x)$ and $L_2(x)$ is converging

to 0.5, as $M \rightarrow \infty$. Following this idea it was observed that the correlation factor for a particular portfolio plays an important role in estimating the expectation of the default ratio. Moreover, an increase of the correlation factor governing a particular subportfolio causes a decrease of the expectation of its default ratio. This observation is of a significant importance to the market. The presented model provides the new approach in linking the correlation with estimating the default ratio in complex portfolios assuming that the correlation between particular subportfolios is known. The last observation made for the section validating the numerical implementations arise from the pure logic- the higher the threshold the higher the probability of default for a single security. The numerical implementation mimics also this observation in a neat manner.

The last part of this work consist of results mostly regarding changes in the correlation factor, however for the joint behavior. It is of a crucial importance to investigate the expected default jointly and try to find proper patterns based on existing correlations between them. In this approach one starts with a single portfolio and than moves to a single portfolio with two subportfolios. For example, it can be viewed as an investor's portfolio divided into securities traded in the USA and securities traded in Asia. However, the concept presented by this model can be easily extended to more than two subportfolios. The numerical code was designed in such a way that adding additional numbers of subportfolios or increasing the number of securities in each of them is relatively easy. The biggest difficulty lies in obtaining the correlation factors governing the relations between the particular subportfolios and their securities. Having these correlation factors available from the market and utilizing the model presented in this work it is possible to simulate the default ratios and propose the solutions for particular portfolios with smaller default rates.

Next, the simplified correlation factor for two subportfolios in a single portfolio is analyzed. It is observed that the joint distribution $L_{12}(x)$ strongly depends on the

correlation factor ρ . For numerous sets of 1000 Monte Carlo simulations it can be clearly seen that the increase of the correlation factor decreases the expected default ratio of two subportfolios considered jointly. Knowing the correlation factors between large amount of subportfolios it would be possible to find the best option with the smallest probability of default. It is observed and clearly stated that the largest difference occurs for the tranche level in the range $[0.2, 0.8]$ whereas for smaller tranches it does not matter that significantly. This is another interesting observation and information to the market that this research affects mainly the investors insuring the tranches above 0.2, since then the difference becomes more significant. The investors insuring the tranche up to a certain level less than 0.2 are not going to be affected by these results. This pattern further analyzed and confirmed by numerous of numerical simulations shows that despite the amount of sets of 1000 Monte Carlo simulations or changes in the correlation factor within the particular subportfolio ρ_1 and ρ_2 the general results, i.e. the expected value of the default of the securities in the two subportfolios considered jointly decreases with the increase of the correlation factor ρ , is valid for all cases. Finally, it is observed that the correlation between the default ratios of the subportfolios is within close vicinity of the correlation factor ρ .

Bibliography

- [1] U. Cherubini, E. Luciano and W. Vecchiato, *Copula Method in Finance*, John Wiley & Sons, 2004.
- [2] Ryan Coelho, *Characterization of the Dependency Across Foreign Exchange Markets Using Copulas*, MS Thesis, The Department of Mathematics, Louisiana State University, December, 2006.
- [3] Eric Hillebrand, Ambar N. Sengupta and Junyue Xu, *Temporal Correlation of Defaults in Subprime Securitization*, Working Paper, 2010.
- [4] Hui-Hsiung Kuo, *Introduction to Stochastic Integration*, Universitext, Springer, 2000.
- [5] D. Li, *On Default Correlation: A Copula Function Approach*, Journal of Fixed Income, Vol. **9**, 2000. p. 43-54.
- [6] Chao Meng, *Stochastic and Copula Models for Credit Derivatives*, PhD Thesis, The Department of Mathematics, Louisiana State University, August, 2008.
- [7] Chao Meng and Ambar N. Sengupta, *CDO Tranche Sensitivities in the Gaussian Copula Model*, Preprint, 2008.
- [8] R. B. Nelsen, *An Introduction to Copulas*, Lecture Notes in Statistics, Springer, New York, first edition, 2000.
- [9] Ambar N. Sengupta, *Pricing Derivatives*, McGraw-Hill, 2005.
- [10] Ambar N. Sengupta and Morten Andersen, *A CDO Primer*, Notes, 2005.
- [11] A. Sklar, *Random variables, joint distribution, and copulas*, Kybernetika, Vol. **9**, 1973. p. 449-460.
- [12] Daniel Totouom and Margaret Armstrong, *Dynamic Copula Processes: A New Way of Modelling CDO Tranches*, Advances in Econometrics: Econometrics of Risk Management, Vol. **22**, 2007.

Appendix: Numerical Algorithm

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%      COPULA METHOD IN FINANCE      %%
%%      MSC MATH                      %%
%%      by                            %%
%%      Adam Lodygowski               %%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
format long
clear
clc
%
% BEGINNING
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% PART 1, generate standard gaussian variables Z, u_1 and u_2 to obtain Z_1
% and Z_2 and later num_sec of eps(num_cdo,num_sec), num_cdo = 2
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% INPUT DATA %%%%%%%%%
% number of subportfolios
num_cdo = 2;
% Number of securities
num_sec = 100;
% Number of Monte Carlo simulations
N = 1000;
% set correlation factor rho
rho_init = 0.5;
% Define rho(1,2,...,num_cdo), for this case num_cdo = 2
rho(1) = 0.5;
rho(2) = 0.5;
% create threshold X_star
X_star = 0;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% END INPUT DATA %%%%%%%%%
%
% BEGINING OF SINGLE MONTE CARLO SIMULATION, k runs from 1 to N, where N is predefined
% total number of simulations
% Initialize matrix of defaults L for N simulations and num_cdo number of
% subportfolios
L = zeros(N,num_cdo);
for k = 1:N
    % First create standard gaussian variables Z~N(0,1), U~N(0,1) and eps~N(0,1)
    % Z_init is constant for all subportfolios, U is different for each
```

```

% subportfolio U_init(num_cdo)
% Create common risk factor
Z_init = randn(1,1);
for i = 1:num_cdo
    U_init(i) = randn(1,1);
    Eps(:,i) = randn(num_sec,1);
end
%x = -3:0.1:3;
%hist(Eps(:,1),x);
%hist(Eps(:,2),x);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% PART 2, Generate num_cdo Z's based on common Z and Eps(1,:) and Eps(2,:)
% respectively
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% NOTE this case is for 2 subportfolios but can be easily extended to more
% subportfolios
% Define risk factors for num_cdo subportfolios (here 2)
% i stays for number of subportfolios
for i = 1:num_cdo
    Z(k,i) = sqrt(rho_init) * Z_init + sqrt(1-rho_init) * U_init(i);
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% PART 3, Calculate the state variable X for each subportfolio
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% number of securities in each subportfolio as defined above num_sec
% Generate num_sec individual state variables X(i)
% use gaussian Z and Eps(num_cdo) calculated in PART 1
% initiate sum_D - number of X(i) being less than X_star and hence number
% of defaults, number of Losses
% initiate matrix L representing number of defaults in each Monte Carlo
% simulation for each subportfolio (column 1- defaults in subportofio 1,
% column 2- defaults in subportfolio 2 and so on...)
% i is again the number of subportfolio, whereas j stays for the number of
% current security
X = zeros(num_sec,num_cdo);
for i = 1:num_cdo
    sum_D = 0;
    for j = 1:num_sec
        X(j,i) = sqrt(rho(i)) * Z(k,i) + sqrt(1-rho(i)) * Eps(j,i);
        if X(j,i) < X_star
            D = 1;
        else
            D = 0;
        end
    end
end

```

```

        end
        sum_D = sum_D + D;
    end
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    % PART 4, Calculate number of defaults for each subportfolio
    % Default if  $X < X_{star}$  -->  $D = 1$ , otherwise  $D = 0$ 
    %  $L = \sum_{i=1}^{\{num\_sec\}} D(i)$ 
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    L(k,i) = sum_D;
end
% where L(1), L(2), ..., L(num_cdo) are the default rates of subportfolios
end
% END OF N MONTE CARLO SIMULATION
% Uncomment following 3 lines to check the distribution of X
%x = -3:0.1:3;
%hist(X(:,1),x);
%hold on
%hist(Eps(:,2),x);
% CHECK 1: the correlation between Z_1 and Z_2 for all Monte Carlo
% Simulation should be close to the given correlation factor rho
Check_rho = corr2(Z(:,1),Z(:,2));
% CHECK 2: since X is standard gaussian  $X \sim (0,1)$  hence for  $X_{star} = 0$  and the
% same rho;  $\rho_1 = \rho_2$  the probability of default should be 50%
Check_default_rate = mean(L);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% PART 5, calculate the minimum value, for each
% x in [0,1]  $L_x(1) = \min\{L(1)/N.x\}$ 
% this corresponds to the standard tranches mainly used as 3,7,10 and 30%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% initialize L_x
num_div = 100;
L_x = zeros(N,num_cdo);
L_single = zeros(num_div,num_cdo);
% define the standard tranche x_tranch
x_tranch = 1;
% take the minimum value of losses and standard tranche
% k stays for number of Monte Carlo simulations and i stays for number of
% subportfolios
% in following loop j stays for changing x_tranch from 0 to 1 with 0.01
% interval
for j = 1:num_div
    for k = 1:N
        for i = 1:num_cdo

```

```

        L_x(k,i) = min(L(k,i)/num_sec,j/num_div);
        % single simulation, comment lines 125 and 131 and change
        % plotting options (read lines 137, 143 and 148)
        L_single(j,i) = min(L(1,i)/num_sec,j/num_div);
    end
end
    x(j) = j/num_div;
    L_x_plot(j,:) = mean(L_x);
end
plot(x,L_x_plot(:,1),'k')
% for single simulation uncomment line below and comment line above
% plot(x,L_single(:,1),'k')
hold on
plot(x,L_x_plot(:,2),'k','linewidth',2);
% for single simulation uncomment line below and comment line above
% plot(x,L_single(:,2),'k','linewidth',2);
legend('L_1(x)', 'L_2(x)',0);
xlabel('x'),ylabel('L(x)');
title(['Default Ratio for ',num2str(N),' Monte Carlo Simulations ...
(\rho =',num2str(rho_init),', \rho_1 = ',num2str(rho(1)),',...
\rho_2 = ',num2str(rho(2)),').']);
% for single simulation uncomment line below and comment line above
% title(['Default Ratio for one Monte Carlo Simulations ...
% (\rho=',num2str(rho_init),', \rho_1 = ',num2str(rho(1)),',...
% \rho_2 = ',num2str(rho(2)),').']);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% PART 6, calculate the correlation between L_x(1) and L_x(2)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
L_x_correlation = corr2(L_x(:,1),L_x(:,2));
%L_x_correlation = corr2(L_single(:,1),L_single(:,2))

```

Vita

Adam Lodygowski was born in July 1981, in Poznan City, Poland. He finished his undergraduate studies at Poznan University of Technology and earned a master of science degree in civil engineering from both University of Hanover and Poznan University of Technology in May 2005. In August 2006 he came to Louisiana State University to pursue graduate studies in civil engineering. He is currently a candidate for the degree of Doctor of Philosophy in civil engineering which will be awarded in May 2010. In spring 2008, in addition to the doctoral program, he enrolled in the graduate studies in mathematics with concentration in finance and currently he is a candidate for the degree of Master of Science in mathematics, which will be awarded in May 2010.