Univerzita Karlova v Praze
Fakulta sociálních věd

Institut ekonomických studií

Diplomová práce

2007 Petr Jablonský

Univerzita Karlova v Praze
Fakulta sociálních věd

Institut ekonomických studií

DIPLOMOVÁ PRÁCE

Measuring credit risk for portfolios with heavy-tailed risk factors

Vypracoval:	Bc. Petr Jablonský
Vedoucí:	Doc. Ing. Miloslav Vošvrda, CSc.
Akademický rok:	2006-2007

Hereby I declare that I compiled this thesis independently, using only the listed literature and resources.

Prague, May 17, 2007

Petr Jablonský

ACKNOWLEDGMENTS

I would like to express my gratitude to Miloslav Vošvrda of the Faculty of Social Sciences, Charles University for supervising my work on this thesis. I also thank Jan Kodera for providing me with consultations on Matlab software and Ludmila Rosová for valuable stylistic revision and emendation.

ABSTRACT

Measuring and managing credit risk constitute one of the most important processes within bank risk management. Classical credit risk models assume multivariate normality for distribution of underlying risk factors. Resulting methods offer analytical simplicity and computational efficiency but disregard of extreme joint events since their probability is too small. Recently several studies have doubted multivariate normality assumption saying that if we accept this assumption we might seriously underestimate downside risk of given credit portfolio.

The master thesis provides with an insight into the problem of modelling credit risk under assumption of heavy tailed risk factors. We first present necessary mathematical preliminaries of copula functions which stand for an alternative method of modelling multivariate dependence structures. Next we introduce a credit risk model for bond portfolio with heavy tailed risk factors. At last we carry out several simulations on portfolios of different riskiness and compare to what extent the results from both mentioned models differ.

ABSTRAKT

Měření a řízení kreditního rizika představuje jeden z nejdůležitějších procesů v rámci řízení bankovních rizik. Klasické modely kreditního rizika předpokládají vícerozměrnou normalitu pro rozdělení rizikových faktorů. Takto definované modely nabízejí analytickou jednoduchost a výpočetní efektivitu, avšak odhlížejí od extrémních sdružených událostí, neboť pravděpodobnost nastání takových jevů se zdá býti příliš malou. V poslední době několik studií zpochybňuje vícerozměrnou normalitu a říká, že takovýto předpoklad vážně podhodnocuje výsledný odhad kreditního rizika portfolia.

Diplomová práce se zabývá modely kreditního rizika za předpokladu rizikových faktorů s těžkými chvosty. Nejdříve popisujeme nezbytné matematické metody kopula funkcí, které představují alternativní přístup modelování vícerozměrných závislostních struktur. Poté představujeme vlastní model kreditního rizika portfolia obligací s rizikovými faktory s těžkými chvosty. Závěrem popisujeme několik simulací provedených na obligačních portfoliích o různé rizikovosti a porovnáváme do jaké míry se výsledky z obou zmíněných modelů liší.

TABLE OF CONTENT

1. INTRODUCTION

In recent development of financial markets we observe an ever-rising importance and popularity of risk management science. Interest in deeper understanding of uncertainty and threats, which financial institutions are exposed to, comes not only from boards of such institutions but as well from directive decisions of regulatory authorities. European Basel Committee on Banking Supervision currently recognizes three major type of risks. These are market risk, credit risk and operational risk. Among all of them, the credit risk have always constituted the major concern for all banks and lending institutions around the world.

As a credit risk we understand an uncertainty in a counterparty's ability to fully meet its obligations on financial as well as non-financial contracts[1]. An effort to capture and quantify the risk of loss due to credit has led to establishment of numerous amount of credit risk models. Despite using different methodological approaches, all models are driven by three basic common factors. First one is the **probability of default** which simply means assessing the chance that undesirable outcome would ever happen. Secondly, the **credit exposure** stands for amount of money which are in danger of loss in case of default. It is important to be aware that this might not always be easily defined. Default can occur practically anytime from the first to the last year of the contract duration and therefore for instance various instalments and discount rates have to be considered in the analysis. Lastly if default really happens, it is important to determine a fraction of the claims which could be recovered through bankruptcy proceedings – so called **recovery rate** stand for this.

Within the framework of **credit risk modelling**, the probability of default is the most challenging from all the mentioned factors. The analysis becomes even more involved when we consider a **portfolio of credit instruments**. In such case mutual dependency of portfolio segments needs to be taken into account and modelling such dependence structures poses the key element in portfolio construction and credit risk management. Traditionally, a **linear correlation** serves as the most popular measure of dependence between random variables and therefore it is not too much surprising that leading commercial credit risk models are built upon the assumption of that. Let us first describe the methodology of most popular models and then discuss their shortcomings.

Probably the most influential credit risk model ever was introduced by Robert C. Merton in 1974 and is called in his name. Merton's both revolutionary and intuitive idea was to see

---

[1] http://www.riskglossary.com/link/credit_risk.htm

firm's equity as a call option on company's assets[2] **and defining default as a case when firm's equity value falls bellow some predetermined threshold**. Studying probability of default on the base of such defined mark-to-market models gained a lots of popularity and two of the currently most important commercial industry models are the followers of Merton's thoughts. **KMV's model** (1993) continues in further development of Merton's option pricing theory, while **CreditMetrics^{TM}** model (introduced by J. P. Morgan in 1997) brings new methodology based on estimation of obligor's credit rating and credit rating migration. The both models are, hence, of a latent variable type[3].

From the portfolio point of view, the dependence between defaults of portfolio segments is driven by mutual dependence structure between latent variables which are underlying these segments. Both KMV's and CreditMetrics^{TM} models assume that those **underlying dependence structures are of multivariate normal class of distribution**. However recently there has been interesting amount of studies criticizing such approach for being too restrictive. In the following chapter we give the main arguments.

1.1 LITERATURE SURVEY

Apparently the most rational reason for assuming multivariate normal dependence structure for underlying risk factors might be found in **analytical simplicity of the resulting models**[4]. It might be well shown that the shape of multivariate normal distribution is driven entirely by structure of its covariance matrix. Since the only parameters of covariance matrix are the variances of marginal distributions and the value of their linear correlation, which might be both extracted from the latent variable of equity returns, we get a very appealing and easily assessable model. But the main question remains. **Is it admissible to assume just multivariate normal distribution structures between latent variables? And how inaccurate would be exercitation of such models on data sets where the multivariate normality assumption would be not fulfilled?**

Mashal (2002) finds that assumed normal dependence structure of joint equity returns had never been tested. Apart from Mashal's study, there has been an impressive number of studies negating the CreditMetrics^{TM} and KMV's corner-stone. Nyfeler (2001) states that

---

[2] As any other option, this might be valuated by using **Black-Scholes option pricing formula**. Hence we have a five determinants of price - strike value, asset market value, volatility of underlying asset, time to maturity and risk free rate.

[3] As a **latent variable** we understand the value of obligor's assets and the default occurs when the latent variable falls bellow some given threshold.

[4] Embrechts et al. (1999)

there is no compelling reason for choosing multivariate normal (Gaussian) distribution for asset values and Embrechts (2002) adds that empirical research in finance and insurance shows that distributions of real world assets are seldom in the class of multivariate normal distributions. Domingues (2004) continues with declaring that while it is legitimate to assume normality of the portfolio changes due to market risk, it is no longer the case for credit returns. Domingues (2004) further shows that **portfolio of credit instruments has highly skewed distribution which exhibits heavy tails** which is in direct contradiction with normal dependence structure[5].

Taken from the another point of view, the multivariate normal distribution automatically incorporates linear correlation as the measure of dependence between given random variables. However, as we will show, such a thinking might be misguided. Embrechts (1999) sees correlation to be not only one of the **most ubiquitous but as well the most misunderstood concepts in modern finance**. Correlation is often associated with thought of being able to measure any dependence. Linear correlation, however, is only the measure of linear dependence. Embrechts continues that **correlation is canonical measure only in the world of multivariate normal class of distribution**. Moreover Longin and Solnik (2001) show that linear correlation is **not able to capture dependence between low probable but extreme events**.

1.2 STUDY CASE

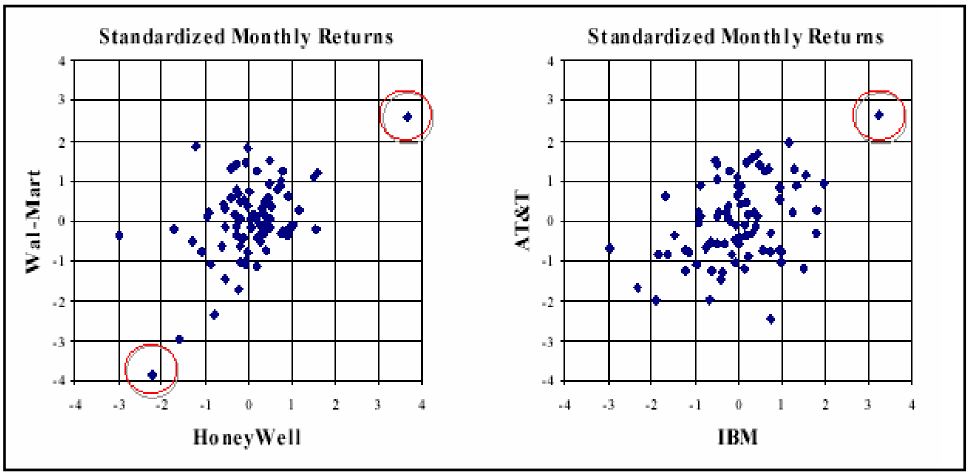
For purposes of studying heavy tailed risk factors, the **main pitfall of multivariate normality assumptions is the small probability of extreme joint events**. We next present a study case which gives empirical evidence that extreme events occur and that they occur more often than what would be predicted under the multivariate normal distribution structures.

Mashal (2002) takes and plots bivariate standardized monthly returns of four US firms and highlights the extreme joint events. Out of 84 observations depicted in Figure 1, we can find two, respectively one occurrence of extreme joint realisations. Probability of presence each of highlighted events in bivariate normal distribution is approximately once in 100,000 observations. Given this fact Mashal concludes that for investigated bivariate distribution to be normal, the quantity of extreme joint realizations is too high and that equity returns have fat tails not only in their marginal distributions but as well in higher dimensions.

---

[5] For more information about **multivariate normality pitfalls** see: Embrechts: Correlation and dependence in risk management: properties and pitfalls , 2002

Figure 1: Standardized US equity returns, August 1994-July 2001



Source: Mashal (2002)

1.2.1 EVIDENCE FROM CZECH STOCK MARKET

We executed the same research as Mashal on the data from Czech stock market. We used GARCH(1,1)[6] filtered monthly returns of **firms ČEZ, Unipetrol, Telefonica O_2 and Komerční banka** from period January 2001 to February 2007. The results are depicted in Figure 2. We **highlighted some of observations whose occurrence, given assumption of bivariate normality, should not be larger than once per five hundred events** (and actually should be lower, for extreme cases even much lower). We stress the fact that each graph contains mere 77 observations[7].

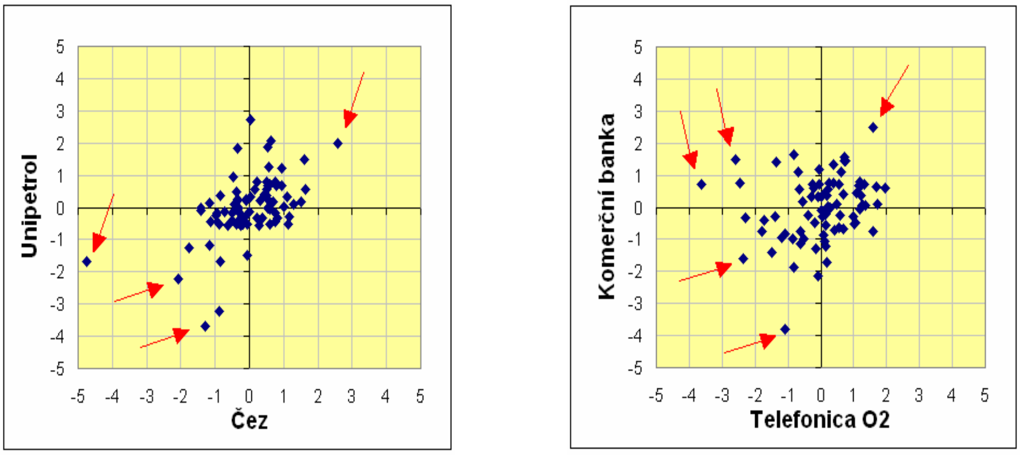
Comparing to the Mashal's results from US data, we find **even higher evidence of extreme joint realizations**. Our personal opinion is that this might be due to the specific characteristics of small and still developing stock market in Czech republic. Nevertheless the results give us support in an effort to investigate the issue of extreme joint events more closely.

---

[6] We give more information about GARCH (1,1) in chapter 5.1

[7] note: implied **assets linear correlation**: Corr(CEZ,Unipetrol)=0.39; Corr(TelefonicaO2, KB)=0.37

Figure 2: GARCH (1,1) filtered Czech equity returns, January 2001-February 2007



1.3 STRUCTURE OF DIPLOMA THESIS

Summing up the previous chapters, we showed that leading industry credit risk models are built in the set-up of multivariate normal distributed variables which are underlying the models' risk factors. We presented the evidence that such assumption might not be appropriate. First of all, the marginal assets' distributions do not have to be normally distributed and actually in reality, often, they are not. Secondly multivariate normal distribution implies linear correlation as a measure of dependence and such defined dependence structure is not able to fully capture possible occurrence of extreme joint events. Or better to say, the probability of extreme joint events under linear correlation dependence structures and multivariate normal distributions is so small that it makes almost no sense to consider such events to ever happen. However disregarding of extreme joint events might be found irresponsible as the potential losses are huge and might bring severe consequences.

In this diploma thesis we are presenting models which offer an alternative to the commercial credit risk models of KMV or CreditMetrics^{TM} type and study and incorporate the chance of realization of low probable extreme joint events. We comparatively study, theoretically as well as practically using the Czech stock market data, to what extent the results from presented models differ.

We start with presentation of necessary mathematical background. Beside the linear correlation there exist many alternative models assessing dependence structures.

Considering risk management, model of copula functions presents the state of art of current science and belongs recently to the most popular ones. We give a little insight to the copula functions and introduce alternative dependence measures to linear correlation in the chapter 2. Chapter 3 is dedicated to the CreditMetrics^{TM} model which is further used as a benchmark for comparative study of our results. In chapter 4 we finally present the credit risk model with incorporated heavy tailed risk factors and in chapter 5 discuss its results given the real data from Czech stock market as well as couple of artificial portfolios of different riskiness. All simulations in this thesis were carried out with program Matlab 6.1 with Statistics Toolbox. Lastly, necessary program codes and some supplementary files are provided on CD which is enclosed to the thesis.

2. MATHEMATICAL PRELIMINARIES

One of the fundamental problems with portfolio credit risk management is perceived as **how to aggregate individual obligor's risks with respect to specific dependence structures between them**. In the previous chapter we stated that models, which are built upon multivariate normal distributions, could lead to results which underestimate the probability of joint defaults of credit obligors. The alternative way how to describe the dependence structure of random variables might be found in **the concept of latent variable copula functions**. Due to their properties, which are presented later in the chapter, copula functions comprise a powerful and very flexible tool for multivariate dependence structure analyses which enable us to describe the examined data in more detail.

Going back through the history we find that the concept of copula functions was firstly introduced by Abe Sklar in 1959. However it was not until late 90's when copulas started to attract attention of financial fields. At these days, copulas are extensively used to investigate dependencies between random variables on the fields of operational, market and most of all the credit risk management. We introduce **copula function as an alternative method for modelling dependence structure between random variables**. From the point of view of classical statistical science, the dependence structures of random variables are completely describe by a joint distribution. If we have F(x) to be a **joint distribution of random variables** $X_1, \ldots, X_n$, we mean:

$$F(x) = P(X_1 < x_1, \ldots, X_n < x_n) \tag{2.1}$$

The motivation standing behind the copula concept of dependency is **to separate the joint distribution function F(x) into two parts**. First one for representing the margins[8] and second one, described by copula function, for modelling the dependence structure between those margins. We dedicate this whole chapter to the copula functions starting with basic definition and important theorems, going through alternative concepts for measuring dependence structures of random variables and finishing with models fitting the copula functions onto given data sets.

---

[8] Throughout the thesis as **margins** we understand the distribution functions of the marginal random variables.

2.1 PROPERTIES OF COPULA FUNCTIONS

Definition 1 (Copula function)

For $n \in N/\{0\}$ a n-dimensional distribution function with margins uniformly distributed on [0,1] is called a copula.

Copula can be alternatively fully defined to be any function which can be used as a distribution function on $[0,1]^n$:

*A copula is **any** function $C : [0,1]^n \to [0,1]$ which fulfil the following properties:*

1. $C(x_1,\dots,x_n)$ *is increasing in each component* x_i
2. $C(1,\dots,1,x_i,1\dots,1) = x_i \quad \forall i \in [1,\dots,n], x_i \in [0,1]$
3. $\forall\ (a_1,\dots,a_n), (b_1,\dots,b_n) \in [0,1]^n$ *with* $a_i \le b_i$ *we have*

$$\sum_{i_1=1}^{2}\dots\sum_{i_n=1}^{2}(-1)^{i_1+\dots+i_n} C(x_{1i_1},\dots,x_{ni_n}) \ge 0$$ [9]

where $x_{j1} = a_j,\ x_{j2} = b_j \quad \forall\ j \in \{1,\dots,n\}$

There are several propositions about the properties of copula functions but none of them has reached the importance of Sklar's theorem which sets up the basis of copula dependence modelling. The theorem clearly shows that **a multivariate distribution function can be separated and completely described by its univariate margins and a multivariate dependence structure**. Moreover, such multivariate dependence structure is entirely characterized by copula. The significance of the Sklar's theorem is found not only in its primary meaning but as well in its several propositions which further enlarge the general knowledge about the copula function modelling.

Sklar's theorem

Let H be an n-dimensional distribution function with margins $F_1,\dots, F_n$. Then there exists an n-copula C such that $\forall x \in R^n$:

$$H(x_1,\dots,x_n) = P(X_1 < x_1,\dots,X_n < x_n) = C(F_1(x_1),\dots,F_n(x_n)). \tag{2.2}$$

If $F_1,\dots,F_n$ are all continuous, then C is unique.

Conversely, if C is an n-copula and $F_1,\dots,F_n$ are distribution functions, then the function H defined as above is an n-dimensional distribution function with margins $F_1,\dots,F_n$.

---

[9] The expression in the sum might be interpreted as $P[a_1 \le X_1 \le b_1 ,\dots, a_n \le X_1 \le b_n]$

Proposition 1

Let H be a n-dimensional distribution function with continuous margins $F_1,\ldots,F_n$ and copula C. Then for $\forall x \in [0,1]^n$ we have:

$$C(x_1,\ldots,x_n) = H(F_1^{-1}(x_1),\ldots,F_n^{-1}(x_n)), \tag{2.3}$$

where $F_i^{-1}(x_i), i = 1,\ldots,n,$ denotes the generalized inverse function.

Hence the first corollary gives an **advice how the concrete form of copula function might be derived** under specified multivariate distribution function and its margins. We use this proposition and show its practicality in the next chapter 2.2, where we derive particular form for several most popular copula functions. From the general properties of distribution functions and Sklar's theorem arises another proposition.

Since $P(X_1 < x_1,\ldots,X_n < x_n) = \prod_{i=1}^{n} P(Y_i < y_i)$ holds if and only if the components of X are independent, we come to a particularly simple but conceptually important case which is summarised in the next statement.

Proposition 2

Let X be a random vector with continuous margins, then the components of X are independent if and only if the copula representation C has the following form:

$$C(u_1,\ldots,u_n) = \prod_{i=1}^{n} u_i \tag{2.4}$$

Last proposition says that **copula representation shows a lots of universality as it is invariant under strictly increasing and continuous transformations of the margins**. Such invariance allows for high flexibility of copula models. For instance if we are building a model based on equity returns we can easily transform the input data into log-values, which is in this type models a common practice, without any need for changing the copula function as logarithmic transformation of margins is strictly increasing and continuous operation and therefore has no impact on copula representation.

Proposition 3

If $(F_1,\ldots,F_n)T$ has copula C and $T_1,\ldots,T_n$ are increasing and continuous functions, then $(T_1(F_1),\ldots,T_n(F_n))T$ also has copula C.

Proof.

Let $(U_1, \dots, U_n)^T$ *have distribution function C. In the case of continuous margins* F_{X_i} *we take* $U_i = F_{X_i}(X_i)$. *Then we may write:*

$$\begin{aligned} C\left(F_{T_1(X_1)}(x_1),\dots,F_{T_n(X_n)}(x_n)\right) &= P\left(U_1 \le F_{T_1(X_1)}(x_1),\dots,U_n \le F_{T_n(X_n)}(x_n)\right) \\ &= P\left(F^{-1}_{T_1(X_1)}(U_1) \le x_1,\dots,F^{-1}_{T_n(X_n)}(U_n) \le x_n\right) \\ &= P\left(T_1 \circ F^{-1}_{X_1}(U_1) \le x_1,\dots,T_n \circ F^{-1}_{X_n}(U_n) \le x_n\right) \\ &= P\left(T_1(X_1) \le x_1,\dots,T_n(X_n) \le x_n\right) \end{aligned}$$

■.

The aim of this chapter was to illustrate the basic properties of copula functions. For those readers who are interested in deeper knowledge of this subject matter we recommend any publications of P. Embrechts or R. Frey. Both of them give very relevant studies on the field of dependent default and copula modelling as well as on copula theory itself.

2.2 EXAMPLES OF MOST COMMON COPULA FORMS

Sklar's theorem shows how copula *C* creates a connection from the margins $F_1,\dots,F_n$ to the joint distribution *H*. Therefore if we keep the margins $F_1,\dots,F_n$ fixed and change the functional form of copula C, we can simulate different kind of joint dependencies between given random variables[10]. In what follows we present very concrete forms of most popular copula concepts.

However before we describe the mentioned copula concepts we outline a terminology which is associated with copula classification. First of all, we distinguish between **implicit copula** whose density function is given by a known multivariate distribution (e.g. multivariate normal or t-distributions are very commonly used) and **explicit copula** whose distribution function has a closed-form expression[11]. Further according to the particular form of copula representation we recognise several copula subclasses. From those, the most important ones are the **subclasses of elliptical copulas** and **Archimedean copulas**. We describe both of them in the following sub-chapters.

---

[10] If not otherwise stated, we suppose that all **risk factors** $X_1,\dots,X_n$ have univariate N(0,1) distribution. As well by saying Gaussian or normal distribution we mean N(0,1) distribution, multivariate or univariate.

[11] Wikipedia explains that an equation or system of equations is said to have a **closed-form solution** if, and only if, at least one solution can be expressed analytically in terms of a bounded number of certain well-known elementary functions..

Nevertheless first of all, the easiest form offers the **copula of independent variables**. We already showed in Proposition 2 of the last chapter that in such case the copula is represented just by simple product of its margins. On the other hand, probably the most often used is so called **Gaussian or normal copula**. This is nothing else than the copula representation for multivariate normal distribution $N_n(0,I)$, i.e. the distribution of independent standard normal variables. Following Proposition 1, which says how to extract copula function from given joint distribution function, we get the formula :

$C_\rho^{Ga}(u_1,u_2) = H_\rho^{Ga}(\phi^{-1}(u_1),\phi^{-1}(u_2))$.

Where $\rho \in [0,1]$ stands for linear correlation parameter, ϕ is univariate N(0,1) distribution function and H_ρ stands for bivariate $N_2(0,1)$ distribution function with correlation coefficient ρ. Combining Proposition 1 with the definition of bivariate normal distribution[12] we then get standard Gaussian copula function as a following integral:

$$\boxed{C_\rho^{Ga}(u_1,u_2) = \int_{-\infty}^{\phi^{-1}(u_1)} \int_{-\infty}^{\phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\left(-\frac{(s^2-2\rho st+t^2)}{2(1-\rho^2)}\right) dsdt} \qquad (2.5)$$

2.2.1 SPHERICAL AND ELLIPTICAL DISTRIBUTIONS

The theory of spherical and elliptical distributions is clearly described in Embrechts et al. (1999). The **spherical distribution can be best understood as a natural extension of multivariate $N_n(0,I)$ normal distributions** which, as we said, stand for distributions of independent (and hence as well uncorrelated) standard normal variables. Following this the spherical distribution provides the family of symmetric distributions for random vectors with **zero mean** which are no more independent but still uncorrelated.

Definition 5 (Spherical distribution)

A random vector $X = (X_1, \dots, X_n)^T$ *is said to have a spherical distribution if for every orthogonal map* $U \in R^{nxn}$ [13] *it holds that* $UX \overset{dist}{=} X$

---

[12] For the **form of bivariate normal distribution** see for instance: http://en.wikipedia.org/wiki/Bivariate_normal_distribution

[13] **Orthogonal maps** are the maps satisfying $UU^T=U^TU=I_{nxn}$

It might be shown that if X has a density $f(x) = f(x_1,...,x_n)$ then for some function $g: R_+ \to R_+$ this is equivalent to $f(x) = g(x^T x) = g(x_1^2 + ... + x_n^2)$. Hence the **spherical distributions might be interpreted as the distributions with constant on spheres**.

Multivariate normal distribution, which was introduced in the last chapter, belongs to the class of spherical distribution and is the only spherical distribution with independent random variables. Apart from that the multivariate t-distribution and the logistic distribution are other two popular spherical distributions. In this paper, we will use only t-distribution and thus we do not provide exact formula for logistic distribution. However for logistic distribution we refer to Embrechts et al. (1999).

Concept of t-copula constitutes for our modelling the biggest importance. Firstly, comparing to Gaussian copula, t-copula **exhibits heavy tails as it is the representation of multivariate student's t-distribution**. Secondly in case of higher degrees of freedom, we might approximate t-copula by Gaussian copula which allows us to create a conjunction with models using multivariate normal distribution. But most importantly, Mashal (2002) shows that considering an ability to fit the empirical data, the **t-copula is generally superior to the Gaussian copula**. For all those reasons, t-copula belongs to the most popular copula concepts. We can derive t-copula form in the same way as the Gaussian copula following the Proposition 1. Standard bivariate t-copula function with ν degrees of freedom then looks as follows.

$$\boxed{C_{\nu,\rho}^{t}(u_1,u_2) = \int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{1}{2\pi\sqrt{(1-\rho^2)}} \left(1 + \frac{s^2 - 2\rho st + t^2}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}} ds dt} \tag{2.6}$$

where ρ = linear correlation parameter $-1 < \rho < 1$

t_ν = univariate standard t-distribution function with ν degrees of freedom

To sum up, the spherical distribution is a natural extension of multivariate $N_n(0, I)$ distribution. If we are looking for even more generalization we, simply, come to the **elliptical distributions**.

We are mentioning the elliptical distribution since they represent the next logical step in the distribution analysis. However as we will not use anything else than spherical distributions, we leave the elliptical distribution with saying that they extend the multivariate $N_n(\mu, \Sigma)$

normal distribution[14] and that they are the affine map of spherical distributions in R^n. Again for those who are interested in deeper knowledge about elliptical distribution, we refer to Embrechts et al. (1999). In the next chapter we introduce the second famous class of copula functions – Archimedean copulas.

2.2.2 ARCHIMEDEAN COPULA FUNCTIONS

Definition 6 (Archimedean copula function)

An Archimedean copula function $C : [0,1]^n \to [0,1]$ *is such a copula function which might be represented in the following form:*

$$C(u_1,...,u_n) = \psi^{-1}[\psi(u_1) + ... + \psi(u_n)]\,,$$

where ψ *is a decreasing convex function satisfying* $\psi : [0,1] \to R_+$ *with* $\psi(1) = 0,\ \psi(0) = \infty$. *Such a function* ψ *is then called a generator of the copula.*

From the definition it is evident that **for the selection of Archimedean copula, it is sufficient to identified its generator**. In what follows we present the three most elementary as well as most popular Archimedean copula functions - the concepts of **Gumbel copula, Clayton copula and Frank copula**. Here we only stay with theoretical background and in the Box 1 bellow present generators for all the mentioned copula functions, the range of used parameters and, of course, the resulting copula function in a bivariate form.

Box 1: Archimedean copula functions

Name	*Parameter*	*Generator (p)*	*Bivariate copula function*
Gumbel	$\theta \in [1,\infty)$	$(-\ln p)^\theta$	$exp\left(-\left[(-\ln u_1)^\theta + (-\ln u_2)^\theta\right]^{1/\theta}\right)$
Clayton	$\theta \in (0,\infty)$	$\frac{(p^{-\theta} - 1)}{\theta}$	$max\left[\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}, 0\right]$
Frank	$\theta \in (-\infty,\infty) \setminus 0$	$-\ln\frac{e^{-\theta p} - 1}{e^{-p} - 1}$	$-\frac{1}{\theta}\ln\left(1 + \frac{\left(e^{-\theta u_1} - 1\right)\left(e^{-\theta u_2} - 1\right)}{e^{-\theta} - 1}\right)$

---

[14] Hence all spherical distributions are indeed the elliptical distributions.

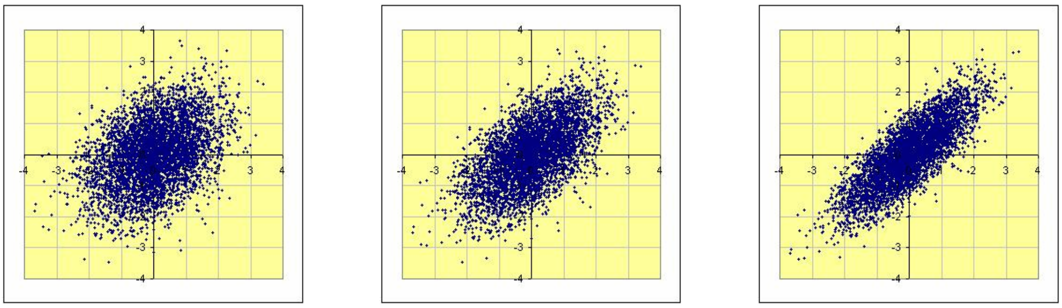
2.2.3 GRAPHICAL DEMONSTRATION

We conclude the whole chapter by **graphical examples of joint distributions of random variables generated by bivariate functional forms of just introduced copula functions**. We assume normally N(0,1) distributed margins for underlying risk factors X_1, X_2 and for each copula function present three graphs which differ in linear correlation between X_1 and X_2. Always, first graph correspond to $\rho = 0.4$, the second to $\rho = 0.6$ and the third to $\rho = 0.8$. Each graph depicts 5000 independent scenarios. In addition to this, on the enclosed CD we give all files underlying these graphs. Files are in MS Excel format and allow to change copula parameters in any way. Hence any form of given copula might be created. Lastly in Appendix A we give algorithms according to each of presented copulas might be generated.

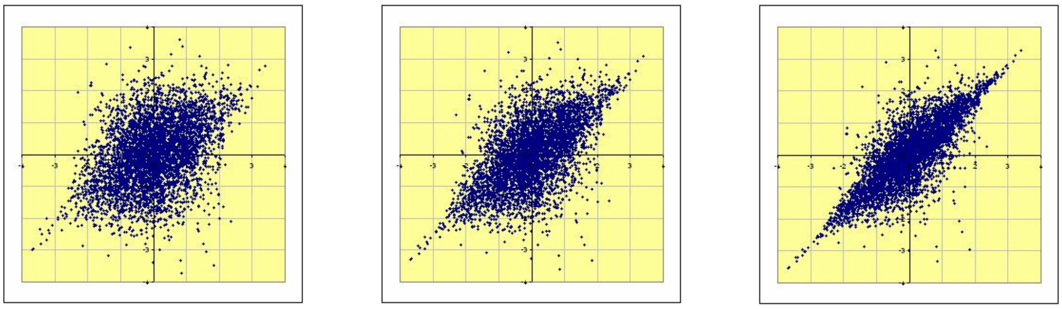
From the graphs we might clearly see that the number of joint events expectedly rises with higher correlation. Moreover we might notice that Gumbel copula exhibits relatively more dependencies in the first quadrant while Clayton copula shows the same in third quadrant and t-copula in both of them. On the other hand we do not observe any of such properties in cases of Gaussian or Frank copulas. We give the proper definition and measure of so called **tailed dependence** in the chapter 2.4.2.

Figure 3: Examples of copula functions[15]

Gaussian copula function



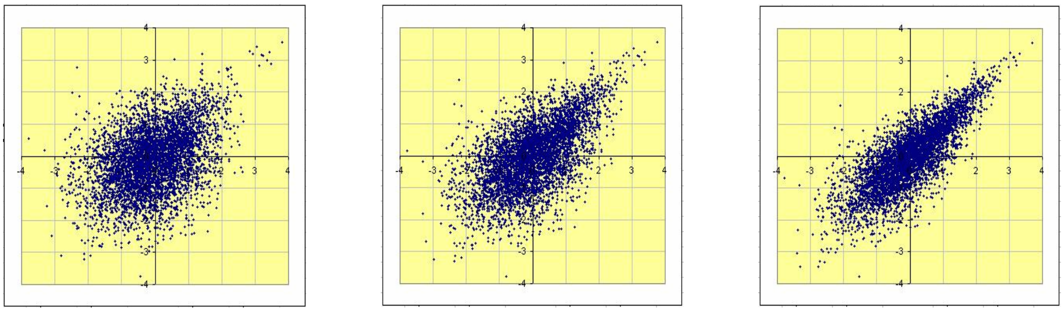
T-copula function



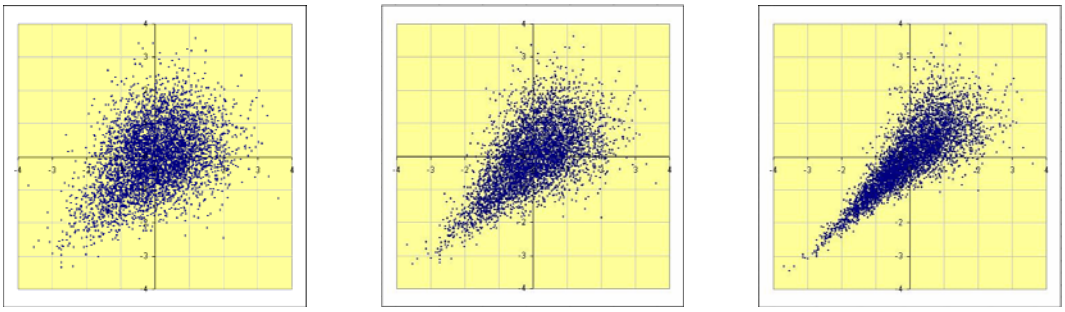
---

[15] x-axis: value for underlying risk factor X_1 ; y-axis: value for underlying risk factor X_2

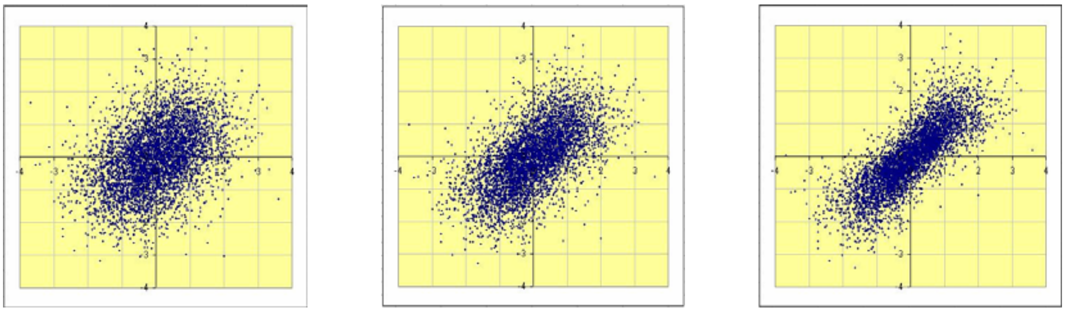
Gumbel copula function



Clayton copula function



Frank copula function



2.3 CONCEPT OF LINEAR CORRELATION

The study of heavy tailed credit risk factors is closely related to the problem of modelling and measuring dependence structures. We are briefly introducing the most common measure of dependence between random variables - the concept of linear correlation. Later we use the concept of linear correlation as a benchmark for comparative studies of copula functions. Now we recall some of the linear correlation basic properties and discuss the shortcomings which this dependence concept brings. We mention the main arguments from Embrechts et al. (1999).

Definition 2 (Linear correlation)

Let X and Y be two random variables with finite nonzero variances Var(X) and Var(Y), the linear correlation between X and Y is defined as:

$$\boxed{\rho(X,Y) = \frac{cov(X,Y)}{\sqrt{Var(X)Var(Y)}}} \tag{2.7}$$

where cov(X,Y) is the covariance between X and Y, given as Cov(X,Y) = E[XY] - E[X]E[Y]; and Var(X) and Var(Y) denotes variances of X and Y.

We recall several properties of linear correlation:

(i) In the case of perfect linear dependence $Y = aX + b$ (where $a, b \in R$ and $a \neq 0$) we get $|\rho(X,Y)| = 1$. More importantly, the statement is equivalence.

This might be shown by considering the following representation:

$$\rho(X,Y)^2 = \frac{Var(Y) - min\,E\left[(Y - (aX + b))^2\right]}{Var(Y)} \tag{2.8}$$

The above equation describes a connection between correlation and simple linear regression. Hence the linear correlation might be interpreted as the relative reduction in the variance of variable Y due to linear regression on X.

(ii) In case of independent random variables, we have $\rho(X,Y) = 0$. However this statement is not anymore equivalence.

(iii) It holds that $\boxed{\rho(aX + b, cY + d) = sign(ac)\rho(X,Y)}$. Thus linear correlation is invariant under strictly increasing linear transformations.

(iv) Linear correlation is easily manipulated under linear operations. Under affine linear transformations $A : R^n \to R^m, x \mapsto Ax + a$ and $B : R^n \to R^m, x \mapsto Bx + b$ where $A, B \in R^{mxn}, a, b \in R^m$ we have: $Cov[AX + a, BY + b] = ACov[X,Y]B^T$

(v) On the base of introduced relationships it might be finally showed that variance of linear combination of random variables is fully determined by their pairwise covariance. In the mathematical notation we get[16]:

for $\alpha \in R^n$, $\sigma^2(\alpha^T X) = \alpha^T Cov(X)\alpha$

---

[16] To see the exact analytic derivations and more about the properties of linear correlation, see Embrechts et al. (1999)

In practice the concept of linear correlation is the canonical measure of dependence in the world of multivariate normal distributions and the natural measure of dependence within the class of elliptical distributions which to a large extent resemble multivariate normal case[17]. However for purposes of **heavy-tailed risk factors modelling the concept of linear correlation defined as in (2.7) is not the appropriate one.** We present the couple of reasons why is it like that.

(i) **Linear correlation is not defined when the variances of X and Y are not finite**. This can be the case of t-distribution with small degrees of freedoms which we will use latter in the modelling section.

(ii) **Independence of X and Y always implies zero linear correlation, but this relationship does not necessarily hold other way round**. In fact it holds in cases of multivariate normal distribution but is no longer valid when only marginal distributions are normal while joint distribution is non-normal or in other general cases. This again will be the case of our model.

(iii) While the concept of linear correlation is invariant under strictly increasing linear transformations, it is **no longer invariant under non-linear strictly increasing transformations**. Hence in general case for $T : R \to R \quad \rho(T(X),T(Y)) \neq \rho(X,Y)$. We might be liking some more general invariance, for example invariance under logarithmic transformations which are so popular within the stock market models.

(iv) **The estimator (2.7) lack robustness in cases when marginal distributions are asymmetric or exhibits heavy tails**.

Summing up, the concept of linear correlation, defined as in (2.7), is undoubtedly well-established, simple and straightforward tool for analysing the dependence structures. It is even the best possible measure of dependence given the assumption that examined random variables are multivariate normally distributed. Further, correlation as well serves as one of the parameters determining multivariate elliptically distributed functions (e.g. t-distribution).

However the quality of estimator (2.7) crucially depends on the marginal distributions of given random variables (see points (i) and (iv) from above). Hence it might pay off to search for models that are able to capture dependence structure between random variables in more general cases, i.e. without restricting assumptions on the shape of modelled distributions.

---

[17] For proper definition of elliptical distribution see chapter 2.2

2.4 ALTERNATIVE DEPENDENCE MEASURES

On several places of previous chapters we stated that as far as we move away from multivariate normal variables and elliptical distributions, the concept of linear correlation stops to be a proper method for dependence structure analyses. On this place we introduce the concept of **Kendall's tau which alternates the linear correlation as a measure of dependence**. Besides that in a mutual comparison to linear correlation, the Kendall's tau shows more generality and can be without any restrictions used not only within the classes of multivariate normal variables but as well within many other classes. Simultaneously we introduce **measures of tail dependence as a tool for analysing extreme events** of bivariate distributions, which are considering this paper of a special interest.

2.4.1 KENDALL'S TAU

Definition 3 (Kendall's tau)

Let (X,Y) *and* (X',Y') *be bivariate random vectors with continuous and common margins F (of X and X') and G (of Y and Y'). We say that* (X,Y) *and* (X',Y') *are concordant if and only if* $(X-X')(Y-Y')>0$ *and discordant if and only if* $(X-X')(Y-Y')<0$.

As ***Kendall's tau*** *we understand the difference between the probabilities of concordance and discordance for vectors* (X,Y) *and* (X',Y').

$$\tau = P((X-X')(Y-Y')>0)-P((X-X')(Y-Y')<0) \tag{2.9}$$

Kendall's tau is known as a **measure of concordance** for bivariate random vectors. It can be further shown[18] that **probabilities of concordance and discordance** (2.9) **might be evaluated by integrating over the distribution of** (X,Y)**.** This transition enables us to rewrite the definition of Kendall's tau into notation of the following integral.

$$\tau = 4\int_0^1\int_0^1 C(u,v)\,dC(u,v)-1 \tag{2.10}$$

where C(u,v) stands for copula associated to (X,Y). Equation (2.10) represents a remarkable transformation as it shows that **Kendall's tau is of copula property**. This for instance

---

[18] Nelsen (1999), p. 127 or Embrechts et al. (2001)

means that, on the contrary to linear correlation, Kendall’s tau is invariant under any strictly increasing and continuous transformations of the margins.

We would like to stress that Kendall’s tau, no matter how often being used, is not the only measure of concordance. Very popular are for instance models of **Spearman’s ρ** or **Schweizer and Wolff’s σ**[19]. It can be shown again that both concepts are of copula property. What have further all three methods in coincidence is that they are all defined for bivariate random variables. However a **dimension of the analysis can be simply enlarged** in a completely same way as for linear correlation, i.e. we can achieve higher dimension analyses just by writing pair-wise correlation into *n x n* matrix.

KENDALL’S TAU AND CORRELATION FOR ELLIPTICAL COPULA FUNCTIONS

From the chapter 2.2 it is evident that distribution of Gaussian as well as t-copula functions are determined by correlation coefficient ρ. In the last chapter we introduced correlation estimator (2.7) and showed its limited robustness. Nevertheless within the models with elliptical distributions with absolutely continuous margins we might construct a **robust linear correlation estimator** using the following relationship between linear correlation and Kendall’s tau[20]:

$$\boxed{\rho(X,Y) = sin\left(\frac{\pi}{2}\,\tau(X,Y)\right)} \tag{2.11}$$

And a consistent Kendall’s tau estimate for a sample of *n* observations of variables defined as above might be obtain by:

$$\hat{\tau}(X,Y) = \binom{n}{2}^{-1} \sum_{k=1}^{n} \sum_{i}^{n} sign\left[\left(X^{(k)} - X^{(i)}\right)\left(Y^{(k)} - Y^{(i)}\right)\right] \tag{2.12}$$

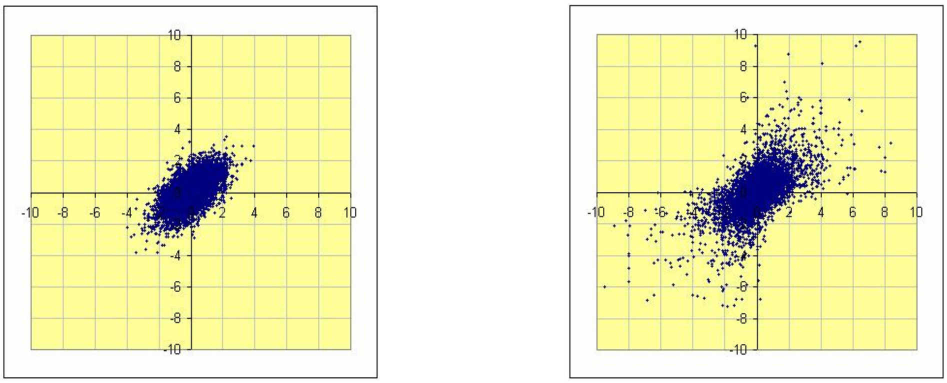
In following Figure 4 we present bivariate mixtures of normal and t-3 distributions which were all built up upon the same value of copula linear correlation parameter 0.6. For each figure we calculated correlation by means of estimator (2.7) and (2.11) – correlation ρ_1 and ρ_2. Expectedly both estimators give almost the same results for bivariate Gaussian distribution (a) and the most distinct results for Gaussian copula with fat tail margins (b). Further we might notice that figures (c) and (d) have evidently more dependence in tails than figures (a) and (b). However none of the employed estimators is able to capture this difference.

---

[19] J.J.Quasada-Molina, What are copulas, 2003, p. 4

[20] Proper derivation of relationship between Kendall’s tau and linear correlation coefficient and further description of this approach might be found for instance in Lindskog et al. (2002).

Figure 4: Copula mixtures of normal and t-distributions[21]

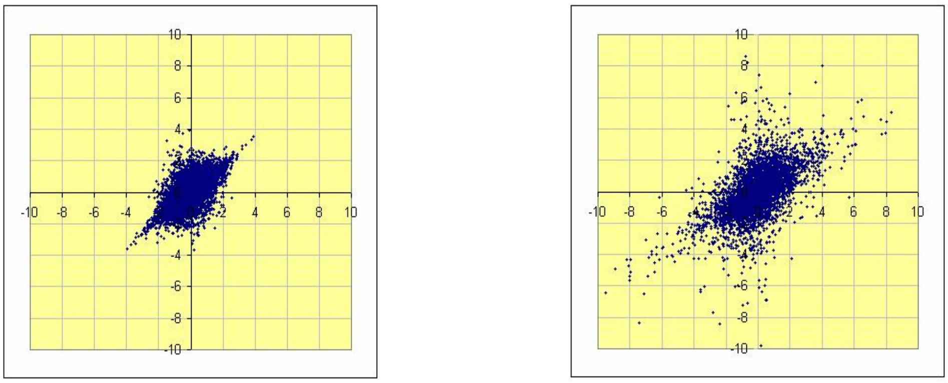


(a) Bivariate Gaussian distribution

ρ_1 = 0.6018 ρ_2 = 0.6013

(b) Gaussian copula with t-3 margins

ρ_1 = 0.5440 ρ_2 = 0.6013



(c) t-3 copula with Gaussian margins

ρ_1 = 0.6059 ρ_2 = 0.6140

(d) Bivariate t-3 distribution

ρ_1 = 0.6102 ρ_2 = 0.6140

2.4.2 TAIL DEPENDENCE

The concept of tail dependence has a close relation to the problems of extreme joint events since it provides an **asymptotic measure of dependencies in upper- and lower- quadrant tail of bivariate distributions**. In another words the model assumes two random variables and is based on measuring probability of extreme realisation of the first variable given the assumption that extreme event of the second variable happened. According to the interests we distinguish between coefficients of upper and coefficients of lower tail dependence.

---

[21] x-axis: value for underlying risk factor X_1 ; y-axis: value for underlying risk factor X_2 ; Each graph contains 5000 independent scenarios.

Definition 4 (Coefficient of tail dependence)

*Let X and Y be two continuous random variables with distribution functions F, G and copula C. A **coefficient of lower tail dependence** of X and Y is then defined as:*

$$\boxed{\lambda_L = \lim_{u \to 0} P\left[Y < G^{-1}(u) \middle| X < F^{-1}(u) \right]} \tag{2.13}$$

if the limit exists. Further if $\lambda_L \in (0,1]$ we say that X and Y are asymptotically dependent in the lower tail and similarly if $\lambda_L = 0$ we say that X and Y are asymptotically independent in the lower tail.

*A **coefficient of upper tail dependence** is likewise defined for the upper right quadrant:*

$$\boxed{\lambda_U = \lim_{u \to 1} P\left[Y > G^{-1}(u) \middle| X > F^{-1}(u) \right]} \tag{2.14}$$

if such a limit exists.

As with the measures of concordance it is again possible to show that coefficients of lower and upper tail dependence are both of copula property and again depend only on copula C of variables X and Y. The **copula-based variation of lower / upper coefficient of tail dependence** then have following forms:

$$\lambda_L = \lim_{u \to 0} \frac{C(u,u)}{u} \qquad \text{and} \qquad \lambda_U = \lim_{u \to 1} \frac{C(u,u) - 2u + 1}{1 - u} \tag{2.15, 2.16}$$

It is good to realize that **for elliptically symmetric copula functions**, which is for instance the case of Gaussian copula or t-copula, **the lower and upper coefficients of tail dependence concur** and consequently are simply denoted by λ. Further we might prove that for Gaussian copula we always get zero tail dependence and that for t-copula the coefficient of tail dependence equals the following expression[22].

$$\lambda = 2t_{\nu+1}\left(-\sqrt{\nu+1}\frac{\sqrt{1-\rho}}{\sqrt{1+\rho}} \right) \tag{2.17}$$

It is obvious that t-copula's coefficient of tail dependence λ is increasing in ρ and decreasing in ν. Further as ν rises, the t-copula converges to the Gaussian copula and λ simultaneously converges to zero. We recall the Figure 4 in chapter 2.4.1 and adds some λ calculations for given values of ρ and ν. The situation where ν equals to infinity corresponds to the case where t-copula perfectly converges to Gaussian copula. On the other hand we might see that

---

[22] For the particular proofs, see Nyfeler (2000), p.45

apart of cases with high positive margins' correlation we get a good approximation of Gaussian copula's tails even for t-copula with only about twenty degrees of freedom. We might as well see that similarly to univariate t-distribution, the higher are the t-copula function's degree of freedom, the more it resembles the Gaussian distribution. And if a t-copula has more than 30 degrees of freedom, it is supposed to be in well coincidence with the Gaussian case.

Table 1: Coefficient of tail dependence λ for t-copula

$\nu \backslash \rho$	**-0.5**	**0**	**0.5**	**0.9**	**1**
2	0.06	0.18	0.39	0.72	1
5	0.01	0.05	0.21	0.59	1
10	0.00	0.01	0.08	0.46	1
20	0.00	0.00	0.02	0.31	1
∞	0.00	0.00	0.00	0.00	1

2.5 TEST STATISTICS FOR THE COPULA FITNESS

Great property of copula based models is that different copula functional forms could be used in order to simulate any dependence structure between random variables that scientists have ever thought about. On the other hand the large variety of plausible functional forms brings the major drawback of copula's wider usage in practice as there usually exists **no clear guidance of what parametric copula form to use**. Moreover there is just one consensus about the selection of the right functional form saying that this process belongs to the most undiscovered parts of copula modelling. The fact whether a chosen copula family appropriately models the dependence structure of an observed data set therefore constitutes a fundamental problem.

When the data set is to be analysed, the problem to be solved generally consists of three parts. First using the experience from the similar type models, we suggest several copula functional forms which are believed to fit the given data. As a second step we estimate the parameters of all proposed functions. Lastly, using goodness-of-fit tests, we analyse which of the particular copulas seems to be most appropriate. If the analysis is done properly, the final output is a conveniently fitted copula function which by means of Monte-Carlo simulation further serves for detail investigation of likely behaviour of examined variables.

We present **two methods for copula parameters estimation**. According to the treatment with copula margins we differ on so called parametric and semi-parametric models. Parametric models specify and estimate parametric forms of margins as well as of copula function. On the contrary the semi-parametric models put no assumption on margins' parameters and substitute marginal distributions by the univariate empirical cumulative distribution functions. We describe both technique in detail in the next two sections and then close the whole chapter with introduction to several goodness-of-fit tests in subchapter 2.6.

2.5.1 PARAMETRIC METHOD OF COPULA ESTIMATION

Parametric copula estimations are accomplished by the method of maximum likelihood. The method assumes that **empirical observations are independent and identically distributed** and that **particular functional forms for copula and margins are somehow specified.** In the procedure, whose description just follows, we principally try to find a set of parameters for the given copula and marginal distributions which would maximize the likelihood of occurrence of observed realisation of examined random variables.

Suppose we have T i.i.d. empirical data observations and *C* is an n-copula with unknown parameter α and continuous margins $F_1,\dots,F_n$ with unknown parameters $\nu_1,\dots,\nu_n$. Then from the Sklar's theorem we have an n-dimensional distribution function *H* such that:

$$H(x_1,\dots,x_n)=C(F_1(x_{1t},\nu_1),\dots,F_n(x_{nt},\nu_n);\alpha).$$

We can derive a corresponding **density function** which then looks like as follows:

$$h(x_1,\dots,x_n)=c(F_1(x_{1t},\nu_1),\dots,F_n(x_{nt},\nu_n);\alpha)\prod_{i=1}^{n} f_i(x_i,\nu_i),$$

where $f_i(x_i,\nu_i)$ stands for the density of the margins and $c(F_1(x_{1t},\nu_1),\dots,F_n(x_{nt},\nu_n);\alpha)$ represents the density of copula *C* and is defined as: $c(u_1,\dots,u_n)=\dfrac{\partial^n C(u_1,\dots,u_n)}{\partial u_1\cdot\dots\cdot\partial u_n}$.

The desired parameters ν and α are finally simultaneously estimated by maximising the density function *h*. The **maximisation procedure is however easier performed under logarithmic structures** and therefore we first carry out a logarithmic transformation[23] of density function *h*. We get a **log-likelihood function *l*(ν, α)** and maximise that.

$$\boxed{l(\nu,\alpha)=\sum_{t=1}^{T}\ln c(F_1(x_{1t},\nu_1),\dots,F_n(x_{nt},\nu_n);\alpha)+\sum_{t=1}^{T}\sum_{i=1}^{n}\ln f_i(x_{it},\nu_i)} \qquad (2.18)$$

---

[23] This is an **increasing continuous transformation** and therefore has no impact on the results.

Joint estimation of margins' and copula parameters is, however, computationally intensive and in a reaction Joe and Xu proposed in 1996 to use the method of inference function for margins (IFM). The base of this approach is to divide maximisation procedure into two computationally simpler parts and estimate parameters for margins and for copula separately. In this case we first evaluate margins' parameters ν by maximising the expression:

$$l(\nu) = \sum_{t=1}^{T} \ln f_i(x_{it}, \nu_i) \tag{2.19}$$

Lastly, we use the results from (2.19) to estimate the copula parameters α:

$$l(\alpha) = \sum_{t=1}^{T} \ln c(F_1(x_{1t}, \nu_1), \ldots, F_n(x_{nt}, \nu_n); \alpha) \tag{2.20}$$

2.5.2 SEMI-PARAMETRIC METHOD OF COPULA ESTIMATION

Semi-parametric method has a close relation to the method of maximum likelihood and therefore is sometimes as well called as a **method of canonical maximum likelihood**. Unlike the approach from the last chapter 2.5.1, the **CML method is based on the estimation of copula parameters without specifying the parametric form for margins**. The method relies on the fact that if $X_1, \ldots, X_n$ are random variables with continuous cumulative distribution functions $F_1, \ldots, F_n$, then variables $u_1, \ldots, u_n$ defined as $u_i=F_i(x_i)$ have each a uniform marginal distribution defined on interval [0,1].

The problem under consideration is how to work with margins $F_1, \ldots, F_n$ when we should not introduce any distributional assumptions for them. The solution comes with incorporating the **empirical cumulative distribution function** defined for T observations x_{Ti} of variables $X_1, \ldots, X_n$ as:

$$F_E(x) = \frac{1}{T+1} \sum_{j=1}^{T} I(x_{ji} \le x) \tag{2.21}$$

where i=1,…,n and ***I(...)*** is the **indicator function** defined as 1 if expression in the parenthesis holds and 0 otherwise. Further as number of observations T converges to infinity, the empirical distribution function converges to the actual one. Therefore in case of large data samples we can substitute $u_i = F_i(x_i)$ from the first paragraph by $u_i = F_{Ei}(x_i)$.

At the end are the transformed variables used as parameters of likelihood function $l(\alpha)$, which is then maximized and gives a wanted copula parameter α.

$$l(\alpha) = \sum_{t=1}^{T} \ln c(u_1,...,u_n;\alpha). \qquad (2.22)$$

The last question to be asked is which of the two introduced models for copula estimation is more appropriate. The cardinal difference between both models is in the specification of marginal distributions $F_1,\ldots,F_n$. It is actually this fact that makes the parametric method of copula estimation vulnerable as a wrong assumption about the margins' parametric distribution functions causes that copula parameters estimates are biased. Therefore **if we are not completely sure about the marginal distribution functions the semi-parametric method should be preferred**.

2.6 GOODNESS-OF-FIT TEST

In the previous sections we introduced two methods how to estimate parameters for given type of copula functional form. Having determined the parameters, one might naturally arise a question how good such resulting model is, i.e. how much appropriately an assumed family of copula could model the dependency structure of an observed data set.

Nevertheless taken from the practical point of view, in order to fit the given data, first usually several different plausible copula functional forms are proposed and their parameters are estimated. The question then is no more how good the individual models are, but which model outperforms the others and could be most likely relied on. Goodness-of-fit tests comprises a method which should provide a direction for answering these questions.

2.6.1 AKAIKE'S INFORMATION CRITERION

The first, straightforward, computationally easy, but unfortunately only very rough guideline on copula selection might be obtained by **Akaike's Information Criterion**. The criterion is based on the maximum likelihood estimates (recall the part 2.5.1) and is defined as:

$$\boldsymbol{AIC = -2L(\nu, \alpha, x)+2q} \qquad (2.23)$$

where q stands for the number of parameters of the copula family being fitted.

The simple rule says that we **should use such copula that returns the smallest value of AIC**. Evidently the higher the value of likelihood function and the lesser the number of copula parameters, the better. In another words the Akaike's model measures the pay-off between the quality of copula models and their complexity. The point to be stressed is the fact that there is no hypothesis to be tested and that we work only with the selection criterion. This represents the crucially weak point of Akaike's method as it **fails to provide any information about the size or power of decision rule** that is employed. Hence if we are about to choose the copula more consistently, we need to look for some other methods.

2.6.2 UNIVARIATE MODELS

When we are to decide about appropriateness of an univariate model, we can simply rely on several well established **goodness-of-fits statistics for univariate distribution**. These models are built on the assumption that if a hypothesised model is constructed correctly, then the empirical cumulative distribution F_E would almost surely converge to the hypothesised cumulative distribution F_H[24]. The quality of modelled distribution F_H can therefore be tested by measuring its distance deviations from empirical cumulative distribution F_E.

In the literature we can usually find **two test statistics - Kolmogorov-Smirnov's and Anderson-Darling's**. From a certain point of view, the two tests mutual supplement each other as Kolmogorov-Smirnov's exhibits higher sensitivity to deviations in the centre of distribution while Anderson-Darling's approach puts more weight on the deviation in the tails of the distribution. However as our interests are in copula heavy tailed risk factors modelling, usually the Aderson-Darling's test is applied. In what follows, we give four test statistics from Kole (2006). Let x_t be a realization of the random variable X out of sample T realization, then:

- Kolmogorov-Smirnov $D_{KS}^{m} = \max_{t} \left| F_E(x_t) - F_H(x_t) \right|$
- Average Kolmogorov-Smirnov $D_{KS}^{a} = \int_x \left| F_E(x_t) - F_H(x_t) \right| dF_H(x)$
- Anderson-Darling $D_{AD}^{m} = \max \frac{\left| F_E(x_t) - F_H(x_t) \right|}{\sqrt{F_H(x_t) \cdot (1 - F_H(x_t))}}$
- Average Anderson-Darling $D_{AD}^{a} = \int_x \frac{\left| F_E(x_t) - F_H(x_t) \right|}{\sqrt{F_H(x_t) \cdot (1 - F_H(x_t))}} dF_H(x)$

---

[24] E.Kole et al., Selecting copulas for risk management, 2006, p.6

2.6.3 MULTIVARIATE MODELS

Given a multivariate case we are trying to find and impose some transformation which would **project the multivariate decision problem into univariate scales**. If we manage to find such transformation, we can then apply any method from the last chapter of univariate models. Breymann et al. (2003) suggest a technique which is found as computationally very efficient but the way they project from multivariate to univariate problem makes the results inconsistent, i.e. meaning that the resulting test statistics is not strictly increasing for every deviation from the null hypothesis. In this respect, Berg and Bakken (2005) deepen the original Breymann's idea and come with numerically efficient and unbiased goodness-of-fit test which moreover already proves consistency.

Before we describe the Berg's and Bakken's test in detail, we look at a technique of **probability integral transform**[25] which represents the important part of Berg's and Bakken's method as well as the part of original model by Breymann. Given the right multivariate distribution, the method of PIT serves as a tool for transmuting a set of observed dependent variables into a new set of independent variables with uniform U(0,1) distribution.

Hence we could take a hypothesised multivariate distribution represented by copula function and carry out the PIT method. **If the hypothesised distribution fits the observed data set appropriately**[26]**, the PIT should give us a set of i.i.d. U(0,1) variables**. The task for goodness-of-fit tests is then to find out whether the transformed variables are really i.i.d. U(0,1) or not. The technique of PIT works as follow, Berg (2005):

Let $X=(X_1,\dots,X_n)$ *be a random vector with marginal distribution functions* $F_i(x_i)=P(X_i\le x_i)$ *and conditional distributions* $F(X_i\le x_i|X_1=x_1,\dots,X_{i-1}=x_{i-1})$ *for i=1,…,n. As a* ***probability integral transform of X*** *we understand the vector* $T(X)=(T_1(X_1),\dots,T_d(X_d))$*, where* $T_i(X_i)$ *is defined as:*

$$\boxed{\begin{aligned}
&T_1(X_1)=P(X_1\le x_1)=F_{X_1}(x_1)\\
&T_2(X_2)=P(X_2\le x_2|X_1=x_1)=F_{X_2|X_1}(x_2|x_1)\\
&\vdots\\
&T_n(X_n)=P(X_n\le x_n|X_1=x_1,\dots,X_{n-1}=x_{n-1})=F_{X_n|X_1\dots X_{n-1}}(x_n|x_1,\dots,x_{n-1})
\end{aligned}}$$

---

[25] Further in the text we call the **Probability Integral Transform** as PIT.

[26] This case we call a **zero hypothesis**. An **alternative hypothesis** is that chosen copula does not fit the given data.

The such obtained random variables $Z_i = T_i(X_i)$, *for i=1,…,d are then uniformly and independently distributed on* $[0,1]^n$.

The original Breymann's model then **decreases dimensions** of the analysis and gets the multivariate problem into univariate scales by computing $Y = \sum_{i=1}^{n} \Phi^{-1}(Z_i)^2$. It is right this projection that makes the test inconsistent[27]. In order to avoid this problem, Berg and Bakken propose firstly to transform the variables Z_i into Z_i^* and afterwards to apply the same projection method as in Breymann, i.e. we have $Y = \sum_{i=1}^{n} \Phi^{-1}\left(Z_i^*\right)^2$. Z_i^* is defined as:

$$\boxed{Z_i^* = P\left(r_i \le \tilde{Z}_i \middle| r_1,\dots,r_{i-1}\right) = \left(1 - \left(\frac{1-\tilde{Z}_i}{1-r_{i-1}}\right)^{n-(i-1)}\right)} \tag{2.24}$$

for $i = 1,..,n$; where $\tilde{Z} = \left(\tilde{Z}_1,\dots,\tilde{Z}_n\right)$ is the sorted version of Z and r_i is rank variable i from Z[28].

Berg and Bakken claim that *if we find the probability, under H_0, that the variable with rank i, given the variables with rank 1,…, i-1, will be smaller or equal to the observed variable with rank i, $\tilde{Z}_i$, then we can project the multivariate problem to a univariate case consistently. In a mathematical notation, we can rewrite this idea as a following probability:*

$P\left(r_i \le \tilde{Z}_i \middle| r_1,\dots,r_{i-1}\right) = 1 - P\left(r_i > \tilde{Z}_i \middle| r_1,\dots,r_{i-1}\right)$. *The only way r_i could be greater than $\tilde{Z}_i$ is when all remaining n–(i-1) variables are greater than $\tilde{Z}_i$. Since the remaining n-(i-1) variables are independent, the probability of all being greater than $\tilde{Z}_i$ equals to the product over the probabilities of each r_k, $k \in [0,1]$, being greater than $\tilde{Z}_i$, thus:*

$$P\left(\tilde{Z}_i < r_k < 1 \middle| r_k > r_{i-1}\right) = \frac{P\left(r_k > \tilde{Z}_i \cap r_k > r_{i-1}\right)}{P\left(r_k > r_{i-1}\right)} = \frac{P\left(r_k > \tilde{Z}_i\right)}{P\left(r_k > r_{i-1}\right)} = \frac{1-\tilde{Z}_i}{1-r_{i-1}}, \qquad k \in [i,n] \qquad \blacksquare.$$

Having created the variable $\tilde{Z}_i$, we can introduce the **projection function** Y and transform the multivariate analysis into an univariate problem in a way $Y = \sum_{i=1}^{n} \gamma(Z_i;\alpha) \cdot \Phi^{-1}\left(Z_i^*\right)^2$, where $\gamma(Z_i;\alpha)$ is a **weighting function** with weight parameter α.

---

[27] For more information about this problem see **Breymann et al. (2003)**

[28] i.e. r_i comes from ascendingly ordered Z.

Function $\gamma(Z_i\,;\alpha)$ represents another valuable extension from Breymann's model as it enables us to use **much more general weighting procedure** than just by implicit mean of inverse normal function as in $Y=\sum_{i=1}^{n}\Phi^{-1}(Z_i)^2$ [29]. For instance we might wish to weight more the tail accuracy of modelled copula function. For this sake, Berg and Bakken introduce several examples of such functions:

(i) power tail weighting: $\gamma(Z_i\,;\alpha)=(Z_i-1/2)^{\alpha}$ for α be some positive, integer, even number.

(ii) Left power tail weighting: $\gamma(Z_i\,;\alpha)=1-Z_i^{1/\alpha}$

(iii) Right power tail weighting: $\gamma(Z_i\,;\alpha)=1-(1-Z_i)^{1/\alpha}$

(iv) Inverse student's t tail weighting: $\gamma(Z_i\,;\alpha)=t_{\nu}^{-1}(Z_i)^2$

Finally the Berg's and Bakken's test is defined as: $B(w)=P(F_Y(Y)\leq w)$, where F_Y is a cumulative distribution function of Y and $w\in[0,1]$. We might as well derive empirical version of $B(w)$, then for T observation of n-dimensional vector Z we have:

$$\boxed{\hat{B}(w)=\frac{1}{T+1}\sum_{j=1}^{T}I(F_{Yj}(Y)\leq w),\qquad w=\frac{1}{T+1},\ldots,\frac{T}{T+1}} \tag{2.25}$$

Under the null hypothesis the function $B(w)$ has an uniform distribution and hence $B(w)=w$ and the corresponding density function is $b(w)=1$ for all $w\in[0,1]$. We have obtained all information which we need for performing the univariate analysis. We carry out the Anderson-Darling test and reject or do not reject the null hypothesis about copula function appropriateness.

---

[29] Moreover the **square root** in the function causes that **Breymann's approach** weights larger deviations relatively more than the small ones.

3. MODEL CREDITMETRICS^{TM}

CreditMetrics^{TM} represents one of the most influential financial models of the last decade and is believed to stand at the beginning of quantitative credit risk modelling at a portfolio level. Originally, the model was introduced by JP Morgan in 1997 as a **mark-to-market model[30] which simulates the full distribution of future value for any credit portfolios over a given time horizon[31]**. For purposes of our study we will use the fact that CreditMetrics^{TM} model is the representative of credit risk models using an assumption of multivariate normal distributions for underlying risk factors. Moreover CreditMetrics^{TM} model is very popular in the commercial practice. Primarily for those reasons **we find CreditMetrics^{TM} model as the ideal benchmark** for comparison with our model supposing the heavy tailed risk factors.

The framework of the model is based on evaluation of credit quality of given obligors. An elegant idea standing behind the model is that any change in credit rating of particular credit instrument - bonds, loans, certificates or other financial claims with derivable loss distribution - including up- and downgrades as well as default, has a direct impact on the value of such instrument. Therefore in CreditMetrics^{TM} model the **risk comes not only from default but as well from credit quality changes** and the proper analysis of credit quality migration[32] provides an effective tool for estimating the **portfolio value at risk due to credit** as a final output. In the next three chapters we introduce the rating systems and transition probabilities, valuation mechanism within the CreditMetrics^{TM} model and finally analytic background for simulating loss distribution.

3.1 RATING SYSTEMS AND TRANSITION PROBABILITIES

A rating system[33] with probabilities of migrating from one rating class to another constitutes the key component of the overall model. CreditMetrics^{TM} introduce a flexible approach as it puts no restrictions on the origin of such rating system nor on the number of implied rating classes. Hence risk managers and financial institutions might use any established

---

[30] **Market-to-market models** are models assigning a position held in a financial instrument on the base of the current market price for such instrument.

[31] Even though the **time** can be taken arbitrary, it is often set fixed as a one year.

[32] As a **credit quality migration** or credit rating migration we understand the probability of moving from one credit quality to another, including default.

[33] A **credit rating** assesses the credit worthiness of an individual or corporation. Credit ratings are calculated from financial history and current assets and liabilities. Typically, a credit rating tells a lender or investor the probability of the subject being able to pay back a loan.

commercial system like Standard and Poor's or Moody's or imply their own internal rating systems which are recently so popular in the light of new legislative directives resulting from Basel II Committee. However taken **from the portfolio view it is essential for the credit rating system to be homogeneous over all portfolio segments** otherwise the analysis would not be consistent. In our study we will use solely system and statistics developed by Standard and Poor's with eight rating classes given as AAA, AA, A, BBB, BB, B, CCC and default[34].

In practice, actually, the issue is not about to choose a rating system but to choose such a system which provides full description of **probability of credit quality migration** for any obligor in examined portfolio to any possible states of the world at given risk horizon. Having right data, one could easily calculate the desired probabilities by means of observing the historical patterns. However such a data set should be large enough with both respect to time (at least five years) as well as portfolio volume and thus it is seldom facile to obtain an appropriate data file[35]. As a consequence of data scarcity one might incline to use information published by large credit rating agencies. Nevertheless it is important to bear in mind that such statistics usually represent average data across different industries and over several business cycles and therefore their interpretation should be taken carefully.

If we carry out an embracive analysis and estimate the probability of credit quality change of any rated obligor for all possible circumstances over a given time horizon X, we get so called **X-year credit transition matrix**. These matrices are regularly issued by large rating agencies. In the table bellow we present an example of transition matrix created by S&P.

Table 2: Average One-Year N.R.-Removed[36] Transition Rates, 1981 to 2005 (%)

From/To	AAA	AA	A	BBB	BB	B	CCC	Default
AAA	91.42	7.92	0.51	0.09	0.06	0.00	0.00	0.00
AA	0.61	90.68	7.91	0.61	0.05	0.11	0.02	0.01
A	0.05	1.99	91.43	5.86	0.43	0.16	0.03	0.04
BBB	0.02	0.17	4.08	89.94	4.55	0.79	0.18	0.27
BB	0.04	0.05	0.27	5.79	83.61	8.06	0.99	1.20
B	0.00	0.06	0.22	0.35	6.21	82.49	4.76	5.91
CCC	0.00	0.00	0.32	0.48	1.45	12.63	54.71	30.41

Source: Standard & Poor's Global Fixed Income Research; Standard & Poor's CreditPro® 7.02 .

---

[34] For more information see http://www.standardandpoors.com

[35] For methodology of **modelling the credit migration matrices** see Measurement and Estimation of Credit Migration Matrices by Til Schuermann and Yusuf Jafry, 2003

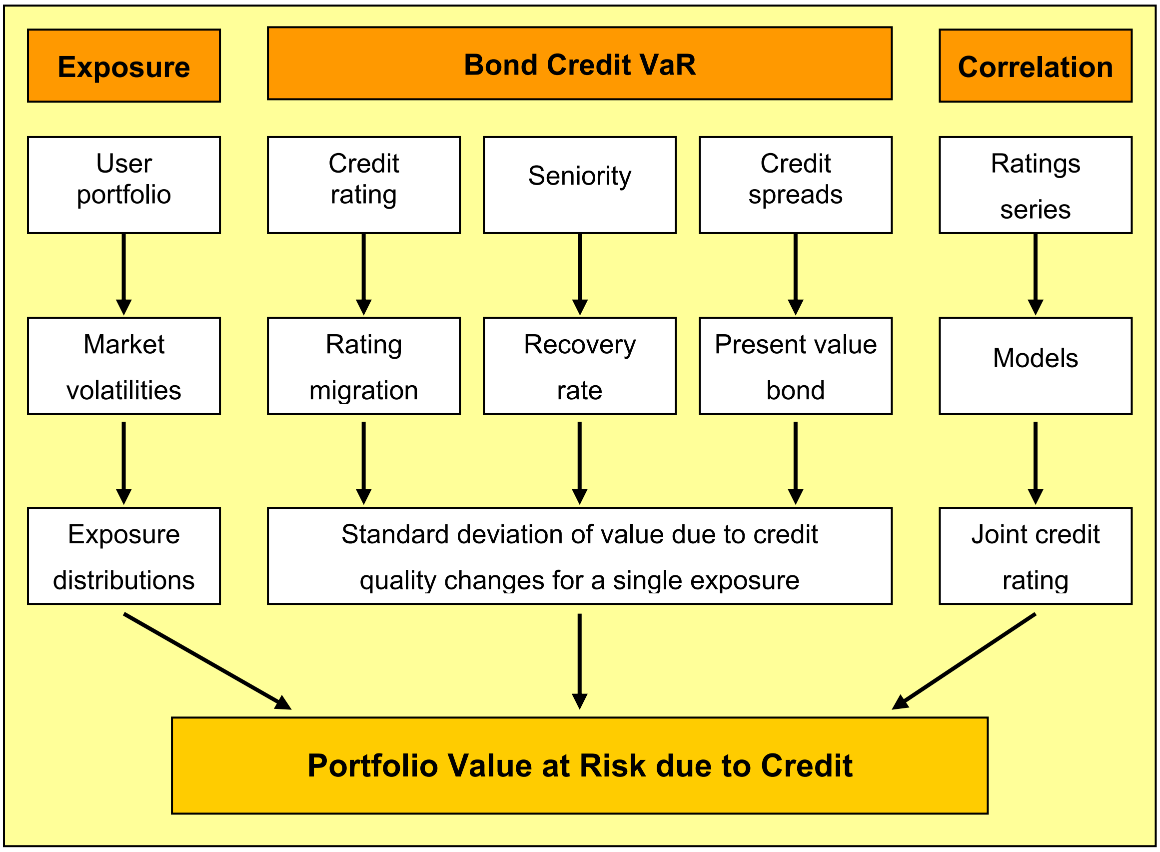
[36] N.R. Removed stands for **Not Rated Removed**, i.e. companies, which are not rated at the end of investigated horizon are removed from the whole analysis

From the transition matrix we might observe that for instance a A-rated obligor has in a one year time 91.43% chance to remain within the A class, 1.99% chance to be upgraded to AA class and 5.86% chance to be downgraded to BBB class of the rating etc. We might as well notice that the biggest chance for all obligors is to remain within the same rating class which they are in at the beginning of the analysis. **Therefore the credit quality changes might be regarded as rather rare events.**

3.2 VALUATION IN THE SENSE OF CREDITMETRICS^{TM}

The complete **process of credit risk evaluation in sense of CreditMetrics^{TM} model** is depicted in Figure 5. We see that portfolio Value at Risk due to Credit is determined by three factors - exposures the portfolio has to face, credit VaR of every single instrument that portfolio has and finally by correlation between such instruments.

Figure 5: Scheme of CreditMetrics^{TM}



We present the **valuation method on the example of bond portfolio**. The reason for such choice is very clear – it is just straightforward and simple. However we stress the fact that

CreditMetrics^{TM} model is not restricted only to the valuation of bond instruments and that very similar analysis might be performed on any credit instrument which has a derivable loss distribution. We present the valuation process in two steps - valuation of single bond and valuation of overall bond portfolio where mutual dependence structures need to be taken into account.

3.2.1 VALUATION OF A SINGLE BOND

The basic thought of valuation procedure is that any examined rated bond can during time of analysis change its credit quality to any state of the world with probability given from transition matrix (see chapter 3.1). The theory then says that **the bond value is given by its value at all possible credit quality migration states weighted by the probability of migrating to such state**. Hence in order to evaluate the bond, one might proceed like that:

1. First, the credit rating system, the credit rating of particular obligor and transition matrix needs to be selected and determined.
2. A time horizon, within which the analysis is considered, is set up. This is often taken almost automatically as a one year.
3. The bond is evaluated at all possible credit quality states. In case of S&P system of rating this includes seven rating classes states and one case for default.
4. The bond value is given as a weighted average of value at all states with weights given from transition matrix.
5. Finally the Value at Risk due to credit is calculated.

The most complex part is hidden behind point 3 (evaluation) and we will discuss it now in detail. Generally the evaluation procedure might be separated into one case representing the default and another case representing credit quality upgrades and downgrades.

(i) If the obligor falls into **default** we are interested in amount of our claims which could be recovered through bankruptcy proceedings. Original CreditMetrics^{TM} model point out that under bankruptcy proceedings the creditors are settled according to the seniority class of their claims. Hence obviously it should be the factor of seniority that determines the value of debt in default. However quite often we come to the recent studies which simplify the analysis by incorporating **beta distribution**[37] **with 50% mean and 20% variance** for all the cases[38]. The residual value is than given as in the following formula:

$$\textbf{Residual Value (Debt Value in default)} = f(x;\alpha,\beta) \textbf{ x Face Value}$$

(ii) In order to obtain **value at risk horizon corresponding to rating up- and down-grades**, we need to perform straightforward present value bond revaluation, CMDT (1997). The question are the discount rates. Theory says that while in general case for loans the discount rates for each rating category and maturity would be given by credit spreads, **in case of bonds the rates are given by forward zero curves**. Bonds corresponding to different rating category are discounted by rates from different forward zero curves, respecting the expected assets riskiness. The discount rates which we employ in calculations throughout the thesis are presented in following Table 3.

Table 3: Forward zero curves[39]

Rating	Year 1	Year 2	Year 3	Year 4
AAA	3,01%	3,27%	3,46%	3,56%
AA	3,02%	3,28%	3,46%	3,57%
A	3,03%	3,30%	3,49%	3,61%
BBB	3,16%	3,49%	3,71%	3,86%
BB	3,67%	4,31%	4,74%	4,96%
B	6,33%	7,18%	7,46%	7,45%
CCC	22,89%	16,32%	13,52%	11,96%

---

[37] **Beta distribution** is defined with probability density function $f(x,\alpha,\beta)=\frac{x^{\alpha-1}\cdot(1-x)^{\beta-1}}{\int_0^1 u^{\alpha-1}\cdot(1-u)^{\beta-1}du}$ and has mean $E(X)=\frac{\alpha}{\alpha+\beta}$ and variance $Var(X)=\frac{\alpha\cdot\beta}{(\alpha+\beta)^2\cdot(\alpha+\beta+1)}$.

[38] **CreditMetrics™ Technical Document** (1997) recognises **5 seniority classes** each determining different mean recovery rate and standard deviation. Nevertheless senior secured and unsecured assets are calibrated to recovery rates with means of 53.8% and 51.13% and variances of 26.86% and 25.45%. Hence our approximation with beta distribution, supposing some better quality assets, is justifiable.

[39] Presented **forward zero curves** were derived specially for Czech environment given transition matrix in Table 2 and information about forward interest rates of Czech national bank (www.cnb.cz) and selected major Czech commercial banks. The proper technique and calculations are described on enclosed CD in file Forward_zero_curves.xls

Given the discount rates, the general **formula for bond revaluation in a one year time** then looks as follows. *C* stands for coupon, *Face* for bond face value and f_i for discount rates determined by forward zero curves from Table 3.

$$V = C + \frac{C}{1+f_1} + \frac{C}{(1+f_2)^2} + \frac{C}{(1+f_3)^3} + \ldots\ldots + \frac{Face + C}{(1+f_n)^n} \tag{3.1}$$

We introduced all necessary requisites for bond evaluation. The calculating procedure of the whole model is presented in the next section, where **volatility of value due to credit quality changes** and **credit Value at Risk** are both given as a final output of CreditMetrics^{TM} risk measurement.

3.2.2 EXAMPLE OF BOND VALUATION

We will follow the valuation procedure which was introduced in this chapter. We assume a BBB-rated bond with 5 years to maturity - a classical example from CreditMetrics^{TM} Technical Document. In a one year time the bond might change its credit quality to any state of the world for which we carry out the bond revaluation as in the chapter 3.2.1. Next, corresponding probabilities of migration might be extracted from transition matrix. The sum of probability weighted values at all the states finally gives the mean (i.e. expected) value of the bond. At the end the variance and standard deviation are calculated. Resulting numbers are presented in Table 4.

Table 4: Evaluation of single bond

State in one year	Probability at the state	Value at the state	Probability weighted value	Difference from mean	Difference square	Prob. weighted difference square
AAA	0.02%	110.35	0.02	1.55	2.42	0.00
AA	0.17%	110.31	0.19	1.51	2.28	0.00
A	4.08%	110.18	4.50	1.38	1.91	0.08
BBB	89.94%	109.24	98.25	0.44	0.20	0.18
BB	4.55%	105.30	4.79	-3.50	12.26	0.56
B	0.79%	96.87	0.77	-11.93	142.35	1.12
CCC	0.18%	83.01	0.15	-25.79	665.00	1.20
Default	0.27%	50.00	0.14	-58.80	3457.01	9.33
		Mean	108.80		**Variance**	12.47
					St. dev.	3.53

We conclude the analysis by comparing 99% Value at Risk and the first percentile of the calculated value distribution. For establishing the first percentile it is necessary to sum up the probability of default, CCC and B rating which counts for 1.24% and for corresponding value 96.87. However we need first percentile. We can use the fact that default and CCC counts for 0.45% and value 83.01. If we apply the method of linear polarization we come to the first percentile at value of 92.66 which is 16.14 points bellow the average value of 108.80.

Corresponding 99% VaR calculated from standard normal distribution equals to Var(99%)=2.33*3.53=8.22. This means that assumed BBB rated bond exhibits a long downside tail because corresponding first percentile of value distribution equals 16.14 which is almost the double of 99% VaR. Hence again, **assuming normal distribution structure might not be appropriate**.

3.2.3 VALUATION OF BOND PORTFOLIO

The area of main interest of CreditMetrics^{TM} model might be found, of course, not in the evaluation of one single instruments but in the assessment of portfolio of credit instruments. Together with condition of homogeneous rating system over all portfolio segments, which we already mentioned in the chapter 3.1.1, a portfolio CreditMetrics^{TM} model is built upon **three core assumptions**, Nyfeler (2000):

(i) all bond issuers are credit-homogenous within the same rating class and therefore they share the same transition probabilities
(ii) the transition probability of every firm depends only on the rating category the company is in at the moment of analysis
(iii) the transition probabilities are stationary, i.e. not time dependent

The question under consideration is then how to merge a behaviour of every single portfolio instrument into one common model so that **obligors specific properties as well as their joint dependencies would be respected**. If we assumed independence, i.e. regardless the correlation, the probabilities of assets joint movements would be given only by product of particular marginal probabilities of transition matrix relating to these assets. However in the globalised world with high market interference, this assumption is truly naive. For example, it is broadly accepted[40] that correlation is higher for companies working within the same industry sector and geopolitical region or that correlation significantly vary within the state of economy.

Simply says, joint asset correlation matters and because of that a **simulation of joint distribution of portfolio assets** represents a challenging problem. We might theoretically reckon the correlation for instance with historical defaults data or bond spreads. However in practice both estimates are inaccurate with wide confidence levels respectively suffer from scarcity of data. In a reaction, CMTD (1997) presents its own **Asset value model which links company's credit rating evolution with changes in its asset value** (stock returns)[41]. In another words **the evolution of obligors' credit quality is tied to evolution of their asset returns and simultaneously the dependence structure between the obligors is given by dependence structure of those returns.**

The model further assumes that the asset returns are **jointly correlated and have multivariate normal distribution**. The shape of such distribution is from definition driven by its variance-covariance matrix Σ whose only parameters are variances of its marginal variables and their correlation (see Example 1). Therefore variances and correlation are the only needed model inputs. Moreover, both can be extracted from particular equity returns which stand in the model as proxy variables. We follow this property and in the next section conclude the chapter with model simulation set-up.

Example 1

Here we just recall the form of **variance-covariance matrix of bivariate normal distribution**:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho_{1,2}\,\sigma_1\,\sigma_2 & \rho_{1,3}\,\sigma_1\,\sigma_3 & \cdots & \cdots & \rho_{1,n}\,\sigma_1\,\sigma_n \\ \rho_{2,1}\,\sigma_2\,\sigma_1 & \sigma_2^2 & \rho_{2,3}\,\sigma_2\,\sigma_3 & \cdots & \cdots & \rho_{2,n}\,\sigma_2\,\sigma_n \\ \rho_{3,1}\,\sigma_3\,\sigma_1 & \rho_{3,2}\,\sigma_2\,\sigma_3 & \sigma_3^2 & \cdots & \cdots & \rho_{3,n}\,\sigma_2\,\sigma_n \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ \rho_{n,1}\,\sigma_n\,\sigma_1 & \rho_{n,2}\,\sigma_n\,\sigma_2 & \rho_{n,3}\,\sigma_n\,\sigma_3 & \cdots & \cdots & \sigma_n^2 \end{bmatrix}$$

$\rho_{i,j} \in [0,1]$ stands for linear correlation parameter, i,j=1,…,n

σ_i^2 are variances of random variables.

From the definition, it holds that $\rho_{i,j} = \rho_{j,I}$ *for all i,j=1,…,n* and hence the matrix is symmetric.

---

[40] L. Illová (2005)

[41] We recall the methodological resemblance with model of **Robert C. Merton** from 1974 which was mentioned in the introductory chapter.

3.3 SIMULATION

As we already stated, **CreditMetrics^{TM} model's approach is based on distribution of portfolio value.** Considering this distribution, the model relies on **Monte Carlo simulation** – *a process which generates numerous independent scenarios in which the future credit rating of each obligor in the portfolio is set and correlations are reflected so that for example highly correlated obligors default in the same scenario more frequently than the less correlated.* Within each simulated scenario, obligors' credit ratings fully determine the portfolio value. Finally, gathering all generated independent scenarios together we can estimate the shape of the distribution for the portfolio as a whole.

For the simulation, lets assume the following:

- we have a portfolio with ***n*** obligors which are all rated according to S&P rating system with eight categories of credit quality (i.e. from AAA to default).
- In a one year time of our analysis the credit quality might change to any other seven states including default.
- In accordance with CreditMetrics‘ Asset value model lets further suppose that change in obligors' rating quality depends only on value ***V*** of obligors' asset.
- Finally lets assume that obligors' assets value ***V*** follows multivariate normal distribution.

In order to carry out a simulation in sense of CreditMetrics^{TM} model, we need to first **generate scenarios of multivariate normal distributions and ensure that such generated random variables respect the correlation structure of portfolio assets**. We proceed that we randomly generate n i.i.d. N(0,1) distributions, we mark them η_i for i=1,..,n. Next we interconnect η_i variables so that desired dependence structure given by variance-covariance matrix Σ is met. This might be manage by so called **decomposition of variance-covariance matrix**. There exist two well established techniques for decomposition – Cholesky factorisation and eigenvector-eigenvalue decomposition – both of them lead to the same result. However since the computer Matlab program, which we will use for modelling session, uses Cholesky factorisation, we describe this and refer for instance to http://mathworld.wolfram.com/EigenDecomposition.html for the second method.

3.3.1 CHOLESKY DECOMPOSITION

For *i=1,..,n* let η_i be i.i.d. N(0,1) distributions and ε_i be N(0,1) random variables respecting the dependencies from some variance-covariance matrix Σ. Then the **Cholesky factorisation** might be written as

$$\boxed{\begin{aligned} &\varepsilon_1 = a_{1,1}\,\eta_1 \\ &\varepsilon_2 = a_{2,1}\,\eta_1 + a_{2,2}\,\eta_2 \\ &\vdots \\ &\varepsilon_n = a_{n,1}\,\eta_1 + a_{n,2}\,\eta_2 + \ldots + a_{n,n-1}\,\eta_{n-1} + a_{n,n}\,\eta_n \end{aligned}}$$

Where a_i are unknown coefficient which might be obtained by solving the following system of equations:

1. $Var(\varepsilon_i) = 1 \;\forall i = 1,\ldots,n$
2. $Cov(\varepsilon_i, \varepsilon_j) = \rho_{ij} \;\forall i, j = 1,\ldots,n \wedge i \neq j$
3. We recall that η_i are N(0,1) i.i.d. and hence uncorrelated

Example 2

We assume a bivariate case of assignment from above. The Cholesky factorisation then takes a following form:

$$\varepsilon_1 = a_{1,1}\,\eta_1$$

$$\varepsilon_2 = a_{2,1}\,\eta_1 + a_{2,2}\,\eta_2$$

Since $Var(\varepsilon_1) = Var(a_{1,1}\,\eta_1) = a_{1,1}^2 Var(\eta_1) = 1$ we immediately see that $a_{1,1}$ equals 1.

Secondly $Var(\varepsilon_2) = Var(a_{2,1}\,\eta_1 + a_{2,2}\,\eta_2) = 1$. Since η_i are i.i.d. we can write:

$$Var(a_{2,1}\,\eta_1 + a_{2,2}\,\eta_2) = a_{2,1}^2 \cdot Var(\eta_1) + a_{2,2}^2 \cdot Var(\eta_2) = 1$$

$$a_{2,1}^2 \cdot 1 + a_{2,2}^2 \cdot 1 = 1$$

We substitute $a_{2,1}^2 = \rho^2$ and get the result as $a_{2,1} = \rho \text{ and } a_{2,2} = \sqrt{1-\rho^2}$

To sum up, the Cholesky factorisation in case of two assets has the following representation:

$$\varepsilon_1 = \eta_1$$

$$\varepsilon_2 = \rho \cdot \eta_1 + \sqrt{1-\rho^2} \cdot \eta_2$$

3.3.2 MODELLING CREDIT MIGRATION

Having generated correlated multivariate normal distributions which are describing the evolution of assets' value **V**, the last step of our analysis is to **incorporate a process identifying changes in assets' credit quality**. In order to do so, we slice the generated distributions into several ranges, each one referring to one particular credit quality state.

Since we use eight rating classes, we need to set up seven asset value thresholds $\mathbf{Z_{AAA},\dots,Z_{CCC}}$ which is just enough for a full determination of rating quality of particular obligor given value of its asset. Hence for instance if we have $\mathbf{V<Z_{CCC}}$ then the corresponding rating is default and if we have $\mathbf{Z_{BBB}<V<Z_{A}}$ then the corresponding rating is BBB. Since CreditMetrics^{TM} model assumes normal distribution for asset value ***V***, we can compute the probability of each credit rating state to come just as follows:

$$P[default] = P[V < Z_{CCC}] = \Phi(Z_{CCC} / \sigma)$$

$$P[BBB] = P[Z_{BBB} < V < Z_A] = \Phi(Z_A / \sigma) - \Phi(Z_{BBB} / \sigma)$$

It is crucial to ensure that everything will be in accordance to transition probabilities of transition matrix. The way how this might be manage is called **slicing the x-axis of marginal normal distributions for every obligor into bands which are reflecting a rating category** after possible credit quality migration. The method is described in the table bellow on example of a BBB rated bond.

Table 5: Selecting thresholds

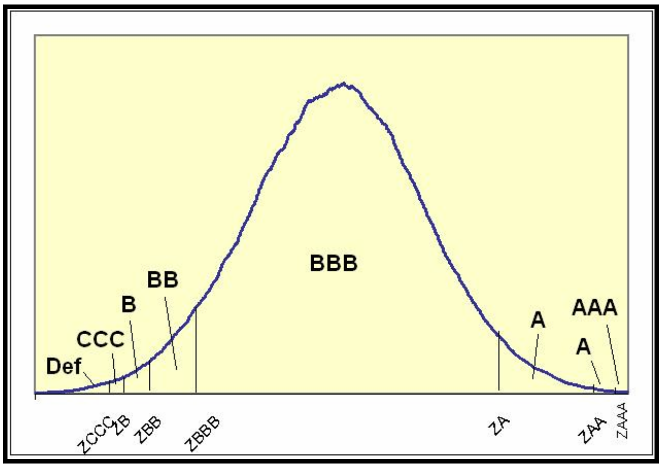
Rating	Probability from the transition matrix	Probability according to the asset value model
AAA	0.02%	$1-\Phi(Z_{AAA} / \sigma)$
AA	0.17%	$\Phi(Z_{AAA} / \sigma)-\Phi(Z_{AA} / \sigma)$
A	4.08%	$\Phi(Z_{AA} / \sigma)-\Phi(Z_{A} / \sigma)$
BBB	89.94%	$\Phi(Z_{A} / \sigma)-\Phi(Z_{BBB} / \sigma)$
BB	4.55%	$\Phi(Z_{BBB} / \sigma)-\Phi(Z_{BB} / \sigma)$
B	0.79%	$\Phi(Z_{BB} / \sigma)-\Phi(Z_{B} / \sigma)$
CCC	0.18%	$\Phi(Z_{B} / \sigma)-\Phi(Z_{CCC} / \sigma)$
Default	0.27%	$\Phi(Z_{CCC} / \sigma)$

Next we calculate the value for each threshold, for instance: $Z_{CCC} = \Phi^{-1}(0.27\%) \cdot \sigma$

Finally when we solve the whole system of equation from Table 5 we get the concrete values for desired thresholds. We present them in the Table 6 together with graphical example of corresponding "sliced normal distribution".

Table 6: Thresholds value and corresponding graph

Threshold	Value
Z_{AAA}	3.26 σ
Z_{AA}	2.83 σ
Z_{A}	1.73 σ
Z_{BBB}	-1.57 σ
Z_{BB}	-2.24 σ
Z_{B}	-2.58 σ
Z_{CCC}	-2.73 σ



At this stage the simulation is completed – Monte-Carlo method generates correlated N(0,1) distributions for which the thresholds are set-up so that assets' credit quality is completely determined. Once we know the credit rating, it is easy to evaluate given assets (see chapter 3.2) and calculate portfolio loss distribution, its volatility and value at risk due to credit.

Using proxy equity returns correlations to evaluate obligor-to-obligor asset correlations has several **shortcomings**. For many obligors there exist a certain scarcity of data which does not enable proper analysis. As well a storage of huge correlation matrices seems to be at least an uneasy problem. Possible solution might be found in resorting obligors and assessing their correlation on basis of **sector index correlation**[42]. However the general quality of this approach could be doubtful as it well depends on the strength of correlation between the sectors.

In our modelling session **we will not manage to avoid the shortcomings relating the data availability** mentioned in the last paragraph and we keep on depending on stock market returns. Nevertheless the main pitfall of Asset value model remains still its assumption of multivariate normal distribution for the assets returns and this is what we are about to change towards more generalized approach in the last chapters.

---

[42] In such approach, correlation between two assets is fully determined by correlation between two sectors which the assets correspond to. On portfolio level, each asset is first assigned particular sector index label. The correlation structure of portfolio having *n* assets is then fully described by **vector of sector indices** (n x 1) and **sector correlation matrix** (m x m) where *m* is a number of incorporated sectors. Since *m*<<*n* the resulting dimension is much smaller then for obligor-to-obligor corr. approach.

4. HEAVY TAILED RISK FACTORS CREDIT RISK MODEL

We are going to introduce a credit risk model which is build up on the foundations of industrial CreditMetrics™ methodology of credit rating migration. However on the contrary to the original CreditMetrics™ approach we are leaving the path of models which are based on assumption of multivariate normal distributions for underlying risk factors. Instead of this we suggest to look at the dependence structure of risk factors from the perspective of copula functions. Such approach enables us to operate with much larger scale of plausible dependence concepts which we utilize for analysing the occurrence of low probable but extreme joint events – i.e. for the analysis of credit portfolio model with heavy tailed risk factors.

4.1 CREDIT RISK MODEL

We consider the model which was presented in **Kostadinov (2006)**. Suppose we have a credit portfolio with *n* obligors ($n \in N$) and $X = (X_1,...,X_n)$ is a random vector with discrete integer margins, all having the same range {1, 2,…, K} and standing for credit rating of i^{th} obligor at the given time horizon T of our analysis (for some *i=1,…,n*). Since throughout the thesis we are using the S&P system of rating with eight rating classes, we set K=8 and we suppose that 1 means default while 8 means rating class AAA. The value of our considered credit portfolio might be then simply written down as:

$$\boxed{V(T) = \sum_{i=1}^{n} e_i \cdot V_i(T)} \tag{4.1}$$

where e_i stands for exposure of the i^{th} obligor and $V_i(T)$ is a particular value given the time horizon *T*. We assume that $V_i(T)$ entirely depends only on credit rating X_i, i.e. once the X_i is determined, the ratings and states of other obligors have no impact on value determination of i^{th} obligor.

Since it is a credit rating migration that determines a value of the whole portfolio, we are naturally mostly interested in finding a way how we might describe its behaviour. In what follows we describe the joint distribution of vector $X = (X_1,...,X_n)$. Once we manage to cope with that, computing the loss distribution should be straightforward and easy.

First we denote the marginal probability of default and credit quality migration by $P(X_i = k) = p_{i,k}$. So we have $P(X_i \le s) = \sum_{k=1}^{s} p_{i,k} = p_i^s\ ,\ s,k = 1,...,K, i = 1,...,n$. (4.2)

Incorporation of the Asset value model of CreditMetrics^{TM} is a next logical step, i.e. we introduce a random vector of (log-)returns of obligors assets $Z = (Z_1,...,Z_n)$ and by means of that will model the dependence structure of vector $X = (X_1,...,X_n)$. We suppose that the considered random vector $Z = (Z_1,...,Z_n)$ has a continuous marginal distributions F_i and a copula function C. In another words the multivariate distribution function of Z is given as:

$$F_Z(z_1,...,z_n) = C(F_1(z_1),...,F_n(z_n)) \tag{4.3}$$

Following the CreditMetrics^{TM} model, the process of matching credit quality migration with asset value development is determined by following expression. For all *i=1,…,n* :

$$X_i = k \quad \Leftrightarrow \quad F_i^{-1}(p_j^{k-1}) < Y_i \le F_i^{-1}(p_j^k)\ ,\ k = 1,...,K \tag{4.4}$$

Kostadinov (2006) gives a proof that the distribution function of $X = (X_1,...,X_n)$ is uniquely determined by the marginal probabilities (4.2) and the copula C of vector Z. Hence in a next step we focus on the dependence structure of the asset returns $(Z_1,...,Z_n)$. We suppose that they follow a linear one factor model with a multiplicative random shock:

$$\boxed{Z_i = \sqrt{\rho} \cdot W \cdot X + \sqrt{1 - \rho} \cdot W \cdot \varepsilon_i} \tag{4.5}$$

where:

- X is a standard normally distributed variable standing for a **common risk factors**
- W is some positive random variable, independent of X. It represents a **global shock** affecting all assets across the portfolio.
- ε_i represents **idiosyncratic risk**[43]. It has i.i.d. N(0,1) distributions for *i=1,…,n* and is independent of X and W.
- ρ is a **factor loading** or, if you want, a correlation coefficient.

It comes out that for W=1 we get $Z = (Z_1,...,Z_n)$ as a Gaussian vector with correlation matrix Σ (i.e. $Z \in N_n(0,\Sigma)$). A popular alternative is to set $W = \sqrt{\frac{\nu}{S_\nu}}$ *where* $S_\nu \in \chi_\nu^2$ - chi-square distribution with ν degrees of freedom. In such case we have a t-model $Z \in t_n(0,\Sigma,\nu)$. More

---

[43] **Idiosyncratic risk** means obligor specific risk.

importantly the Gaussian version $Z \in N_n(0,\Sigma)$ describes exactly what CreditMetrics^{TM} model is. On the other hand the t-model $Z \in t_n(0,\Sigma,\nu)$ stands for our desired heavy tailed risk factors model. Now we turn to the computation of risk of the credit portfolio.

Once we determine the distribution of portfolio values we define **distribution of portfolio losses L**: $L_i = \overline{V} - V(i)$ for $i=1,\dots,n$ (4.6)

The risk of credit portfolio is then generally computed with a value at risk (VaR) measure given at some confidence level α. For our purposes we define it as the α-percentile of the loss distribution L given by relationship (4.6). Hence as a Credit VaR we understand:

$$Credit\ Var(\alpha) = inf\{L : P\{L(t) \le L\} \ge \alpha\} \tag{4.7}$$

Another point of view on the value of exposed credit risk might be obtained by calculating **expected shortfall for confidence level** α. This measurement is defined as a mean value of particular distribution for realizations happening behind the threshold specified by the confidence level α of value at risk measure (4.7). For example, 95% expected shortfall is the mean of values under remaining last 5% of the distribution.

$$Expected\ SF(\alpha) = E\left[X \mid X > VaR(\alpha)\right] \tag{4.8}$$

4.2 SELECTION OF COPULA REPRESENTATION

From a professional point of view the computational aspects of resulting model is one big issue. We might say that we are looking for a **compromise between the model preciseness and difficulty of parameter estimation and duration of model simulation**. We are particularly interested in such compromise for analysis in higher dimensions as industrial portfolios tend to have numerous elements. Archimedean copula functions are not really tractable in higher dimensions and for our analysis would lead to a computationally difficult problem. On the other hand **Gaussian as well as Student-t copula have both the properties which make it easy to estimate their parameters and straightforward to carry out the simulation** (for more information about both copula functions recall the chapter 2.2.1).

Since our concern is to introduce a model of heavy tailed risk factors we give the equation (4.3) representation of the **survival student-t copula function**. Nevertheless our interests are as well in comparison of resulting model with CreditMetrics^{TM} practice and hence we introduce a model with Gaussian survival copula as well. With reference to the previous paragraph we **disregard of Archimedean copula functions** because we find them for our purposes not tractable for analysis within more than two dimensions. Such restrictive assumption is not only convenient for computational efficiency of resulting simulations but further offers very **easy and appealing goodness-of-fit** test. For that we might use a well known fact that t-distributions having around 30 degrees of freedom and more might be approximated by Gaussian distribution. Hence when we are assessing some data set, we might first estimate parameters for t-copula function and then review the value of obtained degrees of freedom. If the value is high enough we might consider that appropriate copula function is a Gaussian one (and t-copula otherwise). Here “high enough” is generally referring to the value around 30 degrees of freedom and more. For purposes of our simulations we decided to set the threshold up to the value of 100 degrees of freedom. Hence our Gaussian approximation should be valid without any need for construction of other test statistics.

In another words, our approach does not require any of goodness-of-fit tests which we introduced in the chapter 2.6. Nevertheless we found it consistent with the copula theory to introduce those test as in practice, where more sophisticated models are implemented, Archimedean copulas and thus as well goodness-of-fit tests are in use.

5. SIMULATION

Throughout the diploma thesis we dwell on the theory of heavy tailed credit risk factors, copula functions and the relating problems. In the remaining part of the thesis we will try to show for a change a bit of practical use of the presented models. For that we carried out three simulations[44]:

- We took several companies from primary market of Prague stock exchange (SPAD) and tested which elliptical copula function[45] would best fit their dependence structure.
- Obtained results were used in the second simulation where we investigated how big influence has a choice of copula function on the number of defaulted obligors within a homogeneous exchangeable model.
- Last simulation is done in the sense of CreditMetrics^{TM} model. We simulated value distribution of bond portfolio, given four portfolios of different riskiness and four elliptical copula functions having different weight in their tails.

5.1 ELLIPTICAL COPULA FUNCTION FOR CZECH STOCK MARKET

As a first one we introduce a **simulation showing which elliptical copula functions might be found in the real world**. We concentrated on the Czech stock market, concretely on the **SPAD market of Prague stock exchange**, and chose all companies which have price history from 7th May 2007 to at least 1st January 2001. However since SPAD market is really small only five companies fulfilled our criteria – these companies are Čez, Unipetrol, Komerční banka, Phillip Morris and Telefonica O2. In order to reduce influence of short term fluctuations we worked with monthly log-returns which makes up 79 observations. This is not much and therefore we as well carried out the same analysis for weekly returns of the identical sample of companies. However considered data sample might not be regarded as i.i.d. due to conditional heteroskedasticity. Therefore we first employ univariate GARCH(1,1)[46] filter and then proceed the simulations on the result.

---

[44] **Algorithms** for each simulation are readable from Matlab codes which are presented in Appendix B.

[45] Here and further we restrict the view of elliptical copula functions to Gaussian copula and t-copula functions only.

[46] Generalized autoregressive conditional heteroskedasticity, Garch (p,q) model**,** is for **GARCH(1,1)** and asset returns described by following system of equations:

(1) $x_t = \mu + \varepsilon_t$ where x_t denotes asset returns, μ mean value and ε_t is a random error defined as:

(2) $\varepsilon_t = \sigma_t \cdot u_t$ where u_t follows i.i.d. N(0,1) distribution and variance σ_t^2 is modelled as:

(3) $\sigma_t^2 = \kappa + \alpha \cdot \varepsilon_{t-1}^2 + \beta \cdot \sigma_{t-1}^2$, where $\kappa \geq 0, \alpha \geq 0, \beta \geq 0$

Having data for five companies we set up four types of portfolios differing in size – one portfolio contains all 5 companies, five portfolios are of size 4, ten portfolios of size 3 and ten portfolios of size 2. Further we refer to chapter 4.2 where we due to computational efficiency decided to carry out our simulations only for Gaussian and student t-copula functions. We applied the **semi parametric method of copula estimation** (see chapter 2.5.2) and present the results of our analysis in Table 7.

Table 7: Copula estimation

Sample	**Monthly returns**		**Weekly returns**	
	Copula	MLE	Copula	MLE
Cez-Telef-Uni-PM-KB	**Gauss**	25.651	t-12	103.598
Cez-Telef-Uni-PM	t-27	17.025	t-14	66.415
Cez-Telef-Uni-KB	**Gauss**	20.938	t-9	89.248
Cez-Telef-PM-KB	**Gauss**	12.150	t-13	73.821
Cez-Uni-PM-KB	t-35	21.381	t-9	68.429
Telef-Uni-PM-KB	t-14	9.148	t-16	46.193
Cez-Telef-Uni	**Gauss**	14.420	t-8	55.949
Cez-Telef-PM	**Gauss**	4.388	t-16	37.356
Cez-Telef-KB	**Gauss**	8.427	t-10	59.785
Cez-Uni-PM	t-9	14.084	t-12	36.566
Cez-Uni-KB	**Gauss**	16.790	**t-6**	54.273
Cez-PM-KB	**Gauss**	8.293	t-7	41.025
Telef-Uni-PM	t-11	4.397	t-39	17.460
Telef-Uni-KB	t-13	4.348	t-10	34.500
Telef-PM-KB	t-22	5.870	t-23	30.070
Uni-PM-KB	t-9	5.352	t-9	22.134
Cez-Telef	**Gauss**	2.531	t-10	26.759
Cez-Uni	**Gauss**	11.520	**t-5**	26.704
Cez-PM	**Gauss**	1.289	t-10	9.851
Cez-KB	**Gauss**	4.827	**t-4**	27.331
Telef-Uni	t-24	1.801	t-8	10.469
Telef-PM	t-18	1.107	**Gauss**	4.911
Telef-KB	t-8	2.621	t-12	19.884
Uni-PM	**t-5**	2.278	**Gauss**	5.159
Uni-KB	t-8	0.550	**t-5**	11.378
PM-KB	t-25	3.167	t-8	9.956

Lets look at the table not from the point of view of what distribution fits the particular sample of data but distributions which might be in the table observed. Considering the monthly returns **12 out of 25 samples might be best represented by Gaussian copula function**. In all these cases company CEZ is included in the examined portfolio. Moreover only three out of 15 portfolios including company CEZ are not represented by a Gaussian copula. Hence taken the monthly returns, presence of CEZ in portfolio has probably a big influence on copula representation. Quite interestingly, none of the 12 portfolios having Gaussian copula representation for monthly returns keeps the Gaussian copula when we consider weekly returns. Moreover out of the 25 portfolios with weekly returns, only two has a Gaussian copula. We might conclude that **weekly returns exhibit higher tail dependence and hence tend more to joint extreme events than monthly returns**.

Nevertheless our main interest was to find out which copula function was represented most frequently. From Table 7 **we identified four copulas which will be referred to in simulations** following in chapters 5.2 and 5.3. Those copulas are **Gaussian, student t-20** standing for average of upper bounds of t-copulas, **student t-10** representing average of lower bounds and **student t-5** representing the lower extreme for t-copulas from Table 7. Gaussian copula further as well serve as a benchmark for comparison of CreditMetrics^{TM} model with models exhibiting heavy tailed risk factors.

5.2 HOMOGENEOUS EXCHANGEABLE MODEL

As a homogeneous exchangeable model we simple understand a portfolio with assets all having the same marginal distribution, sharing the same correlation and most importantly sharing the same exposure. We set up three homogeneous exchangeable portfolios differing in credit quality and **investigate to what extent a copula representation influence number of defaulted obligors within these portfolios**. For the own model we incorporate the concept described in chapter 4.2. We set K=2 and thus work only with two rating categories – default and non-default. Nice property of homogenous exchangeable models is that since all assets are the same, mere number of affected assets is enough to determine the value (or loss) resulting from given model.

Finally portfolios used in our study are defined in Table 8 which is as in Frey (2001). R. Frey presents three realistic homogeneous scenarios (A, B, C) which were made up after discussion with UBS Switzerland. We compared the default threshold of presented scenarios with information from transition matrix in chapter 3 and supplemented each portfolio with its expected credit rating – hence we have portfolios of credit quality AA, BBB and B.

Table 8: Homogeneous exchangeable portfolios

Group	Default threshold	Correlation	Corresponding rating
A	0.01%	2.58%	~AA
B	0.50%	3.80%	~BBB
C	7.50%	9.21%	~B

For each portfolio we supposed 1000 obligors and simulated 5000 independent scenarios for Gaussian, student t-5, t-10 and t-20 copula functions. First idea about intensity of defaults might be provided by coefficient of tail dependence (see chapter 2.4.2). Even though it is defined for a bivariate case only we can use it since homogeneity of our samples. We calculated bivariate coefficient for every copula and portfolio and present the results in first part of Table 9. It is important to bear in mind that this approach generally provides only a rough guidance as number of defaulted events is determined not only by tail dependence of copula function but as well by default threshold which is varying with different samples (however in our case default thresholds and assets correlation are positively correlated).

Further we carried out the own simulations. Within each scenario the numbers of defaulted obligors were summed up and distribution of defaulted obligors was created. The 99th and 95th percentiles of this distribution for each scenario and copula function are presented again in Table 9 bellow. The overall summary is accompanied by graphical evidence of upper percentiles of mentioned - number of defaults – distributions which is given in Figure 6.

Table 9: Homogenous exchangeable model – default distribution

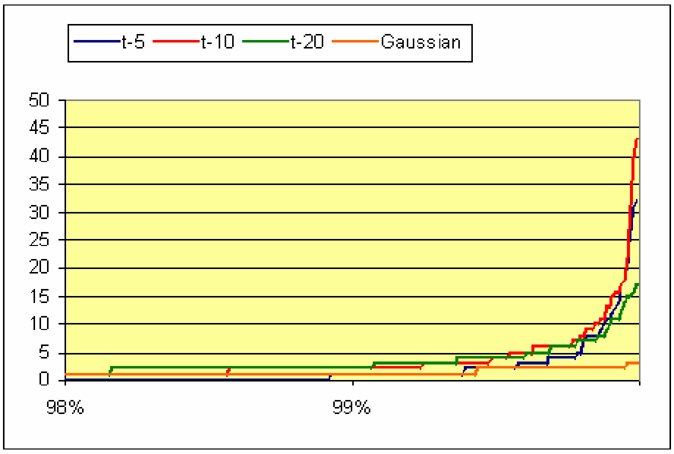
Sample	t-5	t-10	t-20	Gaussian	Gauss / t-5
Tail dependence (X,Y)					
A	0.0542	0.0080	0.0002	0.0000	
B	0.0564	0.0086	0.0002	0.0000	
C	0.0670	0.0116	0.0004	0.0000	
99th percentile					
A	1	2	2	1	1,00
B	99	62	41	18	**5,50**
C	377	316	271	229	1,65
95th percentile					
A	0	0	1	1	0,00
B	27	24	20	12	**2,25**
C	244	209	190	167	1,46

Number degrees of freedom has a massive influence on default distribution. Apart from case A we might clearly see that as we move from t-5 to Gaussian copula, number of default events expectedly falls, sometimes even quite drastically. In the last column of Table 9 we compare number of default events under extreme cases of Gaussian and t-5 copula functions. Immediately we might notice that **relatively biggest difference in outcomes brings Group B representing assets of average credit quality BBB**.

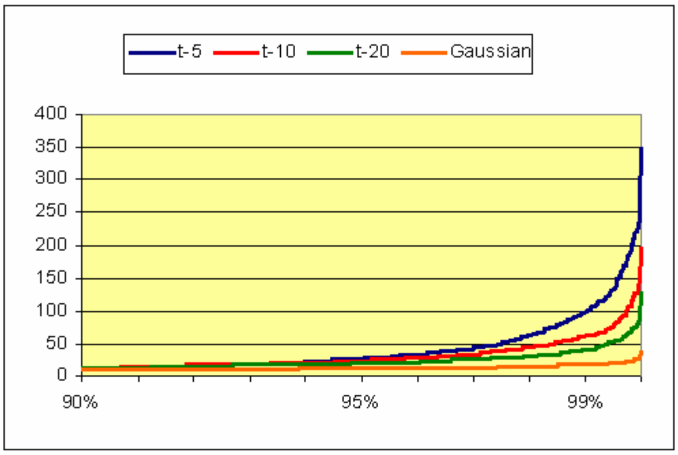
Quite interestingly, considering the Group A, which shows the highest credit quality, we see that even 99th percentile is not high enough for t-copula functions to show more defaults than Gaussian case. Default distribution within Group A is depicted in Figure 6 in illustration (a) with graph's axis beginning at 98th percentile and ending in 100. The top observation of t-5 copula functions showed 293 defaults (the second largest was 32), the same for t-10 function was 54 (and the second largest 43). These observations are such outliers that we left them out and build the graph without them. Other two graphs (b) and (c) are constructed on the scales 90-100 percentiles and no outliers were omitted at this time, hence they show the full distribution.

<u>**Figure 6:** Homogenous exchangeable model – default distribution</u>

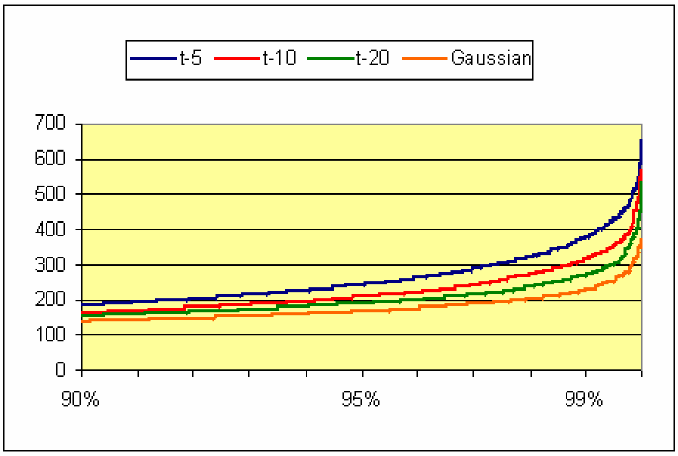
(a) Group A (x-axis: 98-100 percentile of default distribution; y-axis: number of defaults)



(b) Group B (x-axis: 90-100 percentile of default distribution; y-axis: number of defaults)



(c) Group C (x-axis: 90-100 percentile of default distribution; y-axis: number of defaults)

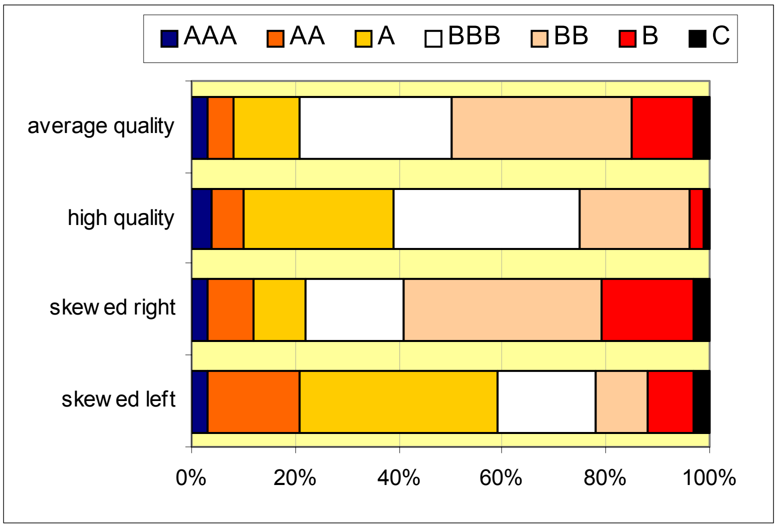


5.3 BOND PORTFOLIO VALUE DISTRIBUTION

The last, third, simulation set-up represents our main interest – value distribution for portfolios of heavy tailed credit instruments. We employ portfolios of bond instruments, the methodology given by CreditMetrics^{TM} model and the credit model described in chapter 4.1. Hence credit rating and rating migration constitutes a fundament of our analysis. However **apart from original CreditMetrics^{TM} model we work not only with dependence structure of multivariate normal distribution but as well with t-copula functions**, again having 5, 10 and 20 degrees of freedom.

For the simulation we considered four bond portfolios of different credit quality which are described in Bank of Canada Working Paper 2004-30. So called average and high quality portfolios refer to survey of U.S. bank holdings in 2002. Other two portfolios called skewed left and skewed right are just hypothetical. The percentage representation for each credit quality and portfolio is depicted in Figure 7.

Figure 7: Quality distribution of contemplated portfolios



Own simulations consisted of 10,000 scenarios calculated for portfolios defined as above and having 200 bond instruments each. Moreover all **bonds are homogenous considering time to maturity (5 years), coupon (5%) and face value of 100 units**. On the other hand, bonds of course differ in their credit quality and thus share different probabilities of credit rating migration. Finally, a correlation matrix (serving as a parameter of used copula functions) is for each portfolio modelled using artificial sector correlation indices[47]. The proper algorithm for model simulation is again described in Appendix B.

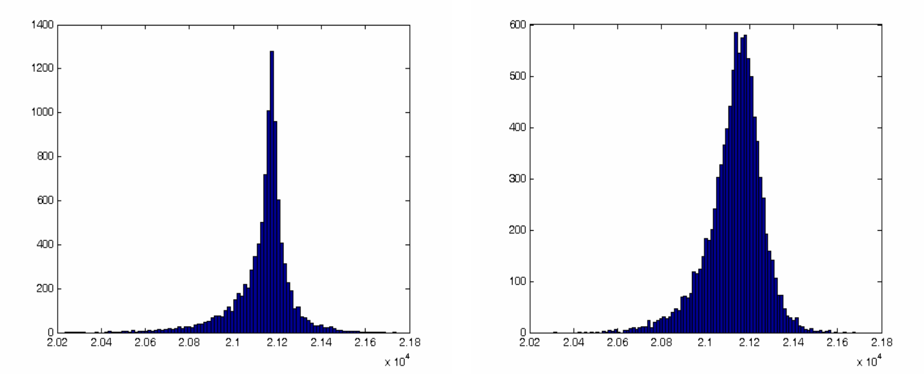
First of all in Figure 8 we present histograms for portfolio value distribution upon simulations with student t-5 a Gaussian copula functions which again represents two extreme scenarios. The remaining histograms for t-10 and t-20 copula functions always lied somewhere in between presented graphs. In histograms we might see that **apart of skewed right portfolio, all cases show a heavy left (downside) tail**. The most evident downside tail might be observed for high quality and skewed left portfolios. For both of those two portfolios the tail is much longer for t-5 copula distribution than Gaussian one. Looking at quality distribution of these portfolios in Figure 7, we see that both have a large proportion of A and BBB rating quality bonds. This is quite in accordance with conclusions from last chapter 5.2 where we said that among AA, BBB a B rated bonds, the BBB are the ones which exhibits the highest sensitivity of outcome to changes in copula representation.

---

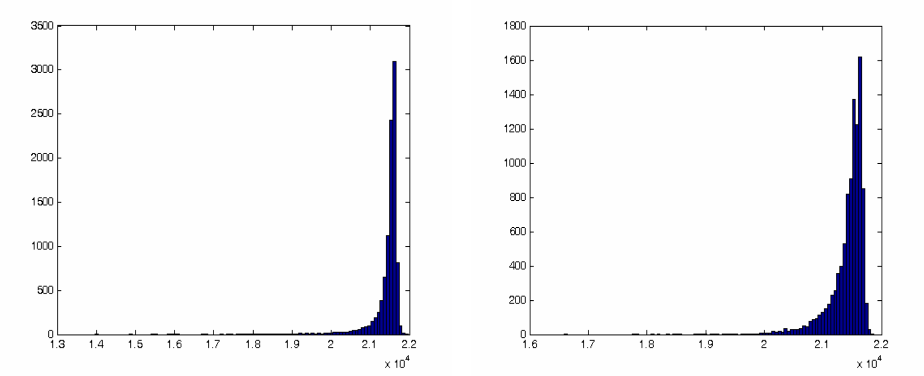
[47] To our analysis we imply four sectors as presented in Jouanin et al. (2003), p.12. Sector indices were randomly distributed along considered portfolio, further portfolio assets correlation matrix was calculated and obtained results were kept fix for all simulations.

Figure 8: Histograms of portfolio value distribution *(t-5 and Gaussian copulas)*[48]

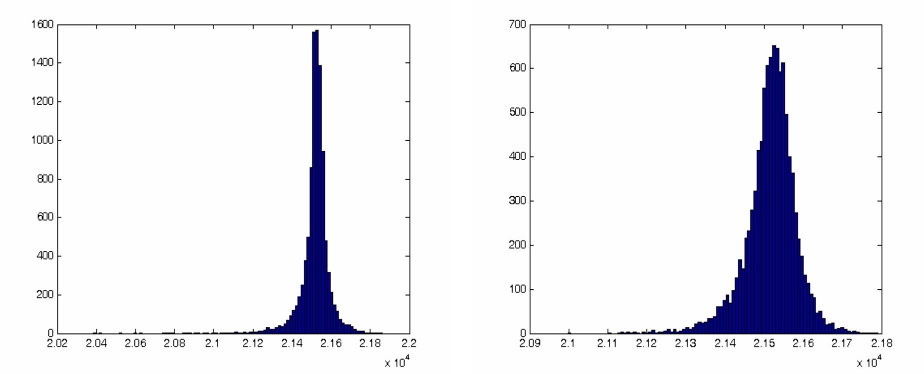
(a) Average quality portfolio



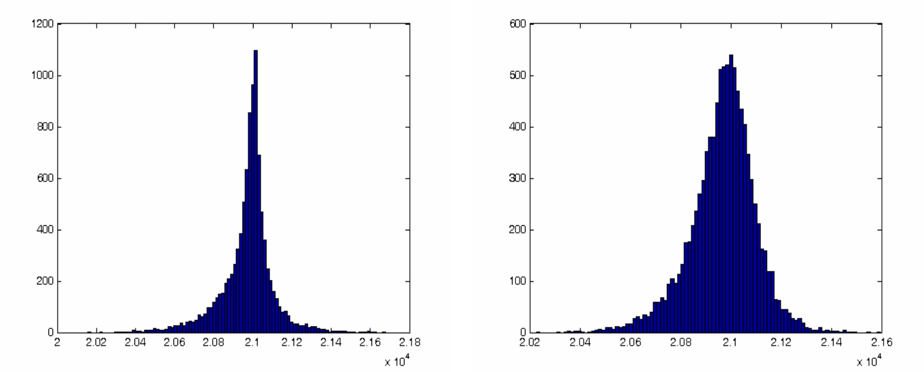
(b) High quality portfolio



(c) Skewed left portfolio



(d) Skewed right portfolio



---

[48] x-axis: portfolio value in "units" ; y-axis: number of realizations

We have four copula functions and four bond portfolios, hence 16 scenarios. For each of them we calculated **expected portfolio value**, 99th and 95th percentiles of **expected portfolio loss** distribution (i.e. value at risk) and values for 99% and 95% **expected shortfall**. All values are presented in “units” with given portfolio having 200 bonds of 100 units value each. Everything is again summarized in table bellow and again we supplement the summary by relative difference of results for extreme t-5 and Gaussian copula functions which might be found in the last column.

Table 10: Portfolios value and loss distribution

	t-5	**t-10**	**t-20**	**Gaussian**	t-5/Gauss
Average quality portfolio					
Exp. value	21138.14	21134.09	21134.62	21133.44	1.00
Loss - 99	479.06	438.64	416.17	391.38	**1.22**
Loss - 95	244.00	247.60	239.18	232.11	1.05
Exp SF-99	593.15	544.48	516.46	474.22	**1.25**
Exp SF-95	384.24	364.76	349.67	327.17	1.17
High quality portfolio					
Exp. value	21417.58	21423.28	21421.54	21418.64	1.00
Loss - 99	2251.54	1766.03	1474.23	1259.78	**1.79**
Loss - 95	689.80	628.78	619.25	597.54	1.15
Exp SF-99	3388.31	2894.35	2271.31	1816.18	**1.87**
Exp SF-95	1644.47	1388.24	1187.30	1029.39	1.60
Skewed left portfolio					
Exp. value	21517.34	21516.62	21515.65	21516.43	1.00
Loss - 99	278.08	268.23	254.62	210.86	**1.32**
Loss - 95	117.42	120.77	117.41	117.02	1.00
Exp SF-99	444.94	393.53	363.94	277.26	**1.60**
Exp SF-95	227.47	214.27	203.49	175.27	1.30
Skewed right portfolio					
Exp. value	20972.75	20968.39	20969.51	20967.94	1.00
Loss - 99	449.21	404.56	400.27	379.87	**1.18**
Loss - 95	245.09	249.00	238.20	233.98	1.05
Exp SF-99	537.30	498.90	485.68	458.87	**1.17**
Exp SF-95	361.85	349.19	337.35	323.41	1.12

Looking at the mean values we might notice that taking each portfolio separately all copula functions always give almost the same results (and of course when comparing different portfolios, the means are different). However, **when we consider percentiles of expected loss distribution, we see huge disparities**. Looking at extremes, at 99th percentile of loss distribution for highly quality portfolio we see that value implied by t-5 copula function is about 79% higher than for Gaussian copula representation. The same case for skewed left portfolio counts for 32% difference. Other two portfolios then show the difference of about 20%. On the other hand considering the 95th percentile values, the disparities are not so evident – once we see 15%, twice 5% and once even zero difference. Quite similar analysis might be done for **values of expected shortfall**. Here large disparities are observable even for 95% level; nevertheless it is not unexpectable since value of 95% ES is to a large extent influenced by extreme values caught in 99% VaR which were already discussed above.

Considering **sensitivity of quality of portfolio assets to changes in copula representation** we see that the most sensitive are two portfolios marked as "high quality" and "skewed left". Both consist of a large proportion of A and BBB rated assets. From these two, the high quality portfolio is more sensitive and contains more BBB rated assets than skewed left portfolio. This again is in accordance with conclusions of chapter 5.2.

Finally we might sum up results of the simulation into the following points:

(i) Selection of copula representation has a massive influence on determination of loss distribution of credit portfolio.

(ii) 95th percentile of loss distribution seems to be the breakpoint where heavy tails of t-copula functions start to prevail over the tails of Gaussian copula

(iii) For 99th percentile of loss distribution we observe huge differences between the results implied by t-copula functions and Gaussian copula.

(iv) Portfolios consisting of large share of BBB rated assets tend to be more sensitive to changes in copula representation than the others.

6. CONCLUSION

Measuring and managing credit risk constitute one of the most important processes within bank risk management. Classical credit risk models assume multivariate normality for underlying risk factors. Resulting methods offer analytical simplicity and computational efficiency but disregard of extreme joint events since their probability is too small. Recently several studies have doubted multivariate normality assumption saying that if we accept this assumption we might seriously underestimate downside risk of given credit portfolio.

The master thesis provided with an insight into the problem of modelling credit risk under assumption of heavy tailed risk factors. In the first part, we discussed the concept of linear correlation in the light of optimal method describing dependence structure between random variables. We introduced copula functions as the alternative method for modelling dependence structures and supplemented this concept with test statistics for the copula fitness and several copula goodness-of-fits tests. The second part of the thesis was dedicated to CreditMetrics^{TM} model which represents the example of models which incorporate criticized assumption of multivariate normality for underlying risk factors. Further, the methodology of this model is widely used in practice. For both these reasons we identified CreditMetrics^{TM} model as the ideal benchmark for comparative analysis with our heavy tailed risk factors model which was introduced later in third part of the thesis. Finally after a short discussion we decided to proceed in remaining sections with Gaussian and t-copulas only.

In the fourth and last part, we introduced and carried out three different simulations. The first simulation estimated appropriate copula functional form and its corresponding parameters given the stock market data for few companies from primary market of Prague stock exchange. Obtained results were used in the second simulation where we investigated how big influence had a choice of copula function on the number of defaulted obligors within a homogeneous exchangeable model. Last simulation was done in the sense of CreditMetrics^{TM} model. We simulated value distribution of bond portfolio, given four portfolios of different riskiness and four copula functions having different weight in their tails.

The first simulation was carried out given GARCH(1,1) monthly and weekly returns for few companies from Czech stock market for time period from 1st January 2001 to 7th May 2007. Considering this we found that weekly returns exhibit higher tail dependence and hence tend more to joint extreme events than monthly returns. Moreover we identified Gaussian, student t-20, t-10 and t-5 copula functions as the optimal ones for the remaining simulations.

For the second simulation we set up homogeneous exchangeable model for three portfolios corresponding to bonds with credit rating AA, BBB and B respectively. We investigated the number of defaulted obligors within these scenarios at 95% and 99% confidence level. At last we compared obtained results for extreme t-5 and Gaussian copula representations We conclude that the relatively biggest difference in outcomes brought second portfolio representing assets of average credit quality BBB.

Last simulation represented our biggest concern – measuring credit risk for bond portfolio. For that we introduced four portfolios of different riskiness from which two corresponded to commercial practice (so called average and high quality portfolios) and two were hypothetical (skewed left and skewed right portfolios). For each portfolio we derived value distribution, calculated expected loss at 95% and 99% confidence level (value at risk), 95% and 99% expected shortfall and again compared the values given by extreme cases of t-5 and Gaussian copula functions. The results of our analysis might be concluded into the following points:

(i) Selection of copula representation has a massive influence on determination of loss distribution of credit portfolio.

(ii) 95th percentile of loss distribution seems to be the breakpoint where heavy tails of t-copula functions start to prevail over the tails of Gaussian copula

(iii) For 99th percentile of loss distribution we observe huge differences between the results implied by t-copula functions and Gaussian copula.

(iv) Portfolios consisting of large share of BBB rated assets tend to be more sensitive to changes in copula representation than the others.

Clearly, multivariate normal distribution is in general cases not the appropriate method of modelling dependence structures for risk factors underlying credit portfolios. Instead of this, more credible dependence structures might be obtained by incorporating the concept of copula functions. Nevertheless there are as well several shortcomings relating to the copula modelling approach. One of the biggest one is perceived as how to choose a correct copula functional form out of many available options. Researchers agree that this problem is to a large extent still undiscovered and leaves an open space for further studies.

APPENDIX A

ALGORITHMS FOR GENERATION OF COPULA FUNCTIONS

In the first appendix we present algorithms which might be used for generating variables following dependencies from different copula distributions which were described in the chapter 2.2. Step by step we introduce the algorithms for Gaussian copula, student t-copula, Gumbel copula, Clayton copula and Frank copula. All algorithms, as presented here, were borrowed from study of C. Cech (2006).

GAUSSIAN COPULA DISTRIBUTION

Simulation of a m-dimensional Gaussian distribution with n simulated realisations and copula parameters given by correlation matrix Σ (m x m):

1. Simulate *m* independent standard normal random variables $z_j \sim N(0,1)\ j=1,..,m$
2. Use a Cholesky factorisation to transform the independent random variables z_j into dependent variables x_j according to the copula parameter matrix Σ.
3. Transform the standard normally distributed variables x_j into variables u_j that are uniformly distributed on (0,1) scales. For that define $u_j = \Phi(x_{j,i})$ for j=1,…,m and i=1,…,n where $\Phi(\cdot)$ is the standard normal distribution function.
4. Compute the simulated joint realizations of the Gaussian distribution a_j as $a_{j,i} = F_j^{-1}(u_{j,i})$, $j = 1,2$, $i = 1,...,n$ where F_j^{-1} is the functional inverse of the j^{th} marginal distribution function F_j.

T-COPULA DISTRIBUTION

Simulation of a m-dimensional student t distribution with n simulated realisations and copula parameters given by correlation matrix Σ (m x m) and ν degrees of freedom:

1. Simulate *m* independent standard normal random variables $z_j \sim N(0,1)\ j=1,..,m$ and one chi-square distributed random vector (1xn) with ν degrees of freedom $s \sim \chi_\nu^2$.

2. Use a Cholesky factorisation to transform the independent random variables z_j into dependent variables x_j according to the copula parameter matrix Σ.
3. Transform the standard normally distributed variables x_j into variables u_j that are uniformly distributed on (0,1) scales. For that define $u_j = t_\nu \cdot \left(x_{j,i} \cdot \sqrt{\nu / s_i} \right)$ for j=1,…,m and i=1,…,n where $t_\nu(\cdot)$ is the Student t distribution function with ν degrees of freedom.
4. Compute the simulated joint realizations of the Student t-copula distribution a_j as $a_{j,i} = F_j^{-1}\left(u_{j,i}\right),\ j = 1,2\ ,\ i = 1,...,n$ where F_j^{-1} is the functional inverse of the j^{th} marginal distribution function F_j.

GUMBEL COPULA DISTRIBUTION

Simulation of a bivariate Gumbel distribution with *n* simulated realizations and copula parameter θ:

1. Simulate 2 independent uniformly distributed random variables $s,\ t_1 \sim U(0,1)$.
2. Use numerical methods to find the elements of the variable t_2 such that
   $$t_{2,i} \cdot \left(1 - \frac{\ln t_{2,i}}{\theta} \right) = s_i \qquad i = 1,...,n$$
3. Compute the dependent uniformly distributed variables $u_{1,2} \sim U(0,1)$ as
   $$u_{1,i} = \exp\left(t_{1,i}^{1/\theta} \cdot \ln t_{2,i} \right)$$
   $$u_{2,i} = \exp\left(\left(1 - t_{1,i}\right)^{1/\theta} \cdot \ln t_{2,i} \right) \qquad i = 1,...,n$$
4. Compute the simulated joint realizations of the Gumbel distribution a_j as $a_{j,i} = F_j^{-1}\left(u_{j,i}\right),\ j = 1,2\ ,\ i = 1,...,n$ where F_j^{-1} is the functional inverse of the j^{th} marginal distribution function F_j.

CLAYTON COPULA DISTRIBUTION

Simulation of a bivariate Clayton distribution with *n* simulated realizations and copula parameter θ:

1. Simulate 2 independent uniformly distributed random variables $s, u_1 \sim U(0,1)$.
2. Compute the dependent uniformly distributed variable $u_2 \sim U(0,1)$ as

   $$u_{2,i} = \left(u_{1,i}^{\theta} \cdot \left(s_i^{-\theta/(\theta+1)} - 1\right) + 1\right)^{-1/\theta} \qquad i = 1,\dots,n$$

3. Compute the simulated joint realizations of the Clayton distribution a_j as $a_{j,i} = F_j^{-1}(u_{j,i})$, $j = 1,2$, $i = 1,\dots,n$ where F_j^{-1} is the functional inverse of the j^{th} marginal distribution function F_j.

FRANK COPULA DISTRIBUTION

Simulation of a bivariate Frank distribution with *n* simulated realizations and
copula parameter θ:

1. Simulate 2 independent uniformly distributed random variables $s, u_1 \sim U(0,1)$.
2. Compute the dependent uniformly distributed variable $u_2 \sim U(0,1)$ as

   $$u_{2,i} = -\frac{1}{\theta} \cdot \ln\left(1 + \frac{s_i \cdot \left(1 - e^{-\theta}\right)}{s_i \cdot \left(e^{-\theta \cdot u_{1,i}} - 1\right) - e^{-\theta \cdot u_{1,i}}}\right) \qquad i = 1,\dots,n$$

3. Compute the simulated joint realizations of the Frank distribution a_j as $a_{j,i} = F_j^{-1}(u_{j,i})$, $j = 1,2$, $i = 1,\dots,n$ where F_j^{-1} is the functional inverse of the j^{th} marginal distribution function F_j.

APPENDIX B

MATLAB CODES (All presented programs are available on the enclosed CD)

1. ESTIMATING ELLIPTICAL COPULA FUNCTION FOR CZECH STOCK MARKET

```
%File SPAD.m provides a program code for semi-parametric method of copula estimation for some
%given data file x. The data file should contain log-returns on some stocks or stock indices. We
%provide such monthly and weekly data for Czech SPAD stock market. The program is restricted to
%work out only t-copula estimates. If the given results exhibit high value for degree of freedom (>100)
%we substitute t-copula by Gaussian copula.
%
%The program goes as follows. First Kendall's tau is estimated and then using its relation with
%correlation, the correlation parameter rho is set up. Next the data from matrix x are transpose by
%their empirical distributions onto [0,1] scales. Finally log-likelihood estimates are constructed and the
%maximum log-likelihood function value is selected.
%
%In the INPUT part you might set the range for degrees of freedom within which the optimum is found
%in the MAIN program section. The most of this program code (particularly Kendall's tau and mle
%estimates) was borrowed from resources of Mr. Saket Sathe - this and more might be found on
%http://www.ee.iitb.ac.in/uma/~saket/ . For high dimensions the program might take time. In order to
%show progress during computation some 'disp' messages introducing just processing sections were
%incorporated into the code.

%INPUTS:
%--------------------------------------------------------------------------------
%--------------------------------------------------------------------------------
clc;
x1=xlsread('E:\Matlab_read_data.xls','monthly'); %Here must be a correct path to data files!!!
x2=xlsread('E:\Matlab_read_data.xls','weekly'); %Here must be a correct path to data files!!!
t=input('Please insert 1 for monthly log returns or 2 for weekly log returns: ');
do=1;
while do==1
lowdof=input('Insert lower bound for degrees of freedom (integer number): ');
updof=input('Insert upper bound for degrees of freedom (integer number): ');
disp('--------------------------------------------');

%MAIN SECTION:
%--------------------------------------------------------------------------------
%--------------------------------------------------------------------------------
if t==1  x=x1;
else  x=x2;
end;

disp('Estimating Kendalls tau');
[sim,n] = size(x);
tau = eye(n); %eye(n) = identity matrix(n)
for row = 1:(n-1)
   for col = (row+1):n
      s = 0;
         for i = 1:sim
            s = s + sum(sign(x(i:sim,row)-x(i,row)).*sign(x(i:sim,col)-x(i,col)));
         end;
      tau(row,col) = s.*2./(sim.*(sim-1));
      tau(col,row) = tau(row,col);
   end;
end;
```

```
disp('Computing parameter for correlation');
rho=sin((pi/2)*tau);

disp('Working on empirical distributions');
for i=1:1:n
   for j=1:1:sim
      index=0;
      for k=1:1:sim
         if x(j,i)>=x(k,i)
            index=index+1;
         end;
      end;
      emp(j,i)=index/(sim+1); %division by (sim+1) ensures that emp(i,j)<1 even for index=sim
   end;
end;

disp('Working on maximum likelihood estimation');
i=0;
for dof=lowdof:1:updof
   tmatrix=tinv(emp,dof);
   i=i+1;
   f1=sim*log((gamma((dof+n)/2))/gamma(dof/2))-sim*n*log((gamma((dof+1)/2))/gamma(dof/2))-
(sim/2)*log(det(rho));
   temp1=0;
   temp=0;
   for t=1:sim
         for n=1:n
         temp1=temp1+log((dof+(tmatrix(t,n)^2))/dof);
         end;
      temp=temp+log((dof+(tmatrix(t,:)*inv(rho)*tmatrix(t,:)'))/dof);
   end;
   f2=-((dof+n)/2)*temp;
   f3=((dof+1)/2)*temp1;
   f = f1 + f2 + f3;
   out(i,1)=dof;
   out(i,2)=f;
end;

[L,i]=max(out(:,2));
DoF=out(i,1);

disp('---------------------------------------------');
disp(sprintf('Maximum log-likelihood function value: %f at DoF %d',L,DoF));

disp('---------------------------------------------');
if DoF>100
   disp('Best fit is Gaussian copula');
else
   disp(sprintf('Best fit is t-%d copula',DoF));
end;

disp('---------------------------------------------');
if DoF==lowdof | DoF==updof
   disp('Given results lies on the boundary of plausible value for DoF');
   disp('Very possible, some better results might be obtained by enlarging the set for DoF');
   do=input('Try again? Please enter 1 for YES or 0 for NO: ');
   disp('---------------------------------------------');
else do=0;
end;
end;
```

2. HOMOGENEOUS EXCHANGEABLE MODEL

```
%File exchangeable.m contains a program code for description default events within some
%homogeneous exchangeable credit portfolio model of size 'n'. All 'n' assets are correlated by same
%correlation parameter 'corr' and as well default threshold 'thrs' has a constant value along all portfolio
%segments. Such assumptions are restrictive however model might be used for observing influences
%of different copulas on default behavior for assets of particular riskiness.

%First, program computes correlation matrix. Given that, four scenarios of portfolio distribution are
%simulated - one for Gaussian copula and three for t-copula functions with degrees of freedom t1, t2
%and t3. Amount of default events is determined and resulted value is given as final output for each
%scenarios at 95 and 99 confidence level.

%In the first, INPUT, section you might select all mentioned parameters, i.e. n, sim (number of
%simulations, should be rather large), corr, thrs, t1, t2 and t3. The second section represents own
%program. For high dimensions the program might take time. In order to show progress during
%computation some 'disp' messages introducing just processing sections were incorporated into the
%code.


%INPUTS:
%------------------------------------------------------------------------------------
%------------------------------------------------------------------------------------
clear;
clc;
n=input('Insert number of portfolio assets: ');
sim=input('Insert number of scenarios to be simulated (must be divisible by 100): ');
corr=input('Insert assets correlation (constant among all assets): ');
thrs=input('Insert default threshold (constant among all assets): ');
t1=input('Insert value for degrees of freedom of first t-copula: ');
t2=input('Insert value for degrees of freedom of second t-copula: ');
t3=input('Insert value for degrees of freedom of third t-copula: ');
disp('---------------------------------------------');


%MAIN SECTION:
%------------------------------------------------------------------------------------
%------------------------------------------------------------------------------------
disp('Working on correlation matrix');
sigma=zeros(n,n);
for i=1:1:n;
 for j=1:1:n;
  if i==j
   sigma(i,j)=1;
  else
   sigma(i,j)=corr;
 end;
end;
end;

disp('Generating multivariate variables');
ndist=mvnrnd(zeros(n,1),sigma,sim);
t1dist=mvtrnd(sigma,t1,sim);
t2dist=mvtrnd(sigma,t2,sim);
t3dist=mvtrnd(sigma,t3,sim);

disp('Computing copula functions');
disp('ncop');
ncop=normcdf(ndist,0,1);
disp('t1cop');
t1cop=tcdf(t1dist,t1);
```

```matlab
disp('t2cop');
t2cop=tcdf(t2dist,t2);
disp('t3cop');
t3cop=tcdf(t3dist,t3);

disp('Computing number of defaulting events');
for i=1:1:sim;
    ndefault=0;
    t1default=0;
    t2default=0;
    t3default=0;
    for j=1:1:n;
        if ncop(i,j)<thrs ndefault=ndefault+1;
        end;
        if t1cop(i,j)<thrs t1default=t1default+1;
        end;
        if t2cop(i,j)<thrs t2default=t2default+1;
        end;
        if t3cop(i,j)<thrs t3default=t3default+1;
        end;
    end;
    ndef(i,1)=ndefault;
    t1def(i,1)=t1default;
    t2def(i,1)=t2default;
    t3def(i,1)=t3default;
end;

disp('Summarising');
nvar=sort(ndef);
t1var=sort(t1def);
t2var=sort(t2def);
t3var=sort(t3def);

order99=sim*0.99;
order95=sim*0.95;

disp('--------------------------------------------');
disp('--------------------------------------------');
disp('Number of defaults on 95 confidence level is:');
disp(sprintf('t-%d copula  : %d',t1,t1var(order95,1)));
disp(sprintf('t-%d copula : %d',t2,t2var(order95,1)));
disp(sprintf('t-%d copula : %d',t3,t3var(order95,1)));
disp(sprintf('Gaussian    : %d',nvar(order95,1)));

disp('--------------------------------------------');
disp('Number of defaults on 99 confidence level is:');
disp(sprintf('t-%d copula  : %d',t1,t1var(order99,1)));
disp(sprintf('t-%d copula : %d',t2,t2var(order99,1)));
disp(sprintf('t-%d copula : %d',t3,t3var(order99,1)));
disp(sprintf('Gaussian    : %d',nvar(order99,1)));


%last part is here just for graphical evidence for master thesis
    %summarized(:,1)=nvar;
    %summarized(:,2)=t1var;
    %summarized(:,3)=t2var;
    %summarized(:,4)=t3var;
    %wk1write('homo.wk1',summarized)
```

3. BOND PORTFOLIO VALUE DISTRIBUTION

```
%File creditmodel.m is designed to calculate distribution of values for portfolio of bond instruments
%on 95 and 99 confidence level. Undertaking analysis is based on credit rating migration - technique
%introduce by JP Morgan in CreditMetricsTM model. Similarly to this model we assume 7 rating
%classes (from AAA to CCC). However on the contrary to JP Morgan's model we employ not only
%Gaussian dependence structures but as well structures implied by t-copula functions. For the
%comparison, each program execution calculates 99 and 95 percent value at risk for one Gaussian
%and three t-copula functions (given degrees of freedom t1, t2, t3).
%
%The program works as follows. In INPUT section, parameters, which are required for the model, are
%set up. This includes portfolio size 'n', number of simulations 'sim' (should be rather large), degrees
%of freedom for considered t-copula functions 't1, t2, t3' and distribution of credit rating 'r1,...,r7'
%(r1 stands for AAA and r7 for CCC). For ratings, you might either select from four predetermined
%options or create your own distribution.
%In MAIN section first correlation matrix 'sigma' is derived given the sector correlation matrix and
%distribution of sectors among the assets (very first this is randomly assigned). Next, four scenarios of
%portfolio distribution are simulated - one for Gaussian copula and three for t-copula functions with
%degrees of freedom t1, t2 and t3. Under these migration scenarios, new credit rating for each asset
%is determined. Given this, assets are evaluated, whole portfolio value computed and resulted value
%is given as final output for each scenarios at 95 and 99 confidence level.
%
%For high dimensions the program might take time. In order to show progress during computation
%some 'disp' messages introducing just processing sections were incorporated into the code.
%
%Program works with three matrices which were determined out of this model - these are sector
%correlation matrix 'sector_corr', transition matrix 'trans_matrix' and matrix for asset value given its
%rating 'value'. These matrices are either describe in diploma thesis or in enclosed.xls files.


%INPUTS:
%-------------------------------------------------------------------------------------
%-------------------------------------------------------------------------------------
clc;
n=input('Insert number of assets (must be divisible by 100): ');
sim=input('Insert number of scenarios to be simulated (must be divisible by 100): ');
t1=input('Insert value for degrees of freedom of first t-copula: ');
t2=input('Insert value for degrees of freedom of second t-copula: ');
t3=input('Insert value for degrees of freedom of third t-copula: ');

%Rating distribution:
in=0;
while (in~=1 & in~=2 & in~=3 & in~=4 & in~=5)
disp('--------------------------------------------------');
disp('Select distribution of ratings: ');
disp('   1...average quality portfolio');
disp('   2...high quality portfolio');
disp('   3...skewed left portfolio');
disp('   4...skewed right portfolio');
disp('   5...different');
in=input('Portfolio number: ');
end;

if in==1
   r1=0.03; r2=0.05; r3=0.13; r4=0.29; r5=0.35; r6=0.12; r7=0.03;
end;

if in==2
   r1=0.04; r2=0.06; r3=0.29; r4=0.36; r5=0.21; r6=0.03; r7=0.01;
end;
```

```
if in==3
   r1=0.03; r2=0.18; r3=0.38; r4=0.19; r5=0.10; r6=0.09; r7=0.03;
end;

if in==4
   r1=0.03; r2=0.09; r3=0.10; r4=0.19; r5=0.38; r6=0.18; r7=0.03;
end;

if in==5
   i=0;
   while i==0
   r1=input('Insert share of assets with AAA rating: ');
   r2=input('Insert share of assets with AA rating: ');
   r3=input('Insert share of assets with A rating: ');
   r4=input('Insert share of assets with BBB rating: ');
   r5=input('Insert share of assets with BB rating: ');
   r6=input('Insert share of assets with B rating: ');
   r7=1-(r1+r2+r3+r4+r5+r6);
      if r7<0
         disp('Share of CCC rating assets is negative. Please try again.');
         i=0;
      else i=1;
      end;
   end;
end;

rating_sum=[r1 r2 r3 r4 r5 r6 r7];
disp('-------------------------------------------------');
disp('Selected rating distribution is (aaa/aa/a/bbb/bb/b/ccc):');
disp(sprintf('%g ',rating_sum));
disp('-------------------------------------------------');

%sector_correlation matrix
%We incorporate 4 sectors but any dimension of sector analysis is generally possible.
sector_corr=[0.3 0.2 0.1 0; 0.2 0.4 0.3 0.2; 0.1 0.3 0.5 0.1; 0 0.2 0.1 0.6];

%transition matrix
trans_matrix=[0.9142    0.0792  0.0051  0.0009  0.0006  0       0       0;
0.0061  0.9068  0.0791  0.0061  0.0005  0.0011  0.0002  0.0001;
0.0005  0.0199  0.9144  0.0586  0.0043  0.0016  0.0003  0.0004;
0.0002  0.0017  0.0408  0.8994  0.0455  0.0079  0.0018  0.0027;
0.0004  0.0005  0.0027  0.0579  0.8360  0.0806  0.0099  0.012;
0       0.0006  0.0022  0.0035  0.0621  0.8249  0.0476  0.0591;
0       0       0.0032  0.0048  0.0145  0.1263  0.5471  0.3041];

%matrix of asset value given its rating (taken from CM.xls, as well presented in Table 4 in chapter 3.2)
value=[110.35 110.31 110.18 109.24 105.30 96.87 83.0 100*betacdf(rand,15,15)];


%MAIN SECTION:
%------------------------------------------------------------------------------------------
%------------------------------------------------------------------------------------------
disp('Generating sectors for underlying assets');
%Each single one asset out of 'n' is randomly assgined some sector.
sector_index=rand(1,n);
size_sc=size(sector_corr,1); %i.e. size_sc is number of sectors incorporated into model
for i=1:1:n
   j=0;
   ind=0;
   while ind==0
```

```
        j=j+1;
        if (((j-1)/size_sc)<=sector_index(1,i) & sector_index(1,i)<(j/size_sc))
        sector_index(1,i)=j;
        ind=1;
        end;
        if j==size_sc
        ind=1; %i.e. for utterly rare cases when a(1,i) = 1, we set sector_index=1
        end;
    end;
end;


disp('Constructing correlation matrix');
for i=1:1:n;
  for j=i:1:n;
   row=sector_index(1,i);
   column=sector_index(1,j);
    if i==j
      sigma(i,j)=1;
    else
     sigma(i,j)=sector_corr(row,column);
     sigma(j,i)=sector_corr(row,column);
    end;
  end;
end;


disp('Assigning credit rating values for underlying assets');
%Here we combine information about rating distribution from rating_sum with number
%of assets 'n'.
i=0;
cumul=0;
for j=1:1:7
    for i=(i+1):1:(rating_sum(1,j)*n+cumul)
        rating(1,i)=j;
    end;
    cumul=i;
end;
rating(1,n)=7;


disp('Generating multivariate variables');
ndist=mvnrnd(zeros(n,1),sigma,sim);
t1dist=mvtrnd(sigma,t1,sim);
t2dist=mvtrnd(sigma,t2,sim);
t3dist=mvtrnd(sigma,t3,sim);


disp('Computing copula functions');
ncop=normcdf(ndist,0,1);
t1cop=tcdf(t1dist,t1);
t2cop=tcdf(t2dist,t2);
t3cop=tcdf(t3dist,t3);

%First row in ncop and txcop is sacrified and will further refer to credit quality of
%assets in given column.
ncop(1,:)=rating;
t1cop(1,:)=rating;
t2cop(1,:)=rating;
t3cop(1,:)=rating;
```

```
%We create cumulative transition matrix which will be used for credit matching in next section.
for i=1:1:7
cumul_trans_matrix(i,1)=0;
   for j=2:1:9
      if j==2
      cumul_trans_matrix(i,j)=trans_matrix(i,j-1);
      else
      cumul_trans_matrix(i,j)=trans_matrix(i,j-1)+cumul_trans_matrix(i,j-1);
   end;
end;
end;
cumul_trans_matrix(:,9)=cumul_trans_matrix(:,9)+0.001; %This is necessary for the next section
                                                %in order to unambiguously determine the credit rating.


disp('Determining new credit rating');
for i=1:1:n
   for j=2:1:sim %We go from j=2 as j=1 correspond to the initial rating.
      r=ncop(1,i); %This identified original asset rating, next new rating is determined.
      for k=1:1:8
         if (cumul_trans_matrix(r,k)<=ncop(j,i) & ncop(j,i)<cumul_trans_matrix(r,k+1))
            ncat(j,i)=k;
         end;
         if (cumul_trans_matrix(r,k)<=t1cop(j,i) & t1cop(j,i)<cumul_trans_matrix(r,k+1))
            t1cat(j,i)=k;
         end;
         if (cumul_trans_matrix(r,k)<=t2cop(j,i) & t2cop(j,i)<cumul_trans_matrix(r,k+1))
            t2cat(j,i)=k;
         end;
         if (cumul_trans_matrix(r,k)<=t3cop(j,i) & t3cop(j,i)<cumul_trans_matrix(r,k+1))
            t3cat(j,i)=k;
         end;
      end;
   end;
end;


disp('Evaluating assets under new credit ratings');
%Given the 'value' matrix from input section, we evaluate each asset.
%Asset values on each row are then sumed up to give portfolio value.
for i=2:1:sim
   for j=1:1:n
      rat=ncat(i,j);
      nval(i,j)=value(1,rat);
   end;
   nport_val(i,1)=sum(nval(i,:));
end;

for i=2:1:sim
   for j=1:1:n
      rat=t1cat(i,j);
      t1val(i,j)=value(1,rat);
   end;
   t1port_val(i,1)=sum(t1val(i,:));
end;

for i=2:1:sim
   for j=1:1:n
      rat=t2cat(i,j);
```

```
        t2val(i,j)=value(1,rat);
    end;
    t2port_val(i,1)=sum(t2val(i,:));
end;

for i=2:1:sim
    for j=1:1:n
        rat=t3cat(i,j);
        t3val(i,j)=value(1,rat);
    end;
    t3port_val(i,1)=sum(t3val(i,:));
end;


disp('Computing Value at Risk');
nvar=sort(nport_val(2:sim,1));
t1var=sort(t1port_val(2:sim,1));
t2var=sort(t2port_val(2:sim,1));
t3var=sort(t3port_val(2:sim,1));

order99=sim*0.01;
order95=sim*0.05;

%Expected (mean) value
nmean=sum(nvar)/size(nvar,1);
t1mean=sum(t1var)/size(t1var,1);
t2mean=sum(t2var)/size(t2var,1);
t3mean=sum(t3var)/size(t3var,1);
A=[t1mean t2mean t3mean nmean];

%VaR
Var99=A-[t1var(order99,1) t2var(order99,1) t3var(order99,1) nvar(order99,1)];
Var95=A-[t1var(order95,1) t2var(order95,1) t3var(order95,1) nvar(order95,1)];


disp('-----------------------------------------------');
disp('-----------------------------------------------');
disp('Value at risk on 95 confidence level is:');
disp(sprintf('t-%d copula  : %g',t1,Var95(1,1)));
disp(sprintf('t-%d copula : %g',t2,Var95(1,2)));
disp(sprintf('t-%d copula : %g',t3,Var95(1,3)));
disp(sprintf('Gaussian    : %g',Var95(1,4)));

disp('-----------------------------------------------');
disp('Value at risk on 99 confidence level is:');
disp(sprintf('t-%d copula  : %g',t1,Var99(1,1)));
disp(sprintf('t-%d copula : %g',t2,Var99(1,2)));
disp(sprintf('t-%d copula : %g',t3,Var99(1,3)));
disp(sprintf('Gaussian    : %g',Var99(1,4)));


%last part is here just for graphical evidence for master thesis
    %summarized(:,1)=nvar;
    %summarized(:,2)=t1var;
    %summarized(:,3)=t2var;
    %summarized(:,4)=t3var;
    %x=hist(summarized,100)
    %wk1write('CM_var.wk1',summarized)
```

PROJECT OF MASTER THESIS

Author: Bc. Petr Jablonský
Leader of the master thesis: Doc. Ing. Miloslav Vošvrda, CSc.
Term of master examinations: Summer semester 2006-2007

Title: *Measuring credit risk for portfolios with heavy-tailed risk factors*

Characteristics of the theme:

The aim of the master thesis is to introduce a model for measuring portfolio credit risk under the assumption of possible joint default events. The portfolio credit risk will be first modelled in sense of method CreditMetrics^{TM} which represents a benchmark of commercial practice. The methodology of examined model will be further extended by the assumption of heavy tails for underlying risk factors. Such assumption causes much higher occurrence of extreme joint events than what might be observed under the CreditMetrics^{TM} model. To find the extend of diversity in model outcomes, both approaches are confronted using the Monte Carlo simulation on artificial portfolios of different riskiness. At the end we discuss the consequences of assumed model for decisioning process of credit risk management.

The master thesis will try to find the answer to following questions:

- What consequences does assumption of joint default bring to portfolio credit risk management?
- Are classic credit risk models able to deal with assumption of joint default?
- What are optimal methods for modelling heavy tailed risk factors?
- To what extent are outcomes resulting from models with and without an assumption of heavy tailed risk factor different?
- What kind of recommendation for measuring and managing credit risk might new model bring?

The basic outline:

1. Introduction
2. Mathematical preliminaries
3. Model CreditMetrics^{TM}
4. Heavy tailed risk factors credit risk model
5. Monte Carlo simulation and model comparisons
6. Resulting consequences for decisioning process of credit risk management
7. Conclusion

Main resources:

Morgan, J.P.: CreditMetricsTM – Technical Document, J.P. Morgan, New York, version: April 2, 1997

Kostadinov, K. K.: Portfolio Credit Risk Modelling With Heavy-Tailed Risk Factors, Munich University of Technology, 2006

Nyfeler, M. A.: Modelling Dependencies in Credit Risk Management, Swiss Federal Institute of Technology Zurich, 2000

Schonbucher, P. J. , Schubert, D.: Copula-Dependent Default Risk in Intensity Models, Bonn University, Economics Faculty, Department of Statistics, 2001

In Prague 2nd May 2007

Signature of the consultant: Signature of the author:

BIBLIOGRAPHY

BANGIA, A., DIEBOLD, F.-X. and T. SCHUERMANN [2000], Ratings Migration and the Business Cycle, With Applications to Credit Portfolio Stress Testing, *Department of Economics, University of Pennsylvania.*
Available at: http://www.ssc.upenn.edu/~fdiebold/papers/paper37/bds.pdf

BERG, D. and H. BAKKEN [2006], A copula goodness-of-fit test based on the probability integral transform, *The Norwegian Computing Center and University of Oslo, Oslo*
Available at: http://www.danielberg.no/publications/Btest.pdf

BERG, D. and H. BAKKEN [2006], Copula Goodness-of-fit Tests: A Comparative Study, *The Norwegian Computing Center and University of Oslo, Oslo*
Available at: http://www.danielberg.no/publications/CopulaGOF.pdf

BLUM, P., DIAS, A. and P. EMBRECHTS [2002], The ART of Dependence Modelling: the Latest Advances in Correlation Analysis, *Depertment of Mathematics, Swiss Federal Institute of Technology, Zurich*
Available at: http://www.math.ethz.ch/~baltes/ftp/art.pdf

BOUYÉ, E. [2000], Copulas for Finance A Reading Guide and Some Applications, *City University Business School, Financial Econometrics Research Centre, London*
Available at: http://gro.creditlyonnais.fr/content/wp/copula-survey.pdf

CECH, C. and I. FORTIN [2006], Investigating the dependence structure between market and credit portfolios' profits and losses in a top-down approach using institution-internal simulated data, *University of Applied Sciences bfi Vienna, Study paper*
Available at: www.fh-vie.ac.at/files/Study.pdf

CECH, C. [2006], Copula-based top-down approaches in financial risk aggregation, *University of Applied Sciences of bfi Vienna, Vienna, Working Paper Series, n. 32*
Available at: www.fh-vie.ac.at/files/workingpapers/Cech_WP32.pdf

CROUHY, M., GALAI, D. and R. MARK [2000], A comparative analysis of current credit risk models, *Journal of Banking & Finance 24 (2000) 59-117*
Available at: http://www.creditriskresource.com/papers/paper_101.pdf

DEMARTA, S. and A. J. MCNEIL [2004], The t Copula and Related Copulas, *Department of Mathematics, Federal Institute of Technology, Zurich*

Available at: http://www.math.ethz.ch/~mcneil/ftp/tCopula.pdf

DJEHICHE, B., LIV, S. and H. HULT [2004], An Introduction to Copulas with Applications, *Svenska Aktuarieföreningen, Stockholm*

Available at: http://www.aktuarieforeningen.com/afsidor/utbildning-filer/studiecirkel%202004/Copulas%20presentation%20-%20Boualem%20och%20Hult.pdf

DURRLEMAN, V, NIKEGHBALI, A. and T. RONCALLI [2001], What are the Most Important Copulas in Finance?, *Groupe de Recherche Opérationnelle, Crédit Lyonnais, Presentation*

Available at: http://gro.creditlyonnais.fr/content/wp/copula-tunisia-march17.pdf

EMBRECHTS, P., MCNEIL, A.-J. and D. STRAUMANN [1999], Correlation and Dependence in Risk Management: Properties and Pitfalls, *Risk Management: Value at Risk and Beyond, ed. M.A.H. Dempster, Cambridge University Press, Cambridge, pp. 176-223*

Available at: http://www.math.ethz.ch/%7Estrauman/preprints/pitfalls.pdf

EMBRECHTS, P., LINDSKOG, F. and A. MCNEIL [2001], Modelling Dependence with Copulas and Application to Risk Management, *Department of Mathematics, ETHZ, Zurich*

Available at: http://www.math.ethz.ch/~baltes/ftp/copchapter.pdf

FREES, E.-W. and E.-A. VALDEZ [1998], Understanding Relationships Using Copulas, *School of Business, University of Wisconsin–Madison, Winsconsin*

Available at: http://library.soa.org/library/naaj/1997-09/naaj9801_1.pdf

FREY, R. and A.-J. MCNEIL [2001], Modelling Dependent Defaults, *Department of Mathematics, Federal Institute of Technology, Zurich*

Available at: www.math.ethz.ch/~mcneil/ftp/defaults.pdf

FREY, R., MCNEIL, A.-J. and M.-A. NYFELER [2001], Copulas and credit models, *Swiss Banking Institute, Zurich*

Available at: http://www.math.uni-leipzig.de/~frey/copula-credit.pdf

FREY, R. [2006], Credit Risk Management, *Universität Leipzig, Leipzig, Lecture Notes*

HOFF, P.-D. [2006], Marginal set likelihood for semiparametric copula estimation, *Departments of Statistics, Biostatistics and the Center for Statistics and the Social Sciences, University of Washington, Seattle*

Available at: http://www.stat.washington.edu/www/research/reports/2006/tr505.pdf

ILLOVÁ, L. [2005], Stress testing of bank risks, *Master Thesis, IES FSV, Charles University in Prague*

Available at: http://ies.fsv.cuni.cz

ILLING M. and G. PAULIN [2004], The New Basel Capital Accord and the Cyclical Behaviour of Bank Capital, *Bank of Canada Working Paper 2004-30*

Available at: http://dsp-psd.pwgsc.gc.ca/Collection/FB3-2-104-30E.pdf

J.P. MORGAN [1997], CreditMetrics^{TM} – Technical Document, *J.P. Morgan, New York*

Available at: http://www.creditriskresource.com/papers/paper_125.pdf

JOUANIN, J.-F., RIBOULET, G. and T. RONCALLI [2003], Financial Applications of Copula Functions, *Groupe de Recherche Opérationnelle, Crédit Lyonnais, France*

Available at: http://gro.creditlyonnais.fr/content/wp/copula-szego.pdf

KÄÄRIK, E. [2006], Imputation Algorithm Using Copulas, *Metodološki zvezki, Vol. 3, No. 1, 2006, 109-120*

Available at: http://mrvar.fdv.uni-lj.si/pub/mz/mz3.1/kaarik.pdf

KOLE, E., KOEDIJK, K. and M. VERBEEK [2006], Selecting Copulas for Risk Management, *Erasmus University Rotterdam, Rotterdams*

Available at: http://people.few.eur.nl/kole/SSRN-id929448.pdf

KOSTADINOV, K.-K. [2006], Portfolio Credit Risk Modelling With Heavy-Tailed Risk Factors, *Center of Mathematical Sciences, Munich University of Technology, Munich, Dissertation*

Available at: http://deposit.ddb.de/cgi-bin/dokserv?idn=979981778&dok_var=d1&dok_ext=pdf&filename=979981778.pdf

LARDY, J.-P. [2002], A Simple Model to Assess Default Probabilities from Equity Markets, *JPMorgan Credit Derivatives Conference, JPMorgan*

Available at: http://www.gloriamundi.org/picsresources/jpl.pdf

LINDSKOG, F., MCNEIL, A. and U. SCHMOCK [2001], Kendall's Tau for Elliptical Distributions, *RiskLab, Switzerland*

Available at: http://www.math.kth.se/~lindskog/papers/KendallsTau.pdf

LONGIN, F. and B. SOLNIK [2001], Extreme Correlation of International Equity Markets, *Forthcoming in The Journal of Finance*.

Available at: http://www.hec.fr/hec/fr/professeurs_recherche/upload/cahiers/CR705.pdf

LUCAS, A., KLAASEN, P., SPREIJ, P. and S. STRAETMANS [2001], Tail Behavior of Credit Loss Distributions for General Latent Factor Models, Vrije Universiteit, Amsterdam

Available at: http://www.fdewb.unimaas.nl/finance/faculty/straetmans/amf.pdf

MASHAL, R. and M. NALDI [2002], Pricing Multiname Credit Derivatives: Heavy Tailed Hybrid Approach, *Graduate School of Business, Columbia University*

Available at: http://www.faculty.idc.ac.il/roy/Pub/Mashal_Naldi_Hybrid.pdf

MERTLÍK, J. [2007], Kopuly a České Státní Dluhopisy, *Article*

NYFELER, M. A. [2000], Modelling Dependencies in Credit Risk Management, *Swiss Federal Institute of Technology Zurich, Zurich, Diploma Thesis*

Available at: http://www.risklab.ch/ftp/papers/DiplomaThesisNyfeler.pdf

PANCHENKO, V. [2006], Estimating and Evaluating the Predictive Abilities of Semiparametric Multivariate Models with Application to Risk Management, *CeNDEF, University of Amsterdam, Amsterdam*

Available at: http://www.univie.ac.at/vwl/QED/Meeting2006/panchenko.pdf

PANCHENKO, V. [2005], Goodness-of-fit test for copulas, CeNDEF, *University of Amsterdam, Amsterdam*

Available at: http://www1.fee.uva.nl/cendef/publications/papers/artcopula.pdf

PANZA, A. and V. SCOCCIA [2004], Applying a Parametric method for Credit Risk VAR: a Comparison between One-Factor and Multifactor Approach, *Rome, Presentation*

Available at: http://www.gloriamundi.org/picsresources/vsap.pdf

PURCARU, O. and Y. GOEGEBEUR [2003], Modelling Dependence through Copulas, *Institute de statistique & Institut de sciences actaurielles, Université Catholique de Louvain, Belgium*

Available at: http://www.gloriamundi.org/picsresources/peflam.pdf

RONCALLI, T. [2003], Some Pratical Issues on Credit Risk, *Séminaire CREST / LFA, Paris, Groupe de Recherche Opérationnelle, Crédit Lyonnais*

Available at: http://www.gro.creditlyonnais.fr/content/wp/creditrisk-crest.pdf

SALMON, M. [2004], The Econometrics of Copula Functions, *Warwick Business School, University of Bologna, Lecture Notes.* Available at: http://www2.warwick.ac.uk/fac/soc/wbs/subjects/finance/finance_faculty/mark_s/bologna.pdf

SCHÖNBUCHER, J.-P. [2002], Taken to the Limit: Simple and not-so-simple Loan Loss Distributions, *Department of Statistics, Bonn University, Bonn.* Available at: http://www.defaultrisk.com/_pdf6j4/Taken_To_the_Limit-Smpl_n_Nt-s-smpl_Ln_Lss_Dstrbtns.pdf

SCHÖNBUCHER, J. P. and D. SCHUBERT [2001], Copula-Dependent Default Risk in Intensity Models, *Department of Statistics, Bonn University, Bonn*

Available at: http://www.creditriskresource.com/papers/paper_116.pdf

TRÜCK, S. and S. T. RACHEV [2005], Credit Portfolio Risk and PD Confidence Sets through the Business Cycle, *Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe*

Available at: http://www.statistik.uni-karlsruhe.de/download/tr_credit_portfolio_risk.pdf

VRIES, A. [2000], The Value at Risk, *WestLB, D-40199 Düsseldorf, Germany*

Available at: http://citeseer.ist.psu.edu/cache/papers/cs/31813/http:zSzzSzwww.ruhr-uni-bochum.dezSzfinanzmathezSzVaR.pdf/devries00value.pdf

WEHRSPOHN, U. [2002], Credit Risk Evaluation, Modeling – Analysis – Management, *Heidelberg University, Inaugural-dissertation*

Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=670884

ŽIKEŠ, F. [2007], Dependence Structure and Portfolio Diversification on Central European Stock Markets, *IES Working Paper: 2/2007, Charles University in Prague*

Available at: http://ies.fsv.cuni.cz