CHALMERS | GÖTEBORG UNIVERSITY

MASTER'S THESIS

Pricing Bivariate Rainbow Options Using a Copula-Based Approach

ELISABET GULEVSKA
TOMAS WESTMAN

Department of Mathematical Statistics
CHALMERS UNIVERSITY OF TECHNOLOGY
GÖTEBORG UNIVERSITY
Göteborg, Sweden 2011

Thesis for the Degree in Master of Science in Mathematical Statistics
with specialization in Financial Mathematics

Pricing Bivariate Rainbow Options Using a Copula-Based Approach

Elisabet Gulevska
Tomas Westman

CHALMERS | GÖTEBORG UNIVERSITY

Department of Mathematical Statistics
Chalmers University of Technology and Göteborg University
SE – 412 96 Göteborg, Sweden
Göteborg, November 2011

Pricing Bivariate Rainbow Options Using a Copula-Based Approach

Elisabet Gulevska
Tomas Westman

2011-11-11

Abstract

Copula functions can be used for modelling the dependence structure between two random variables. In this thesis we use a GARCH model in order to model the logreturns of different stock prices. We combine the GARCH model with the Copula theory in order to price bivariate rainbow options. In order to validate the model we consider stocks from different time periods and stock markets.

The aim of this thesis is to evaluate the estimated option price when the dependence structure between two stock prices is modelled by different copula functions.

Key words: Rainbow options, copula, GARCH, dependence.

Acknowledgement

We would like to thank our supervisor Patrik Albin for his support during this project. Also, a special thank you to Niclas Delfs for sharing his in-depth knowledge in programming.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The usage of options can be dated back to ancient times. At that time options were manly used in agriculture. One famous example is *Thales of Miletus* whom, according to legend, bought the rights to use a number of olive presses the following year, since he had strong reasons to believe the harvest would turn out to be better than usual. Today options are still used in agriculture, however, the usage has spread widely and is now covering everything from real estate to stocks.

There are several reasons for using options. One reason is in the sense Thales of Miletus used them, namely to profit after making predictions about the future. For example, if one has believes that a stock price will increase by 10% during the following month, an option contract can be signed giving the holder of the option the right but not the obligation to purchase the stock for, say 3% higher than today's price. Thus, the holder can make a 7 percentage point gain if the prediction turns out to be correct. However, there is always a price for signing an option contract and if the prediction turns out to be wrong, the holder risks having to pay the price of the contract, without exercising the contract. The reason for not just purchasing the stock today and waiting for the stock price to increase, is that by signing an option contract one reduces the maximum possible loss to the price of the contract. If not signing an option contract and instead purchase the stock straight away, the maximum possible loss is equal to the stock price, since there is always a possibility for bankruptcy.

A second reason for the use of options is if one has an obligation to purchase an asset in the future. In such a situation one might want to reduce risk. This can be done by agreeing on a fixed price in the future through an option contract. In the same way as for the first reason, the maximum possible loss is the price of the contract.

The perhaps most common option strategy is the so called *delta hedging strategy*. In this strategy a portfolio is created in such a way that small changes of the underlying asset does not change the total price of the portfolio. In order to create such a portfolio one can use different options which offset the value change of the asset. Thus if combined properly options can cancel out an asset price change on a portfolio.

In this thesis we will aim to price options that depend on two stocks. The stock price will be modelled using a special case of the *General AutoRegressive Conditional Heteroskedasticity* model. Often, in option theory, the variance of the log returns are assumed to be constant, e.g. in Black-Scholes theory. Many studies has shown that this is not the case, thus a more sophisticated model, such as the GARCH, is needed in order to model a stochastic variance process. By using this model we assume that the log returns have standard normal residuals.

Since our option price depend on two stocks, we need to describe the dependence structure between the residuals of the stock prices. We are able to do this by using the *concept of copulas*, thus joining the distribution of each log return into a joint probability distribution. Copulas are a rather modern phenomenon in financial mathematics. In the last decade the usage for them have rapidly increased. As for all new applications, some critique has been pointed at people who quite casually use copulas without understanding what exactly is being modelled. Some people even take it further and claim that the financial crisis of the late 2000s was a direct consequence of the use of copulas. At that time Gaussian copulas were used to model the credit risk of collateralized debt obligations. However, as David X. Li, the person who introduced the copula concept in credit risk, puts it [15]

"The most dangerous part is when people believe everything coming out of it".

Chapter 2

Options

An option is a financial derivative that gives a holder the right but not the obligation to make a transaction on an underlying asset at a pre-specified price, called the *strike price* or *exercise price*, at a given (or up to a) date. Since we in this thesis will be working with stocks, we will be referring to the asset as a stock and drop the more general term.

The simplest type of options are the so called *European put* or *call option*. These two options gives the holder of the option the right but not the obligation to sell or buy, respectively, the underlying stock at a given date and price. With this in mind it is easy to find that the payoff function, $g(x)$, for a European call option with expiration date T is

$$g(S(T)) = \max(S(T) - K, 0))$$

where $S(T)$ is the underlying stock price at expiration date T and K is the pre-specified strike price of the stock. If an option can be exercised at any time during its lifetime of the option it is called an *American option* and they are more common on exchanges than their European counterpart [9].

2.1 Pricing an option

Assuming that a stock price, $S(T)$, is arbitrage free and that the market is complete, the *Fundamental theorem of asset pricing* [5] states that there exists a probability measure $\mathbb{Q}$ such that the discounted stock price, $e^{-r(T-t)}S(T)$, is a martingale under $\mathbb{Q}$ and $\mathbb{Q}$ is also equivalent[1] to the real world probability measure $\mathbb{P}$. The fact that the discounted stock price is a martingale under $\mathbb{Q}$ is essential to option pricing since the fair price of the option boils down to the expected value of the options payoff.

[1] $\mathbb{Q}$ is equivalent to $\mathbb{P}$ if they have the same null set [10].

Definition 2.1.1 *Let t be todays date and T the date of maturity, thus (T-t) is the time to maturity. Furthermore, let S be a stock price as before and $g(S)$ be the payoff function. The fair option price is*

$$V(t) = e^{-r(T-t)} \boldsymbol{E}^{\mathbb{Q}}[g(S(T)) \mid \mathcal{F}_t] \tag{2.1.1}$$

where $\mathcal{F}_t$ is the filtration containing all information about stock S up to time t and r is the risk-free interest rate.

From an economical point of view this can be seen as an investor is risk-neutral, i.e. an investor do not need any extra payment for the risk she takes [9]. Thus, the expected return on the stock is the risk-free interest rate. Since investors are viewed as risk-neutral, $\mathbb{Q}$ is therefore referred to as the risk-neutral probability measure.

In particular, equation 2.1.1 gives us that when the payoff function is $g(S(T)) = S(T)$ we get that the option price simply is the value of the stock price today, i.e. $V(t) = S(t)$. This follows directly from the measurability and since the expected return of the stock under $\mathbb{Q}$ is the risk-free interest rate.

2.2 Bivariate options

Options consisting of two or more underlying stocks are called *rainbow options*. The price of these options are dependent of the dependence structure between the stocks since the fair option price still is the expected value. We define the fair price of a bivariate rainbow option as

Definition 2.2.1 *Let S_1 and S_2 be two stocks traded on a complete and arbitrage free market. Let t be the present time and T the time of maturity, then the price of an option with payoff function $g(S_1, S_2)$ is*

$$\begin{aligned} V(t, S_1, S_2) &= e^{-r(T-t)} \boldsymbol{E}^{\mathbb{Q}}[g(S_1(T), S_2(T)) \mid \mathcal{F}_t] \\ &= e^{-r(T-t)} \int_0^\infty \int_0^\infty g(x,y) f^{\mathbb{Q}}_{S_1,S_2}(x,y) dx dy. \end{aligned} \tag{2.2.1}$$

Here $f^{Q}_{S_1,S_2}$ is the joint probability distribution of the two stocks under the risk-neutral probability measure $\mathbb{Q}$ and as in the one-dimensional case $\mathcal{F}_t$ is the filtration containing all information about the two stocks.

The bivariate option with payoff function $g(S_1(T), S_2(T)) = S_1(T) + S_2(T)$ has the fair price of $S_1(t) + S_2(t)$ since the expectation is a linear function, thus the problem is reduced to the one-dimensional case.

2.3 Different bivariate payoffs

There are obviously numerous different payoff functions for rainbow options. The payoff function can be constructed to fit the clients specific needs. The options listed below are the ones we will focus on and investigate in this paper.

- One important rainbow option is the so called *Exchange option*. This option gives the holder the right to exchange stock 1 for stock 2. This means that the option will be exercised only if stock 2 is worth more than stock 1. Thus we get the following payoff function

$$g(S_1(T), S_2(T)) = \max\left(S_2(T) - S_1(T), 0\right).$$

- The second option we will examine is the *Maximum option*. It gives the holder the highest valued stock at expiration date. With this definition, the owner of the option is guaranteed a positive cash-flow at maturity except for the extreme event when both stocks have become worthless. The payoff function is

$$g(S_1(t), S_2(t)) = \max\left(S_1(T), S_2(T)\right).$$

- The third option is the call on the Maximum option, and it is called the *Call on max*. Here we essentially have the same option as before only now we also demand that the highest valued stock exceeds the strike price K. The Call on max option has the following payoff function

$$g(S_1(T), S_2(T)) = \max\left[\max\left(S_1(T), S_2(T)\right) - K, 0\right].$$

- The final option is the *Average spread option*. It is a so called path-dependent option, i.e. the payoff function depend on the path of the stocks. Thus, it requires the daily prices of both stocks for the entire life of the option. If the average spread between the two stocks exceeds the strike price the option is executed. The payoff function is

$$g(S_1(T), S_2(T)) = \max\left(\frac{1}{T-t}\sum_{i=t}^{T}(S_1(i) - S_2(i)), K\right).$$

Chapter 3

GARCH

The GARCH process (*Generalized Autoregressive Conditional Heteroskedasticity*) was first introduced by Tim Bollerslev in 1986 [1] and is as its name implies an extended version of Engle's ARCH process. It is one of the most well-known processes for conditional variances and it captures features such as mean reverting and time dependence. Empirical studies have shown that stock prices have these properties [9]. The following introduction to the GARCH model follows Bollerslev's work, however since this thesis is built on financial time series we will use financial terms.

3.1 GARCH(1,1)

Given a stocks price history $\{S_t\}_{i=0}^{n}$, its corresponding log returns, R_t, are defined as

$$R_t = \log(S_t) - \log(S_{t-1}), \quad t = 1, 2, \dots, n. \tag{3.1.1}$$

Here each t can, without loss of generality, be viewed as one trading day. The definition of the conditional variance, h_t, under the general GARCH(p, q) model is given by

Definition 3.1.1

$$\begin{aligned} R_t &= \epsilon_t \sqrt{h_t} \\ \epsilon_t &\sim N(0,1) \\ h_t &= \alpha_0 + \sum_{i=1}^{q} \alpha_i R_{t-1}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \\ &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 h_{t-i} + \sum_{i=1}^{p} \beta_i h_{t-i} \end{aligned}$$

where $t = 1, 2, \dots, n$.

The p is the the number of previous variance-terms affecting the next days conditional variance and q is the number of previous log returns. The constants α_1 and β are weights given to the previous information and α_0 is associated with the long-run mean[2]. In order for the process to make sense they are given the following constraints

$$\begin{aligned} \alpha_0 &> 0 \\ \alpha_1 &\geq 0 \\ \beta &\geq 0 \\ \alpha_1 + \beta &< 1. \end{aligned}$$

The last constraint is of great importance since if the weights sums up to one the variance process will explode.

3.2 Duan's model

We will use an alternative model, suggested by Duan in 1995 [6]. In this model the process of the asset's log returns are changed but the structure of the conditional variance is kept. We will call it the DGARCH process and will use the DGARCH$(1, 1)$, e.g. p and q are set to 1. The definition of this process is as follows

Definition 3.2.1 *Let r be the risk-free interest rate and $\lambda > 0$. Under the DGARCH process the log returns, R_t for $t = 1, 2, \dots, n$, are given by*

$$R_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \sqrt{h_t}\epsilon_t \tag{3.2.1}$$

$$\epsilon_t \sim N(0, 1)$$

$$h_t = \alpha_0 + \alpha_1\epsilon_{t-1}^2 h_{t-1} + \beta h_{t-1} \tag{3.2.2}$$

The constants, α_0, α_1 and β, have the same constraints as before. Duan developed this model in order to price options and hence we chose this model instead of the original GARCH model. There are plenty of other types of GARCH processes, including the Heston-Nandi [8] which also was developed in order to price options.

3.3 DGARCH under risk-neutral measure

In order to price options we require the risk-neutral measure $\mathbb{Q}$. To obtain $\mathbb{Q}$, Duan introduces the so called *locally risk-neutral valuation relationship*, or LRNVR, and gives the following definition of $\mathbb{Q}$

[2]Long-run mean $= \alpha_0/(1 - \alpha_1 - \beta)$

Definition 3.3.1 *A measure $\mathbb{Q}$ is said to satisfy the LRNVR if measure $\mathbb{Q}$ is mutually absolutely continuous with respect to measure $\mathbb{P}$ (real world). Under $\mathbb{Q}$, we have*

$$\boldsymbol{E}^{\mathbb{Q}}\left[\frac{S_t}{S_{t-1}} \mid \mathcal{F}_t\right] = e^r$$

and

$$Var^{\mathbb{Q}}(R_t \mid \mathcal{F}_t) = Var^{\mathbb{P}}(R_t \mid \mathcal{F}_t).$$

This means that the conditional variance in the consecutive time step is the same for both measures, thus we can estimate the parameters in Equation 3.2.2 under $\mathbb{P}$. With this definition Duan proves that under the local risk-neutral measure $\mathbb{Q}$, the DGARCH model from Definition 3.2.1 becomes

$$\begin{aligned} R_t &= r - \frac{1}{2}h_t + \sqrt{h_t}\epsilon_t^* \\ \epsilon_t^* &\sim N(0,1) \\ h_t &= \alpha_0 + \alpha(\epsilon_{t-1}^* - \lambda\sqrt{h_{t-1}})^2 + \beta h_{t-1}. \end{aligned} \tag{3.3.1}$$

Chapter 4

The Basic Concepts of Copulas

One tool to model the depencence between two random variables is to use the concept of copulas. The word *copula* origins from the Latin noun which means *a link* or *a tie*. The word was firstly used in mathematical statistics by Sklar in 1959. He introduced the term in a theorem on how there exists functions that joins one-dimensional distribution functions to form a multivariate distribution function.

In this chapter we present the basic concepts of copulas and we will follow the notations used in Nelsen [13], if not stated otherwise. We will restrict ourselves to the two dimensional case, however, the interested reader can turn to Nelsen for the generalization in n dimensions. The details and proofs of this chapter may also be found therein.

4.1 Definitions and Basic Properties

By having known marginal distributions, the goal with using a copula is to connect these marginal distributions to form a multivariate distribution. Primarily, the data is transformed into uniformly distributed random variables. Thus, we gain a dependence structure which depends on the uniformly distributed random variables. We start off this section by defining a two-dimensional copula, which we from now on only will refer to as a copula.

Definition 4.1.1 *A copula is a function $C : \boldsymbol{I}^2 \to \boldsymbol{I}$, where $\boldsymbol{I}$ is the unit interval, with the following properties*

1. *For every u, $v \in \boldsymbol{I}$*

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

2. *For every u, $v \in \boldsymbol{I}$*

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

3. *For every* u_1, u_2, v_1, v_2 *in* $\boldsymbol{I}$ *such that* $u_1 \leq u_2$ *and* $v_1 \leq v_2$

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

A function that fulfills Property 1 is named *grounded* and a function that fulfills Property 3 is said to be *2-increasing* [3].

Theorem 4.1.2 *Let* C *be a copula. For any* $v \in \boldsymbol{I}$*, the partial derivative* $\partial C(u,v)/\partial u$ *exists for almost all*[3] u*. The same is true for the partial derivative* $\partial C(u,v)/\partial v$*, where* $u \in \boldsymbol{I}$

Theorem 4.1.3 *Let* C *be a copula. If* $\partial C(u,v)/\partial v$ *and* $\partial^2 C(u,v)/\partial u \partial v$ *are continuous on* $\boldsymbol{I}^2$ *and* $\partial C(u,v)/\partial u$ *exists for all* $u \in (0,1)$ *when* $v = 0$*, then* $\partial C(u,v)/\partial u$ *and* $\partial^2 C(u,v)/\partial v \partial u$ *exist in* $(0,1)^2$ *and* $\partial^2 C(u,v)/\partial u \partial v = \partial^2 C(u,v)/\partial v \partial u$.

This leads us to the following definition [3] for the density function of a copula

Definition 4.1.4 *The density c(u, v) associated to a copula C(u, v) is*

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}.$$

One important theorem in statistics is *Sklar's Theorem*, as we mentioned earlier. It plays a central role in the theory of copulas and describes how copulas connect marginal distributions to their joint distribution.

Theorem 4.1.5 *(Sklar's Theorem). Let* F *be be a joint distribution function with margins* F_X *and* F_Y*. Then there exists a copula* C *such that* $\forall x, y \in \tilde{\mathbb{R}}$*, where* $\tilde{\mathbb{R}}$ *denotes the extended real line* $[-\infty, \infty]$*,*

$$F(x,y) = C(F_X(x), F_Y(y)). \tag{4.1.1}$$

If F_X *and* F_Y *are continuous, then* C *is unique; otherwise,* C *is uniquely determined on* $Ran(F_X) \times Ran(F_Y)$*, where* $Ran(F_X)$ *and* $Ran(F_Y)$ *denote the range of* F_X *and* F_Y *respectively. Conversely, if* C *is a copula and* F_X *and* F_Y *are distribution functions, then the function* F *is a joint distribution function with margins* F_X *and* F_Y *as defined in Equation 4.1.1.*

[3] In the sense of Lebesgue measure, thus the set for which this does not hold is a null set.

Since the second order derivative is defined for a copula we can, by using Sklar's theorem and the chain rule, obtain an expression for the density of a copula. It is given by

$$c\big(F_X(x), F_Y(y)\big) = \frac{f(x,y)}{f_X(x)f_Y(y)}. \tag{4.1.2}$$

By taking the inverse of each margin in Equation 4.1.1 we can instead obtain an expression for the copula in terms of the joint distribution function and the two margins. However, if a margin is not strictly increasing the inverse of the margin does not exist. Therefore, in those cases, the definition of a *quasi-inverse* of a distribution functions will be necessary to obtain an equation for the copula in terms of the joint distribution function and its margins.

Definition 4.1.6 *(Quasi-Inverse). Let F be a distribution function. Then a quasi-inverse of F is any function $F^{(-1)}$ with domain $\boldsymbol{I}$ such that*

1. $\forall t \in Ran(F)$,

$$F(F^{(-1)}(t)) = t$$

2. $t \notin Ran(F)$, *then*

$$F^{(-1)}(t) = \inf\{x \mid F(x) \geq t\} = \sup\{x \mid F(x) \leq t\}.$$

With this definition in mind we, by inverting 4.1.1, obtain

$$C(u,v) = F(F_X^{(-1)}(u), F_Y^{(-1)}(v)).$$

However, if we have strictly increasing margins the usual inverse exists and the copula is given by

$$C(u,v) = F(F_X^{-1}(u), F_Y^{-1}(v)). \tag{4.1.3}$$

For a two-dimensional sample an *empirical copula* can be obtained by the following definition

Definition 4.1.7 *(Empirical Copula). Let $\{(x_k, y_k)\}_{k=1}^n$ denote a sample of size n from a continuous bivariate distribution. The empirical copula is the function C_n given by*

$$C_n\Big(\frac{i}{n}, \frac{j}{n}\Big) = \frac{\textit{number of pairs } (x,y) \textit{ in the sample with } x \leq x_{(i)}, y \leq y_{(j)}}{n}.$$

4.2 Copula Families

In this section we will present some families of copulas that will be relevant for this paper. We will follow the notation used in Umberto et al. [3] if not stated otherwise.

The bivariate Gaussian copula

Definition 4.2.1 *The Gaussian copula C_{G,ρ_P} is defined as*

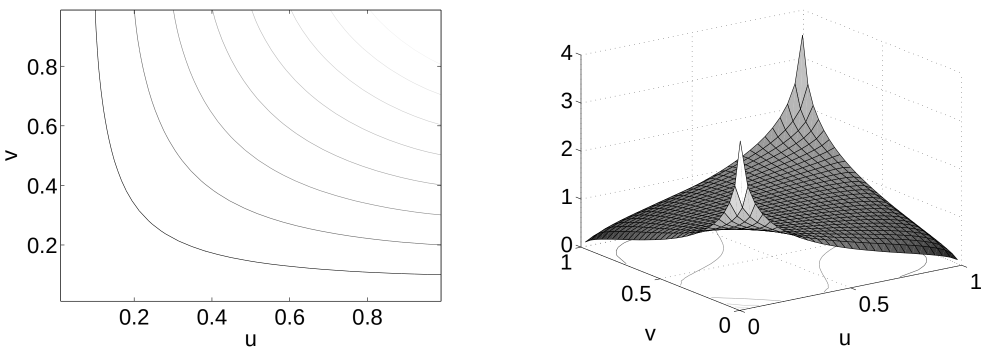
$$C_{G,\rho_P}(u,v) = \Phi_{\rho_P}\left(\Phi^{-1}(u), \Phi^{-1}(v)\right)$$

where Φ is the standard normal distribution function and Φ_{ρ_P} is the joint distribution function of a two-dimensional standard normal vector, with Pearson's linear correlation coefficient ρ_P.

Through the defintion we have the following expression for the Gaussian copula

$$\Phi_{\rho_P}\left(\Phi^{-1}(u), \Phi^{-1}(v)\right) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho_P^2}} \exp\left(\frac{2\rho_P st - s^2 - t^2}{2(1-\rho_P^2)}\right) ds dt.$$

Figure 4.1 displays a contour plot of the cumulative distribution function, as well as the probability density function for a Gaussian copula.



(a) Contour plot of cumulative distribution function.

(b) Probability density function.

Figure 4.1: The Gaussian copula with $\rho_P = 0.3$

The bivariate Student's t copula

In order to be able to define the Student's t copula we first need to remind ourselves of the univariate Student's t distribution function. Let $t_\nu : \mathbb{R} \to \mathbb{R}$ be the (central) univariate Student's t distribution, with ν degrees of freedom

$$t_\nu(x) = \int_{-\infty}^{x} \frac{\Gamma\big((\nu+1)/2\big)}{\sqrt{\pi\nu}\;\Gamma(\nu/2)}\left(1+\frac{s^2}{\nu}\right)^{-\frac{\nu+1}{2}} ds$$

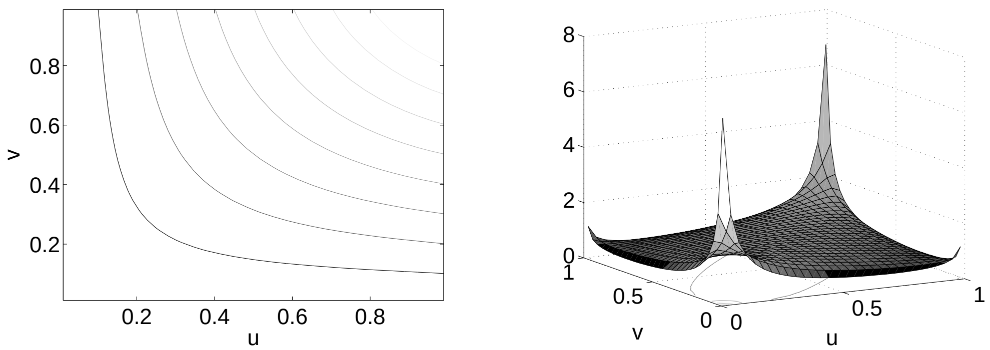
where Γ is the Gamma function. For $\rho_P \in [-1,1]$ and $t_{\rho_P,\nu}$ the bivariate distribution corresponding to t_ν is given by

$$t_{\rho_P,\nu}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} \frac{1}{2\pi\sqrt{1-\rho_P^2}}\left(1-\frac{2\rho_P st - s^2 - t^2}{\nu(1-\rho_P^2)}\right)^{-\frac{\nu+2}{2}} dsdt.$$

Definition 4.2.2 *The bivariate Student's t copula $C_{t,\rho_P,\nu}$ is defined as*

$$C_{t,\rho_P,\nu} = t_{\rho_P,\nu}\Big(t_\nu^{-1}(u), t_\nu^{-1}(v)\Big) = \int_{-\infty}^{t_\nu^{-1}(u)}\int_{-\infty}^{t_\nu^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho_P^2}}\left(1-\frac{2\rho_P st - s^2 - t^2}{\nu(1-\rho_P^2)}\right)^{-\frac{\nu+2}{2}} dsdt.$$

A contour plot of the cumulative distribution function and a probability density plot are presented in Figure 4.2. Remember that as $\nu \to \infty$ the Student's t distribution approaches the standard normal distribution. Thus, in this case the Student's t copula is equal to the Gaussian copula.



(a) Contour plot of cumulative distribution function.

(b) Probability density function

Figure 4.2: The Student's t copula with $\rho_P = 0.3$ and $\nu = 5$.

Archimedean Copulas

A special class of copulas are the *Archimedean Copulas.* Due to the ease of which these copulas can be constructed, the Archimedean copulas are very important in applications. Also, thanks to the many different families of copulas in this class, the Archimedean copulas can be applied in a wide range of areas [13].

Consider the continuous, decreasing and convex function $\phi : \mathbf{I} \to [0,\infty]$, and such that $\phi(1) = 0$. We call such a function ϕ a *generator.*

Definition 4.2.3 *The pseudo-inverse of ϕ is defined as*

$$\phi^{[-1]}(x) = \begin{cases} \phi^{-1}(x) & 0 \le x \le \phi(0) \\ 0 & \phi(0) \le x \le 0 \end{cases}$$

Furthermore, we have $\phi^{[-1]}\big(\phi(x)\big) = x$ for every $x \in \boldsymbol{I}$.

Definition 4.2.4 *Given a generator ϕ and its pseudo-inverse $\phi^{[-1]}$, an Archimedean copula C_A is generated as follows*

$$C_A(u,v) = \phi^{[-1]}\big(\phi(u) + \phi(v)\big).$$

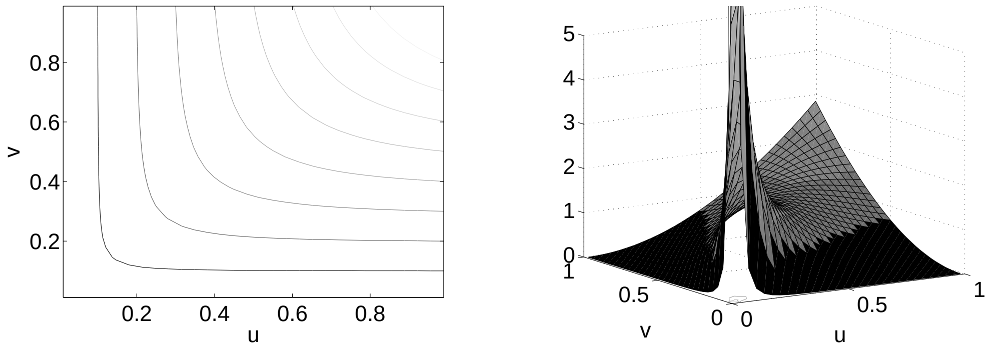
Whenever $\phi(0) = +\infty$ we call ϕ a *strict generator*. In this case, the pseudo-inverse equal the usual inverse, $\phi^{[-1]} = \phi^{-1}$. The *strict* Archimedean copula C_A is then given by $C_A(u,v) = \phi^{-1}\big(\phi(u) + \phi(v)\big)$.

In this thesis we will only consider the *one-parameter families of Archimedean copulas*. Such a copula is constructed by using a generator $\phi_\theta(x)$, for a real valued θ. Table 4.1 describes the Archimedean copulas we will use in this thesis.

Name	Generator $\phi_\theta(x)$	Range for θ	$C(u,v)$
Clayton	$\frac{1}{\theta}(x^{-\theta} - 1)$	$[-1,0) \cup (0,+\infty)$	$\max\big[(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, 0\big]$
Frank	$-\log \frac{\exp(-\theta x)-1}{\exp(-\theta)-1}$	$(-\infty,0) \cup (0,+\infty)$	$-\frac{1}{\theta}\log\left(1 + \frac{(\exp(-\theta u)-1)(\exp(-\theta v)-1)}{\exp(-\theta)-1}\right)$
Gumbel	$-(\log(x))^\theta$	$[1,+\infty)$	$\exp\big\{-[(-\log(u))^\theta + (-\log(v))^\theta]^{1/\theta}\big\}$

Table 4.1: Some Archimedian copulas

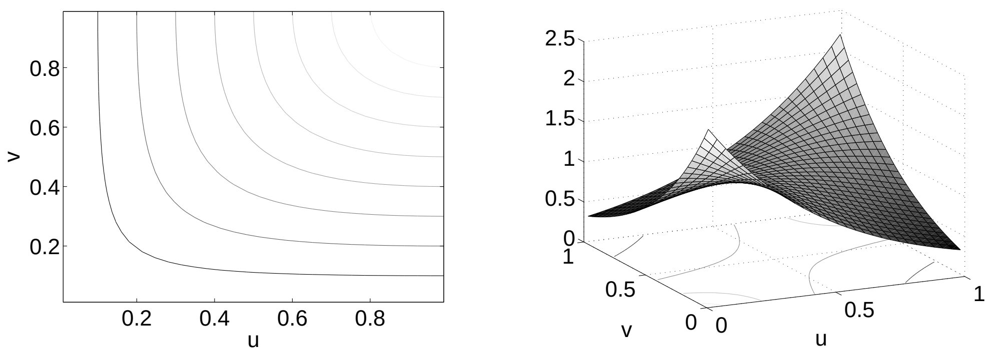
Figure 4.3, Figure 4.4 and Figure 4.5 displays the contour plot of the cumulative distribution function and the probability density function for the Clayton, Frank and Gumble copula respectively. The density plots of the Clayton and the Gumble copula have been cut off at 5 and 7 respectively for a better illustration of the behaviour of the probability densities.



(a) Contour plot of cumulative distribution function.

(b) Probability density function, cut off at 5.

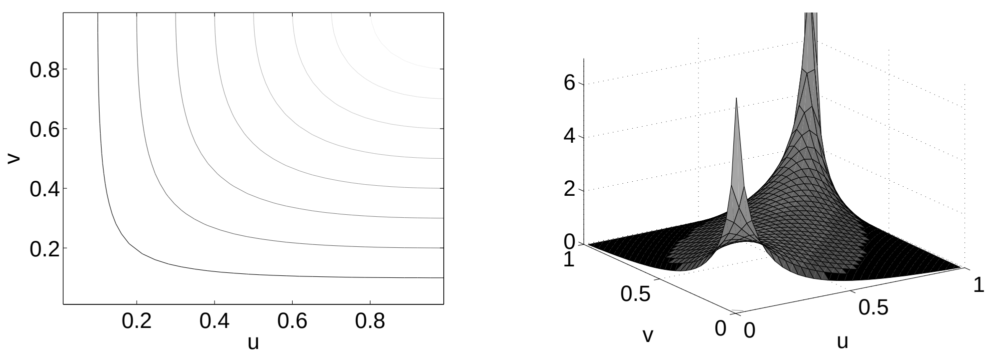
Figure 4.3: The Clayton copula with $\theta = 2$



(a) Contour plot of cumulative distribution function.

(b) Probability density function.

Figure 4.4: The Frank copula with $\theta = 2$



(a) Contour plot of cumulative distribution function

(b) Probability density function, cut off at 7

Figure 4.5: The Gumbel copula with $\theta = 2$

Chapter 5

Measures of Dependence

In this section we will present some different approaches on how to measure the dependence between random variables. One of the most well-known measures of dependence is *Pearson's linear correlation coefficient*, ρ_P. However, it turns out that this correlation mainly is useful when dealing with the Gaussian copula. Thus, for the other copulas presented in this thesis other measures of association are needed, such as *Kendall's* tau, τ_K, and *Spearman's* rho, ρ_S.

5.1 Pearson's measure of linear correlation

Consider two random variables X and Y, then the correlation coefficient ρ_P is given by

$$\rho_P = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}(X)\mathrm{var}(Y)}}.$$

For a normal distribution this measure is important since a zero correlation, in this case, implies independence between the random variables.

When using the correlation in copula theory one has to be a bit cautious. This due to the fact that the correlation depends on the choice of marginal distributions for the copula. This can be shown by considering

$$\begin{aligned} \mathrm{cov}(X,Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)\Big[c\big(F_X(x),F_Y(y)\big)-1\Big]dxdy \\ &= \int_{-1}^{1}\int_{-1}^{1} F_X^{-1}(u)F_Y^{-1}(v)\Big[c(u,v)-1\Big]dudv \end{aligned}$$

where the last equation is obtained by setting $x = F_X^{-1}(u)$ and $y = F_Y^{-1}(v)$, which means $dx = du/f_X(x)$ and $dy = dv/f_Y(y)$. Note here that $c(u,v)$ is the probability density function of the known copula.

5.2 Measures of Association

In this section we will follow the notation used in Nelsen [13], if not stated otherwise. Before presenting two other measures of association we need to clarify the meaning of the word *concordance*. We say that (x_i, y_i) and (x_j, y_j) are concordant if $x_i < x_j$ and $y_i < y_j$, or $x_i > x_j$ and $y_i > y_j$. We say that (x_i, y_i) and (x_j, y_j) are discordant if $x_i < x_j$ and $y_i > y_j$, or $x_i > x_j$ and $y_i < y_j$. This means that the pairs are concordant if $(x_i - x_j)(y_i - y_j) > 0$ and discordant if $(x_i - x_j)(y_i - y_j) < 0$.

Kendall's tau

Consider a random sample of n observations, $\{(x_i, y_i)\}_{i=1}^n$. Let n_c be the number of concordant pairs and let n_d denote the number of discordant pairs. For such a sample, that thus has $\binom{n}{2}$ distinct pairs, Kendall's tau for is defined as

$$\tau_K = \frac{n_c - n_d}{n_c + n_d} = (n_c - n_d) \Bigg/ \binom{n}{2} = \frac{(n_c - n_d)}{0.5n(n-1)}.$$

Thus, for a sample, Kendall's tau is the probability of the difference between concordant pairs and discordant pairs. When considering a population, Kendall's tau is defined similar to the one for a sample. Consider a vector (X, Y) of continuous random variables with joint distribution function F. Now, let(X_1, Y_1) and (X_2, Y_2) be i.i.d. random vectors, each with the joint distribution function F. Kendall's tau τ_K for a population is then defined as

$$\tau_K = \mathrm{P}\big[(X_1 - X_2)(Y_1 - Y_2) > 0\big] - \mathrm{P}\big[(X_1 - X_2)(Y_1 - Y_2) < 0\big].$$

It turns out that when X and Y are continuous random variables with a copula C, Kendall's tau is given by

$$\tau_K = 4 \int_{-1}^{1} \int_{-1}^{1} C(u, v) dC(u, v) - 1.$$

The interested reader can turn to Nelsen for the proof.

Spearman's rho

Another way to measure the concordance and the discordance of a population is to use Spearman's rho, ρ_S. Let (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) be three independent random vectors with a common joint distribution F and a copula C. Spearman's rho is given by

$$\rho_S = 3\Big(\mathrm{P}\big[(X_1 - X_2)(Y_1 - Y_3) > 0\big] - \mathrm{P}\big[(X_1 - X_2)(Y_1 - Y_3) < 0\big]\Big).$$

As for Kendall's tau we can get an expression for Spearman's rho in terms of a copula. For the continuous random variables X and Y whose copula is C, Spearman's rho is given by

$$\rho_S = \int_{-1}^{1}\int_{-1}^{1} uvdC(u,v) - 3.$$

Chapter 6

Method

We now have all the tools we need in order to price a bivariate rainbow option.

- The DGARCH model gives us the marginal distributions for each stock price.
- We use copula theory in order to join each marginal distribution into one bivariate joint distribution.
- The option pricing theory is used to get the fair price of a bivariate rainbow option.

From Equation 2.2.1 and Equation 4.1.2 we have, under the risk-neutral probability measure $\mathbb{Q}$,

$$\begin{aligned} V(t, S_1, S_2) &= e^{-r(T-t)} \int_0^\infty \int_0^\infty g(x,y) f^{\mathbb{Q}}_{S_1,S_2}(x,y)\,dxdy \\ &= e^{-r(T-t)} \int_0^\infty \int_0^\infty g(x,y) f^{\mathbb{Q}}_{S_1}(x) f^{\mathbb{Q}}_{S_2}(y) c(F^{\mathbb{Q}}_{S_1}(x), F^{\mathbb{Q}}_{S_2}(y))\,dxdy \end{aligned} \tag{6.0.1}$$

where g is the payoff function.

Given two vectors, R_1, R_2 containing the log returns for the two stocks, see Equation 3.1.1, we will now present a way to price a bivariate option using the DGARCH model, copulas and Monte Carlo simulations. We have chosen to use a numerical method to calculate the option price due to the complexity of the DGARCH model and copulas. This procedure has been inspired by Chiou and Tsay [4].

1. For each vector of log returns, use maximum likelihood to estimates the parameters α_0, α_1, β_1 and λ in Equation 3.2.2. Thus, the problem is to maximize

$$\log \mathcal{L}(\mu, h_t) = -\frac{n}{2}\left[\log(2\pi) + \frac{1}{n}\sum_{t=1}^{n}\left[\log(h_t) + \frac{(x_t-\mu)^2}{h_t}\right]\right]$$

with respect to the parameters. In the DGARCH framework the x_t's are the last term in Equation 3.2.1, namely $\epsilon_t\sqrt{h_t}$. According to the theory the x_t's should have zero mean, thus reducing the problem considerably. The maximization can be done by using standard optimization tools, e.g. $fmincon$ in Matlab.

2. Use the estimated parameters to calculate h_t and ϵ_t in Equations 3.2.2 and 3.2.1 for each stock.

3. In order to fit the data to copulas, we need to transform the data into uniformly distributed random variables. Thus we transform the ϵ_i, $i = 1, 2$, obtained in step 2 for each stock into uniformly distributed variables, by $u_i = \Phi(\epsilon_i)$, where Φ is the standard normal cumulative distribution function.

4. Use maximum likelihood to fit a copula to $[u_1, u_2]$. That is, for the transformed data $\{u_{1,t}, u_{2,t}\}_{t=1}^{n}$, estimate the copula parameters θ_{copula}

$$\hat{\theta}_{copula} = \mathrm{ArgMax}_{\theta_{copula}} \sum_{t=1}^{n} \log \Big[c\big((u_{1,t}, u_{2,t}); \theta_{copula}\big)\Big]$$

where θ_{copula} are the parameters for the copula function C, and where c is the probability density function for the given copula.

5. Now it is time to calculate the option price using Monte Carlo simulations. First generate a sample $\{u^*_{1,t}, u^*_{2,t}\}_{t=1}^{T}$ from a uniform marginal distribution from the chosen copula, see Appendix A. Here T is the time to maturity for the option.

6. For each time step transform the generated margins to standard normal margins, in the risk-neutral world, by

$$\epsilon^*_{i,t} = \Phi^{-1}(u^*_{i,t}), \quad for\ i = 1, 2.$$

7. Using the $\epsilon^*_{i,t}$ calculate the conditional variances by Equation 3.3.1 and the parameters estimated is step 1. The two future stock prices at time T are

$$S_i(T) = S_i(0) \exp\Big[rT - 0.5 \sum_{t=1}^{T} h_{i,t} + \sum_{t=1}^{T} \sqrt{h_{i,t}}\epsilon^*_{i,t}\Big], \quad for\ i = 1, 2.$$

8. Repeat step 5 to 7 for N runs and let $g(S_1(T), S_2(T))$ be the payoff function for the option. Thus we obtain the Monte Carlo option price as

$$V(t) = \frac{e^{-r(T-t)}}{N} \sum_{i=1}^{N} g(S_{1,i}(T), S_{2,i}(T)).$$

Chapter 7

Analysis of the data

For the bivariate rainbow options, the two stocks chosen were the food processing company Kraft Foods (KFT) and the technology company Hewlett-Packard (HPQ), both traded on the New York Stock Exchange. Both stocks are components of the Dow Jones Industial Average, which is a stock market index containing 30 of the major companies in America. The data was collected from the Internet at Yahoo Finance[4].

We consider a one year time period from 1 June 2010 to 31 May 2011. We use daily close prices in USD that are adjusted for dividends and splits. During this time period there were 253 trading days. We choose a one year period since it is a rule of thumb in financial applications to use the same data length as the length of the later generated sample, i.e. the time to maturity for the option. The one year interest rate for this period was set to be approximately equal to the yield to maturity for a one year government bond in the U.S. Thus, the yearly risk-free interest rate was set to 0.25%.

Figure 7.1 and Figure 7.2 displays the stock prices and the log returns for KFT and HPQ respectively, over the chosen time period. Note that HPQ experiences some strong increase and decrease in the stock price from one day to another. This is especially clear when observing the log returns for HPQ. One possibility for these spikes could be that HPQ are launching a new product or technology. KFT is less volatile in that sense. The log returns for KFT has a mean of 9.06×10^{-4} and a standard deviation of 0.0087. The mean for the log returns for HPQ is -7.56×10^{-4} and the standard deviation is 0.0171.

In Table 7.1 different measures of dependence between the log returns of KFT and HPQ are displayed. The table also displays the p-values for the null hypothesis that there is no autocorrelation. Thus, we clearly see that the two different stocks

[4] http://finance.yahoo.com/

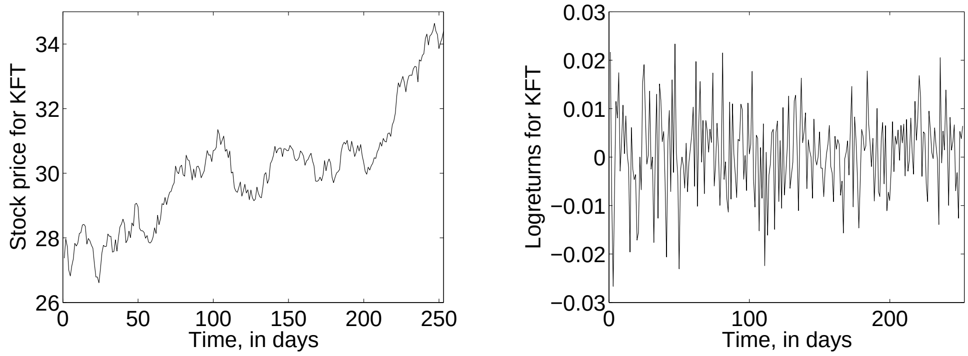


Figure 7.1: Daily stock prices and log returns for KFT over the sample period.

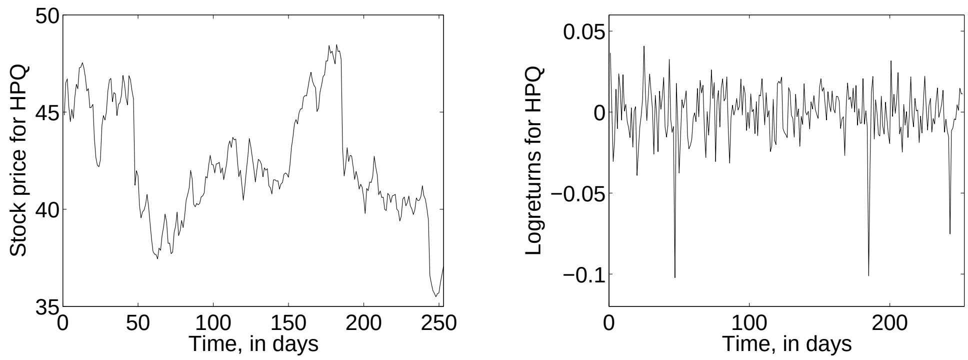


Figure 7.2: Daily stock prices and log returns for HPQ over the sample period.

are statistically dependent.

Figure 7.3 displays the histograms for the log returns of KFT and HPQ. In many elementary models in finance the log returns are assumed to be normally distributed. However, even if one can not conclude whether or not the log returns for the stocks are normally distributed, it is common knowledge that this most often is not the case. Thus, a more sophisticated model, such as DGARCH, is necessary for simulating future stock prices.

	Spearman's rho, ρ_S	Kendall's tau, τ_K	Pearson's rho, ρ_P
Correlation	0.231	0.158	0.162
P-value	2.104×10^{-4}	1.872×10^{-4}	0.010

Table 7.1: Measures of dependence between log returns of KFT and log returns of HPQ.

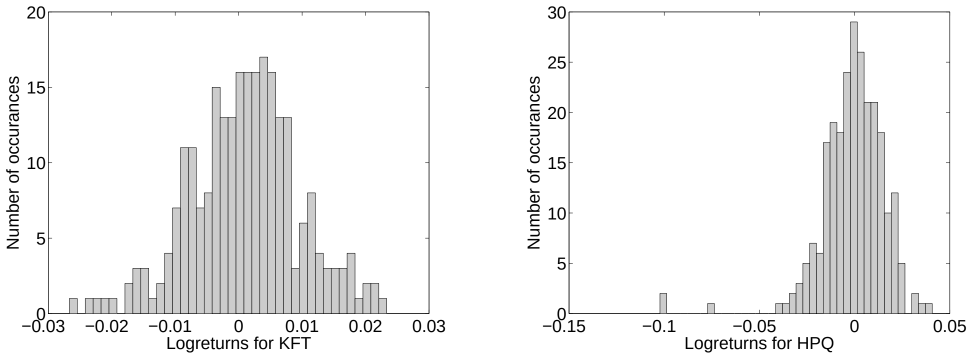


Figure 7.3: Histograms of the log returns for KFT and HPQ with 40 bins.

Chapter 8

Analysis of the DGARCH model

In this section we will implement the method described in Chapter 6 in order to obtain the parameters in the DGARCH model. Tests are carried out in order to validate the model.

8.1 The parameters

The parameters obtained when using MLE on the DGARCH model for the stocks KFT and HPQ are given in Table 8.1. As we can see, the parameters fulfills the required constraints in the DGARCH model.

Stock	α_0	α_1	β_1	λ
KFT	3.43×10^{-7}	1.44×10^{-2}	0.978	0.114
HPQ	2.85×10^{-4}	8.72×10^{-3}	1.05×10^{-4}	4.72×10^{-6}

Table 8.1: DGARCH parameters

8.2 The residuals

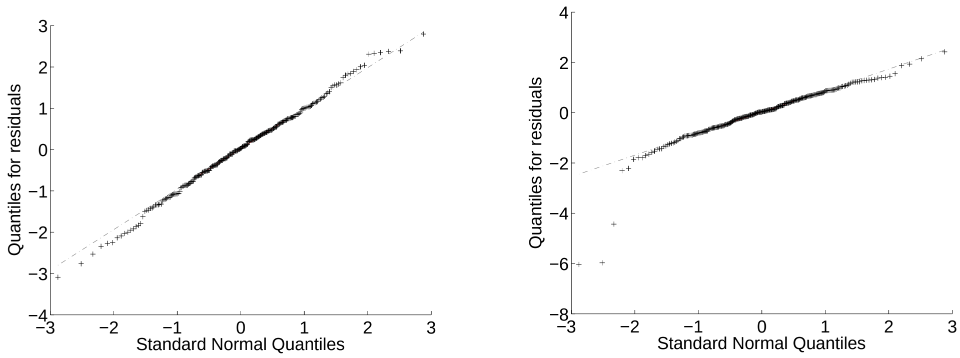
The residuals in the DGARCH model, i.e. ϵ_t, should in theory follow a standard normal distribution. Furthermore they should also have independent increments.

The *Kolmogorov-Smirnov test* or K-S test, see Appendix B, are applied for each stocks residual with a 5% significance level. The residuals empirical distribution are tested against the standard normal distribution and the p-values and test statistics are presented in Table 8.2.

	KFT	HPQ
p-value	0.926	0.209
Test statistic	0.034	0.066

Table 8.2: P-values and K-S distance for standard normality of the residuals.

Clearly, we can not reject the null hypothesis that they are indeed standard normal distributed. For a graphical interpretation the QQ-plots for the residuals are also displayed, see Figure 8.1.



(a) KFT's residuals (b) HPQ's residuals

Figure 8.1: QQ-plots for the residuals of the stocks.

The independence property of the increments is tested with the so called *Ljung-Box test*, see Appendix B. The null hypothesis is that there is no autocorrelation. For standard normal variables, this implies independent increments. Again we have a 5% significance level and, as can be seen in Table 8.3, we can not reject the null hypothesis that they are independent.

	KFT	HPQ
p-value	0.621	0.527
Test statistic	17.515	18.926

Table 8.3: P-values for independence of the residuals using Ljung-Box test.

Chapter 9

Fitting a copula to the data

In this section we will fit our data to the copulas of our choice. In order to validate the parameters obtained in the MLE fit we will combine three approaches to determine how well the different copulas fit the data. This due to the technical challange in preforming a bivariate, two-sample Kolmogorov-Smirnov test. Thus, we will start of by using a graphical approach. Then we will examine the distances between the empirical copula and the estimated copula. Finally, we will examine how well each copula manages to preserve the different measures of dependence from the data.

First, the standard normal residuals ϵ_t obtained from the DGARCH, are transformed into uniformly distributed variables. We use maximum likelihood, as presented in Chapter 6, to estimate the parameters for the different copulas. The parameters obtained are presented in Table 9.1.

Copula	Gaussian	Student's t	Clayton	Frank	Gumbel
Parameter	0.152	$(0.271, 10.991)$	0.070	1.670	1.196

Table 9.1: The estimated parameters for the diffrerent copulas.

In Figure 9.1 we have illustrated how the cumulative distribution function for each copula behaves in comparison to the empirical copula for the data. From these contour plots, the Student's t copula and the Frank copula are the best fit. Also, the estimated Clayton copula seems to be the copula that is furthest away from the empirical copula.

In the next step we examine the distances between the empirical copula and the estimated copulas. We do this by finding the maximum difference between the copulas in the same way as when using a two-sample K-S test, see Appendix B.

The distance is given by

$$D = \max(|C(u,v) - C_{emp}(u,v)|)$$

where C is the estimated copula in each point (u,v) and C_{emp} is the empirical copula in (u,v).

Copula	Gaussian	Student's t	Clayton	Frank	Gumbel
Distance	0.073	0.069	0.080	0.069	0.070

Table 9.2: The distances between empirical copula and estimated copulas.

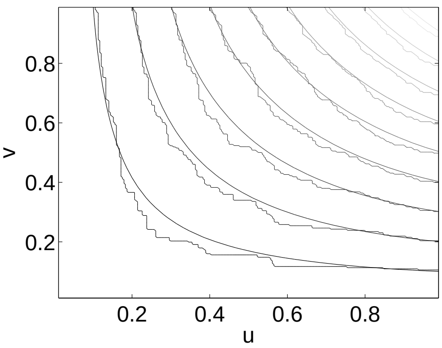
As one could already see in the countour plots, the K-S distances gave us that the Student's t copula and the Frank copula are the estimated copulas closest to the empirical copula and the Clayton copula is the one that is furthest away from the empirical copula.

However, the question wheather or not the copula manages to capture the dependence between the log returns remains. We examine this by observing the measures of dependence, Spearman's rho, Kendall's tau and Pearson's rho as presented in Table 9.3. These measures are calculated on the residuals of the data.

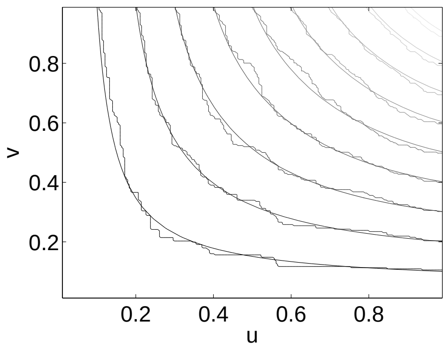
	Data	CI	Gaussian	Student's t	Clayton	Frank	Gumbel
ρ_S	0.230	[0.107, 0.355]	0.145	0.255	0.051	0.268	0.241
τ_K	0.156	[0.069, 0.244]	0.097	0.174	0.034	0.181	0.164
ρ_P	0.152	[-0.037, 0.329]	0.152	0.269	0.056	0.256	0.260

Table 9.3: Measures of depencence for the residuals and different copulas.

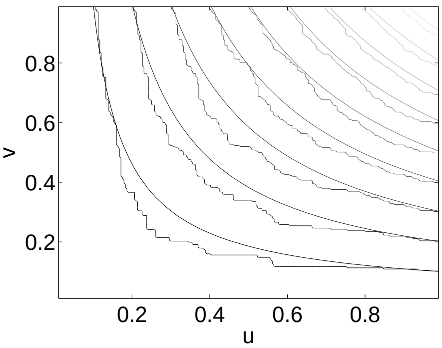
In theory, for the Gaussian copula the linear correlation ρ_P captures the dependence between the variables. In Table 9.3 we see that this is the case, since we for the Gaussian copula gained the same ρ_P for the residuals as for the estimated copula. However, for the other copulas the linear correlation will not capture the entire dependence, thus for those copulas ρ_S and τ_K will be of greater importance. We also notice that these measures of dependence are not accurate for the Gaussian copula in comparison to the data. The Clayton copula, once again, seems to be the least good fit, since it for two measures fall outside the confidence interval. For the Student's t, Frank and Gumbel copula ρ_P and ρ_S are of the same magnitude as for the residuals of the data and it seems like these copulas are a good fit.



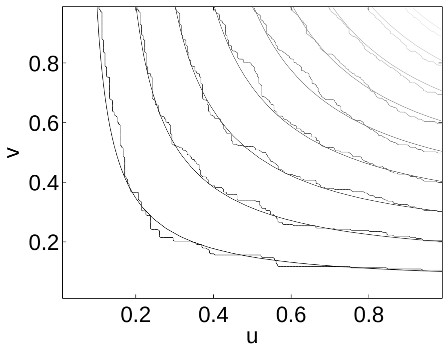
(a) Gaussian copula



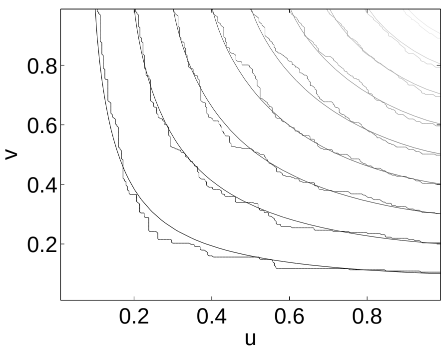
(b) Student's t copula



(c) Clayton copula

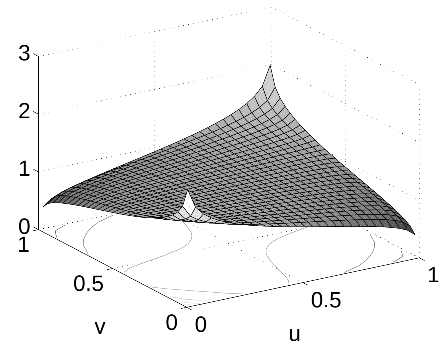


(d) Frank copula

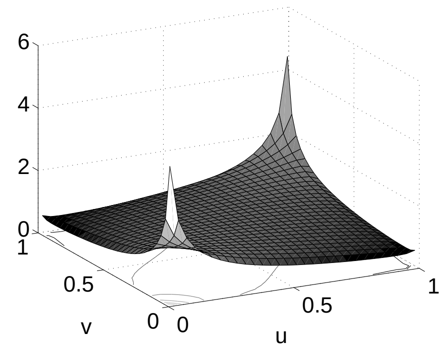


(e) Gumbel copula

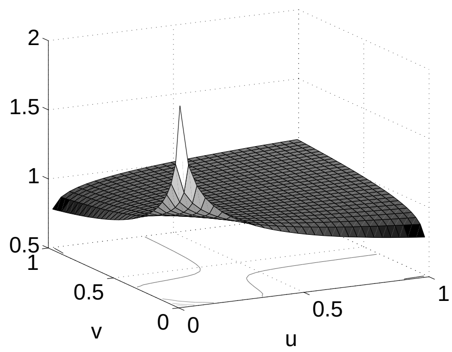
Figure 9.1: Contour plots for the estimated cumulative distribution functions of the copula and the empirical copula. The smooth contour lines represent the estimated copula and the crooked lines represent the empirical one.



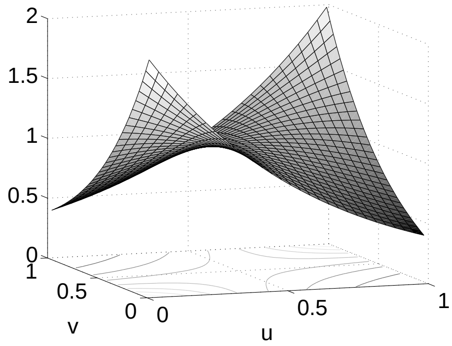
(a) Gaussian copula



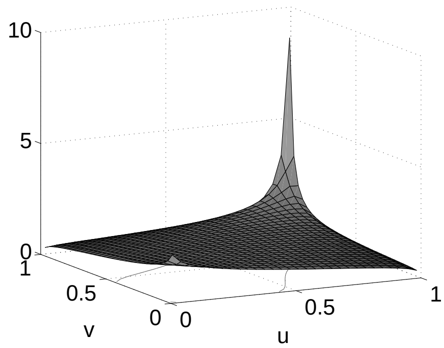
(b) Student's t copula



(c) Clayton copula



(d) Frank copula



(e) Gumbel copula

Figure 9.2: The probability density functions for the different copulas, with the estimated parameters.

Chapter 10

Simulating future stock prices

Using the DGARCH parameters that where estimated with MLE and presented in Table 8.1, together with the fitted copulas for the residuals, Table 9.1, we can simulate future stock prices and the corresponding conditional variances in the risk-free world. In Figure 10.1 and Figure 10.2, at the right hand side of the vertical line, possible stock paths and their daily conditional variances are shown using simulated data from the fitted Gaussian copula. At the left hand side of the vertical line the historical stock price path and its corresponding conditional variance are show. The same is displayed in Figure 10.3 and Figure 10.4, but with the Frank copula. It is simulations like these that will be used in the Monte Carlo simulation of the option price.

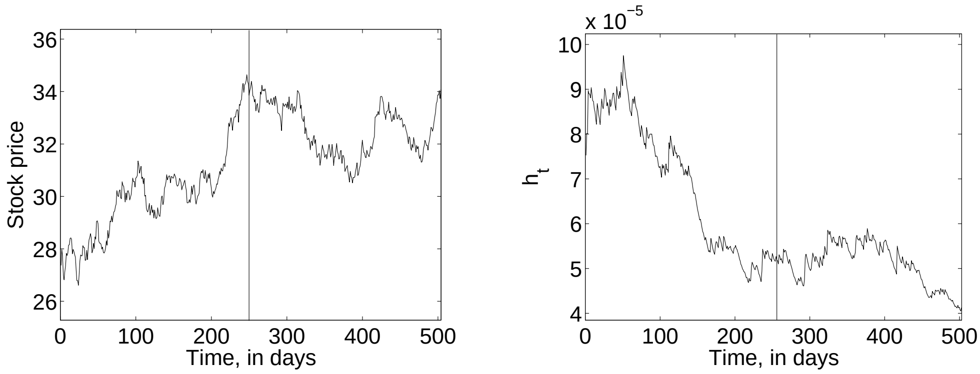


Figure 10.1: Simulated future stock price and conditional variance for KFT, using a Gaussian copula.

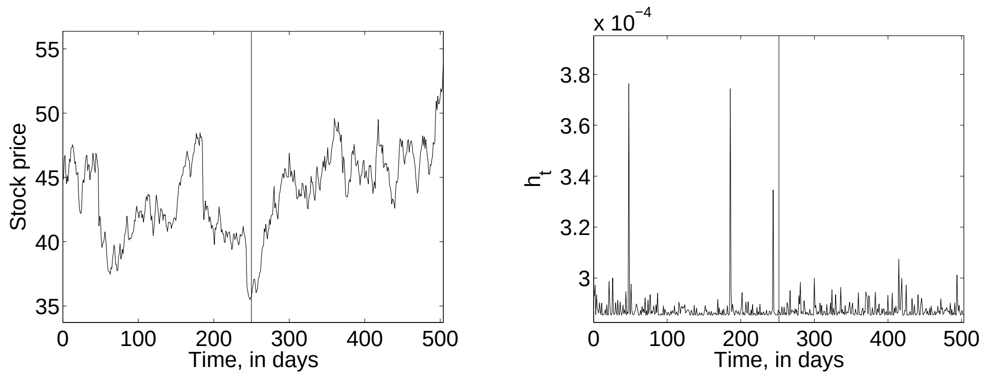


Figure 10.2: Simulated future stock price and conditional variance for HPQ, using a Gaussian copula.

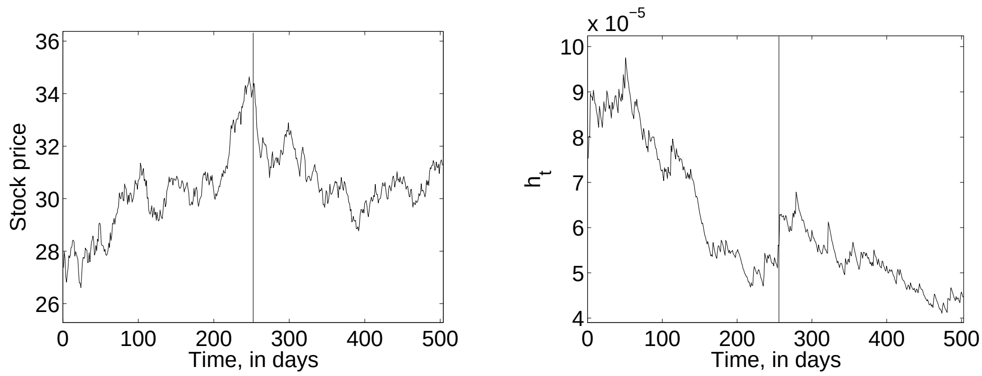


Figure 10.3: Simulated future stock price and conditional variance for KFT, using a Frank copula.

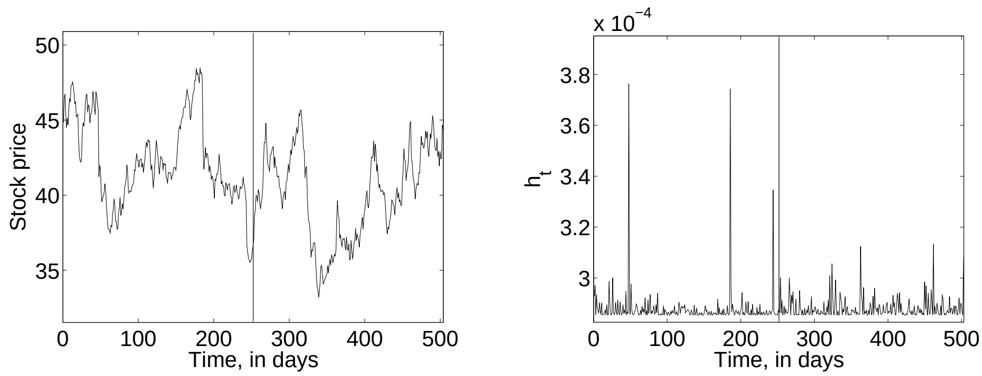


Figure 10.4: Simulated future stock price and conditional variance for HPQ, using a Frank copula.

Chapter 11

Calculating the option price

We have now reached the point where we are able to calculate option prices for the stocks of KFT and HPQ. First, however, we would like to investigate whether or not our model generates correct option prices. In previous sections we have noted that we get good fits to our model, but how well does our model perform when calculating option prices?

Before pricing options with more advanced payoff functions, we, as a final check, find the fair price for the option presented in Chapter 2.2. This is done since this is the only option for which we have a known price. We remind the reader that for this bivariate option, the payoff function at maturity T is $g(S_1(T), S_2(T)) = S_1(T) + S_2(T)$. The real option price today, that is in time t, is then $S_1(t) + S_2(t)$. For the KFT stock and the HPQ stock the real option price is \$71.45. If we obtain estimated option prices close to this real price, the model works. In Table 11.1 we present the percental differences between the real option price and the estimated option prices gained from Monte Carlo simulations.

Gaussian	Student's t	Clayton	Frank	Gumbel
2.63×10^{-4}	-2.01×10^{-4}	-3.57×10^{-4}	-1.43×10^{-4}	0.57×10^{-4}

Table 11.1: Percental differences in estimated option price with different copulas.

Clearly, the percental differences are very small, thus, our model seems to be fitting the theoretical option price well. These option prices were estimated using 100 000 Monte Carlo simulations. When performing the calculations we noticed that the model requires a large amount of simulations and a minimum of 50 000 iterations should be used.

The different types of options presented in Chapter 2.3 were calculated, again using 100 000 iterations. The result are presented in Table 11.2. We can see that the option prices when using the Student's t, Frank and Gumbel copulas are very close to eachother. Since the Gaussian and the Clayton copulas gave a worse fit for the data in comparison to the other three copulas, as we saw in Chapter 9, they probably will give a larger error in the option prices examined in Table 11.2.

	Gaussian	Student's t	Clayton	Frank	Gumbel
Exchange	5.38	5.20	5.52	5.23	5.22
Maximum	39.77	39.59	39.91	39.61	39.61
Call on max	5.13	5.06	5.19	5.07	5.06
Average spread	2.32	2.22	2.40	2.23	2.22

Table 11.2: Option prices using different copulas

Chapter 12

Options prices using other stocks

In order to ensure the quality of our model, we will now examine the results we obtain when choosing other underlying stocks for our options. Additional stocks were chosen from the components of Dow Jones Industial Avarage, but also from OMX Stockholm 30, OMXS30. This is a stock market index of the 30 most-traded stocks on the Stockholm Stock Exchange. The Swedish data was also collected daily[5] and is stated in SEK. For the American stocks we examine two different time periods, 1 January 2006 to 31 December 2006 and also 1 June 2010 to 31 May 2011. The number of trading days were 251 and 253 respectively. For the Swedish stocks we chose only to look at the time period 1 June 2010 to 31 May 2011, which was 255 trading days. In Table 12.1 we summarize the chosen stocks.

Stock	Label	Industry	Index
Alcoa	AA	Aluminium	Dow
Atlas Copco - stock A	AC	Manufacturing	OMXS30
Ericsson - stock B	ER	Telecommununications	OMXS30
General Electric Company	GE	Conglomerate	Dow
The Coca-Cola Company	KO	Beverage	Dow
Svenska Handelsbanken	SH	Financial Services	OMXS30

Table 12.1: Additional stocks used for option pricing.

For all the stocks, we could not reject the null hypothesis regarding the residuals in the DGARCH model, that is that the residuals are i.i.d. standard normal. In Table 12.2 we present the K-S distances between the empirical copula and the

[5]This data was collected at http://www.nasdaqomxnordic.com/

estimated copula for the American stocks during 2010-2011. In Appendix D the results for the Swedish data and the other time period for the American stocks can be found.

In Table 12.3 we present the measures of dependence for the residuals of the data of AA and GE. The measures of dependence for the other stock pairs can be found in Appendix C. We get similar results for all pairs, namely that the Gaussian copula corresponds well with the linear correlation ρ_P and that the Clayton copula underestimates the dependence between the stocks. The Clayton copula is in fact the only copula that show tendencies to fall outside the confidence interval. This is the same result we found for KFT and HPQ.

We will start with finding the fair option price of the bivariate rainbow option with payoff function $g\big(S_1(T), S_2(T)\big) = S_1(T) + S_2(T)$, at maturtity T, as we did in Chapter 11. In Table 12.4, the percental difference between the theoretical and the estimated option price are presented. The tables for the other time period and stock market can be found in Appendix E. We used 50 000 Monte Carlo simulations when estimating the option price. For the examined stocks, we once again receive small differences which indicated that the model works well. However, there are no conclusive result regarding whether or not the Clayton and the Gaussian copula are less accurate than the other copulas.

Stocks	Gaussian	Student's t	Clayton	Frank	Gumbel
AA & GE	0.065	0.060	0.100	0.065	0.065
AA & KO	0.099	0.096	0.128	0.093	0.090
GE & KO	0.089	0.083	0.118	0.085	0.083

Table 12.2: The maximum distance between the estimated copula and the empirical copula for the American stocks in 2010-2011.

	Data	CI	Gaussian	Student's t	Clayton	Frank	Gumbel
ρ_S	0.611	[0.525, 0.701]	0.577	0.593	0.438	0.649	0.530
τ_K	0.443	[0.371, 0.514]	0.407	0.425	0.305	0.463	0.375
ρ_P	0.596	[0.484, 0.706]	0.596	0.617	0.461	0.619	0.554

Table 12.3: Measures of dependence for the residuals of the data and different copulas for the AA and GE stocks in 2010-2011.

As we did for KFT and HPQ, we priced the bivariate options with more advanced payoff functions. The result for the option prices with AA and GE as the underlying assets are presented in Table 12.5. For the Gaussian, the Student's t and the Frank copula the option prices are very close to eachother. In general, this was the result we obtained for the other stock pairs, see Appendix E.

Stocks	Gaussian	Student's t	Clayton	Frank	Gumbel
AA & GE	2.64×10^{-4}	1.74×10^{-3}	-2.43×10^{-4}	9.70×10^{-4}	7.82×10^{-5}
AA & KO	2.42×10^{-5}	8.40×10^{-4}	-1.34×10^{-4}	-3.58×10^{-4}	7.68×10^{-4}
GE & KO	2.19×10^{-4}	-4.28×10^{-4}	-6.77×10^{-4}	-2.00×10^{-4}	1.63×10^{-4}

Table 12.4: Percental differences in real and estimated option prices with different copulas for the American stocks in 2010-2011.

	Gaussian	Student's t	Clayton	Frank	Gumbel
Exchange	3.06	3.03	3.23	3.02	3.12
Maximum	19.82	19.80	20.00	19.80	19.87
Call on max	2.58	2.56	2.69	2.56	2.60
Average spread	0.85	0.83	0.99	0.83	0.89

Table 12.5: Option prices for AA and GE 2010-2011.

Chapter 13

Conclusions

In this thesis we have studied rainbow options that depend on two underlying stocks. We have modelled the stock prices using a special type of GARCH process, that we have chosen to call the DGARCH model. In order to estimate the depencende structure between the stocks we have adapted copula theory. We have chosen to examine the Gaussian, the Student's t, the Clayton, the Frank and the Gumbel copula. Various tests have been carried out in order to validate the suggested model. Data from different stock markets and time periods were tested in the model. In order to price the options we have adapted a Monte Carlo approach.

We have found evidence that the DGARCH model fits well to our data, in the sense that we could not reject that the residuals of the DGARCH model were i.i.d. standard normal. This is a crusial property since we need these distributions of the residuals for our estimation of the copula function.

Throughout our copula estimation, using maximum likelihood, we found that the Clayton copula was furthest away from the empirical copula distribution when considering the Kolmogorov-Smirnov distance. However, one can argue whether or not this results are due to the fact that the data sets were fairly small. In general, one needs large data sets in order to obtain accurate estimations of bivariate distributions. Since we only had access to daily stock prices, we could not use larger data sets without violating the rule of thumb that is to only to use historical data for the same time length as the future prediction. We also found that the Clayton copula did not manage to capture the measures of dependence as well as the other copula functions. This was true except for one case where the Gumbel copula was slightly less accurate. For our data, the Clayton copula showed tendencies to underestimate the measures of dependence.

No matter what copula we chose, we got very satisfactory results when pricing the option with the most basic payoff function. Despite the indications that the

Clayton copula did not manage to capture the dependence structure of the data as well as the other copulas, it still gives very pleasing results. This did not come as a surprise since the payoff function itself did not consider the depencence stucture between the stocks. Therefore, when pricing options with more advanced payoff functions we had strong reason to believe that the Clayton copula might not be as accurate as the other copula functions.

We chose to examine four options with more complicated payoff functions, namely the Exchange, the Maximum, the Call on Max and the Average Spread option. The payoff function where we suspected that the option price was most sensitive to the choice of copula was for the path-dependent Average Spread option. This since the estimated option price depends on the simulated stock prices for each day. We saw indications that this was in fact the case. Most often the price difference between the options when using different copulas was higher for the Average Spread option, in comparison to the other options. For the other options we saw no clear trends in the behaviour of the option price. In most cases, when using the Clayton copula we obtained option prices that differed slightly more, in comparison to the other copulas.

Since our Monte Carlo simulation require a vast amount of iterations the procedure quickly becomes computationally heavy. If one wants to price rainbow options for more than two stocks the computation might in fact become too time consuming, thus useless for practical purposes.

Thus, when pricing an option with a fairly simple payoff function the price is less affected by the choice of copula. However, as soon as the payoff function becomes more complicated the choice of copula is of greater importance and one needs to be more careful when choosing the copula that describes the dependence structure.

Appendix A

Generate random variate from a copula

For computational resons one might be interested in generating random draws from a copula. For the copulas presented in this paper, the procedure is quite straight forward. Throughout this section we will follow the notation used in [7].

Algorithm for the Gaussian copula

1. Find the Cholesky decomposition[A] A of R (linear correlation matrix)
2. Simulate n independent random variates $z_1, ..., z_n$ from a standard normal distribution.
3. Set $\mathbf{x} = A\mathbf{z}$.
4. For $i = 1, ..., n$, set $u_i = \Phi(x_i)$.
5. Then $(u_1, ..., u_n)^T \sim C_{Gaussian}$

Algorithm for the Student's t copula

1. Find the Cholesky decomposition A of R.
2. Simulate n independent random variate $z_1, ..., z_n$ from a standard normal distribution.
3. Simulate a random variate s from χ^2_ν independent of $z_1, ..., z_n$.
4. $\mathbf{y} = A\mathbf{z}$.

[A]"The Cholesky decomposition of R is the unique lower-triangular matrix L with $LL^T = R$" [7]

5. $\mathbf{x} = \frac{\sqrt{\nu}}{\sqrt{s}}\mathbf{y}$.

6. For $i = 1, ..., n$, set $u_i = t_\nu(x_i)$.

7. Then $(u_1, ..., u_n)^T \sim T_{\rho,\nu}$

Algorithm for an Archimedean copula

Consider an Archimedean copula with generator ϕ.

1. Simulate two independent random variates s and q from a uniform distribution on $[0, 1]$.

2. Set $w = K_C^{-1}(q)$, where $K_C(t) = t - \frac{\phi(t)}{\phi'(t)}$.

3. Set $u = \phi^{[-1]}(s\phi(w))$ and $v = \phi^{[-1]}\big((1-s)\phi(w)\big)$.

Appendix B

Goodness of Fit

B.0.1 Kolmogorov-Smirnov Test for a single sample

A *Kolmogorov-Smirnov test* for *a single sample* is used to compare a data set $F(x)$ to a known cumulative distribution function $G(x)$. The null hypothesis it that $x \sim G$, and the Kolmogorov-Smirnov statistic is given by [11]

$$D_{KS_1} = \max\big(|F(x) - G(x)|\big).$$

B.0.2 Kolmogorov-Smirnov Test for two samples

The Kolmogorov-Smirnov test can also be used to compare the distributions of *two different samples* with cumulative distribution functions $F_1(x)$ and $F_2(x)$ respectively. The null hypothesis is that the two samples come from the same distribution and the statistic is give by [14]

$$D_{KS_2} = \max\big(|F_1(x) - F_2(x)|\big).$$

Note that this does not test whether or not the two samples come from a known cumulative distribution function $G(x)$. It merely tests how similar the two samples are distributed.

B.0.3 Ljung-Box Test for independence

Using the *Ljung-Box test* we can test the null hypothesis that the residuals of a sample have an autocorrelation that is zero. In this case, for normally distributed residuals, we get that the residuals are independent [2]. The Ljung-Box statistic is given by [12]

$$Q = N(N+2)\sum_{k=1}^{M}\frac{\rho_k^2}{N-k}$$

where N is the sample size, M is the number of autocorrelation lags and ρ_k is the autocorrelation at lag k. Under the null hypothesis the statistic is asymptotically $\chi^2(M)$.

Appendix C

Tables of the measures of dependence

	Data	CI	Gaussian	Student's t	Clayton	Frank	Gumbel
ρ_S	0.395	[0.280, 0.512]	0.401	0.399	0.248	0.460	0.369
τ_K	0.279	[0.193, 0.363]	0.275	0.278	0.168	0.317	0.255
ρ_P	0.418	[0.286, 0.554]	0.417	0.419	0.268	0.439	0.393

Table C.1: Measures of dependence for the residuals of the data and different copulas for the AA and KO stocks in 2010-2011.

	Data	CI	Gaussian	Student's t	Clayton	Frank	Gumbel
ρ_S	0.460	[0.355, 0.570]	0.460	0.478	0.341	0.535	0.407
τ_K	0.326	[0.245, 0.407]	0.318	0.337	0.234	0.373	0.283
ρ_P	0.479	[0.351, 0.606]	0.478	0.501	0.364	0.510	0.432

Table C.2: Measures of dependence for the residuals of the data and different copulas for the GE and KO stocks in 2010-2011.

	Data	CI	Gaussian	Student's t	Clayton	Frank	Gumbel
ρ_S	0.200	[0.075, 0.328]	0.202	0.204	0.105	0.230	0.176
τ_K	0.139	[0.051, 0.225]	0.136	0.138	0.070	0.155	0.119
ρ_P	0.212	[0.083, 0.340]	0.211	0.214	0.114	0.220	0.192

Table C.3: Measures of dependence for the residuals of the data and different copulas for the AA and GE stocks in 2006.

	Data	CI	Gaussian	Student's t	Clayton	Frank	Gumbel
ρ_S	0.240	[0.121, 0.362]	0.197	0.213	0.122	0.262	0.184
τ_K	0.166	[0.084, 0.249]	0.133	0.145	0.082	0.177	0.125
ρ_P	0.207	[0.076, 0.338]	0.207	0.225	0.133	0.251	0.201

Table C.4: Measures of dependence for the residuals of the data and different copulas for the AA and KO stocks in 2006.

	Data	CI	Gaussian	Student's t	Clayton	Frank	Gumbel
ρ_S	0.443	[0.340, 0.548]	0.417	0.431	0.387	0.470	0.349
τ_K	0.306	[0.231, 0.383]	0.286	0.299	0.266	0.324	0.240
ρ_P	0.435	[0.304, 0.566]	0.434	0.451	0.410	0.449	0.372

Table C.5: Measures of dependence for the residuals of the data and different copulas for the GE and KO stocks in 2006.

	Data	CI	Gaussian	Student's t	Clayton	Frank	Gumbel
ρ_S	0.428	[0.322, 0.536]	0.380	0.434	0.333	0.478	0.354
τ_K	0.298	[0.218, 0.377]	0.260	0.303	0.228	0.330	0.244
ρ_P	0.396	[0.260, 0.529]	0.396	0.456	0.355	0.456	0.378

Table C.6: Measures of dependence for the residuals of the data and different copulas for the AC and ER stocks in 2010-2011.

	Data	CI	Gaussian	Student's t	Clayton	Frank	Gumbel
ρ_S	0.488	[0.376, 0.603]	0.504	0.505	0.374	0.567	0.476
τ_K	0.356	[0.270, 0.441]	0.351	0.359	0.257	0.398	0.334
ρ_P	0.523	[0.412, 0.636]	0.523	0.529	0.397	0.541	0.501

Table C.7: Measures of dependence for the residuals of the data and different copulas for the AC and SH stocks in 2010-2011.

	Data	CI	Gaussian	Student's t	Clayton	Frank	Gumbel
ρ_S	0.392	[0.282, 0.505]	0.337	0.414	0.286	0.461	0.322
τ_K	0.272	[0.192, 0.353]	0.230	0.289	0.195	0.317	0.221
ρ_P	0.353	[0.149, 0.546]	0.352	0.435	0.308	0.440	0.345

Table C.8: Measures of dependence for the residuals of the data and different copulas for the ER and SH stocks in 2010-2011.

Appendix D

Tables of Kolmogorov-Smirnov distances

Stocks	Gaussian	Student's t	Clayton	Frank	Gumbel
AA & GE	0.062	0.061	0.075	0.064	0.061
AA & KO	0.081	0.078	0.093	0.076	0.074
GE & KO	0.068	0.064	0.087	0.066	0.064

Table D.1: The maximum distance between the estimated copula and the empirical copula for the American stocks in 2006.

Stocks	Gaussian	Student's t	Clayton	Frank	Gumbel
AC & ER	0.100	0.094	0.122	0.094	0.089
AC & SH	0.094	0.088	0.126	0.084	0.080
ER & SH	0.100	0.090	0.123	0.086	0.086

Table D.2: The maximum distance between the estimated copula and the empirical copula for the Swedish stocks in 2010-2011.

Appendix E

Tables of option prices

Stocks	Gaussian	Student's t	Clayton	Frank	Gumbel
AA & GE	-3.71×10^{-4}	5.71×10^{-4}	4.05×10^{-4}	-2.40×10^{-6}	3.63×10^{-4}
AA & KO	-7.25×10^{-4}	1.14×10^{-3}	5.14×10^{-4}	3.81×10^{-5}	-6.29×10^{-4}
GE & KO	2.41×10^{-4}	-2.62×10^{-4}	5.89×10^{-4}	4.67×10^{-4}	3.76×10^{-4}

Table E.1: Percental differences in option prices between different copulas for the American stocks in 2006.

Stocks	Gaussian	Student's t	Clayton	Frank	Gumbel
AC & ER	1.31×10^{-4}	-9.28×10^{-4}	-8.01×10^{-4}	-4.57×10^{-4}	2.43×10^{-4}
AC & SH	-4.69×10^{-4}	-7.61×10^{-4}	9.08×10^{-4}	-3.34×10^{-4}	-7.61×10^{-4}
ER & SH	-5.88×10^{-5}	1.32×10^{-3}	-9.25×10^{-4}	-5.09×10^{-4}	-3.60×10^{-4}

Table E.2: Percental differences in option prices between different copulas for the Swedish stocks in 2010-2011.

	Gaussian	Student's t	Clayton	Frank	Gumbel
Exchange	49.10	49.11	49.11	49.11	49.11
Maximum	65.87	65.87	65.88	65.88	65.87
Call on max	24.65	24.65	24.66	24.66	24.65
Average spread	1.78	1.77	1.93	1.75	1.80

Table E.3: Option prices for AA and KO 2010-2011.

	Gaussian	Student's t	Clayton	Frank	Gumbel
Exchange	46.56	46.59	46.59	46.59	46.59
Maximum	65.87	65.87	65.88	65.88	65.88
Call on max	23.40	23.40	23.40	23.40	23.40
Average spread	1.71	1.69	1.82	1.68	1.75

Table E.4: Option prices for GE and KO 2010-2011.

	Gaussian	Student's t	Clayton	Frank	Gumbel
Exchange	5.19	5.18	5.30	5.17	5.21
Maximum	32.54	32.54	32.66	32.54	32.57
Call on max	3.97	3.97	4.05	3.96	3.98
Average spread	1.69	1.69	1.77	1.68	1.70

Table E.5: Option prices for AA and GE 2006.

	Gaussian	Student's t	Clayton	Frank	Gumbel
Exchange	14.76	14.76	14.79	14.74	14.76
Maximum	42.12	42.12	42.15	42.11	42.12
Call on max	7.63	7.62	7.65	7.61	7.63
Average spread	1.75	1.73	1.81	1.70	1.75

Table E.6: Option prices for AA and KO 2006.

	Gaussian	Student's t	Clayton	Frank	Gumbel
Exchange	11.01	11.01	11.01	11.01	11.02
Maximum	41.94	41.94	41.94	41.94	41.95
Call on max	5.79	5.78	5.79	5.79	5.79
Average spread	1.04	1.02	1.06	1.02	1.09

Table E.7: Option prices for GE and KO 2006.

	Gaussian	Student's t	Clayton	Frank	Gumbel
Exchange	0.42	0.31	0.52	0.32	0.45
Maximum	162.71	162.61	162.85	162.61	162.76
Call on max	39.29	39.23	39.39	39.23	39.31
Average spread	9.20	8.82	9.46	8.82	9.30

Table E.8: Option prices for AC and ER 2010-2011.

	Gaussian	Student's t	Clayton	Frank	Gumbel
Exchange	48.21	48.18	49.37	48.08	48.47
Maximum	210.53	210.47	211.65	210.38	210.74
Call on max	31.87	31.83	32.65	31.77	31.95
Average spread	9.23	9.17	10.35	9.05	9.45

Table E.9: Option prices for AC and SH 2010-2011.

	Gaussian	Student's t	Clayton	Frank	Gumbel
Exchange	116.27	116.22	116.23	116.19	116.23
Maximum	207.57	207.53	207.53	207.51	207.53
Call on max	59.27	59.22	59.22	59.21	59.22
Average spread	9.10	8.59	9.37	8.58	9.17

Table E.10: Option prices for ER and SH 2010-2011.

Bibliography

[1] Bollerslev, T. (1986) *Generalized Autoregressive Conditional Heteroskedasticity.* Journal of Econometrics 31 pp. 307-327. North-Holland.

[2] Box, G.E.P., Pierce, A. (1970). *Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models.* Journal of the American Statistical Association. Vol. 65, No. 332, p. 1509-1526.

[3] Cherubini, U., Luciano, E., Vecchiato, W. (2004). *Copula Methods in Finance.* Hoboken: John Wiley & Sons.

[4] Chiou, S.C., Tsay, S.R. (2008). *A Copula-based Approach to Option Pricing and Risk Assessment.* Journal of Data Science. Vol. 6, p. 273-301.

[5] Delbaen, F. et al. (1994). *A general version of the fundamental theorem of asset pricing.* Mathematische Annalen 300.

[6] Duan, J.C (1995). *The GARCH Option Pricing Model.* Mathematical Finance, Vol. 5, No. 1, pp. 13-32. Oxford: Blackwell Publishers.

[7] Embrechts, P., Lindskog, F., McNeil, A. (2001). *Modelling Dependence with Copulas and Applications in Risk Management.*

[8] Heston, S.L. and Nandi, S. (2000). *A Close-Form GARCH Option Valuation Model.* The Review of Financial Studies, Vol. 13 No. 3 pp. 585-625. Oxford University Press.

[9] Hull, J.C. (2009). *Options, futures and other derivatives.* 7th ed. New Jersey: Pearson Education, Inc.

[10] Klebaner, F.C. (2005). *Introduction to stochastic calculus with applications.* 2nd ed. London: Imperial College Press.

[11] Massey, F.J. (1951). *The Kolmogorov-Smirnov Test for Goodness of Fit.* Journal of the American Statistical Association. Vol. 46, No. 253, p. 68-78.

[12] McLeod, A.I., Li, W.K. (1983). *Diagnostic Checking ARMA Time Series Models Using Squared-Residual Autocorrelations"*. Journal of Time Series Analysis. Vol. 4, p. 269-273.

[13] Nelsen, R.B. (2006). *An Introduction to Copulas*. 2nd ed. New York: Springer Science+Business Media, Inc.

[14] Peacock, J. A. (1983). *Two-dimensional Goodness-of-Fit Testing in Astronomy*. Royal Astronomical Society, Monthly Notices. Vol. 202, p.615-627.

[15] Salmon, F. (2009). *Recipe for Disaster: The Formula That Killed Wall Street*. Wired Magazine: 17.03 http://www.wired.com/techbiz/it/magazine/17-03/wp_quant?currentPage=all 2011-10-24.

[16] Shreve, S.E. (2008). *Stochastic Calculus for Finance II Continuous-Time Models*. 8th printing. New York: Springer Science + Business Media, LLC.