The Norwegian Stock Market:
- A Local Gaussian Perspective

by
Andreas Lura

Master's thesis in Statistics
Financial Theory and Insurance Mathematics

University of Bergen
Department of Mathematics

June 3, 2013

Acknowledgement

I would like to thank my supervisor Dag Bjarne Tjøstheim and secondary supervisor Bård Støve for the opportunity and help to write this thesis. The process of working with the topic of Local Gaussian Correlation, and see some of its uses in finance has been exciting and educational. I would especially want to thank Dag Bjarne for an amazing guidance while working on this thesis, the supervision has been of great motivation and help.

I would also like to thank my family, friends and co-students for support and motivation when working on this thesis. I could not have completed this thesis without your support,

Thank You.

Abstract

In this thesis, using daily returns from 18 stocks, oil price, exchange rates and the main index of the Oslo Stock Exchange over a period of 5 years, we investigate how the Local Gaussian Correlation can be used to describe the change in the relationship between stocks and the market and how it can extend already established theory in finance.

Topics covered in this thesis are; risk estimation by conventional risk measures and a method based on Local Gaussian Correlation, the Capital Asset Pricing Model (CAPM), copulas and GARCH as a description of volatility and as a description of the marginal distributions for copulas.

Value at Risk and Expected Shortfall are well established risk measurements in finance. They are dependent on a good description for the distribution in the tail, which can be challenging. These measures only provide one single number as a description of the risk, this might be appealing, but does not really provide detailed information.

By using the theory of CAPM there has been some attempt to describe the change in risk by using the so-called conditional moments of the observations. This approach might be biased, as the conditional moment fails to describe the constant correlation and variances of the Gaussian distribution. By rather using the local parameters found when calculating the Local Gaussian Correlation as a local description of the beta on our data, there seem to be higher risk in the upper tail and the lower than in the middle. However, what differs from the results found by the previously mentioned approach is that the risk in the upper tail seems to be higher than in the lower. This might be explained by very large gains for the stock market might be followed by a possible stock market downturn or even a crash (bubble), while negative values for the market is less likely to resolve in a sudden positive boost for the market.

Contents

Chapter 1

Introduction

The fact that correlation is not necessarily constant between factors in finance is observed in articles like Patton (2004) and Silvapulle & Granger (2001).

In Patton (2004), the authors investigate the dependence on stock markets, and finds evidence suggesting that models based on non-constant correlation between indices is a better description than models where the correlation is assumed to be constant.

In Silvapulle & Granger (2001), the authors group their data into three different categories, the lower tail, the middle and the upper tail. By conditioning on the group of the observations, they get different values for the variance and the correlation in these groups.

These different values for the variance and correlation are then used as a description of how the risk is changing between the groups. However, as pointed out in Boyer *et al.* (1997), the approach of conditional correlation is biased. The method fails to give a constant value for the correlation in a bivariate Gaussian distribution, thereby violating the definition of the Gaussian correlation.

As for calculation on the Norwegian market, there has been done some analysis of the influence of factors like oil price, examples of this are Bjørnland (2009) and Gjerde & Saettem (1999) which both use a vector autoregression model in analysis on the Norwegian market. Bjørnland (2009) analyses the relationships between oil price shocks and stock market booms on the Norwegian market, while Gjerde & Saettem (1999) analyses the relationship between stock returns and macroeconomic variables.

In this thesis, we will investigate changes in correlation on the Norwegian Stock Market, using daily data over a period of 5 years. We will be using the Local Gaussian Correlation to describe the change in correlation for the factors. This approach is not like the method of vector autoregression models, restricted to linear dependence and does not suffer from the bias that is found in conditional correlation. Additionally we will be using theory for risk management and copulas together with the Local Gaussian Correlation to describe and calculate the risk on the Norwegian Stock Market.

Since the Local Gaussian Correlation is a new measure of correlation, there has been limited use of it on observed data. It is therefore of interest to compare the results of this method to results found by other methods, and try to give an interpretation of these findings.

The structure of this thesis can be summarized as follows:

- Finance:
  A quick introduction to logarithmic differences, some distributions used in finance, risk estimation and GARCH as a description of volatility.

- Dependence Measures:
  Since the chapter of finance was restricted to the univariate case, we argue for the use of several variables in finance and give a recap of Pearson's correlation coefficient.

- Copulas:
  Copulas offers a way of binding together random variables under different kind of dependence structures. This chapter gives an introduction to copulas and introduce some different copula models.

- The Capital Asset Pricing Model:
  CAPM is a model, which states a link between the risk and expectation for investments under some assumptions. We will investigate an attempt to model the change in risk by the use of conditional correlation and see why results based on this method might be biased. The problem with conditional correlation motivates for the next chapter.

- Local Gaussian Correlation:
  The Local Gaussian Correlation is a model that lets several multivariate Gaussian distributions approximate an observed distribution f in separate neighborhoods. We then use the correlation of each of the Gaussian distributions as a description of the correlation for f in that particular part of the plane.

- Analysis of Dependence and Risk in the Norwegian Stock Market:
  By using the previously introduced theory, we are able to describe change in correlation on the Norwegian Stock Market and compare our results to results found by others. We use the risk estimators introduced earlier as well as one based on Local Gaussian Correlation to investigate the risk on the Oslo Stock Exchange.

- Conclusion:
  Short summary of the results and some interpretation of these, followed by some suggestions for further work.

Chapter 2

Finance

In this chapter, we will cover some of the fundamental aspects of financial modeling. The theory in this chapter is mostly based on the book of McNeil *et al.* (2005), which offers a thorough introduction to risk management. Another reference of influence that should be mentioned is the Bachelor's thesis, Lura (2011), as it covers some of the theory we will introduce in this chapter, in an investigation of the validity of the assumption of a Gaussian return distribution in finance.

We will start by explaining what financial returns are, how they are calculated and we will investigate some parametric assumptions regarding these. This is followed by a quick introduction to some models used in risk management and a model for describing volatility, all of which are important topics in the field of finance.

When modeling financial data an often used approach is to start out with what is known as the logarithmic difference or the returns for the data. When investigating the logarithmic differences, our interest is the change in the stock price rather than its actual value. The logarithmic difference is defined as follows

Definition 2.0.1. *Logarithmic difference*

$$X_t = 100(log(S_{t+1}) - log(S_t))$$

where S_t is the observed price of a stock.

It is easily seen from the definition of logarithmic difference that when S_{t+1} and S_t is equal, X_t is zero. When S_{t+1} is larger than S_t, the value of X_t will be positive, finally when S_{t+1} is less than S_t the value of X_t will be negative.

2.1 Return Data for the Statoil Stock

As an example of return data, we will be using real-life stock data for the Norwegian oil company Statoil. The Statoil stock is notated on the Oslo Stock Exchange and the New York Stock Exchange. Statoil was partially privatized and listed on the stock exchanges the 18 of June 2001, but the Norwegian government is still its major shareholder, and own approximately 2/3 of its stocks.

Our data for Statoil are based on daily prices of the stock at closing time. The dataset is collected from the Oslo Stock Exchange's homepage, and ranges over a five years period(from October 11, 2007 until October 11, 2012) resulting in 1260 values for the price, or 1259 values for the logarithmic difference. The following table summarizes some of the observed properties of the returns for the stock.

	Mean	Var	Median	Kurtosis	Skewness
Statoil	-0.02	5.04	0	4.04	-0.46

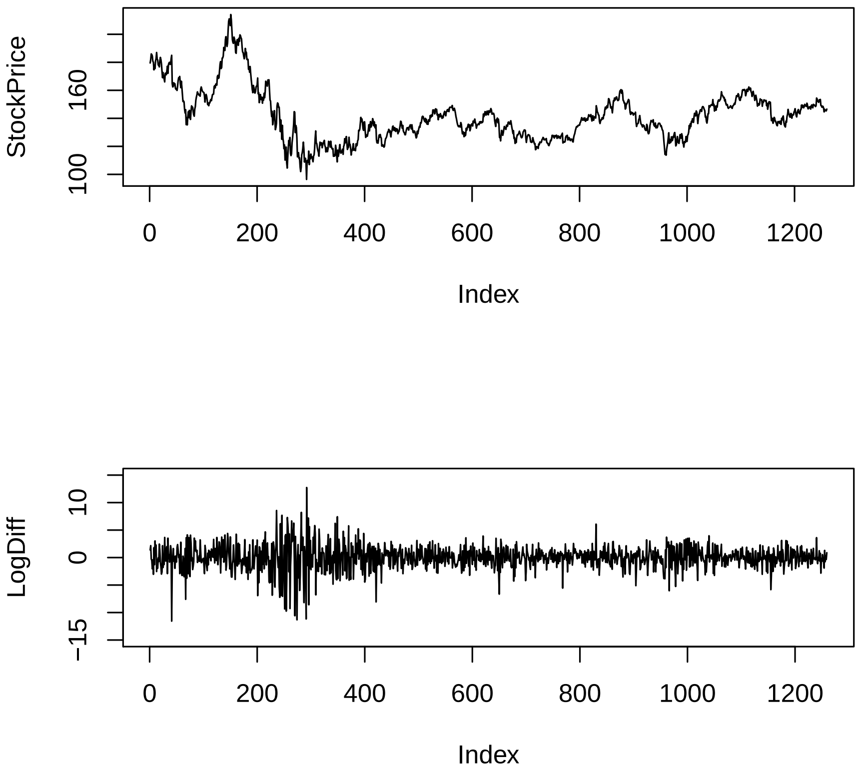


Figure 2.1: Logarithmic differences for Statoil

Note that in Figure 2.1, and for the rest of this thesis the following definition for kurtosis[1], sometimes referred to as excess kurtosis is used.

$$Kurt[X] = \frac{E[X^4]}{Var[X]^2} - 3.$$

2.2 Distribution Models for Financial Returns

By moving on to the parametric parts of statistics, we can take advantage of more advanced models, moreover, we might be able to simplify the calculations. In the following sections we will make and analyze some of the conventional assumptions regarding the distribution of the returns. There are numerous reasons why some of these assumptions should be avoided or at least be treated with caution. Falsely assuming that the return is described by a given distribution might lead to severe underestimation of risk.

Because of this, we will look at some examples where we assume a given distribution and to try to understand some of the pitfalls we should be aware of. In this section, we will be restricting ourselves to the univariate case, but will later see how the theory of copulas can be used to join marginal distributions together in different kinds of dependence.

2.2.1 Gaussian Return Distribution

The most assumed distribution for the returns is probably the Gaussian distribution. With its simple structure, low number of parameters and relation with the central limit theorem it is an appealing choice of distribution. Its univariate density function is

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\{\frac{(x-\mu)^2}{2\sigma^2}\}. \tag{2.1}$$

The expressions for the expectation and the variance for the Gaussian distribution are given

$$E[X] = \mu$$
$$Var[X] = \sigma^2.$$

When assuming a Gaussian return distribution we only need values for the expectation and variance, as the kurtosis and skewness are both zero for the Gaussian distribution. This results in its famous bell form and the symmetry around the mean.

[1]For an explanation of what kurtosis is and what it describes, DeCarlo (1997) might be an article of interest.

2.2.2 Student t Return Distribution

Student's t distribution is a distribution function often used in introductory courses on hypothesis testing on the mean. Nevertheless, this is not the only use for this distribution, as it turns out the t distribution may often be a better description for logarithmic differences than the Gaussian distribution. We say that T has a t distribution with υ degrees of freedom if T is given as follows

$$T = \frac{Z}{\sqrt{\frac{U}{\upsilon}}}. \tag{2.2}$$

Here Z is a standard Gaussian, and U is Chi squared with υ degrees of freedom, and independent of Z Hogg & Tanis (2005). Its density function can be given as

$$f(x|\upsilon) = \frac{\Gamma(\frac{\upsilon+1}{2})}{\sqrt{\upsilon\pi}\Gamma(\frac{\upsilon}{2})}(1+\frac{x^2}{\upsilon})^{-\frac{\upsilon+1}{2}}. \tag{2.3}$$

The expectation for a chi-square random variable is equal to its value for the degrees of freedom, this means that when υ gets large, Equation 2.2 will approach Z, i.e. when the degrees of freedom get large a t-distributed variable will converge to a standard Gaussian distribution.

In Hogg & Tanis (2005) υ is restricted to positive integer values, however we will follow in a similar manner as in Ruppert (2010) where the only restriction for the degrees of freedom is that $\upsilon > 0$, when the t-distribution is used as a model for data.

Like the Gaussian distribution, the t-distribution is symmetric, but it has heavier tails, and therefore, more extreme values are more likely with the t-distribution. Casella & Berger (2002) summarize the moments for the t distribution as

$$E[X^n] = \begin{cases} \frac{\Gamma(\frac{n+1}{2})\Gamma(\frac{\upsilon-n}{2})}{\sqrt{\pi}\Gamma(\frac{\upsilon}{2})}\upsilon^{n/2} & \text{if } n < \upsilon \text{ and even} \\ 0 & \text{if } n < \upsilon \text{ and odd} \\ \text{Undefined} & \text{otherwise} \end{cases} \tag{2.4}$$

From this, it follows that the mean is 0 for $\upsilon > 1$,
variance is $\frac{\upsilon}{\upsilon-2}$ for $\upsilon > 2$ and ∞ for $1 < \upsilon \leq 2$,
skewness is 0 for $\upsilon > 3$ and for $\upsilon > 4$, the kurtosis is $\frac{6}{(\upsilon-4)}$.

2.2.3 Normal Inverse Gaussian Return Distribution

The last distribution we will introduce is the Normal Inverse Gaussian distribution from the family of generalized hyperbolic distributions. This family was introduced in Barndorff-Nielsen (1978) and according to McNeil *et al.* (2005) some of the reason it has been embraced to a large extent in the field of financial-modeling is because of its relation to Lévy processes. Another advantage for the Normal Inverse Gaussian(NIG) distribution is that the description for both the skewness and kurtosis are more flexible than for the t and Gaussian distribution. The density function of the NIG distribution as parameterized in Barndorff-Nielsen (1997) is

$$f(x|\alpha, \beta, \mu, \delta) = a(\alpha, \beta, \mu, \delta)q(\frac{x-\mu}{\delta})^{-1}K_1\{\delta\alpha q(\frac{x-\mu}{\delta})\}e^{\beta x}, \tag{2.5}$$

where K_1 is the modified Bessel function of order three and index 1, which in Bølviken & Benth (2000) is written on the form from Abramowitz *et al.* (1964)

$$K_1(z) = z\int_1^{\infty} e^{-zt}\sqrt{t^2-1}dt$$

$$q(x) = \sqrt{1+x^2},$$

$$\text{and } a(\alpha, \beta, \mu, \delta) = \pi^{-1}\alpha e^{\delta\sqrt{(\alpha^2-\beta^2)}-\beta\mu}.$$

Following are some moments and properties for the distribution as given in Bølviken & Benth (2000)

$$E[X] = \mu + \delta\frac{\beta/\alpha}{(1-(\beta/\alpha)^2)^{1/2}}$$

$$Var[X] = \delta^2\alpha^{-1}\frac{1}{(1-(\beta/\alpha)^2)^{3/2}}$$

$$Skew[X] = 3\alpha^{-1/4}\frac{\beta/\alpha}{(1-(\beta/\alpha)^2)^{1/4}}$$

$$Kurt[X] = 3\alpha^{-1/2}\frac{1+4(\beta/\alpha)^2}{(1-(\beta/\alpha)^2)^{1/2}}.$$

Due to more parameters and the appearance of the Bessel function the NIG distribution is harder to fit to data than both the student t distribution and the Gaussian distribution. We will on two different occasions compare the three distributions introduced and try to determine if the flexibility of the NIG distribution makes it a better fit for the return distribution than the two simpler distributions.

2.2.4 Comparison of the Distributions on the Statoil Stock

In our dataset for Statoil the maximum value for the return is 12.74 while the minimum is -11.56. The probability for a Gaussian random variable with parameters matching the one given for Statoil in Figure 2.1 having a value of less than 11.5 is 0.9999999. Moreover,

$$(1 - 0.9999999) \times 1259 \approx 0.000181,$$

this means that we would expect approximately zero observations from the given Gaussian distribution exceeding this limit during the period, but counting the minimum value we observed two exceedances of this magnitude for the Statoil stock. The probability of observing the maximum value(less than 1-0.9999999) is so low that we should question the use of the Gaussian distribution as a description for these returns.

tFit from the package[2] fBasics suggest 1.75 degrees of freedom based on the observations of the returns, in this case the expectation is zero, the variance if ∞, while the kurtosis and skewness are undefined. But even with a variance of ∞ the t distribution might be a better choice for the return distribution than the Gaussian distribution, the probability of observing a value less than 11.5 with the given degrees of freedom is 0.9940744. This means that the probability of observing the maximum value is low, but it is definitely higher than for the Gaussian distribution. Moreover

$$(1 - 0.9940744) \times 1259 \approx 7.46$$

which means that with the t distribution with the given degrees of freedom we will expect some exceedances of this size during a period of 1259 days. We only observed two occasions where the absolute value of the return exceeded 11.5 during our sample period, nevertheless it seems like the t distribution might be a better choice than the Gaussian distribution, at least from a risk perspective.

Similar the fit.NH function from QRMlib was used to measure parameters for the NIG distribution, and the parameters are summarized in the table below.

delta	alpha	beta	mu
2.01634871	0.41189402	-0.03914205	0.17637789

[2] The calculations and the figures in this thesis is created by the use of **R**, additional packages that have been used in the calculations are mentioned in the text.

The probability for the NIG distribution with the parameters given above having a value less than 11.5 is 0.9996756.

$$(1 - 0.9996756) \times 1259 \approx 0.408375,$$

which is closer to the observed value of 2 exceedance of this size found in the sample, than the Gaussian and t distribution.

Note that this comparison is rather simplified, as the maximum value for the returns was 12.74 and not 11.5, the means for the Gaussian and the NIG distribution are not zero, which means that $P(X < x)$ is not equal to $1 - P(X < -x)$ for these, and the NIG distribution is skewed for these parameters, which is a factor to consider as well. Nevertheless, it provides some information on how well the extreme observations are described by the distributions. We will return to a more reasonable comparison of these distributions later on.

In the following plot there are three random samples, these are from the Gaussian, t and NIG distribution, where the parameters are as given above. For each of the plot the y-axis is restricted from -15 to 15, but it should be noted that the t distribution exceeds this limit on several occasions, and in the random sample from the t distribution our observed maximum was 34.51224 and minimum -31.69406. A visual comparison of these distributions and the plot of the logarithmic difference in Figure 2.1, might suggest either the t or the NIG distribution as a choice for the distribution of the Statoil returns.

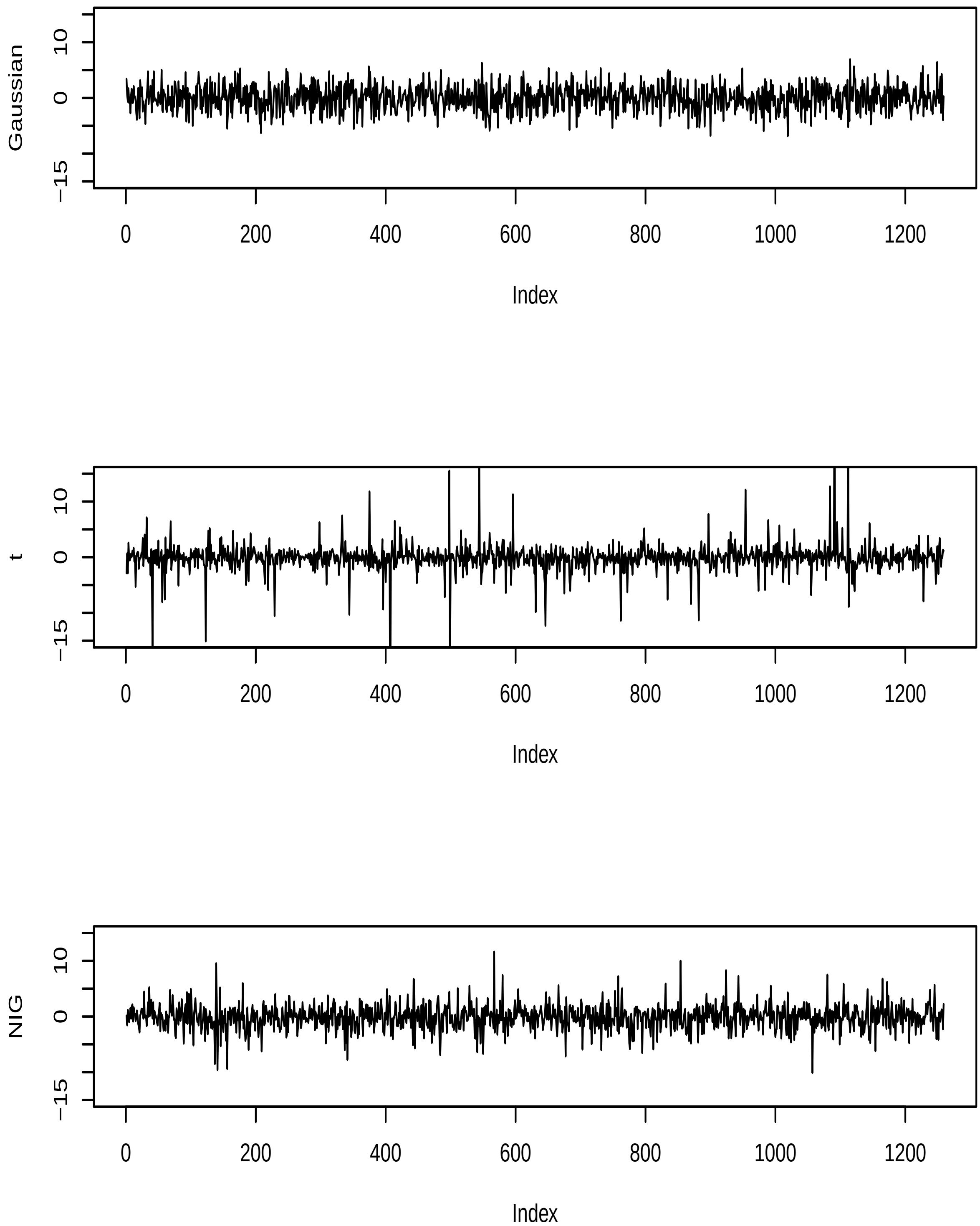


Figure 2.2: Random samples, with parameters based on returns from Statoil

2.3 Risk Measures

If we are interested in the probability of a standard Gaussian distributed variable being less than 1.65 we can easily look up that the probability of this is approximately 0.95. This kind of problem arise in finance, where an insurance company need to verify that they have enough money available to cover potential losses, or someone who has invested, need to know that they can afford the potential loss. We will in this section take a look at two different ways of estimating a limit for the loss, these are Value at Risk and Expected Shortfall. Before defining a risk measure, it is natural to define and have a clear understanding what we mean by loss.

Definition 2.3.1. ***The Loss function***
We let $L(t) = V(t) - V(t+1)$ be the loss over a period of 1 steps and V(t) denotes the given value of our object at time t.

As we can see from the definition a positive value for the loss function $L(t)$ means that the value of interest has decreased from time t to $t+1$. Here we have restricted ourselves to a period of 1 step, although this could easily be defined for a general numbers of steps, moreover, we will refer to the loss function only by L. The reason for this simplification is that the time t is not really relevant in the following definitions and how the theory can be extended for time periods longer than 1 step are briefly mentioned later.

We will throughout this introduction to risk measures use the Gaussian distribution when assuming a distribution for the loss, this might not be a realistic assumption, but it will hopefully provide the necessary understanding on how these calculations can be done under other choices for the loss distribution.

2.3.1 Value at Risk and Expected Shortfall

Value at Risk, or often just VaR is a measure of the quantile function for the Loss L defined above. We will in the case of Value at Risk and Expected Shortfall stick to definition from McNeil *et al.* (2005).

Definition 2.3.2. ***Value at Risk***

$$VaR_\alpha = inf\{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\} = inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\},$$

where $\alpha \in (0,1)$ gives the confidence level. L gives the loss, and $F_L(l) = P(L \leq l)$.

This means that VaR_α is the limit that our loss will not exceed with a probability of α.

The definition of Expected Shortfall is closely related to the definition of VaR, and is

Definition 2.3.3. ***Expected Shortfall***

$$ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 q_u(F_L)\,du,$$

where $q_\alpha(F_L)$ is the quantile function of F_L defined as
$q_\alpha(F_L) = inf\{l \in R : F_L(l) \geq \alpha\}$, $\alpha \in (0,1)$.

Expected Shortfall like VaR is a risk measure often used in finance. One difference between these is that, while VaR only give us a limit, we will not exceed with a given probability, Expected Shortfall gives us the expected value for the loss conditioned on that we have exceeded this limit. The relationship between VaR and ES is easily seen in Definition 2.3.3 where $q_u(F_L)$ equals the definition of VaR given in Definition 2.3.2.

As methods of calculating VaR and ES, we will in this section be using the historical and parametric method, although methods based on Monte Carlo simulation can also be used as shown in the later analysis.

2.3.2 Historical Method for Measuring VaR and ES

In order to use the historical method, we arrange our observed losses $l_1, \dots, l_n$ from the distribution L in increasing order, which is called the order statistics. $l_{(n)}$ is the largest loss and $l_{(1)}$ is our smallest loss, since the Loss function gives a positive value for the loss. When calculating VaR we want to find the smallest loss l where

$$P(L \leq l) \geq \alpha.$$

In other words if we let $\alpha = 0.95$, we want to find the smallest loss l_i , that 95% of our values for the returns exceed. While for the Expected Shortfall, we calculate the mean of the 5% largest losses.

Example 2.3.4. Historical VaR and ES for Statoil Stock
As mentioned earlier we had 1259 values in this dataset for the Statoil stock. We want to find the value our loss will not exceed with a probability of 95%.

$$1259 \times 0.95 = 1196.05,$$

this means that our value of interest is the $l_{(1197)}$, the reason for the use of $l_{(1197)}$ instead of $l_{(1196)}$ is that when calculating a risk it is usually better to overestimate than underestimate. We then have that according to the historical VaR we will with 95% probability not lose more than $l_{(1197)} = 3.536$. Or in other words, our value for the return will with 95% probability exceed -3.536.

$$1259 \times 0.99 = 1246.41$$

and $l_{(1247)} = 7.058$ is the historical VaR with $\alpha = 0.99$.

When calculating the Expected Shortfall we are interested in the loss $l_{(1197)}$, however, we are also interested in the other values exceeding the limit, namely $l_{(1198)}, \ldots, l_{(1259)}$. We therefore calculate the mean of the values $l_{(1197)}, \ldots, l_{(1259)}$, which is 5.557. And we say that if our loss exceeds the 95% limit we expect it to be -5.557.

	$VaR_{0.95}$	$VaR_{0.99}$	$ES_{0.95}$	$ES_{0.99}$
Statoil	3.536	7.058	5.557	9.2153

2.3.3 Parametric Methods for Risk Measures

As mentioned we will limit this introduction of parametric methods to the assumption of a Gaussian distribution for the loss, the reason for this is that it is essential similar for other distributions. Parametric Value at Risk calculates the limit that our loss will not exceed by a given probability conditioned on our loss being distributed by the given distribution, similar Expected Shortfall is the expected value conditioned on that we have exceeded this limit.

Theorem 2.3.5. ***VaR for a Gaussian Distributed Loss***

$$VaR_\alpha = \mu + \sigma\Phi^{-1}(\alpha)$$

where Φ denotes the standard Gaussian cumulative distribution function, Φ^{-1} is its inverse, the quantile function and L is Gaussian distributed with variance σ^2 and mean μ.

This expression is easily verified by Definition 2.3.2 and the fact that

$$P(L \leq \mu + \sigma\Phi^{-1}(\alpha)) = \Phi(\Phi^{-1}(\alpha)) = \alpha.$$

The close relation between VaR and Expected Shortfall combined with Theorem 2.3.5 implies

Theorem 2.3.6. ***Expected Shortfall for Gaussian distribution***
For an Gaussian loss distribution L,

$$ES_\alpha = \mu + \frac{\sigma}{1-\alpha}\phi(\Phi^{-1}(\alpha)),$$

where ϕ is the standard Gaussian density function and Φ^{-1} is the standard Gaussian quantile function. This can be verified by combining Definition 2.3.3 with the expression given for the VaR in Theorem 2.3.5.

Example 2.3.7. Parametric VaR and ES with Gaussian Loss
The 0.95 quantile for a standard Gaussian distribution is $\Phi^{-1}(0.95) \approx 1.645$, and for the already mentioned Statoil stock we had that $\hat{\mu} = -0.02$ and $\hat{\sigma} = 2.245$. This means that if we assume that the loss variable is distributed by a Gaussian distribution, and use Theorem 2.3.5 and 2.3.6 we get the following values for the Expected Shortfall and Value at Risk.

	$VaR_{0.95}$	$VaR_{0.99}$	$ES_{0.95}$	$ES_{0.99}$
Statoil	3.71	5.24	4.650	6

Note that the mean of the Loss distribution L has the opposite sign of the mean for the return.

2.3.4 Discussion of Risk Measures

As we have seen, there are several ways of estimating risk. The historical way of estimating Value at Risk and Expected Shortfall is quite easy, but they are also rather naive. Assuming that our observed values for the loss are a good representation for potential loss might be considered foolish. Similar parametric VaR might be somewhat risky too, at least with a Gaussian distribution, because as we have seen the Gaussian distribution is probably a bad choice for the return distribution and is known for underestimating the risk. This can be seen in our data when $\alpha = 0.99$, then the historical VaR and ES are much higher than the parametric results.

To calculate the risk over a period N, $N > 1$, a natural approach is to use data of length N, for instance using weekly data if we want to estimate the risk over a week. Unfortunately, this will lead to a reduction in the number of observations, an alternative approach as suggested in Hull (2009) is to assume

$$N\text{-day VaR } = 1\text{-day VaR} \times \sqrt{N}$$

This approach makes particular sense if we assume that the loss distribution L is given by a Gaussian distribution with expectation zero, since $N(0, \sigma^2)$ then $\sqrt{n} \times L$ is $N(0, \sigma^2 N)$.

In the choice between VaR and ES one advantage of ES over VaR is as mentioned, that while VaR gives you information of the risk with a certainty α and leaves the rest as an uncertainty, Expected Shortfall tells us what to expect in the tail. Another advantage is that while Expected Shortfall is a so-called coherent risk measure, this is not always the case for Value at Risk, this is shown in an example of defaultable bonds in McNeil *et al.* (2005).

A coherent risk measure is said to be a risk measure satisfying the four axioms of translation invariance, subadditivity, positive homogeneity, and monotonicity. These axioms are defined in both Artzner *et al.* (1999) and McNeil *et al.* (2005). Of the four axioms the subadditivity property is really the one of interest in our case, as this does not necessarily apply for VaR. We will explain this axiom in a less formal way than in the references given above. If we let L_1 and L_2 be two loss distributions and $\xi(L)$ be a function for the risk of a given loss. Subadditivity then requires that we have

$$\xi(L_1 + L_2) \leq \xi(L_1) + \xi(L_2).$$

This essentially means that a merger of multiple risks should not create additional risk.

We have seen that the t or NIG distribution might be a better choice for the loss distribution than the Gaussian, but since the loss is probably not independent and identical distributed, these assumptions might be unrealistic, this leads to the topic of volatility.

2.4 Volatility and Volatility Measures

As we can see by comparing Figure 2.2 and Figure 2.1, it is more than just the size of the more extreme values that suggest against the assumption that the returns of the stock price is given by iid observations from a given distribution. In the article Mandelbrot (1963), Mandelbrot described the phenomenon, now known as volatility clustering, as "large changes tend to be followed by large changes -of either sign- and small changes tend to be followed by small changes.". This phenomenon is easily seen in Figure 2.1, where 28 of the 45 observations having absolute value exceeding 5 is in the interval [200,300] of the 1259 values.

Volatility can be seen as the conditional standard deviation, which is dependent on the past of our process. This means that the volatility itself is not directly observable, but we can estimate it, and a way of doing this is by the GARCH model.

2.4.1 GARCH

GARCH is short for Generalized Autoregressive Conditional Heteroscedasticity, and was introduced in Bollerslev (1986). GARCH is a generalization of the ARCH model, and the original ARCH model was developed by Robert F Engle in Engle (1982). Both the ARCH and its generalization GARCH are models which describes the future volatility, based upon the observed values of the past.

We will in this thesis stick to the GARCH model, which in McNeil *et al.* (2005) is defined as follows

Definition 2.4.1. ***GARCH Process***
Let Z_t be a strict white noise process SWN(0,1), which means that it is independent and identical distributed with expectation 0, and variance 1. The process X_t is a GARCH(p,q) process, if it is strictly stationary for all $t \in \mathbb{Z}$ and X_t is given by $\sigma_t Z_t$. Where

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2,$$

$$\alpha_0 > 0, \alpha_1 \geq 0, \ldots, \alpha_p \geq 0 \text{ and } \beta_1 \geq 0, \ldots, \beta_q \geq 0$$

The sum involving the alphas is often referred to as the ARCH term and the value p let us choose how many lags in $\{X_t\}$ we want to use, i.e. for how long time we will let an historic value influence the volatility.

Similar the sum involving the betas is known as the GARCH term, and q let us choose for how long the previous values for the volatility should influence the present volatility. Obviously this means that a GARCH(p,0) is equal to an ARCH(p) process.

GARCH Filtering

As we saw in Figure 2.1, the observations do not seem to be independent. Later on we will need to assume that, these observations are in fact given by independent and identical distributed observations, we will therefore be using the GARCH process to try to filter out the day to day dependence in financial data.
The logarithmic differences X_t is then described as

$$X_t = \mu + Z_t \sigma_t$$

where $Z_t \sigma_t$ is a GARCH process as defined earlier. And the standardized residuals, where r_t is the observed return is calculated as

$$\hat{a}_t = \frac{(r_t - \hat{\mu})}{\hat{\sigma}_t} \tag{2.6}$$

This filtration is described and used in Støve *et al.* (2012), and in that case, it is concluded that the modification done by this filtration does not significantly affect the conclusions regarding financial contagion in financial data. We will use the same way of filtering financial data, but note that manipulation of data in this manner should always be done with caution.

Fitting of GARCH Processes to Data

Maximum likelihood is one of the ways we usually fit the models to observed data for a GARCH model, we will introduce the approach as described in McNeil *et al.* (2005). Since in a GARCH(p,q) model σ_t is calculated by the p last values of X_t and the q last values of σ_t we are interested in the joint distribution of $X_p, \dots, X_n$ conditioned on $X_0, \dots, X_{p-1}$ and $\sigma_0, \dots, \sigma_{q-1}$. In the case of GARCH(1,1) this is

$$\begin{aligned} & f_{X_1,\dots,X_n|X_0,\sigma_0}(X_1, \dots, X_n|X_0, \sigma_0) \\ & = \Pi_{t=1}^{n} f_{X_t|X_{t-1},\dots,X_0,\sigma_0}(X_t|X_{t-1}, \dots, X_0, \sigma_0). \end{aligned}$$

We can write the conditional likelihood for a GARCH(1,1) model on the following form

$$L(\alpha_0, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q|\mathbf{X}) = \Pi_{t=1}^{n} \frac{1}{\sigma_t} g(\frac{X_t}{\sigma_t}) \tag{2.7}$$

where σ_t is as defined in Definition 2.4.1, and $g(z)$ is the density of a strict white noise process, e.g. a standard Gaussian distribution.

In practice this can be estimated by choosing a starting value for σ_0, and finding the values that maximizes the logarithm of the likelihood. This can be done numerically by systematically trying different values for the parameters in the likelihood, until the values that maximizes the likelihood are found.

This is time-consuming, but can be simplified by using the approach of variance targeting[3] as described in Hull (2009). We then reduce the numbers of parameters we have to estimate. This approach is to set α_0 equal to the following equation.

$$\alpha_0 = V_L(1 - \alpha_1 - \dots - \alpha_p - \beta_1 - \dots - \beta_q).$$

Here V_L, which is known as the long-run average variance is set equal to the sample variance calculated from the data. In the case of a GARCH(1,1) model, we then only have to estimate values for α_1 and β_1.

However, in most cases it is more convenient and more effective to use packages that is available for this purpose, and in the rest of this paper all parameters in the GARCH model is found by the function garchFit from the package fGarch.

[3]A discussion of the possible advantages and disadvantages with variance targeting in GARCH models can be found in Francq *et al.* (2009).

Example 2.4.2. Volatility estimated by GARCH(1,1) on the Statoil stock
For this example we have used the same dataset as showed in Figure 2.1. The coefficients as estimated on this stock, by fGarch for an GARCH(1,1), with a Gaussian strict white noise process are

	α_0	α_1	β_1
GARCH(1,1)	0.0067757	0.0585787	0.9289262

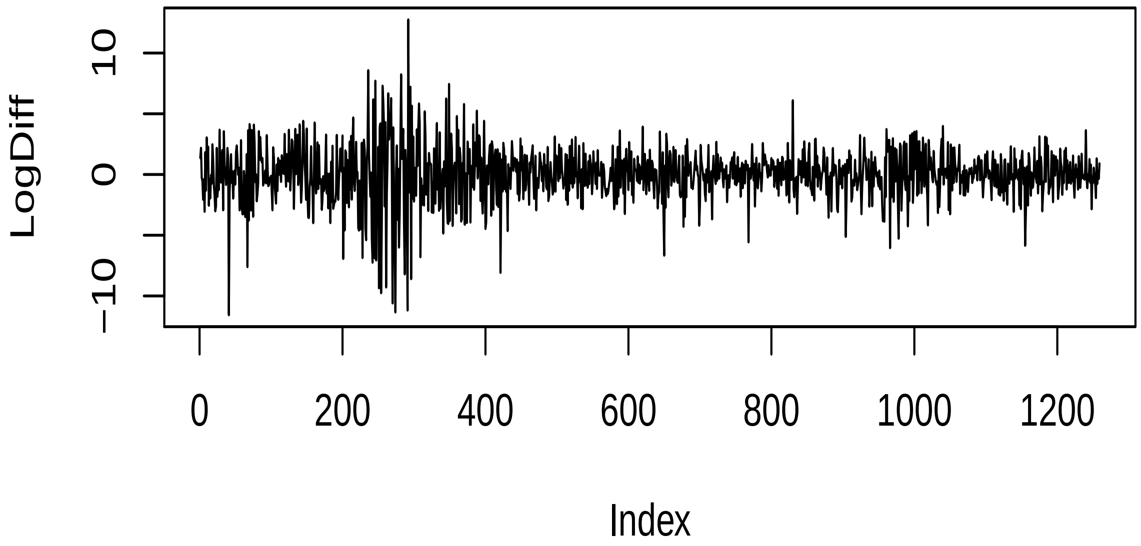
Figure 2.3 is the plot of the logarithmic difference and σ_t from the GARCH(1,1) process described above. It should be noted that the first values estimated will be influenced by the starting value of σ_0. Throughout this thesis σ_0 is set to the empirical standard deviation of the sample. In this plot is quite apparent how periods with large values for the logarithmic difference, so called volatility clusters, is the same area where the GARCH process describes the volatility to be large.

To compare the different choice of distribution for the strict white noise process, Figure 2.4 has plots of three GARCH processes based on the Statoil stock, where the SWN samples are drawn from a Gaussian, t and a NIG distribution.

In the case of a Gaussian and NIG distributed Z_t, the expectation and variance for the distribution is set to 0 and 1. However when Z_t is a t distribution, we have to apply a standardization of the t distribution where,

$$Z_t = \sqrt{\frac{\upsilon - 2}{\upsilon}} Y_t.$$

Here Y_t is t distributed with υ degrees of freedom, then $E[Z_t] = 0$ and $Var[Z_t] = 1$ when $\upsilon > 2$. The degrees of freedom, and the additional values for the NIG distribution are calculated by the garchFit function. It is obvious that the plot in Figure 2.4 resembles the logarithmic difference in Figure 2.1 much better than in the case of independent and identical distributed random variables seen in Figure 2.2. We will later on return to a discussion and test for which of the distribution is the best choice for the white noise process for some stocks available at the Norwegian Stock Market.



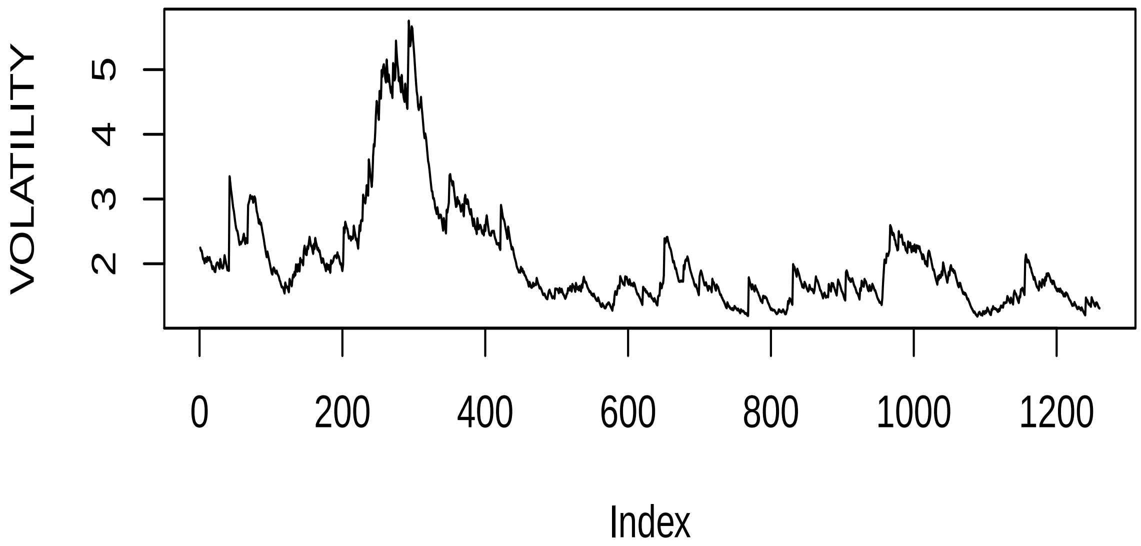


Figure 2.3: Volatility estimated by GARCH(1,1) on the logarithmic differences

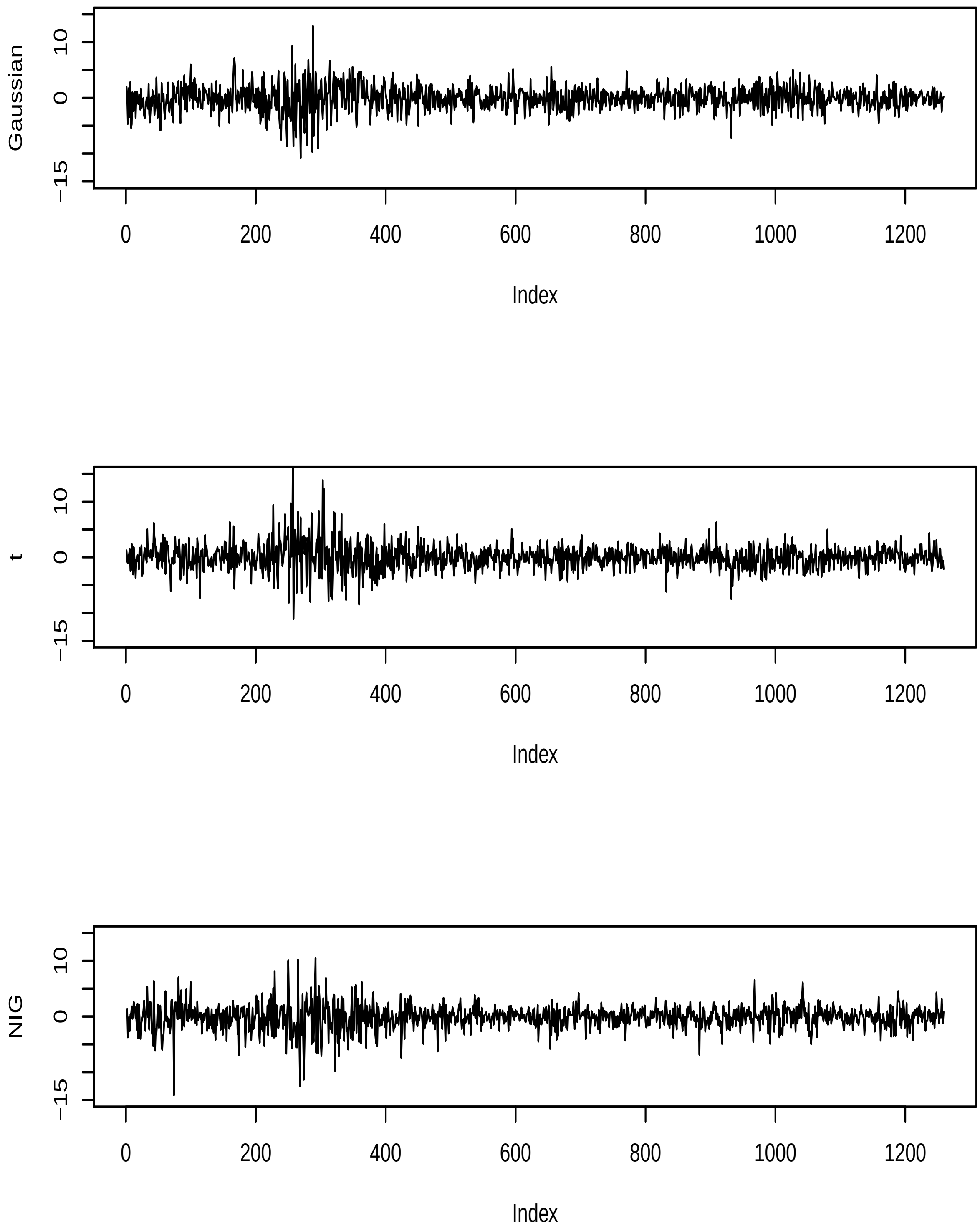


Figure 2.4: GARCH(1,1) processes based on the Statoil stock, where the white noise samples are drawn from a Gaussian, t and a NIG distribution.

Chapter 3

Dependence Measures

Now that we have covered some topics regarding finance and risk management in the univariate case, a natural step further is to consider the case of multiple random variables.

As an example of multiple random processes, let us consider the Norwegian Stock Market. Say we decide to invest in the already introduced stock of Statoil, it is often assumed that Oslo Stock Exchange is highly influenced by the price of oil. One possible relation might be that a high oil price would be good news for Norwegian based companies since several companies are involved with drifting of platforms and export of oil. This could lead to a strong Norwegian currency, however, a strong Norwegian currency would make the Norwegian based companies less competitive on international market, which again could be bad for the Norwegian Stock Market.

As it turns out there are numerous factors to take into consideration when trying to describe the relationship between different processes. Luckily for us there are several methods for describing correlation and dependence. The following chapters will focus on methods of describing and modeling dependence, starting off with a recap of Pearson's correlation coefficient which is a usual part of the curriculum of introductory courses in statistics.

3.1 Pearson's Correlation Coefficient

Pearson's correlation coefficient is together with expectation, variance and covariance, a natural topic in introductory courses in statistic. It is a measure for linear correlation on the interval [-1,1], where -1 or 1 means the observations are on a straight line.

Following is the Pearson's correlation coefficient for the random variables X and Y with their respective standard deviations σ_x and σ_y as given in Casella & Berger (2002)

$$\rho = \frac{E[(X - E[X])(Y - E[Y])]}{\sqrt{Var[X] \times Var[Y]}} = \frac{Cov[X,Y]}{\sigma_x \sigma_y} \tag{3.1}$$

where $\sigma_x < \infty$ and $\sigma_y < \infty$.

Or in the case of a sample observed from X and Y we have the estimate

$$\hat{\rho} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{3.2}$$

There are some properties concerning the Pearson's correlation coefficient and independence that should be mentioned. These are also often part of the curriculum of introductory courses, but their importance cannot be stressed enough. Not being aware of these might very well lead to incorrect conclusions.

Lemma 3.1.1. ***Independence implies zero correlation***
Let X and Y be to independent random variables. Then the correlation coefficient $\rho = 0$.

The next property is a result of Pearson's correlation coefficient being an estimator of linearly dependence. This means that we may experience ρ being small for heavily dependent observations. In the extreme case we may actually experience ρ being 0, even if the random variables are strongly dependent. A classical example of this is given below.

Example 3.1.2. Zero correlation does not imply independence
Let X be given by an N(0,1) distribution and $Y = X^2$ which means that it is given by a $\chi^2(1)$ distribution.
From the distributions we have the following:

$$E[X] = E[X^3] = 0$$
$$E[Y] = 1.$$

Clearly X and Y are dependent, but what about their correlation?

$$Cov[X,Y] = E[(X - E[X])(Y - E[Y])]$$
$$= E[X(Y-1)) = E[XY] = E[X^3] = 0$$

Pearson's correlation coefficient ρ is therefore zero

Because of this, independence should not be concluded by ρ being zero, but $\rho \neq 0$ does imply the variables being dependent.

3.1.1 Pearson's Correlation Coefficient and the Gaussian Distribution

We will be using the same parameterizing for the following two definitions of the multivariate Gaussian distribution as used in Rizzo (2007). In the case of a bivariate distribution it is given

$$\psi(x,y|\theta) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} exp\Big\{-\frac{1}{2(1-\rho^2)}\Big[\big(\frac{x-\mu_x}{\sigma_x}\big)^2 - 2\rho\big(\frac{x-\mu_x}{\sigma_x}\big)\big(\frac{y-\mu_y}{\sigma_y}\big) + \big(\frac{y-\mu_y}{\sigma_y}\big)^2\Big]\Big\}. \tag{3.3}$$

where θ is $[\mu_x, \mu_y, \sigma_x, \sigma_y, \rho]$

As you can see the correlation coefficient ρ is a part of the bivariate Gaussian distribution, and this is the case in in higher dimensions too.

Similar the density function for a multivariate Gaussian distribution is given

$$\psi(\mathbf{x}|\Sigma,\mu) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} exp\{-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)\} \tag{3.4}$$

where $\mathbf{x} = x_1, \ldots, x_d$, Σ is a $d \times d$ nonsingular covariance matrix and $\mu = (\mu_1, \ldots, \mu_d)^T$.

Each of the marginal distributions in a multivariate Gaussian distribution is Gaussian distributed $N(\mu_i, \sigma_i^2)$. Similar two marginal distributions from a multivariate Gaussian distributed are distributed by a bivariate Gaussian distribution. So in a multivariate Gaussian distribution the same relation as described in the bivariate case applies for each of the possible connection of the margins. Therefore, in the case of multivariate Gaussian distributions Pearson's correlation coefficient is a natural part of the distribution.

Following are some plots of samples generated in R with mvrnorm, from the library MASS. These represent six separate bivariate random variables, all of them with a Gaussian distribution. the parameters given below, and ρ as given in the plot.

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \mu = \begin{pmatrix} 0 & 0 \end{pmatrix}$$

As we can see when the correlation is 1 or -1, the sample is on a straight line.
Here $\rho = 1$ results in the line increasing along the $x = y$-axis, while $\rho = -1$ makes high values in the x-axis result in low values at the y-axis and vice versa. This strong dependence is relaxed as the correlation goes to zero.

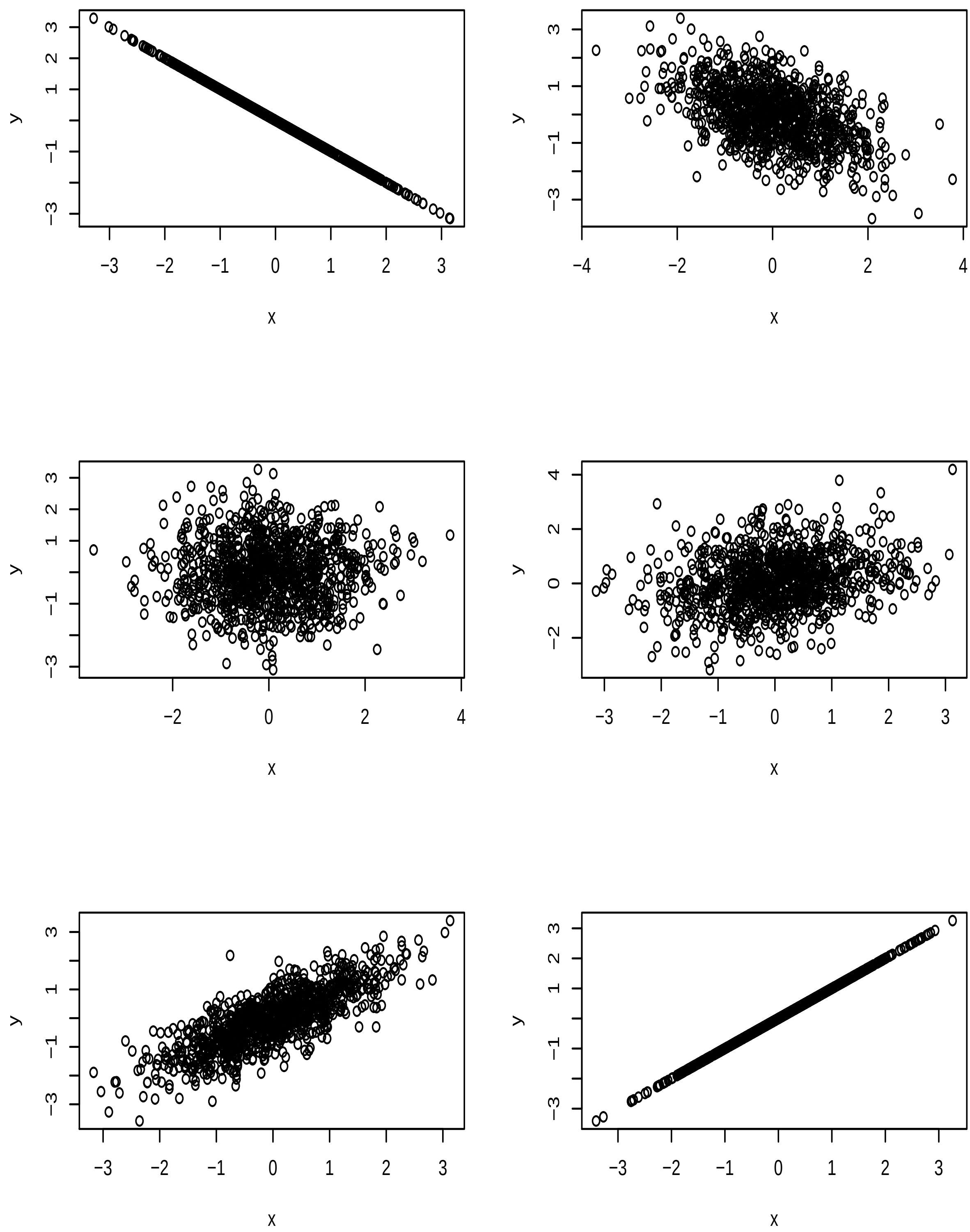


Figure 3.1: Sample from bivariate Gaussian distributed variables with $(-1, -0.5, 0, 0.2, 0.8, 1)$ as corresponding values for ρ

Chapter 4

Copulas

The copula is a topic that has had growing popularity in recent time, especially in the field of finance. The first appearance of the word copula as a way of joining different marginal distribution together in a multivariate distribution was in Sklar (1959). This article was also the first appearance of the theorem known as Sklar's theorem, the theorem that has made the copula into what it is today.

However it should be mentioned that mathematicians like Féron and Fréchet deserve some acknowledgment. Sklar knew that Fréchet was working on theory of connecting multivariate distributions with their margins, after reading Feron (1956), where Féron introduced related theory in three dimensions, Sklar extended it to a general dimension and wrote to Fréchet about it. It is essentially this correspondence between Sklar and Fréchet which were published as Sklar (1959). The history of copulas as described by Sklar himself, can be found in Sklar (1996).

Copulas allow us to choose between different dependence structures for our marginal distribution. Say that we have a portfolio consisting of multiple stocks, and we find that the Gaussian distribution describes the movement for these stocks in a satisfying way, however the relationships between the stocks are not Gaussian, which means that the multivariate Gaussian distribution may not be a good fit for modeling this portfolio. The copula offers a way of combining the univariate margins, where their relationship is described by copulas. With this approach, we can choose between endless combinations of marginal distributions and copulas, and this is one of the main reasons why copulas have really been embraced in the field of finance.

We will use the definition of copulas as given in McNeil *et al.* (2005)

Definition 4.0.3. *Copulas*
A d-dimensional copula is a cumulative distribution function on $[0,1]^d$ *with standard uniform marginal distributions.*

As a consequence of this a copula inherits the following properties:

- $C(u_1, \dots, u_d)$ is increasing in each component u_i
- $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$ for all i $\in \{1, \dots, d\}, u_i \in [0,1]$.
- For all $(a_1, \dots, a_d)$,$(b_1, \dots, b_d) \in [0,1]^d$ with $a_i \leq b_i$ we have $\sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+\cdots+i_d} C(u_{1i_1}, \dots, u_{di_d}) \geq 0$, where $u_{j1} = a_j$ and $u_{j2} = b_j$ for all $j \in \{1, \dots, d\}$

The last one is probably the least obvious one, but it ensures that the probability of observing a vector from its copula is non-negative. These properties are found at McNeil *et al.* (2005).

Theorem 4.0.4. *(Sklar's Theorem)*
Let $\mathbf{F}$ *be a d-dimensional cumulative distribution function with univariate margins* $F_1, \dots, F_d$. *Then there exists a copula* $C : [0,1]^d \to [0,1]$ *such that, for all* $x_1, \dots, x_d \in \bar{\mathbb{R}} = [-\infty, \infty]$,

$$\mathbf{F}(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$$

In the case of continuous margins, the copula is unique, and in the discrete case it is uniquely determined on Range $F_1 \times$ *Range* $F_2 \times \cdots \times$ *Range* F_d.

Sklar's theorem as stated above is from Durante & Sempi (2010), for a proof of the theorem in the case of continuous margins see McNeil *et al.* (2005).

There are numerous versions of copulas of interest available, but this text will be restricted to a few of them. We will introduce the independence, the comonotonicity and the countercomonotonicity copula from the class of fundamental copulas, followed by the Gaussian and t copula from the category of implicit copulas, and Gumbel, Clayton and Frank as examples of Archimedean copulas. The definitions will be given for a multivariate case, but for the different copulas our applications will be restricted to the bivariate case. Note that the definition of the copulas mentioned above is based on the form of McNeil *et al.* (2005) if not stated otherwise.

4.1 Fundamental Copulas

Fundamental copulas are special cases of copulas on their own, but may also appear as special cases for other copulas. The Comonotonicity copula and the Countermonotonicity which are two of the fundamental copulas we will introduce are closely related to what is known as Fréchet bounds. Fréchet bounds as described in McNeil *et al.* (2005) states that for every copula $C(u_1, \ldots, u_d)$ we have the following bounds for the dependence:

$$max\{\sum_{i=1}^{d} u_i + 1 - d, 0\} \leq C(\mathbf{u}) \leq min\{u_1, \ldots, u_d\} \tag{4.1}$$

The Independence copula is defined as

$$C(u_1, \ldots, u_d) = \Pi(u_1, \ldots, u_d) = \Pi_{i=1}^{d} u_i \tag{4.2}$$

As the name suggest there are no relationship between the margins in the independent copula. The upper bound in Fréchet bounds is a multivariate copula by itself and is known as the Comonotonicity copula

$$C(u_1, \ldots, u_d) = M(u_1, \ldots, u_d) = min\{u_1, \ldots, u_d\}. \tag{4.3}$$

In the same way the lower bound in Fréchet bounds is a copula, but only in the bivariate case. It is known as the Countermonotonicity copula

$$C(u_1, u_2) = W(u_1, u_2) = max\{u_1 + u_2 - 1, 0\} \tag{4.4}$$

An extension of the countermonotonicity concept for a dimension higher than two is not possible, this is shown in McNeil *et al.* (2005).

4.2 Implicit Copulas

Let $\mathbf{X}$ be a multivariate random vector with joint cdf denoted by $\mathbf{F}(\mathbf{x})$ with $F(x)$ as the corresponding cdf for the margins, then implicit copulas are copulas on the form

$$C(\mathbf{u}) = \mathbf{F}(F^{-1}(u_1), \ldots, F^{-1}(u_d)). \tag{4.5}$$

The most common implicit copulas are probably the Gaussian copula and the t copula.

From the equation above, the bivariate Gaussian copula can be written

$$C_{\rho}^{Ga}(u_1,u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi(1-\rho^2)^{1/2}} exp\{\frac{-(s_1^2 - 2\rho s_1 s_2 + s_2^2)}{2(1-\rho^2)}\} ds_1 ds_2. \quad (4.6)$$

where Φ^{-1} is the quantile function of a standard Gaussian distribution. Similar the bivariate student-t copula is given

$$C_{\upsilon,\rho}^{t}(u_1,u_2) = \int_{-\infty}^{t_{\upsilon}^{-1}(u_1)} \int_{-\infty}^{t_{\upsilon}^{-1}(u_2)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \{1 + \frac{(s_1^2 - 2\rho s_1 s_2 + s_2^2)}{\upsilon(1-\rho^2)}\}^{-\frac{\upsilon+2}{2}} ds_1 ds_2 \quad (4.7)$$

Where t_{υ}^{-1} is the quantile function of a standard univariate t distribution.

Example 4.2.1. Relationship between the bivariate Gaussian copula and Fundamental copulas
If the correlation ρ is equal to 0 in Equation 4.6, we have that

$$\begin{aligned} C_{0}^{Ga}(u_1,u_2) &= \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi} exp\{\frac{-(s_1^2 + s_2^2)}{2}\} ds_1 ds_2 \\ &= \int_{-\infty}^{\Phi^{-1}(u_1)} \frac{1}{\sqrt{2\pi}} exp\{\frac{-s_1^2}{2}\} ds_1 \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{\sqrt{2\pi}} exp\{\frac{-s_2^2}{2}\} ds_2 \\ &= u_1 u_2, \end{aligned}$$

which is equal to the bivariate independence copula defined above.
Similar if ρ is 1, the bivariate Gaussian copula is equal to the comonotonicity copula and in the case of ρ equal to -1, it gives the countercomonotonicity copula.

4.2.1 Sample from Implicit Copulas

The form of the implicit copulas will in many cases make them quite easy to sample from. If $\mathbf{X}$ is a random vector from a multivariate distribution, then

$$\mathbf{U} = (F(x_1), \ldots, F(x_d))$$

where $F(x)$ is the cdf of the margins, gives a sample from the given copula.
This means that if we can sample from its multivariate distribution and calculate the cdf we can generate a sample from its implicit copula.

Following is two figures, the first is of six different Gaussian copulas, while the second one is six different student-t copulas. Each of the plots consists of 1000 observations from each copula. The original sample from the multivariate Gaussian distribution was generated by the function mvrnorm in the MASS library, with the parameters given below, and ρ as given in the plot.

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \mu = \begin{pmatrix} 0 & 0 \end{pmatrix}$$

The sample from the student-t distribution was generated by the function rmvt from the mvtnorm library, with 1 degree of freedom and with

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

and ρ as given in the plot.
Notice that Σ does not correspond to the covariance matrix, for this distribution and the variance is not even defined when the number of degrees of freedom is less than 2.

When the rho parameter is 0.8 for the copulas, the Gaussian copula start to resemble a line in the $x = y$-axis, while the t copula still has quite a few observations along the opposite diagonal. In the case of rho being zero, we see that the Gaussian copula has observations all over the plane, while the t copula has less values in the area of say $x = 0$ and $y = 0.5$. This is because unlike the Gaussian copula, $\rho = 0$ does not lead to independence between the marginals for the t copula. This can be verified by setting $\rho = 0$ in Equation 4.7.

Gaussian Copula

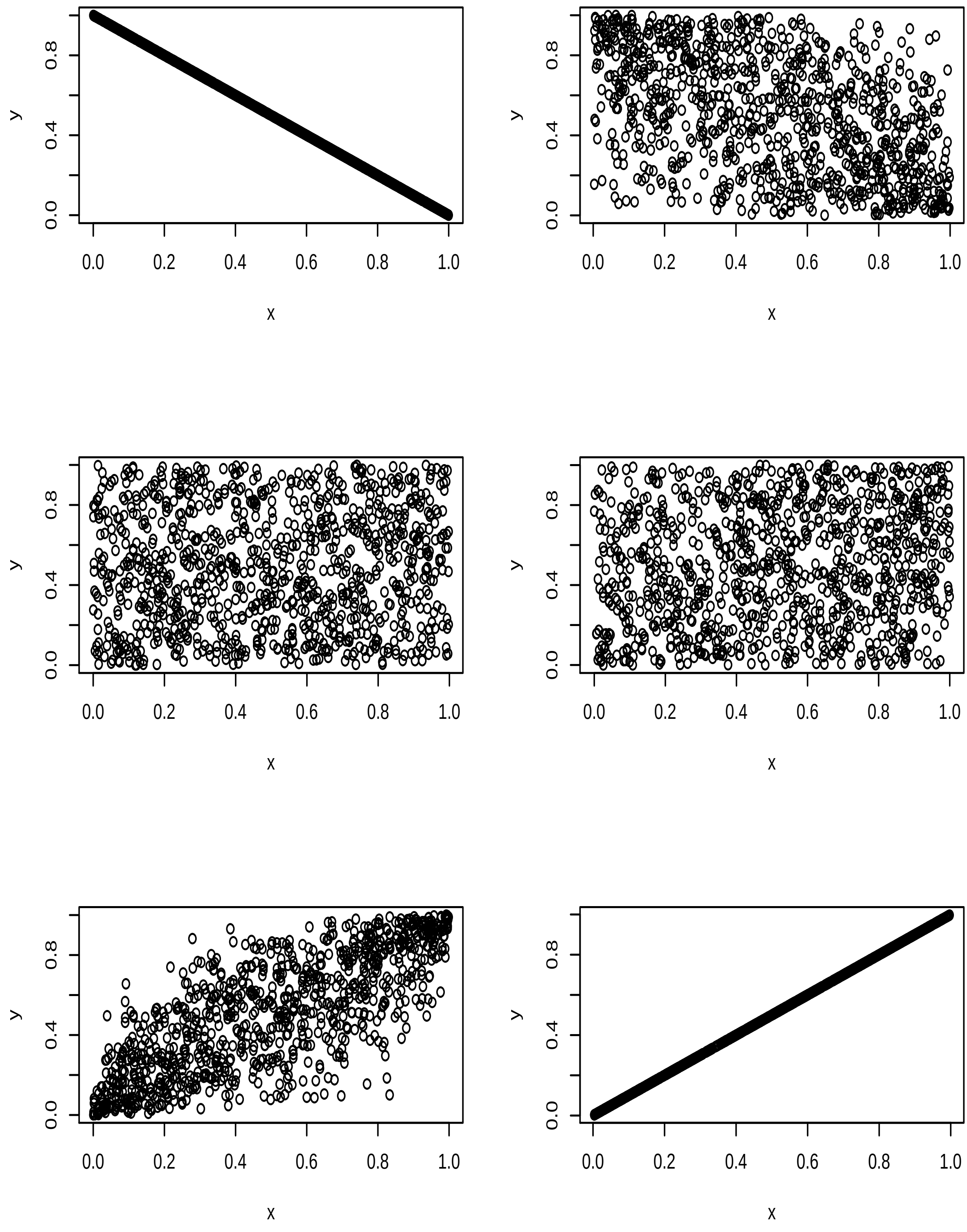


Figure 4.1: Sample from bivariate Gaussian copulas with $(-1, -0.5, 0, 0.2, 0.8, 1)$ as corresponding values for ρ

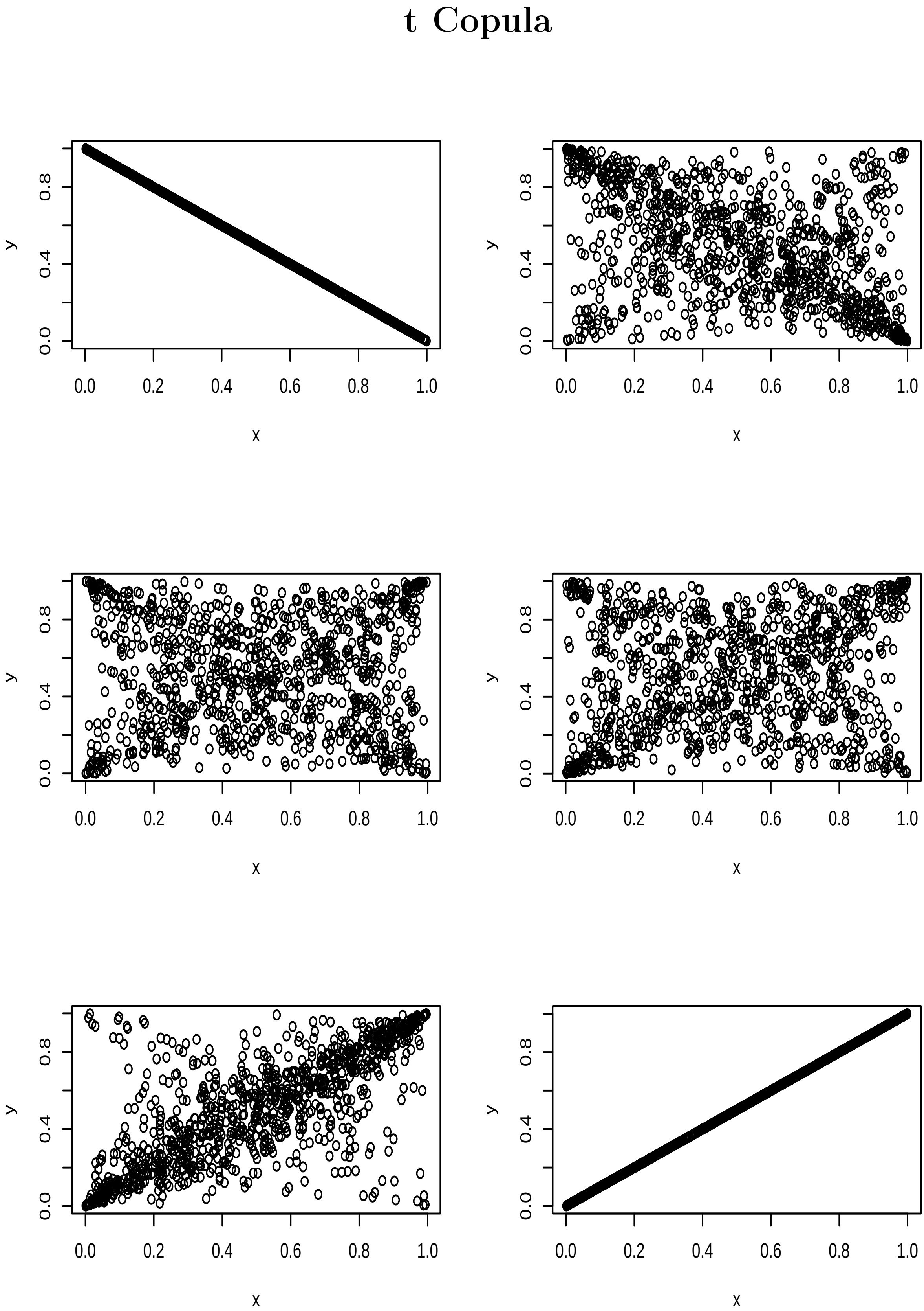


Figure 4.2: Sample from bivariate t copulas with $(-1, -0.5, 0, 0.2, 0.8, 1)$ as corresponding values for ρ

4.3 Archimedean Copulas

Archimedean copulas are copulas on the form

$$C(u_1, \ldots, u_d) = \phi^{-1}\{\phi(u_1) + \ldots + \phi(u_d)\} \tag{4.8}$$

Where ϕ is known as the generator function of the copulas. The following restrictions to the generator function apply:

1. ϕ is a continuous strictly decreasing and convex function mapping $[0, 1]$ onto $[0, \infty]$
2. $\phi(0) = \infty$
3. $\phi(1) = 0$

The second restriction for the generator function may in some cases be relaxed. An example of this is shown in the table below. Here $\theta \geq -1$ are possible values for θ in the Clayton copula. But for $\theta < 1$ the generator function does not satisfy $\phi(0) = \infty$ and is therefore not called strict.

We will in this thesis restrict ourselves to the Archimedean copulas; Gumbel, Clayton and Frank as given below. For a large list over different Archimedean copulas and their generators see Durante & Sempi (2010) where they have examples of 22 different generator functions.

Table 4.3.1. *Some Archimedean bivariate copulas*

Copula	$C(u,v)$	$\phi(t)$	*Parameter range*	*strict*
Gumbel	$exp(-[(-log(u))^\theta + (-log(v))^\theta]^{1/\theta})$	$(-log(t))^\theta$	$\theta \geq 1$	*yes*
Clayton	$[max(u^{-\theta} + v^{-\theta} - 1, 0)]^{-1/\theta}$	$\frac{1}{\theta}(t^{-\theta} - 1)$	$\theta \geq -1$	$\theta \geq 0$
Frank	$-\frac{1}{\theta}log(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1})$	$-log(\frac{e^{-\theta t}-1}{e^{-\theta}-1})$	$\theta \in \mathbb{R}$	*yes*

Unfortunately sampling from Archimedean copulas are a little bit more complex than in the case of implicit copulas, since these are not based on well known multivariate distributions that we can sample from. However, there are methods of doing this, McNeil *et al.* (2005) offers an algorithm based on laplace-stieltjes transformations of known distributions to sample from the given copula, while an approach based on conditional copula for simulating from different copulas are covered in Genest & MacKay (1986).

Example 4.3.2. Frank copula

As an example of Archimedean copulas we will take a closer look at the Frank copula in the bivariate case. We will derive the expression for the copula given in Table 4.3.1, where the generator function is $-log(\frac{e^{-\theta t}-1}{e^{-\theta}-1})$. We then have to derive the inverse of the generator function, this is straight forward because of the first requirement of the generator function.

$$
\begin{aligned}
U &= -log(\frac{e^{-\theta t}-1}{e^{-\theta}-1}) \\
e^{-U} &= \frac{e^{-\theta t}-1}{e^{-\theta}-1} \\
log(e^{-U}(e^{-\theta}-1)+1) &= -\theta t \\
t &= \frac{-log(e^{-U}(e^{-\theta}-1)+1)}{\theta} \\
&= \phi^{-1}(U)
\end{aligned}
$$

With this expression inserted in the formula for the Archimedean copulas given in Equation 4.8, we end up with the formula for the Frank copula given in Table 4.3.1.

Unfortunately the algorithm for simulating from a Frank copula described in McNeil *et al.* (2005) does not allow the parameter θ to be negative.
This is restricting because $\theta \in \mathbb{R}$ for the Frank copula. And $\theta < 0$ result in the Frank copula having negative correlation, which is quite an important factor when investigating dependence. Luckily the algorithm described in Romano (2002) allows θ to be negative.

The following plot is generated by the algorithm for the Frank copula given in Romano (2002), which is based on Genest & MacKay (1986). Each plot consist of 1000 values where θ is as given in the plot. As we can see large values of θ, both positive and negative makes the values appear on a straight line. When theta is 1 the Frank copula has some similarities with the Gaussian copula when the correlation is zero. There are some similarities when theta and the correlation gets larger as well, however the Frank copula appears to have its values sampled equally on a straight line, while the Gaussian copula seems to have more values appearing around origo and $x = y = 1$.

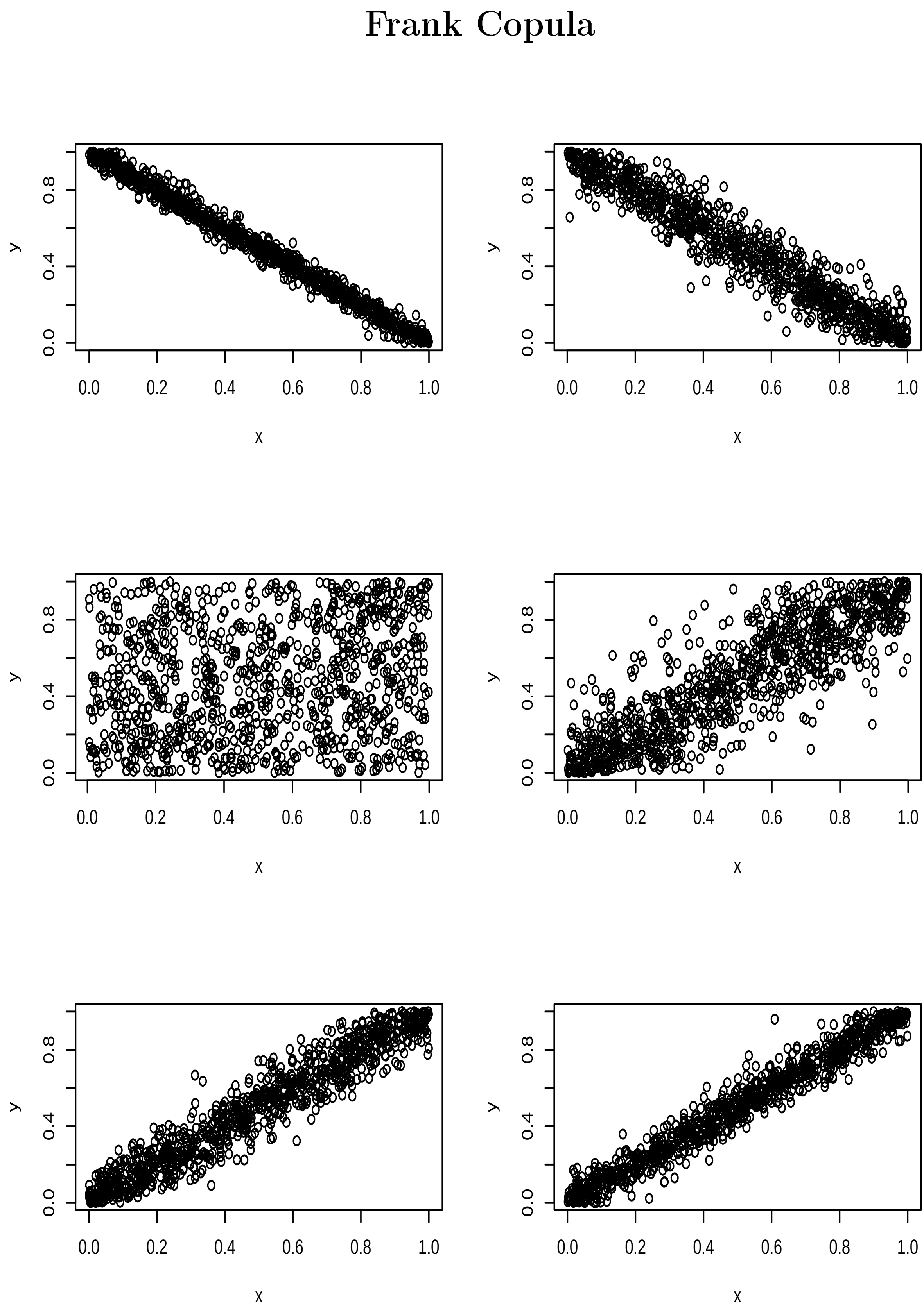


Figure 4.3: Sample from bivariate Frank copulas with $(-50, -20, 1, 10, 20, 30)$ as corresponding values for θ

4.4 Meta-Distributions

An interesting application of Sklar's theorem is the possibility to create so called meta-distributions. Sklar's Theorem 4.0.4 states the existence of a copula binding the univariate margins together so that it equals the joint distribution.

An additional property from this theorem, is the reverse statement:
Let C be a copula and $F_1, F_2, \ldots, F_d$ be univariate distributions, then the cumulative distribution function $\mathbf{F} = C(F_1, F_2 \ldots F_d)$ is a d-dimensional joint distribution with $F_1, F_2, \ldots, F_d$ as its margins.

This means that we can create distribution that is a combination of a copula with the margins of our choice. Giving the possibility to create for example a combination of Gaussian copula with exponential margins. This particular combination is used in what is known as Li's model, Li (1999). This model has gotten an awful lot of critic following the latest financial crisis, Financial times even got as far as publishing an article about the model with the title "'The formula that felled Wall St"' Jones (2009).

The problem with this model is that the Gaussian copula may have been a good description of the dependence under normal circumstances, but the Gaussian copula has a constant correlation. Unfortunately the correlation is not always constant, and it is often the case that the correlation is stronger when things go bad than in good times. This means that this model might work when modeling periods of positive returns, but will underestimate the potential risk in bad times. Whether it is the model itself or the people apparently using it without knowing its limitations that deserve the blame is an open question. However, what is certain is that meta-distribution is a useful way of constructing multivariate distributions with different combinations of copula and margins.

Example 4.4.1. Meta-Gaussian distribution
If we want to sample from a meta-Gaussian distribution with say exponential(β) distributed margins, we first need to get a sample $\mathbf{X}$ from a multivariate Gaussian distribution, this can be done in **R** with the function mvrnorm from the package MASS.
Then $F^{-1}(\Phi(X_i))$ where $F^{-1}(y)$ is the univariate quantile function for an exponential(β) distribution and $\Phi(Z)$ is the cumulative distribution function for a univariate Gaussian distribution is one observation from the meta-Gaussian distribution with exponential distributed margins.

Following are some random samples consisting of 1000 observations from the distribution described above. The exponential distributions parameter equals 1 and the correlation parameter rho is given in the plot.

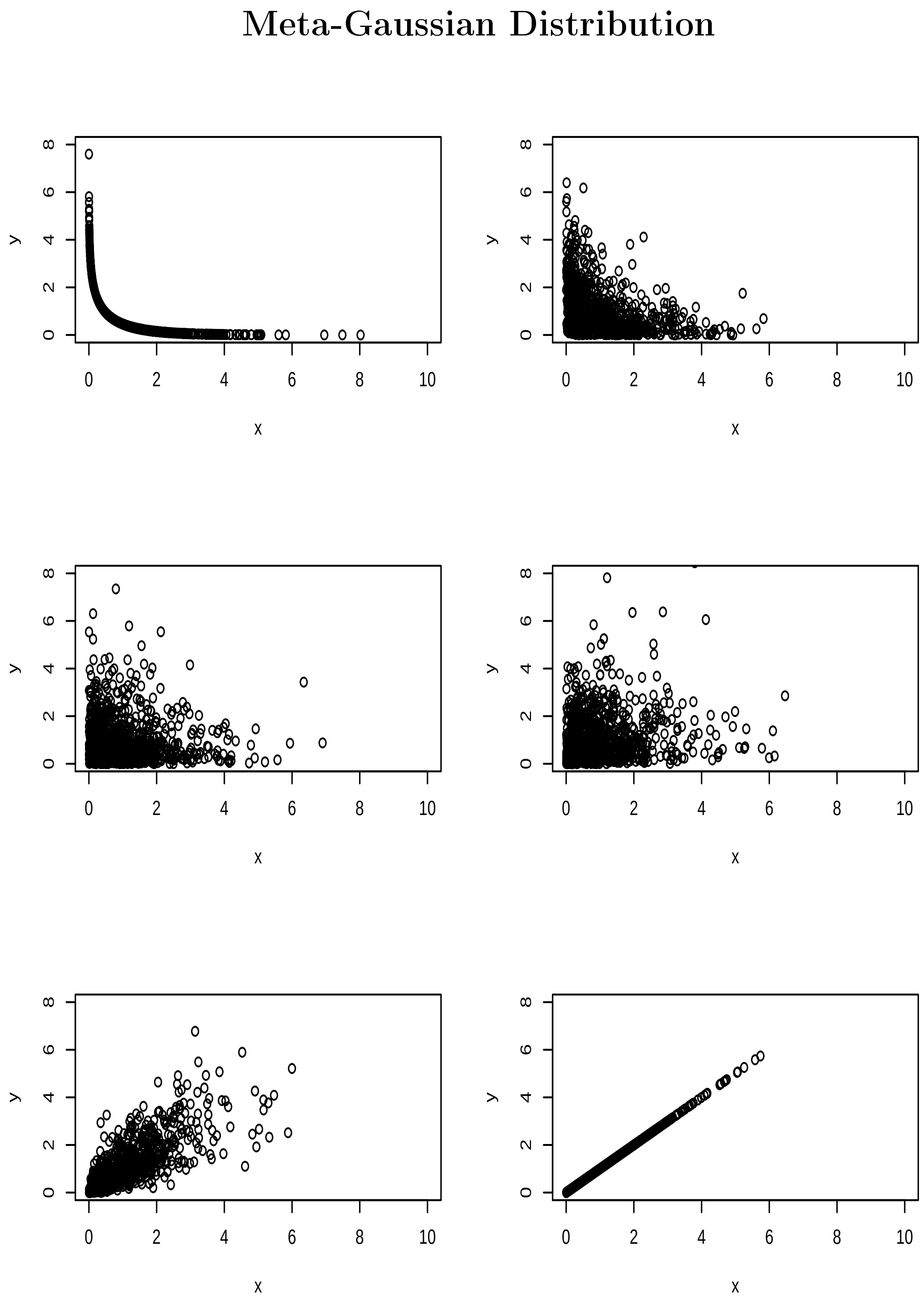


Figure 4.4: Samples from 6 different meta distributions constructed from Gaussian copulas with exponential(1) margins and with $(-1, -0.5, 0, 0.2, 0.8, 1)$ as corresponding values for ρ

4.5 Rank Correlation

Rank correlation coefficients are measures of dependence that are based on the rank of the observations, where the rank for a value Y_i from a sample $Y_1, \ldots, Y_n$ is

$$rank(Y_i) = \sum_{j=1}^{n} I(Y_j \leq Y_i). \tag{4.9}$$

This might make rank correlation measures a preferred choice over Pearson's correlation coefficient when fitting copulas to data. While Pearson's correlation coefficient is influenced by the margins of the copulas, coefficients based on rank correlation only depend on the copula itself. This is because rank correlation under a transformation by increasing monotonic functions preserve the samples rank correlation, and as a consequence of this the correlation coefficients based on rank correlation only depend on the copula. This lack of influence by the marginals makes rank correlation useful when fitting copulas to data.

An important concept when it comes to rank correlation is concordance and disconcordance. This is a property of the relationship between two points, and we will use the definition given in Roger (2006).

Definition 4.5.1. *Concordance and disconcordance*
- two points in $\mathbb{R}^2$, (X, Y) and (X^*, Y^*) are said to be concordant if $(X - X^*)(Y - Y^*) > 0$ or they are said to be disconcordant if $(X - X^*)(Y - Y^*) < 0$

As examples of rank correlation coefficients we will introduce the measures Kendall's tau and Spearman's rho. As with Pearson's correlation coefficient, both the Kendall's tau correlation coefficient and the Spearman's rho correlation coefficient are on the interval [-1,1].

We will in the following introduction of Kendall's tau and Spearman's rho use the definition as given in Ruppert (2010).

4.5.1 Kendall's tau

Let (X, Y) be a bivariate random vector and let (X^*, Y^*) be an independent copy. Then Kendall's tau for (X, Y) are as follows

Definition 4.5.2.

$$\begin{aligned}\rho_\tau(X,Y) &= P((X-X^*)(Y-Y^*)>0) - P((X-X^*)(Y-Y^*)<0)\\ &= E(sign((X-X^*)(Y-Y^*)))\end{aligned}$$

Which is actually the probability of concordance - the probability of discordance for (X, Y) and (X^*, Y^*).

For an bivariate sample $(Y_{i,1}, Y_{i,2})$, where $i = 1, \ldots, n$, the sample Kendall's tau is

$$\hat{\rho}_\tau(\mathbf{Y}_1, \mathbf{Y}_2) = \binom{n}{2}^{-1} \sum_{i=1}^{n-1} \sum_{j=i}^{n} sign((Y_{i,1} - Y_{j,1})(Y_{i,2} - Y_{j,2})).$$

4.5.2 Spearman's rho

Definition 4.5.3. *Spearman's rho*
For the two random variables X and Y with the cumulative distribution functions F_X and F_Y Spearman's rho is given by

$$\rho_S(X,Y) = \frac{Cov(F_X(X), F_Y(Y))}{\sqrt{Var(F_X(X))Var(F_Y(Y))}} \tag{4.10}$$

This means that Spearman's rho is actually Pearson's correlation on the copula in the continuous case, and in the case of discrete margins Spearman's rho is Pearson's correlation on the transformed random variables.

Or for an observed bivariate sample $(Y_{i,1}, Y_{i,2})$, where $i = 1, \ldots, n$, Spearman's correlation coefficient can be computed as follows.

$$\hat{\rho}_s(\mathbf{Y}_1, \mathbf{Y}_2) = \frac{12}{n(n^2-1)} \sum_{i=1}^{n} \{rank(Y_{i,1}) - \frac{n+1}{2}\}\{rank(Y_{i,2}) - \frac{n+1}{2}\} \tag{4.11}$$

This is Pearson's sample correlation for a sample on the ranks of $Y_{i,1}$ and $Y_{i,2}$.

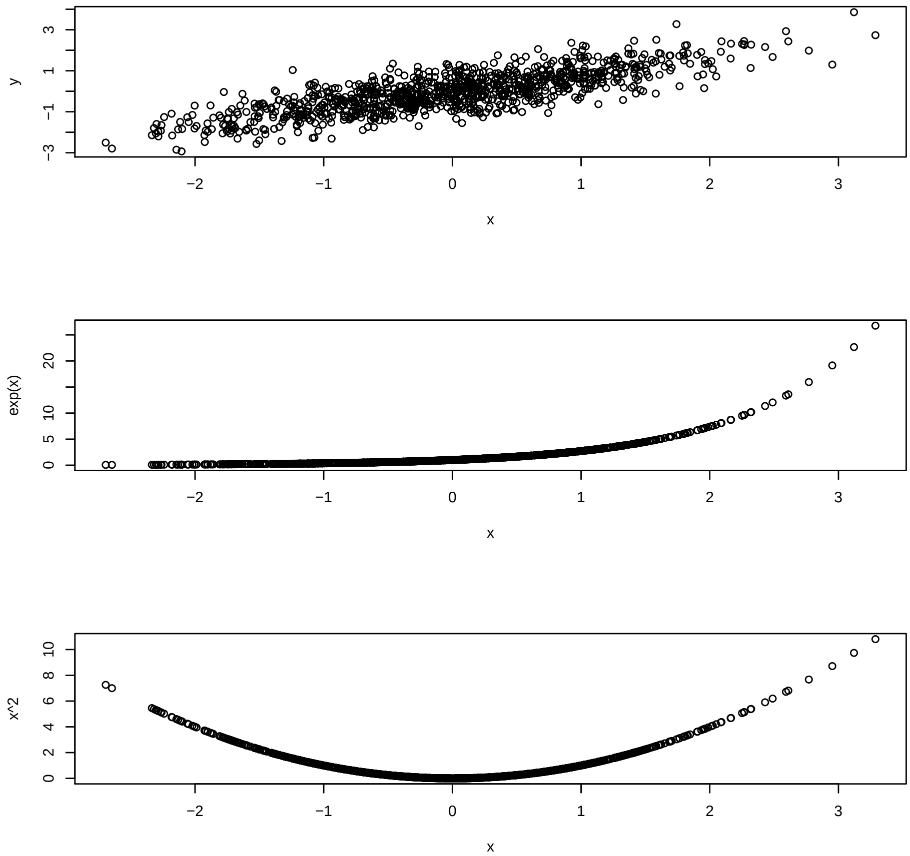
4.5.3 Examples on Spearman's rho and Kendall's tau

To get a picture of how Pearson's, Spearman's and Kendall's correlation coefficients compares to one another, the following is a table of three different examples based on the Gaussian distribution and the correlation calculated with the three coefficients. The actually calculation of the correlation was done in **R** by the corr function, where you can choose between Pearson's, Kendall's and Spearman's coefficient. The following plot is given as the first plot correspond to Test1, second plot to Test2 and third plot to Test3, where Test1, Test2 and Test3 are defined as below.

- Test1 = {'Sample of 1000 values from bivariate Gaussian with correlation 0.8'}
- Test2 = {**X** ='Sample of 1000 values from standard Gaussian' and **Y** $= exp(X)$}
- Test3 = {**X** ='Sample of 1000 values from standard Gaussian' and **Y** $= X^2$}

We can see in Figure 4.5 that all of the above captures the correlation quite well, when it is sampled from a bivariate Gaussian with correlation 0.8, although Kendall is a bit lower than the other two. In the case of an exponential transformation of the sample both Spearman and Kendall estimates the correlation to be one, while Pearson's which is restricted to linear correlation only give correlation approximately equal to 0.77. And in the last example where Y is the square of the Gaussian sample all of the correlation coefficients are close to zero. This coincides with the theoretical values which means that correlation equal to zero does not imply independence for either of the three introduced correlation coefficients.

Examples on different correlation coefficients



	Pearson	Kendall	Spearman
Test1	0.80	0.58	0.77
Test2	0.77	1.00	1.00
Test3	0.03	-0.03	-0.03

Figure 4.5: Pearson Kendall and Spearman on random samples
The upper plot corresponds to Test1, the middle to Test2 and the lower to Test3.

Chapter 5

The Capital Asset Pricing Model

Portfolio theory and the Capital Asset Pricing Model(CAPM) has been credited with making the field of finance into a field of science. The following quote from Merton (1990) describes the field before the introduction of the mentioned theory:
"As recently as a generation ago finance theory was still little more than a collection of anecdotes, rules of thumb, and manipulations of accounting data. The most sophisticated tool of analysis was discounted value and the central intellectual controversy centered on whether to use present value or internal rate of return to rank corporate investments."

One would believe that Merton knew what he was talking about for in 1997 he was, together with Scholes awarded the Nobel Memorial Prize in Economic Sciences, for their work on the famous Black-Scholes formula, Jarrow (1999).

The Capital Asset Pricing Model is a model used in order to establish a connection between the risk of an asset and its expected return. If we are considering investing in a specific stock with a given risk, CAPM can be used to decide if the expected return is worth the risk. Before going into the mathematical parts of this model, we will introduce some necessary definitions and clarify some crucial assumptions that are needed for the CAPM to be valid.

5.1 Essential Definitions for the Model

We will give a short introduction to some of the necessary definitions used in the Capital Asset Price Model, based on the definition given in Lee (2006), which also offers additional properties of the following definitions.

Market Portfolio

The market portfolio is comprised of all risky assets weighted in proportion to their market value. The market portfolio has no unsystematic risk. We will later be doing analysis on the Norwegian Stock Market and will then be using the OSEBX index as an approximation for the market portfolio.

Efficient Portfolio

A portfolio is efficient if no other portfolio has the same expected return at a lower variance of returns.

Riskless Rate

The riskless or risk-free rate is defined as the interest rate that can be earned with certainty. There is no risk associated with the riskless rate(at least in theory).

5.2 Assumption for the CAPM

The validity of the CAPM is dependent on some assumptions stated below, these are more or less equal to the version found in Ruppert (2010)

1. The market prices are based on the fact that supply equals demand.
2. Everyone has the same forecast of expected returns and risks.
3. All investors have a portfolio consisting of combinations of risky assets as well as the risk-free asset.
4. The market rewards people for assuming unavoidable risk, but there is no reward for needless risk.

It is obvious that in a real life situations all of these assumptions will not hold, however with a little caution the following theory might be helpful in asset speculation and risk measuring. We will start the introduction of CAPM by derivation of the Capital Market Line and The Security Market Line.

The following derivations of these are based on the derivation given in Ruppert (2010), with some minor changes based on Berk & DeMarzo (2007).
The Capital Market Line let us estimate the expected value of an efficient portfolio based on the risk of the efficient portfolio, the risk-free rate and the expectation and variance of the market portfolio. This particular line will not be of major importance later on and will with the exception of the example given below only be used as an part of the derivation of the Security Market Line, which is our main focus from the theory of CAPM.

The Security Market Line is used to give a direct link between the expected return and the risk of a risky asset. We will limit the examples on real life data in this section to some discussion about its use in Silvapulle & Granger (2001), however we will return to the CAPM when doing empirical analysis on the Norwegian Stock Market.

5.3 The Capital Market Line

Consider an efficient portfolio that portion a part $w \in [0,1]$, of its assets to the market portfolio and (1-w) to the risk-free asset. We write the return of this portfolio as

$$R = wR_M + (1-w)\mu_f = \mu_f + w(R_m - \mu_f).$$

Where R is the return of the efficient portfolio, μ_f is the risk-free rate and R_M is the return on the market portfolio. We let

$$\mu_R = E[R],\ \sigma_R = \sqrt{Var[R]},$$

$$\mu_M = E[R_M] \text{ and } \sigma_M = \sqrt{Var[R_M]}.$$

Then the expectation of this portfolio can be written

$$\mu_R = \mu_f + w(\mu_M - \mu_f).$$

while the standard deviation is $\sigma_R = w\sigma_M$ and we have that $w = \frac{\sigma_R}{\sigma_M}$.

The value for w inserted in the formula for the expectation yields

$$\mu_R = \mu_f + \frac{\sigma_R}{\sigma_M}(\mu_M - \mu_f). \tag{5.1}$$

And this formula for the expectation of the return of the efficient portfolio is known as The Capital Market Line. It is a link in the μ_R - σ_R plane, and the slope of the line is

$$\frac{\mu_M - \mu_f}{\sigma_M}, \tag{5.2}$$

which is the ratio of the risk premium to the standard deviation of the market portfolio.

We can easily see that when $\sigma_R = 0$, i.e. the standard deviation or volatility of the efficient portfolio is zero, the expectation of the efficient portfolio equals the expectation of the risk-free asset. This makes sense, if we refuse to take any risk our only choice is the risk-free asset.

When $\sigma_R = \sigma_M$ the expectation of the efficient portfolio equals the expectation of the market portfolio. This means that the Capital Market Line is the line from the risk-free rate trough the market portfolio. And it gives us the value of the highest possible expected return for any level of volatility.

Example 5.3.1. Investing by the CML
Say we have invested an amount in a stock where we expect a return of 10 with a volatility of 15%. By saving the same amount in a saving account our rate is 5%, but we suppose this is risk-free i.e. the volatility is 0. We may also put our money in the market portfolio, which has expected return of 12 and volatility 10%.

We want to duplicate the expected return of the stock by rather investing our money in the savings account and the market portfolio, by Equation 5.1 we have

$$10\% = 5\% + (12\% - 5\%)\frac{\sigma_R}{10}$$

$$\sigma_R = 50/7 \approx 7.14$$

This means that, if we invest 0.714 of our amount in the market portfolio and the remaining 0.286 we put in the savings account, the expected return matches the expected return for the stock, but the volatility of 7.14%, is lower than the volatility for the stock, hence, this is a better investment than the stock.

5.4 Security Market Line

Consider a portfolio P consisting of two assets, the market portfolio and the i-th risky asset. The return of this portfolio is

$$R_P = w_i R_i + (1 - w_i) R_M.$$

Where w_i is a weight $\in [0, 1]$, R_i is the i-th risky asset and R_M is the market portfolio. We let

$$\mu_i = E[R_i],\ \sigma_i = SD[R_i],$$

$$\mu_M = E[R_M] \text{ and } \sigma_M = SD[R_M].$$

The expectation of this portfolio can be written

$$\mu_P = w_i \mu_i + (1 - w_i) \mu_M.$$

While the variance for the portfolio is

$$\sigma_P^2 = w_i^2 \sigma_i^2 + (1 - w_i)^2 \sigma_M^2 + 2 w_i (1 - w_i) \sigma_{i,M}.$$

Where $\sigma_{i,M}$ is the covariance between two assets.

The slope of the portfolio is given

$$\frac{\delta \mu_p}{\delta \sigma_P} = \frac{\delta \mu_P}{\delta w_i} \frac{\delta w_i}{\delta \sigma_P} = \frac{(\mu_i - \mu_M) \sigma_P}{w_i \sigma_i^2 - \sigma_M^2 + w_i \sigma_M^2 + \sigma_{i,M} - 2 w_i \sigma_{i,M}}.$$

Which reduces to $\frac{(\mu_i - \mu_M)\sigma_P}{\sigma_{i,M} - \sigma_M^2}$ when $w_i = 0$.
Setting $w_i = 0$ makes our portfolio equal to the market portfolio and should then be equal to the slope of the CML (5.2).

$$\frac{(\mu_i - \mu_M)\sigma_P}{\sigma_{i,M} - \sigma_M^2} = \frac{\mu_M - \mu_f}{\sigma_M}$$

We rearrange and get the following expression for the SML

$$\mu_i - \mu_f = \frac{\sigma_{i,M}}{\sigma_M^2} (\mu_M - \mu_f) = \beta_i (\mu_M - \mu_f). \tag{5.3}$$

Under the CAPM all the individual securities plotted according to their expected return and the beta should fall along the SML. This means that by estimating the individual β's, for the stocks in a given portfolio, we can calculate the expected return of the portfolio under the given assumptions. We will return with an example of this and some discussion on the SML in our empirical analysis later on.

5.5 Conditional Correlation and Betas

In Silvapulle & Granger (2001) the authors investigates the possibility of portfolio diversification, i.e. the possible advantage of spreading the risk in several stocks rather than a few. Their approach is based upon conditional correlation, which is an approach of analyzing asymmetry. The conditional correlation between two random variables X and Y is

$$\rho_A = corr[X, Y|A] \tag{5.4}$$

where $A = \{[a \leq X \leq b \text{ and } c \leq Y \leq d]\}$.

In other words, the conditional correlation is the correlation between two random variables under a truncation. A truncated distribution is a distribution where the values are required to be in an interval which does not necessarily cover the whole support of the distribution. As with the Gaussian distribution where its values are defined on the whole $\mathbb{R}$, a truncated Gaussian is defined in a sub region of $\mathbb{R}$. A more thorough introduction to truncated distributions and derivation of the moments used in this section for the bivariate Gaussian case can be found in Appendix A.

Silvapulle & Granger (2001) use a dataset consisting of 30 Dow Jones industrial stocks in the period of 1991 to 1999. They divide the values of the stock in three separate categories; the lower quantile, the middle values and the upper quantile. They find signs of the average conditional correlation of the 30 stocks to be higher in the lower quantile than in the middle group, but they find no signs of difference between the upper quantile and the middle group.

Another important aspect of Silvapulle & Granger (2001) is the conditional values of the betas. The authors use the approach of conditional correlation, together with the beta in Equation 5.3 to create a conditional beta, which is given

$$\beta_A = \frac{Cov[R_i, R_M|A]}{Var[R_M|A]}. \tag{5.5}$$

The authors then use the conditional beta as a measure of the risk under different conditions for the market.

Unfortunately, there might be some fallacy in the conclusion done in Silvapulle & Granger (2001), due to the approach of conditional correlation. In Boyer *et al.* (1997) it is shown that the approach of splitting the values into different groups may actually give the impression of reduction in correlation, even when this is not the case. This is done by analyzing the bivariate Gaussian distribution, dependent on one of its marginal distribution. The bivariate Gaussian distribution has a constant correlation for the whole sample, however by conditioning the bivariate Gaussian on the region of one of its marginals they show that the value for the conditional correlation between the values are influenced by the truncation of the region.

We will therefore use some of the same approach as in Boyer *et al.* (1997), and replicate the calculations in Silvapulle & Granger (2001) by assuming a Gaussian return distribution, and see how this matches their results. Since the method used in Boyer *et al.* (1997) is restricted to truncation on one of the marginal distribution we will be using a truncated bivariate Gaussian distribution. We will only cover the case of a conditional beta in this section, an evaluation of the conditional correlation, variance and expectation is covered in Appendix A.

Figure 5.1 gives two plots of the movement for the conditional beta calculated on some bivariate Gaussian distributions. The upper plot has an upper truncation on both of the margins, which is moved further and further into the tails, i.e. the size of A is increased. Meanwhile the lower plot has a lower truncation on both its margins, which starts at 0.9 for both of its margins and is increasing, leading to a decrease in the size of A. Notice that the values for the beta in the lower plot would be similar for a upper truncation in the lower tail because of the symmetry of the Gaussian distribution. For both the upper and the lower plot the Gaussian distributions has zero valued expectations and the following matrixes as covariance matrixes in the given order respectively.

$$\Sigma_1 = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}, \Sigma_2 = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 3 \end{pmatrix}, \Sigma_3 = \begin{pmatrix} 2 & 0.5 \\ 0.5 & 1 \end{pmatrix} \text{ and } \Sigma_4 = \begin{pmatrix} 2 & 0.5 \\ 0.5 & 3 \end{pmatrix}.$$

Figure 5.1 does not correspond to what is seen in Silvapulle & Granger (2001), where they reported an average beta of 0.836 for the lower 0.05 quantile, 0.752 in the middle and 0.512 for the upper 0.95 quantile of the values. By looking at the upper plots in Figure 5.1 when $\Pr(A)$ is approaching 1, we see that the value for beta is higher than in the left part of the lower plots. This does not match the results in Silvapulle & Granger (2001), where they found higher values for the beta in the lower quantile than in the middle. Also the authors in Silvapulle & Granger (2001) points out that the standard error of the betas is large, 0.416 for the lower 0.05, 0.148 in the middle and 0.252 for the upper 0.95 quantile of the values, which means that they should be evaluated with caution.

That our values for the beta in Figure 5.1 does not seem to change equal to the ones in Silvapulle & Granger (2001), is not really that surprising. As we have seen and will see more of later, the Gaussian distribution is not necessarily a good fit for financial returns. But as seen in Figure 5.1, the betas calculated with conditional correlation is not constant, as they should be because of the constant correlation and variance in the Gaussian distribution. This means that the findings in Silvapulle & Granger (2001) might also be influenced by this bias. Because of this, we will investigate an additional measure of correlation in the following chapter.

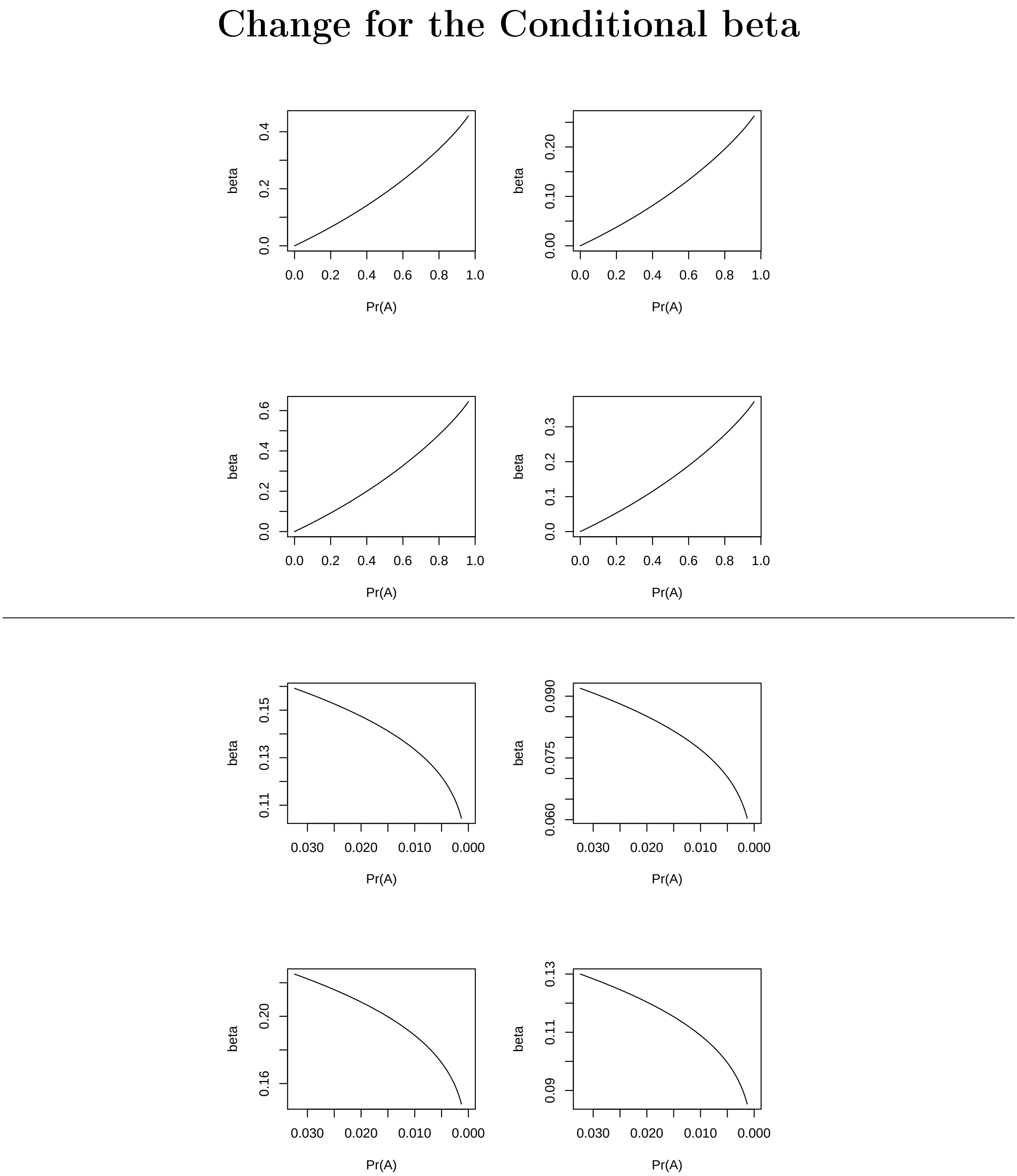


Figure 5.1: Change for the conditional beta with two different truncations on 4 different bivariate Gaussian distribution

Chapter 6

Local Gaussian Correlation

Local Gaussian Correlation is really the main subject of this thesis. It is a new dependence measure with an approach that differs drastically from the other ways of estimating correlation that we have explored earlier. It does produce a value on the interval [-1,1] reflecting the dependence, but instead of letting this value represent the correlation as a constant it measure the correlation locally. This enables us to describe how changes in one variable affect the change in the other variable.

As mentioned under the section of correlation, the Gaussian distribution has a close connection with Pearson's correlation coefficient. This means that linear dependence is a natural part of a Gaussian distribution in multiple dimensions. As seen when analyzing data for the Statoil stock assuming a distribution for the returns often simplify calculations and allow for more sophisticated analyses. One example would be to assume the distribution for the returns is given by a Gaussian distribution. By estimating its variance and expectation by the observed sample we were able to calculate the Value at Risk and Expected Shortfall. Unfortunately the assumption that the return distribution is given by a Gaussian distribution is poor at best, and as we have seen, risk estimation based on this approach might underestimate the risk.

Assuming that the observed sample is from a Gaussian distribution should be avoided, or at least be made with caution. Nevertheless, there are ways for us to get to use some of the properties of the Gaussian distribution even without having to assume too much about the actual return distribution. And this is where Local Gaussian Correlation comes into play.

Consider a point (x, y) given on the plane $\mathbb{R}^2$. We are interested in a bivariate Gaussian distribution with density function denoted ψ so that it approximates the given density function f in a neighborhood A, around (x, y). Or in other words the probability for an observation in A should be approximately the same for our Gaussian approximation and the observed distribution.

By doing this over the region where f is defined, or at least where we have observed values from the distribution of f, we end up with a family of Gaussian distributions, each approximating f in its own neighborhood. We denote the Gaussian approximation in the neighborhood of (x, y) as $\psi_{x,y}$. The Gaussian approximation will in its neighborhood be a good approximation for f, and we can use the correlation of $\psi_{x,y}$ as a measure of the local dependence in this neighborhood.

6.1 Derivation of the Method

Local Gaussian Correlation as described above may give a slight idea of how it is calculated. However, we will also walk trough the derivation in a more technical sense. This will hopefully give an understanding of how the Local Gaussian Correlation actually is estimated and its properties. For a deeper explanation of the theory, Tjøstheim & Hufthammer (2012) which is the article that first introduced this correlation measures might be of interest. Another article of interest may be Hjort & Jones (1996) where related theory is used for density estimation.

As already mentioned in Equation 3.3 the probability density function for a bivariate Gaussian distribution is given by

$$\psi(x, y|\theta) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\}$$

where θ is $[\mu_x, \mu_y, \sigma_x, \sigma_y, \rho]$

We are interested in the bivariate Gaussian that approximate f in A. Because of this, we let the distribution $\psi_{x,y}$, which gives the probability for u and v in the neighborhood of (x, y) depend on x and y.

$$\begin{aligned}\psi_{x,y}(u, v|\theta(x, y)) = &\frac{1}{2\pi\sigma_1(x,y)\sigma_2(x,y)\sqrt{1-\rho(x,y)^2}} exp\left\{-\frac{1}{2(1-\rho(x,y)^2)}\right.\\ &\left.\times\left[\left(\frac{u-\mu_1(x,y)}{\sigma_1(x,y)}\right)^2 - 2\rho(x,y)\left(\frac{u-\mu_1(x,y)}{\sigma_1(x,y)}\right)\left(\frac{v-\mu_2(x,y)}{\sigma_2(x,y)}\right) + \left(\frac{v-\mu_2(x,y)}{\sigma_2(x,y)}\right)^2\right]\right\}\end{aligned} \tag{6.1}$$

Where $\theta(x, y) = [\mu_1(x, y), \mu_2(x, y), \sigma_1(x, y), \sigma_2(x, y), \rho(x, y)]$

To estimate the values for $\theta(x, y)$ that makes $\psi_{x,y}$ approximate f, we turn to a combination of local likelihood and kernel estimation.

6.1.1 Local Likelihood

The ordinary likelihood function for a bivariate sample $(X_{1,1}, X_{1,2}), (X_{2,1}, X_{2,2}), \ldots, (X_{N,1}, X_{N,2})$ is given as

$$L(\theta) = \Pi_{i=1}^{N} f(X_{i,1}, X_{i,2}|\theta)$$

The approach of maximum likelihood is finding the values of θ that maximize the given likelihood. This is an ordinary way of estimating parameters in statistics. The maximation is usually done with setting the derivative on the logarithm of the likelihood function with the respect to the wanted parameter equal to zero and solving for our parameter.

To estimate the parameters locally we will use something closely related to this, known as local likelihood. This version of local log likelihood was described in Hjort & Jones (1996), where it was used for semi-parametric density estimation.

Definition 6.1.1. *Local log likelihood*

$$\begin{aligned} l = log(L) = \frac{1}{N}\sum_{i=1}^{N} K_{b_1}(X_i - x)K_{b_2}(Y_i - y)log(\psi_{x,y}(X_i, Y_i|\theta(x, y))) \\ - \int K_{b_1}(u - x)K_{b_2}(v - y)\psi_{x,y}(u, v|\theta(x, y))dudv. \end{aligned}$$

where $K_{b_1}(u - x) = b_1^{-1}K(b_1^{-1}(u - x))$ *and similarly for* K_{b_2}

K_{b_1} and K_{b_2} are used in what is known as kernel estimation[1]. By letting $N \to \infty$, i.e. when the number of observations goes to infinity, the first part of the local likelihood will converge to the integral form almost surely[2].

$$\begin{aligned} \frac{1}{N}\sum_{i=1}^{N} K_{b_1}(X_i - x)K_{b_2}(Y_i - y)log(\psi_{x,y}(X_i, Y_i|\theta(x, y))) \\ \to \int K_{b_1}(u - x)K_{b_2}(v - y)log(\psi_{x,y}(u, v|\theta(x, y)))f(u, v)dudv \end{aligned} \tag{6.2}$$

This follows from the law of large numbers or the ergodic theorem.

[1]Kernel estimation is an estimation of a density function given an observed sample. An introduction to kernel estimation can be found in Rizzo (2007)

[2]Casella & Berger (2002) offers an introduction to different types of convergence in statistics

By letting the number of observation go to infinity the local likelihood defined above converge to the integral form almost surely, so that

$$l \to \int K_{b_1}(u-x)K_{b_2}(v-y)[log(\psi_{x,y}(u,v|\theta(x,y)))f(u,v) - \psi_{x,y}(u,v|\theta(x,y))]dudv$$

It then follows that the derivative of the local likelihood $\frac{\delta l}{\delta \theta_j}$ will almost surely converge to

$$\int K_{b_1}(u-x)K_{b_2}(v-y)\frac{\delta \log(\psi_{x,y}(u,v|\theta(x,y)))}{\delta \theta_j}[f(u,v) - \psi_{x,y}(u,v|\theta(x,y))]dudv$$

Following the normal approach of maximizing a likelihood, we restrict $\frac{\delta l}{\delta \theta_j}$ to be zero, and by letting b_1 & $b_2 \to 0$ we have that

$$\frac{\delta \log(\psi_{x,y}(x,y|\theta(x,y)))}{\delta \theta_j}[f(x,y) - \psi_{x,y}(x,y|\theta(x,y))] + O(b_1^2 + b_2^2) = 0$$

Summarized we have that when the number of observations goes to infinity and the bandwidth is small, $\psi_{x,y}$ equals the density f in that neighborhood.
There are some restrictions to bandwidth in the kernel estimation that also may apply. We will limit our discussion about bandwidth to an example of Local Gaussian Correlation estimated under different bandwidths. For an more thorough discussion on the importance of bandwidth see Tjøstheim & Hufthammer (2012).

In order to estimate $\theta(x,y)$, we take the derivative on the local likelihood and solve it numerically for

$$[\mu_1(x,y), \mu_2(x,y), \sigma_1(x,y), \sigma_2(x,y), \rho(x,y)]$$

where the derivative of the local likelihood with respect to θ_j is.

$$\frac{1}{N}\sum_{i=1}^{N} K_{b_1}(X_i-x)K_{b_2}(Y_i-y)\frac{\delta log(\psi_{x,y}(X_i,Y_i|\theta(x,y)))}{\delta \theta_j}$$
$$-\int K_{b_1}(u-x)K_{b_2}(v-y)\frac{\delta log(\psi_{x,y,\theta}(u,v|\theta(x,y)))}{\delta \theta_j}\psi_{x,y}(u,v|\theta(x,y))dudv \tag{6.3}$$

6.2 Distribution of the Parameters

In Tjøstheim & Hufthammer (2012) they show that the estimator $\hat{\theta}_b(x,y)$ for $\theta(x,y)$ is asymptotically given by a Gaussian distribution.

$$(Nb_1b_2)^{1/2}[\hat{\theta}_{N,b}(x,y) - \theta_b(x,y)] \xrightarrow{d} N(0, J_b^{-1}M_b(J_b^{-1})^T) \tag{6.4}$$

$$\begin{aligned}
\text{where } J_b = &\int K_{b_1}(u-x)K_{b_2}(v-y)w(u,v,\theta_b(x,y)) \\
&\times w^T(u,v,\theta_b(x,y))\psi(u,v|\theta_b(x,y))dudv \\
&- \int K_{b_1}(u-x)K_{b_2}(v-y)\nabla w(u,v,\theta_b(x,y)) \\
&\times [f(u,v) - \psi(u,v|\theta_b(x,y))]dudv
\end{aligned}$$

$$\begin{aligned}
M_b = b_1b_2 &\int K_{b_1}^2(u-x)K_{b_2}^2(v-y)w(u,v,\theta_b(x,y)) \\
&\times w^T(u,v,\theta_b(x,y))f(u,v)dudv \\
-b_1b_2 &\int K_{b_1}^2(u-x)K_{b_2}^2(v-y)w(u,v,\theta_b(x,y))f(u,v)dudv \\
\times &\int K_{b_1}^2(u-x)K_{b_2}^2(v-y)w^T(u,v,\theta_b(x,y))f(u,v)dudv \\
\text{and } w(u,v,\theta_b(x,y)) &= \frac{\delta log(\psi_{x,u}(u,v|\theta_b(x,y))}{\delta\theta_j}.
\end{aligned}$$

To actually calculate the variance of the estimator $\hat{\theta}$ is unfortunately a little bit harder than it may seem. The return distribution f is still unknown, so we cannot compute J_b or M_b in a simple manner. One way of computing the integral is by using numerical integration based on the observed sample from the distribution f to compute the unknown part of the integral. Another possible method is by using bootstrap[3], to generate the variance and expection for $\hat{\theta}$.

Both these methods are quite dependent of the original observed sample to be a good representation for the distribution f. However, the method based on bootstrap also requires the assumption that the sample is iid, which may not be true. One way to avoid this problem might be to use a GARCH-filtering as described under the introduction of GARCH.

[3] Bootstrap as a method for nonparametric distribution estimation was introduced in Efron (1979). If we have an observed sample from a distribution f we can by resampling obtain a random sample from the distribution $f_n(x)$, which is an estimator of f. For an introduction of bootstrap, see Rizzo (2007)

The method based on bootstrap is the method that has been used in the calculations in this thesis. This has been done by the code written by Karl Ove Hufthammer 'kode-loclkb-gr'.

Throughout the rest of this thesis the calculation of the Local Gaussian Correlation is done with the bandwidths set equal to the empirical standard deviation of the observations if nothing else is stated. Similarly, when the Local Gaussian Correlation is calculated on non-simulated data the data has been filtered with a GARCH-filter under the assumption of a t distribution. The plots along the $x = y$-axis is calculated on 50 points uniformly spaced with a confidence interval of 90% and the bootstrap is calculated by 500 replicates if nothing else is stated.

6.3 Examples and Uses of Local Correlation

We will later in this thesis use Local Gaussian Correlation on some financial data sets to investigate some of the properties Pearson's correlation and other global dependence measures may have problem detecting. Like for example how the correlation between two stocks may be affected by the value of the stocks. The following section will be dedicated to calculation of the local correlation on some simulated data to see the influence of bandwidth, and some visualization of the dependence for different copulas.

6.3.1 Constant Correlation

As mentioned the relationship between the margins in a multivariate Gaussian distribution is constant. This means that for a multivariate Gaussian distribution the theoretical local correlation will be constant and equal to the theoretical Pearson's correlation coefficient. We will in this example estimate the Local Gaussian Correlation on a random sample from a bivariate Gaussian with three different bandwidths; 0.5, 1 and 2. The sample consist of 1000 observations from a Gaussian with the following parameters

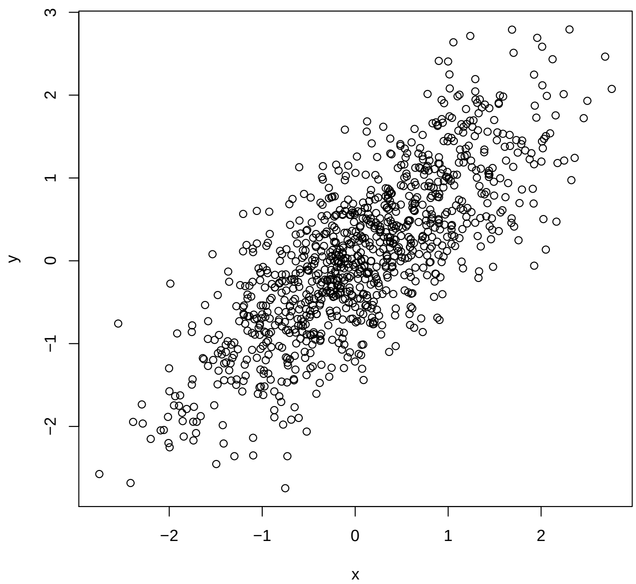
$$\Sigma = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix} \text{ and } \mu = \begin{pmatrix} 0 & 0 \end{pmatrix}.$$

Figure 6.1 contains a scatterplot of the sample and the Local Gaussian Correlation with the bandwidths equal 0.5, Figure 6.2 is the Local Gaussian Correlation with the bandwidths equal to 1 and 2.

What we can see in these plots, is that calculations with a small bandwidth is heavier influenced by each observation, making it possible to observe properties, we might not see with larger bandwidths. However, this also makes it more exposed to random fluctuations. On the other hand, in the case of too large bandwidth we may not observe important local properties in the sample. When the bandwidth is tending to infinity, the correlation of each area will be equal to Pearson's correlation coefficient for the whole sample.

	Pearson	Kendall	Spearman
Gaussian sample	0.77	0.56	0.76

Local Gaussian Correlation on a random Gaussian sample



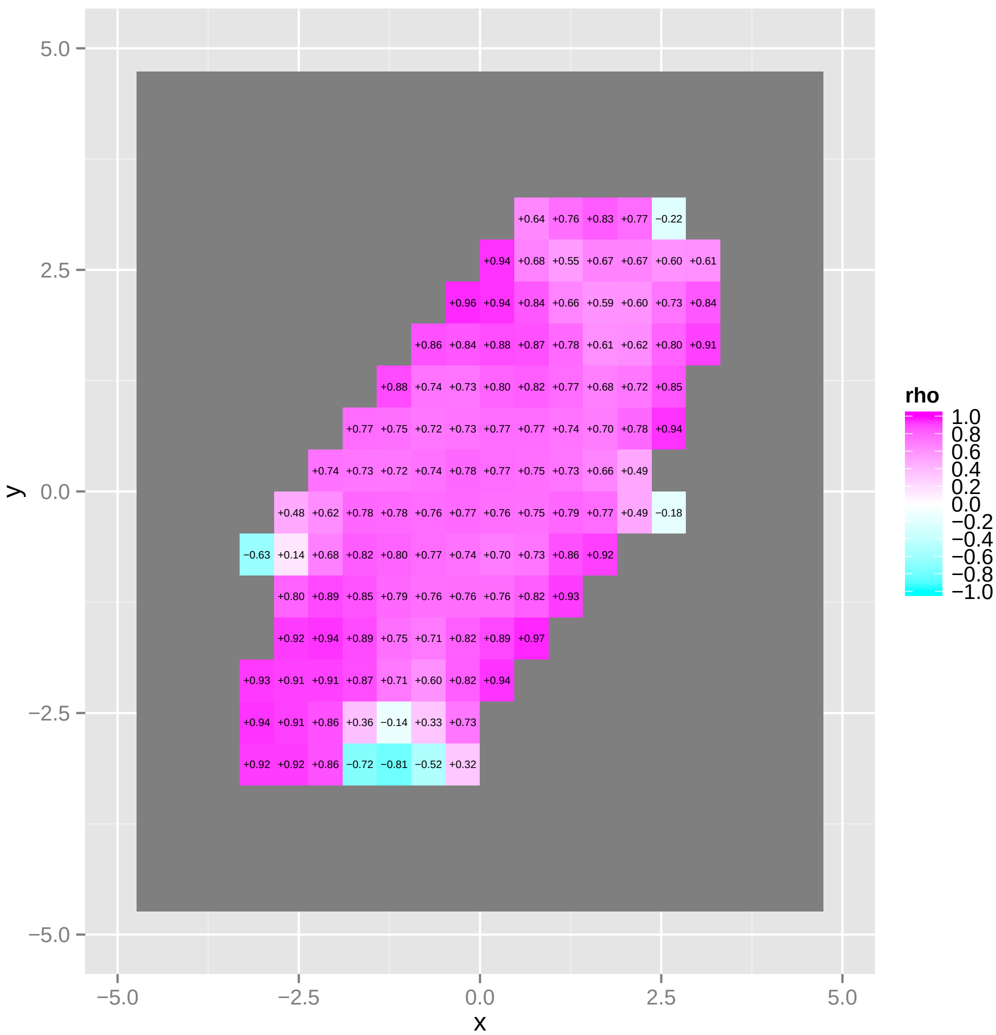


Figure 6.1: Scatterplot of the sample and calculation with bandwidth=0.5

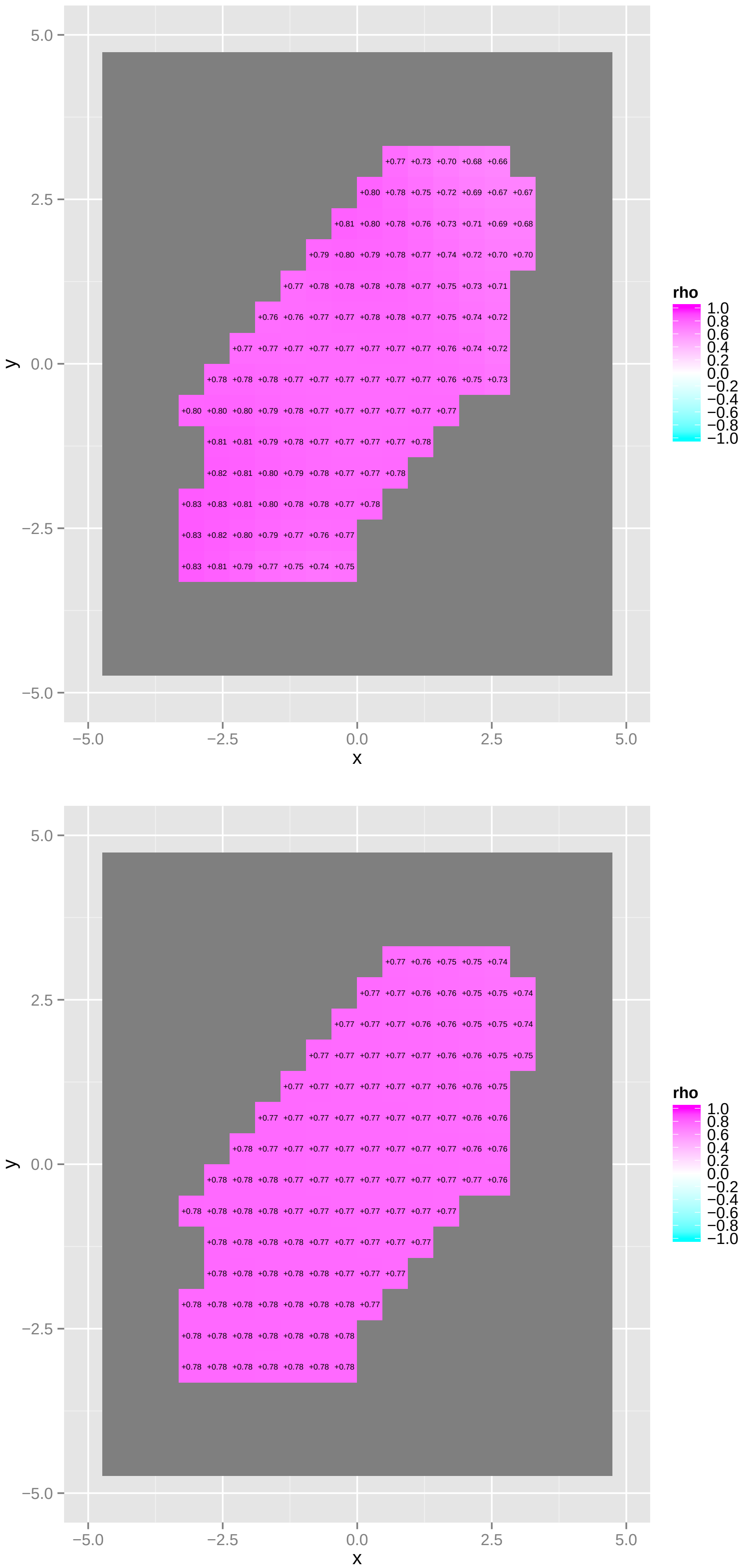


Figure 6.2: Upper plot has bandwidth=1 and lower plot has bandwidth=2

6.3.2 Local Gaussian Correlation Visualization of Copulas

As mentioned, we will later be using Local Gaussian Correlation to describe actual financial data. And in this setting, it is natural to compare the observed dependence structures with known copulas. We will therefore be estimating the Local Gaussian Correlation along the diagonal of $x = y$ on random samples of the already introduced copulas: Gaussian, t, Gumbel, Clayton and Frank. This approach has been followed in Berentsen *et al.* (2012), where tails and other important features of the copulas are calculated analytical.

Notice that in the following visualization of the copulas we have restricted ourselves to the interval of -2 to 2, even though the copula has important features outside this limit. The reason for this is that these plots are made by sampling with a Gaussian as the marginal distributions, which gives that approximately 95% of the observations will be between this limit. Estimating too far out might give influenced by rare observations, which might not necessarily be a good description of the copula.

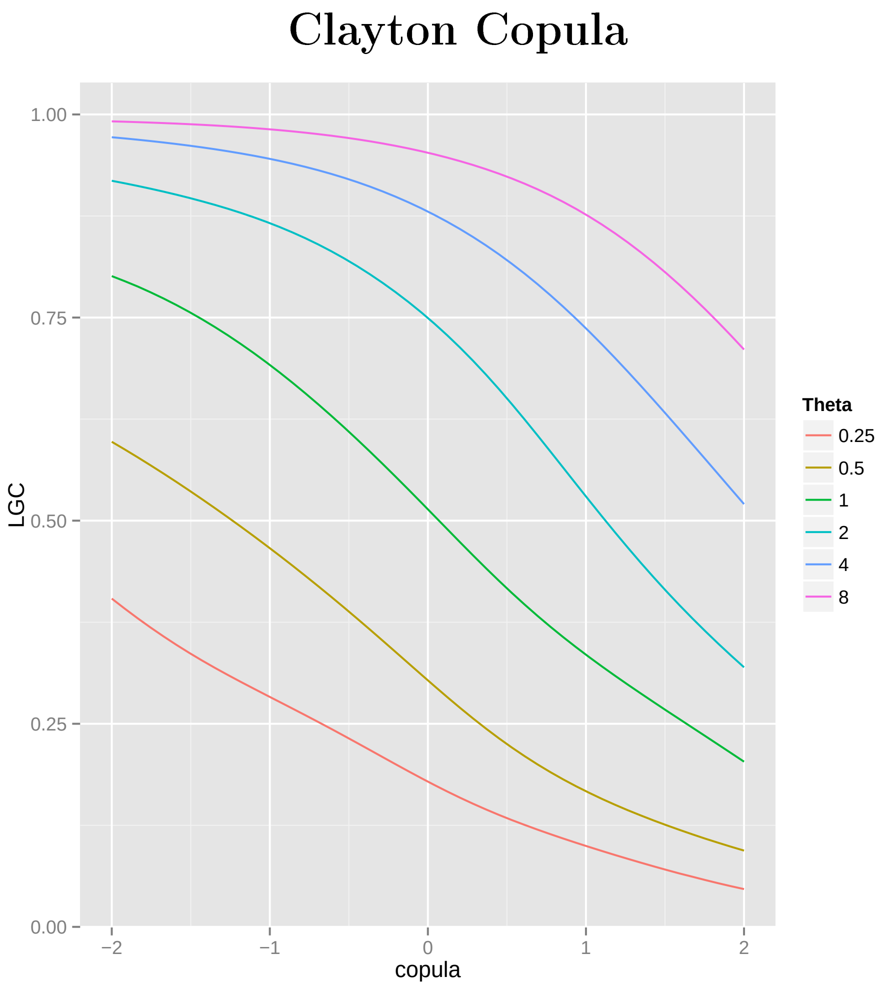
Also, note that there is slight deviation of the estimates of the plots from the theoretical value of the Local Gaussian Correlation due to the randomness of sampling, this is easily seen on the Gaussian copula where the theoretical value is constant equal to the parameter of ρ, but our plots shows some minor variation.

Each of the samples consists of 10000 observations, where the parameters in the distribution are given in the plot. And the bandwidth in the estimation is set equal to the standard deviation of the random variables, which means that the bandwidth is approximately equal to 1 due to its standard Gaussian marginal distributions.

The equation for the Clayton, Gumbel and Frank copula can be found in Table 4.3.1. As seen in Figure 6.3, Figure 6.4 and Figure 6.5 an increase in the value of θ for these copulas increase the correlation between the marginals.

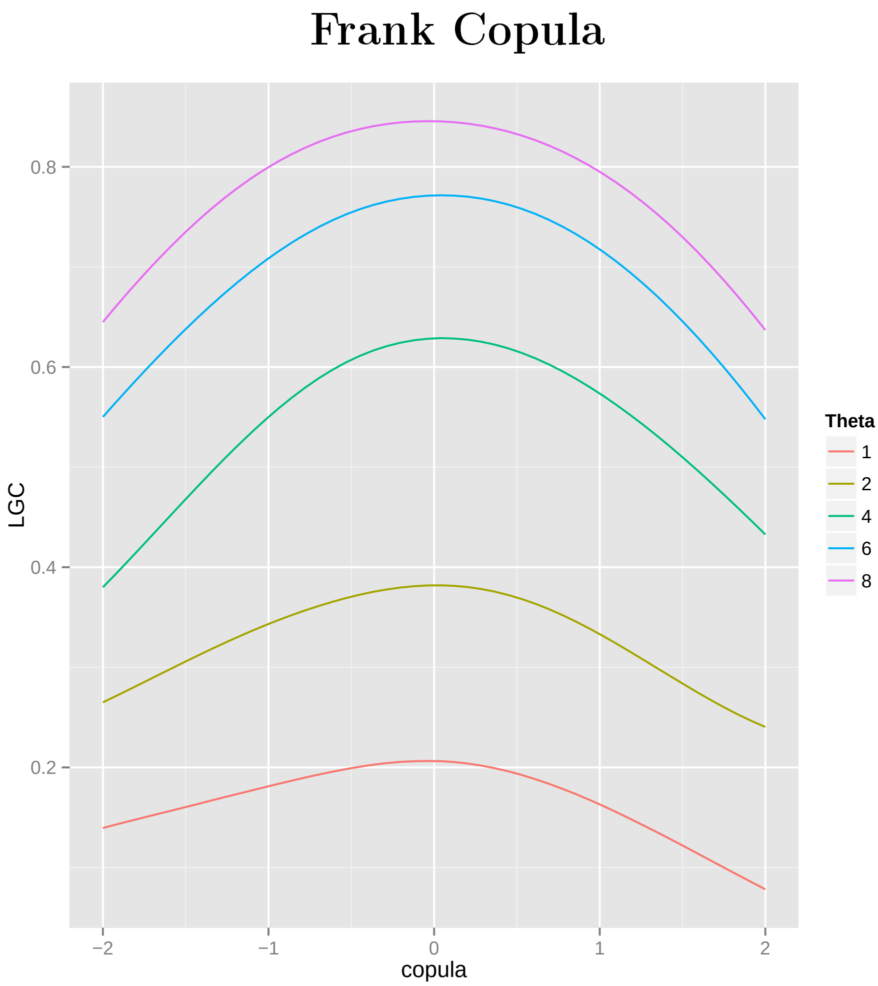
For the Gaussian copula defined in Equation 4.6, the ρ parameter correspond to the global correlation, and we see in Figure 6.6, that the estimated values for the Local Gaussian Correlation are approximately constant equal to the global correlation, in fact Berentsen *et al.* (2012) shows that the analytical value for the Local Gaussian Correlation is constant for a Gaussian copula with arbitrary marginals.

The t copula is defined in Equation 4.7 and the correlation is covered in Figure 6.7 and Figure 6.8. In Figure 6.7 we see the influence of the correlation parameter ρ, and we see that there is an increase in correlation as ρ increase. Note that there might be some deception in the meaning of the parameter ρ in Figure 6.7, as ρ does not correspond to the global correlation for a t copula where $\upsilon < \infty$. However, as seen in Figure 6.8, a t copula with parameter ρ approximates a Gaussian copula with global correlation ρ as the degrees of freedom increase.



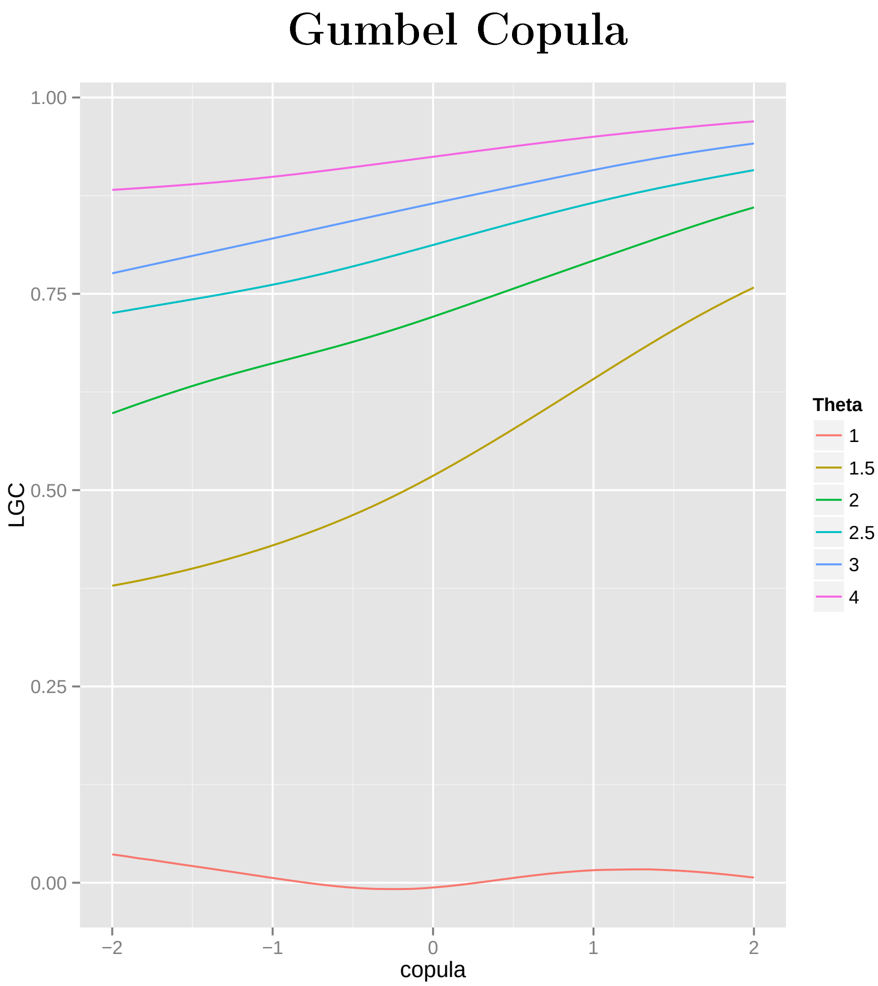
θ	Pearson	Kendall	Spearman
0.25	0.19	0.12	0.18
0.5	0.32	0.20	0.30
1	0.50	0.34	0.48
2	0.69	0.51	0.69
4	0.83	0.67	0.85
8	0.92	0.80	0.94

Figure 6.3: Local estimated Gaussian correlation, Kendall and Spearman on samples from Clayton copula with parameter θ and Gaussian marginal distributions



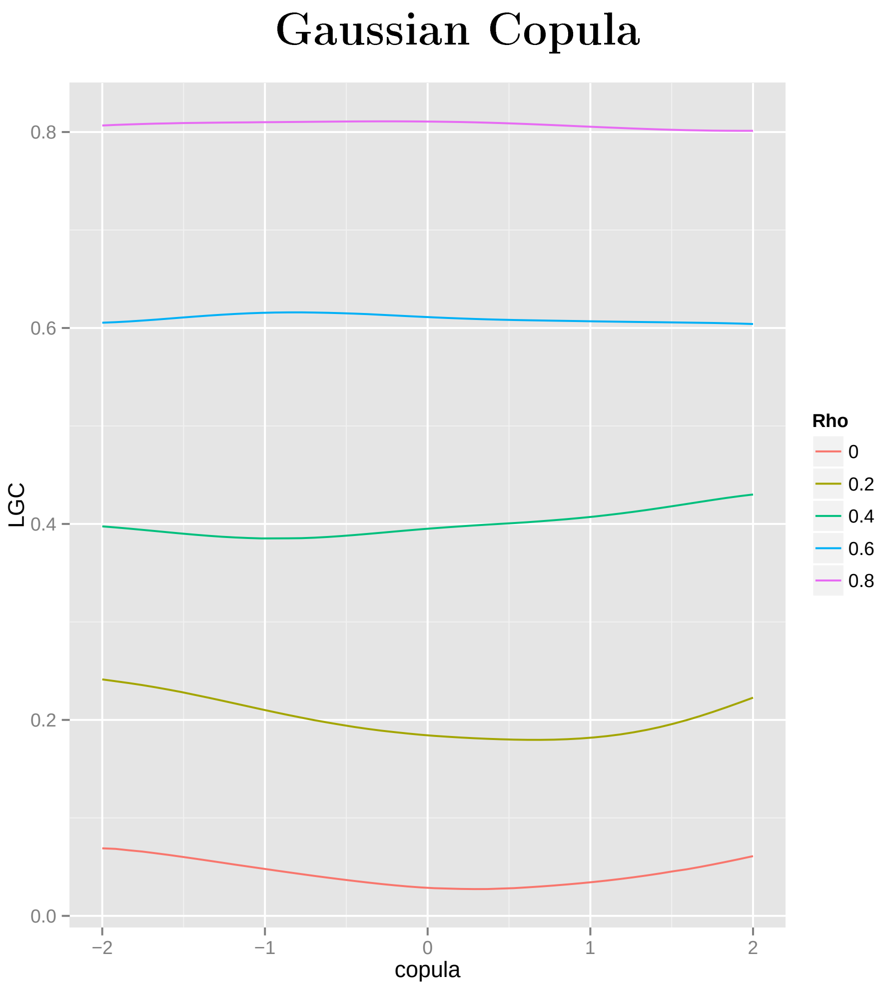
θ	Pearson	Kendall	Spearman
1	0.16	0.11	0.17
2	0.31	0.22	0.32
4	0.52	0.38	0.55
6	0.67	0.52	0.71
8	0.76	0.60	0.80

Figure 6.4: Local estimated Gaussian correlation, Kendall and Spearman on samples from Frank copula with parameter θ and Gaussian marginal distributions



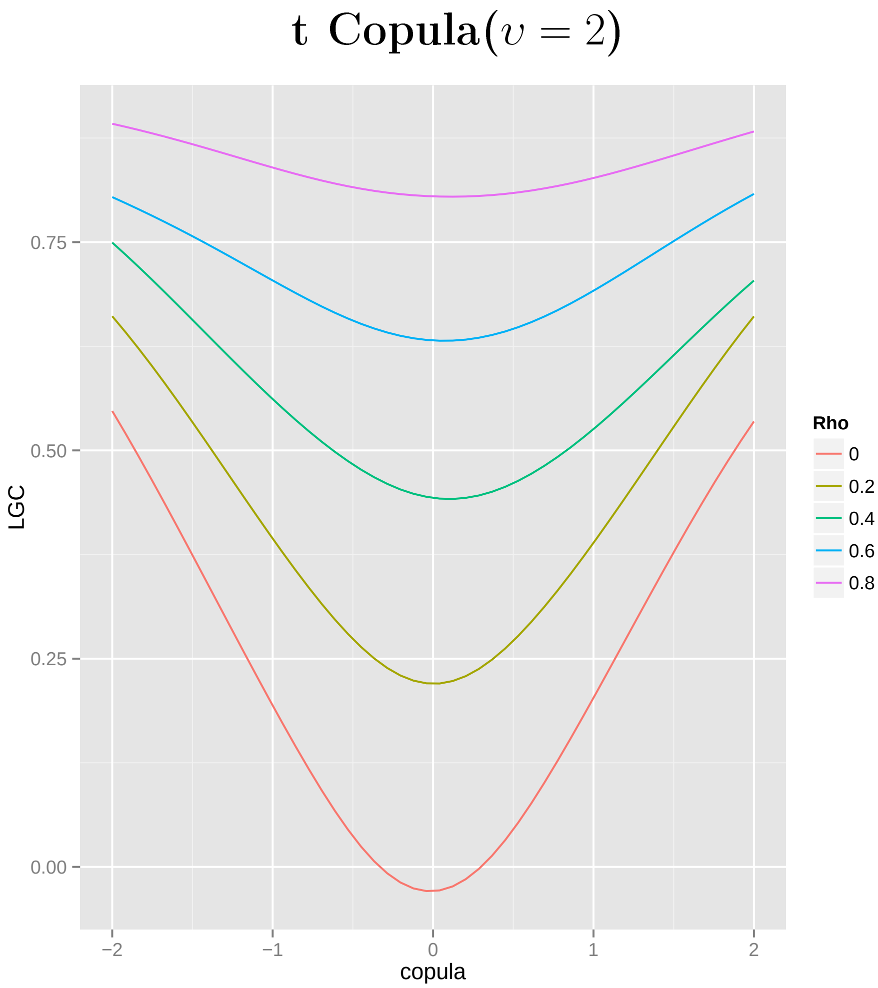
θ	Pearson	Kendall	Spearman
1	0.01	0.00	0.00
1.5	0.51	0.34	0.49
2	0.71	0.51	0.69
2.5	0.80	0.60	0.79
3	0.86	0.67	0.85
4	0.92	0.75	0.91

Figure 6.5: Local estimated Gaussian correlation, Kendall and Spearman on samples from Gumbel copula with parameter θ and Gaussian marginal distributions



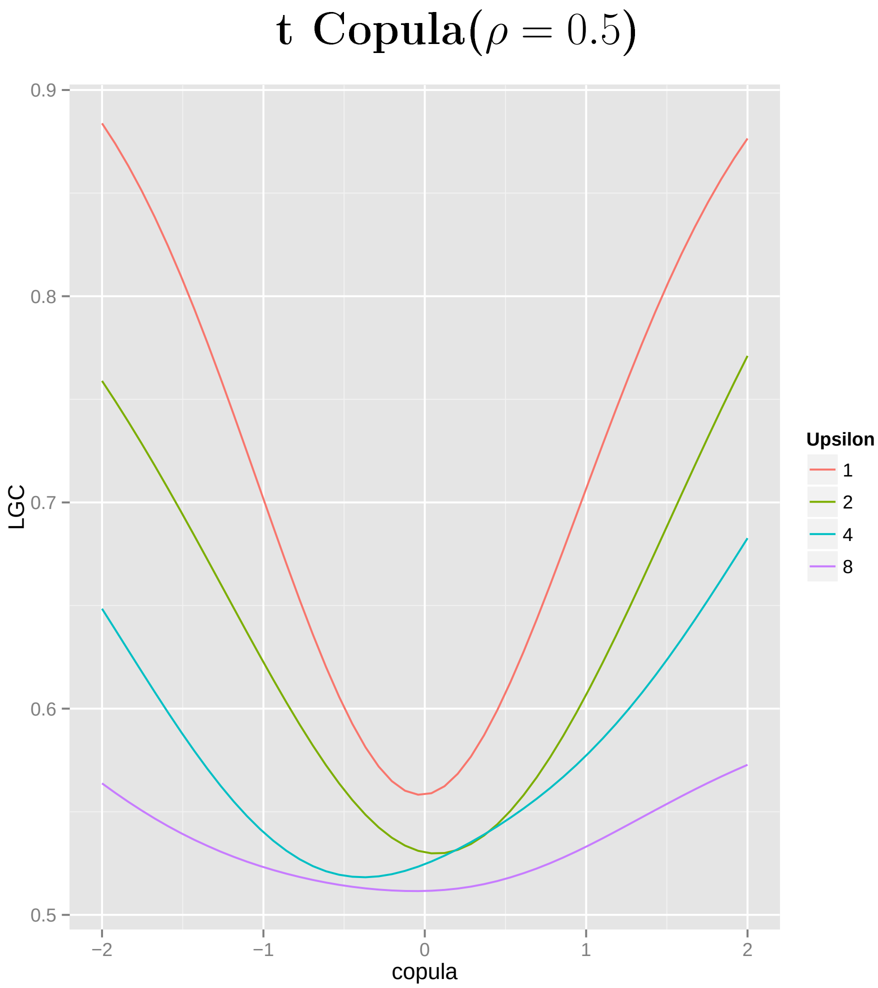
ρ	Pearson	Kendall	Spearman
0	0.02	0.01	0.02
0.2	0.19	0.12	0.18
0.4	0.39	0.26	0.38
0.6	0.60	0.42	0.59
0.8	0.81	0.60	0.80
1	1.00	1.00	1.00

Figure 6.6: Local estimated Gaussian correlation, Kendall and Spearman on samples from Gaussian copula with global correlation ρ and Gaussian marginal distributions



ρ	Pearson	Kendall	Spearman
0	-0.04	-0.02	-0.03
0.2	0.19	0.13	0.18
0.4	0.37	0.26	0.35
0.6	0.58	0.41	0.56
0.8	0.78	0.59	0.75
1	1.00	1.00	1.00

Figure 6.7: Local estimated Gaussian correlation, Kendall and Spearman on samples from t copula with correlation parameter ρ and Gaussian marginal distributions



υ	Pearson	Kendall	Spearman
1	0.46	0.35	0.45
2	0.48	0.34	0.47
4	0.50	0.34	0.48
8	0.49	0.33	0.47

Figure 6.8: Local estimated Gaussian correlation, Kendall and Spearman on samples from t copula with υ degrees of freedom and Gaussian marginal distributions

Chapter 7

Analysis of Dependence and Risk in the Norwegian Stock Market

Now that we have covered the theory for this thesis, we can start with the analysis of the Norwegian Stock Market. We will start with an introduction of the market and some analysis that has previously been done, before we calculate the Local Gaussian Correlation between the main index of the Oslo Stock Exchange and some macroeconomic factors to get an idea of what drives the Norwegian market. Finally, we will use the previously introduced theory of CAPM, copulas, and risk measures together with the Local Gaussian Correlation to describe dependence and risk in the Norwegian Stock Market. We assume that the OSEBX index, which is the Oslo Stock Exchange's main index, is a good description for the overall stock market. It is an investable index consisting of weighted values of chosen stocks available on the Oslo Stock Exchange, where the stocks are selected semiannually.

7.1 Introduction to The Norwegian Stock Market

Oslo Stock Exchange is a small and volatile stock market. It is often assumed to be heavily influenced by the oil prices, since many of the larger companies are directly associated with the oil industry. This hypothesis has previously been tested in articles as Bjørnland (2009) and Gjerde & Saettem (1999). These are studies on macroeconomic factors like oil price, interest rate, unemployment and their relationship with the OSEBX stock exchange index.

The analysis in Bjørnland (2009) uses monthly data from 1993 to 2005 to analyses the effects of oil price shocks on stock returns in Norway. She uses a vector autoregression to describe the relation between oil prices and macroeconomic behavior. She finds that oil price shocks explain almost 20 percent of the variation for the OSEBX index in the time horizon of half a year. Additionally she finds that following a 10 percent increase in oil prices, an immediately increase of 2 to 3 percent for the stock returns is observed.
Moreover, Gjerde & Saettem (1999) conclude with a strong dependence between the Norwegian Stock Market and the oil price by using a similar model and monthly observations

from 1974 to 1994.

In our investigation of factors that influence the Norwegian Stock Market, we will use a different approach than the one in Gjerde & Saettem (1999) and Bjørnland (2009). We will calculate the Local Gaussian Correlation to describe the change in correlation between the OSEBX index, the price of Brent Oil and the value of the USD in NOK as the market changes.

7.1.1 Introduction of Data Concerning the Stock Market

Bjørnland (2009) results indicate a strong dependence between the Norwegian Stock Market and the monetary policy, but in order of keeping this introduction of the stock market fairly simple, we limit ourselves to the change in oil price and the change in the exchange rate between NOK and USD.

The data for the exchange rates are found on the Central Bank of Norway's homepage, while the values for the OSEBX index and the oil price is taken from the Norwegian Stock Exchange's homepage.

The oil price is the price for Brent crude and is given in USD. The articles mentioned above mainly focuses on monthly observations, while we use daily closing time data from October 11, 2007 until October 11, 2012. The datasets are synchronized to give 1214 numbers of logarithmic differences and a short summary of the datasets are given in the table below.

	Mean	Var	Median	Kurtosis	Skewness
Oslo Børs Benchmark Index	-0.01	4.39	0.12	3.81	-0.50
Brent Oil	0.03	6.01	0.02	2.17	-0.15
USD/NOK	0	0.95	-0.02	2.25	0.02

7.1.2 Estimating the Local Gaussian Correlation Between Different Factors and the OSEBX Index

It is clear from the results found in Gjerde & Saettem (1999) and Bjørnland (2009) that there is a strong relationship between the Norwegian Stock Market and oil price.
One might expect that the influence of the oil price shock estimated in Bjørnland (2009) might be better described by longer time horizons e.g. monthly, rather than for daily observations due to the possible time difference in the opening hours of the Norwegian Stock Exchange and other markets, and due to trends that is seen over several days. This might make some of the results found in previously mentioned articles hard to replicate by our approach, although some correlation is expected.

The following Figures has estimated local correlation in the plane for areas where the observed sample suggest the probability for an observation is larger than 0.0001. Note that in the scatterplot the observations has been GARCH-filtered and we have restricted the size of the axis to [-3,3].

By looking at the scatterplot in Figure 7.1 it is apparent that we should probably try to keep our analysis in the range of [-2, 2] for both of the variables, as there are a very limited numbers of values exceeding this limit. In the Local Gaussian Correlation plot we find a positive local correlation along both the $x = y$ and the $x = -y$-axis. And for both the axis there is clearly higher correlation when the logarithmic difference for the oil price is negative, than for positive values. For positive values for the oil price along the $x = y$-axis the correlation stays at approximately 0.25 from origo throughout the tails, where there seems to be a slight increase in correlation. While at the $x = -y$-axis the local correlation quickly drops and seems to end at slight above 0 in the tail. This shows that as expected there is a clear correlation between the Norwegian Stock Market and the oil price on daily basis, although the correlation is noticeable weaker when the stock exchange has negative and the oil price has positive values.

As mentioned the oil price in Figure 7.1 is given in USD. And since the exchange rates is not constant this might influence the correlation between the stock market and the oil price e.g. a fall in the stock market could possible lead to the price for USD in NOK to increase. Figure 7.2 gives a scatterplot and a plot over the local correlation between the OSEBX index and the price for a USD in NOK notated USD/NOK. The scatterplot clearly shows a negative correlation between the factors and that we should probably avoid making conclusions outside the range of [-2,2].
Along the $x = -y$-axis there is a high negative value for the correlation with a minimum approximately around origo, with a value of -0.4.

Since it is clear that there is a strong correlation between both the exchange rate and the price of oil(in USD). One might be interested in how much of the correlation between the oil price and the market that can be explained by the exchange rate.

Figure 7.3 is a scatterplot and the plot over the local correlation given in Figure 7.1. Here the oil price has been transferred to NOK to match the currency of the Norwegian Stock Market. The plot over the local correlation shows a clear decrease in the correlation along both of the axis. And an observation of negative values for the Norwegian Stock Market and positive values for the oil price now results in a negative correlation.

Figure 10.9, Figure 10.10 and Figure 10.11 in Appendix B gives the corresponding 90% confidence interval based on bootstrap replicates, for Figure 7.1 Figure 7.2 and Figure 7.3. The confidence interval seems to show the same trend as mentioned above. However, for positive value for Brent oil in Figure 10.11 and in Figure 10.9 the confidence interval has some cases of opposite sign. This is seen for positive values for Brent Oil and negative values for the OSEBX in Figure 10.9, and for positive values for the Brent Oil and both negative and positive values for the OSEBX in Figure 10.11. This might suggest that there might not be a clear trend for the market when the Brent Oil has positive returns and the OSEBX has negative returns.

7.1.3 Interpretation of the Correlation

Figure 7.1 show that there is clear dependence between the oil price(USD) and the Norwegian Stock Market for positive values for the stock market, while an increase in oil price seems to keep the correlation constant or decrease it.

This may be explained by a small and volatile stock market in Norway, where the oil is a crucial part of the economy, changes in oil price influence the stock market way more than the stock market of Norway influence the oil price.

A large part of the correlation between the stock market and the oil price disappears when taking the currency into account, since there is high negative correlation between the Norwegian Stock Market and the price of USD in NOK. A possible explanation for this might be that, negative values for the oil price will probably lead to negative values for the Norwegian Stock Market, which again leads to a decrease for the value of the NOK.

Local Gaussian Correlation: OSEBX - Brent Oil price(USD)

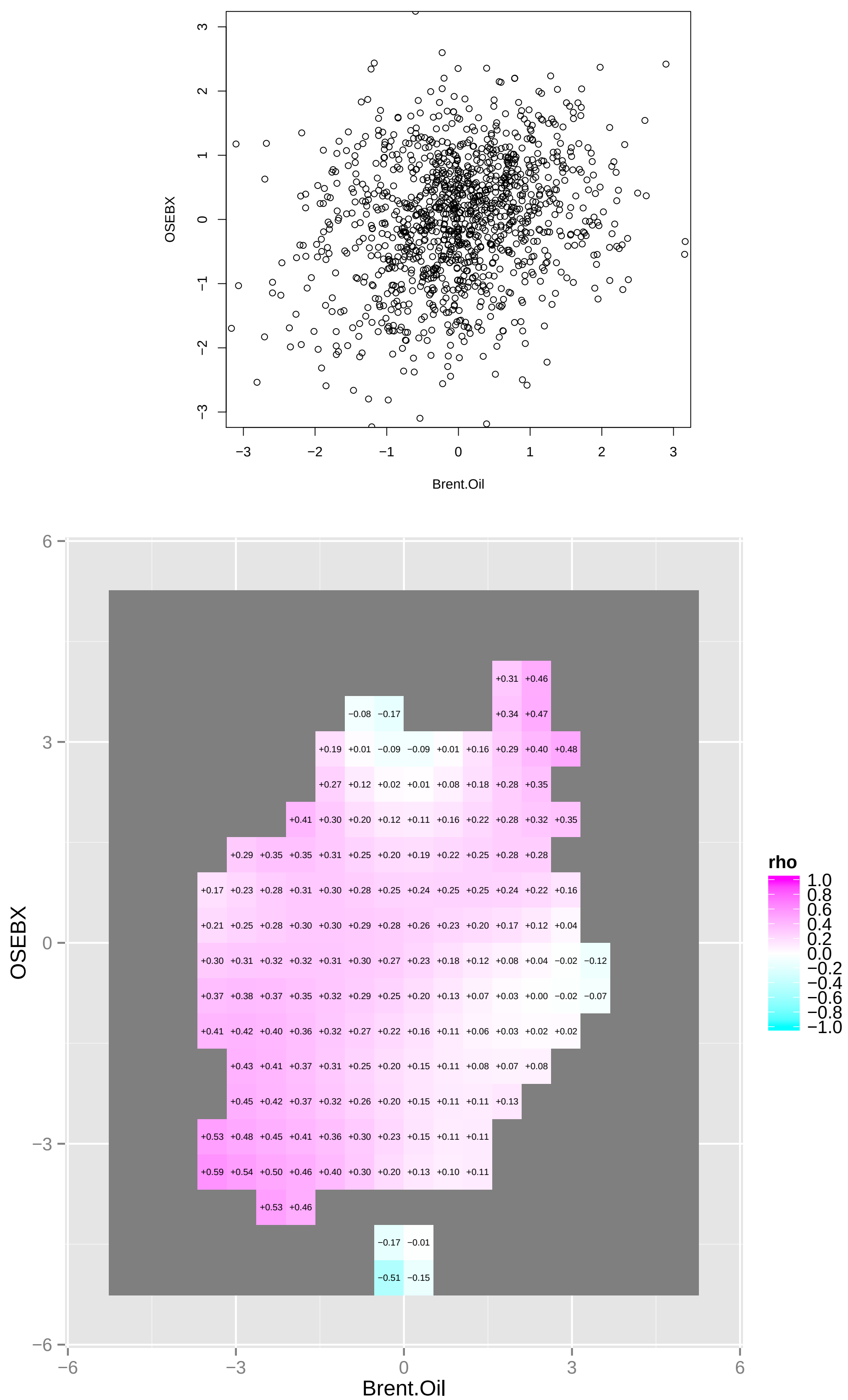


Figure 7.1: Local Gaussian Correlation between the OSEBX index and the price of Crude Brent(USD)

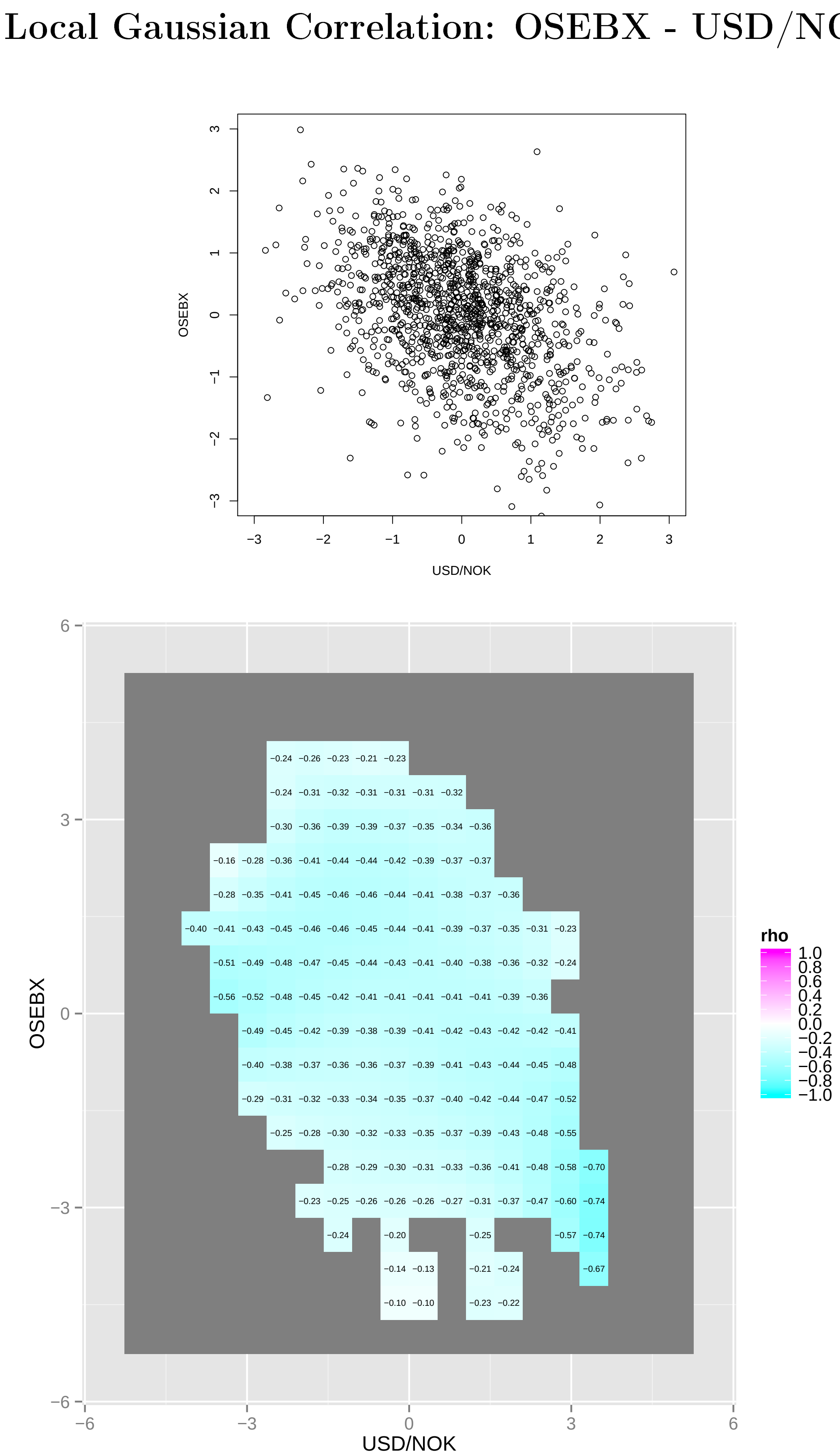


Figure 7.2: Local Gaussian Correlation between the OSEBX index and the USD/NOK

Local Gaussian Correlation: OSEBX - Brent Oil price(NOK)

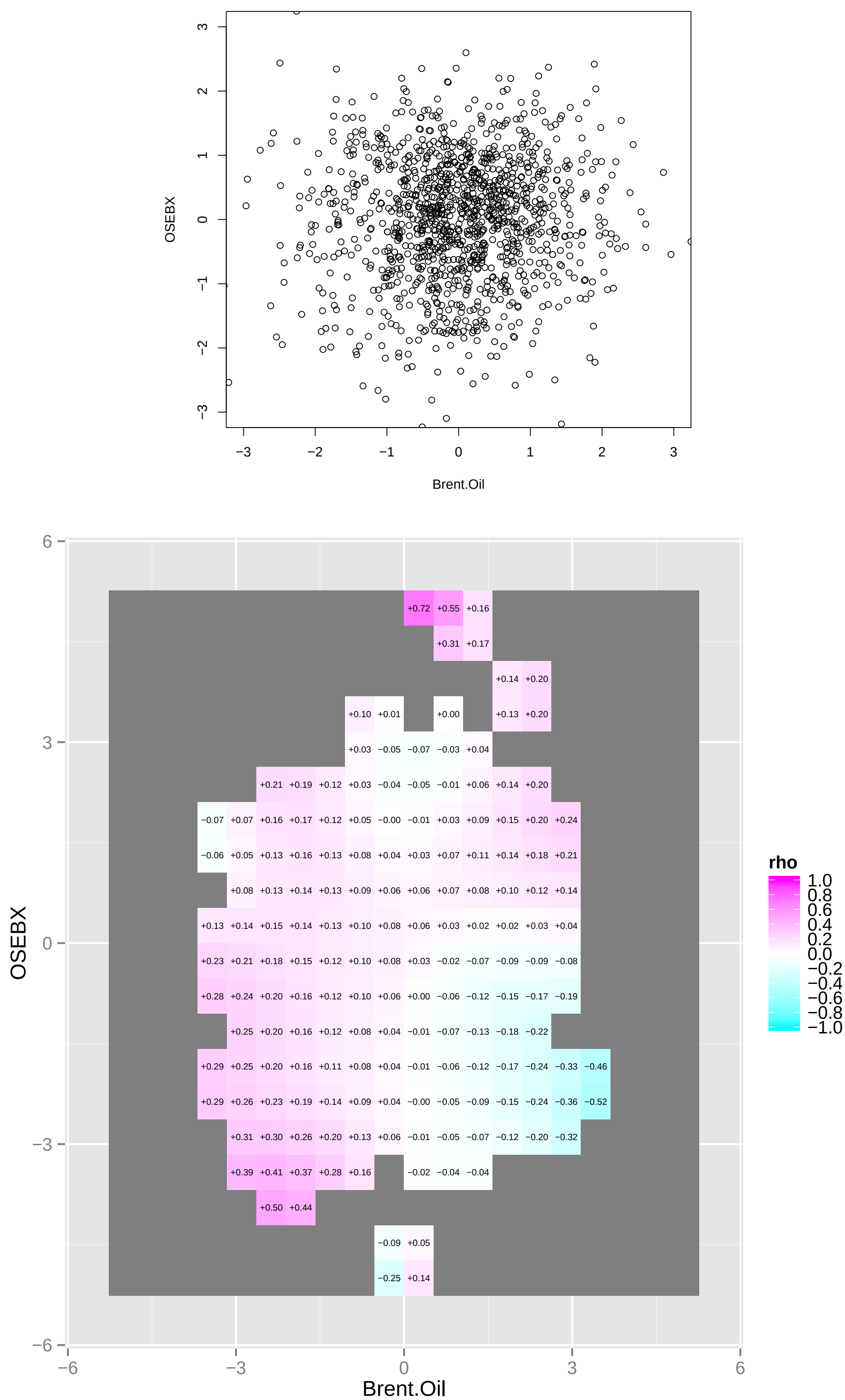


Figure 7.3: Local Gaussian Correlation between the OSEBX index and the price of Crude Brent

7.2 Oslo Stock Exchange Data

Now that we have established an image of the Norwegian Stock Market and seen how it is influenced by some factors, we will investigate the relationship between the stock exchange and its stock.

In the following calculation we will be using data for the OSEBX index and 18 different stocks on Oslo Stock Exchange. The data is collected from Oslo Stock Exchange's homepage. The data for each of the individual stocks ranges from October 11, 2007 until October 11, 2012, leaving a number of 1260 values of daily prices for the stock at closing time, which results in 1259 values of the logarithmic differences.
A short summary of the returns is given in the table below.

	Mean	Var	Median	Kurtosis	Skewness
Oslo Børs Benchmark Index	-0.01	4.25	0.12	3.98	-0.51
Aker Solutions	-0.03	15.87	0	6.58	-0.62
DNB	-0.01	10.92	-0.06	5.93	-0.07
Fred Olsen Energy	-0.01	6.53	0	2.94	-0.36
Frontline	-0.2	18.26	-0.18	27.71	-1.38
Norsk Hydro	-0.08	9.7	0	3.68	-0.22
Orkla	-0.06	5.92	0	6.49	0.01
Petroleum Geo Services	-0.04	14.94	0	3.92	-0.32
Prosafe	-0.01	8.12	0	6.11	-0.48
Renewable Energy Corporation	-0.38	25.1	-0.4	4.14	-0.23
Royal Caribbean Cruises	-0.02	13.61	0	4.25	0.21
Schibsted	0	9.64	0	3.48	-0.02
Seadrill	0.04	10.62	0.19	14.02	-0.77
Songa Offshore	-0.14	21.84	0	79.97	-4.35
Statoil	-0.02	5.04	0	4.04	-0.46
Storebrand	-0.08	15.98	-0.08	4.08	-0.17
Telenor	0	5.8	0.05	21.51	-1.43
TGS NOPEC Geophysical Company	0.06	13.27	0.06	4.79	-0.02
Yara International	0.04	11.06	0.07	3.64	-0.38

Table 7.1: Summary of data concerning the Oslo Stock Exchange

These stocks are chosen because they are all available on the Oslo stock exchange throughout the period mentioned above, the companies are of different size, and business, although 9 of the companies are directly related to the petroleum industry in Norway. All of the stocks are frequently traded, which means that there is not too many observations found where the value of the return is zero caused by lack of trade.

Figure 10.1, Figure 10.2 and Figure 10.3 in Appendix B shows the scatterplots of the stocks plotted against the OSEBX index. Note that the x-axis, representing the individual stocks have different scales in some cases. In the figures we see there are some extreme values for Songa Offshore and Frontline, which is probably some of the cause of their high kurtosis in Table 7.1. These are real observations and has not been filtered out, one might argue for and against filtering out extreme observations as it probably does not represent normal behavior for the stock, but from a risk perspective observations like these are important and should not be filtered.

7.3 Distribution of the Data

As mentioned, assuming that a specific distribution such as the NIG distribution, is a good description of our data may simplify calculations. Because of this, we will search for a distribution which describe the introduced data satisfactory.

In Bølviken & Benth (2000) the authors used the NIG distribution to evaluate the VaR for some stocks on the Norwegian Stock Market. They use eight Norwegian stocks which are found on the Oslo Stock Exchange and Norsk Hydro, which then was found on the New York Stock Exchange. The data for Norsk Hydro were from January 2, 1990 until December 31, 1998, and had a total of 2274 values, while the data from Oslo Stock Exchange consisted of 506 closing time prices from October 16, 1997 until November 17, 1999. They compared the Value at Risk for the Gaussian distribution and the NIG distribution with the result of a non-parametric method based on kernel estimation and conclude that the NIG distribution gives a better fit than the Gaussian.

As we have seen in our examples when introducing the theory of GARCH, models based on a time dependent standard deviation clearly gives a better description of the data than when the standard deviation is assumed constant.

We will therefore be using a test similar to the one used in Aas *et al.* (2005), where the authors uses VaR exceedance as a measure of distributional fit in an comparison of a multivariate NIG-GARCH, Gaussian-GARCH, symmetric t-GARCH and a skewed t-GARCH model. They find that for a one-day period the multivariate NIG-GARCH model is a better fit than the a multivariate model based on a Gaussian-GARCH, symmetric Student t-GARCH and a skew Student t-GARCH distribution.

Table 7.2 summarize a similar test as the one in Aas *et al.* (2005) and Bølviken & Benth (2000) to see if a Gaussian-GARCH, NIG-GARCH or a t-GARCH model is the best description for the return distribution on our stocks from the Oslo Stock Exchange.

We are interested in the numbers of exceedances of the VaR for each of the three models. The quantile is chosen to be the one of $VaR_{0.95}$ and $VaR_{0.99}$.
The GARCH parameters and additional parameters for each of the GARCH-model is cal-

culated on the whole sample. On this test we have removed the first 10 values of the time series as these might be influenced by the first value used for the volatility. 'Diff from Expected' is calculated as the sum of the absolute value of the difference between the expected values of exceedance($1249 \times (1-\alpha)$) and the observed by the different models for each stock.

Not surprisingly it seems from the values in 'Diff from Expected' that overall the NIG-GARCH is the best match for our data. And that both the t-GARCH and the NIG-GARCH model is a better choice than the Gaussian-GARCH model, especially for the $VaR_{0.99}$.

There is a censoring for the NIG distribution in the results, which should be mentioned as the garchFit function for fGarch was not able to calculate the parameters for some of the stocks[1] (these are denoted by (*)). These have not been taken into the calculation of 'Diff from Expected'.

Because of the mentioned error with the estimation of GARCH parameters for some of data we will therefore be using the t-GARCH as our model of choice for further analysis. Even if may be the next best choice as a description of our data, it does seem to explain the data satisfactorily and the GARCH coefficients are found and seem stable for all of the stocks.

[1] Aker and Frontline failed to converge by the nlminb routine but did by the lbfgsb algorithm when assuming a NIG distribution. Frontline converged with a warning message while Schibsted and Telenor failed to converge for all of the 8 combinations of algorithms for maximum likelihood estimation and evaluations of the Hessian matrix.

Exceedance of $VaR_{0.95}$

	Expected	t	Gaussian	NIG
DNB	62.5	60	58	60
Fred Olsen Energy	62.5	65	59	60
Norsk Hydro	62.5	66	64	62
Orkla	62.5	62	54	54
Oslo Børs Benchmark Index	62.5	84	87	74
Petroleum Geo Services	62.5	68	63	64
Prosafe	62.5	60	59	61
Renewable Energy Corporation	62.5	49	45	50
Royal Caribbean Cruises	62.5	61	60	62
Seadrill	62.5	71	69	62
Songa Offshore	62.5	60	55	52
Statoil	62.5	71	66	63
Storebrand	62.5	58	56	58
TGS NOPEC Geophysical Company	62.5	59	54	57
Yara International	62.5	71	71	65
Diff from Expectated	0	89.5	107.5	65.5
Aker Solutions	62.5	63	57	60*
Frontline	62.5	58	51	87*
Schibsted	62.5	67	63	?*
Telenor	62.5	66	54	?*

Exceedance of $VaR_{0.99}$

	Expected	t	Gaussian	NIG
DNB	12.5	13	22	13
Fred Olsen Energy	12.5	18	24	13
Norsk Hydro	12.5	9	14	9
Orkla	12.5	13	15	11
Oslo Børs Benchmark Index	12.5	13	19	11
Petroleum Geo Services	12.5	13	18	11
Prosafe	12.5	12	18	13
Renewable Energy Corporation	12.5	8	11	8
Royal Caribbean Cruises	12.5	11	15	11
Seadrill	12.5	16	26	13
Songa Offshore	12.5	13	19	11
Statoil	12.5	12	17	10
Storebrand	12.5	7	15	9
TGS NOPEC Geophysical Company	12.5	8	15	8
Yara International	12.5	15	18	14
Diff from Expected	0	34.5	81.5	29.5
Aker Solutions	12.5	13	22	13*
Frontline	12.5	10	13	24*
Schibsted	12.5	11	15	?*
Telenor	12.5	10	15	?*

Table 7.2: Numbers of exceedance of $VaR_{0.95}$ $VaR_{0.99}$ for three GARCH models.
* Are observations where there were problems with the coefficient estimation.

7.4 Local Gaussian Correlation on the Norwegian Stock Market

Patton (2004) investigate the difference in dependence structure between indices for monthly data, during so-called bear and bull markets, where bear market is an overall negative trend in the market, while bull is more optimistic with a positive trend. They compare the use of a bivariate Gaussian distribution with time dependent parameters with models based on different copulas(including the Gaussian copula) with Skewed student's t distribution with time dependent parameters as marginal distributions. They find evidence that models able to describe skewness and asymmetric dependence were a better choice when making portfolio decisions than models based on the bivariate Gaussian marginal distributions and/or copula.

Similarly, we have seen by the use of VaR exceedance that the t-GARCH model clearly beats the Gaussian-GARCH model as a description of the marginal distributions for the introduced stocks. We will continue in a similar manner as Patton (2004) and try to describe the dependence structure between the introduced stocks and the index. The relationship between the stock and the index is important because it gives us an image of how the stock is affected by the movement of the rest of the market. We will also calculate the local correlation between some stocks, to see how they are affected by each other.

7.4.1 Local Gaussian Correlation Stock to Index

Figure 10.1, Figure 10.2 and Figure 10.3 in Appendix B shows the scatterplots of the stocks against the OSEBX index. Since the limit of the axes are not standardized and these plots does not show the GARCH-filtered returns, it might not be easy to tell, however, as we did earlier, we should keep our investigation inside the limit of -2 to 2 where we have enough observations.

Figure 7.5, Figure 7.6 and Figure 7.7 contain plots of the Local Gaussian Correlation between each of the introduced stocks and the OSEBX index along the $x = y$-axis. Overall, there are high correlation values between the stocks and the index, and for every one of the stocks the correlation is between the interval of 0.9 and 0.35 for the whole axis. Figure 7.4 shows the mean of the local correlation for the 18 stocks where the confidence interval is calculated by bootstrap with 500 replicates. This confirms the strong correlation seen between the individual stocks and the index. It is also apparent how the negative tail for the stocks and the index has noticeable stronger correlation than the positive tail, as there is a difference on approximately 0.1

Several of the stocks even has as correlation as high as 0.8 when both the returns has an value of about -2 and there seems to be a trend of lower correlation in the positive tail than in the negative.

The highest value for the Local Gaussian Correlation is found in Statoil, here the local correlation is between 0.9 and 0.8 along the whole axis. Similarly Petroleum Geo Services, TGS NOPEC and Aker Solutions are not far behind. Finding a strong correlation between these oil related companies is not surprising, because as we have seen there is quite a strong dependence between the oil price and the Norwegian Stock Market.

Strong correlation is not limited to oil related industries, Norsk Hydro and Orkla is examples of companies that is not directly related to the oil industry but has a high positive correlation with the OSEBX index.

Another company worth mentioning is Renewable Energy Corporation, which belongs to an industry quite different from the oil industry. Renewable Energy Corporation has a strong correlation when both the company and the stock market has large losses, but its correlation is weakened dramatically for positive values for the returns.
This might also be partly explained by the Norwegian Stock Markets relation with the oil industry. In bad times for the stock market, one would expect people to be nervous and more restrictive with their investments, which is bad news for the individual stocks and the stock market as a whole. Good times for the oil industry and the industry of renewable energy might not necessarily be related, as these are, at least to some extent counterparts. This might explains why the correlation is lower for positive values.

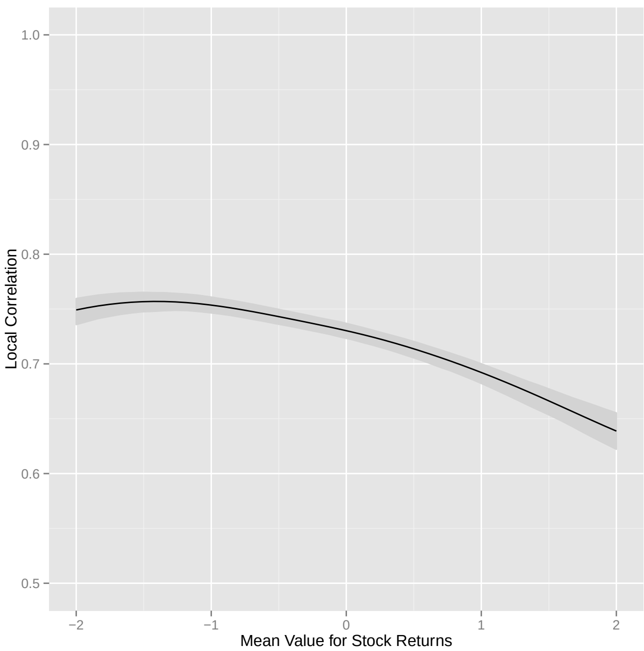


Figure 7.4: Mean estimated Local Gaussian Correlation between stocks and the OSEBX index

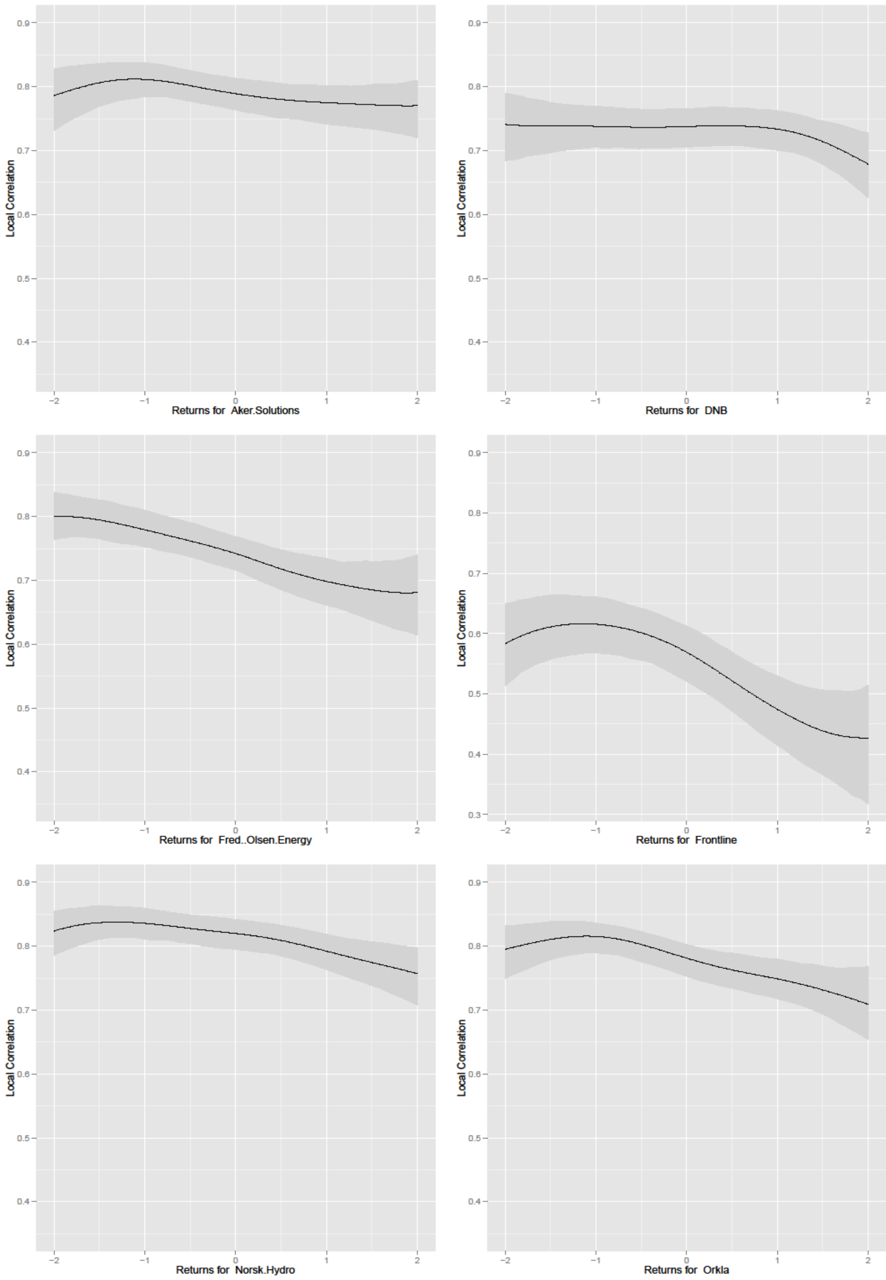


Figure 7.5: Local Gaussian Correlation between stocks and the OSEBX index(1/3)

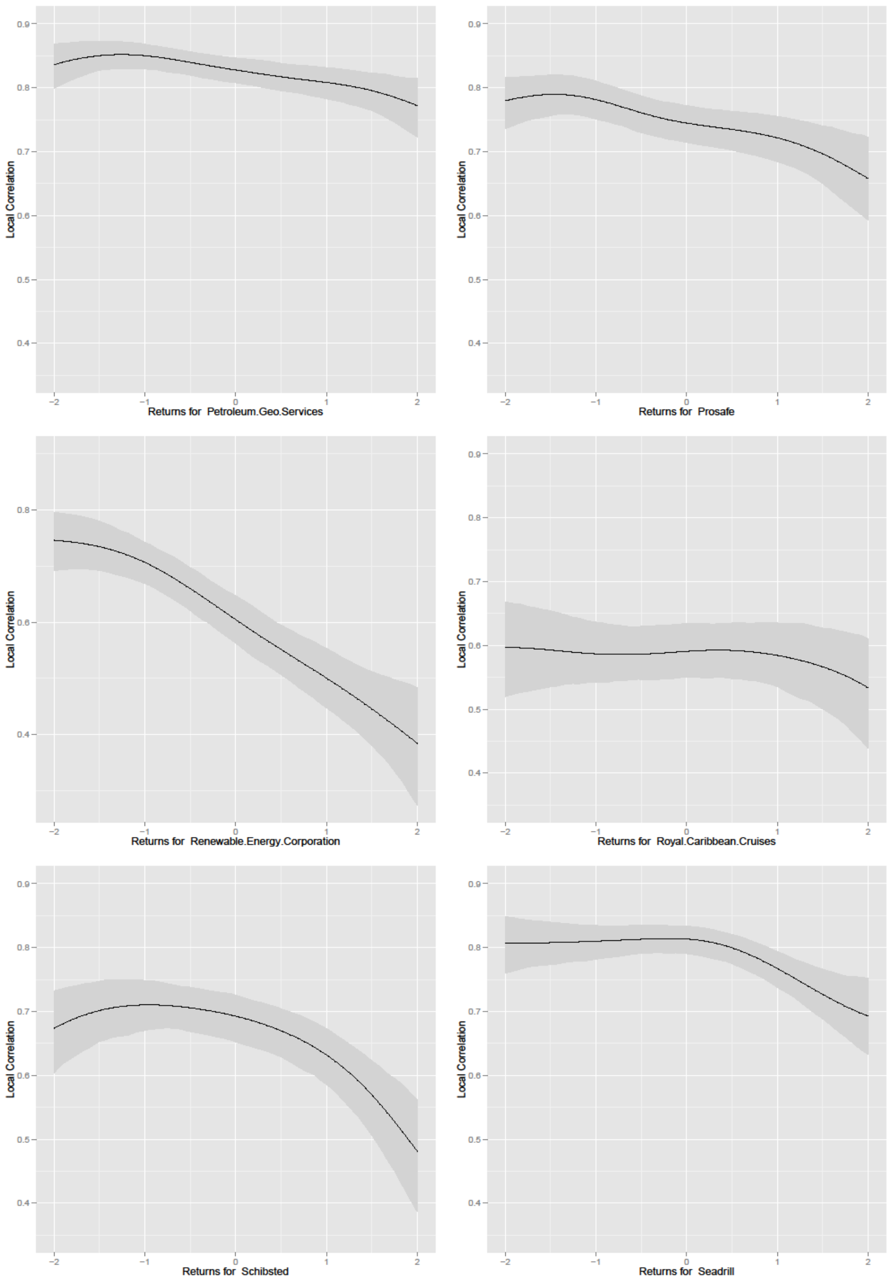


Figure 7.6: Local Gaussian Correlation between stocks and the OSEBX index(2/3)

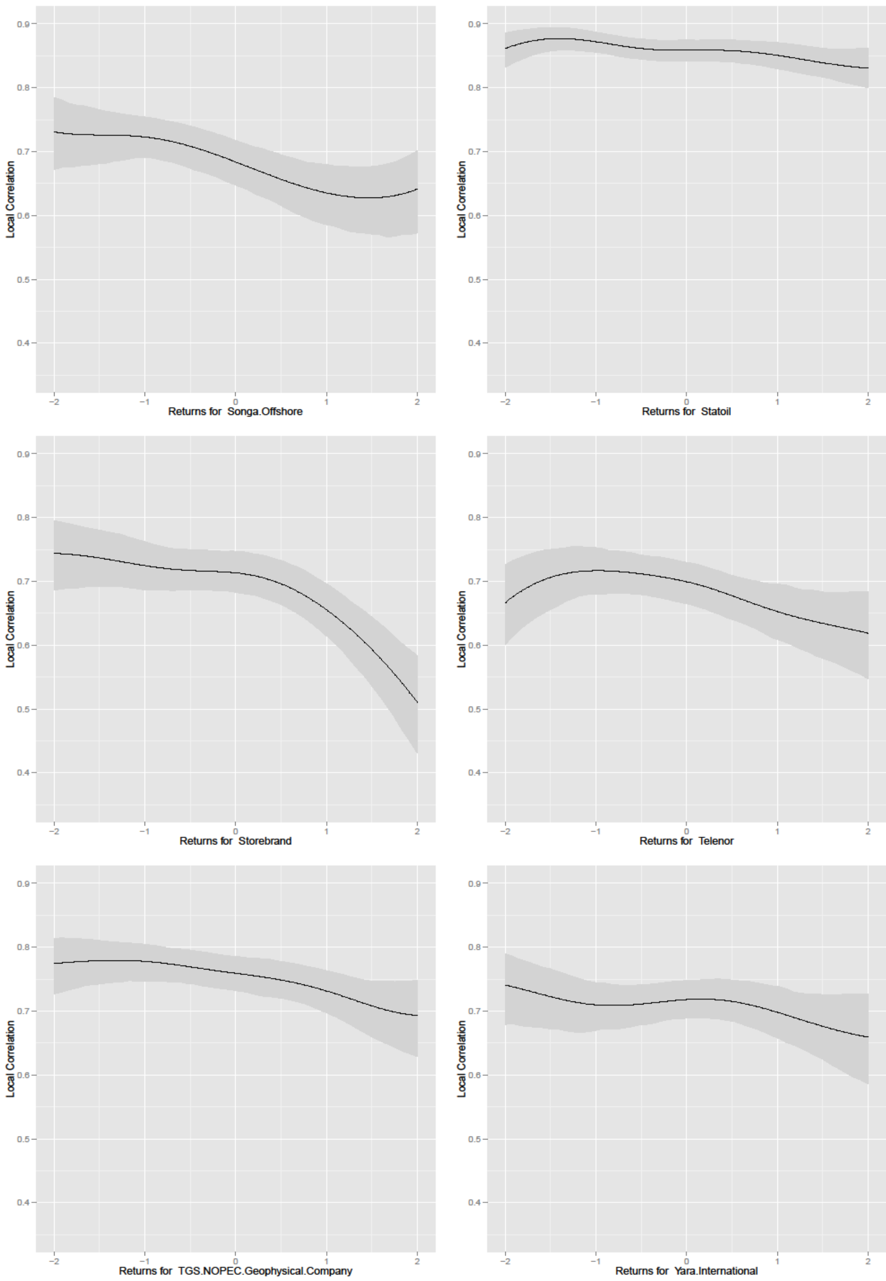


Figure 7.7: Local Gaussian Correlation between stocks and the OSEBX index(3/3)

7.4.2 Local Gaussian Correlation Between Stocks

In this investigation we will limit ourselves to the first 5 of the 18 stocks introduced in Table 7.1, namely Aker Solutions, DNB, Fred Olsen Energy, Frontline and Norsk Hydro.

Figure 7.8 and Figure 7.9 shows the Local Gaussian Correlation on the $x = y$-axis for the 10 possible combinations of these stocks. Overall, there seems to be quite a strong correlation between the stocks along the whole axis. If we exclude Frontline, the lowest value for the confidence interval of the correlation is right beneath 0.4, which is quite a strong correlation.

The lowest value of the correlation along the whole of the axis seems to be between Frontline and DNB. While Norsk Hydro and Aker Solutions seems to be strongest correlated with the different stocks, and the strongest correlation is found between them.

Due to the size of the confidence interval for the local correlation the claim of a constant correlation can probably not be disproved for all of the possible combinations, this is seen in for example the correlation plot between in DNB and Aker Solutions.

There are several of the plots strongly suggesting against a constant dependence structure, e.g. Fred Olsen Energy - Aker Solutions, Frontline - Aker solutions and Frontline - Fred Olsen Energy.

In Frontline - Fred Olsen Energy the correlation falls from almost 0.5 in -2 to right above 0.2 in 2, which is the largest change in the correlation among the stocks, but also note that the confidence interval is rather large for this combination.

By focusing on the observed values and not too much on the confidence interval there seems to be a trend of an increase in local correlation from -2, it reaches a reaches a peek somewhere around -1, before it decrease to a value similar to or lower than its value in -2. This peak is not seen between the individual stocks and the index.

This means that the correlation between the stocks is strengthened for small loses for the stocks, but not for larger loses. This might be explained by small loses for the stocks might be linked to small loses for the whole market, but large loss for either of the stocks might mean bad news for the particular company and not the market as a whole, thus weakening the correlation.

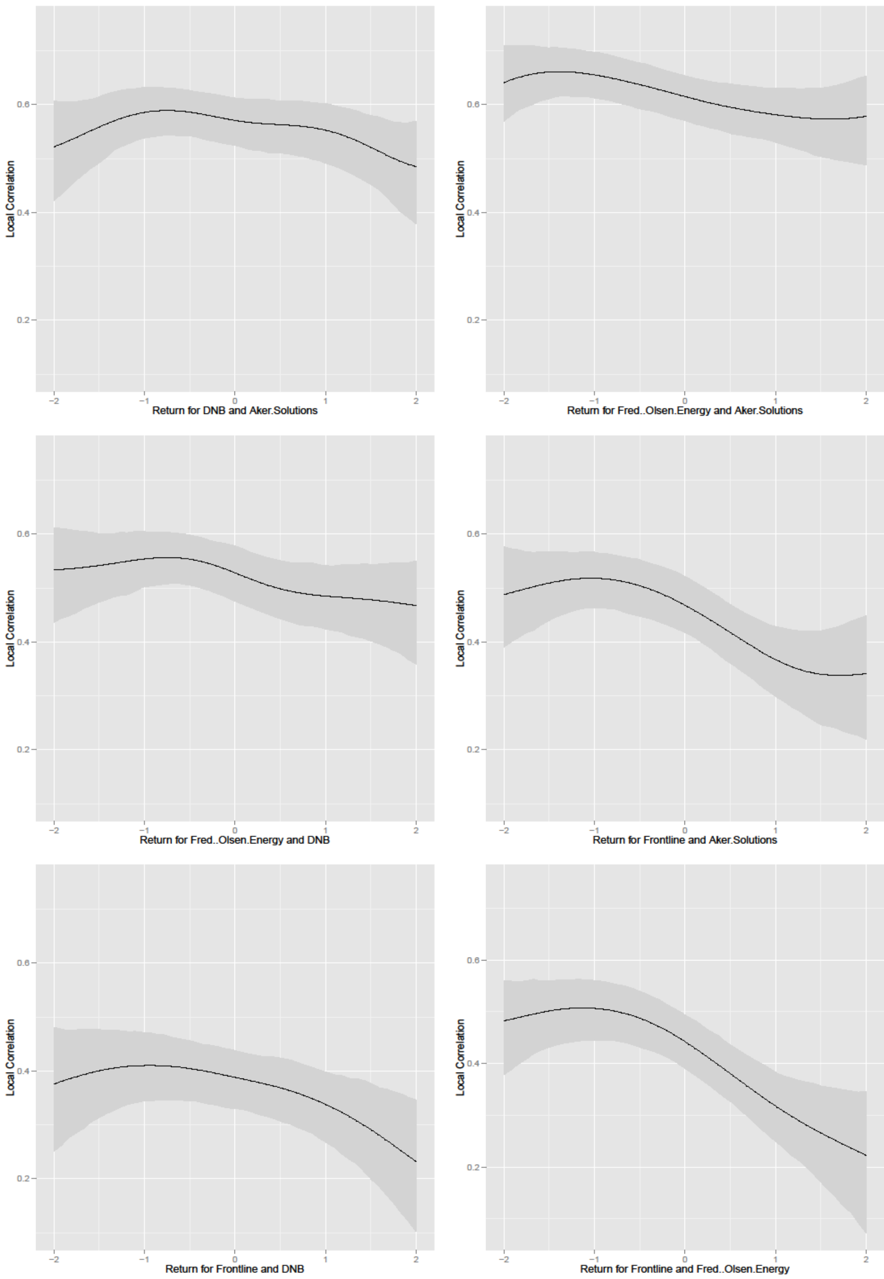


Figure 7.8: Local Gaussian Correlation between stock returns(1/2)

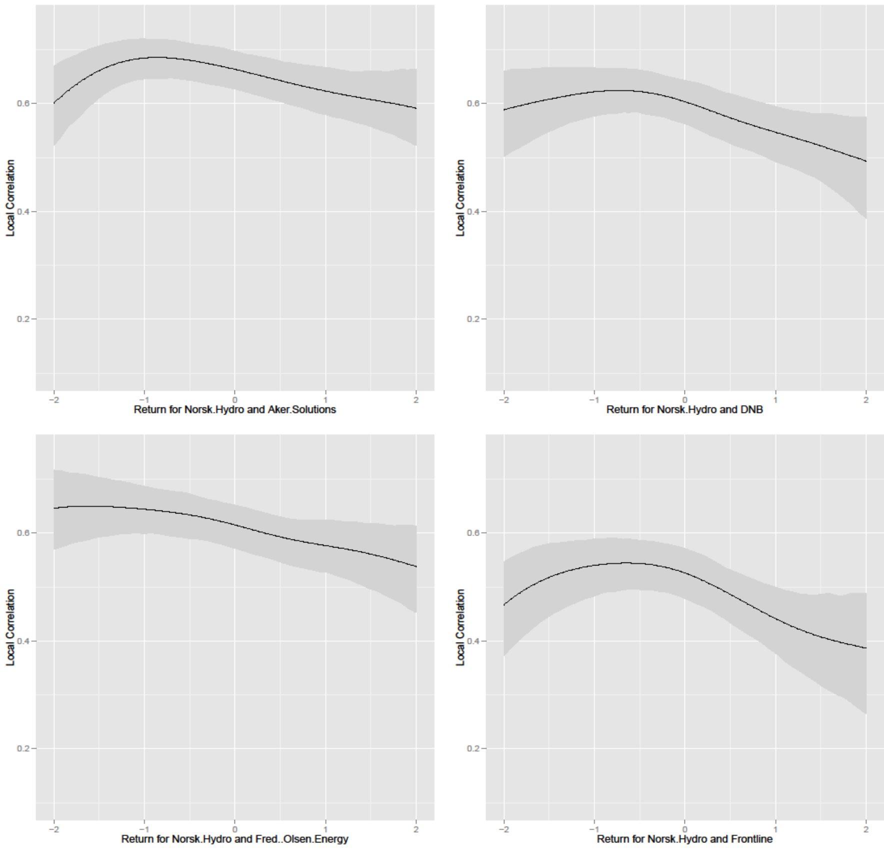


Figure 7.9: Local Gaussian Correlation between stock returns(2/2)

7.5 The CAPM on the Norwegian Stock Market

We have introduced some theory for the Capital Asset Pricing Model and seen how it(in theory) can be used under strict assumptions to describe the relationship between the expected returns and the risk for stocks.

By moving the return of the risk-free rate over to the opposite side of the equal sign in Equation 5.3 we have that the expectation for the stock can be written as follows

$$\mu_i = \mu_f + \frac{\sigma_{i,M}}{\sigma_M^2}(\mu_M - \mu_f) = \mu_f + \beta_i(\mu_M - \mu_f). \tag{7.1}$$

By letting $\mu_f > 0$ and μ_M be the estimate of the mean for the OSEBX index, which from Table 7.1 is -0.01, we easily see that a positive value for beta will result if μ_i is less than μ_f.

One of the assumptions for the CAPM as stated in Ruppert (2010) is that "‘All investors choose portfolios optimally according to the principles of efficient diversification."’ This means that no one would be investing in a stock where the expected return is less than the risk-free rate. This means that no one would be interested in investing in stocks, since $\mu_M < \mu_f$ leads to $\mu_i < \mu_f$. This is one of several reasons why the validity of the CAPM model has been questioned, some important parts in this debate has been the dispute between Black (1993) and Fama & French (1992).

In Fama & French (1992) the authors do not find the expected relation between average return and betas on observed data. Black (1993) questions the use of past average returns as a measure of the expected returns and criticize the authors of Fama & French (1992) among other researchers for questionable selection of data and misleading presentation of results as an attempt to disprove the CAPM model. This dispute is covered in several articles from the mentioned authors and others, but the important one for us will be the one of Pettengill *et al.* (1995).

In Pettengill *et al.* (1995) the authors investigate the topic of negative values for the market portfolio and the validity of CAPM due to this. They find that even though they cannot say anything about the validity of the CAPM. They do find results that support the continued use of beta as a measure of market risk.

Following is a table of values for the expected return for the stocks already introduced with a corresponding plot of the security market line where the risk-less rates is set to be 0.0054, and the expected return for the market portfolio is set to the observed mean of the OSEBX index(-0.01). The value for the risk-less rate is chosen since

$$100 \times log(1.02^{(1/365)}) \approx 0.0054,$$

which means 0.0054 is approximately the daily value for the logarithmic difference corresponding to a yearly rate of 2%.

Betas for Norwegian Stocks

	Expected Return	beta
Aker Solutions(1)	-0.016	1.477
DNB(2)	-0.011	1.146
Fred Olsen Energy(3)	-0.007	0.902
Frontline(4)	-0.010	1.080
Norsk Hydro(5)	-0.012	1.221
Orkla(6)	-0.008	0.935
Petroleum Geo Services(7)	-0.017	1.564
Prosafe(8)	-0.010	1.057
Renewable Energy Corporation(9)	-0.016	1.491
Royal Caribbean Cruises(10)	-0.008	0.923
Schibsted(11)	-0.008	0.957
Seadrill(12)	-0.013	1.308
Songa Offshore(13)	-0.016	1.528
Statoil(14)	-0.008	0.940
Storebrand(15)	-0.013	1.293
Telenor(16)	-0.005	0.739
TGS NOPEC Geophysical Company(17)	-0.014	1.352
Yara International(18)	-0.011	1.177

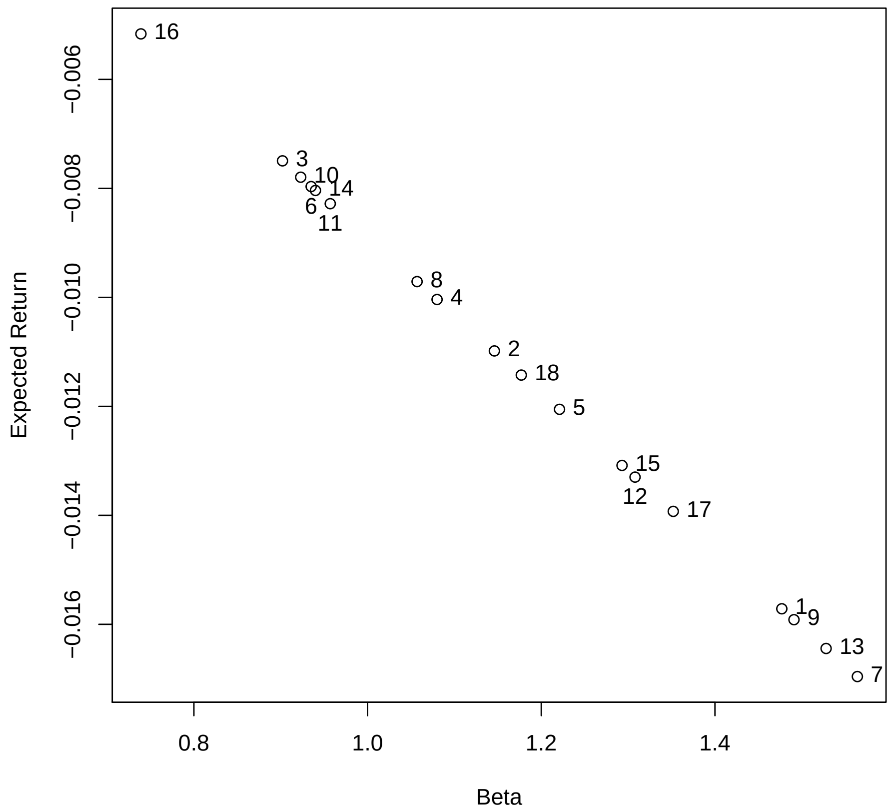


Figure 7.10: Values of betas and Expected Returns(according to Equation 7.1) for Norwegian stocks

7.5.1 Interpretation of the Betas

As already mentioned the above values for the expected returns does not necessarily make that much sense, because the average value of the market portfolio(OSEBX index) is negative for the considered time period, thus making it smaller than any reasonable value of the risk-less rate.

For the particular choice of value for the risk-less rate and the expected value for the market portfolio as given above one could clearly choose to invest in Telenor rather than the other stocks, as it has the highest expectation and the lowest risk for the stocks. As seen in Equation 7.1 the expectation for all the stock is bound to be less than the risk-free rate when the expected return for the market portfolio is less than the return for the risk-free rate, making the risk-free rate the most reasonable choice of investment.

The negative average value of the OSEBX index is not really that big a surprise. The period of our data is known for its financial difficulties, a period of losses, debt and unemployment around the world. Some of the background of the financial crisis can be found in the article Jones (2009) mentioned before, which focuses on collateralized debt obligation (CDO) mispricing. And for the choice of the value for the expectation of the market portfolio Black (1993) is probably right. The negative observed mean does necessarily match what the investors expected, and this is really the expectation of interest in the CAPM. The fact that people have been investing throughout this period, shows that they have clearly not been expecting to lose money. However, we will not focus on the relationship between the expected returns and the betas as stated in Equation 7.1, but rather use the beta as a measure of risk on is own.

The definition for beta in Equation 5.3 can be written as

$$\beta = \frac{\sigma_{i,M}}{\sigma_M^2} = \frac{\rho_{i,m}\sigma_i}{\sigma_M} \tag{7.2}$$

where $\sigma_{i,M} = Cov[R_i, R_M]$.
In the case of observed values, this can also be written as

$$\hat{\beta}\sum_{j=1}^{n}(R_{M,j} - \hat{\mu}_M)(R_{M,j} - \hat{\mu}_M) = \sum_{j=1}^{n}(R_{i,j} - \hat{\mu}_i)(R_{M,j} - \hat{\mu}_M)$$

were $R_{i,j}$ and $R_{M,j}$ are the observed j-th return for the i-th stock and the market portfolio with the corresponding expectations μ_i and μ_M .

By having a adequate number of observations, a beta of 1 will then suggest that the volatility of the stock is similar to the one of the market portfolio, while a beta larger than one gives us a risky stock, and a value for the beta which is less than one gives us a risk which is less than the market portfolio.

By matching the value of the beta in Table 7.10 by the empirical value for the mean given in Table 7.1, rather than the one calculated Equation 7.1 in Figure 7.10,
we find that TGS NOPEC Geophysical Company has the highest mean of all of the introduced stocks, but is a less risky stock(lower beta) than Aker, Renewable Energy Corporation, Songa Offshore and Petroleum Geo Services, suggesting for TGS NOPEC Geophysical Company to be a better investment than the more risked, but lower mean valued stocks mentioned above.

7.5.2 Local Betas and the Norwegian Stocks

As an example of estimating the Local Gaussian Correlation, we also obtain local values for the other parameters given in the bivariate Gaussian distribution. Since all we need to compute values for the betas is values for the correlation and the two standard deviations, we are able to calculate local values for the betas. This approach does not suffer from the previously mentioned bias for the conditional correlation described in Chapter 5.5.

Figure 7.12, Figure 7.13 and Figure 7.14 give local values for the betas for each of the introduced stocks, where the market portfolio is assumed to be described by the OSEBX index. The value for the beta and the confidence interval is calculated on the same format as set as a standard under the chapter of Local Gaussian Correlation, only here the value of the beta is calculated on 30 points equally spaced on the diagonal instead of 50.

In the mentioned figures we see that the all the local values for the betas for each stocks are between 1 and 0.4, however, there are large differences in these. In our observations there seems to be both symmetric and asymmetric changes in risk, where Fred Olsen Energy, Orkla, Petroleum Geo services seems to be approximately symmetric. Aker Solutions, DNB Norsk Hydro and Statoil are examples of asymmetry with higher value for betas and therefore higher risk in the positive tail.

Our observations differs mainly from what is found in Silvapulle & Granger (2001) by the fact that they find that there seem to be a trend of higher values for the beta in the lower tail than in the upper tail. Yara and Renewable Energy Corporation and Schibsted might suggest some minor decrease for the beta. Aker Solutions, DNB, Telenor, TGS NOPEC and Prosafe are examples, which suggest otherwise.

Figure 7.11 shows the mean and the corresponding confidence interval of the local values for the betas calculated on the 18 stocks, and according to this there do seem to be a trend of lower values for the betas for returns somewhere between -1 and 0 compared to the tails. The value of higher beta in the lower quantile than in the middle correspond to the result found in Silvapulle & Granger (2001). For the positive tail the returns seem to differ, as we observe higher values for the beta for large positive values than in the middle and lower quantile, whereas Silvapulle & Granger (2001) finds no noticeable difference between the middle and upper quantile.

That the lower tail has higher values for the betas than in the middle agree to what is seen by volatility, where large values of both sign, for the returns seem to appear in clusters. A larger beta associated with positive values for the returns than negative might suggest that very large gains for the stock market might be followed by a possible stock market downturn or even a crash (bubble), while negative values for the market is less likely to resolve in a sudden positive boost for the market.

Discussion on the Local Beta

In Figure 7.12, Figure 7.13 and Figure 7.14 the local value of beta is less than 1 along the whole diagonal on our interval. This corresponds to the values found in Silvapulle & Granger (2001), where they are less than one as well. We calculated the local values for the betas in these figures in a similar way as we calculated the local correlation, which means that we have used a GARCH-filter. Under the assumption that our return is described by a GARCH process, applying a GARCH-filter reduces our variable to a strict white noise(SWN) process.

Table 10.1 in Appendix B shows the summary of our data after we have GARCH-filtered our values. In the table, it is obvious that the OSEBX has the lowest value for the kurtosis in the table. In Tjøstheim & Hufthammer (2012) it is shown that for a distribution having thicker tail than the Gaussian distribution, the local variance goes to infinity in the tail. The lower kurtosis of the OSEBX index than the stocks will then result in the $\frac{\sigma_i(x,y)}{\sigma_M(x,y)}$, part of the formula for the beta in Equation 7.2 will be increasing throughout the tail, because the tail of the stocks is heavier than the one for the index. This corresponds with how our local values for the beta is increasing in the tail, even though the local correlation may be decreasing.

Figure 10.8 in Appendix B gives the local betas without a GARCH-filter for the first 5 stocks, and a mean calculated on all the 18 stocks. These values seem to match the values of the global betas in Figure 7.10, better than the one using a GARCH-filter. The fact that the local values for the betas is above and below 1 allows us to group the data in higher than and lower than market-risk, this is a nice feature. The mean in Figure 10.8 is calculated equal to the one in Figure 7.11, except from the filter and the number of replicates in the bootstrap procedure(5000 replicates in Figure 10.8 and 500 in Figure 7.11). Obvious it is quite risky to try to describe a trend when the values has not been filtered as there are large uncertainty in the values, as shown by the confidence interval. Additionally, the values for the local betas does not seem reasonable, a lower risk for large loss than in the middle does seem rather strange, and might suggest for another way of interpreting the non-filtered results.

There are some more reasons why we would prefer applying a GARCH-filter when calculating the local betas on the returns. The first reason is that the SWN process is assumed iid, which means that we can use bootstrap to calculate a confidence interval. This is not a major concern since it can probably be done similar by using the approach of block bootstrap introduced in Kunsch (1989). The second one is of larger concern and is also related to the fact that our returns are probably not iid. This lead to the local betas without a GARCH-filter is not calculated under similar conditions, and the value of beta is highly influenced by the volatility and is probably the reason why the non-filtered values did not seem reasonable, hence we will keep the calculation of the non-GARCH-filtered values just as a comparison to the filtered values.

It then follows that the local beta as calculated with the GARCH-filter is probably the best of the two approaches for a local value for the betas for non-iid data. This approach shows the change in risk for each of the stocks when the volatility is removed. The results with increasing risk in the tail correspond to how risk is expected to behave. That the values are less than one along the diagonal, at least in our limited interval might suggest for an interpretation different than the one for the global beta where $\beta < 1$ gives a low-risk stock, and $\beta > 1$ gives a high-risk stock. Since the local variance for the market portfolio will be similar for the stocks we compare to, the local beta will probably act better as a relative risk estimator between the stocks, than as a risk categorization criterion.

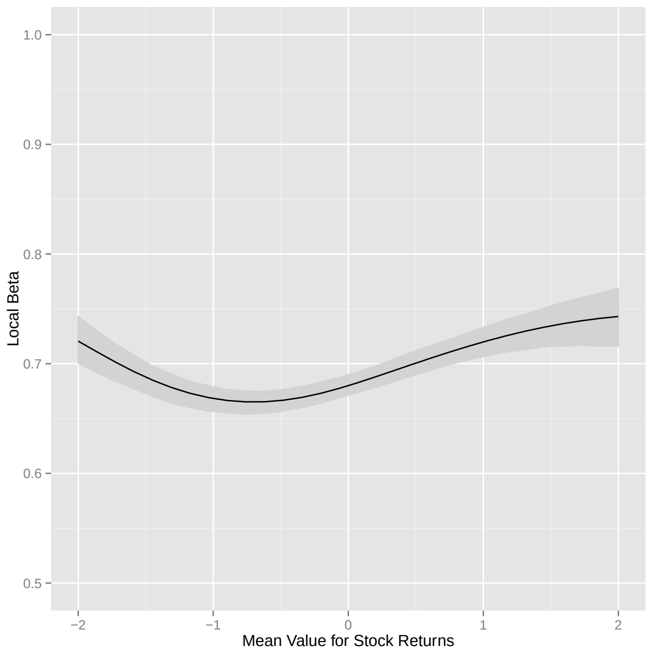


Figure 7.11: Mean estimated local beta between stocks and the OSEBX index

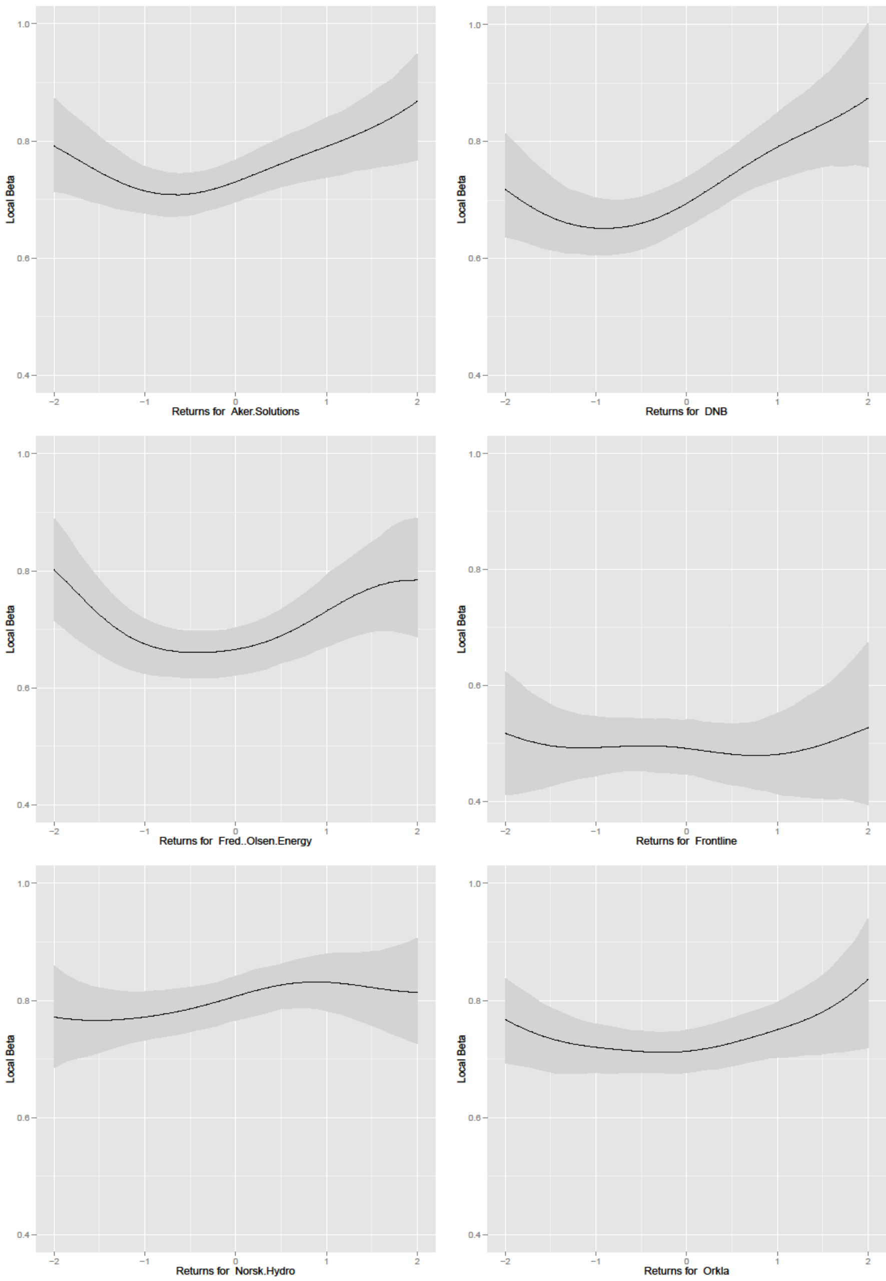


Figure 7.12: Change for the local beta for stock returns(1/3)

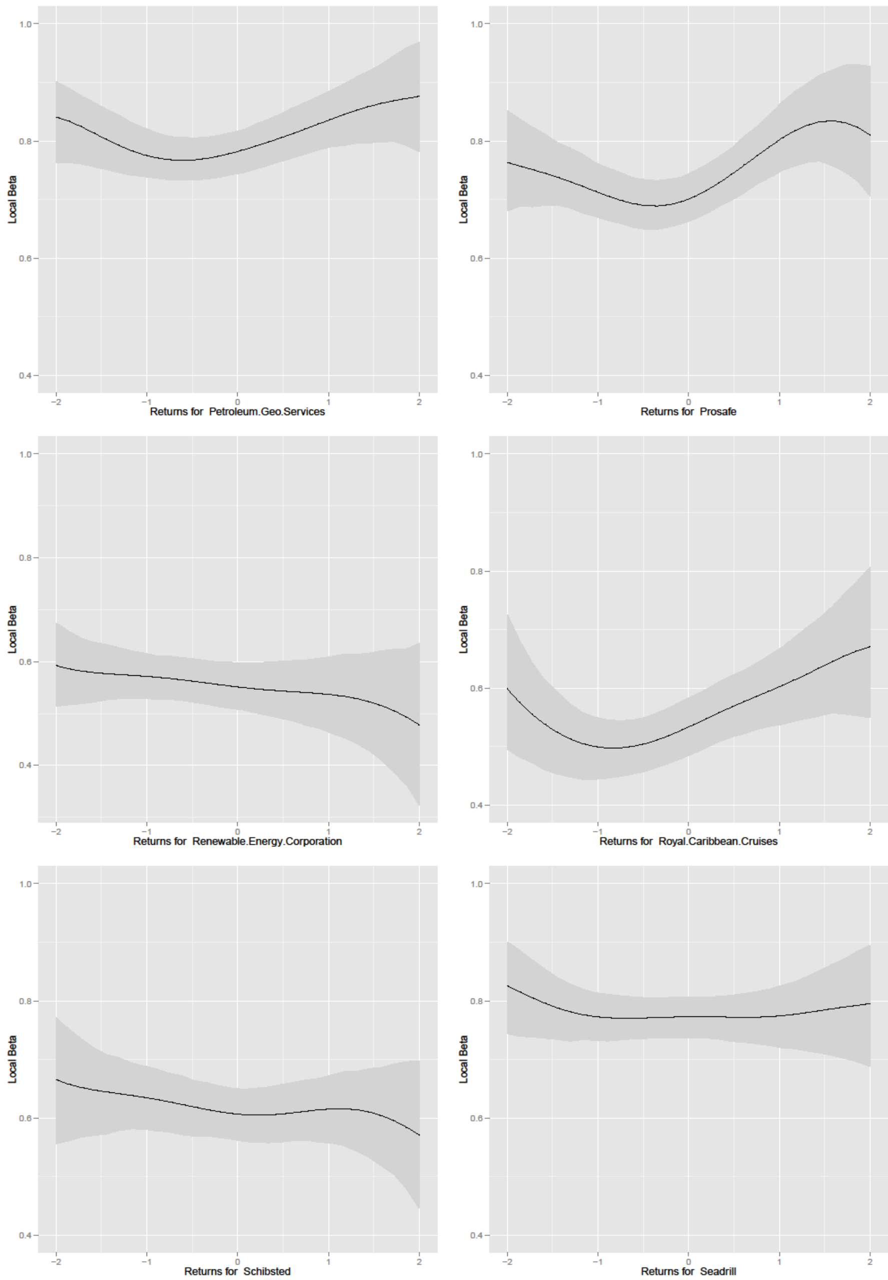


Figure 7.13: Change for the local beta for stock returns(2/3)

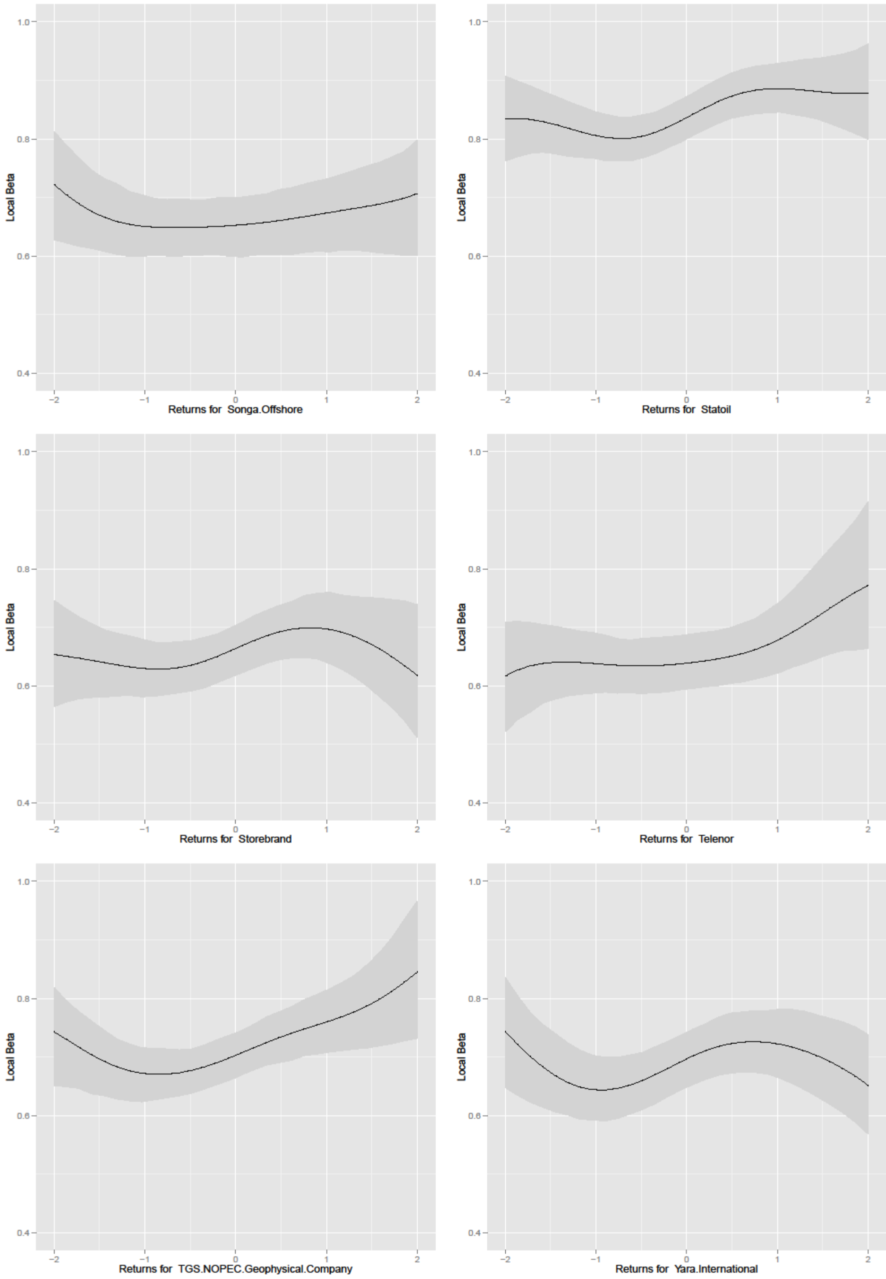


Figure 7.14: Change for the local beta for stock returns(3/3)

7.6 A Copula-GARCH Description of the Data

Under the chapter of Local Gaussian Correlation we estimated the change in correlation along the $x = y$-axis of random samples from the introduced copulas with standard Gaussian marginal distributions.

Berentsen *et al.* (2012) introduces a goodness-of-fit test for bivariate copula models based on the Local Gaussian Correlation, we will in a simpler manner try to find a copula which describe our stocks by visualization.

Since our plots of the Local Gaussian Correlation on the copulas are given with a Gaussian marginal, and our data does not seem to be well described by a Gaussian distribution, comparing the plots of the copulas with Figure 7.8 and Figure 7.9 might not be a good approach. This is because the local correlation in the plots is influenced by the marginal distributions. Because of this, we make the following transformation of the marginal distributions on the data.

$$x_t = \Phi^{-1}(F(\delta_t|\upsilon)) \tag{7.3}$$

$$\delta_t = \sqrt{\frac{\upsilon}{\upsilon - 2}}\frac{r_t - \hat{\mu}}{\hat{\sigma}_t}$$

Here Φ^{-1} is as usually the univariate Gaussian quantile function, σ_t is the volatility, μ_t is the mean, r_t is the return and $F(x|\upsilon)$ is the distribution function of a univariate t distribution with υ degrees of freedom.

This transformation will under the assumption that the t-GARCH model describes our data well, give us a meta distribution with the copula similar to the observed, but with standard Gaussian marginal distributions.
Note that when we apply the marginal distribution transformation we have not done an additional filter of heteroscedasticity by a GARCH-filter, because the heteroscedasticity is filtered out by the transformation.

Figure 7.15 and Figure 7.16 shows the Local Gaussian Correlation calculated on the marginal-transformed observations along the $x = y$-axis.
Comparing these plots with Figure 7.8 and Figure 7.9 shows that the marginal transformation does not drastically change the local correlation, but still the transformation removes a possible source of error.

A comparison between the original, GARCH-filtered and marginal transformed variables can be found in Appendix B, Figure 10.1 gives the 5 stocks plotted against the OSEBX index. Figure 10.4 and Figure 10.5 gives the GARCH-filtered returns plotted against each other and Figure 10.6 and Figure 10.7 is the marginal transformed returns. Notice that Figure 10.6 and Figure 10.7 has standardized interval on the axis while Figure 10.1, Figure 10.4 and Figure 10.5 does not.

7.6.1 Description of the Dependence Between the Stocks

Comparing Figure 7.15 and Figure 7.16 with the Frank copula plotted in Figure 6.4, seems to be a rather poor fit, as the Frank copula is characterized by a peak in $x = y = 0$, the observed data has a peak, but this is found on the negative side of the axis.

Similar the Gumbel copula in Figure 6.5 does not seem to match as it is increasing along the axis, which is a bad description for our data along most of the axis, where the correlation at -2 seems to higher than in 2.

The t copula in Figure 6.8 and Figure 6.7 is also a bad fit regardless of its degrees of freedom and correlation due to its symmetry and that its minimum value for the correlation is found in zero.

This leaves us with the options of Gaussian and Clayton copula. The Gaussian copula in Figure 6.6 with its constant correlation is probably not a good choice, but it seems like it would probably stay inside the confidence interval for some of our observed values.

But the best fit of our copulas is probably found by the Clayton copula in Figure 6.3.
It does not offer the peak found at -1 at the axis, but it matches the tail dependence observed in our stocks, where the correlation is higher at the negative tail than in the positive.

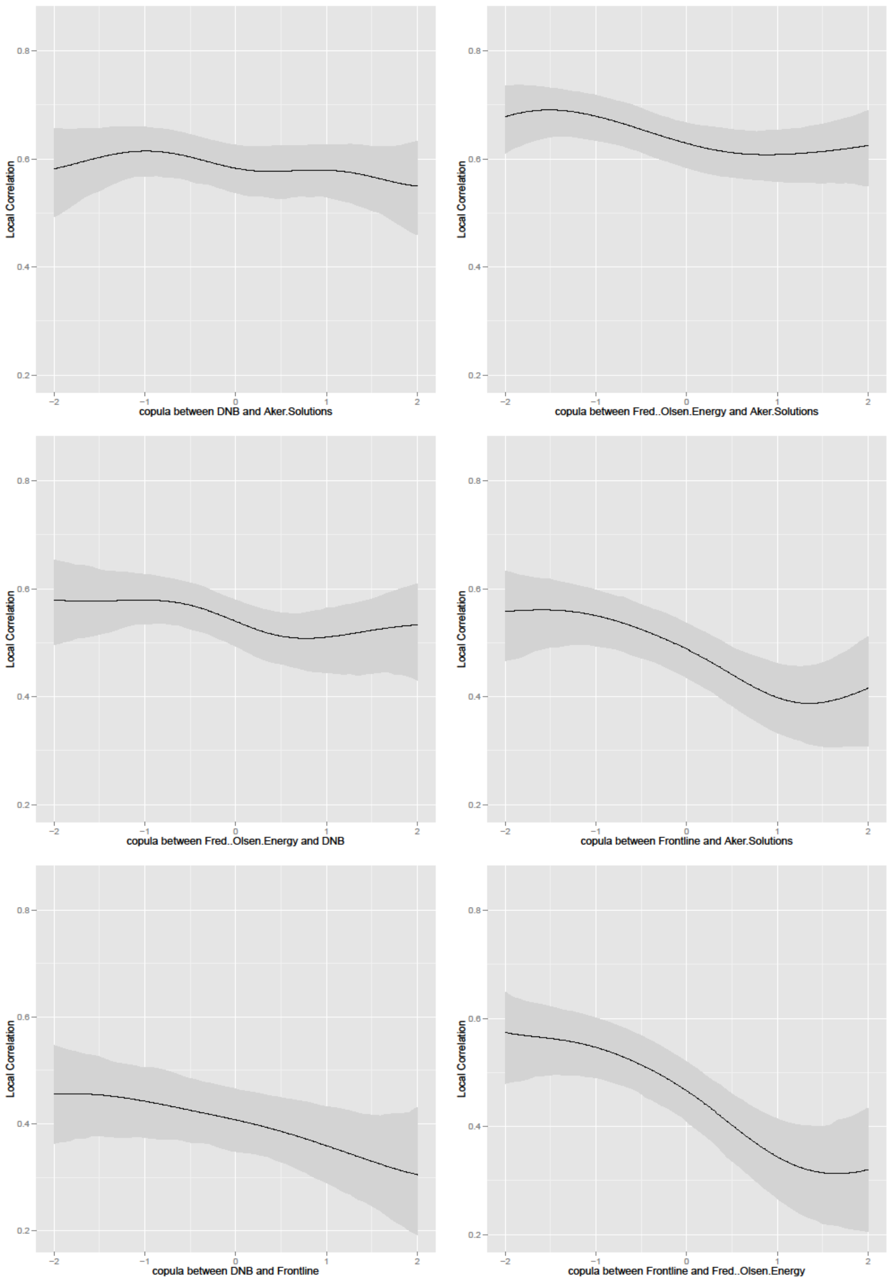


Figure 7.15: Local Gaussian Correlation on stock returns with transformed marginal distributions

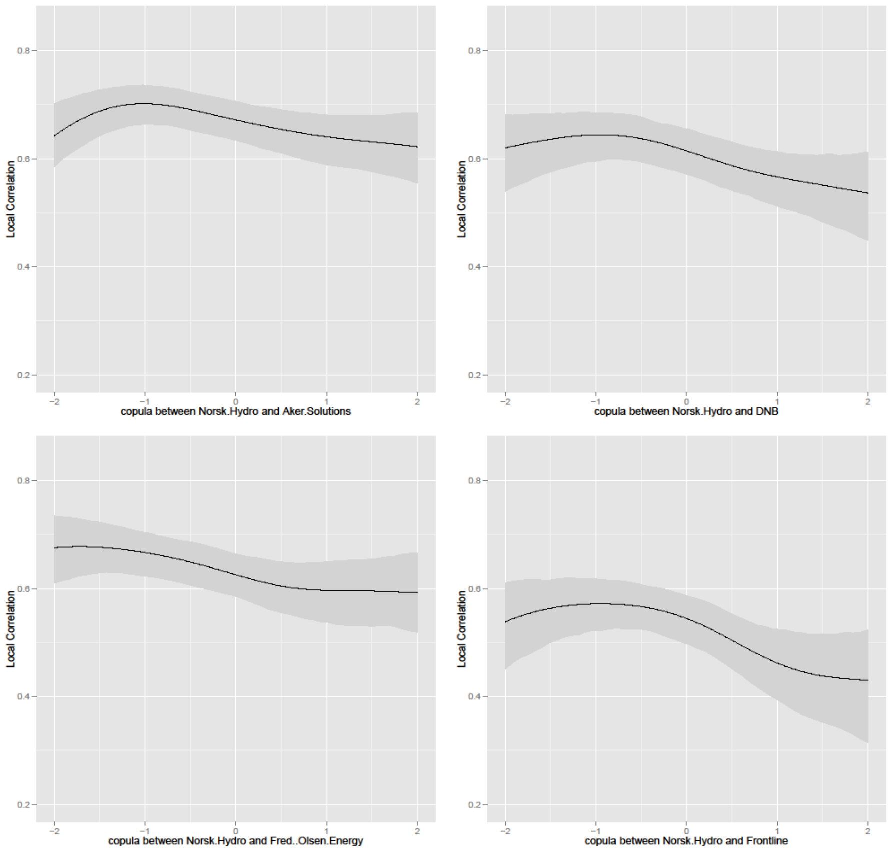


Figure 7.16: Local Gaussian Correlation on stock returns with transformed marginal distributions

7.6.2 Risk Estimation for a Copula-GARCH Model

Now that we have found a description for both the marginal distribution and the dependence structure for our stocks we can use it in risk estimation.
Table 7.3 gives the number of degrees of freedom(υ), the volatility for 'tomorrow' and the VaR for the stocks. The 'volatility' is the last value for the volatility calculated on our 1259 returns for the stock based on a GARCH(1,1) distribution with a standardized t distribution with υ degrees of freedom. The VaR is a parametric VaR for 'tomorrow' with the distribution described by the GARCH process and is calculated by

$$\mu_L + \sigma_{1260}\sqrt{\frac{\upsilon - 2}{\upsilon}} t_{\upsilon}^{-1}(0.95)$$

where μ_L is the mean of the loss, σ_{1260} is the last value for the volatility calculated on our 1259 returns for the stock and the numbers of degrees of freedom is given in the table.

	υ	Volatility	$VaR_{0.95}$
Aker Solutions	6.85	2.18	3.51
DNB	7.63	1.67	2.70
Fred Olsen Energy	6.23	1.71	2.73
Frontline	4.94	3.20	5.19
Norsk Hydro	10.00	2.24	3.71

Table 7.3: Degrees of freedom, volatility and VaR calculated for mentioned stocks

We have found that the Clayton copula seems to be the best description of the dependence between our data. McNeil *et al.* (2005) covers theory on how to estimate parameters for copulas. In this case, we have used a method based on the inverse of Kendall's tau to find the parameter that describes the dependence between two stocks. This has been done by the function fitCopula from the copula package.

Calculating the Value at Risk in the bivariate case is a little more complex than for the univariate, as a bivariate probability function has several potential values, which all give the probability of α. When estimating the Value at Risk for a copula consisting of two of our stocks we will therefore be using a Monte Carlo based risk calculation, which is covered in Hull (2009) for the univariate case.

By sampling a bivariate sample (X_i, Y_i), where $i = 1, \dots, 10000$, from our estimated Clayton copulas with t-GARCH models as a description of the marginal distributions we get a sample from a Meta Clayton-t-GARCH distribution.

To simplify our calculation we weight our stocks equally, and define the loss of the portfolio as the sum of the loss of the stocks and we have that

$$L_i = \mu_{L_1} + \mu_{L_2} + \sigma_{L_1,1260}\sqrt{\frac{\upsilon_1 - 2}{\upsilon_1}}t_{\upsilon_1}^{-1}(X_i) + \sigma_{L_2,1260}\sqrt{\frac{\upsilon_2 - 2}{\upsilon_2}}t_{\upsilon_2}^{-1}(Y_i)$$

Here L_i is the total loss for the equally weighted portfolio, μ_{L_1} is the mean, $\sigma_{L_1,1260}$ is the volatility and υ_1 is the number of degrees of freedom for one of the stocks, and corresponding for the second stock.

$L_{(500)}$ is the $VaR_{0.95}$ for this portfolio, while the mean of $L_{(1)} \cdots L_{(500)}$ is the $ES_{0.95}$. Table 7.4 gives the VaR and the ES for different combinations of the stocks based on the observed sample consisting of 10000 observations from the distribution, similar to how we previously calculated historical VaR and ES.
Similar to Table 7.3 where Norsk Hydro and Frontline, were the most risky stocks by itself, Table 7.4 shows that the portfolio consisting of these stocks is the most risky portfolio.

We have previously seen that these stocks are not independent, which means that the approach in Table 7.4 is probably a better approach than to summate the individual risk found in Table 7.3, at least under the assumption that the dependence is well described by the Clayton copula and the marginals by t-GARCH models. Here we have restricted ourselves to the bivariate case, this can off course be done in dimensions higher than two and the approach of pair-copula mentioned in Aas *et al.* (2009) is then a natural place to start.

$VaR_{0.95}$ for a bivariate portfolio

	DNB	Fred Olsen Energy	Frontline	Norsk Hydro
Aker Solutions	-5.994	-5.826	-7.996	-6.875
DNB		-5.096	-7.008	-6.085
Fred Olsen Energy			-7.074	-6.27
Frontline				-8.31

$ES_{0.95}$ for a bivariate portfolio

	DNB	Fred Olsen Energy	Frontline	Norsk Hydro
Aker Solutions	-8.355	-8.014	-11.293	-9.275
DNB		-7.075	-10.159	-8.243
Fred Olsen Energy			-10.133	-8.456
Frontline				-11.926

Table 7.4: Monte Carlo based VaR and ES for a bivariate Portfolio

7.7 Discussion of Risk Management on the Norwegian Stock Market

Under the theory of CAPM there is a relationship between the risk and the expected value of the returns. However, as seen in Figure 7.10 this does not seem to make much sense when we use the observed average as expectation for the market portfolio, which might suggest for other ways of estimating the expected returns of the market portfolio. Regardless of the validity of the relationship between the expected value and the betas we have seen that the theory of the CAPM can be used to describe the risk associated with stocks.

As the last part of the analysis on the Norwegian Stock Market, we will compare the results found by global values for the betas and VaR for the first 5 stocks in Table 7.1 moreover, we will discuss how these compares to the local beta.

Table 7.5 gives a summary for the risk associated with these stocks, the VaR is calculated under the assumption of a t-GARCH description of the return distribution and the beta is the global values. It should be noted that comparing these risk measures might be somewhat deceptive as they calculate different kind of risk, but we will return to this.

	beta	$VaR_{0.95}$
Aker Solutions	1.48	3.51
DNB	1.15	2.70
Fred Olsen Energy	0.90	2.73
Frontline	1.08	5.19
Norsk Hydro	1.22	3.71

Table 7.5: Values for beta and VaR calculated for mentioned stocks

For the global betas, Fred Olsen Energy is the less risky stock followed by Frontline, DNB, Norsk Hydro and finally Aker Solutions.

For the parametric t-GARCH VaR we have that DNB, Fred Olsen Energy, Aker Solutions, Norsk Hydro and finally Frontline has the VaR in increasing order.

We see that there is some difference between which of these risk measures suggest as the riskiest stocks. This might be explained by the fact that VaR focuses on the risk found in the quantile, while the global beta is a description of the overall risk for the asset.

Value at Risk, Expected Shortfall and the betas has the appealing form of giving us a simple number for the risk associated with the stock, but this simplicity is also a part of its weakness. Adding GARCH for a description of the volatility or other similar extensions does not change the fact that VaR and ES only gives us a number for the risk, which overall provides very little information.

Even under the assumption that our model describe the risk well, we are bound to exceed the $VaR_{0.95}$ approximately 12.6$(1260/5 \times 0.05)$ times a year. And for the Expected Shortfall assuming to have a good description for the mean in the tail is probably not very realistic. The global beta is calculated by constant values for the correlation and the variance, and as we have seen neither the dependence or the variance seem to be constant for our data.

This suggests that local betas could be used as a description for risk rather than focusing on the global risk or the risk found in the quantile. By looking at the values of the local betas for the risk, we are able to describe the change in risk against the change in the returns for the stocks. This gives the possibility of choosing stocks where the risk of the portfolio matches the one we are interested in, or in the spirit of the CAPM, carefully selecting stocks related to different industries, in an attempt of cancel out the risk.

Chapter 8

Conclusion

We have seen that there is a strong correlation between the main index of Oslo Stock Exchange and the price of oil in USD for daily data. The correlation is weakened by transforming the currency of the oil price over to NOK. Since the market of the NOK is mostly restricted to Norway, while the EURO or the USD is used in a larger market, the change in NOK is highly dependent on the change for the market. This again amplifies the correlation found between the oil price and the Norwegian market.

Overall there is a high local correlation between the stock and the index for the whole $x = y$-axis. From the mean value of the local correlations, calculated on the stocks, it is clear that the correlation between the stocks and the market is stronger for losses, than for positive returns. This might be explained by that in bad times for the stock market, one would expect people to be nervous and not invest that much, which is bad news for the individual stocks and the stock market as a whole, while positive values for the stock might often be explained by individual factors.

Statoil have the strongest correlation of our stocks and seems to have rather constant values for the correlation inside the interval of 0.8 to 0.9. While Renewable Energy Corporation has a high correlation when both the stock and the market index has large negative values, but the correlation is drastically weakened for large positive values. These findings might be seen as a support of the already seen tendency of large correlation between the Norwegian Market and the petroleum industry.

By using the local parameters for the correlation and the variances for the market portfolio(OSEBX index) and the stock, we can use the theory of CAPM to get a local value for the beta. Plotting the change in the beta along the $x = y$-axis allow us to see how the risk for the stocks changes along the axis. Analyzing the mean values for the stocks, there seems to be a trend of higher values for the betas in the lower tails than in the middle, but the highest values for the betas is often found in the positive tail.

The beta calculated on this approach differs from findings found in the article Silvapulle & Granger (2001), but our results seem to make sense when considering the nature of the stock market, where very large gains for the stock market might be followed by a possible stock market downturn.

We have seen that models based on non-constant volatility is probably a better description for the stock returns than models where the returns are assumed to be iid. Further we have seen that a NIG-GARCH is probably a better description than both the t-GARCH and Gaussian-GARCH model. This means that our results support previously mentioned assumptions of non-constant correlation, volatility and risk on the stock market.

Under the assumption that our data is described well by a t-GARCH model the Clayton copula seems to be the best description for the dependence between 5 chosen stocks, when choosing between the Gaussian, t, Clayton, Gumbel and Frank copula.
And we have seen how Monte Carlo theory can be used to calculate the VaR and ES under the assumption of a copula-GARCH model.

For further work on this subject, expanding the limit of the interval is a natural place to start. The choice of [-2,2] for the variables might have been a little bit too cautious, and one would expect interesting findings out in the tail. This thesis has focused on daily data and analysis on data of different time periods might also be of interest, as some trends are seen over longer time periods.

Another feature of CAPM that might be of interest is the Security Characteristic Line, which can be found in Ruppert (2010). The Security Characteristic Line gives an expression for the undiversifiable risk of the stock, moreover, a local value of this can be used to test for the benefit of portfolio diversification under different behavior on the market.

Overall, we find that the Local Gaussian Correlation brings an interesting new aspect to the field of finance and statistical modeling.

Chapter 9

Appendix A: Additional Theory

9.1 The Truncated Gaussian Distribution

If we let X be a bivariate Gaussian distributed variable and define the random variable Z as

$$Z = X \times I$$

Where I is an indicator function defined below

$$I = \begin{cases} 1 & \text{if } a \leq x \leq b \text{ and } c \leq y \leq d \\ 0 & \text{otherwise} \end{cases}$$

We get that the variable Z is a truncated bivariate Gaussian variable, with the following density

$$f(x, y|A) = \begin{cases} \frac{f(x,y)}{Pr(A)} & \text{if } a \leq x \leq b \text{ and } c \leq y \leq d \\ 0 & \text{otherwise} \end{cases} \tag{9.1}$$

Where

$$A = \{[a \leq x \leq b] \text{ and } [c \leq y \leq d]\} \text{ and } Pr(A) = \int_a^b \int_c^d f(x, y) dx dy$$

Rosenbaum (1961) offers the moments of interest for a Gaussian truncated bivariate on the interval where $a \leq x \leq \infty$ and $c \leq y \leq \infty$, however, we need them for $A = \{[a \leq x \leq b] \text{ and } [c \leq y \leq d]\}$. We will therefore derive the moments $E[Y|A]$, $E[Y^2|A]$ and $E[XY|A]$ for a general interval.

9.1.1 Moments for the Truncated Bivariate Gaussian

In the following section we let $\phi(x)$ be the univariate Gaussian density and $\Phi(x)$ its cumulative distribution function. We let $f(x, y)$ be the bivariate Gaussian density and $F(x, y)$ be the bivariate cumulative distribution function. At first, note that the integral over a non standardized Gaussian density can be written on the form

$$
\begin{aligned}
&\int_a^b\int_c^d \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} exp\Big(\frac{-1}{2(1-\rho^2)}\Big((\frac{x-\mu_1}{\sigma_1^2})^2 + \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + (\frac{y-\mu_2}{\sigma_2^2})^2\Big)\Big)dydx \\
&= \int_{\frac{a-\mu_1}{\sigma_1}}^{\frac{b-\mu_1}{\sigma_1}}\int_{\frac{c-\mu_2}{\sigma_2}}^{\frac{d-\mu_2}{\sigma_2}} \frac{1}{2\pi\sqrt{1-\rho^2}} exp\Big(\frac{-1}{2(1-\rho^2)}(u^2+2\rho uv+v^2)\Big)dudv \\
&= \int_{\frac{a-\mu_1}{\sigma_1}}^{\frac{b-\mu_1}{\sigma_1}}\int_{\frac{c-\mu_2}{\sigma_2}}^{\frac{d-\mu_2}{\sigma_2}} \frac{1}{\sqrt{1-\rho^2}}\phi(v)\phi(\frac{u-\rho v}{\sqrt{1-\rho^2}})dudv
\end{aligned}
\tag{9.2}
$$

Since the limits of the integral can hold the information regarding the variance and expectation, we only need to derive the following moments for the standardized distribution. We write the first partial derivative of $f(x,y)$ with respect to y as

$$
\begin{aligned}
\frac{\delta}{\delta y}f(x,y) &= \frac{\phi(x)}{\sqrt{1-\rho^2}}\phi(\frac{y-\rho x}{\sqrt{1-\rho^2}})\frac{\rho x - y}{(1-\rho^2)} \\
&= \frac{\rho x}{(1-\rho^2)}\frac{\phi(x)}{\sqrt{1-\rho^2}}\phi(\frac{y-\rho x}{\sqrt{1-\rho^2}}) - \frac{y}{(1-\rho^2)}\frac{\phi(x)}{\sqrt{1-\rho^2}}\phi(\frac{y-\rho x}{\sqrt{1-\rho^2}})
\end{aligned}
\tag{9.3}
$$

Similar the partial second partial derivative of $f(x,y)$ with respect to y then x is

$$
\begin{aligned}
\frac{\delta^2}{\delta x\delta y}f(x,y) &= \frac{\delta}{\delta y}(\frac{\rho x}{(1-\rho^2)}\frac{\phi(x)}{\sqrt{1-\rho^2}}\phi(\frac{y-\rho x}{\sqrt{1-\rho^2}}) - \frac{y}{(1-\rho^2)}\frac{\phi(x)}{\sqrt{1-\rho^2}}\phi(\frac{y-\rho x}{\sqrt{1-\rho^2}})) \\
&= \frac{\delta}{\delta x}(\frac{\rho x}{(1-\rho^2)}\frac{\phi(x)}{\sqrt{1-\rho^2}}\phi(\frac{y-\rho x}{\sqrt{1-\rho^2}})) - \frac{y(\rho y - x)}{(1-\rho^2)}\frac{\phi(x)}{\sqrt{1-\rho^2}}\phi(\frac{y-\rho x}{\sqrt{1-\rho^2}})
\end{aligned}
\tag{9.4}
$$

The First Moment for the Truncated Bivariate Gaussian Distribution

By taking the double integral and some simple algebra on Equation 9.3, we end up with the following expression for the first moment.

$$
\begin{aligned}
E[Y|A]Pr(A) &= E[X|A]Pr(A)\rho - \Big[_{y=c}^{y=d}(1-\rho^2)\int_a^b f(x,y)dx\Big] \\
&= E[Y|A]Pr(A)\rho^2 - \Big[_{x=a}^{x=b}\rho(1-\rho^2)\int_c^d f(x,y)dy\Big] - \Big[_{y=c}^{y=d}(1-\rho^2)\int_a^b f(x,y)dx\Big] \\
&= \Big[_{x=a}^{x=b}\Big[_{y=c}^{y=d} - \phi(y)\Phi(\frac{x-\rho y}{\sqrt{1-\rho^2}}) - \rho\phi(x)\Phi(\frac{y-\rho x}{\sqrt{1-\rho^2}})\Big]\Big]
\end{aligned}
\tag{9.5}
$$

Where the second step follows by symmetry between $E[X|A]$ and $E[Y|A]$.

The Second Moments for the Truncated Bivariate Gaussian Distribution

The second moments is a little bit more complicated than the first moment, we will solve the expression of $E[XY]$ and $E[Y^2]$ simultaneously.
From Equation 9.4 we have that

$$(1-\rho^2)\int_a^b\int_c^d \frac{\delta^2}{\delta x\delta y}f(x,y)dxdy = \Big[_{x=a}^{x=b}\Big[_{y=c}^{y=d}\rho\phi(x)\Phi(\frac{y-\rho x}{\sqrt{1-\rho^2}})\Big]\Big] - \frac{\rho}{1-\rho^2}E[Y^2|A]Pr(A)$$
$$+\frac{1}{1-\rho^2}E[XY|A]Pr(A)$$

And by multiplying in a factor of y into the first derivative in Equation 9.3, we have that

$$\begin{aligned}(1-\rho^2)\int_a^b\int_c^d y\frac{\delta}{\delta y}f(x,y)dydx &= \rho E[XY|A]Pr(A) - E[Y^2|A]Pr(A)\\ &= \Big[_{x=a}^{x=b}\Big[_{y=c}^{y=d}(1-\rho^2)y\phi(y)\Phi(\frac{x-\rho y}{\sqrt{1-\rho^2}}) - (1-\rho^2)F(x,y)\Big]\Big]\end{aligned}$$

And by merging the two equations and some algebra we get the following expressions for the moments

$$\begin{aligned}E[XY|A]Pr(A) =&\Big[_{x=a}^{x=b}\Big[_{y=c}^{y=d}\sqrt{1-\rho^2)}\phi(x)\phi(\frac{x-\rho y}{\sqrt{1-\rho^2}})\\ &-\rho x\phi(x)\Phi(\frac{y-\rho x}{\sqrt{1-\rho^2}}) - \rho y\phi(y)\Phi(\frac{x-\rho y}{\sqrt{1-\rho^2}})\Big]\Big] + \rho Pr(A)\end{aligned} \tag{9.6}$$

$$\begin{aligned}E[Y^2|A]Pr(A) =&\Big[_{x=a}^{x=b}\Big[_{y=c}^{y=d}\rho\sqrt{1-\rho^2}\phi(x)\phi(\frac{x-\rho y}{\sqrt{1-\rho^2}})\\ &-\rho^2 x\phi(x)\Phi(\frac{y-\rho x}{\sqrt{1-\rho^2}}) - y\phi(y)\Phi(\frac{x-\rho y}{\sqrt{1-\rho^2}})\Big]\Big] + Pr(A)\end{aligned} \tag{9.7}$$

Probability of Observing the Truncated Area A

$$\int_a^b\int_c^d \frac{\phi(x)}{\sqrt{1-\rho^2}}\phi(\frac{y-\rho x}{\sqrt{1-\rho^2}})dydx$$
$$= \Big[_{x=a}^{x=b}\int_c^d x\frac{\phi(x)}{\sqrt{1-\rho^2}}\phi(\frac{y-\rho x}{\sqrt{1-\rho^2}})dy\Big] - \int_a^b\int_c^d x\frac{\rho y-x}{1-\rho^2}\frac{\phi(x)}{\sqrt{1-\rho^2}}\phi(\frac{y-\rho x}{\sqrt{1-\rho^2}})dydx$$
$$= \Big[_{y=c}^{y=d}\Big[_{x=a}^{x=b}x\phi(x)\Phi(\frac{y-\rho x}{\sqrt{1-\rho^2}})\Big]\Big] - \frac{Pr(A)}{1-\rho^2}(\rho E[XY|A] - E[X^2|A])$$

$$Pr(A) = \frac{\Big[_{y=c}^{y=d}\Big[_{x=a}^{x=b}x\phi(x)\Phi(\frac{y-\rho x}{\sqrt{1-\rho^2}})\Big]\Big]}{1+\frac{1}{1-\rho^2}(\rho E[XY|A] - E[X^2|A])} \tag{9.8}$$

Evaluation of the Truncated Moments

In order of get an understanding of how truncation affects the moments we will show the change for the moments by plotting some examples similar to what we did with the betas in Figure 5.1. We will consider the truncation on two different bivariate Gaussian distributions, the standardized distribution and one where the variances are set to two and three. Both of the distribution has correlation equal to 0.5 and unconditioned expectations equal to zero. The moments are calculated in **R** by using the expression for the moments above, while $\Pr(A)$ is calculated by pmnorm from the package mnormt.

Figure 9.1 shows the truncated standardized Gaussian distribution with correlation equal to 0.5. The upper plot shows the change in correlation and the variances for the standardized bivariate Gaussian distribution as the lower truncated is moved from the 0.95 quantile towards 1 on both its margins, while the lower plot gives the expectations under similar conditions. We only see one value for the expectation and the variances since these are equal throughout the whole axis.

Figure 9.2 has the same truncation as the one in Figure 9.1, but the covariance matrix is changed. Now the value for the correlation is still equal to 0.5, but the values for the variance is 2 and 3. From these plots it is obviously that when the lower truncation is increasing the value for the expectation is also increasing and the values for the correlation and variances is decreasing.

And finally Figure 9.3 shows the change in correlation and the variances for the distribution mentioned above, only now the truncation is given with both upper and lower truncation and is relaxed along the x-axis resulting in increasing probability of A. In the last case we have not plotted the expectations as these have a constant value of zero under the truncation described above. As we can see on the plots the moments is approaching the global values for the moments as the truncation are removed.

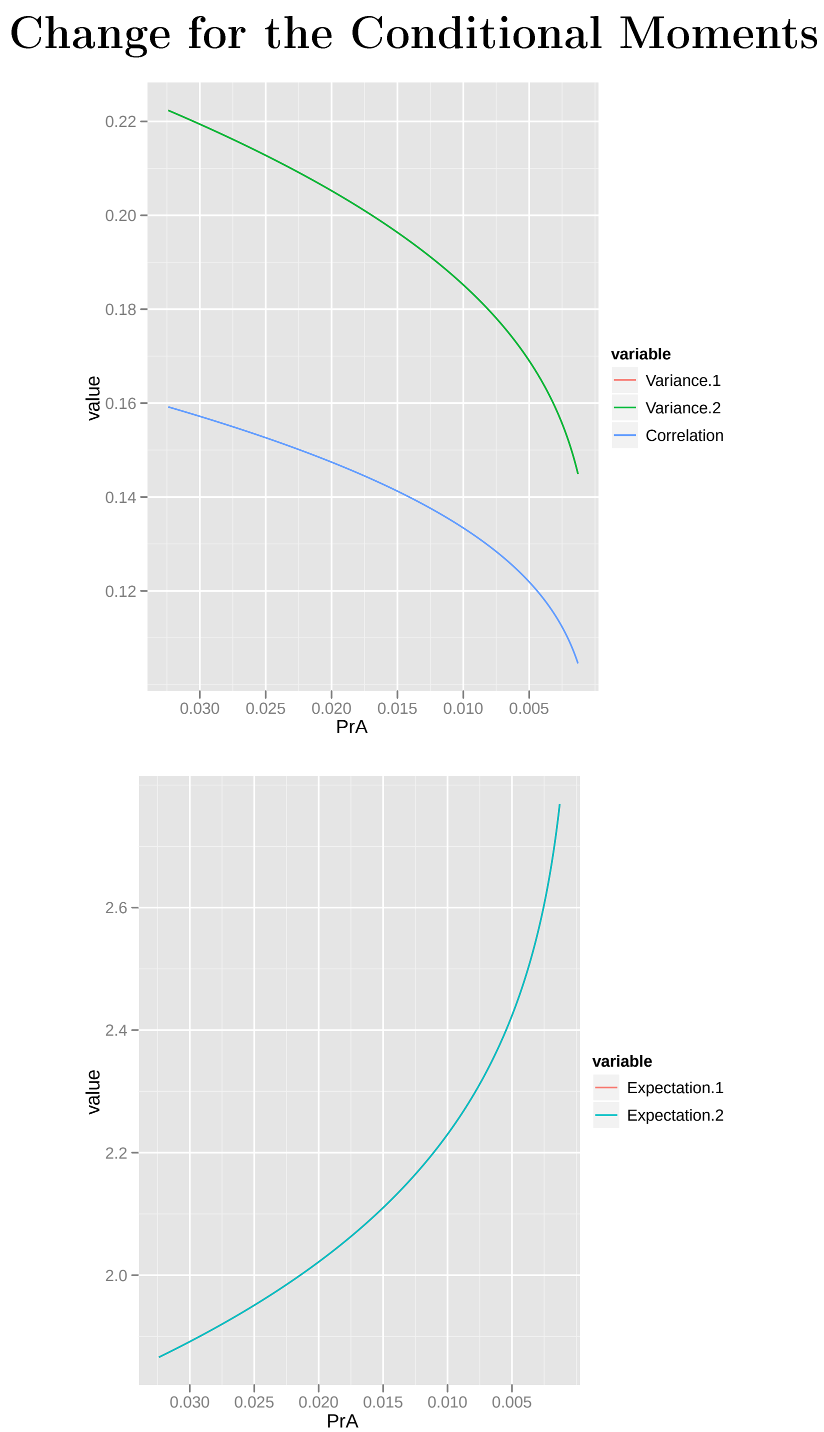


Figure 9.1: Change in moments and lower truncation on a standardized Gaussian distribution

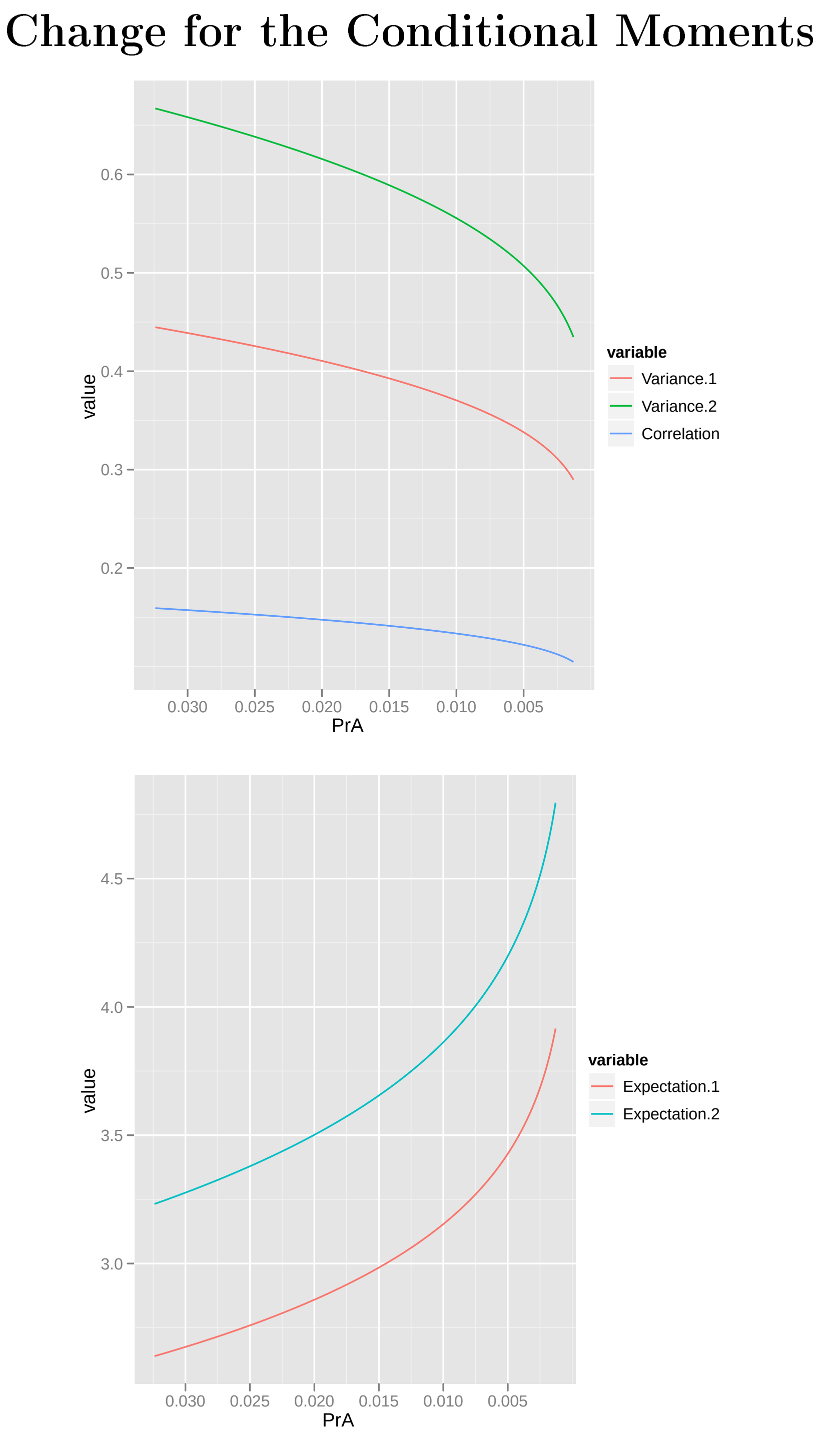


Figure 9.2: Change in moments and lower truncations on a bivariate Gaussian distribution

Change for the Conditional Moments

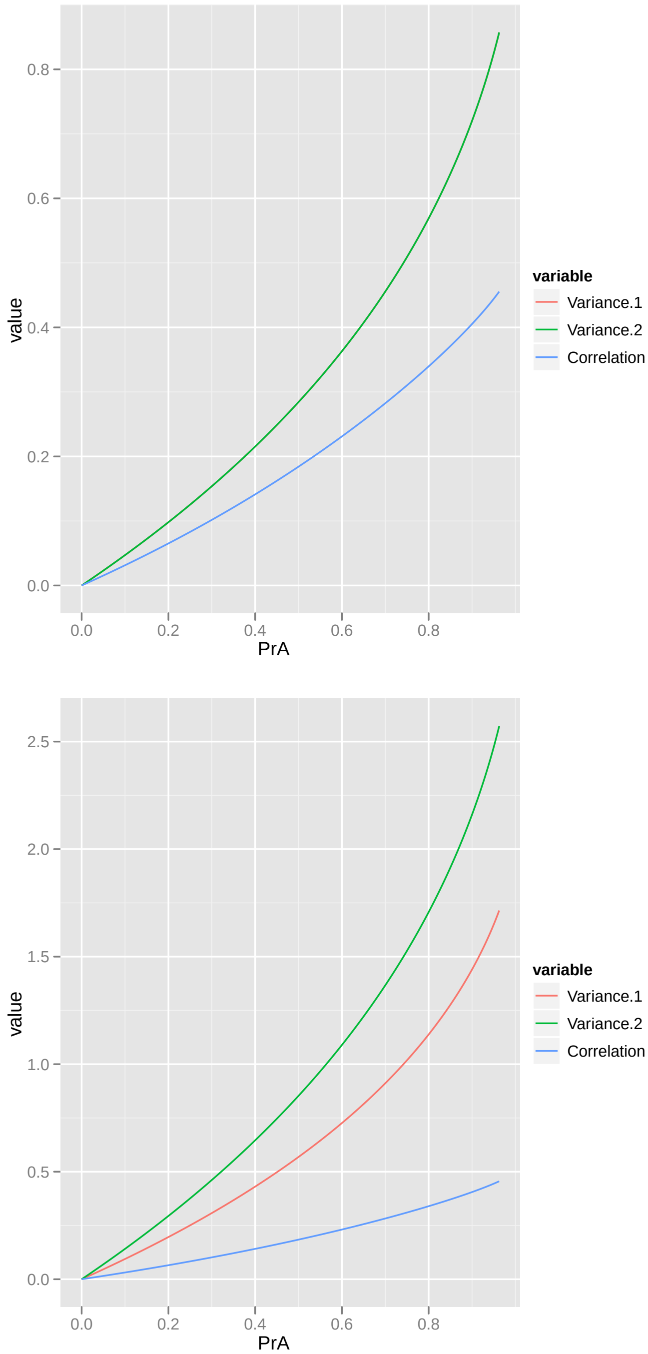


Figure 9.3: Change in moments and truncations(upper and lower) on two bivariate Gaussian distributions with different variance and correlation

Chapter 10

Appendix B: Figures and Tables of Less Importance

	Mean	Var	Median	Kurtosis	Skewness
Oslo Børs Benchmark Index	0.00	0.96	0.09	0.10	-0.27
Aker Solutions	0.01	1.01	0.01	2.71	0.15
DNB	-0.00	0.98	-0.02	1.19	0.04
Fred Olsen Energy	0.00	0.99	0.00	2.01	-0.42
Frontline	0.00	1.01	0.01	8.37	-0.40
Norsk Hydro	0.01	0.98	0.03	0.58	-0.10
Orkla	0.00	1.01	0.04	3.54	-0.42
Petroleum Geo Services	0.02	0.99	0.01	0.93	-0.06
Prosafe	0.00	0.98	0.01	0.45	-0.01
Renewable Energy Corporation	0.01	1.01	-0.01	4.13	-0.36
Royal Caribbean Cruises	0.00	0.98	0.01	1.73	0.13
Schibsted	0.00	0.98	0.00	2.76	0.15
Seadrill	0.00	0.99	0.07	0.65	-0.19
Songa Offshore	0.01	1.04	0.04	8.60	-0.92
Statoil	0.00	1.01	0.01	1.93	-0.42
Storebrand	0.00	1.00	-0.00	2.12	-0.04
Telenor	0.01	1.01	0.03	3.55	0.18
TGS NOPEC Geophysical Company	0.01	1.03	0.00	4.96	-0.43
Yara International	0.00	0.98	0.01	0.59	-0.19

Table 10.1: Summary of GARCH-filtered data concerning the Oslo Stock Exchange

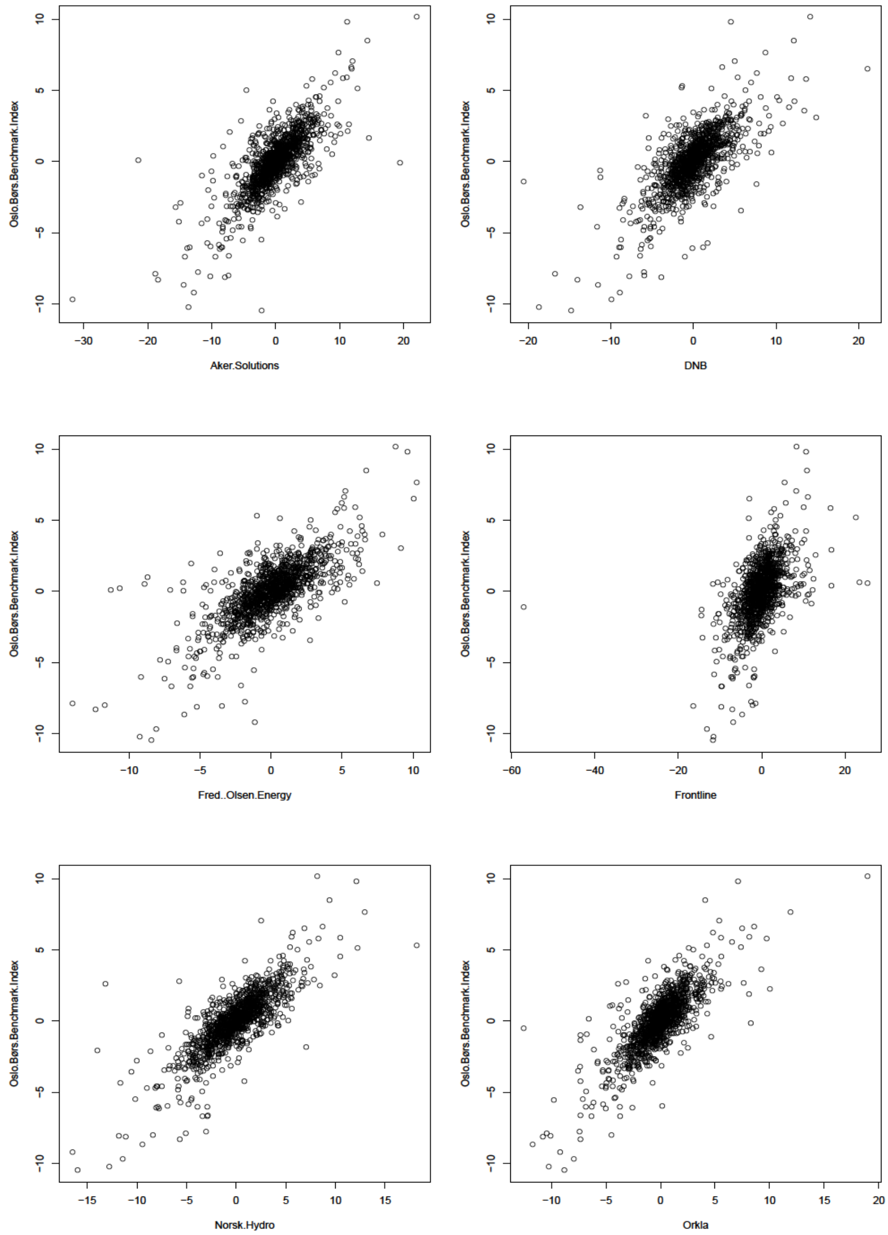


Figure 10.1: Scatterplot between OSEBX and stocks (1/3)

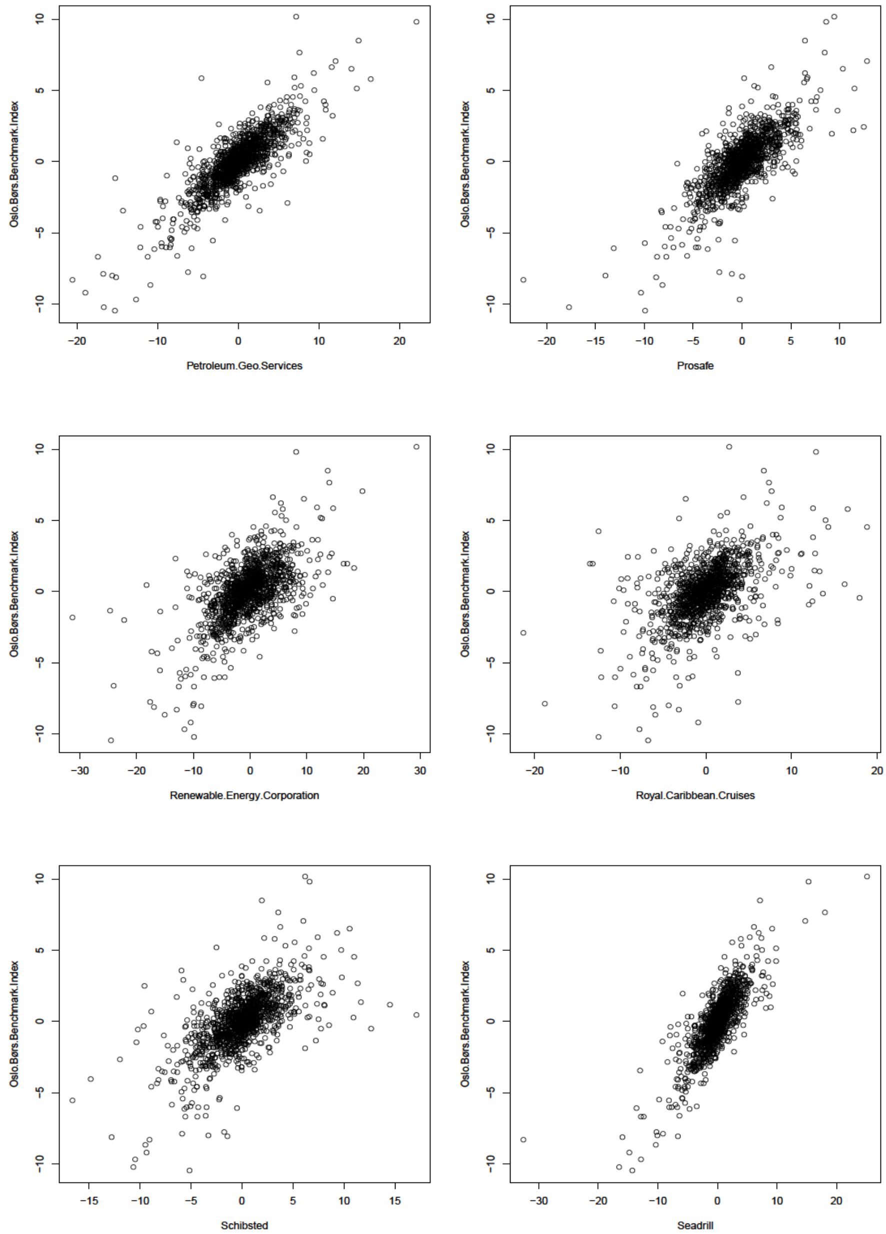


Figure 10.2: Scatterplot between OSEBX and stocks (2/3)

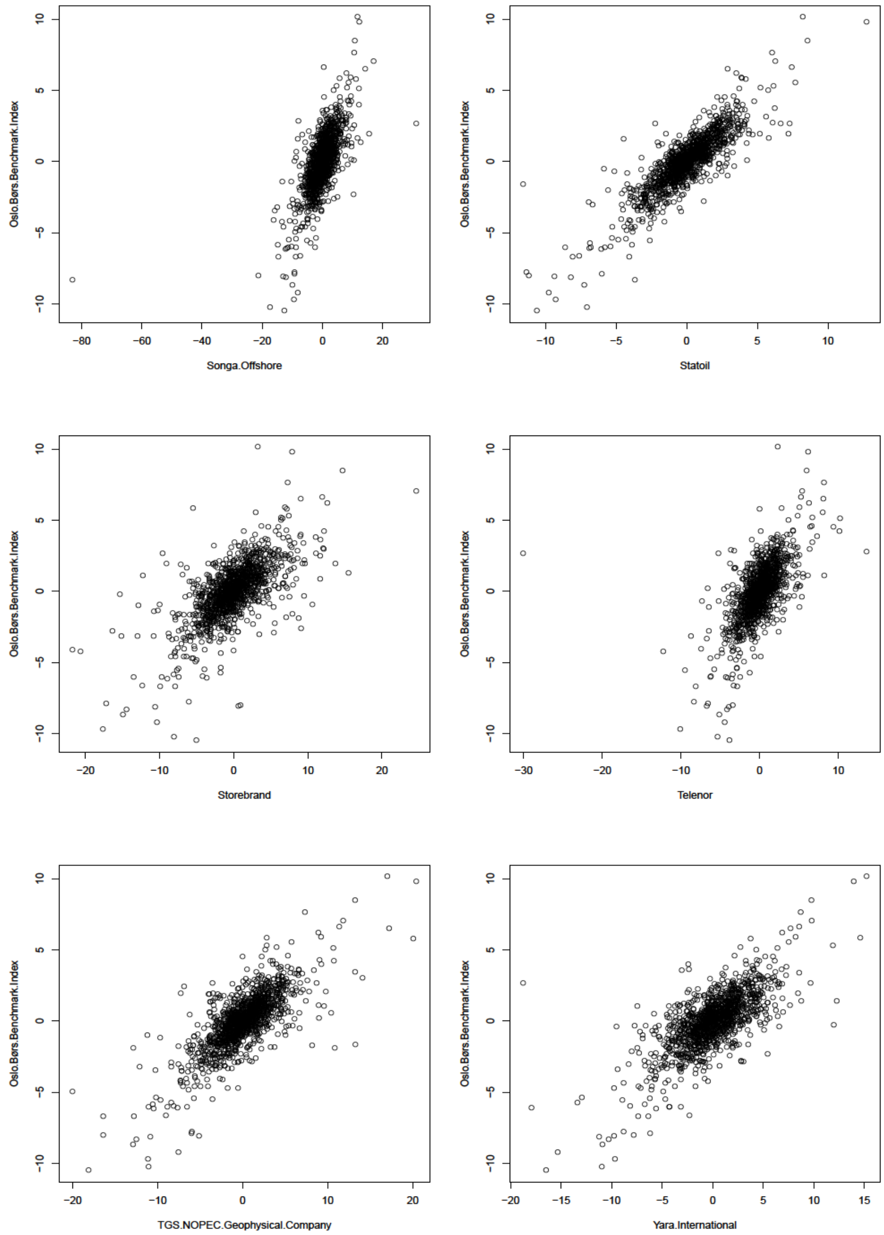


Figure 10.3: Scatterplot between OSEBX and stocks (3/3)

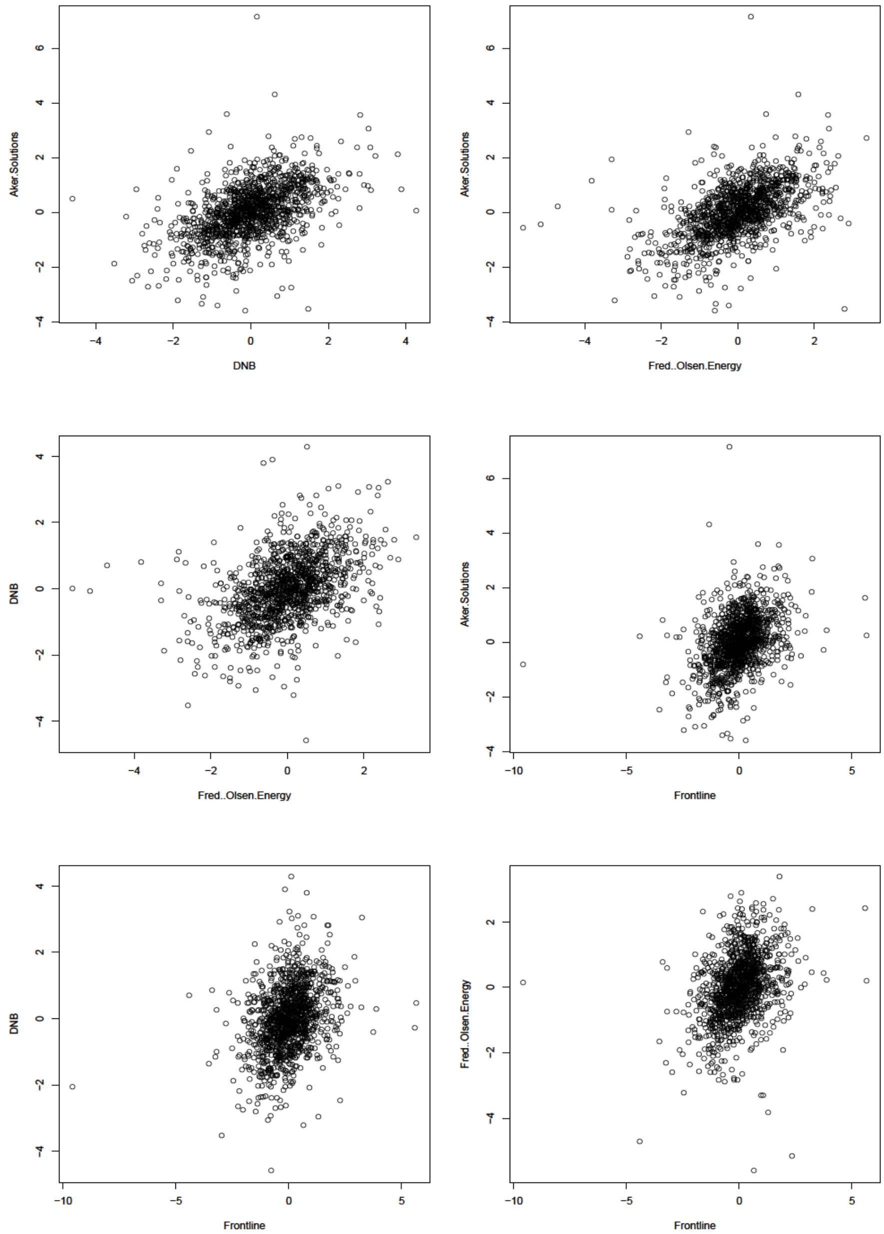


Figure 10.4: Scatterplot between GARCH-filtered returns for stocks (1/2)

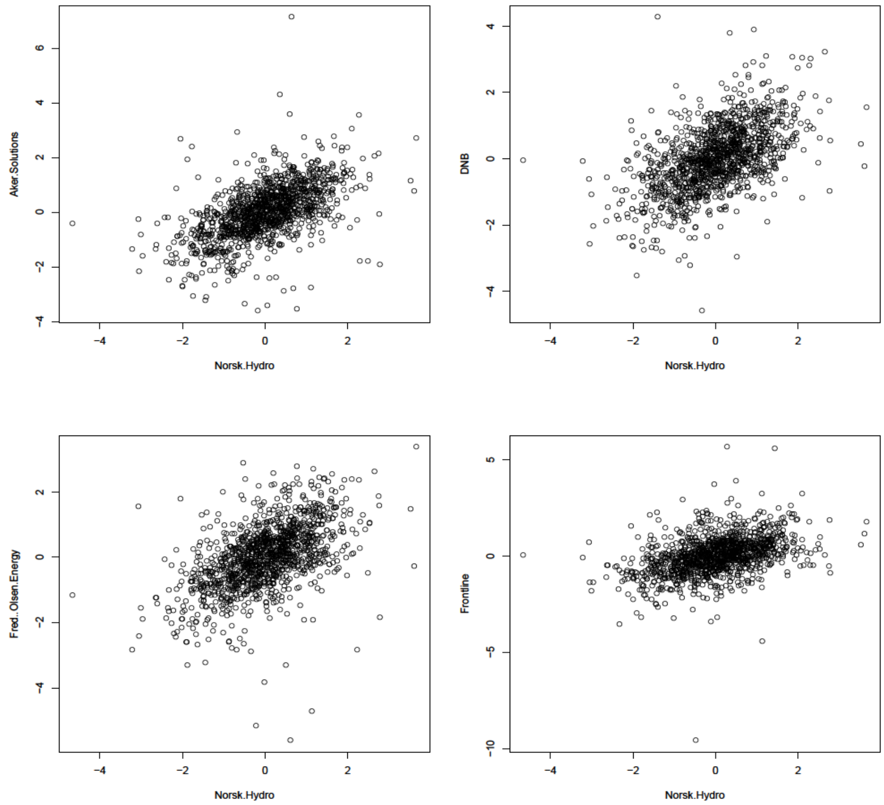


Figure 10.5: Scatterplot between GARCH-filtered returns for stocks (2/2)

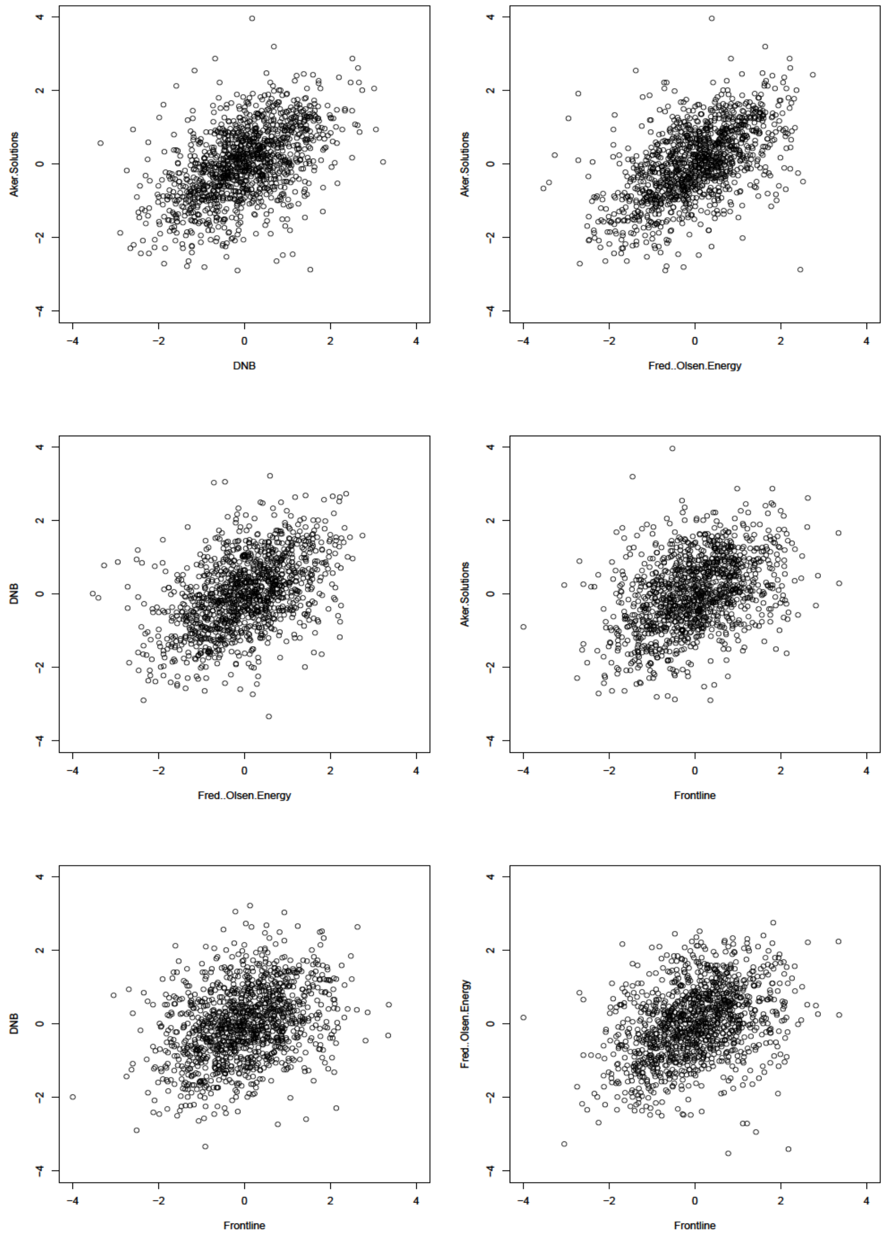


Figure 10.6: Scatterplot between marginal transformed returns for stocks (1/2)

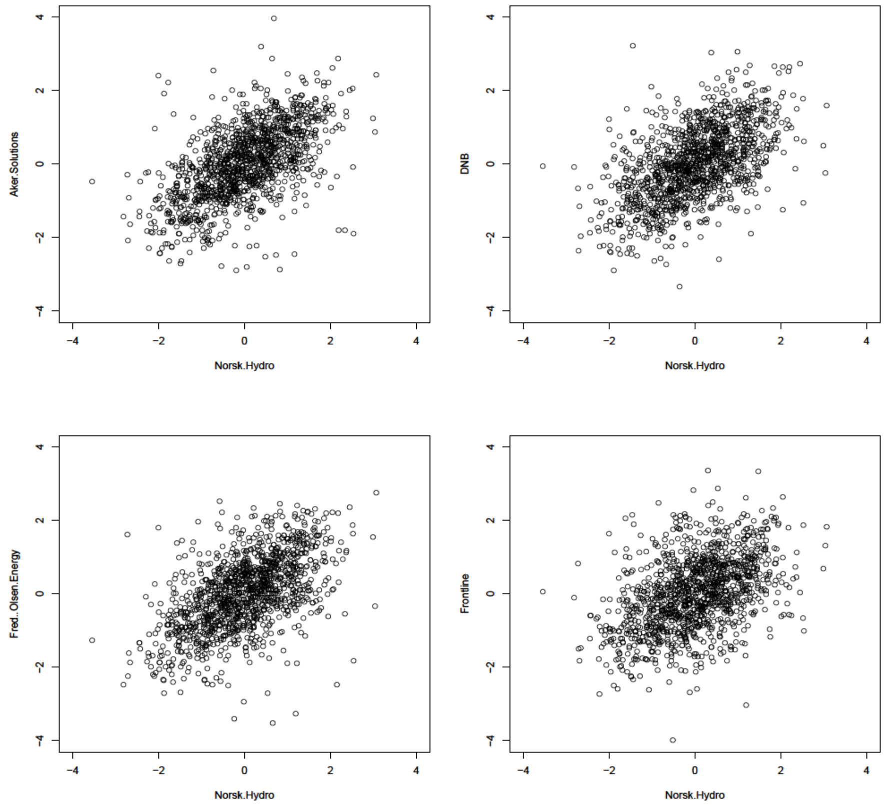


Figure 10.7: Scatterplot between marginal transformed returns for stocks (2/2)

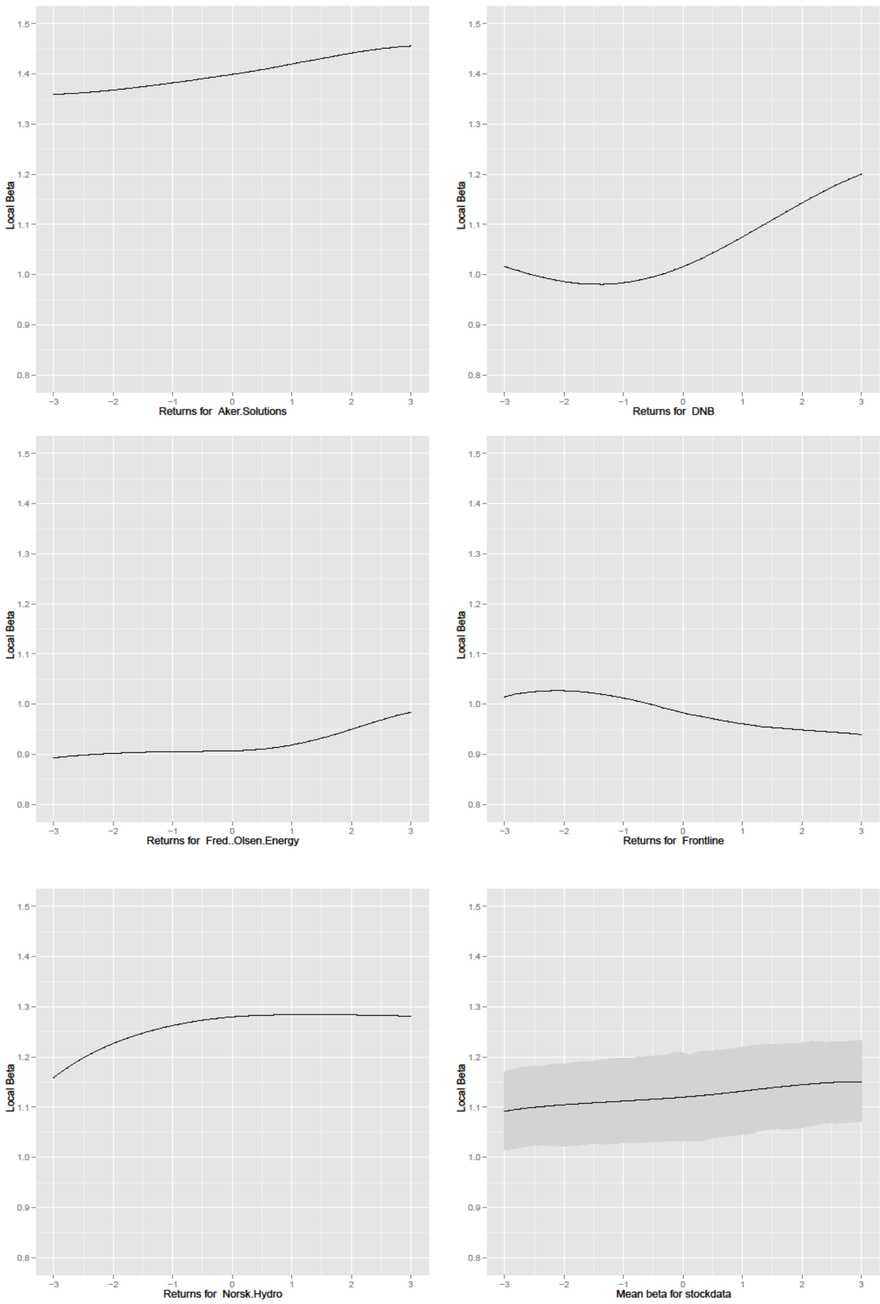
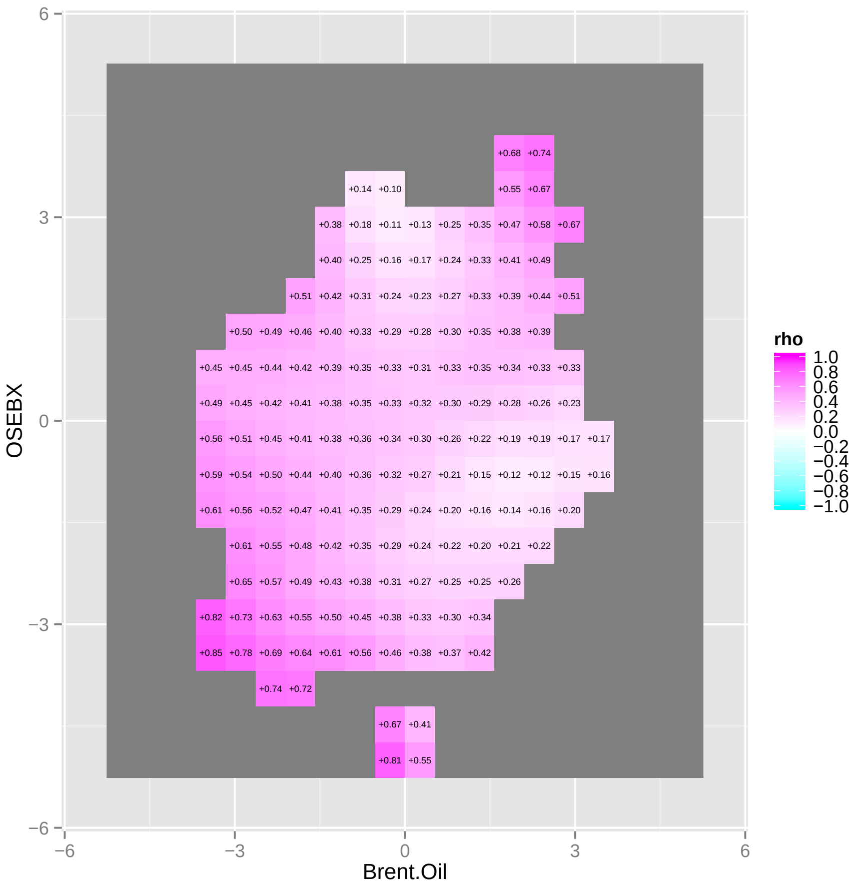


Figure 10.8: Change for the local beta on non GARCH-filtered stock returns for 5 stocks and mean calculated on 18 stocks.



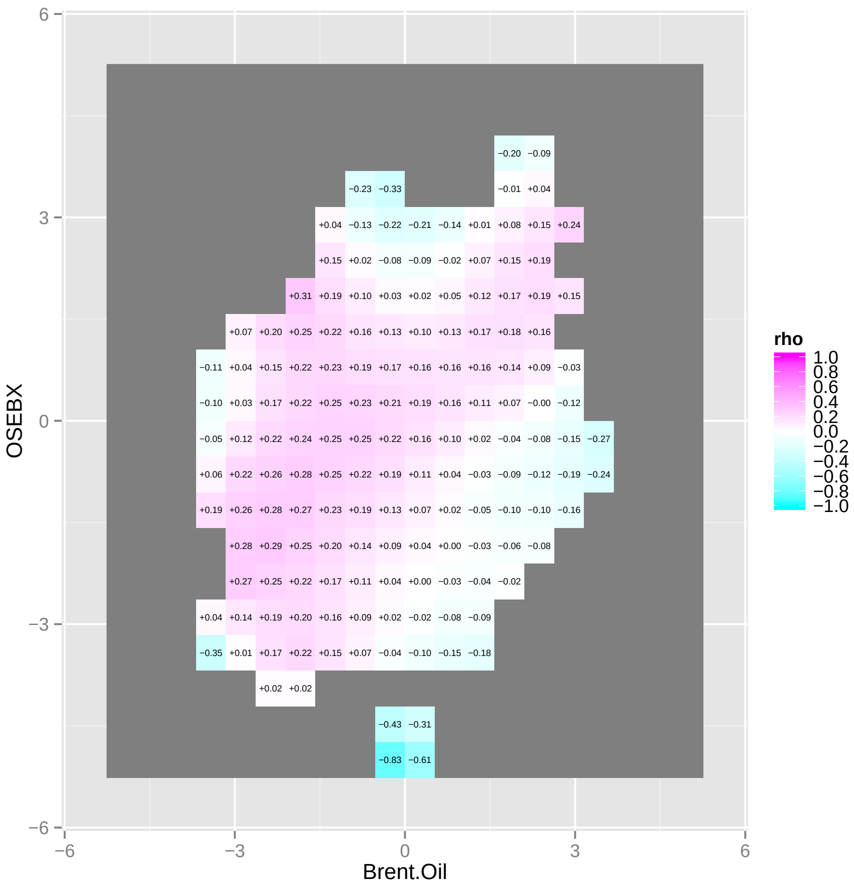


Figure 10.9: A 90% confidence interval for the Local Gaussian Correlation between the OSEBX index and the price of Crude Brent(USD)

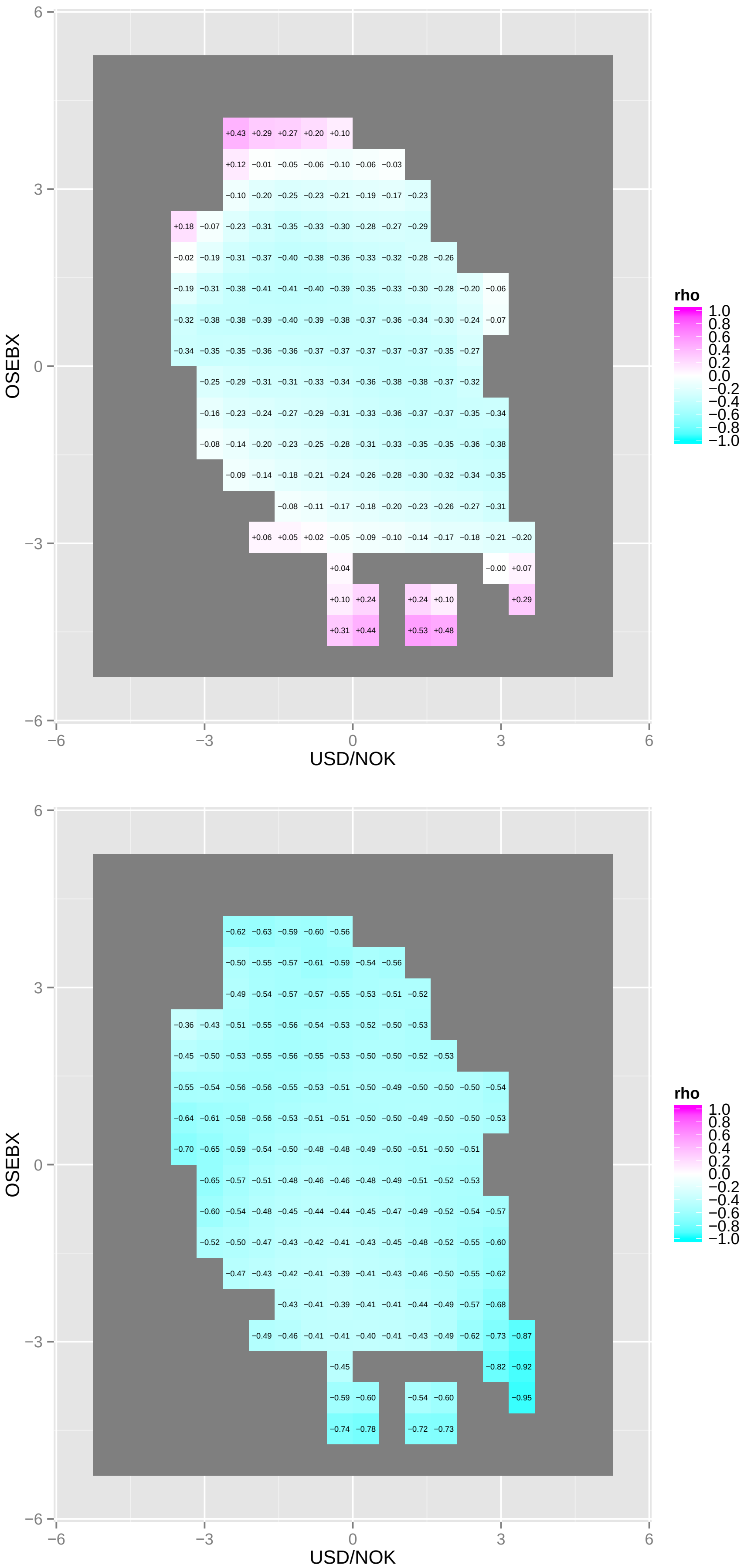


Figure 10.10: A 90% confidence interval for the Local Gaussian Correlation between the OSEBX index and the USD/NOK

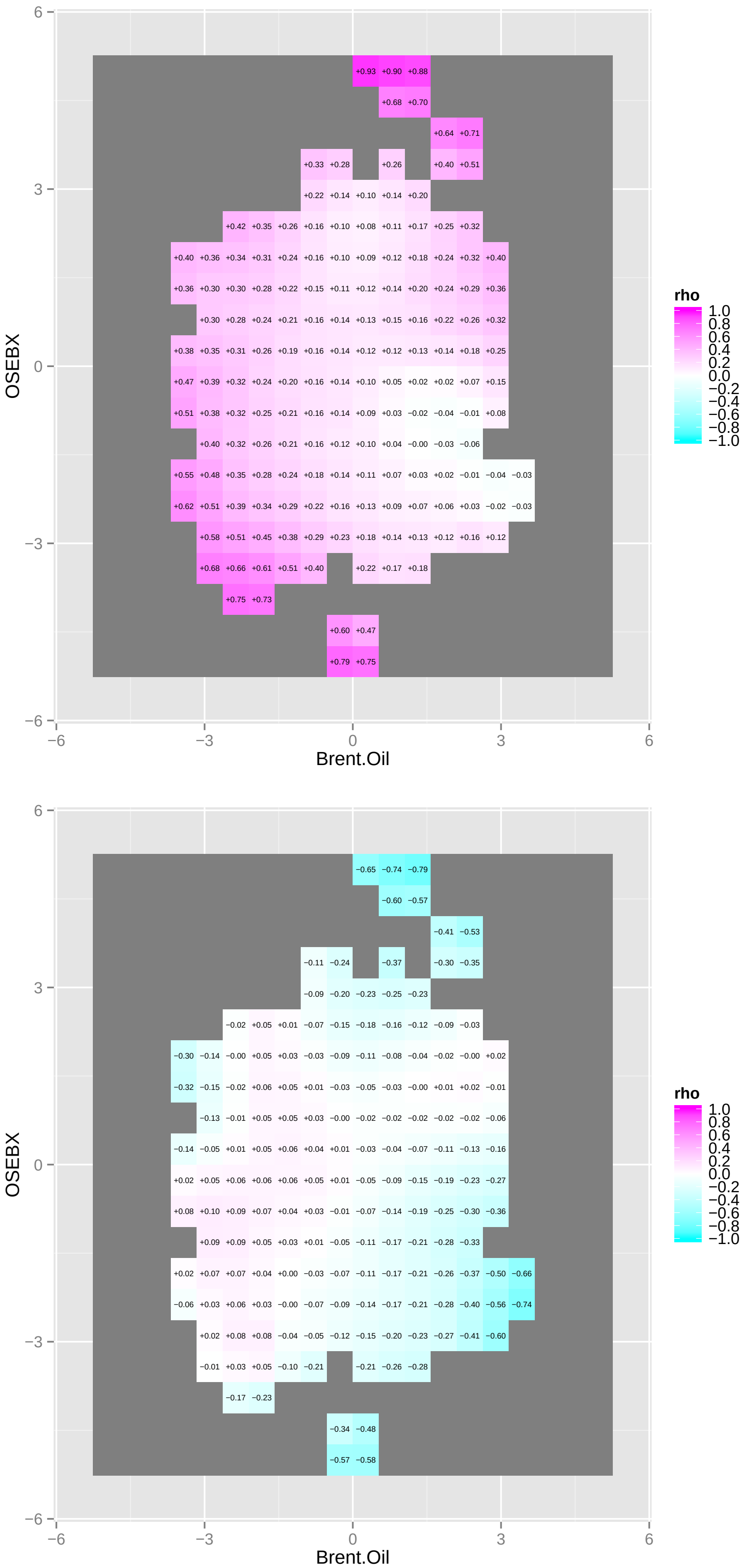


Figure 10.11: A 90% confidence interval for the Local Gaussian Correlation between the OSEBX index and the price of Crude Brent(NOK)

Bibliography

Aas, Kjersti, Haff, Ingrid Hobæk, & Dimakos, Xeni Kristine. 2005. Risk estimation using the multivariate normal inverse gaussian distribution. *Journal of Risk*, **8**(2), 39.

Aas, Kjersti, Czado, Claudia, Frigessi, Arnoldo, & Bakken, Henrik. 2009. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, **44**(2), 182–198.

Abramowitz, M Ed, *et al.* 1964. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Vol. 55. Courier Dover Publications.

Artzner, P., Delbaen, F., Eber, J.M., & Heath, D. 1999. Coherent measures of risk. *Mathematical finance*, **9**(3), 203–228.

Barndorff-Nielsen, O. 1978. Hyperbolic distributions and distributions on hyperbolae. *Scandinavian Journal of statistics*, 151–157.

Barndorff-Nielsen, Ole E. 1997. Processes of normal inverse Gaussian type. *Finance and stochastics*, **2**(1), 41–68.

Berentsen, G.D., Støve, B., Tjøstheim, D., & Nordbø, T. 2012. Recognizing and visualizing copulas: an approach using local Gaussian approximation. *Manuscript, Department of Mathematics, University of Bergen. See http://folk. uib. no/gbe062/local-gaussian-correlation*.

Berk, J., & DeMarzo, P. 2007. *Corporate finance*. Addison-Wesley.

Bjørnland, Hilde C. 2009. Oil price shocks and stock market booms in an oil exporting country. *Scottish Journal of Political Economy*, **56**(2), 232–254.

Black, Fischer. 1993. Estimating expected return. *Financial Analysts Journal*, 36–38.

Bølviken, E, & Benth, FRED ESPEN. 2000. Quantification of risk in Norwegian stocks via the normal inverse Gaussian distribution. *In: Proceedings of the AFIR 2000 Colloquium, Tromso, Norway*, vol. 8798.

Bollerslev, Tim. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, **31**(3), 307–327.

Boyer, Brian H, Gibson, Michael S, & Loretan, Mico. 1997. *Pitfalls in tests for changes in correlations*. Vol. 597. Board of Governors of the Federal Reserve System.

Casella, George, & Berger, Roger L. 2002. *Statistical inference*. Duxbury Pr.

DeCarlo, Lawrence T. 1997. On the meaning and use of kurtosis. *Psychological Methods*, **2**(3), 292–307.

Durante, F., & Sempi, C. 2010. Copula theory: an introduction. *Copula Theory and Its Applications*, 3–31.

Efron, B. 1979. Bootstrap methods: another look at the jackknife. *The annals of Statistics*, **7**(1), 1–26.

Engle, Robert F. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Fama, Eugene F, & French, Kenneth R. 1992. The cross-section of expected stock returns. *the Journal of Finance*, **47**(2), 427–465.

Feron, R. 1956. Sur les tableaux de correlation dont les marges sont donnees, cas de lespace a trois dimensions. *Publ. Inst. Statist. Univ. Paris*, **5**.

Francq, Christian, Horvath, Lajos, & Zakoian, Jean-Michel. 2009. *Merits and drawbacks of variance targeting in GARCH models*. MPRA Paper 15143. University Library of Munich, Germany.

Genest, C., & MacKay, R.J. 1986. Copules archimédiennes et families de lois bidimensionnelles dont les marges sont données. *Canadian Journal of Statistics*, **14**(2), 145–159.

Gjerde, Øystein, & Saettem, Frode. 1999. Causal relations among stock returns and macroeconomic variables in a small, open economy. *Journal of International Financial Markets, Institutions and Money*, **9**(1), 61–74.

Hjort, N.L., & Jones, MC. 1996. Locally parametric nonparametric density estimation. *The Annals of Statistics*, 1619–1647.

Hogg, R.V., & Tanis, E.A. 2005. *Probability and statistical inference*. Prentice Hall.

Hull, J. 2009. *Options, Futures and other Derivatives*. Pearson Prentice Hall.

Jarrow, Robert A. 1999. In honor of the Nobel Laureates Robert C. Merton and Myron S. Scholes: a partial differential equation that changed the world. *The Journal of Economic Perspectives*, **13**(4), 229–248.

Jones, S. 2009. The Formula that Felled Wall Street. *The Financial Times*, **22**(04).

Kunsch, Hans R. 1989. The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, **17**(3), 1217–1241.

Lee, Cheng-Few. 2006. *Encyclopedia of finance.* Springer.

Li, D.X. 1999. On default correlation: a copula function approach. *Available at SSRN 187289.*

Lura, Andreas. 2011. Risk management: Predicting the unpredictable. *Unpublished Bachelor Thesis(UIB).*

Mandelbrot, B. 1963. The Variation of Certain Speculative Prices. *The Journal of Business*, **36**(4), 394–419.

McNeil, A.J., Frey, R., & Embrechts, P. 2005. *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton Univ Pr.

Merton, Robert C. 1990. *Continuous Time Finance.* Cambridge, MA: Basil Blackwell.

Patton, Andrew J. 2004. On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics*, **2**(1), 130–168.

Pettengill, Glenn N, Sundaram, Sridhar, & Mathur, Ike. 1995. The conditional relation between beta and returns. *Journal of Financial and Quantitative Analysis*, **30**(01), 101–116.

Rizzo, M.L. 2007. *Statistical computing with R.* Chapman & Hall/CRC.

Roger, B.N. 2006. *An introduction to copulas.* Springer.

Romano, C. 2002. Calibrating and simulating copula functions: an application to the Italian stock market. *Risk Management Function, Capitalia, Viale U. Tupini*, **180**.

Rosenbaum, S. 1961. Moments of a truncated bivariate normal distribution. *Journal of the Royal Statistical Society. Series B (Methodological)*, 405–408.

Ruppert, D. 2010. *Statistics and data analysis for financial engineering.* Springer.

Silvapulle, Param, & Granger, Clive WJ. 2001. Large returns, conditional correlation and portfolio diversification: a value-at-risk approach. *Quantitative Finance*, **1**, 542–551.

Sklar, A. 1959. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, **8**(1), 11.

Sklar, A. 1996. Random variables, distribution functions, and copulas: A personal look backward and forward. *Lecture Notes-Monograph Series*, 1–14.

Støve, B., Tjøstheim, D., & Hufthammer, K. 2012. Using local Gaussian correlation in a nonlinear re-examination of financial contagion. *Revised manuscript.*

Tjøstheim, D., & Hufthammer, K.O. 2012. Local Gaussian correlation: A new measure of dependence. *Journal of Econometrics.*