

FORECASTING CONDITIONAL
CORRELATION FOR EXCHANGE RATES
USING MULTIVARIATE GARCH MODELS
WITH HISTORICAL VALUE-AT-RISK
APPLICATION

JOEL HARTMAN
DEPARTMENT OF ECONOMICS

&

JAN SEDLAK
DEPARTMENT OF STATISTICS

SUPERVISOR: LARS FORSBERG

BACHELOR THESIS
UPPSALA UNIVERSITY
SPRING 2013

Abstract

The generalization from the univariate volatility model into a multivariate approach opens up a variety of modeling possibilities. This study aims to examine the performance of the two multivariate GARCH models BEKK and DCC, applied on ten years exchange rates data. Estimations and forecasts of the covariance matrix are made for the EUR/SEK and USD/SEK, whereby the forecasts are used in a practical application: 1-day and 10-day ahead historical simulated Value-at-Risk predictions for two theoretical portfolios, one equally weighted and one hedged, consisting of the two exchange rates. An univariate GARCH(1,1) approach is included in the Value-at-Risk predictions to visualize the diversification effect in the portfolio. The conditional correlation forecasts are evaluated using three measures, OLS-regression, MAE and RMSE, based on an one year evaluation period of intraday data. The Value-at-Risk estimates are evaluated with the backtesting method introduced by Kupiec (1995). The results indicate that the BEKK model performs relatively better than the DCC model, and both these models perform better than the univariate GARCH(1,1) model.

Keywords: multivariate GARCH, exchange rates, conditional correlation, forecasting, Value-at-Risk

Contents

1	Introduction	1
1.1	The purpose of the study	2
1.2	Related research	2
1.3	Further outline	2
2	Economic background	3
2.1	The Exchange Rate Market	3
2.2	Exchange Rate Regimes	3
2.3	The Efficient Market Hypothesis and the market for exchange rates .	4
2.4	The Interest Rate Parity	4
2.5	Portfolio Diversification	5
3	Methodology and Statistical Theory	6
3.1	Data and Conditional Distributions	6
3.1.1	Descriptive statistics	6
3.1.2	The univariate sample distributions	10
3.1.3	The Multivariate Gaussian distribution	12
3.1.4	The Multivariate Student-t distribution	15
3.2	Theoretical Framework	15
3.2.1	Stochastic processes and Stationarity	15
3.2.2	Background of the univariate GARCH model	16
3.2.3	Multivariate GARCH models	16
3.2.4	The VEC model	18
3.2.5	The DVEC model	19
3.2.6	The BEKK model	19
3.2.7	The CCC model	20
3.2.8	The DCC model	21
3.3	The model order selection: AIC and BIC	21
3.4	Estimation Procedure: The Quasi Maximum Likelihood method . . .	22
3.5	Estimation evaluation	23
3.5.1	The Ljung-Box test	23
3.5.2	The ARCH LM test	24
3.5.3	The Baringhaus-Franz test	24
3.5.4	Test of dynamic correlations	25
3.6	Forecasting Procedure and Evaluation	26
3.6.1	The Procedure	26
3.6.2	Forecasting with the BEKK and DCC models	26

3.6.3	Defining a proxy	27
3.6.4	Ordinary Least Squares regression	27
3.6.5	Mean Absolute Error (MAE)	28
3.6.6	Root Mean Square Error (RMSE)	28
3.7	Practical application using the MGARCH forecasts: Historical Value-at-Risk	29
3.7.1	Background of Value-at-Risk	29
3.7.2	Backtesting the Value-at-Risk estimates	29
4	Estimations	31
4.1	BEKK(I,I)	31
4.2	DCC(I,I)	35
4.3	Comparison of the estimated models	38
4.3.1	Differences in the estimations	38
4.3.2	Residual analysis	38
4.3.3	Goodness of fit	39
4.4	Discussion of the estimations	39
5	Forecasts	42
6	Historical Value-at-Risk	45
6.1	Historical VaR Forecasts	45
6.2	Backtesting the VaR Forecasts	46
7	Conclusions	47
8	Further studies	48
	References	49
A	Appendix - Tables	51
B	Appendix - Figures	53

List of Tables

1	Properties of the return series	9
2	Jarque-Bera test for the return series	11
3	Result of Mardia's test	14
4	Order selection for BEKK	22
5	Order selection for DCC	22
6	Full BEKK estimation results	31
7	Diagonal BEKK estimation results	32
8	DCC estimation results	35
9	BEKK forecast evaluation	44
10	DCC forecast evaluation	44
11	1-day VaR estimates (in SEK) for the portfolio at given confidence levels	45
12	10-day VaR estimates (in SEK) for the portfolio at given confidence levels	45
13	ARCH LM test of return series	51
14	Ljung Box test for BEKK	51
15	Ljung Box test for DCC	51
16	Results of the Baringhaus-Franz test	51
17	ARCH LM test for BEKK	51
18	ARCH LM test for DCC	51
19	1-day Kupiec test results	52
20	10-day Kupiec test results	52
21	1-day Bernoulli trial results	52
22	10-day Bernoulli trial results	52

List of Figures

1	Exchange rates for EUR/SEK and USD/SEK	7
2	Logarithmic returns for EUR/SEK	8
3	Logarithmic returns for USD/SEK	8
4	Squared returns for EUR/SEK	9
5	Squared returns for USD/SEK	9
6	Histogram EUR/SEK	10
7	Histogram USD/SEK	10
8	The bivariate distribution	13
9	Estimated conditional correlations using BEKK	33
10	Estimated conditional covariance using BEKK	34
11	Estimated conditional variance for EUR/SEK using BEKK	34
12	Estimated conditional variance for USD/SEK using BEKK	35
13	Estimated correlations using the DCC model	36
14	Estimated conditional covariance using DCC	36
15	Estimated conditional variance for EUR/SEK using the DCC model	37
16	Estimated conditional variance for USD/SEK using the DCC model	37
17	BEKK correlation forecasts	42
18	DCC correlation forecasts	43
19	BEKK forecasts and Realized Correlation ($m = 48$)	43
20	DCC forecasts and Realized Correlation ($m = 48$)	44
21	Absolute returns for EUR/SEK	53
22	Absolute returns for USD/SEK	53
23	BEKK: Standardized residuals for EUR/SEK	54
24	BEKK: Standardized residuals for USD/SEK	54
25	BEKK: ACF for EUR/SEK	55
26	BEKK: ACF for USD/SEK	55
27	BEKK: XCF for the two series	55
28	DCC: Standardized residuals for EUR/SEK	56
29	DCC: Standardized residuals for USD/SEK	56
30	DCC: ACF for EUR/SEK	57
31	DCC: ACF for USD/SEK	57
32	DCC: XCF for the two series	57
33	Spread between the BEKK and DCC estimated correlations	58
34	The DCC conditional correlation for the year 2008	58
35	Key Interest Rates for the United States, Sweden and the Eurozone	59

36	Domestic bank CDS prices for the United States, Sweden and the Eurozone	59
37	Simulated 1-day portfolio value distribution for GARCH(1,1)	60
38	Simulated 1-day portfolio value distribution for BEKK(1,1)	60
39	Simulated 1-day portfolio value distribution for DCC	60
40	Simulated 1-day portfolio value distribution for GARCH, one long and one short position	61
41	Simulated 1-day portfolio value distribution for BEKK, one long and one short position	61
42	Simulated 1-day portfolio value distribution for DCC, one long and one short position	61
43	Simulated 10-day portfolio value distribution for GARCH(1,1)	62
44	Simulated 10-day portfolio value distribution for BEKK(1,1)	62
45	Simulated 10-day portfolio value distribution for DCC	62
46	Simulated 10-day portfolio value distribution for GARCH, one long and one short position	63
47	Simulated 10-day portfolio value distribution for BEKK, one long and one short position	63
48	Simulated 10-day portfolio value distribution for DCC, one long and one short position	63
49	Simulated 1-day portfolio critical value forecast at 1 % for BEKK . . .	64
50	Simulated 1-day portfolio critical value forecast at 1 % for DCC . . .	64
51	Simulated 1-day portfolio critical value forecast at 1 % for GARCH . .	64
52	Simulated 1-day portfolio critical value forecast at 5 % for BEKK . . .	65
53	Simulated 1-day portfolio critical value forecast at 5 % for DCC . . .	65
54	Simulated 1-day portfolio critical value forecast at 5 % for GARCH . .	65
55	Simulated 10-day portfolio critical value forecast at 1 % for BEKK . .	66
56	Simulated 10-day portfolio critical value forecast at 1 % for DCC . . .	66
57	Simulated 10-day portfolio critical value forecast at 1 % for GARCH .	66
58	Simulated 10-day portfolio critical value forecast at 5 % for BEKK . .	67
59	Simulated 10-day portfolio critical value forecast at 5 % for DCC . . .	67
60	Simulated 10-day portfolio critical value forecast at 5 % for GARCH .	67

I Introduction

It is a widely accepted fact that financial time series suffer from heteroscedasticity; the phenomena of volatility clustering. Some time periods are more volatile than others due to turbulence and unexpected events. The first technique to model this heteroscedasticity was developed by Engle in 1982 and named the Auto Regressive Conditional Heteroscedasticity (ARCH) model. The ARCH model was modified by Bollerslev in 1986 and renamed the Generalized ARCH (GARCH) model.¹

Using the aforementioned models enables researchers to estimate the volatility of financial time series. However, a major component of the literature about volatility models is focused on the univariate approach.

Although, univariate models are not sufficient when one wants to estimate correlation, because it is depending on the interaction between regions. Therefore, a model that simultaneously takes more than one time series into account would be more convenient in the procedure of estimating time-varying correlation.

This family of models is referred as multivariate GARCH (MGARCH) models, and during recent years a variety of models has been introduced. The first developed MGARCH models were the VEC and the BEKK. The multivariate estimation approach is, theoretically, a straightforward procedure. However, the implementation with empirical data is of a more complex nature, the number of parameters in the model get easily unmanageable since they increase rapidly. This issue has lead to that a variety of papers have proposed models that are aiming to be more parsimonious. (Sentana, 1998, p. 1)

The fluctuations of exchange rates is a discussed macroeconomic topic, and especially after the termination of the Bretton Woods system.² The debate of exchange rate volatility has concerned various subjects within the macroeconomic field, such as its impact on inflation and international trade. Another field of interest is risk management, where estimations and forecasting of exchange rates is an appealing subject, particularly due to recent innovations in the derivatives market such as volatility and correlation swaps³ and of course the last decades of turbulence on the financial markets. (Suliman & Suliman, 2001, p. 1)

¹The GARCH model is a special case of an infinite-order ARCH and has, in most cases, replaced the ARCH model in applications. (Teräsvirta, 2006, p. 5)

²The Bretton Woods was a monetary system for the major industrial countries, which pegged their currencies against the USD, which itself guaranteed a price in gold. Formally the system collapsed in 1973, although the USD had been floating since 1968.

³As an example, a correlation swap is when the payments of the swap's fixed and floating legs are based on the correlation of an underlying asset, for example exchange rates.

1.1 The purpose of the study

The purpose of this paper is to estimate and forecast time-varying correlation. The purpose is expressed in the following questions:

- To what extent can the multivariate GARCH approach be used to forecast time-varying correlations applied on the exchange rate market?
- To what extent can the forecasts be used in a historical Value-at-Risk simulation?

1.2 Related research

During the last decades, a body of research regarding exchange rate volatility has been presented. Many countries changed from a fixed currency regime to a floating system. Nevertheless, mainly those studies⁴ were performed in order to understand the connections between macroeconomic variables and the movements of exchange rates.

However, a large amount of research has appeared during recent years, that on the contrary has been aiming to study exchange rates in a time series context. A large proportion of this research has focused on a univariate approach for studying volatility clustering, volatility persistence and asymmetric effects⁵ in the returns. A well known research among econometricians is the study performed by Andersen & Bollerslev (1998), where they concluded that univariate standard volatility models in fact are able to perform accurate volatility forecasts. But, only a smaller fraction of the literature has been concentrated on a multivariate approach for conditional correlation forecasting, and the performance of these models.

1.3 Further outline

The next section presents the economic background. Thereafter the chosen approach and the selection of data is presented. This section includes the statistical theory that describes some basic concepts and then gives a description of the used models. Section four presents the results of the estimations and model diagnostics. This part includes a comparison of the models and ends with a discussion of the results. Section five shows the estimated models forecasting performance. The seventh section is the evaluation of the Value-at-Risk forecasts. Section seven contains the final conclusions. Thereafter, the authors give suggestions for further research within the field. This is followed by the list of used references. Finally, the appendixes are attached in the end.

⁴Suliman & Suliman (2011, p. 1) list many examples.

⁵Meaning that downward fluctuations are followed by higher volatility compared to upward movements.

2 Economic background

This section aims to briefly discuss some fundamental concepts of what the exchange rates market is, different systems and also describe a basic model for how the exchange rate between two countries is determined.

2.1 The Exchange Rate Market

The market for exchange rates is one of the most liquid. Trades are performed through market makers, which is the reason for the bid-ask spread since they are buying and selling at different levels. The exchange market does not close in the same way as, for example, the stock market, which simplifies the volatility modeling, since one does not need to take the overnight variance into account. This means the variance that occurs (for instance for stocks) due to that the price changes during nights or weekends.

A exchange rate pair is the quotation of the value of one currency unit against another. For instance, if a EUR/USD transaction is traded at 1.2500, that simply means that 1.0000 Euro is bought for 1.2500 USD.

There is a difference between spot and forward exchange rate, the first one is the current exchange rate and the second one is a rate that is quoted and traded today, however the actual delivery and payment is not today but on a specific date in the future.

2.2 Exchange Rate Regimes

The categorization of exchange rate regimes is based on the flexibility of the system. From a traditional point of view, the regimes are divided into two types, fixed or flexible. However, the currency of a country may be floating, pegged, fixed or a hybrid.

A floating currency is allowed to vary and the market determines the value of the currency in accordance with the forces of supply and demand. This means that the quotes of that currency more quickly adjust. The monetary policy is often performed without further considerations regarding the exchange rate. In managed floating systems the authorities sometimes take actions if, for example, the market exhibits a period of high volatility. (Suliman & Suliman, 2011, p. 218)

A fixed exchange rate means that the home currency is fixed against another currency or against a basket of other currencies. It can also be fixed, e.g. against gold.

A pegged currency is fixed, but also allowed to devalue. For instance, the Chinese RMB was pegged to the USD until 2005 and the Danish currency is currently pegged

against the Euro as a part of ERM II.⁶ Still, some governments keep their currency within a narrow range. As a result, currencies are becoming over-valued or under-valued, causing trade deficits or surpluses.

2.3 The Efficient Market Hypothesis and the market for exchange rates

According to Fama (1970), the Efficient Market Hypothesis (EMH) implies that markets are efficient in the way that a price of a certain asset includes all available information. This indicates that the exchange rate between two currencies is based on all known information, resulting in that the price is the right one. The result of this is that one cannot continuously earn profits in the market using the public available information (Fama, 1970).⁷ Additionally, the theory states that the rates adjust to new information immediately, that prevents potential arbitrages. Interestingly enough, exchange rates do not adjust immediately, but are lagged according to research of Eichenbaum & Evans (1995). Also, findings of Neely (1997) indicate that the EMH fails in describing exchange rates.

In the year 1980, Grossman & Stiglitz studied markets with respect to whether they can be considered as efficient or not. The findings were, not surprisingly, that they are not. This can be summarized in the paradox that the theory indicates that no predictions on future movements are possible to perform on a market that is efficient. But since exchange rates consist of more or less all available information regarding macroeconomics, i.e. interest rates, debt levels of countries, inflation and so on, a trading based on a in depth analysis is both time consuming and expensive.

The fact that EMH may not be applicable to the exchange rate market makes volatility modeling and forecasting more appealing.

2.4 The Interest Rate Parity

A diversity of macroeconomic variables are affecting the exchange rate. One fundamental relationship that provides a basic understanding is the **Interest Rate Parity**, which is an equilibrium under which no arbitrage is possible. However two assumptions exist, first capital mobility must prevail and second there must be perfect

⁶Within the ERM II system, a currency can float within a range of plus/minus 15 % with respect to a central rate against the euro. Using the Danish currency as an example, the exchange rate is kept with a narrower range that equals plus/minus 2.25 %.

⁷The theory of EHM is further discussed by Fama (1970).

sustainability of assets between the countries. The basic idea is illustrated with the **Uncovered Interest Rate Parity**⁸ in the following equation

$$S_t = \frac{(1 + i_a)}{(1 + i_b)} \times E_t(S_{t+k}) \quad (1)$$

The spot exchange rate is S_t at time t , and i_a and i_b are the nominal interest rates for two countries. The term $E_t(S_{t+k})$ is the k periods ahead expected spot exchange rate between the two countries. The major feature of this equilibrium is its basic way of explaining how the domestic and foreign key interest rates affect the exchange rate between the countries.

2.5 Portfolio Diversification

Ever since Markowitz (1952) introduced the concept of portfolio diversification a body of articles have been published in this field. A short description is included since the concept is later used when the VaR forecasts are made. Given a portfolio of asset A and asset B , their corresponding expected returns are $E[r_A]$ and $E[r_B]$. The variances are σ_A^2 and σ_B^2 . The weight of asset A is set to be ω , the weight in B is then $1 - \omega$, where the sum of the weights by definition equals one. If $\omega < 0$, this indicates a short position in asset A , and if $\omega > 1$ this means that the portfolio has a short position in asset B . The portfolio return is the weighted average of the returns for A and B , but the portfolio variance equals

$$\sigma_p^2 = \omega^2 \sigma_A^2 + (1 - \omega)^2 \sigma_B^2 + 2\omega(1 - \omega) \sigma_A \sigma_B \rho_{A,B} \quad (2)$$

where $\rho_{A,B}$ is the correlation between the two assets. As can be seen, the variance of the portfolio depends on the interaction between the assets. When there are no correlation between the assets, the portfolio variance is a weighted average of the assets individual variances.

To summarize, one can expect different outcomes depending on the portfolios composition, and whether you decide to include the diversification effect or not.

⁸In the **Covered Interest Rate Parity**, the expected future spot exchange rate $E_t(S_{t+k})$ is replaced by the forward exchange rate at time t since forward contracts are available for investors. The Covered Interest Rate Parity is one way of explaining the current forward exchange rate.

3 Methodology and Statistical Theory

This chapter first describes the data, and then the approach and procedure are being explained. Thereafter, it describes first the univariate GARCH model, whereafter the multivariate approach and the common models are described more profoundly. Then the theoretical procedure of performing the forecasts is explained. In addition, the chosen tests are presented and the techniques that are used to perform the estimations, forecasts and evaluations.

3.1 Data and Conditional Distributions

3.1.1 Descriptive statistics

The analyzes and estimations have been performed with **MATLAB**. In this study the MATLAB toolbox **GARCH** created by Sheppard has been used in the estimations and most of the tests, but a substantial part of the code has been written by the authors.

The series used for the first estimation is for a ten year period, with start date 2000-01-01 and end date 2010-01-01. The second period equals the next year, from 2010-01-01 to 2011-01-01. This second period consist of hourly intraday data and is used to evaluate the conditional correlation forecasts of the series.

Two exchange rate series have been selected, the **EUR/SEK** and the **USD/SEK**. The daily closing price of the spot prices has been taken for each day for the sample period. A contradicting aspect of the choice of exchange rates for estimating time-varying conditional correlations is their common factor, Sweden. Domestic events as for example changes in macroeconomic data or key interest rate cuts obviously have impact on both the USD/SEK and EUR/SEK. Nonetheless, there is still unclear how the conditional correlation changes over time and how well the models suit the data, and the accuracy of the forecasts. **Figure 1** displays the exchange rates movement during the estimation period:

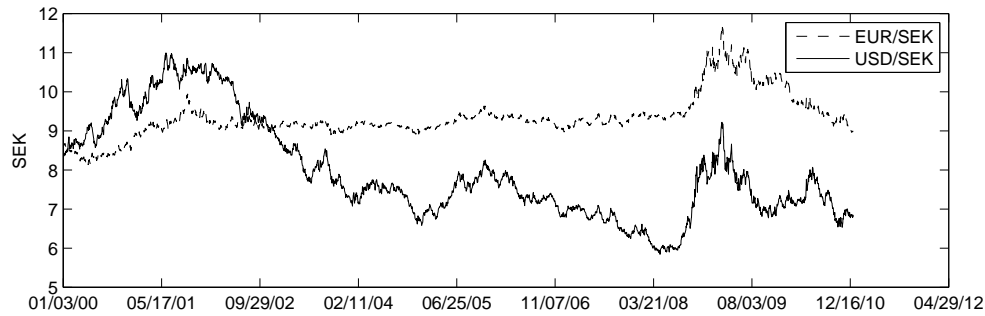


Figure 1: Exchange rates for EUR/SEK and USD/SEK

The first impression is that it seems that the USD/SEK series moves more over time. The EUR/SEK is at a similar exchange rate level in the start and the end of the sample, while the USD/SEK is at a lower level in the end of the sample. They both move in a similar pattern during 2008 where the SEK is depreciating, although the USD/SEK has been in a downward movement for a long period.

The logarithmic daily returns⁹ have been calculated based on the daily closing prices, and these are shown in **Figure 2** to **3**:

⁹The logarithmic returns are calculated in this paper, in accordance with other academic literature, simply due to the main advantage that continuously compounded returns are symmetric, while arithmetic returns are asymmetric. For example, imagine an investment that initially is valued to 100 euro. An increase in price of 50 % results in a value of 150 euro, but a later decrease of 50 % will lead to a current value of the investment equal to 75 euro.

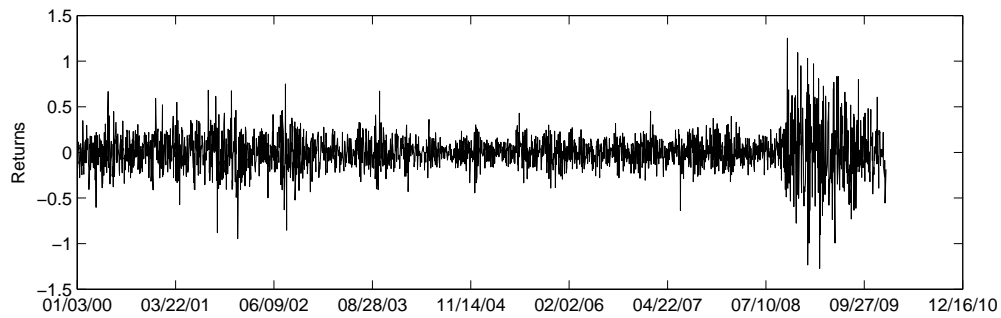


Figure 2: Logarithmic returns for EUR/SEK

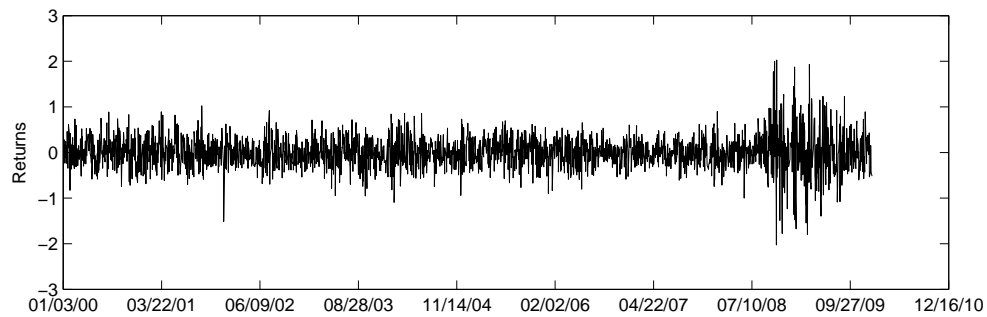


Figure 3: Logarithmic returns for USD/SEK

The EUR/SEK seems, compared to the USD/SEK, to be more volatile in the beginning of the sample period, especially the first half of the year of 2002. Besides that, the last period is highly volatile for both series.

Furthermore, in **Table 1** the descriptive statistics of the returns are presented:

	Mean	Median	Max	Min	Un. var.	Skewness	Kurtosis
EUR/SEK	0,0030	0,0023	1,2517	-1,2711	0,0402	0,0473	7,9432
USD/SEK	0,0032	-0,0135	2,0270	-2,0251	0,1261	0,1474	6,4931

Table 1: Properties of the return series

It is seen that the means are close to each other, even though the median of USD/SEK is negative. This means that there are more negative observations, but the positive observations are relatively higher. The USD/SEK series also exhibit a higher unconditional variance. Another meaningful aspect of the return series is the unconditional correlation, which equals 0.5584 for the chosen sample period. The graphs in **Figure 4** and **5** consist of the squared returns:

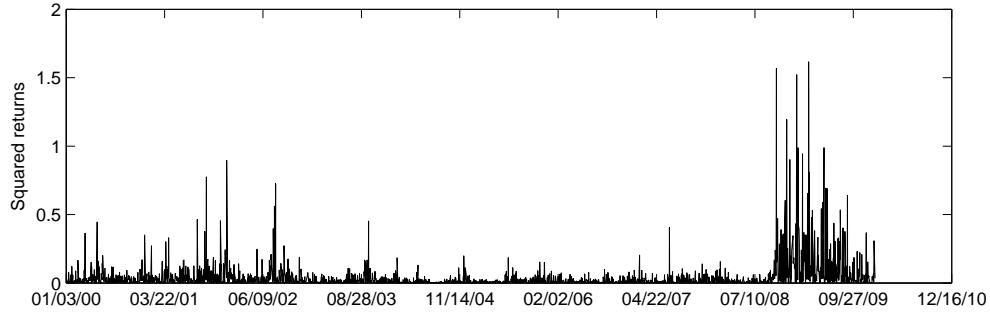


Figure 4: Squared returns for EUR/SEK

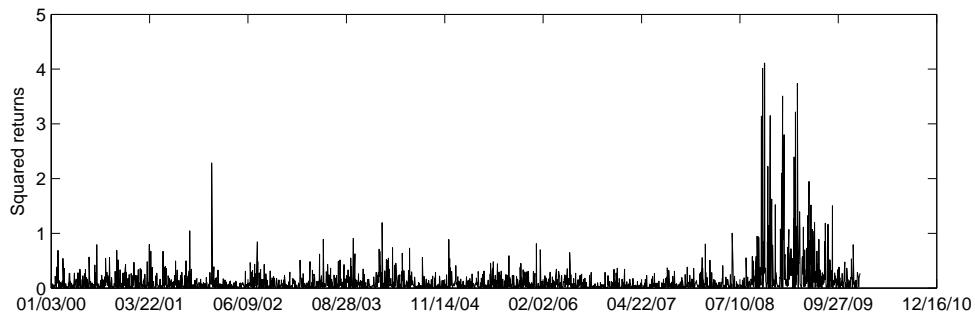


Figure 5: Squared returns for USD/SEK

Both the squared returns graphs have a similar pattern, the last part of the estimation period is a textbook example of volatility clustering. The absolute returns, see **Figure 21** and **22** in **Appendix B**, also indicate volatility clustering. The beginning of the period also shows tendencies of a period with higher volatility for the EUR/SEK, whilst the USD/SEK only has a spike.

Moreover, as **Table 1** visualizes, both the series have a high kurtosis and are positively skewed. These properties are examined in the next section.

3.1.2 The univariate sample distributions

Looking at the distributions of the univariate series (with the theoretical univariate normal distribution marked with a solid line), in **Figure 6** and **7**, they clearly seem to be characterized by leptokurtic distributions.

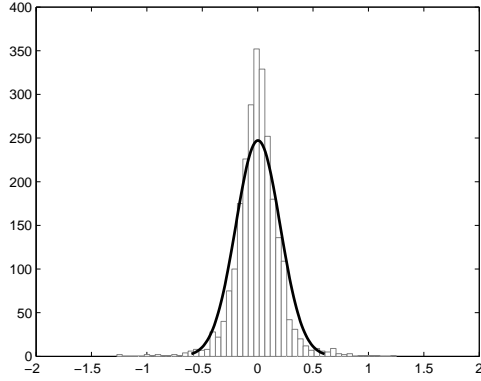


Figure 6: Histogram EUR/SEK

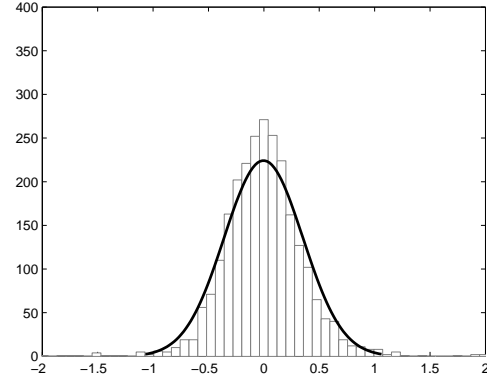


Figure 7: Histogram USD/SEK

However, to confirm that the distributions are non-normal, the Jarque-Bera test is performed. More specifically, the test determines if the sample has an excess kurtosis and a skewness equal to zero or not. This test statistic is defined as

$$JB = \frac{n}{6} \left[S^2 + \frac{(K - 3)^2}{4} \right] \quad (3)$$

where K is the kurtosis of the sample and given by $K = \frac{1}{n} \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{(\hat{\sigma}^2)^2}$ and the skewness represented by S and expressed as $S = \frac{1}{n} \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{(\hat{\sigma}^2)^{3/2}}$. Additionally, $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ where x_i is an individual observation and n is the number of observations.

In **Equation 3** the number of observations is denoted as n , the sample skewness is S and K is the kurtosis of this sample. Moreover, the excess kurtosis equals $EK = K - 3$ where the subtraction is made since three is the kurtosis of a normal distribution. The test hypotheses are defines as

$$H_0 : S = EK = 0$$

$$H_a : S \neq 0 \text{ or } EK \neq 0$$

The null hypothesis is that the distribution is normal, and if it is false there is an indication of a non-normal distribution. The test static JB can be compared with a chi-square distribution with two degrees of freedom, and is rejected if the observed value exceeds the critical value given by the distribution of χ^2 with two degrees of freedom. In **Table 2**, the received χ^2 and p -values from the Jarque-Bera tests are displayed:

Serie	χ^2	p -value
EUR/SEK	2557	0,0000
USD/SEK	1286	0,0000

Table 2: Jarque-Bera test for the return series

The distributions by themselves are according to the test non-normal since H_0 is rejected in both cases.

To summarize the analysis of the univariate descriptive statistics, it is possible to state that the series are characterized by:

- **Volatility clustering:** One of the most known feature that is characterizing financial time series. High returns simply tend to be followed by high returns and low returns by low returns. This clustering tendencies were visualized by the squared returns in **Figure 4** and **5**, but also by the absolute returns in **Figure 21** and **22** attached in **Appendix B**.
- **Leptokurtic and skewed distributions:** As the distribution plots and the descriptive statistics are showing clear signs of high kurtosis, the series are said to be leptokurtic. Additionally, they are positively skewed. These properties could be seen in **Figure 6** and **7** and be confirmed by the results of the Jarque-Bera test.

These results are expected, and the conclusion is that the univariate time series clearly exhibits typical financial time series features.

But modeling with a multivariate approach, an assumption about the unconditional covariance matrix is needed. Next section will describe the **multivariate Gaussian distribution**, and test whether the multivariate distribution of the EUR/SEK and the USD/SEK is multivariate normal or not.

3.1.3 The Multivariate Gaussian distribution

The multivariate Gaussian distribution is a generalization of the univariate normal distribution. Let first

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \in \mathbb{R}^N \text{ and } E(x) = \begin{bmatrix} E(x_1) \\ \vdots \\ E(x_N) \end{bmatrix} \in \mathbb{R}^N$$

then it follows that $Cov(x) = E((x - \mu)(x - \mu)')$. If $x \sim N(\mu, \xi)$ holds, then

$$f_x(x_1, \dots, x_N) = \frac{1}{(2\pi)^{N/2} |\xi|^{1/2}} \exp\left(-\frac{1}{2} (x - \mu)' \xi^{-1} (x - \mu)\right) \quad (4)$$

where $\mu \in \mathbb{R}^N$, $|\xi|$ is the determinant of ξ and $\xi \in \mathbb{R}^{N \times N}$ is a symmetric positive definite matrix. Since this paper utilizes a bivariate¹⁰ volatility model the probability density function simplifies into

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_{x_1}\sigma_{x_2}(1-\rho^2)^{1/2}} e^{\left(-\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_{x_1})^2}{\sigma_{x_1}^2} + \frac{(x-\mu_{x_2})^2}{\sigma_{x_2}^2} - \frac{2\rho(x_1-\mu_{x_1})(x_2-\mu_{x_2})}{\sigma_{x_1}\sigma_{x_2}} \right] \right)} \quad (5)$$

where σ_{x_i} for x_1 and x_2 are strictly positive and ρ is the correlation coefficient. The multivariate central limit theorem states that when the sample size increases, many multivariate statistics converge into a multivariate normal distribution. The multivariate histogram for the EUR/SEK and USD/SEK is:

¹⁰A two dimensional model is by definition bivariate.

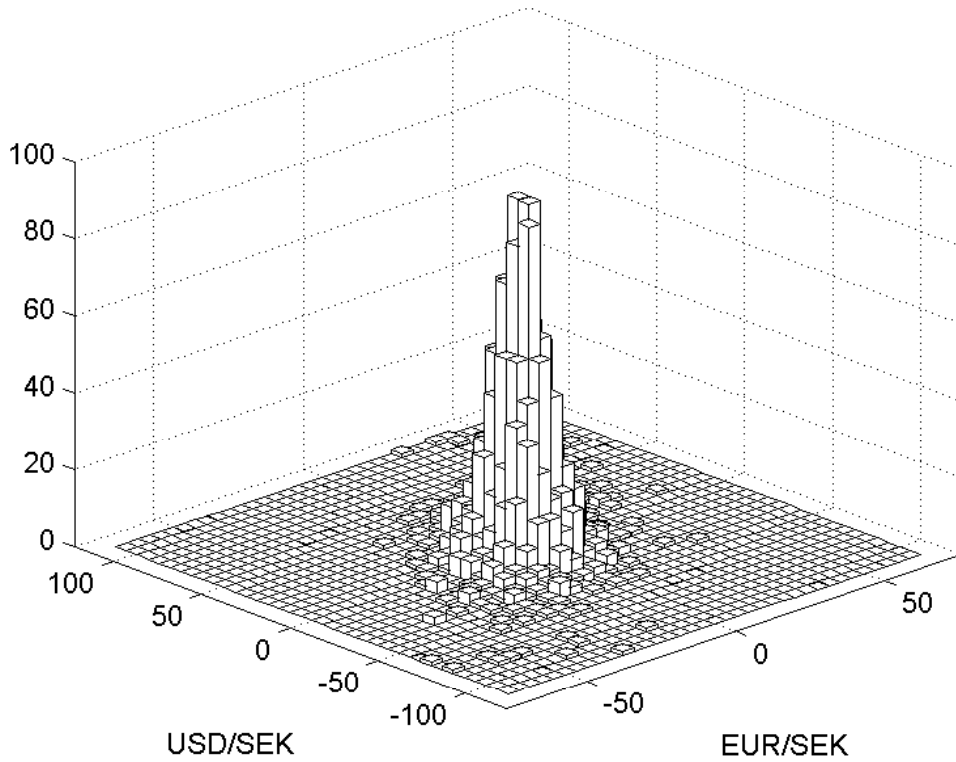


Figure 8: The bivariate distribution

However, in order to determine whether the used samples together follow a multivariate normal distribution or not, the **Henze-Zirkler's Multivariate Normality Test** is performed.¹¹ If the multivariate sample is normally distributed, then the test statistic is approximately lognormal distributed. The H_0 is that the data are multivariate normally distributed, and H_a is that the data are not. Let d be the dimension of the random sample x_i and n the number of observations. The test statistic $T_u(d)$ is then computed as

¹¹This test procedure was introduced 1990, and has been shown to have a good overall power compared to similar tests.

$$T_u(d) = \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n \exp\left(-\frac{u^2}{2} |Y_j - Y_k|^2\right) - 2(1+u^2)^{-\frac{d}{2}} \frac{1}{n} \sum_{j=1}^n \exp\left(-\frac{u^2}{2(1+u^2)} |Y_j|^2\right) + (1+2u^2)^{-\frac{d}{2}} \quad (6)$$

where the parameter u depends on the sample size n as $u_d(n) = \frac{1}{\sqrt{2}} \left[\frac{2d+1}{4}\right]^{\frac{1}{(d+4)}} n^{\frac{1}{(d+4)}}$, the sample covariance matrix is denoted ϑ and $|Y_j - Y_k|^2 = (x_j - x_k)' \vartheta^{-1} (x_j - x_k)$. Testing the multivariate distribution of EUR/SEK and USD/SEK, the result confirms that it is non-normal. The received value of $T_u(d)$ equals 18,0166, and $p \leq 0,0000$, and therefore the null hypothesis of normality is rejected.

Since there are several available tests for multivariate distributions, **Mardia's Multivariate Normality Test** is also used to verify that the achieved results from the test of Henze & Zirkler are correct. In Mardia's test, the multivariate skewness is asymptotically distributed as a χ^2 random variable for large samples. The hypotheses are the same as in the Henze-Zirkler test, H_0 is that the data are multivariate normally distributed and H_a is the data are not. Mardia's skewness statistic is defined as

$$g_{1,d} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n [(x_i - \bar{x})' \vartheta^{-1} (x_j - \bar{x})]^3 \quad (7)$$

and the kurtosis statistic is

$$g_{2,d} = \frac{1}{n} \sum_{i=1}^n [(x_i - \bar{x})' \vartheta^{-1} (x_i - \bar{x})]^2 \quad (8)$$

where ϑ is the covariance matrix and $g_{1,d}, g_{2,d} \sim \chi^2(d(d+1)(d+2)/6)$.¹² The multivariate normality test of Mardia is displayed in **Table 3**:

M_S	P_S	M_K	P_K
16,2393	0,0027	54,5166	0,0000

Table 3: Result of Mardia's test

¹²Let x_i denote the sample vector with mean \bar{x} and ϑ to be the covariance matrix, then $(x_i - \bar{x})' \vartheta^{-1} (x_i - \bar{x})$ is defined as the **Mahalanobis distance** which with a intuitive approach can be seen as the normalized distance of an observation to its mean.

In **Table 3** M_S is the test statistic for the multivariate skewness and P_S is the corresponding p -value. M_K is the multivariate kurtosis test statistic and P_K its p -value. As can be seen, neither the skewness or the kurtosis seem to be normal. As both multivariate normality test reject normality, the conditional distribution will be assumed to follow a **Multivariate Student-t distribution**.

3.1.4 The Multivariate Student-t distribution

There exist a variety of different generalizations of the univariate Student-t distribution into the multivariate case. This paper will use the most simple one recommended by Orskaug (2009), where the joint density of the standardized errors z_t is

$$f(z_t|\nu) = \prod_{t=1}^T \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) [\pi(\nu-2)]^{n/2} |\xi_t|^{1/2}} \left[1 + \frac{z_t^T z_t}{\nu-2}\right]^{-\frac{n+\nu}{2}} \quad (9)$$

where ν represents the degrees of freedom.¹³ The optimization procedure in the estimation is dependent on the chosen conditional distribution. Therefore, the next section describes the way the optimization is done.

3.2 Theoretical Framework

3.2.1 Stochastic processes and Stationarity

The definition of a time series¹⁴ is that the data are an ordered collection of observations throughout a definite length of time. In other words the sample consists of observations such as (x_0, x_1, \dots, x_t) and that $(t < \infty)$. It is established to call this collection of observations a realization of the underlying stochastic process. A time series that is strictly stationary have a joint distribution that does not vary when a shift in time or space is made, i.e. the joint distribution of $(X_{t_1}, \dots, X_{t_n})$ has to be equal to the joint distribution of $(X_{t_1-k}, \dots, X_{t_n-k})$ for all values of t and k , where k is the shift in time. If this is true, then the distributions are identical, which implies that the moments are the same. Furthermore, it is notable that under strict stationary the mean function of the time series is constant over time, due to the fact that the distributions are identical for X_t and X_{t-k} for all possible values of t and k . However, strictly stationary time series are seldom observed. Another definition of stationary exist though:

¹³In **Equation 9** the Gamma function $\Gamma(\cdot)$ is used. For $\forall a \in \mathbb{Z}$ then $\Gamma(a) = \Gamma(a-1)!$ holds.

¹⁴There exist a body of literature that well describes the basic concepts of time series, see for example *Time series analysis: with applications in R* by Chan & Cryer (2008), *Time Series: Data Analysis and Theory* by Brillinger (2001) or *Introduction to Time Series Modeling* by Kitagawa (2010).

covariance stationary. A stochastic process X_t fulfills the requirements of being covariance stationary if the mean function can be observed to be constant over time, but additionally that the covariance only depends on the time lag k . In other words, for a process to be covariance the first two moments of the process cannot vary but must be constant. (Brockwell & Davis, 2009, p. 1ff)

3.2.2 Background of the univariate GARCH model

First of all, let the returns r_{t-j} be expressed as the change in logarithmic spot price over a certain period

$$r_{t-k} = \ln \left(\frac{s_t}{s_{t-k}} \right) \quad (10)$$

where s_t is the spot price at time t and s_{t-k} the spot price at time $t-k$. One of the definitely most common and widely known univariate volatility models is the GARCH(p, q), which is defined as

$$\underbrace{h_t^2}_{\text{Conditional variance}} = \underbrace{c}_{\text{Intercept}} + \underbrace{\sum_{i=1}^q a_i \epsilon_{t-i}^2}_{\text{ARCH terms}} + \underbrace{\sum_{j=1}^p b_j h_{t-j}^2}_{\text{GARCH terms}} \quad (11)$$

The volatility term h_{t-j}^2 denotes the variance and j represents the number of lags.¹⁵ The term ϵ_{t-j}^2 is the squared error for the period $t-j$. In this equation, both the intercept c and the residual coefficient a_j has to be larger than zero in order to ensure a positive variance. (Reider, 2009, p. 4). The a -coefficient of lagged squared returns is interpreted as how fast the model react to, for example, market events. The b -coefficient of lagged conditional variance determines the degree of persistence in the volatility. A large value of a indicates that the conditional variance decays slowly, and that the volatility is persistent. On the other hand, if the a -value is relatively higher than the b -value, then the volatility is more extreme. The sum of a and b is expected to equal a value close to one.

Nevertheless, one region's volatility has impact on volatility of other regions. Making estimations that include these contagion effects require an extended model.

3.2.3 Multivariate GARCH models

A large part of the literature deals with univariate models. But markets interact, and therefore a generalization from the univariate model to a multivariate one is needed.

¹⁵In other words, the number of previous periods with effect on the estimated volatility.

MGARCH models can be categorized¹⁶ into four types:

- **Models of the conditional covariance matrix:** The conditional covariance is computed in a direct way. For example the VEC and BEKK models.
- **Factor models:** The return process is assumed to consist of a small number of unobservable heteroscedastic factors. This approach benefits from that the dimensionality of the problem reduces when the number of factors compared to the dimension of the return vector is small.
- **Models of conditional variances and correlations:** At first the univariate conditional variances and correlations are computed and then used to get the conditional covariance matrix. Some models are for example the Constant Conditional Correlations (CCC) model and the Dynamic Conditional Correlations (DCC) model.
- **Nonparametric and semiparametric approaches:** Models in this class form an alternative to parametric estimation of the conditional covariance structure. The advantage of these models is that they do not impose a particular structure (that can be misspecified) on the data.

This paper uses **Models of the conditional covariance matrix** and **Models of conditional variances and correlations**. But before describing these models more profoundly, some definitions are required. This study is using two exchange rates series, resulting in a bivariate approach

$$r_t = \begin{bmatrix} r_{1,t} \\ r_{2,t} \end{bmatrix} \quad (12)$$

where r_t is the vector of returns, which can be decomposed as

$$r_t = \mu_t + \alpha_t \quad (13)$$

The return at time t can be divided in the two terms μ_t and α_t , where $\mu_t = E(r_t|F_{t-1})$ and $\alpha_t = H_t^{1/2} \cdot \epsilon_t$ respectively.¹⁷ The first term is the expected return for time t given the available information at the previous period, and this information set is denoted F_{t-1} . The covariance of the conditional unpredictable component is defined as

¹⁶This model categorization is in line with Orskaug (2009, p. 20).

¹⁷Obtaining $H_t^{1/2}$ could be somewhat difficult due to the issue of taking the square root of a matrix, whereby the **Cholesky Decomposition** is used. In addition, the likelihood function becomes dramatically simplified.

$$Cov(\alpha_t | F_{t-1}) = H_t = \underbrace{\begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{21,t} & h_{22,t} \end{bmatrix}}_{\text{Conditional variances}} \quad (14)$$

where ϵ_t are independently identically distributed random vectors with mean equal to zero. As can be seen in **Equation 14** the matrix obtained is symmetric since $H_t' = H_t$. Additionally, the matrix H_t has to be positive definite for all t . The approach of multivariate modeling brings complications. Firstly, it may be hard to ensure that H_t is positive for all t . Secondly, it easily becomes too many parameters to estimate, and at last it may be difficult to obtain the stationarity condition for $\sigma_{rt}^2 = E(H_t)$.

The following subsections will describe the theory of the models, where first the VEC, DVEC and CCC models are described in order to give the reader a basic understanding of the predecessors to the BEKK and DCC models.

3.2.4 The VEC model

The VEC model is a generalization of a univariate GARCH, and developed by Bollerslev, Engle and Wooldridge in 1988. In this model the conditional variances and covariances are functions of all the lagged conditional variances and covariances, but also of lagged squared returns and the cross products of the returns. (Silvennoinen & Teräsvirta, 2008, p. 3f) The general VEC model is defined as

$$vech(H_t) = C + \sum_{i=1}^p A_i \cdot vech(\epsilon_{t-i}\epsilon_{t-i}') + \sum_{j=1}^q B_j \cdot vech(H_{t-j}) \quad (15)$$

and the case when $p = q = 1$ is defined as

$$vech(H_t) = C + A \cdot vech(\epsilon_{t-1}\epsilon_{t-1}') + B \cdot vech(H_{t-1}) \quad (16)$$

The left term is the lower diagonal matrix that is transformed into a $N(N+1)/2 \times 1$ vector. For a VEC(1,1) model the left term can be expressed as following:

$$vech(H_t) = \begin{bmatrix} h_{11,t} \\ h_{12,t} \\ h_{22,t} \end{bmatrix} \quad (17)$$

Despite the models flexibility, the number of parameters tend to grow rapidly. The rate of growth can be expressed as $N(N+1)(N(N+1)+1)/2$.

3.2.5 The DVEC model

The diagonal VEC is a simplified VEC, where A_i and B_j are diagonal matrices in aspiration to obtain a positive H_t . This new version was also introduced by Bollerslev, Engle and Wooldridge in 1988. The DVEC(p, q) model is defined as

$$H_t = C + \sum_{i=1}^p A_i \odot (\epsilon_{t-i} \epsilon_{t-i}') + \sum_{j=1}^q B_j \odot H_{t-j} \quad (18)$$

where A_i and B_j are diagonal matrices and \odot is the Hadamard product of two matrices. The DVEC(1,1) is

$$H_t = C + A \odot (\epsilon_{t-1} \epsilon_{t-1}') + B \odot H_{t-1} \quad (19)$$

The model above can be decomposed into univariate GARCH models of variances and covariances.

3.2.6 The BEKK model

The BEKK model¹⁸ is a further development of the DVEC model. The parameters of this model can be configured in different ways, allowing the BEKK model to have a different degree of restrictions. The most restrictive one is the scalar BEKK, with a and b as scalars. A diagonal BEKK has diagonal matrices as parameters and the full BEKK uses $n \times n$ parameter matrices. The general full and diagonal BEKK model is

$$H_t = C \cdot C' + \sum_{i=1}^p \sum_{k=1}^K A_{ik} \cdot (\epsilon_{t-i} \epsilon_{t-i}') \cdot A_{ik}' + \sum_{j=1}^q \sum_{k=1}^K B_{jk} \cdot H_{t-j} \cdot B_{jk}' \quad (20)$$

where A_{ik} and B_{jk} are parameter matrices and C is a lower triangular matrix. (Silvennoinen & Teräsvirta, 2008, 3f) When $p = q = 1$ the model becomes a BEKK(1,1)

$$H_t = C \cdot C' + A \cdot (\epsilon_{t-1} \epsilon_{t-1}') \cdot A' + B \cdot H_{t-1} \cdot B' \quad (21)$$

The advantage of the BEKK model is that H_t by definition is positive.¹⁹ The matrices of parameters are multiplied with an arbitrary symmetrical matrix and the transpose of

¹⁸Where BEKK is an abbreviation for the creators of the model; Baba, Engle, Kraft and Kroner.

¹⁹As the matrix expects to attend real values, for the same reasons as when assuming a real number to be positive when taking the square root, an assumption has to be done, namely that the matrix H_t needs to be positive definite so it is possible to take the power of one half.

the parameter matrix. For example, CIC' where I is the identity matrix. This ensures that each term in the model becomes positive semi-definite.²⁰

One condition has to be fulfilled in order to ensure covariance stationarity, that the absolute eigenvalues of the expression $\sum_{i=1}^p \sum_{k=1}^K A_{ik} \otimes A_{ik} + \sum_{j=1}^q \sum_{k=1}^K B_{jk} \otimes B_{jk}$ have to be less than one.²¹ (Silvennoinen & Teräsvirta, 2009, p. 205)

3.2.7 The CCC model

The Constant Conditional Correlation model was suggested by Bollerslev (1990), where the time varying covariance matrix H at time t is expressed as

$$H_t = D_t R_t D_t \quad (22)$$

The right hand side consists of the conditional correlation matrix R that is time invariant, meaning that $R_t = R$. D is a diagonal matrix of (h_1, \dots, h_k) such as:

$$D_t = \begin{bmatrix} \sqrt{h_{1,t}} & & \\ & \ddots & \\ & & \sqrt{h_{k,t}} \end{bmatrix} \quad (23)$$

where each $h_{i,t}$ follows a univariate GARCH process. The conditional correlation matrix is given by $R = [\rho_{i,j}]$, and the non diagonal elements of H_t are

$$|H_t|_{i,j} = h_{i,t}^{1/2} h_{j,t}^{1/2} \rho_{i,j} \quad \forall i \neq j \quad (24)$$

Since the return process $r_{i,t}$ is modeled with a univariate approach, the desired conditional variances can be expressed in vector form

$$h_t = C + \sum_{i=1}^q A_i \epsilon_{t-i}^2 + \sum_{j=1}^p B_j h_{t-j}^2 \quad (25)$$

The first term C is a vector of the intercepts with a size of $n \times 1$ and the matrices of the coefficients are $n \times n$. Furthermore, $\epsilon_{t-j}^2 = \epsilon_{t-j} \odot \epsilon'_{t-j}$.

The advantage of the CCC model is that the computational procedure is more easily performed, because the correlation matrix $|H_t|_{i,j}$ is constant. However, this means that the model may be too restrictive. (Orskaug, 2009, p. 21f)

²⁰Let $x \in \mathbb{R}^n$ where x' is the transpose of x and let G be an arbitrary matrix. Then a matrix $M = x' G x$ is positive semi-definite if and only if G holds the properties of a Gram matrix. Furthermore, if the determinant of the Gram matrix is nonzero it is positive definite.

²¹Note that \otimes denotes the Kronecker product of two matrices.

3.2.8 The DCC model

This model was developed by Engle & Sheppard (2011). The Dynamic Conditional Correlation model is

$$H_t = D_t R_t D_t \quad (26)$$

where H_t is the covariance matrix and R_t is an $n \times n$ matrix of the conditional correlation of the returns. The diagonal matrix D_t is expressed as

$$D_t = \begin{bmatrix} \sqrt{h_{1t}} & 0 & \cdots & 0 \\ 0 & \sqrt{h_{2t}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{h_{nt}} \end{bmatrix} \quad (27)$$

This matrix consists of the univariate GARCH models. Furthermore, H_t has to be positive definitive, which is automatically obtained while R_t is a correlation matrix that is symmetric by definition. When this matrix is defined, two requirements are needed. Firstly, H_t needs to be positive definite since it is a covariance matrix. Secondly, the parts that belong to R_t need to be less than one. These requirements are met through a decomposition: $R_t = \text{diag}(q_{iit})^{-1} Q_t \text{diag}(q_{iit})^{-1}$ for $i = 1 \dots n$ where

$$Q_t = (1 - a - b) \bar{Q} + a \epsilon'_{t-1} + b Q_{t-1} \quad (28)$$

and $\bar{Q} = \text{Cov}[\epsilon_t \epsilon'_t] = E[\epsilon_t \epsilon'_t]$. Additionally, the parameters a and b are scalars and $\text{diag}(Q)$ is used to rescale the parts of Q_t in order to fulfill that $|\rho_{ij}| = |\frac{q_{ijt}}{\sqrt{q_{iit}q_{jjt}}}| \leq 1$ and where $\sqrt{q_{iit}}$ is the content of the matrix $\text{diag}(q_{iit})$. The estimate of \bar{Q} is

$$\bar{Q} = \frac{1}{T} \sum_{t=1}^T \epsilon_t \epsilon'_t \quad (29)$$

Moreover, the scalars a and b must be larger than zero, but the sum has to be less than one. One may note that these are conditions of the univariate GARCH to be stationary, but which is applied in the DCC model. (Orskaug, 2009, p. 21f)

3.3 The model order selection: AIC and BIC

An important step before making the estimations is to determine the order selection of the models. Theoretically one can do this using the autocorrelation function, but in practice this may be difficult. A more formal way is to use an information criterion

and choose the order that minimizes the criterion value. Two common criteria are the **Akaike Information Criterion** and the **Bayesian Information Criterion**.²² The formulas for these are

$$\begin{aligned} AIC &= -2 \times LLF + 2m \\ BIC &= -2 \times LLF + m \times \ln(N) \end{aligned} \tag{30}$$

where N is the sample size and m is the number of parameters. LLF is an abbreviation for log likelihood function. The reason for using both criteria is that the BIC is consistent but inefficient, and the AIC is the opposite, not consistent but efficient. No criteria is superior the other, but an overall assessment is needed based on the results showed by the criteria. (Brooks, 2008, p. 233ff) The result is presented below:

Model	AIC	BIC
BEKK(1,1)	-38764*	-38711*
BEKK(2,1)	-38754	-38682
BEKK(1,2)	-38756	-38684
BEKK(2,2)	-38746	-38655

Table 4: Order selection for BEKK

Model	AIC	BIC
DCC(1,1)	-39040*	39028*
DCC(2,1)	-39010	38997
DCC(1,2)	-39010	38997
DCC(2,2)	-39006	38992

Table 5: Order selection for DCC

As the tables show, the order of $p = q = 1$ seems to be the best suited for both of the models, and as can be seen a star points out the lowest value in each column. This outcome is a desirable one, since a smaller amount of parameters to estimate remarkably simplifies the computational procedure.

3.4 Estimation Procedure: The Quasi Maximum Likelihood method

A common issue using multivariate data for estimation is that the likelihood function becomes flat. Therefore, the choice of start values is a crucial aspect of the estimation. Nevertheless, the remedy for the problem is simply to perform the estimation with different start values. The starting values should be set in the way that renders the highest possible likelihood. (Orskaug, 2009, p. 33)

A majority of empirical papers has concluded that financial data suffer from fat tailed distributions and this is often corrected by assigning the data with a Student-t

²²The method using these criteria to determine the order selection is employed by for example Porjariev & Polasek (2003).

distribution. However, the maximum likelihood (ML) method relies on the assumed distribution and by appointing an incorrect distribution to the maximum likelihood data then in general the ML is not consistent. But, by using the quasi maximum likelihood (QMLE) the estimation remains consistent even under misspecification. Furthermore when dealing with models with conditional heteroscedasticity, the estimates are known to be asymptotically normal. (Bollerslev & Wooldridge, 1992)

Another thing that may be pointed out regarding the estimation procedure, is that robust standard errors are used since they allow the sample to contain heteroscedasticity. In other words, the robust standard error better deals with non constant variance.

3.5 Estimation evaluation

The evaluation is done with several methods. Firstly, the estimations are compared using the method of three measures: **MAE**, **RMSE** and the explanatory power from an ordinary least squares regression. These measures are explained in **Section 3.6.3**. Secondly, an attempt to determine the goodness of fit of the residuals is done through a univariate approach consisting of an analysis with the **Ljung-Box** test. Thirdly, an **ARCH LM** test is performed on the residuals. A fourth step is the **Baringhaus-Franz** test, and it is performed to validate the multivariate fit of the used models.

3.5.1 The Ljung-Box test

If the estimated output contains autocorrelated residuals, this indicates that there is still variation left to be explained by the model. One way to test for autocorrelation in the residuals is the Ljung-Box test which checks for autocorrelations different from zero in a set of data. The test's hypothesis, with lag length m is defined as

$$H_0 : \gamma_1 = \dots = \gamma_m = 0$$

$$H_a : \text{at least one } \gamma_i \neq 0, i = 1, \dots, m$$

where γ_i denotes the autocorrelation for a certain lag i and the test statistic is

$$Q_m = n(n+2) \sum_{i=1}^m \frac{\hat{\gamma}_i^2}{n-i} \quad (31)$$

where the number of observations is n , m the number of lags and i the lag length. When n gets large Q_m becomes asymptotically χ^2 distributed with m degrees of freedom. The null hypothesis, for significance level α , is rejected if $Q_m > \chi_{1-\alpha, m}^2$ where a rejection means that the hypothesis of random errors is rejected. In practice, several lag lengths are used and compared since different lags may give different results.

3.5.2 The ARCH LM test

This test was introduced by Engle in 1982, and it investigates whether the conditional volatilities of the chosen sample of data contain variation. The test can be used first to determine if there is volatility clustering in the data, and then to determine if there is any ARCH-effects left in the residuals. Firstly, the r_t^2 are regressed as

$$r_t^2 = c + \left(\sum_{i=1}^n a_i \cdot r_{t-i}^2 \right) + \epsilon_t \quad (32)$$

and secondly, the hypotheses are

$$H_0 : a_1 = a_2 = \dots = a_n = 0$$

$$H_a : \text{at least one of } a_i \neq 0$$

If H_0 is true, then homoscedasticity characterizes the variance indicating no volatility clustering. The test statistic is

$$LM = nR^2 \sim \chi^2 \quad (33)$$

where n is the number of observations and R^2 is the explanatory power of the regression model and obtained through the estimated residuals. If n is large, LM is χ^2 -distributed with ν degrees of freedom. The result from the ARCH LM test on the return series and the ARCH LM test of residuals are both presented in **Appendix A**.

3.5.3 The Baringhaus-Franz test

Even though a good fit in the univariate case is received (having a good marginal fit) it does not imply that the multivariate distribution fits well. The multivariate goodness of fit tests can be somewhat tedious. The Baringhaus-Franz multivariate test checks whether one data sample is identically distributed as another sample or not

$$H_0 : X_1 \text{ is distributed as } X_2$$

$$H_a : X_1 \text{ is not distributed as } X_2$$

Given these hypotheses, the test statistic is then

$$T = \frac{mn}{m+n} \left[\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \|X_i - Y_j\| - \frac{1}{2m^2} \sum_{i=1}^m \sum_{j=1}^m \|X_i - X_j\| - \frac{1}{2n^2} \sum_{i=1}^n \sum_{j=1}^n \|Y_i - Y_j\| \right] \quad (34)$$

As can be seen, m and n are the number of observations for X and Y , respectively.²³ Nevertheless, this test can be modified in a way to validate the fit of a model instead of testing two sample distributions. Let the sample X_1, X_2, \dots, X_m be vectors of errors denoted as z_t and Y_1, Y_2, \dots, Y_n be samples from a theoretical distribution.

As with all tests, the calculated value has to be compared to a critical value. This can be obtained by utilizing a bootstrapping method. The statistic in **Equation 34** is bootstrapped with a 95 % level. Then if the observed value of T ends up inside the confidence interval H_0 is accepted and it is possible to state that the model fits the selected sample well. (Orskaug, 2009, p. 42f)

3.5.4 Test of dynamic correlations

Some data have non-constant correlation and ignoring this could bring misleading conclusions. By testing the presence of dynamic correlation could therefore give a guidance of which type of multivariate GARCH model is most suitable for that type of data.²⁴ The hypotheses are constructed as

$$H_0 : R_t = R$$

$$H_a : vech^u(R_t) = vech^u(\bar{R}) + \beta_1 vech^u(R_{t-1}) + \beta_2 vech^u(R_{t-2}) + \dots + \beta_p vech^u(R_{t-p})$$

where $vech^u$ uses the elements above the main diagonal and R is the covariance matrix which under the null hypothesis is given by the identity matrix. First, an auxiliary regression of the elements above the main diagonal, and the corresponding lagged elements, is done. The regression parameters $\hat{\delta}$ are examined, and the test statistic is

$$\frac{\hat{\delta} X' X \hat{\delta}'}{\hat{\sigma}^2} \sim \chi_{\nu+1}^2 \quad (35)$$

where X is the regressor matrix. If the test is rejected, then the model is considered having dynamic correlation whereby a model allowing for this would be more suitable. (Engle & Sheppard, 2001, p. 11f)

²³It may be pointed out that $\|\cdot\|$ is the **Euclidian distance**. This means that in a two dimensional case, the Euclidean distance is the length between two dots in a real plane.

²⁴In other words, a model that assumes the conditional correlation to be constant or a model that allows it to be dynamic.

3.6 Forecasting Procedure and Evaluation

3.6.1 The Procedure

The series of conditional correlation forecasts for the period of 2010-01-01 to 2011-01-01 is compared to the realized values for the same period. This period consists of half hourly intraday data. Starting with an initial estimation for the ten years period of 2000-01-01 to 2010-01-01, a forecast is made for the following day ($T + 1$). The estimation window is moved one day and another forecast is made for ($T + 2$). This procedure continues throughout the forecast period. More generally expressed, the estimation window reaches between the first observation at time t and the last observation of the first period at time T . This mean that the length of the window d is $d = T - t$. The last observation of the evaluation period is denoted T^* . The forecast horizon k is set to be constant as $k = 1$. The forecast loop continues until the starting point t and the ending point T of the first period both have moved a number of steps equal to $T^* - T - 1$, which results in a series of $T^* - (T + 1)$ forecasts.

3.6.2 Forecasting with the BEKK and DCC models

After estimating the models, the forecast procedure can be performed. When forecasting the BEKK(1,1) model one period ahead, one period²⁵ is added to the lagged terms, which renders in the following model

$$\hat{H}_{t+1} = \hat{C} \cdot \hat{C}' + \hat{A} \cdot E_t(\epsilon_t \epsilon_t') \cdot \hat{A}' + \hat{B} \cdot H_t \cdot \hat{B}' \quad (36)$$

As the equation states, the earlier received estimates are used in the forecast. For the DCC model on the other hand, the forecast expression for one step ahead in time is written as

$$H_{t+1} = D_{t+1} R_{t+1} D_{t+1} \quad (37)$$

What that complicates the forecast is that the forecasts for D_{t+1} and R_{t+1} have to be made by themselves. First

$$\hat{D}_{t+1} = E[D_{t+1}|F_t] = \text{diag} \left(\sqrt{E[h_{1,t+1}|F_t]}, \sqrt{E[h_{2,t+1}|F_t]} \right) \quad (38)$$

where F_t is the information available at time t . The second step is to calculate R_{t+1} . Orskaug (2009, p. 35f) suggests two different assumptions in order to achieve unbiased

²⁵The reason for only forecasting one day ahead is due to that when the forecast horizon $k \rightarrow \infty$, the forecast of the conditional correlation converges to the unconditional correlation.

forecasts. The assumption has to be made because $E [\text{diag} (Q_t^{-1}) Q_t \text{diag} (Q_t^{-1})]$ is unknown, meaning that the forecast cannot be made in a direct way, whereby this paper therefore assumes that $E [\epsilon_{t+1} \epsilon'_{t+1} | F_t] \approx E [Q_{t+1} | F_t]$. The desired relationship²⁶ is then

$$\hat{R}_{t+1} = E [R_{t+1} | F_t] \approx \text{diag} (\hat{Q}_{t+1}^{-1}) \hat{Q}_{t+1} \text{diag} (\hat{Q}_{t+1}^{-1}) \quad (39)$$

where it may be pointed out that $\text{diag} (\hat{Q}_{t+1}^{-1})$ is a diagonal matrix. (Orskaug, 2009, p. 35f) When both the estimates \hat{D}_{t+1} and \hat{R}_{t+1} are calculated, then the one step ahead forecast for the conditional covariance matrix can be computed.

3.6.3 Defining a proxy

First of all, the question is how the true correlation ρ is determined. Since it is not observable at time t , an approximation is needed. In accordance with Andersen et al. (2007, p. 528f), the realized variance is used, and for a certain forecast $t + 1$ the proxy for the realized variance is set to be $r_{(m),t+1/m}^2$ where m is the number of intraday samples. For exchange rates, $m = 1$ is daily returns and $m = 24$ is hourly returns. When $m \rightarrow \infty$ the realized variance is getting closer to the integrated variance that occurs with continuous sampling frequency. The estimate of the realized covariance matrix is expressed as

$$VCM_{t,m} = \sum_{m=1}^M R_{m,t} R'_{m,t} \quad (40)$$

and if n is the number of assets, then $R_{m,t}$ is a $n \times M$ matrix for a given day t . By arbitrary²⁷ setting the number of intraday sub-periods M to different values, the above expression for the realized covariance is used in order to calculate the realized correlation which is used as a benchmark for the forecasts.

3.6.4 Ordinary Least Squares regression

In order to determine if the performed estimations and forecasts have been accurate or not, they are compared to the the aforementioned realized correlations. Using an

²⁶The derivation of this relationship is presented by Orskaug (2009, p. 36).

²⁷In fact, a variety of research has been aiming to determining the optimal sampling frequency for the realized variance. One could easily assume that the integrated variance is the best proxy, however microstructure effects in financial data, for example the **bid-ask bounce**, makes the proxy biased when $m \rightarrow \infty$. For further reading about optimal sampling frequency, see Bandi & Russel (2007), and O'Hara (1995) for a general overview of microstructure noise effects.

intuitive method, the estimations and predictions are used in an ordinary least squares regression as an explanatory variable and $RC_{t,m}$ is the dependent variable. From this regression the explanatory power, denoted as R^2 , is extracted and interpreted. The regression is

$$RC_t = a + b\hat{\rho}_t + \epsilon_t \quad (41)$$

where a is the intercept, b the parameter of $\hat{\rho}_t$ and ϵ_t is the error term for the regression. A traditional t -test is used in order to determine if the estimated parameter b in the regression can be considered to be equal to zero or not. It may be pointed out that in the forecasts, one period is added to the time notation, i.e. it is $t + 1$ instead.

To check whether the regression is unbiased, one can examine if the residuals are $E[\epsilon_t] = 0$ and white noise.²⁸ Furthermore, since the residuals in an OLS regression is assumed to be normal, a Jarque-Bera test of the residuals is performed.

3.6.5 Mean Absolute Error (MAE)

One way of determining the goodness of the estimations and forecasts is interpreting the MAE. The approach is to measure how close the received conditional correlations are from their corresponding realized value. The formula is

$$MAE_{i,j} = \frac{1}{n} \sum_{k=1}^n |\rho_{ik} - \hat{\rho}_{ik}| \quad (42)$$

where the proxy is used for ρ_{ik} and the estimated conditional correlation is used as $\hat{\rho}_{ik}$. By comparing the MAE between the estimated models, one is given an indication of which model that makes the best estimations.

3.6.6 Root Mean Square Error (RMSE)

The third measure is the Root Mean Square Error (RMSE), which is defined as

$$RMSE_{i,j} = \sqrt{\frac{1}{n} \sum_{k=1}^n |\rho_{ik} - \hat{\rho}_{ik}|^2} \quad (43)$$

Using these methods, the estimated models can be compared, and using the same measurements for estimations and forecasts, one can determine if the relatively best estimations model also makes the best forecasts.

²⁸The residuals should not be forecastable, i.e. $a = b = 0$ in $\epsilon_{t+1|F_t} = a + b\hat{\rho}_{t+1|F_t} + \epsilon_{t+1}$ where F_t is the available set of information given at time t .

3.7 Practical application using the MGARCH forecasts: Historical Value-at-Risk

3.7.1 Background of Value-at-Risk

Given the received forecasts, it may be convenient to use these in a practical implementation. There are capital requirements²⁹ on financial institutions regarding risk measurement, e.g. the concept of Value-at-Risk (VaR). For example, there are **quantitative parameters** that the Value-at-Risk (VaR) calculations need to satisfy according to the Basel regulations:

- A horizon of 10 trading days (or two calendar weeks).
- A 99 % confidence interval.
- The observation period needs to be based on at least one year of historical data. In addition, it needs to be updated at a minimum every quarter.

First, let define VaR for a given confidence level α as

$$VaR_\alpha = -h_t \times \Theta_\alpha \quad (44)$$

where h_t is the conditional standard deviation, which is multiplied by Θ_α such as it satisfies the condition of $P(r_t < \Theta_\alpha) = \alpha$ where r_t is the return at time t .

This study is making VaR evaluations at 1-day ahead and at a confidence level of 5 %, but due to the earlier stated capital requirements also at 10-days ahead and an α of 1 %. Using a historical VaR approach, the conditional return distribution 10-days ahead is simulated given the historical series of returns.³⁰ To summarize the procedure, the standardized residuals are defined as $e_t^* = \hat{e}_t \div \sqrt{\hat{h}_t^2}$. Then $z_{t+k}^* = e_1^* \times \sqrt{h_{t+k}^2}$, where h_{t+k}^2 is the forecasted conditional variance and k is the forecasting horizon. Finally, the simulated return is $r_{t+k}^* = \hat{\mu} + z_{t+k}^*$. Using a bootstrap method, a sample of future returns is simulated and their distributions calculated.

3.7.2 Backtesting the Value-at-Risk estimates

One way of evaluating whether VaR forecasts have been successful or not is to count the number of occasions (in this case the number of days) when the actual loss of the

²⁹The capital requirements are imposed by The Basel Committee on Banking Supervision.

³⁰This method has a drawback, the future simulated distribution is based on the historical distribution.

portfolio exceeds the VaR estimate. The failure rate is $FR = \frac{x}{N}$, where x is the amount of values that exceeds the VaR estimates and N is the number of observations. If the VaR level is set to be $p = 1\%$, then FR in an optimal scenario is an unbiased measure of p . Since each day causes a violation of the VaR level or not, a series of successes or failures is received. This series is known as a **Bernouli trial**, and the failures x follow the binomial distribution

$$f(x) = \binom{N}{x} p^x (1-p)^{N-x} \quad (45)$$

that can be approximated with the normal distribution, when the number of observations increases, as $z = \frac{x-pN}{\sqrt{p(1-p) \times N}} \approx N(0, 1)$ where the variance of the expected number of failures $p \times N$ is calculated as $p(1-p) \times N$.³¹ (Jorion, 2007, p. 143f)

The test of Kupiec (1995) determines if the number of failures is the same as the confidence level α . H_0 is then that $p = \hat{p} = x/N$ and H_a that it is not. The likelihood ratio is defined as

$$LR = -2\ln \left[(1-p)^{N-x} p^x \right] + 2\ln \left[(1-(x/N))^{N-x} (x/N)^x \right] \quad (46)$$

Given that H_0 is correct, the test statistic is asymptotically χ^2 distributed with $\nu = 1$. If the received value is higher than the critical value given by the χ^2 distribution, H_0 is rejected. When this happens, the conclusion is that the model is inaccurate. It may be pointed out, in order to avoid confusion, that the significance level of the test is not the same as the significance of the VaR model. The evaluation of the test results can be summarized as:

- **Large number of failures in relation to α :** The risk is underestimated.
- **Small number of failures in relation to α :** The risk is overestimated.

The two outcomes above imply a bad model, since it fails to predict the risk. (Jorion, 2007, p. 147)

³¹Note that when a test that either accept or reject an outcome, there is a tradeoff between **Type I error** and **Type II error** that has to be considered. The power of the test increases as the number of observation increases, whereby a large sample is one way to reduce the risk for errors.

4 Estimations

This section will present the results from the estimations of the two selected models, using the exchange rate data of EUR/SEK and USD/SEK. In both the estimated models $p = q = 1$, in accordance with the results of the **AIC** and the **BIC**. The last subsection contains a discussion of the estimation results, where the authors aim to intuitively discuss, based on economic events rather than statistical theory, peculiarities in the results.

4.1 BEKK(1,1)

Starting with estimating a full BEKK model, the following estimation results are achieved:

Parameter	Estimated Coeff.	Std. Error	t-value	p-value
c_{11}	0,7782	0,1482	5,2513	0,0000
c_{21}	0,8787	0,3801	2,3117	0,0209
c_{22}	1,3511	0,2518	5,3659	0,0000
a_{11}	0,2254	0,0181	12,4357	0,0000
a_{12}	0,0001	0,0114	-0,0019	0,9985
a_{21}	-0,0002	0,0047	0,0242	0,9807
a_{22}	0,2252	0,0165	13,6699	0,0000
b_{11}	0,9715	0,0047	208,1979	0,0000
b_{12}	-0,0005	0,0049	-0,0267	0,9787
b_{21}	-0,0001	0,0017	-0,3055	0,7600
b_{22}	0,9712	0,0041	239,4315	0,0000

Table 6: Full BEKK estimation results

When evaluating **Table 6**, it can be observed that the non diagonal elements are non significant which mirrors the theoretical fact that these parameters often are redundant. Due to this, the estimation is instead performed with diagonal BEKK model, trying to achieve significant parameters. When the diagonal BEKK is decomposed into its univariate GARCH models, they are expressed as

$$\begin{aligned}
 h_{11,t} &= c_{11}^2 + a_{11}^2 \epsilon_{1,t-1}^2 + b_{11}^2 h_{11,t-1} \\
 h_{12,t} &= c_{21}c_{11} + a_{11}a_{22}\epsilon_{1,t-1}\epsilon_{2,t-1} + b_{11}b_{22}h_{12,t-1} \\
 h_{22,t} &= c_{21}c_{11} + c_{22}^2 + a_{22}^2 \epsilon_{2,t-1}^2 + b_{22}^2 h_{11,t-1}
 \end{aligned} \tag{47}$$

where it is shown that in the bivariate case, it is now seven parameters to estimate

instead of eleven. The following table shows the results from the estimation of the diagonal BEKK:

Parameter	Estimated Coeff.	Std. Error	t-value	p-value
c_{11}	0,7644	0,1265	6,0433	0,0000
c_{21}	0,8491	0,2127	3,9916	0,0001
c_{22}	1,3046	0,2523	5,1699	0,0000
a_{11}	0,2306	0,0164	14,0400	0,0000
a_{22}	0,2104	0,0141	14,9043	0,0000
b_{11}	0,9703	0,0041	235,0779	0,0000
b_{22}	0,9746	0,0035	278,5663	0,0000

Table 7: Diagonal BEKK estimation results

All of the parameters are now highly significant. Additionally, the remaining parameters are very close to the parameter estimations of the full BEKK model. Despite the restrictions of the model construction, the model seems to suit the selected sample of data. The t -values have been calculated through $t_{\hat{\xi}} = \frac{\hat{\xi}}{SE(\hat{\xi})}$ where SE is an abbreviation for standard error and ξ a parameter to be tested. The following matrices display the parameter estimates

$$C = \begin{bmatrix} 0,7644 & 0 \\ 0,8491 & 1,3046 \end{bmatrix} A = \begin{bmatrix} 0,2306 & 0 \\ 0 & 0,2104 \end{bmatrix} B = \begin{bmatrix} 0,9703 & 0 \\ 0 & 0,9746 \end{bmatrix} \quad (48)$$

To determine whether the model estimates are covariance stationary or not, the condition presented in the theoretical description of the BEKK model is used. Let the Kronecker product (K) of A and B be

$$K = \begin{bmatrix} 0,2306 & 0 \\ 0 & 0,2104 \end{bmatrix} \otimes \begin{bmatrix} 0,2306 & 0 \\ 0 & 0,2104 \end{bmatrix} + \begin{bmatrix} 0,9703 & 0 \\ 0 & 0,9746 \end{bmatrix} \otimes \begin{bmatrix} 0,9703 & 0 \\ 0 & 0,9746 \end{bmatrix} \quad (49)$$

and by setting $\det(K - \lambda I) = 0$ the following eigenvalues are obtained

$$\begin{bmatrix} 0,9941 & 0,9942 & 0,9942 & 0,9947 \end{bmatrix}' \quad (50)$$

The eigenvalues show that the condition of covariance stationarity is fulfilled for the estimated model since all the calculated absolute eigenvalues are less than one. Additionally, if the matrices are expressed as AA' and BB' then

$$AA' = \begin{bmatrix} 0,0531 & 0 \\ 0 & 0,0443 \end{bmatrix} BB' = \begin{bmatrix} 0,9415 & 0 \\ 0 & 0,9498 \end{bmatrix} \quad (51)$$

where the sums of the parameter pairs, respectively, equal $0,0532 + 0,9415 = 0,9947$ and $0,0443 + 0,9498 = 0,9941$. These results also indicate that the estimation has been successful.

The conditional variance has been relatively stable, except the second half of 2008 and the year 2009, due to the financial turbulence. Some higher degree of conditional variance can be spotted in the beginning of the sample for EUR/SEK. The graphs confirm what the received values of B indicate, that the conditional variance is persistent. The next graph visualizes the estimated conditional correlation:

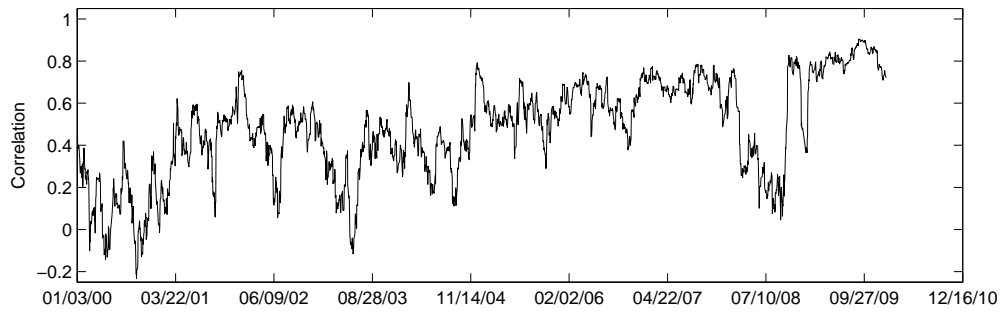


Figure 9: Estimated conditional correlations using BEKK

The estimated conditional correlation graph shows a more non stable pattern, although it seems that the moving pattern is centralized around an upward rising trend, correlating from around 0.2 to about 0.8 except from a structural break at the near end of the sample period. The conditional covariance graph is:

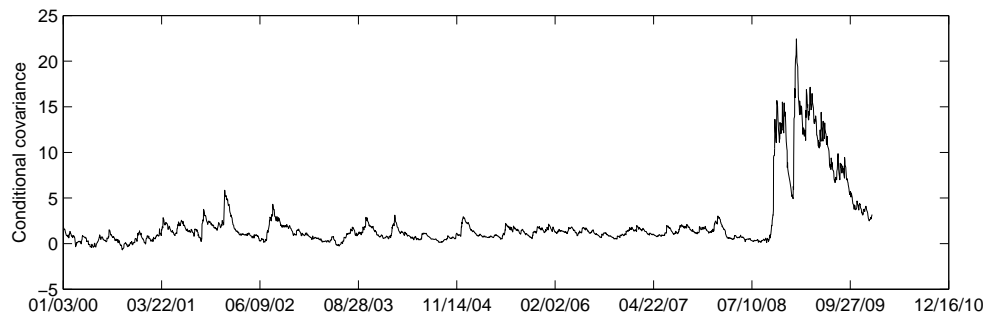


Figure 10: Estimated conditional covariance using BEKK

The next two graphs displays the estimated conditional variance during the sample period:

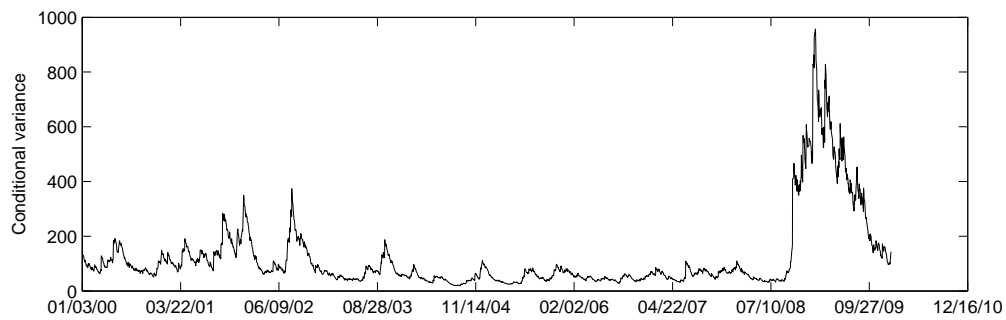


Figure 11: Estimated conditional variance for EUR/SEK using BEKK

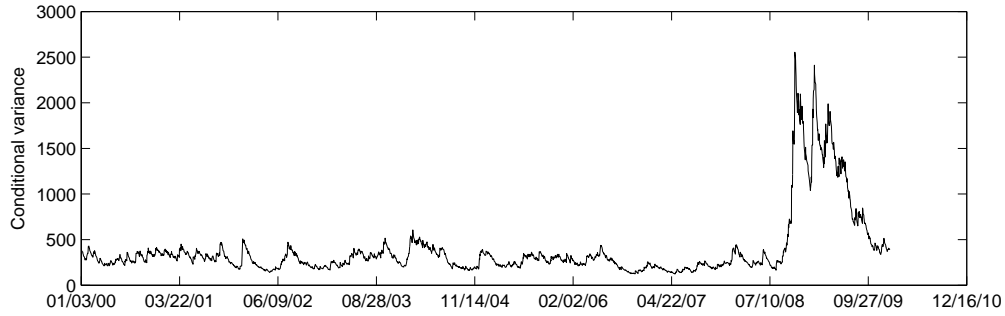


Figure 12: Estimated conditional variance for USD/SEK using BEKK

4.2 DCC(1,1)

To confirm that the DCC model is preferred over the CCC model, a test of dynamic correlation is made in order to see which one of the models that is most suitable for the data. A p -value of less than 0.05 is received, pointing out the absence of constant correlation and the data is thus suitable for using a DCC model compared to the CCC model where the correlation is assumed to be constant. Consequently, an estimation of the DCC model will be done. The table below consists of the estimation results:

Parameter	Estimated Coeff.	Std. Error	t -value	p -value
c_{11}	0,6990	0,2565	2,7251	0,0066
a_{11}	0,0765	0,0127	6,0069	0,0000
b_{11}	0,9190	0,0121	75,9275	0,0000
c_{22}	4,0583	1,4540	2,7912	0,0053
a_{22}	0,0500	0,0094	5,2960	0,0000
b_{22}	0,9380	0,0116	80,7735	0,0000
α	0,0356	0,0062	5,7422	0,0000
β	0,9615	0,0072	133,5250	0,0000

Table 8: DCC estimation results

The last two columns are the α and β of the estimated DDC model. As can be seen all the coefficients are significantly not equal to zero, which implies that they all have a significant impact. A noticeable fact is that the degree of significance seems to follow a particular pattern, where the t -values of the intercepts and the ARCH parameter are relatively lower compared to the GARCH parameters which are highly significant.

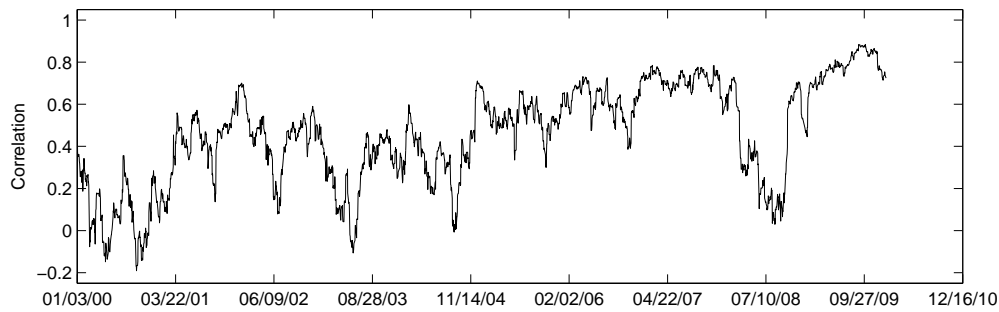


Figure 13: Estimated correlations using the DCC model

In line with the BEKK estimation results, the general impression is that the correlation seems to be trending upwards during this ten years period. A few times the correlations have been negative³², especially in the beginning of the estimated period. Furthermore, another impression is that the graph shows tendencies of correlational drops during periods of economic crises, most easily observed is the period around the year 2009.

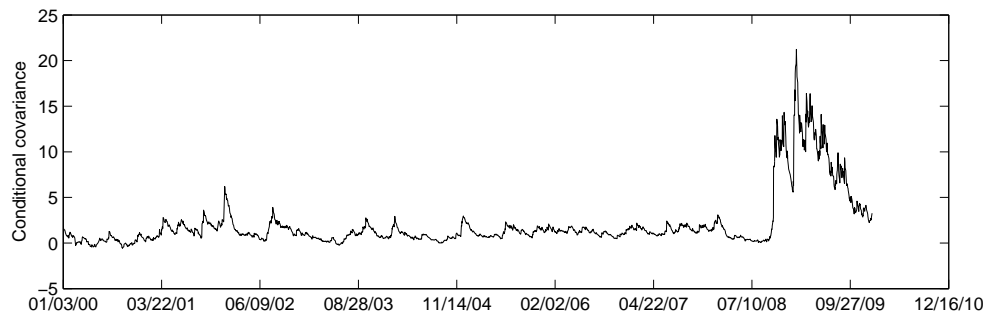


Figure 14: Estimated conditional covariance using DCC

Another thing that stands out is a relatively high value of the second models intercept that equals 4.0583. This indicates that the averages variance of the **USD/SEK** is at an higher level. This fact is confirmed by **Figure 15-16**.

³²The total number of days with negative correlation is 103, where 101 of those days occurred during the first three years of the sample.

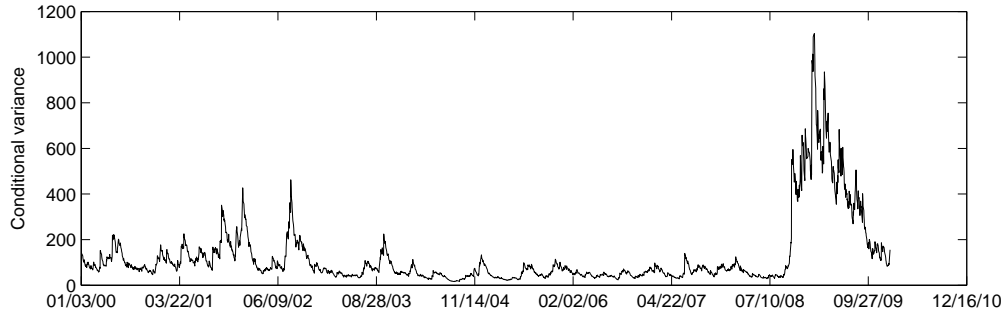


Figure 15: Estimated conditional variance for EUR/SEK using the DCC model

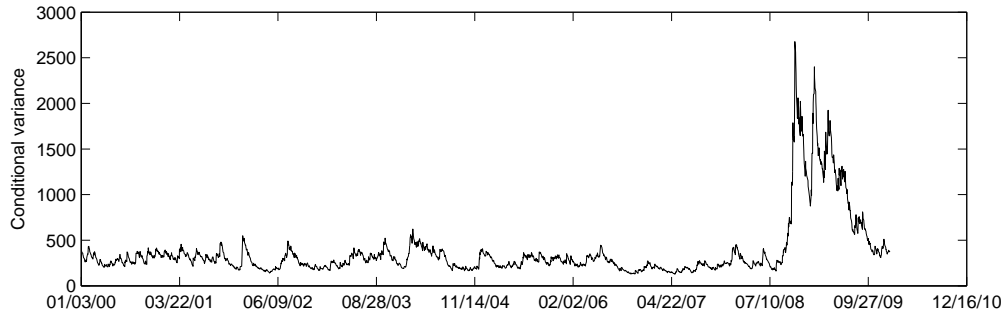


Figure 16: Estimated conditional variance for USD/SEK using the DCC model

As stated in the theoretical description of the DCC model the scalars a and b have to individually be larger than zero but at the same time the sum of them has to be strictly less than one. The results in **Table 8** show that $a + b = 0.0356 + 0.9615 = 0.9971$ and in addition, the parameters of the inherent univariate processes are $a_{11} + b_{11} = 0.0765 + 0.9190 = 0.9955$ and $a_{22} + b_{22} = 0.0500 + 0.9380 = 0.9878$. To interpret this, a high b means that the conditional variance is persistent. A high a means that the volatility is spiky. Since our values of b are high, the conditional variances seem to be more persistent. As can be seen $a_{11} > a_{22}$, indicating a more spiky time series graph. Looking at **Figure 15** and **16**, this is also true. **Figure 13** is the conditional correlation graph for the DCC model:

4.3 Comparison of the estimated models

4.3.1 Differences in the estimations

One way to compare the model's estimation performance is to look at the spread between the estimated correlations. **Figure 33** in **Appendix B** displays this spread. As can be seen, the spread is to be considered relatively small. Interestingly enough, the spread seems to increase during periods of crises, for example the year 2009. The mean of the spread series equals 0.0188. Testing whether this estimated mean is significantly different from zero or not renders in a p -value of approximately 0.000, indicating that the two models in average result in different time series of correlations. Since it is the DCC correlations that have been subtracted from the BEKK estimations, it is with statistical significance possible to state that the BEKK estimated model on average produces higher conditional correlations than the DCC.

4.3.2 Residual analysis

Starting by looking at the residuals they should be following a white noise process. The residual graphs from the estimations are attached in **Figure 23, 24, 28** and **29** in **Appendix B**. Furthermore, a check for autocorrelation in the residuals is needed. The autocorrelation function (ACF) graphs are also attached in **Appendix B**.

By looking at the ACF graphs, it can be determined whether the processed data has any trace of correlation in the residuals. If more than 5 % of the lags in the autocorrelation plot falls outside the confidence interval, the residuals can not be treated as white noise. In both the estimated BEKK and DCC, about five out of 100 lags fall outside the confidence level. Even though that is close to the limit, the critical values seem to appear randomly and do not make a pattern, it can therefore be concluded that the data are uncorrelated and the residuals do not suffer from heteroscedasticity.

Another evaluation method is to look at the XCF graphs, **Figure 27** and **24**. The disadvantage with the autocorrelation function is that only the marginal from the two univariate time series are examined, leaving the cross correlation between the two samples unexplained. It is thus useful to check the correlation between the samples that is left in the residuals. Using the same procedure, a number of five lags outside the confidence interval indicate cross correlated residuals. As there is no pattern in the plot, it is concluded that the residuals are not cross correlated.

As mentioned, the standardized residuals seem to be white noise, but to confirm this the Ljung-Box test is made. **Table 14** and **15** in **Appendix A** contain the test results. Since all p -values exceeds 5 %, the H_0 that there is no autocorrelation is not rejected, meaning that the test confirms that the estimations have been successful,

considering an analysis of the residuals.

In **Appendix A** the results from the ARCH LM test of the return series can be found, which indicate that there is high presence of volatility clustering in the data before the estimation are made. The next model diagnostic step is to determine whether there are still ARCH effects left in the residuals. If there still are unexplained effects left, this means that the chosen model suits the sample poorly, since the return series suffer from heteroscedasticity according to the first ARCH LM test.

As can be seen in **Equation 32**, a desirable outcome for the ARCH LM test of residuals is that all $a_i = 0$ and that it is not possible to reject H_0 . **Table 17** and **18** show insignificant p -values for all lags, even though it seems that the BEKK model overall better handles the ARCH effects. Additionally, the EUR/SEK residuals seem to suffer less from ARCH-effects. The conclusion of the two ARCH LM tests is that both models capture the volatility clustering successfully.

4.3.3 Goodness of fit

An additional step in the model evaluation is to look at the goodness of fit with the Baringhaus-Franz test. If the residuals are distributed as the theoretical distribution, then the model can be considered as successfully suiting the sample. The H_0 is that the two samples are identically distributed, and therefore the aim is to not reject H_0 . However, according to Orskaug (2009, p. 69) it is difficult to achieve a p -value higher than 1 % using empirical data. The Baringhaus-Franz test results are attached in **Table 16** in **Appendix A**. The critical value CV in this table is based on 1000 ordinary bootstrap replicates and T is the test statistic from the test. As pointed out in the table, the BEKK model is relatively better suited compared to the DCC model. Nevertheless, the results are only significant on the 1 % level. But considering the aforementioned concerns stated by Orskaug (2009), the results can still be seen as good. Since the estimations of both models are successfully performed, next step is to investigate their accuracy in forecasts. But before that, a discussion about the estimations is presented.

4.4 Discussion of the estimations

Dividing the conditional variance graphs into pairs, **Figure 11** and **15** for the EUR/SEK and **Figure 12** and **16** for the USD/SEK, all show a similar pattern. In the summer of 2008 the variance of the exchange rates explodes after many years at stable levels. **Figure 9** and **13** show evidence of a structural break in the conditional correlations during 2008, hitting levels close to zero, even though the levels before and after this period fluctuated around 0.8. This is an interesting result, because with an intuitive ap-

proach the correlation between two assets during a crisis should be increasing, simply due to that the majority of asset prices are moving in the same direction.

A probable reason for the movements in the conditional variances and the conditional correlation is the financial crises of 2008. The crises had its origin in an overvalued and over leveraged mortgage market in the United States. To shortly summarize, banks gave loans to households with lower repayment abilities, since the banks assumed that the market value of the assets, the houses, would continue to increase. The idea was that the amortization and interest payments for the loans could be postponed to the future when the asset was expected to possess a higher market value. The loans themselves were packaged and further sold to different investors, which increased the subsequent contagion effect. **Figure 34** in **Appendix B** is the DCC correlation estimates enlarged for the year of 2008. It visualizes that the downward movement from the first marked point to the second point is rather large. The conditional correlation went down from levels around 0.7 at the end of February 2008 to approximately 0.1 at the beginning of June the same year. The conditional correlation stayed at these levels for a couple of months before it went upwards again, back to the previous levels of the beginning of 2008.

However, the graph in section three, **Figure 1**, visualizes that both the USD/SEK and the EUR/SEK had upward movements. During periods of crises smaller currencies tend to suffer more since a larger economy is considered to be better suited to handle the effects of a crisis. Investors therefore prioritize currencies that are believed to be reliable. This is according to the Swedish Central Bank themselves one of the main reasons for the depreciation of the Swedish currency against other currencies during this period. (Penningpolitisk rapport 2008:3, 2008, p. 23)

The sharp shift broke a trend where the USD/SEK moved in a downward pattern for a couple of years. The sharp shift may also depend on the fact that the Swedish central bank initiated a series of aggressive cuts in the key interest rate, from an initial key interest rate of 4.75 % in October in 2008 it ended up at 0.25 % in July 2009. As followed by the interest rate parity theory, the SEK may have depreciated since investors could probably find higher returns in other currencies. **Figure 35** in **Appendix B** is attached to show the levels of the key interest rates for the United States, Sweden and the Eurozone during the estimation period from 2000-01-01 to 2010-01-01. The figure shows that all three regions initiated key interest cuts during the period of the financial crisis, although the American central bank FED started earlier in order to stimulate the failing domestic economy. When FED made their cuts, the system probably expected other central banks to follow, and when the Swedish central bank finally decided to start making cuts, the Swedish currency started to depreciate.

Another way of analyzing the structural break in the correlation between exchange

rates is to evaluate the Credit Default Swap (CDS) movements for domestic banks for each region. One explanation could be that if the Swedish banks CDS prices increased earlier due to turbulence in the Baltic region, this could have impact on the exchange rates since higher CDS prices mean that the market consider the probability of default as higher. That could be a reason for the exchange rate movement.³³

But, **Figure 36** in **Appendix B** reveals that the movements of CDS prices seem to be highly correlated, and this strengthened the hypothesis that the key interest rate cuts may be the main reason for the structural break.

³³The Swedish series is an average of the CDS prices for Swedish banks. The US series is based on the individual CDS:s for Bank of America , Citibank, Goldman Sachs, J.P. Morgan, Lehman Brothers, Morgan Stanley and Wells Fargo & Co (where an adjustment is made for the bankruptcy of Lehman Brothers). Finally, for the Eurozone the ITRAXX index is used.

5 Forecasts

As mentioned in **Section 3.6.1**, the forecasts are made one day ahead throughout the evaluation period. However, due to computational restrictions in the BEKK forecasting procedure, the parameter estimation is made every second day. On the contrary, the DCC model estimation is every day, but made in two steps. First, two univariate GARCH(1,1) estimations are made, and thereafter a DCC parameter estimation based on the first estimation results. This procedure is less computational demanding, hence enabling every day estimation.

First, **Figure 17** and **18** show the forecasts together with the last part of the estimations. The horizontal dotted line marks the zero correlation level, and the vertical dotted line is the breakpoint between estimations and forecasts.

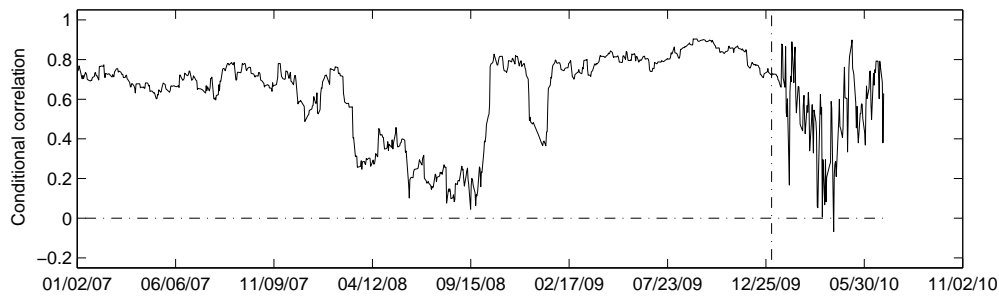


Figure 17: BEKK correlation forecasts

Forecasting with the BEKK model seems somewhat difficult as the forecast falls out to be very volatile. The DCC forecasts on the other hand, seem to be more stable:

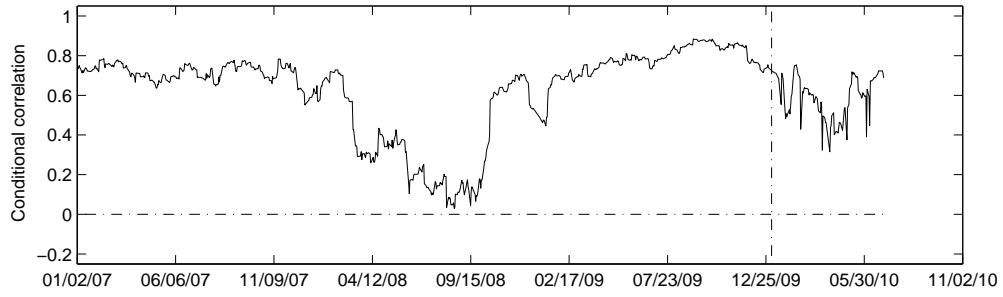


Figure 18: DCC correlation forecasts

Second, **Figure 19** and **20** consist of the forecasting period, where the solid line is the forecasted conditional correlations and the dotted line the realized correlations:

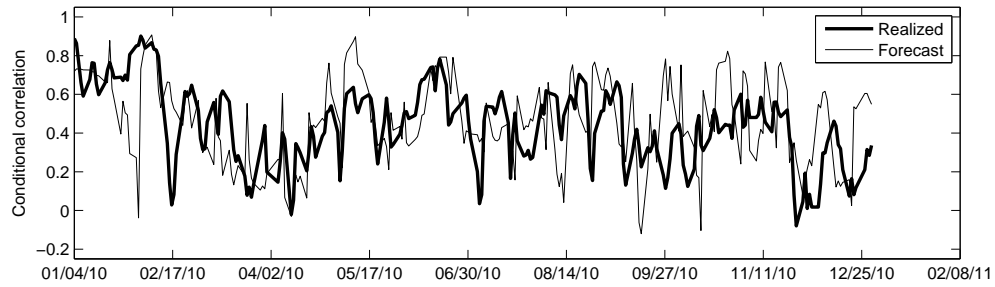


Figure 19: BEKK forecasts and Realized Correlation ($m = 48$)

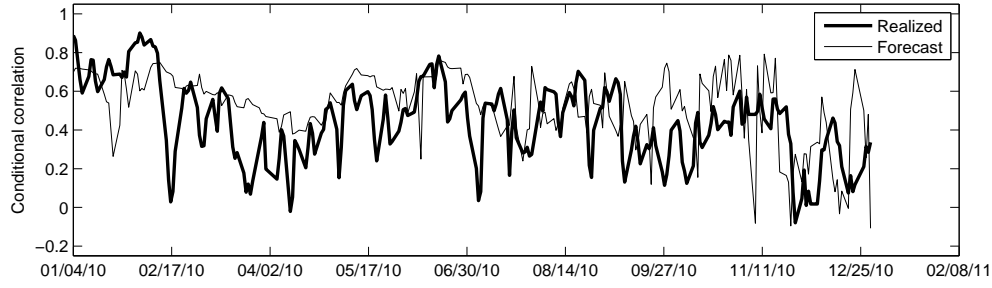


Figure 20: DCC forecasts and Realized Correlation ($m = 48$)

Only evaluating by looking at the graphs may be difficult. However, **Table 9** and **10** presents the earlier mentioned evaluation measures for each model and different amounts of intraday samples for the proxy:

m	R^2	MAE	RMSE
12	9,4%	0,25	0,30
24	16,7%	0,20	0,25
48	17,0%	0,18	0,23

Table 9: BEKK forecast evaluation

m	R^2	MAE	RMSE
12	6,8 %	0,29	0,33
24	10,0%	0,22	0,27
48	11,4%	0,19	0,24

Table 10: DCC forecast evaluation

The intuitive measure of using the explanatory power reveals that the BEKK model performs better than the DCC model for all levels of m . Moreover, a trend where an increase in m results in better forecasts. This is intuitive since the proxy realized correlation get better when the number of intraday samples increases. Unfortunately, 30-minutes intervals of intraday data was the smallest accessible interval. Even smaller intervals would probably result in better forecasts.

6 Historical Value-at-Risk

This section presents the results from the historical VaR simulation. Starting with a theoretical portfolio value of 100 SEK, the future portfolio value distribution is simulated. The accuracy of the predictions are then evaluated.

6.1 Historical VaR Forecasts

Looking at **Table 11** and **12**, it is seen that the VaR forecasts based on the BEKK model in all cases ends up with the lowest critical values.

1-day VaR		
Model	1%	5%
GARCH	66,68	76,96
BEKK	41,76	66,10
DCC	59,10	72,87
GARCH*	68,48	80,20
BEKK*	90,68	93,88
DCC*	87,64	90,87

Table 11: 1-day VaR estimates (in SEK) for the portfolio at given confidence levels

10-day VaR		
Model	1%	5%
GARCH	19,28	27,34
BEKK	7,73	21,83
DCC	11,03	32,48
GARCH*	45,99	53,01
BEKK*	81,13	74,65
DCC*	62,56	73,59

Table 12: 10-day VaR estimates (in SEK) for the portfolio at given confidence levels

Recall **Figure 33**, the BEKK model tends to estimate relatively higher variances compared to the DCC model. Assuming that the portfolio is equally weighted and has an initial value of 100 SEK, the VaR estimates based on BEKK forecasts state that the 1-day probability of ending up having a portfolio worth 41.76 SEK or less is 1%. For 10-day the amount is 7.73 SEK for the same significance level. This equals losses of approximately 58 % and 92 %, respectively. The GARCH(1,1) model is included to show the effect when the diversification effect is excluded, since the covariance between the assets are not taken into account. In the tables, the **GARCH***, **BEKK*** and **DCC*** represent VaR estimates of portfolios where instead of a portfolio of two long positions, the portfolio is hedged with one long and one short position. The initial value is 100 SEK for these portfolios also.

The simulated future distributions of portfolio values for each model can be seen in **Figure 40-48** in **Appendix B**. As they show, the distributions are rather different compared to each other, for example the distribution for 10-day BEKK(1,1) is relatively skewed. The next step is to evaluate these VaR forecasts and see to what extent they can be considered accurate.

6.2 Backtesting the VaR Forecasts

In order to evaluate the accuracy of the VaR, further VaR estimations are needed. Therefore, one year of VaR is simulated and thereafter one can look at the number of critical value violations. As can be seen in the one day ahead VaR figures attached in **Appendix B**, the correlation that the multivariate models take into account, leads to a more cautious VaR estimation, where the univariate GARCH model seems to underestimate the risk. The DCC model seems to follow the portfolio return in a good way, following the bottom of the returns precisely, but risking to underestimate the actual risk. Furthermore, the BEKK has a more volatile pattern, not following the returns as precisely as the DCC, but maybe ending up with the best failure rate. These patterns are even further distinctive when looking at the 10-day ahead VaR figures, which are also found in **Appendix B**.

The portfolio based on GARCH models seems to really underestimate the risk, as well as the DCC model. The BEKK model on the other hand seems to better predict the critical value, but still, when the VaR estimate is exceeded, it does not seem to occur randomly. In order to determine if the models are significant, a Bernoulli trial and a Kupiec test are performed, which results are presented in **Tables 19-22**, that are found in **Appendix A**.

The results reveal that for the 1-day ahead VaR, the BEKK model is a significantly good estimator. Particularly at the 5% level it shows great accuracy. Both the univariate GARCH model and the DCC model has an intolerable failure rate and cannot be used as 1-day ahead VaR estimators. When looking at the 10-day ahead VaR, all three models show insignificant results while they all underestimate the risk. The worst model is the GARCH model with 89 out of 242 critical value violations. This clearly displays the loss when naively ignoring the correlation between assets.

7 Conclusions

The aim has been to investigate to what extent MGARCH models can forecast the conditional correlation of exchange rates. Two models were selected, the diagonal BEKK and the DCC. As an application, historical Value-at-Risk predictions have been evaluated. The historical simulation approach means that the future returns distribution is simulated based on the estimations for the chosen sample. The Value-at-Risk predictions have been conducted assuming a portfolio with two equally weighted long positions, but also one hedged portfolio in order to show diversification effects.

Furthermore, a univariate approach has been included for these Value-at-Risk forecasts in order to quantify the diversification effect due to the assets conditional covariance. The conditional correlation forecasts showed a similar pattern for the both models, when the sampling interval for the proxy, the realized correlation, became smaller the forecasts improved for all the three chosen evaluation measures. More intraday samples for the proxy simply means that the proxy moves towards the true correlation, improving the benchmark. The BEKK showed a better ability of forecasting, since the DCC tends to underestimate the conditional correlation.

The 1-day and 10-day ahead VaR were estimated. The BEKK model showed good estimation of expected loss and could therefore be used as a somewhat accurate measure of Value-at-Risk. When evaluating the 10-day ahead VaR, the models show estimating issues and they all end up as insignificant. Still, even though 10-day ahead is a more relevant horizon in practice, many portfolios are re-hedged more often than that, and therefore a 1-day ahead VaR perspective could be of interest.

Even though the BEKK model results in quite accurate VaR predictions, one need to consider the fact that during structural breaks, when accurate forecasts are most useful, they probably perform the worst. Obviously, it is necessary to not blindly rely on the models and instead be attentive to events with economic impact. **Section 4.4** contributed with a fundamental analysis of the structural break that has been seen.

To summarize the findings, the conclusions are that the forecasts can be considered as reasonable, but could most probably be even better if the intraday sample could be based on smaller intervals than every 30-minutes. The MGARCH models do contribute to better historical Value-at-Risk forecasts. When comparing the BEKK and DCC model, the BEKK seems to perform better than the DCC in both forecasting conditional correlation and predicting VaR. On the contrary, the BEKK is much more computationally demanding, which most certainly would be even more noticeable when the number of assets increase. More simple models may therefore often be preferred to MGARCH models, whereby the use of the MGARCH models is until today more interesting from a theoretical point of view rather than a practical one.

8 Further studies

One improvement that could be made, given that high frequency intraday data is available, is to determine the optimal sampling frequency and see to what extent the forecasts are improved. It may in fact be able to perform rather good predictions if it is possible to sample, for example, every fifth or tenth minute.

If the time series contain a longer period stable volatility, and then a structural break, the forecast values will be lagged compared to the actual values. A way to mitigate this could be a shorter estimation period, and it could be interesting to investigate this issue further by trying different lengths or maybe using an unconditional approach to forecast correlation.

Another interesting idea that has emerged during the work of this paper, but that could not be included, is to use MGARCH models to estimate a conditional time-varying beta in the Capital Asset Pricing Model (CAPM) and to evaluate how good these estimates would be. For example Faff et al. (2000) find that the market model beta to be as good as the conditional time-varying beta, that can be obtained through different MGARCH models. The aim of a possible research to compare the accuracy between models, or if the MGARCH approach tends to be more suitable for certain sectors.

References

Articles:

Anderson, Torben G. & Bollerslev, Tim (1998) Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts, *International Economic Review*, vol. 39, No. 4, November 1998

Andersen, Torben G., Bollerslev, Tim, Christoffersen Peter F. & Diebold, Francis X. (2007), Practical Volatility and Correlation Modeling for Financial Market Risk Management, National Bureau of Economic Research

Baringhaus, L. & Franz, C. (2010), Rigid Motion Invariant Two-Sample Tests, *Statistica Sinica* 20, 1333-1361

Bandi, Federico M., & Russell, Jeffrey R. Russell (2007) Microstructure Noise, Realized Variance, and Optimal Sampling

Bollerslev, Tim, (1990), Modelling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model, *The Review of Economics and Statistics*, MIT Press, vol. 72(3), pages 498-505, August

Bollerslev, T. & J.M. Wooldridge (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, *Econometric Reviews* 1, 143-173

Bollerslev, T., R.F. Engle & J.M. Wooldridge, (1988), A Capital Asset Pricing Model with Time-Varying Covariances, *Journal of Political Economy* 96, 116 -131

Brockwell, Peter J. & Davis, Richard A. (1991), *Time series: theory and methods*. 2. ed. New York: Springer-Vlg

Brooks, Chris (2008), *Introductory Econometrics for Finance*, Second Edition, Cambridge University Press

Engle, Robert F. & Sheppard, Kevin (2001), Correlation Multivariate GARCH : Theoretical and Empirical properties of Dynamic Conditional, NBER Working Paper No. 8554 Issued in October 2001

Engle, Robert F. (1982), Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation, *Econometrica* 50, 987-1008

Eugene F. Fama (1970), Efficient Capital Markets: A Review of Theory and Empirical Work, *The Journal of Finance*, Vol. 25, No. 2, Papers and Proceedings of the Twenty-Eighth Annual Meeting of the American Finance Association New York, N.Y.

December, 28-30, 1969 (May, 1970), pp. 383-417

Faff, R., Hillier, D., & Hillier, J. (2000). Time varying beta risk: An analysis of alternative modelling techniques. *Journal of Business Finance & Accounting*, 27(5- 6):523-554.

Kupiec, P. (1995), Techniques for Verifying the Accuracy of Risk Measurement Models, *Journal of Derivatives*

Jorion, Philippe (2007), *Value at risk: the new benchmark for managing financial risk*. 3. ed. New York: McGraw-Hill

O'Hara, Maureen (1995), Overview: Market structure issues in market liquidity, BIS Papers No 2

Orskaug, Elisabeth (2009), Multivariate DCC-GARCH Model - With Various Error Distributions, Norwegian Computing Center

Pojarliev, Momtchil & Polasek, Wolfgang (2003), Portfolio Construction by Volatility Forecasts: Does the covariance structure matter?, Swiss Society for Financial Market Research (pp. 103-116)

Grossman, Sanford J. & Stiglitz Joseph E. (1980), On the Impossibility of Informationally Efficient Markets, *The American Economic Review*, Volume 70, Issue 3, 393-408

Sentana, Enrique, (1998), The relation between conditionally heteroskedastic factor models and factor GARCH models, *Econometrics Journal*, Royal Economic Society, vol. 1(RegularPa), pages 1-9

Silvennoinen, Annastiina & Teräsvirta, Timo (2008), Multivariate GARCH models, CREATES Research Papers 2008-06, University of Aarhus

Suliman, Zakaria & Suliman, Abdalla (2011), Modelling Exchange Rate Volatility using GARCH Models: Empirical Evidence from Arab Countries, *International Journal of Economics and Finance*, Vol. 4, No. 3, March 2012

Sveriges Riksbank (2008), Penningpolitisk rapport 2008:3

Teräsvirta, Timo (2006), An Introduction to Univariate GARCH Models, SSE/EFI Working Paper in Economics and Finance, No. 646

Electronic sources:

Reider, Rob (2009), Volatility Forecasting I: GARCH Models
URL: http://cims.nyu.edu/~almgren/timeseries/Vol_Forecast1.pdf
(downloaded 25-04-2013)

A Appendix - Tables

Lag	EUR/SEK		USD/SEK	
	LM	<i>p</i>	LM	<i>p</i>
1	115,5377	0,0000	134,8121	0,0000
2	202,3768	0,0000	197,4754	0,0000
3	291,0325	0,0000	280,0207	0,0000
4	315,1721	0,0000	328,8676	0,0000
5	356,6215	0,0000	345,9473	0,0000

Table 13: ARCH LM test of return series

	LB	<i>p</i>
EUR/SEK	28,527	0,098
USD/SEK	22,439	0,317

Table 14: Ljung Box test for BEKK

	LB	<i>p</i>
EUR/SEK	27,093	0,133
USD/SEK	21,008	0,397

Table 15: Ljung Box test for DCC

Model	Simulated <i>CV</i>	<i>T</i>	<i>p</i>
BEKK	1,4710	1,8400	0,0260
DCC	1,2856	1,8802	0,0150

Table 16: Results of the Baringhaus-Franz test

Lag	EUR/SEK		USD/SEK	
	LM	<i>p</i>	LM	<i>p</i>
1	0,051	0,821	0,737	0,390
2	0,252	0,881	2,731	0,255
3	0,403	0,939	2,799	0,423
4	0,714	0,949	3,729	0,443
5	1,681	0,891	4,360	0,498

Table 17: ARCH LM test for BEKK

Lag	EUR/SEK		USD/SEK	
	LM	<i>p</i>	LM	<i>p</i>
1	1,059	0,303	1,880	0,170
2	2,080	0,353	2,757	0,251
3	2,222	0,527	2,958	0,398
4	2,179	0,702	3,435	0,487
5	2,263	0,811	3,950	0,556

Table 18: ARCH LM test for DCC

1-day VaR				
Model	5%		1%	
	<i>LR</i>	<i>p</i>	<i>LR</i>	<i>p</i>
GARCH	76,49	0,0000	81,60	0,0000
BEKK	0,12	0,7282	2,93	0,0871
DCC	12,77	0,0003	38,12	0,0000

Table 19: 1-day Kupiec test results

10-day VaR				
Model	5%		1%	
	<i>LR</i>	<i>p</i>	<i>LR</i>	<i>p</i>
GARCH	215,19	0,0000	222,06	0,0000
BEKK	6,49	0,0108	15,08	0,0001
DCC	48,82	0,0000	47,73	0,0000

Table 20: 10-day Kupiec test results

1-day VaR				
Model	5%		1%	
	<i>x</i>	<i>p</i>	<i>x</i>	<i>p</i>
GARCH	54	0,0000	26	0,0000
BEKK	19	0,0219	2	0,2558
DCC	39	0,0000	10	0,0002

Table 21: 1-day Bernoulli trial results

10-day VaR				
Model	5%		1%	
	<i>x</i>	<i>p</i>	<i>x</i>	<i>p</i>
GARCH	89	0,0000	52	0,0000
BEKK	23	0,0034	11	0,0001
DCC	44	0,0000	20	0,0000

Table 22: 10-day Bernoulli trial results

B Appendix - Figures

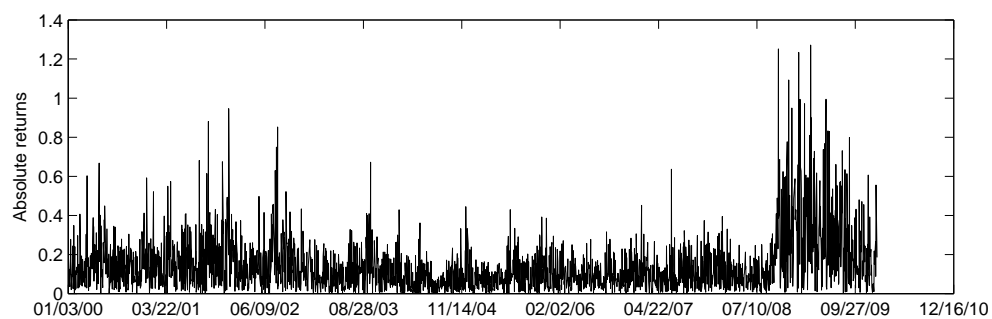


Figure 21: Absolute returns for EUR/SEK

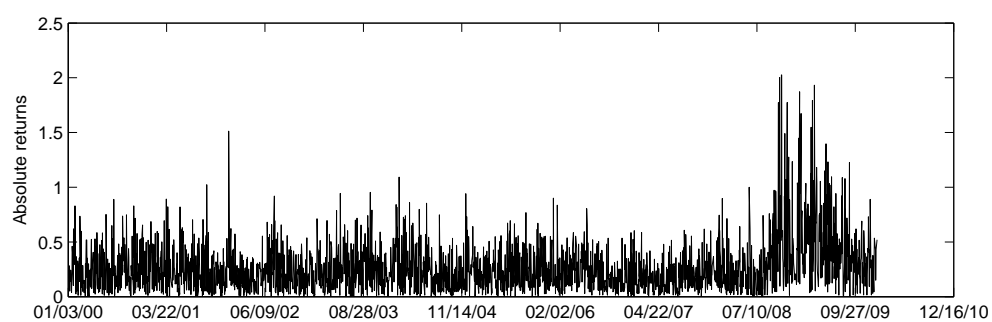


Figure 22: Absolute returns for USD/SEK

BEKK estimation: Standardized residuals

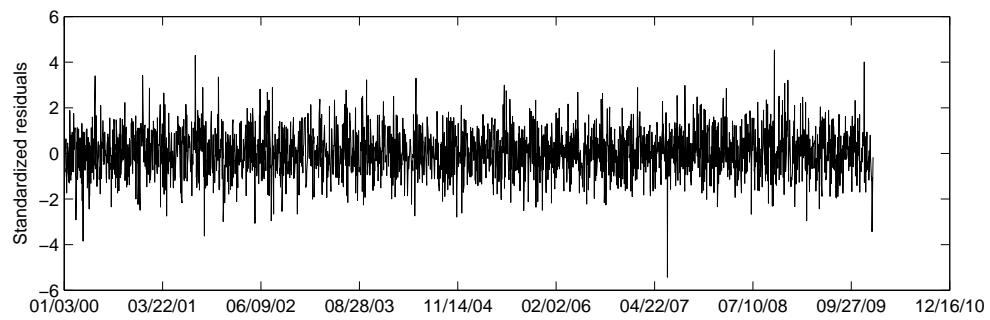


Figure 23: BEKK: Standardized residuals for EUR/SEK

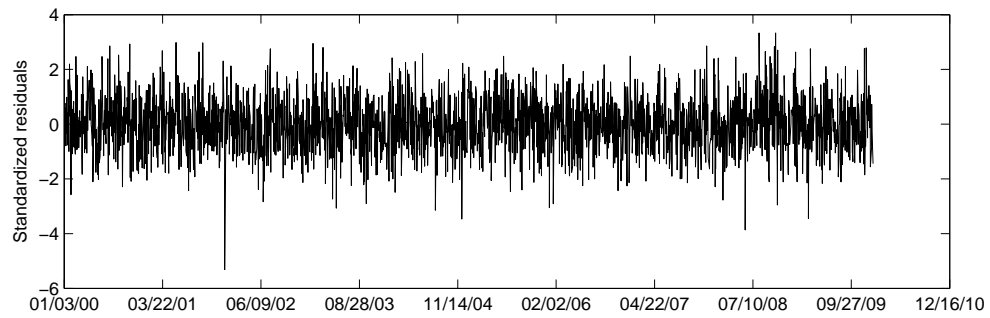


Figure 24: BEKK: Standardized residuals for USD/SEK

BEKK estimation: ACF and XCF

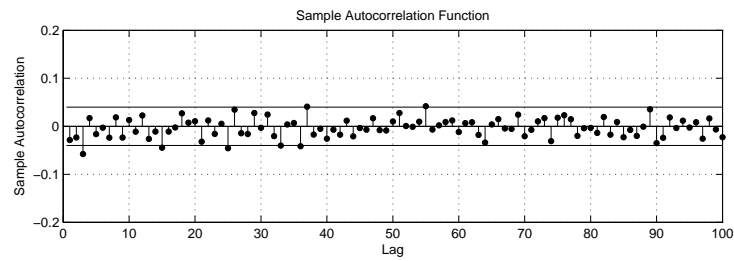


Figure 25: BEKK: ACF for EUR/SEK

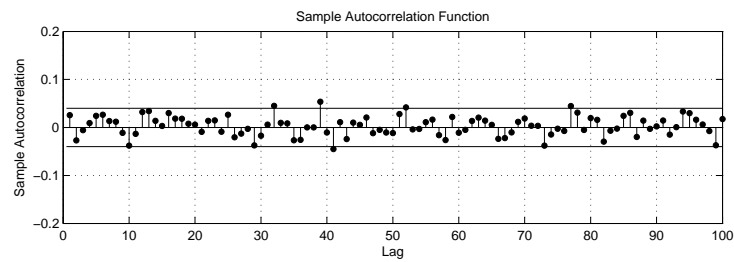


Figure 26: BEKK: ACF for USD/SEK

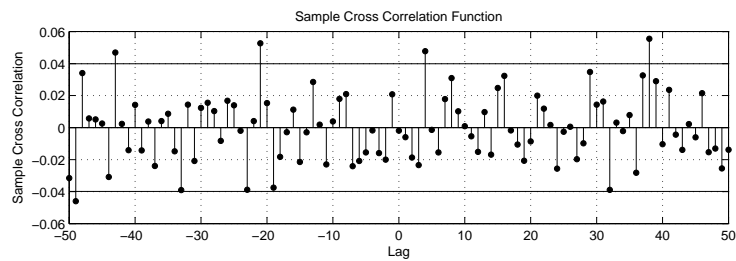


Figure 27: BEKK: XCF for the two series

DCC estimation: Standardized residuals

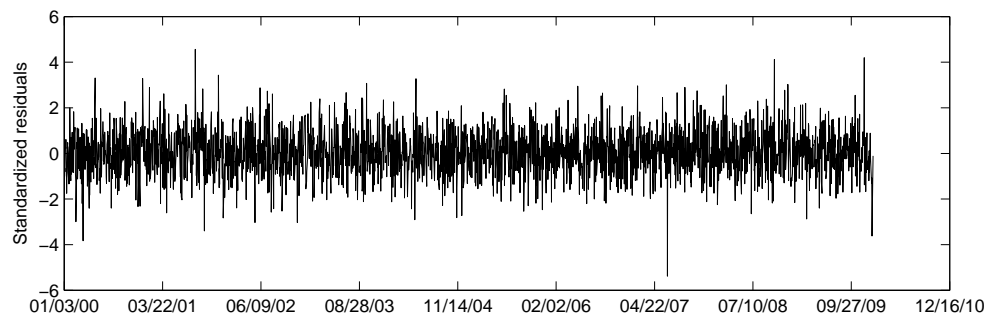


Figure 28: DCC: Standardized residuals for EUR/SEK

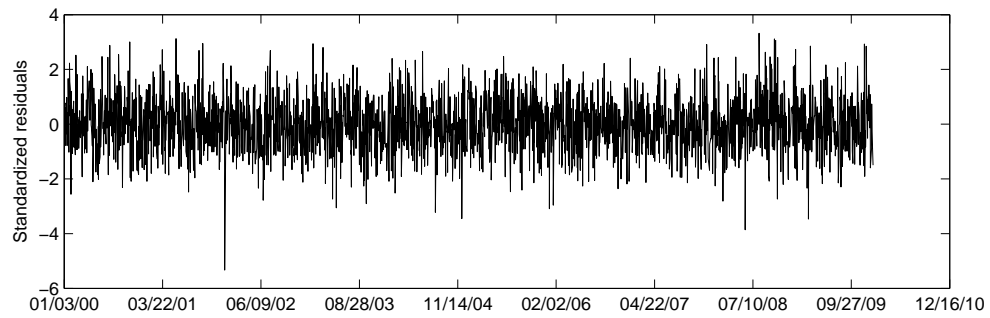


Figure 29: DCC: Standardized residuals for USD/SEK

DCC estimation: ACF and XCF

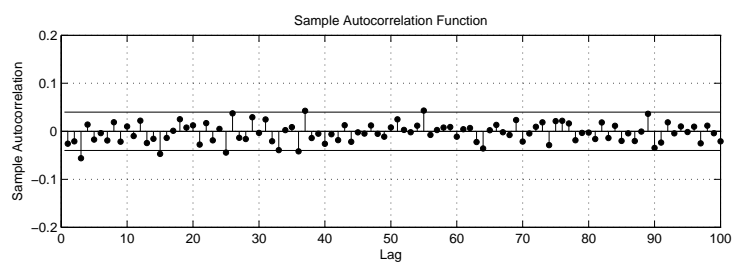


Figure 30: DCC: ACF for EUR/SEK

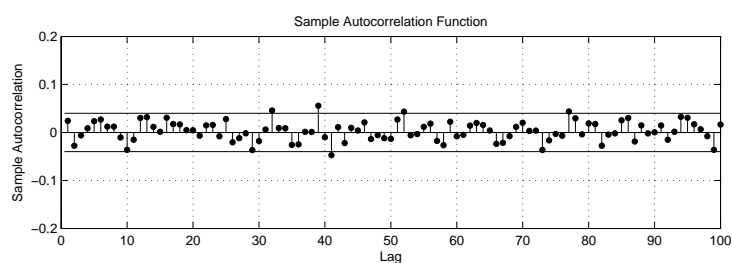


Figure 31: DCC: ACF for USD/SEK

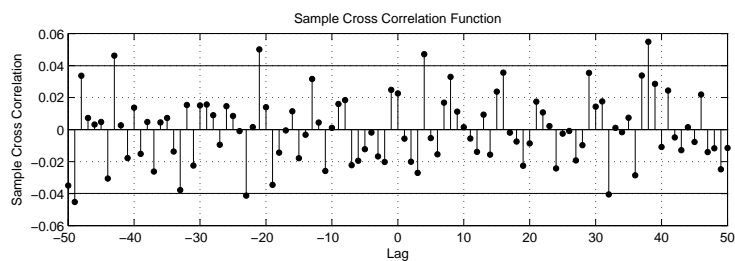


Figure 32: DCC: XCF for the two series

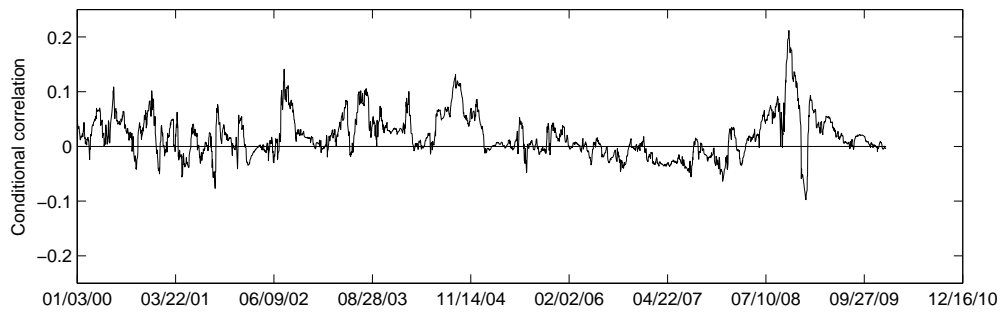


Figure 33: Spread between the BEKK and DCC estimated correlations

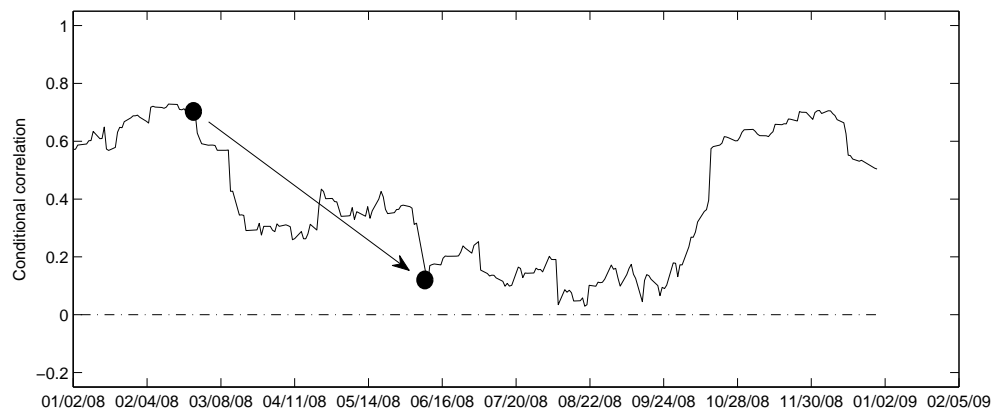


Figure 34: The DCC conditional correlation for the year 2008

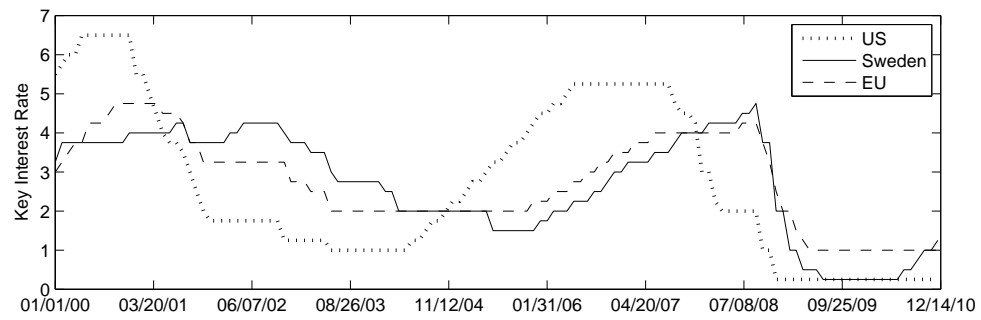


Figure 35: Key Interest Rates for the United States, Sweden and the Eurozone

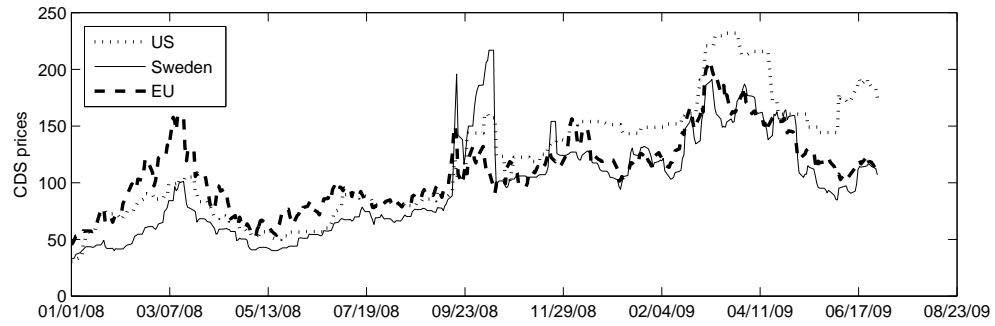


Figure 36: Domestic bank CDS prices for the United States, Sweden and the Eurozone

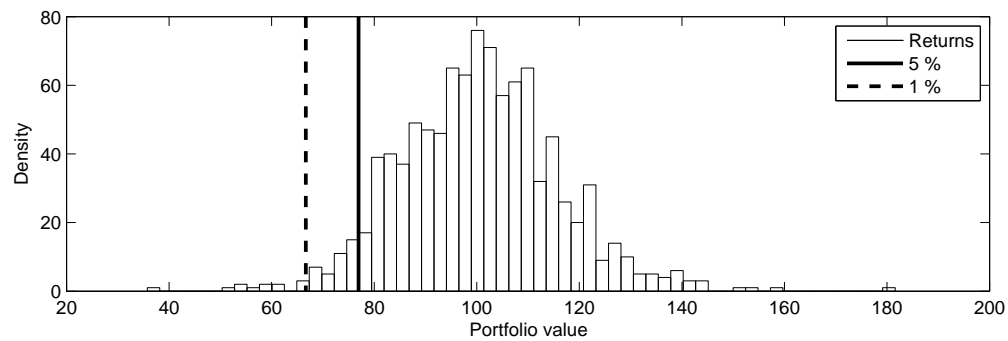


Figure 37: Simulated 1-day portfolio value distribution for GARCH(1,1)

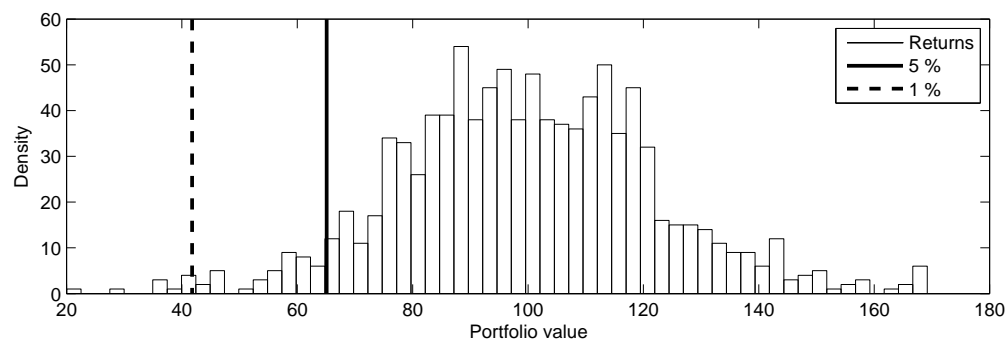


Figure 38: Simulated 1-day portfolio value distribution for BEKK(1,1)

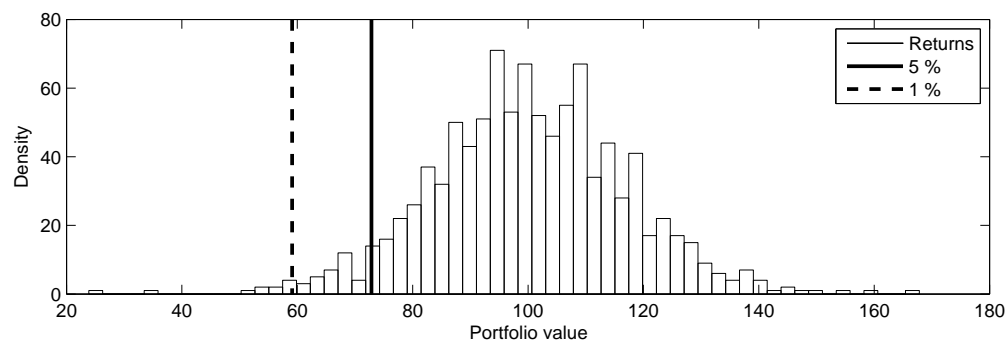


Figure 39: Simulated 1-day portfolio value distribution for DCC

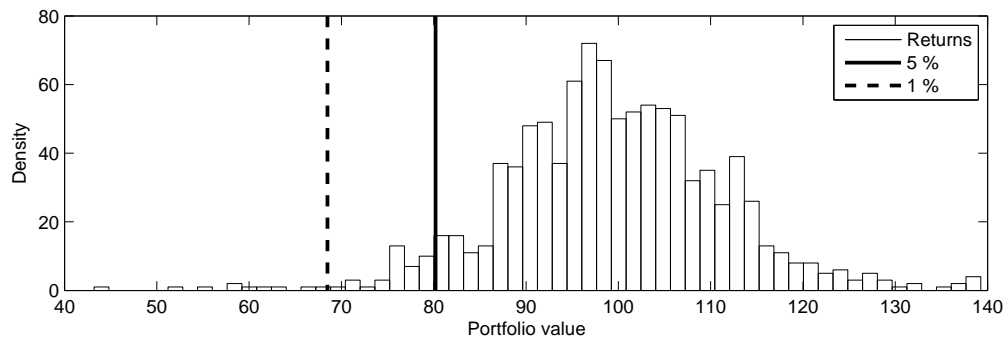


Figure 40: Simulated 1-day portfolio value distribution for GARCH, one long and one short position

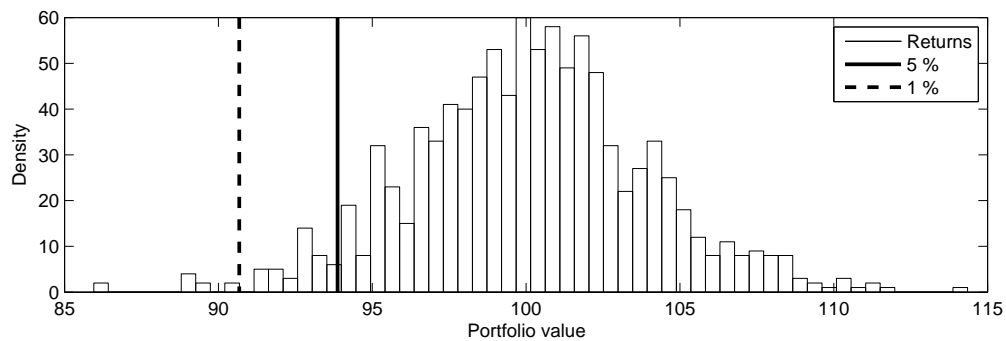


Figure 41: Simulated 1-day portfolio value distribution for BEKK, one long and one short position

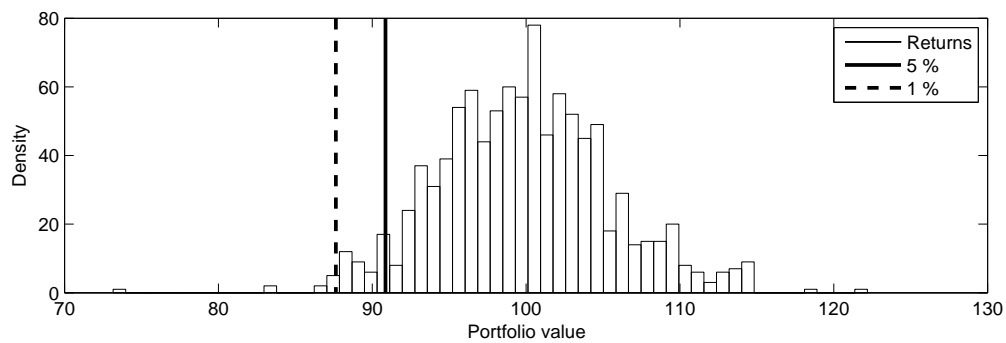


Figure 42: Simulated 1-day portfolio value distribution for DCC, one long and one short position

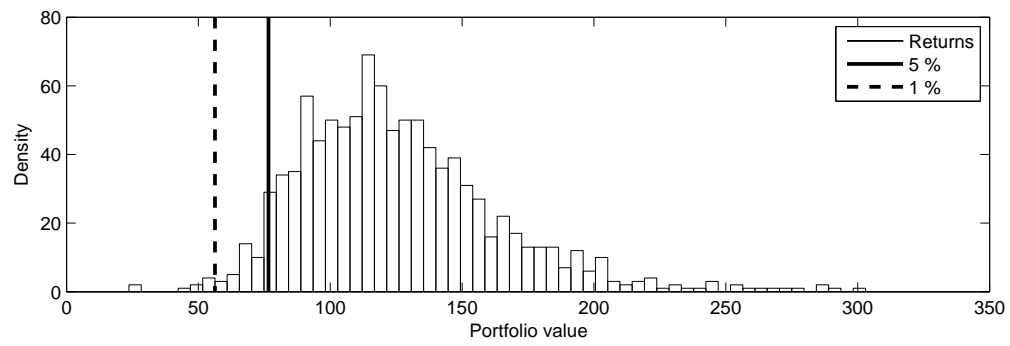


Figure 43: Simulated 10-day portfolio value distribution for GARCH(1,1)

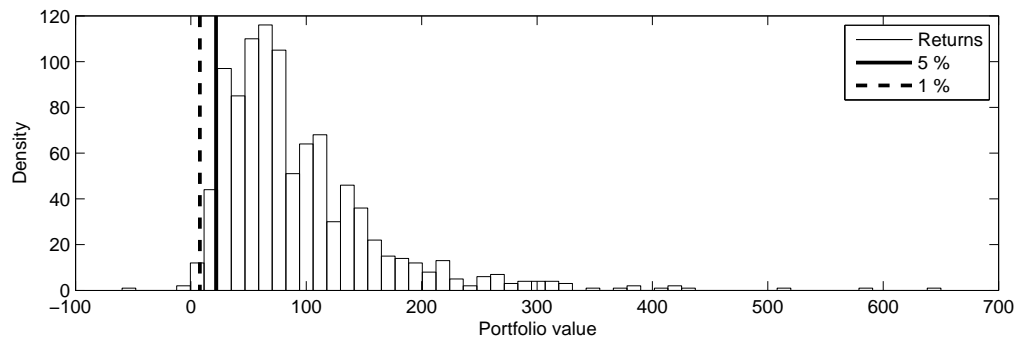


Figure 44: Simulated 10-day portfolio value distribution for BEKK(1,1)

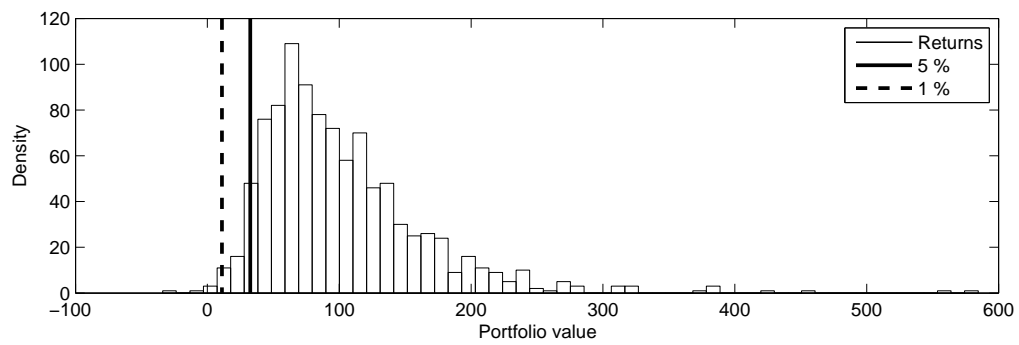


Figure 45: Simulated 10-day portfolio value distribution for DCC

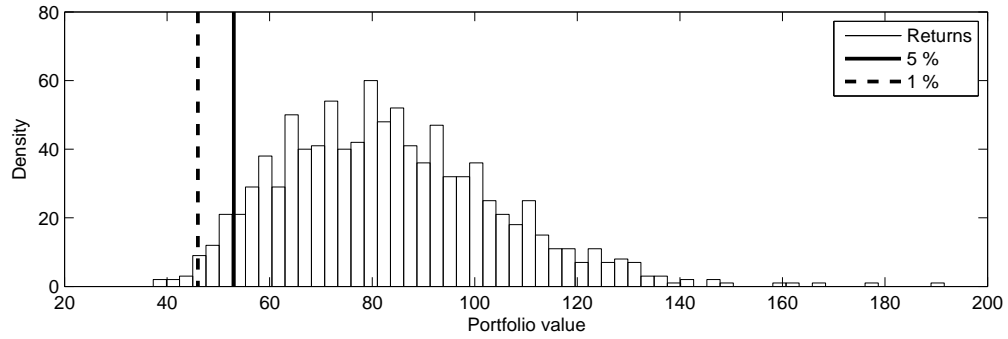


Figure 46: Simulated 10-day portfolio value distribution for GARCH, one long and one short position

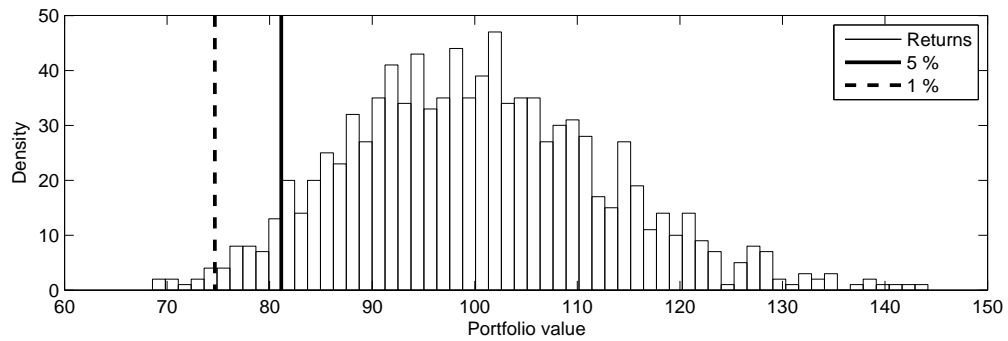


Figure 47: Simulated 10-day portfolio value distribution for BEKK, one long and one short position

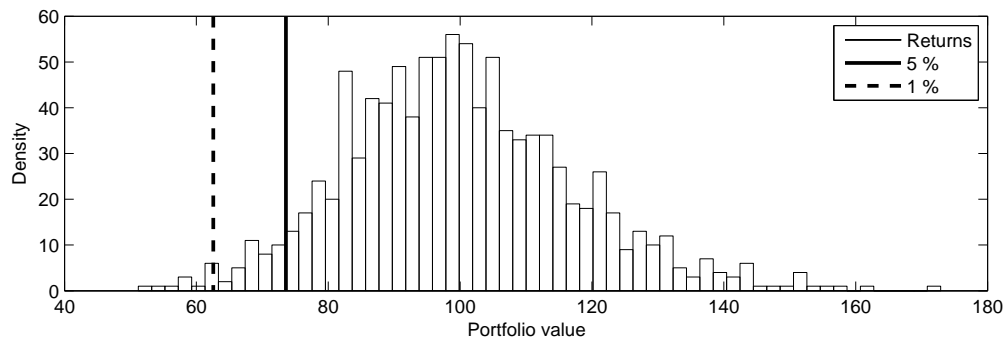


Figure 48: Simulated 10-day portfolio value distribution for DCC, one long and one short position

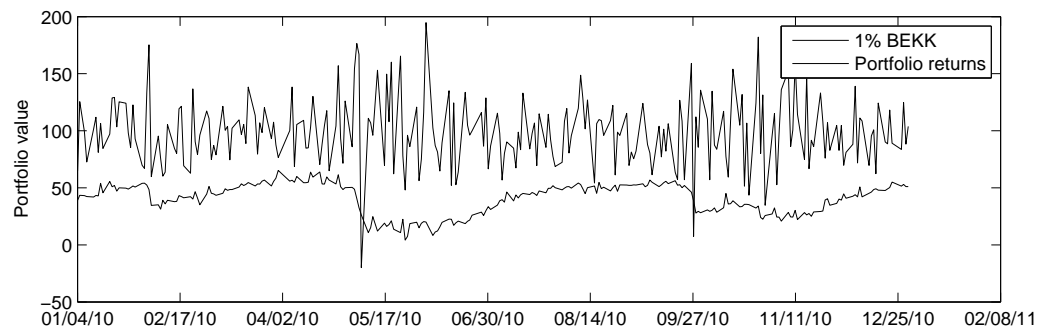


Figure 49: Simulated 1-day portfolio critical value forecast at 1 % for BEKK

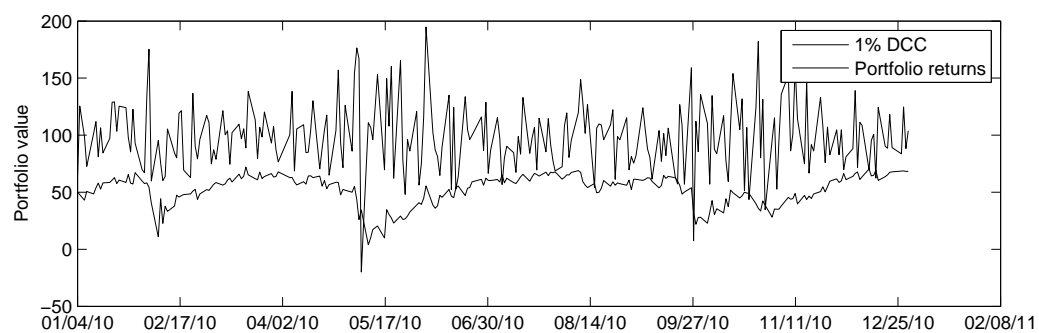


Figure 50: Simulated 1-day portfolio critical value forecast at 1 % for DCC

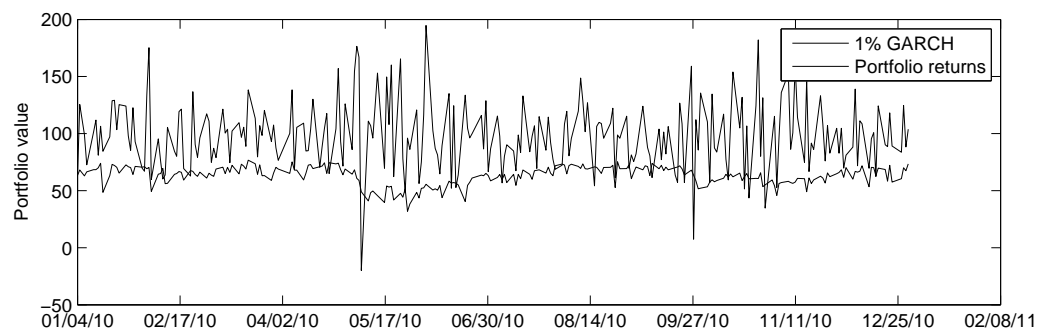


Figure 51: Simulated 1-day portfolio critical value forecast at 1 % for GARCH

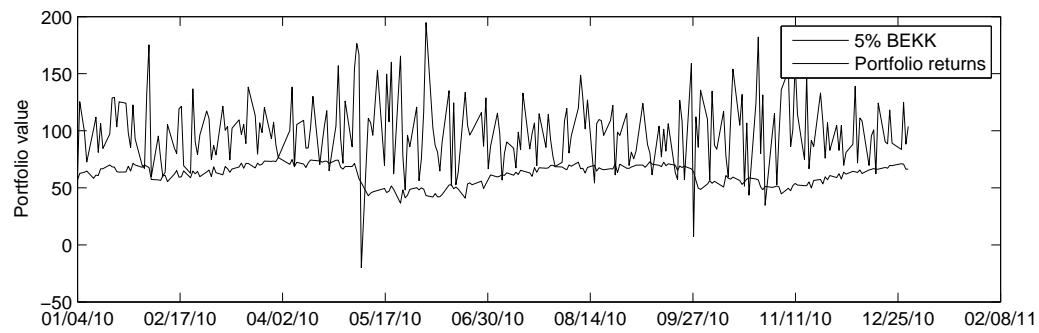


Figure 52: Simulated 1-day portfolio critical value forecast at 5 % for BEKK

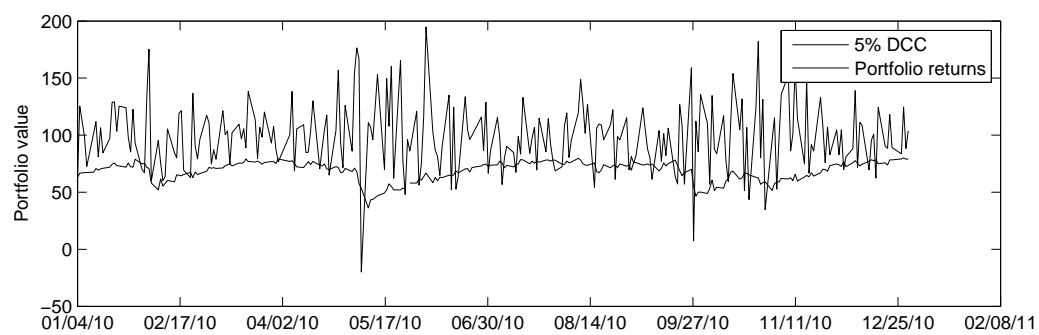


Figure 53: Simulated 1-day portfolio critical value forecast at 5 % for DCC

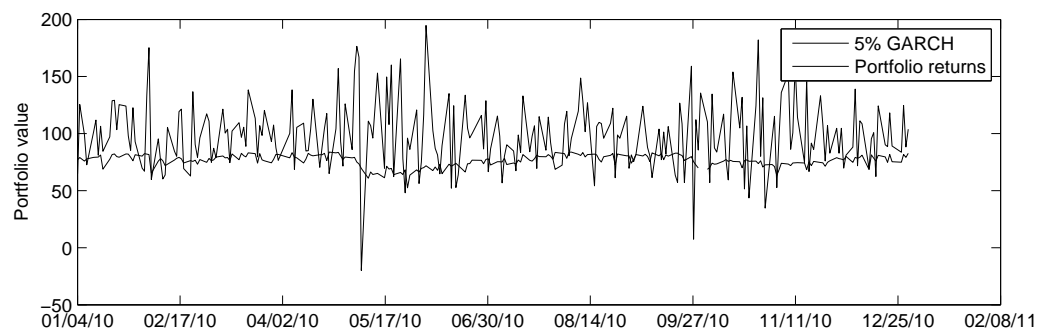


Figure 54: Simulated 1-day portfolio critical value forecast at 5 % for GARCH

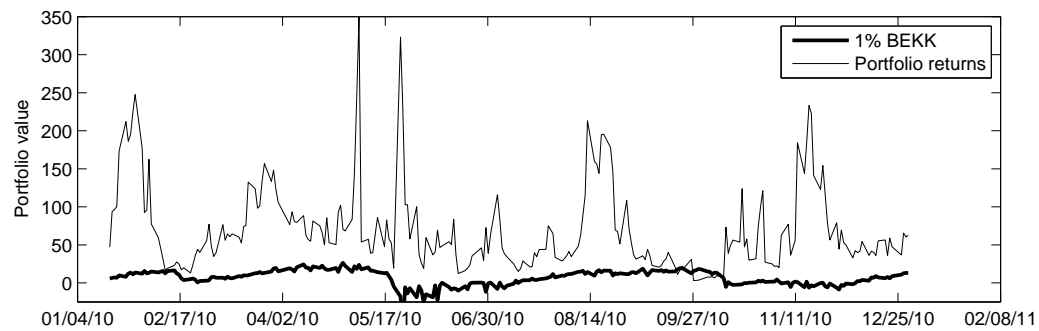


Figure 55: Simulated 10-day portfolio critical value forecast at 1 % for BEKK

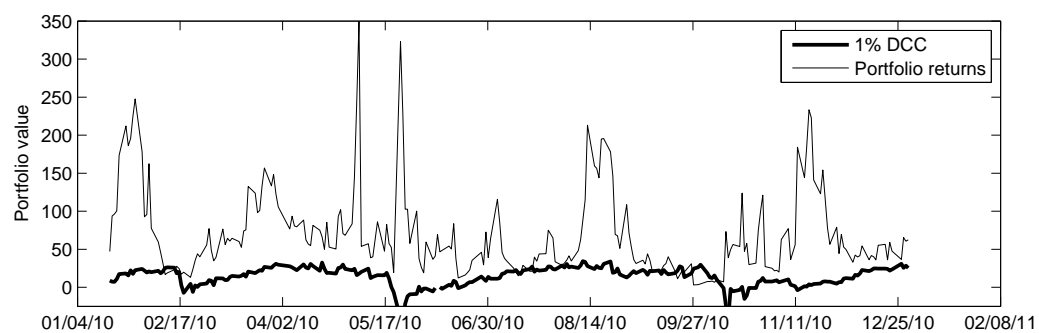


Figure 56: Simulated 10-day portfolio critical value forecast at 1 % for DCC

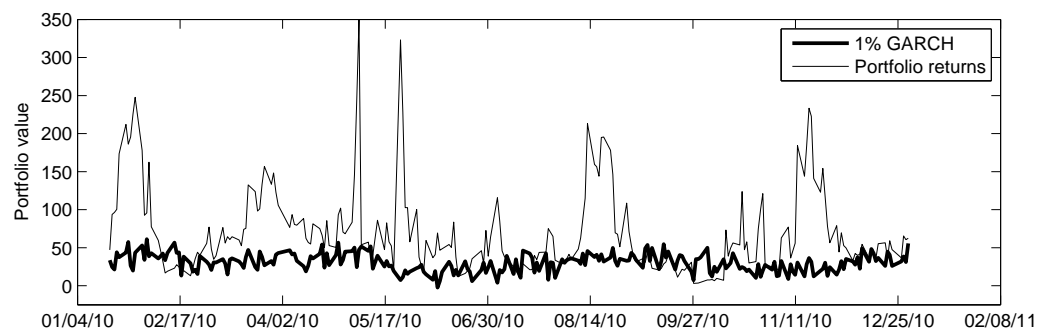


Figure 57: Simulated 10-day portfolio critical value forecast at 1 % for GARCH

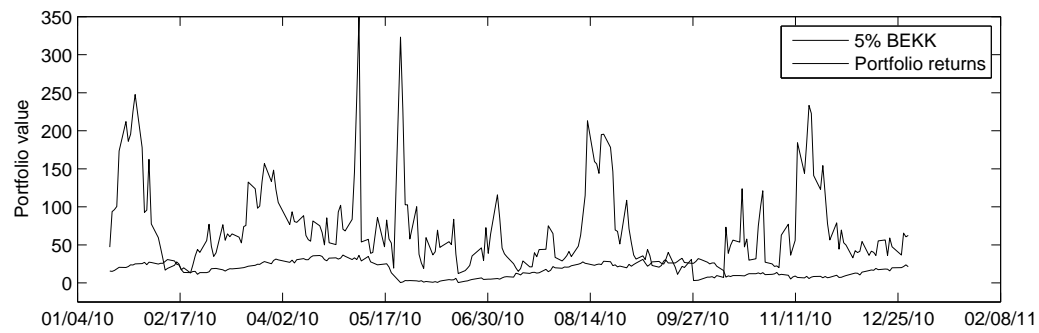


Figure 58: Simulated 10-day portfolio critical value forecast at 5 % for BEKK

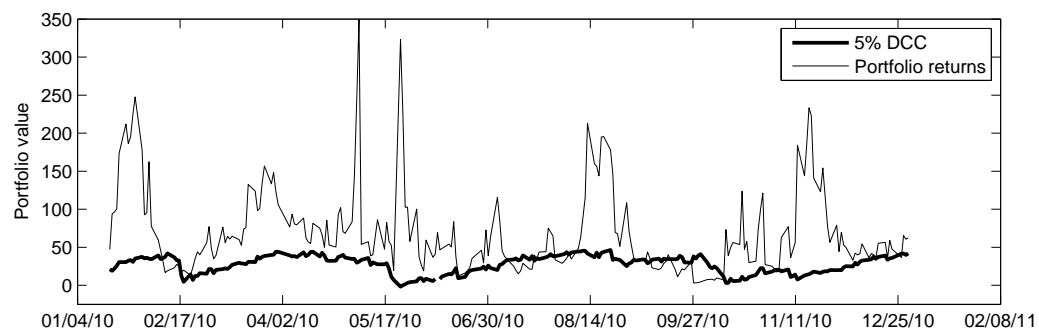


Figure 59: Simulated 10-day portfolio critical value forecast at 5 % for DCC

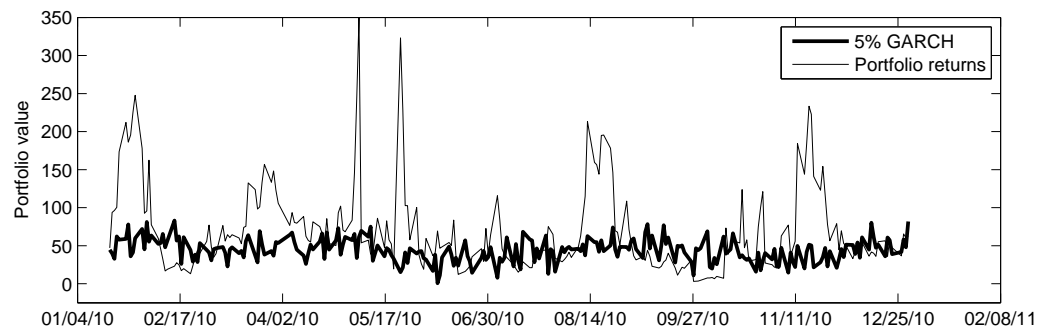


Figure 60: Simulated 10-day portfolio critical value forecast at 5 % for GARCH