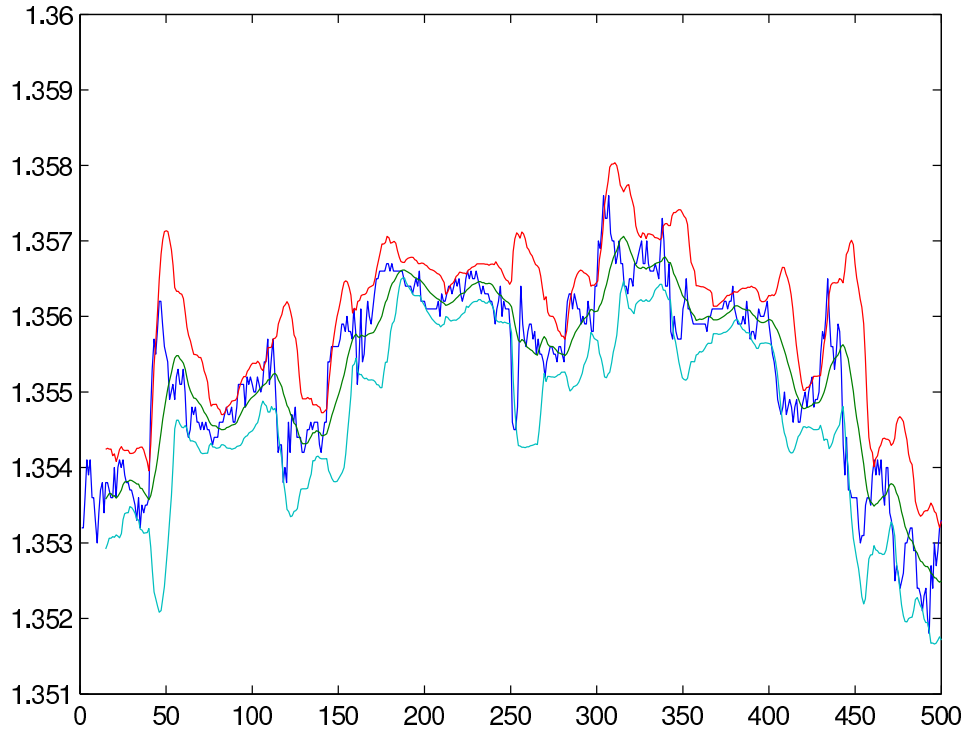
CHALMERS



ARIMA Modeling and Simulation of Currency Pairs

Master's Thesis in Engineering Mathematics and Computational Science

KRISTINA BERNDTSSON

Department of Mathematical Sciences
Chalmers University of Technology
Gothenburg, Sweden 2014

Abstract

In this thesis the currency pairs *USDCHF*, *EURUSD* and *EURSEK* are examined. The aim is to develop a model that describes the pairs in a gratifying way. This has been done with an *ARIMA* model, the decision on this model was made after studying the stationarity or lack there of, of the currency pairs.

Further more the model are used to develop strategies for trading the currencies, when the signals to buy or sell should be fired. The model is also used to simulate values of the currency pairs. The coefficients of the model are simulated via a copula simulation.

Acknowledgements

A special thanks to Patrik Albin for his patients, supervision, and support throughout this project. I would also like to express my gratitude to Box, Jenkins and Reinsel for writing the book that has been in my bag, on the side of my computer, and on my night-stand during this project. Last I want to thank my partner, Fredrik, for his practical and emotional support during this time.

Kristina Berndtsson, Gothenburg, November 2014

Contents

1 Introduction **1**
- 1.1 Purpose and research questions . . . 1
- 1.2 Thesis disposition . . . 2

2 The Data **3**

3 Theoretical Background **5**
- 3.1 Lag operators . . . 5
  - 3.1.1 Difference operator . . . 6
- 3.2 Stationary time series . . . 6
  - 3.2.1 Strictly stationary . . . 6
  - 3.2.2 Weak-sense stationary . . . 6
  - 3.2.3 Stationary and invertibility conditions for a linear process . . . 7
- 3.3 Modelling time series . . . 7
  - 3.3.1 Autoregressive–moving-average model . . . 8
  - 3.3.2 Autoregressive integrated moving average model . . . 10
- 3.4 Degree of the model . . . 13
  - 3.4.1 Autocorrelation function . . . 13
  - 3.4.2 Akaike information criterion . . . 14
- 3.5 Parameter estimation . . . 14
- 3.6 Godness of fit . . . 15
  - 3.6.1 Portmanteau lack-of-fit test . . . 15
- 3.7 Forecasting . . . 16
  - 3.7.1 RMSD and MPE . . . 17
- 3.8 Simulation . . . 17
  - 3.8.1 Copula simulation . . . 17
  - 3.8.2 Multivariate normality test . . . 18

4 Method **19**
4.1 Treating the data . . . 19
4.2 Stationary or non-stationery? . . . 19
4.2.1 Estimating the ACF . . . 20
4.3 Determining the number of autoregressive and moving average terms . . . 20
4.3.1 Algorithm for choosing the degree of the model . . . 21
4.4 Parameter estimation . . . 21
4.4.1 Godness of fit . . . 21
4.5 Forecasting . . . 22
4.6 Simulation . . . 22
4.6.1 Checking for normality . . . 23
4.7 Buy and sell stratigies . . . 23
4.7.1 Strategy 1 . . . 23
4.7.2 Strategy 2 . . . 23

5 Results **25**
5.1 Investigating stationarity and the level of differencing . . . 25
5.2 Finding p and q . . . 25
5.2.1 Godness of fit . . . 29
5.3 Forecasting . . . 30
5.4 Buy and sell strategies . . . 33
5.4.1 Strategy 1 . . . 33
5.4.2 Strategy 2 . . . 33
5.5 Simulation . . . 34
5.5.1 Buy and sell strategies of the simulated data . . . 36

6 Conclusion **39**
6.1 Future Research . . . 40

Appendices **41**

A Results **43**

Bibliography **59**

1

Introduction

In finance, currency pairs describes the relative value of a currency against another currency in the foreign exchange market. This is a stochastic process in that sense that it represents the evolution of a system of random values over time. The aim of this project is to model the process for the purpose of understanding about the behaviour of currency pairs.

Different combinations of autoregressive and moving average models will be investigated to determine if there is one that can describe the currency pairs in a satisfactory way. The models that will be investigated are the ARMA, autoregressive–moving-average model for stationary processes, and the ARIMA, autoregressive-integrated-moving-average, for non-stationary processes.

A practical use of a model for currency rates are the ability to predict future values and from that make decision on whether to buy or sell the currency. And hence, making a profit from exchanging between the currencies in the pair.

If a successful model is found this will be used to try to simulate a currency over a longer period of time. This is an interesting application of the model since it can be used to test out different strategies in buying and selling the currency without having to collect real data.

1.1 Purpose and research questions

The main purpose of this work is to investigate the behaviour of the currency data by finding a model that describes it in a proper way. The model will be used to mimic the data for purposes such as simulation of the currency pairs. To achieve this, the following points are to be investigated in this thesis:

1. Finding the model
2. Investigating the fit of the model

3. Forecasting future values
4. Simulating currency pairs

1.2 Thesis disposition

In Chapter two the background on the data sets used in this work are presented. Chapter three contains the theoretical framework on which the thesis is based on, the method used is presented in Chapter four. The results of the work are presented in Chapter five and the conclusions are in Chapter six.

2

The Data

The object of this thesis is to examine the behavior of the currency pairs EURSEK, EURUSD and USDCHF. The data where collected from the the web site *FX street, The Forex Market* on www.fxstreet.com [5].

The choice of the currency pairs USDCHF, that is United States Dollars as base currency and Swiss Franc as second currency, and EURUSD, Euro as base and United States Dollars as second, are based on that the pairs belong to the so called *Majors*. That is two of the seven currency pairs that constitute about 85% of the foreign exchange market. EURSEK, Euro as base and Swedish Krona as second, have been chosen for the local connection.

The foreign exchange market begin trading 22:00 GMT on Sunday in Sydney and cease at 22:00 GMT on Friday in New York. The data used in this thesis are from 22:00 GMT Sunday to 18:55 GMT Friday, so three hours are missing from the end of the series. This is due to the limitations in the data that *FX street* makes available for public download. The data have a period of five minutes, which gives us 1404 points of data in one week.

The data downloaded from *FX street* are four values for each five minutes, the opening and closing value, and the highest and lowest value of the five minute interval. In this thesis just the closing values of the five minutes interval are used to represent the behaviour of the currencies.

The data where collected for twelve consecutive weeks, in the period 23 of June 2014 to 12 of September 2014, that is week 26 to week 37.

Currency pairs are nice to analyse since there is a whole week of continuous data instead of just one day, which is the case with for example stock prices. This is due to that the same currencies are traded on different markets around the world and are not bounded to a particular markets opening hours. Then data with a five minutes interval one week can be sufficient amount of information to draw conclusions from.

From this the decision to treat the data as individual weeks has been made. Events

during the weekends, when the markets are closed, can influence the price. This can be difficult to account for in a mathematical description of the data, and hence it seems sufficient to treat the weeks individually.

The currency market are the most liquid market in the world, this is interesting from a mathematical viewpoint. That is, the currency can be sold quickly without having to reduce the price; there are always ready and willing buyers and sellers. The high liquidity comes from the huge trade volumes; currencies represent the largest asset class in the world.

The foreign exchange market is close to representing the ideal of perfect competition, which makes for an interesting topic of study.

3

Theoretical Background

MODELLING TIME SERIES can be done by many different methods. Some are very simple and have the benefit of being fast and easy to use. Other models are more complicated but might be more accurate. Which ones to use, depend on the application of the time series. The models used, along side with other theoretical background needed, in this project are presented in this chapter. Further in this text $\check{X}$ will denote $X - \mu$, the deviation of the process about its mean.

3.1 Lag operators

Lag operators are used to simplify the notation of time series. The *backshift operator* operator B are defined as

$$X_t = BX_{t+1} \qquad \text{for all } t \leq 1,$$

and eqvivalently the *forward operatior*, F

$$FX_t = X_{t+1} \quad F = B^{-1}.$$

A polynomial of lag operators are written as

$$\phi(B) = 1 + \sum_{i=1}^{p} \phi_i B^i,$$

where the power of the lag operator are

$$B^k X_t = X_{t-k}.$$

3.1.1 Difference operator

The difference operator is a special case of lag polynomials

$$\nabla X_t = (1 - B)X_t = X_t - X_{t-1},$$

and higher order differences

$$\begin{aligned}\nabla^2 X_t =& \nabla(\nabla X_t) = \nabla X_t - \nabla X_{t-1} \\ =& X_t - X_{t-1} - (X_{t-1} - X_{t-2}) \\ =& X_t - 2X_{t-1} + X_{t-2} \\ =& \{\text{with lag operators}\} \\ =& (1 - B)^2 X_t,\end{aligned}$$

and in general

$$\nabla^i X_t = (1 - B)^i X_t = \sum_{m=0}^{i} (-1)^m \binom{i}{m} X_{t-m}.$$

3.2 Stationary time series

3.2.1 Strictly stationary

A time series $\{X_t\}$ is said to be *strictly stationary* if the distribution of the set $\{X_{t_1}, \ldots, X_{t_k}\}$ is identical to that of $\{X_{t_1+h}, \ldots, X_{t_k+h}\}$ for all h, (k positive integer and $t_1, \ldots, t_k$ a collection of positive integers) [6]. The joint distribution of $\{X_{t_1}, \ldots, X_{t_k}\}$ is invariant under time shift.

For a time series to be strictly stationary there can be no trends, neither in the mean values of the X_t, in their variances or in the relation between successive terms of the series.

3.2.2 Weak-sense stationary

A process $\{X_t\}$ is called weak-sense stationary or wide-sense stationary (WSS) if the expectations

$$E[X_s] \quad \text{and} \quad E[X_{s+t}X_s],$$

are well-defined for all s and t and do not depend on the value of s [6].

3.2.3 Stationary and invertibility conditions for a linear process

Considering a linear filter, whose input is white noise:

$$\begin{aligned}\check{X}_t &= a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j} \\ \check{X}_t &= \left(1 + \sum_{j=1}^{\infty} \psi_j B^j\right) a_t \\ \check{X}_t &= \psi(B) a_t.\end{aligned}$$

This is equivalent to representing the time series as

$$\begin{aligned}\check{X}_t &= \sum_{j=1}^{\infty} \pi_j \check{X}_{t-j} + a_t \\ \left(1 - \sum_{j=1}^{\infty} \pi_j B^j\right) \check{X}_t &= a_t \\ \pi(B)\check{X}_t &= a_t.\end{aligned}$$

$\psi(B)$ and $\pi(B)$ can be regarded as the *generating function* of the ψ and π weights, with B now treated simply as a variable whose jth power is the coefficient of ψ and π. The weights are related as

$$\pi(B) = \psi^{-1}(B).$$

Then the series is stationary if

$$\sum_{j=0}^{\infty} |\psi_j| < \infty,$$

or embodied in the condition that the generating function $\psi(B)$ must converge for $|B| \leq 1$, that is on ore within the unit circle. We shall also say that the series is invertible if the weights π_j are absolutely summable

$$\sum_{j=0}^{\infty} |\pi_j| < \infty.$$

3.3 Modelling time series

The methods used for modelling the time series of currency couples in this project are presented below together with the methods for estimating the unknown parameters of the models.

3.3.1 Autoregressive–moving-average model

The autoregressive-moving-average model, or ARMA for short, is a combination of two simple models, the moving-average and autoregressive models. The following is an excerpt from Tsay [11]

Moving-average model

The notation $MA(q)$ refers to the moving average model of order q:

$$X_t = \mu + a_t + \sum_{i=1}^{q} \theta_i a_{t-i},$$

written with lag polynomials

$$X_t = \mu + \sum_{i=0}^{q} \phi_i B^i a_t,$$

or eqvivalently

$$\check{X} = \theta(B) a_t.$$

Here μ is the mean of the series, $\theta_1, \ldots, \theta_q$ are the parameters of the model and $a_1, \ldots, a_{t-q}$ are white noise error terms assumed to be $a_t \sim$ i.i.d N(0,σ_a^2). This implies

1. $E[a_t] = E[a_t | a_{t-1}, a_{t-2}, \ldots] = 0$
2. $E[a_t a_{t-j}] = Cov(a_t, a_{t-j}) = 0$
3. $Var(a_t) = Var(a_t | a_{t-1}, a_{t-2}, \ldots) = \sigma_a^2$

Since the series

$$\psi(B) = \theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q,$$

is finite there are no restrictions on the parameters of the MA-process to ensure stationarity. To ensure invertibility he conditions are obtained from

$$a_t = \theta^{-1}(B) \check{X}.$$

Expanding

$$\theta(B) = \prod_{i=1}^{q} (1 - H_i B),$$

in partial fractions

$$\pi(B) = \theta^{-1}(B) = \sum_{i=1}^{q} \left(\frac{M_i}{1 - H_i B} \right),$$

shall then converge for the process to be invertible. Equivalently, the weights $\pi_j = -\sum_{i=1}^{q} M_i H_i^j$ are absolutely summable if $|H_i| < 1$, for $i = 1,2,\ldots,q$ Then since the roots of $\theta(B) = 0$ are H_i^{-1} it follows that if the roots of

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q = 0,$$

lie *outside* the unit circle the process is invertible.

Autoregressive model

Autoregressive model of order p, $AR(p)$,

$$X_t = c + a_t + \sum_{i=1}^{p} \phi_i X_{t-i},$$

with lag operators

$$X_t = c + a_t + \sum_{i=1}^{p} \phi_i B^i X_t,$$

or

$$\phi(B)\check{X}_t = a_t.$$

For the general $AR(p)$ process written as $\check{X}_t = \phi^{-1}(B)a_t$ we have

$$\phi(B) = (1 - G_1 B)(1 - G_2 B)\cdots(1 - G_p B),$$

where $G_i^{-1}, \ldots, G_p^{-1}$ are the roots of $\phi(B) = 0$. Expanding $\phi^{-1}(B)$ in partial fractions

$$\check{X}_t = \phi^{-1}(B)a_t = \sum_{i=1}^{p} \frac{K_i}{1 - G_i B} a_t.$$

Then for the $AR(p)$ to represent a stationary series $\psi(B) = \phi^{-1}(B)$ has to be a convergent series for $|B| \leq 1$, that is, the weights $\psi_j = \sum_{i=1}^{p} K_i G_i^j$ are to be absolutely summable, $|G_i| < 1$ for $i = 1,2,\ldots,p$. In conclusion the roots of $\phi(B) = 0$ must lie *outside* the unit circle.

Since the series

$$\pi(B) = \phi(B) = 1 - \phi_i B - \phi_2 B^2 - \cdots - \phi_p B^p,$$

is finite, no restrictions are required on the parameters of an AR-process to ensure invertibility.

ARMA

Given this the ARMA(p,q)-model is given by a combination of the MA(q) and AR(p) models

$$X_t = c + a_t + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=1}^{q} \theta_i a_{t-1}$$
$$\check{X}_t = \phi_1 \check{X}_{t-1} + \cdots + \phi_p \check{X}_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q},$$

and with lag operators

$$\left(1 - \sum_{i=1}^{p} \phi_i B^i\right) X_t = \left(1 + \sum_{i=1}^{q} \theta_i B^i\right) a_t,$$

or

$$\phi(B)\check{X}_t = \theta(B)a_t.$$

This will define a stationary process if, following the reasoning in Section 3.3.1, the characteristic equation $\phi(B) = 0$ has all its roots lying outside the unit circle. Similary, the roots of $\theta(B) = 0$ must lie outside the unit circle if the process is to be invertible.

The stationary and invertible ARMA(p,q) process can be represented as an infinite moving average process:

$$\check{X}_t = \psi(B)a_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}, \tag{3.1}$$

and an infinite autoregressive process

$$\pi(B)\check{X}_t = \check{X}_t - \sum_{j=1}^{\infty} \pi_j \check{X}_{t-j} = a_t,$$

where $\psi(B) = \phi^{-1}(B)\theta(B)$ and $\pi(B) = \theta^{-1}(B)\psi(B)$. The weights ψ_j and π_i are determind from the relations $\phi(B)\psi(B) = \theta(B)$ and $\theta(B)\pi(B) = \phi(B)$

$$\begin{aligned} \psi_j &= \phi_1\psi_{j-1} + \phi_2\psi_{j-2} + \cdots + \phi_p\psi_{j-p} - \theta_j \quad j > 0, \\ \pi_j &= \theta_1\pi_{j-1} + \theta_2\pi_{j-2} + \cdots + \theta_q\pi_{j-q} + \phi_j \quad j > 0, \end{aligned} \tag{3.2}$$

with $\psi_0 = 1$, $\pi_0 = -1$ and $\theta_j = 0$ for $j > q$, $\phi_j = 0$ for $j > p$.

3.3.2 Autoregressive integrated moving average model

Data can behave as through they have no fixed mean but still shows signs of homogeneity, in the sense that apart from local level and trend, one part of the series behaves much like any other part. To form a model that describes such homogeneous non-stationary behaviour can be obtained by an initial step of differencing can be applied to remove the non-stationarity. This entire section is an excerpt from Box, Jenkins and Reinsel [2] and Brockwell and Davis [4].

Definition

The definition of this process from [4]:

The ARIMA(p,d,q) Process 1. *If d is a non-negative integer, then $\{X_t\}$ is a said to be an ARIMA(*p,d,q*) process if $Y_t = (1-B)^d X_t$ is a causal ARMA(p,q) process.*

This means that the process $\{X_t\}$ satisfies a difference equation of the form

$$\phi^*(B)X_t = \phi(B)(1-B)^d X_t = \theta(B)a_t, \qquad \{a_t\} \sim N(0,\sigma^2),$$

where $\phi(z)$ and $\theta(z)$ are polynomials of degrees p and q respectively. Also $\phi(z) \neq 0$ for $|z| \leq 1$ and $\phi^*(z)$ has a zero of order d at $z = 1$, since the corresponding ARMA process

is stationary if the roots of $\phi(B) = 0$ lie outside the unit circle, and exhibits explosive non-stationary behaviour if the roots lie inside the unit circle.

The process $\{X_t\}$ is stationary if and only if $d = 0$, which reduces to an ARMA(p,q) process. The model can be written as

$$\phi(B)\nabla^d X_t = \theta(B)a_t. \tag{3.3}$$

Or equivalently defined by these two equations

$$\phi(B)w_t = \theta(B)a_t, \tag{3.4}$$

and

$$w_t = \nabla^d X_t. \tag{3.5}$$

Then we see that the process can be represented by a stationary, invertible ARIMA process on the dth difference of the series. For $d \geq 1$ inverting (3.5) gives

$$X_t = S^d w_t, \tag{3.6}$$

where

$$\begin{aligned} Sx_t &= \sum_{h=-\infty}^{t} x_h = (1 + B + B^2 + \dots)x_t \\ &= (1 - B)^{-1}x_t \\ &= \nabla^{-1}x_t. \end{aligned}$$

Thus

$$S = (1 - B)^{-1} = \nabla^{-1}.$$

The operator S^2 is similarly defined as

$$\begin{aligned} S^2x_t =& Sx_t + Sx_{t-1} + Sx_{t-2} + \dots \\ =& \sum_{i=-\infty}^{t} \sum_{h=-\infty}^{i} x_h \\ =& (1 + 2B + 3B^2 + \dots)x_t, \end{aligned}$$

and equivalently for higher-order d. Equation (3.6) implies that the process (3.3) can be obtained by summing (or integrating) the stationary process (3.4) d times. That is what the name of the model comes from.

Since the infinite summation operator $S = (1 - B)^{-1}$ dose not converge it can not be used to define the non-stationary ARIMA process. Instead we consider the finite operator S_m, for any positive integer m,

$$S_m = (1 + B + B^2 + \dots + B^{m-1}) \equiv \frac{1 - B^m}{1 - B},$$

and similary

$$\begin{aligned} S_m^{(2)} =& \sum_{j=0}^{m-1} \sum_{i=j}^{m-1} B^i \\ =& (1 + 2B + 3B^2 + \cdots + mB^{m-1}) \\ \equiv& \frac{1 - B^m - mB^m(1-B)}{(1-B)^2}, \end{aligned}$$

then $(1-B)S_m^{(2)} = S_m - mB^m$, and so on. Then the relation between X_t and w_t in terms of values back to some origin $k < t$ can be expressed as

$$X_t = \frac{S_{t-k}}{1 - B^{t-k}} w_t = \frac{1}{1 - B^{t-k}} (w_t + w_{t-1} + \cdots + w_{k+1}),$$

so that $X_t = w_t + w_{t-1} + \cdots + w_{k+1} + X_k$ can be thought of as the sum of a finite number of terms from the stationary process w plus an initializing value of the process X at time k. Hence in the formal definition of the ARIMA process one would need to specify initializing conditions for the process.

General Form of the Autoregressive Integrated Moving Average Process

In the general form of the ARIMA model a constant term is added

$$\phi(B)\nabla^d X_t = \theta_0 + \theta(B) a_t, \tag{3.7}$$

where

$$\begin{aligned} \phi(B) =& 1 - \phi_1 B - \phi_2 B^2 - \ldots \phi_p B^p \\ \theta(B) =& 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q. \end{aligned}$$

In what follows:

1. $\phi(B)$ is called the *autoregressive operator*; assumed to be stationary.
2. $\phi(B)\nabla^d$ is called the *generalized autoregressive operator*; non-stationary operator with d of the roots equal to unity.
3. $\theta(B)$ is called *moving average operator*; assumed to be invertible.

In allowing the constant term θ_0 to be nonzero the ARIMA process is capable of showing deterministic polynomial trend, of degree d. Since

$$E[w_t] = E[\nabla^d X_t] = \mu_w = \frac{\theta_0}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}.$$

For example when $d = 1$ a nonzero θ_0 allows for estimation of possible deterministic linear trend.

3.4 Degree of the model

In deciding the degree of the model, that is the values of p, d and q, we have used two different methods which are presented below. One for deciding the degree of differencing and another for the number of autoregressive and moving average parameters used.

3.4.1 Autocorrelation function

To find the degree of differencing, d in our model, a close study of the autocorrelation function, ACF, of the data can be used. Autocorrelation is the cross-correlation of a signal with it self, the measure of how much a the value of a series at time t depends on the values of the series at times before time t. The autocovariance at lag k, meaning the covariance between X_t and X_{t+k} is defined as

$$\gamma_k = \mathrm{Cov}(X_t, X_{t+k}) = E[(X_t - \mu)(X_{t+k} - \mu)].$$

Under the stationary assumption this must be the same for all t. The autocorrelation at lag k, that is the correlation between X_t and X_{t+k}, is then

$$\rho_k = \frac{E[(X_t - \mu)(z_{t+k} - \mu)]}{\sqrt{E[(X_t - \mu)^2]E[(X_{t+k} - \mu)^2]}} = \frac{\gamma_k}{\sigma_X^2}.$$

Since for a stationary process, the variance $\sigma_X^2 = \gamma_0$ is the same at time t as at time $t + k$ we have that

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \tag{3.8}$$

implying that $\rho_0 = 1$, which corresponds with intuition.

Autocorrelation function of a mixed process

The ACF of a mixed autoregressive-moving average model written as

$$\check{X}_t = \phi_1 \check{X}_{t-1} + \cdots + \phi_p \check{X} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q},$$

may be derived by multiplying by $\check{X}_{t-k}$ and taking expectation:

$$\gamma_k = \phi_1 \gamma_{k-1} + \cdots + \phi_p \gamma_{k-p} + \gamma_{Xa}(k) - \theta_1 \gamma_{Xa}(k-1) - \cdots - \theta_q \gamma_{Xa}(k-q),$$

where $\gamma_{Xa}(k) = E[\check{X}_{t-k} a_t]$, the cross-covariance function between $\check{X}$ and a. $\check{X}_{t-k}$ depends only on shocks that have occurred up to time $t - k$, then we have from (3.1), $\psi(B)a_{t-k} = \sum_{j=0}^{\infty} \psi_j a_{t-k-j}$ that

$$\gamma_{Xa}(k) = \begin{cases} 0 & k > 0 \\ \psi_{-k}\sigma_a^2 & k \leq 0 \end{cases}.$$

Hence the equation for γ_k may be expressed as

$$\gamma_k = \phi_1 \gamma_{k-1} + \cdots + \phi_p \gamma_{k-p} - \sigma_a^2(\theta_k \psi_0 + \theta_{k+1}\psi_1 + \ldots \theta_q \psi_{q-k}),$$

with the convention that $\theta_0 = 1$. (3.2) implies that

$$\gamma_k = \phi_1\gamma_{k-1} + \cdots + \phi_p\gamma_{k-p} \qquad k \geq q+1,$$

and hence

$$\rho = \phi_1\rho k - 1 + \cdots + \phi_p\rho k - p \qquad k \geq q+1,$$

or

$$\phi(B)\rho_k = 0.$$

3.4.2 Akaike information criterion

In model selection there is a trade off between adding more parameters to the model to achieve a better fit to the data and longer calculation times due to a more complex model. A measure of the relative quality of a statistical model in this sense is the Akaike information criterion, AIC, defined as

$$AIC = 2k - 2\ln(L), \tag{3.9}$$

from [1], where k is the number of parameters in the model and L is the maximized value of the likelihood function for the model.

The AIC should be used as a measure for comparing a set of candidate models; the model with the lowest AIC value should be the one preferred. Since increasing the number of parameters in the model almost always improves the goodness of fit the penalty for adding more parameters to be estimated is there to discourage over fitting.

3.5 Parameter estimation

Estimating the parameters in the model used for this thesis has been done by conditional maximum likelihood. Let $N = n + d$ original observations of a time series, where d is the degree of differentiating in the ARIMA model. Then the generated series $\mathbf{w}$ of n differences $w_1, w_2, ..., w_n$, $w_t = \nabla^d X_t$ transforms the problem from fitting the parameters $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ of the ARIMA model to fitting the same parameters to the w's in a stationary invertible ARMA(p,q) model, written as

$$\begin{aligned} a_t =& \tilde{w}_t - \phi_1\tilde{w}_{t-1} - \phi_2\tilde{w}_{t-2} - \cdots - \phi_p\tilde{w}_{t-p} \\ &+ \theta_1 a_{t-1} + \theta_2 a_{t-2} + \ldots \theta_q a_{t-q}, \end{aligned} \tag{3.10}$$

where $\tilde{w}_t = w_t - \mu$, with $E[w_t] = \mu$. μ can be estimated by $\bar{w} = \sum_{t=1}^{n} w_t/n$ or if desired (if the sample size is not big enough) μ may be included as an additional parameter to be estimated. Because of the difficulty if starting up the difference equation in (3.10) p starting values for the w's, $\mathbf{w}_*$, and q starting values for the a's, $\mathbf{a}_*$ must be given, hence the *conditional* maximum likelihood estimate.

Assuming that the a's in (3.3) are normally distributed, their probability density is

$$p(a_1,a_2,\ldots,a_n) \propto \sigma_a^{-n}\exp\left[-\left(\sum_{t=1}^{n}\frac{a_t^2}{2\sigma_a^2}\right)\right],$$

then the log-likelihood associated with the parameter values ($\boldsymbol{\phi}$,$\boldsymbol{\theta}$,σ_a), conditional on the choice of ($\mathbf{w}_*$,$\mathbf{a}_*$), would be

$$l_*(\boldsymbol{\phi},\boldsymbol{\theta},\sigma_a) = -n\ln(\sigma_a) - \frac{S_*(\boldsymbol{\phi},\boldsymbol{\theta})}{2\sigma_a^2},$$

the *sum of squares function*

$$S_*(\boldsymbol{\phi},\boldsymbol{\theta}) = \sum_{t=1}^{n} a_t^2(\boldsymbol{\phi},\boldsymbol{\theta}|\mathbf{w}_*,\mathbf{a}_*,\mathbf{w}),$$

from [2]

3.6 Godness of fit

When accessing how well the estimated model fits the data one useful approach is to examine the residuals of the model:

$$\hat{a}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)\tilde{w}_t,$$

where ($\hat{\boldsymbol{\phi}}$,$\hat{\boldsymbol{\theta}}$) are the maximum likelihood estimates of ($\boldsymbol{\phi}$,$\boldsymbol{\theta}$). The residuals can be computed recursively as

$$\hat{a}_t = \tilde{w}_t - \sum_{j=1}^{p}\hat{\phi}_j\tilde{w}_{t-j} + \sum_{j=1}^{q}\hat{\theta}_j\hat{a}_{t-j} \quad t = 1,2,\ldots,n,$$

using either zero initial values (conditional method) or back-forecasted initial values (exact method), then if the model is adequate

$$\hat{a}_t = a_t + O\left(\frac{1}{\sqrt{n}}\right).$$

3.6.1 Portmanteau lack-of-fit test

If the fitted model is appropriate if the modified Ljung-Box-Pierce statistic:

$$\tilde{Q} = n(n+2)\sum_{k=1}^{K}\frac{r_k^2(\hat{a})}{n-k},$$

from [3], is approximately distributed as $\chi^2(K-p-q)$. Where $r_k(\hat{a})$ are the estimated autocorrelations of $\hat{a}$.

3.7 Forecasting

This section is an excerpt from Box, Jenkins and Reinsel [2]. Forecasting l, $l \geq 1$ time steps into the future when standing at time t will be represented by X_{t+l}. That is said to be an forecast at *orgin* t for *lead-time* l. The generalized ARIMA process (3.7) will be represented as an infinite weighted sum of current and previous shocks

$$X_{t+l} = \sum_{j=0}^{\infty} \psi_j a_{t+l-j}, \tag{3.11}$$

where $\psi_0 = 1$ and the weights may be obtained by

$$\phi(B)(1 + \psi_1 B + \psi_2 B^2 + \dots) = \theta(B).$$

The forecast of X_{t+l} is denoted $\hat{X}_t(l)$. Suppose the best forecast is

$$\hat{X}_t(l) = \psi_l^* a_t + \psi_{l+1}^* a_{t-1} + \psi_{l+2}^* a_{t-2} + \dots,$$

where $\psi_l^*, \psi_{l+1}^*, \psi_{l+2}^*, \dots$ are to be determined. Then together with (3.11) the mean square error of the forecast is

$$E[X_{t+l} - \hat{X}_t(l)]^2 = (1 + \psi_1^2 + \dots + \psi_{l-1}^2)\sigma_a^2 + \sum_{j=0}^{\infty} (\psi_{l+j} - \psi_{l+j}^2)^2 \sigma_a^2,$$

which is then minimized by $\psi_{l+j}^* = \psi_{l+j}$. We then have

$$\begin{aligned} X_{t+l} =& (a_{t+l} + \psi_1 a_{t+l+1} + \dots + \psi_{l-1} a_{t+1}) + (\psi_l a_t + \psi_{l+1} a_{t-1} + \dots) \\ =& e_t(l) + \hat{X}_t(l), \end{aligned}$$

where $e_t(l)$ is the error function of the forecast $\hat{X}_t(l)$ at lead time l. Assuming that the $\{a_t\}$ are a sequence of independent random variables an thus $E[a_{t+j}|X_t, X_{t-1}, \dots] = 0$ for $j > 0$ a few conclusions are made:

1. $$\hat{X}_t(l) = \psi_l a_t + \psi_{l+1} a_{t-1} + \dots = E[X_{t+l}].$$

   Thus the minimum mean square error forecast at origin t, for lead-time l, is the conditional expectation of X_{t+l} at time t.

2. Since
   $$E[e_t(l)|X_t, X_{t-1}, \dots] = 0,$$
   the forecast is unbiased. Also the variance of the forecast error is
   $$V(l) = \mathrm{Var}\,(e_t(l)) = (1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{l-1}^2)\sigma_a^2.$$

3. The one-step-ahead forcast error is

$$e_1(l) = X_{t+1} - \hat{X}_t(1) = a_{t+1}.$$

In conclusion, denoting $E[X_{t+l}|X_t,X_{t-1},\dots]$ as $[X_{t+l}]$ and $E[a_{t+l}|X_t,X_{t-1},\dots]$ as $[a_{t+l}]$, the forecast for orgin t with lead time l is

$$[X_{t+l}] = \hat{X}_t(l) = [a_{t+l}] + \psi_1[at + l - 1] + \dots,$$

and on form we are use to

$$[X_{t+l}] = \hat{X}_t(l) = \phi_1[X_{t+l-1}] + \dots + \phi_{p+d}[X_{t+l-p-q}] - \theta_1[a_{t+l-1}] - \dots - \theta_q[a_{t+l-q}] + [a_{t+l}].$$

3.7.1 RMSD and MPE

To evaluate the forecast error of the model two calculation methods where used, the root-mean-square deviation, RMSD, and the mean percentage error, MPE. The RMDS are calculated as

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^{T}(\hat{X}_t - X_t)^2}{n}}.$$

The MPE is average of percentage errors:

$$\text{MPE} = \frac{100\%}{n}\sum_{t=1}^{T}\frac{\hat{X}_t - X_t}{X_t}.$$

3.8 Simulation

When real world data is hard to find or time consuming to gather, a simulation of the process is useful to test theories and understand behaviour. This is done in this thesis by the means of copula simulation.

3.8.1 Copula simulation

Copula 1. *[9] Let $\{X_1,X_2,\dots,X_d\}$ be a random vector with continuous margins: $F_i(x) = \mathbb{P}[X_i \leq x]$.*

$\{U_1,U_2,\dots,U_d\} = \{F_1(X_1),F_2(X_2),\dots,F_d(X_d)\}$ has then by the probability integral transformation*, uniformly distributed margins.*

The copula of $\{X_1,X_2,\dots,X_d\}$ is then defined as the joint cumulative distribution function of $(U_1,U_2,\dots,U_d)$, which is

$$C(u_1,u_2,\dots,u_d) = \mathbb{P}[U_1 \leq u_1, U_2 \leq u_2, \dots, U_d \leq u_d].$$

3.8.2 Multivariate normality test

Checking the simulated data for similarity to the multivariate normal distribution can be done by the means of the Mardia's test [8] which is based on multivariate extensions of skewness and kurtosis measures. For a sample of $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ of k-dimensional vectors

$$\widehat{\mathbf{\Sigma}} = \frac{1}{k}\sum_{j=1}^{k}(\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})^T$$

$$A = \frac{1}{6k}\sum_{i=1}^{k}\sum_{j=1}^{k}\left[(\mathbf{x}_i - \bar{\mathbf{x}})^T\,\widehat{\mathbf{\Sigma}}^{-1}(\mathbf{x}_j - \bar{\mathbf{x}})\right]^3$$

$$B = \sqrt{\frac{n}{8n(n+2)}}\left\{\frac{1}{k}\sum_{i=1}^{k}\left[(\mathbf{x}_i - \bar{\mathbf{x}})^T\,\widehat{\mathbf{\Sigma}}^{-1}(\mathbf{x}_i - \bar{\mathbf{x}})\right]^2 - n(n+2)\right\}.$$

With the null hypothesis of multivariate normality, $A \overset{a}{\sim} \chi^2$ with $\frac{1}{6}n(n+1)(n+2)$ degrees of freedom, and $B \overset{a}{\sim} N(0,1)$.

4

Method

This chapter describes the study approach used in this thesis. First an appropriate model for the data is found, this is done by investigate whether the data behaves like a stationary or non-stationary process and then looking at the AIC for different degrees of either ARMA or ARIMA models. The coefficients are estimated by the maximum likelihood method. A goodness-of-fit test is used to make a final decision on the degree of the model with the estimated coefficients. The finished model is then used for forecasting and simulation. The forecasting is used in one of the buy and sell strategies which are some what of a test of the usability of the model.

4.1 Treating the data

The data spans from 22:00 GMT Sundays to 18:55 GMT Fridays with a frequency of 5 minutes this gives us 1404 data points in a week. In this thesis the weeks are chosen to be treated separately. That is at the start of every week new coefficients are estimated and only after the initial estimation trading can start. Since the data have a period of five minutes it is possible to re-estimate the coefficients of the model before every new data point is collected. It has been found by trial and error that 400 data point is enough to estimate the model. This gives us about three and a half days left each week to trade.

4.2 Stationary or non-stationery?

The first thing to do when answering the question if a time series is stationary on non-stationary is to inspect the plot of the data, to see if a trend is present or not. It is always good to plot the raw data to get an idea of the behaviour of the process. A more general approach is to investigate the autocorrelation function of the process, how to estimate the ACF is shown in Section 4.2.1. In Section 3.4.1 it is shown that for an

ARMA(p,q) model the autocorrelation function satisfies

$$\phi(B)\rho_k = 0 \qquad k > q.$$

Writing $\phi(B) = \prod_{i=1}^{p}(1 - G_i B)$ the solution to this differencing equation, assuming distinct roots, take the form

$$\rho_k = A_1 G_i^k + A_2 G_2^k + \cdots + A_p G_p^k \qquad k > q - p,$$

the stationary requirements from Section 3.2.3 that the zeros of $\phi(B)$ lie outside the unit circle implies that the roots G_1,G_2,...,G_p lie inside the unit circle. If the process in question is stationary the autocorrelation function will then "die out" quickly for moderate and large k. If a single root, say G_1, approaches unity, $G_1 = 1 - \delta$, $\delta > 0$, the for large k

$$\rho_k \simeq A_1(1 - k\delta),$$

the autocorrelation function will fall of slowly. A similar argument may be applied if more then one of the roots approaches unity. The estimated ACF tends to behave in the same way as the theoretical autocorrelation function so a failure of the estimated ACF to die out rapidly will suggest a non-stationary process in X_t, but possibly as stationary in ∇X_t, or some higher difference. According to Box, et al. [2] in practice the degree of differencing is normally either 0, 1, or 2, and it is usually sufficient to inspect the first 20 or so estimated autocorrelations.

4.2.1 Estimating the ACF

Based on the data one can estimate the autocorrelation function, ρ_k (3.8) by

$$\begin{aligned} r_k &= \frac{c_k}{c_0} \\ c_k &= \frac{1}{N}\sum_{t=1}^{N-k}(X_t - \bar{X})(X_{t-k} - \bar{X}). \end{aligned}$$

4.3 Determining the number of autoregressive and moving average terms

If the data shows signs of non-stationary the ARIMA model will be chosen to represent the data. Then the decision of the order of the model, have to be made. The goal is to have one model for the three different currency pairs and a model that is complex enough to represent the currencies over a long time period. The model has been fitted for three consecutive weeks for the three different currency pairs.

4.3.1 Algorithm for choosing the degree of the model

Deciding the degree of the model, that is determine the values of p and q in (3.7), is done by the means of finding the minimum AIC, defined in (3.9), for different models. Since the model is re-estimated about 1000 times in the length of a week and a total of nine weeks where used to fit the model it was necessary to use a different technique then just estimating the model for all the different combinations of p and q. The algorithm used are formed by Hyndman and Khandakar [7] and are presented below:

Step 1. Start with four possible models:

- ARIMA(2,d,2)
- ARIMA(0,d,0)
- ARIMA(1,d,0)
- ARIMA(0,d,1)

If $d \leq 1$ then the model are fitted with $\theta_0 \neq 0$, otherwise $\theta_0 = 0$. The model with the smallest AIC value is the chosen and called the "current" model.

Step 2. Now consider a few variations from the current model.

- p or q is allowed to vary by ± 1 from the current model.
- p and q both vary by ± 1 from the current model.
- the constant θ_0 is included if the current model has $\theta_0 = 0$ or excluded if the current model has $\theta_0 \neq 0$.

Whenever a model with a lower AIC is found it becomes the new "current" model and step 2 is repeated. This process terminates when no new model has a lower AIC then the current model.

4.4 Parameter estimation

In tandem with the search for the values of p and q, the coefficients of the model will have to be estimated. This is done by the means of the maximum likelihood approach laid out in Section 3.5. This is done with the R function `arima` in the stats package. Where the exact likelihood is computed via a state-space representation of the ARIMA process. The innovations and their variance is found by the means of a Kalman filter. The complete algorithm can be found in *An Algorithm for the Exact Likelihood of a High-Order Autoregressive- Moving Average Process*, by J. G. Pearlman [10].

4.4.1 Godness of fit

The Lack-of-fit test from Section 3.6.1 is preformed for every set of estimated coefficients, that is about 1000 times for one week. The model with the most test statistic that point to that the null hypothesis will not be rejected, will be regarded as the best model.

4.5 Forecasting

When the model is complete it will be used to forecast the future behaviour of the currency pair. Since five minutes are enough time to re-estimate the coefficients and forecast the next value the one-step ahead forecast have been used in this thesis. Then the process is as described in Section 3.7:

$$\hat{X}(1) = \phi_1[X_t] + \cdots + \phi_{p+d}[X_{t+1-p-d}] - \theta_1[a_t] - \cdots - \theta_q[a_{t+1-q}] + [a_{t+1}],$$

where the brackets denote the conditional expectation at time t. Then assuming that data are available starting from time $s = 1$, the necessary a_s's are computed recursively from

$$a_s = X_s - \hat{X}_{s-1}(1) = X_s - \left(\sum_{j=1}^{p+q} \phi_j X_{s-j} - \sum_{j=1}^{q} \theta_j a_{s-j} \right) \quad s = p+d+1, \ldots, t,$$

setting initial a_s's equal to zero, for $s < p + d + 1$.

4.6 Simulation

Since data for currency couples are some what hard to collect it has been another focus of this work to try to use our ARIMA model to simulate data. To do this we need to simulated the, in previous Section estimated, coefficients in the ARIMA model, this is done by the means of a Copula simulation. We have $n = p+q$ coefficients dependent on time: $\{X_1(t), X_2(t), \ldots, X_n(t)\}$. Then for each $X_i(t)$ decide the empirical distribution

$$\tilde{F}(x) = \frac{\#x_i < x}{\#x_i}.$$

By definition 3.8.1 we then have that

$$Y_i(t) = \Phi^{-1}(\tilde{F}_i(X_i(t))) \quad i = 1,2,\ldots,n,$$

have a Gaussian distribution. Then we can model $\{\tilde{Y}_1(t), \tilde{Y}_2(t), \ldots, \tilde{Y}_n(t)\}$ as a Gaussian process with covariance matrix R, with dimension $T \times n$. The covariance matrix is estimated from $\{Y_1(t), Y_2(t), \ldots, Y_n(t)\}$, as

$$R_{ij} = \mathrm{Cov}\,(Y_i, Y_j) = \frac{1}{T} \sum_{k=1}^{T} (y_{i,k} - \hat{\mu_i})(y_{j,k} - \hat{\mu_j}).$$

With $\hat{\mu_i}$ the sample mean. Our model for the coefficients is then $\tilde{X}_i(t) = \Phi(\tilde{F}_i^{-1}(\tilde{Y}_i(t)))$ with $i = 1,2,\ldots n$.

To do the simulation, coefficients from all 12 weeks of data, from each currency pair separately, will be estimated and the covariance matrices calculated from the Y_i. The $\tilde{Y}_i$ will then be simulated from a Gaussian process with an average from all R.

The simulated coefficients are checked that they met the stability and invertibility conditions and those that do not are discarded. The remaining coefficients are then used to simulate an ARIMA(p,d,q) model with the last week of the data as starting values. The simulation is done with the innovations distributed as $N(0,\sigma^2)$, where σ is estimated from an average of the maximum likelihood estimated standard deviations of the data.

4.6.1 Checking for normality

Before modelling the $\tilde{X}_i(t)$, a test that shows if the simulated $\{\tilde{Y}_1(t), \tilde{Y}_2(t), \ldots, \tilde{Y}_n(t)\}$ really are from a Gaussian distribution have to be carried out. This are done with the Mardia's test for normality described in Section 3.8.2.

4.7 Buy and sell stratigies

We will implement two different buy and sell strategies to investigate the models suitability.

4.7.1 Strategy 1

This strategy is the most intuitive way of using our model to set up a buy and sell algorithm. Just looking at the next forecasted value and from that deciding on whether to buy or sell. Looking at the EURSEK pair:

- If $\hat{X}(1) = \{\text{up from the current value}\}$: Buy EUR
- If $\hat{X}(1) = \{\text{down from the current value}\}$: Buy SEK

4.7.2 Strategy 2

The second strategy builds on the theory of Bollinger Bands, that comparing the observed data with three bands:

- a middle band of a n-periodic simple moving average SMA
- an upper band of k times a n-periodic standard deviation, σ, above our middle band: $SMA + k\sigma$
- a lower band of k times a n-periodic standard deviation, σ, bellow our middle band: $SMA + k\sigma$

The simple moving average of period n is calculated as

$$SMA = \frac{X_t + X_{t-1} + \ldots X_{t-(n-1)}}{n}.$$

This strategy dose not use our ARIMA model but is there to have another strategy to evaluate our modelled data, that is based on the ARIMA model. The bands can be

used as indicators of overbought and oversold levels, the strategy would then be, in the example of EURSEK, to sell SEK (eq. buy EUR) when the price cross the upper band and sell SEK when the price cross the lower band.

5

Results

The result of the investigation made in this thesis are presented in this Chapter. The model found that represents the behaviour of the different currency pairs; the implementations of the model in buy and sell strategies, and the simulation of the currencies. The initial three weeks of the three currency pairs are use for to found the model, and then the rest of the data are used for the buy and sell part. The simulation is on the other hand based on all the data. Some of the calculations are made in MatLab but fore the most part the programming language R have been used to produce the results.

5.1 Investigating stationarity and the level of differencing

The plots for the data from week 26, 27, and 28 are shown in Figures A.1 and A.2 in the appendix. Over the course of one week there might be a trend present. To make sure that the process is non-stationary the ACF of the data are plotted in Figures 5.1, 5.2 and 5.3.

That the ACF dose not die out quickly is a sign of non-stationary. Hence the data are differentiated once and the ACF are calculated for the differentiated data, this is shown in Figures 5.4, 5.5 and 5.6. This looks a lot better, the currency pairs shows signs of non-stationarity on a week to week basis but this is redeemed by differencing once, thus an ARIMA(p,d,q) model with $d = 1$ will be used to represent the process.

5.2 Finding p and q

Following the algorithm in Section 4.3.1 to find the values for p and q such that the ARIMA(p,1,q) model gives minimum values for the AIC criteria. The result is presented in Table 5.1. We will investigate three different models, to see which one was the best result in the goodness-of-fit test. The models are: the combined result for USDCHF,

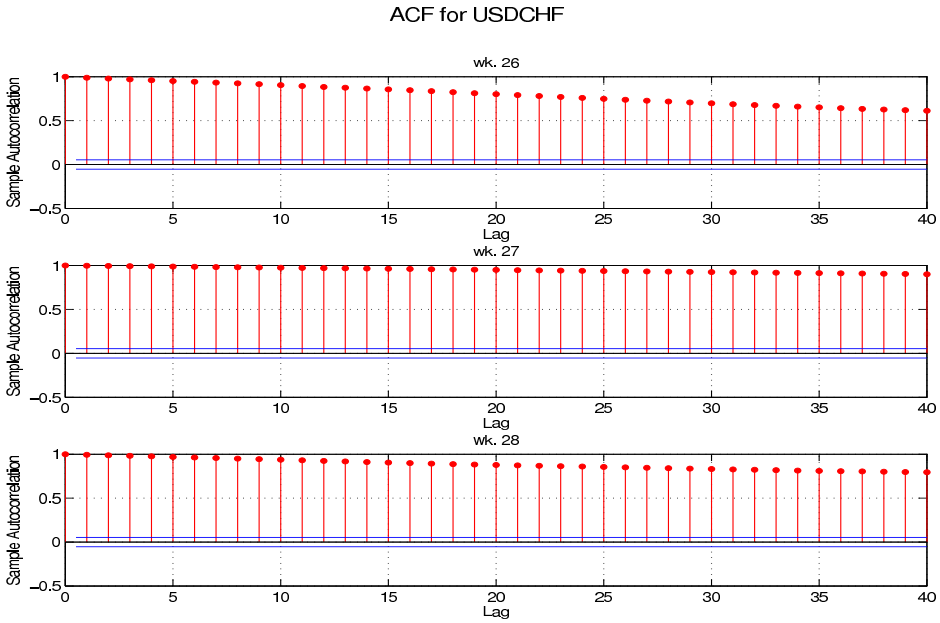


Figure 5.1: The ACF of USDCHF, undifferentiated

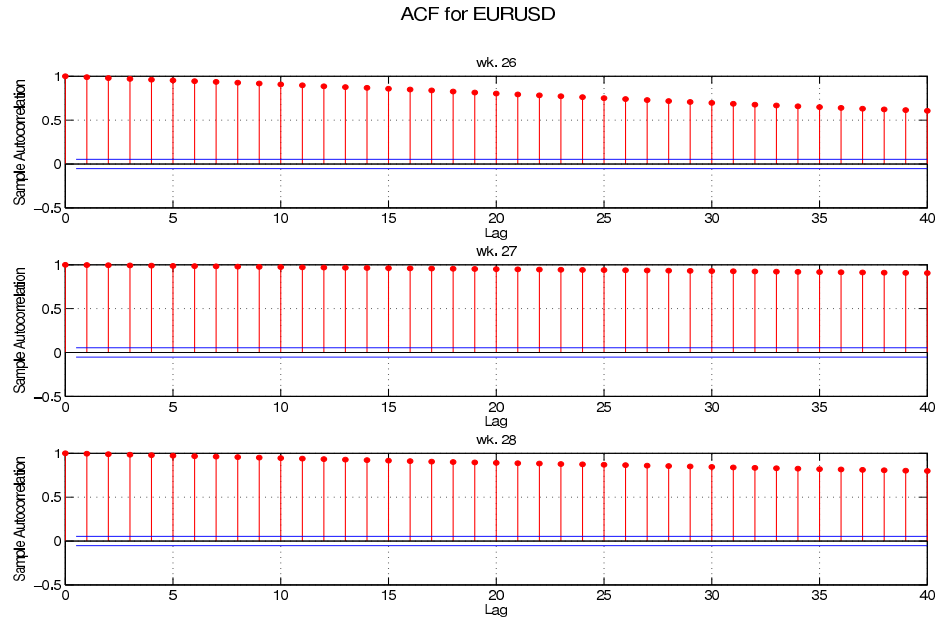


Figure 5.2: The ACF of EURUSD, undifferentiated

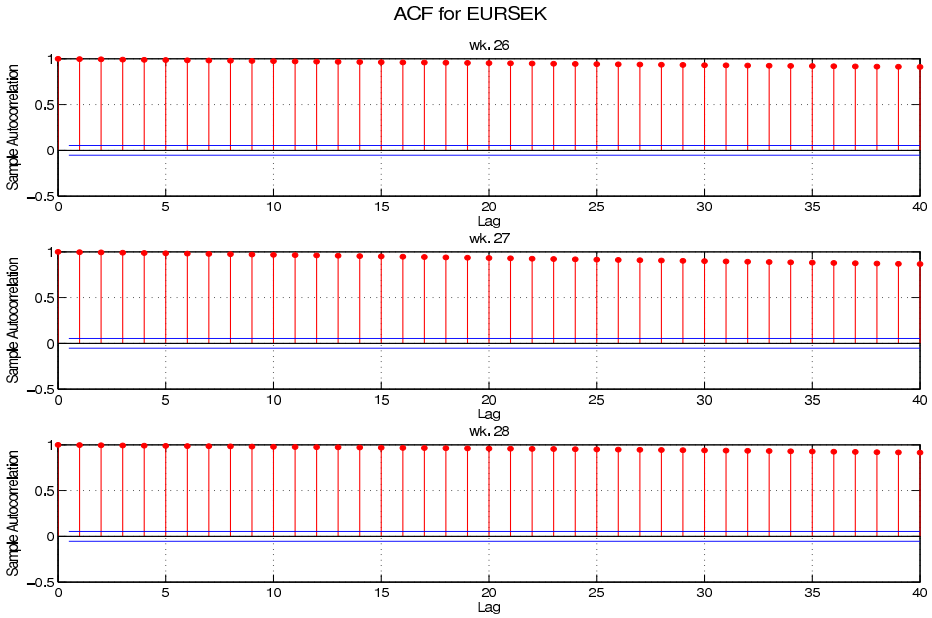


Figure 5.3: The ACF of EURSEK, undifferentiated

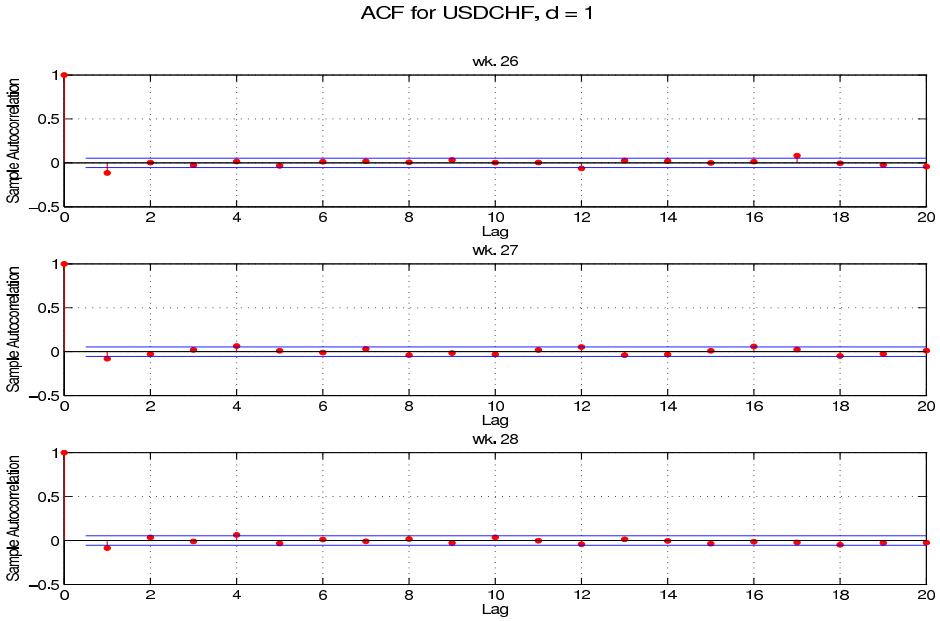


Figure 5.4: The ACF of USDCHF differentiated once.

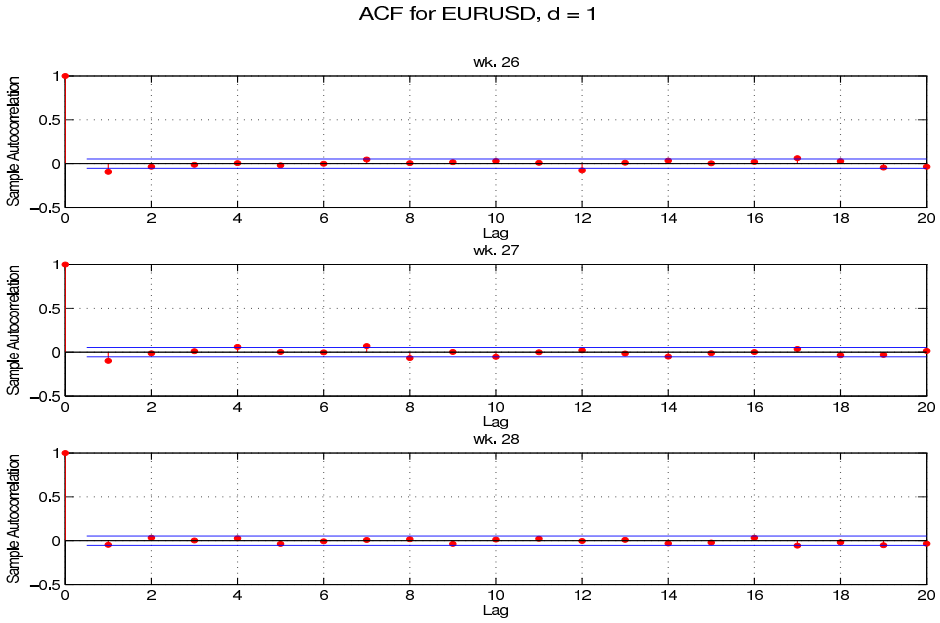


Figure 5.5: The ACF of EURUSD differentiated once.

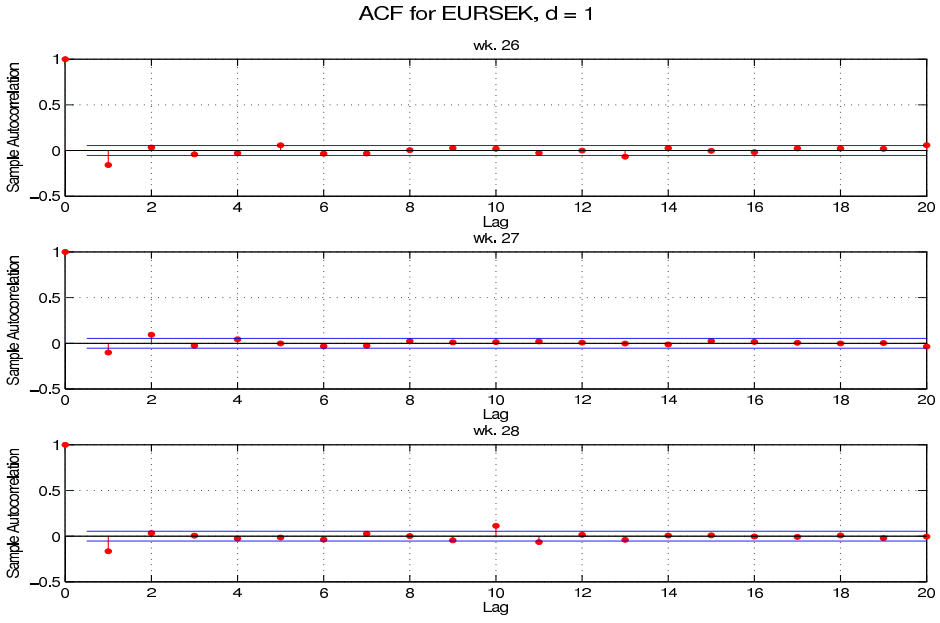


Figure 5.6: The ACF of EURSEK differentiated once.

$p = 10$ and $q = 6$. A combination of the combined result for EURUSD and EURSEK, $p = 8$ and $q = 7$. And a total combination of all the results, $p = 10$ and $q = 7$.

	USDCHF	EURUSD	EURSEK
wk. 26	$p = 6$, $q = 4$	$p = 5$, $q = 4$	$p = 6$, $q = 7$
wk. 27	$p = 10$, $q = 6$	$p = 8$, $q = 6$	$p = 5$, $q = 5$
wk. 28	$p = 6$, $q = 6$	$p = 6$, $q = 3$	$p = 6$, $q = 6$
combined result	$p = 10$, $q = 6$	$p = 8$, $q = 6$	$p = 6$, $q = 7$

Table 5.1: The p and q values with the lowest AIC value

5.2.1 Godness of fit

Deciding on the degree of the model from the three alternatives from the algorithm have been done by the goodness-of-fit test from Section 3.6. This will also be a measure of which model will result in the best estimation of the coefficients. Since the coefficients are re-estimated at every time step the Ljung-Box-Pierce statistic is recalculated at each time step after the initial estimation period of 400 data points, this gives about 1000 values of the statistic. Tables of the interval of these statistics are presented in 5.2, 5.3 and 5.4.

Histograms of the values of the Ljung-Box-Price statistic for the different models with different p and q values are presented in the appendix in Figures A.3, A.4, A.5, A.6, A.7 and A.8. To evaluate which model is the better we will se how many of the p-values of the statistics will fall in the 10% to 90% range of the χ^2-distribution, how many will give support for the null hypothesis that the residuals are independently distributed. This is shown in Table 5.5, the average values of these results are 58.5% for ARIMA(10,1,6), 60.0% for ARIMA(8,1,7) and 57.3% for ARIMA(10,1,7). The best results are then for

	USDCHF	EURUSD	EURSEK
wk. 26	[2.4608, 29.606] median: 8.2556	[2.1951, 27.640] median: 8.5524	[2.8186, 34.732] median: 12.053
wk. 27	[3.072261, 30.01377] median: 11.31192	[2.009874, 26.7828] median: 10.2	[1.194141, 37.85046] median: 5.345968
wk. 28	[1.532838, 27.08716] median: 10.74051	[3.049808, 29.38023], median: 11.13324	[1.758518, 37.15711] median: 7.67111

Table 5.2: Test statistic interval for $ARIMA(10,1,6)$

	USDCHF	EURUSD	EURSEK
wk. 26	[3.734831, 29.95651] median: 9.191576	[2.613274, 31.88653] median: 9.528291	[1.795583, 48.06473] median: 14.47148
wk. 27	[4.162933, 30.06187] median: 12.01964	[2.191619, 42.33729] median: 11.18065	[0.6669133, 39.96872] median: 5.663984
wk. 28	[3.356739, 32.76843] median: 12.69303	[4.152267, 31.1317] median: 12.65702	[1.947663, 37.62582] median: 9.332463

Table 5.3: Test statistic interval for $ARIMA(8,1,7)$

	USDCHF	EURUSD	EURSEK
wk. 26	[2.119914, 24.73587] median: 8.281466	[1.731562, 30.08485] median: 8.856522	[1.710631, 35.61219] median: 11.43474
wk. 27	[3.15774, 32.27852] median: 12.0743	[2.041602, 27.35932] median: 10.08781	[0.9923357, 39.92604] median: 5.496606
wk. 28	[3.493696, 28.90537] median: 10.31406	[2.709189, 37.53186] median: 11.10408	[1.310781, 33.33295] median: 7.31651

Table 5.4: Test statistic interval for $ARIMA(10,1,7)$

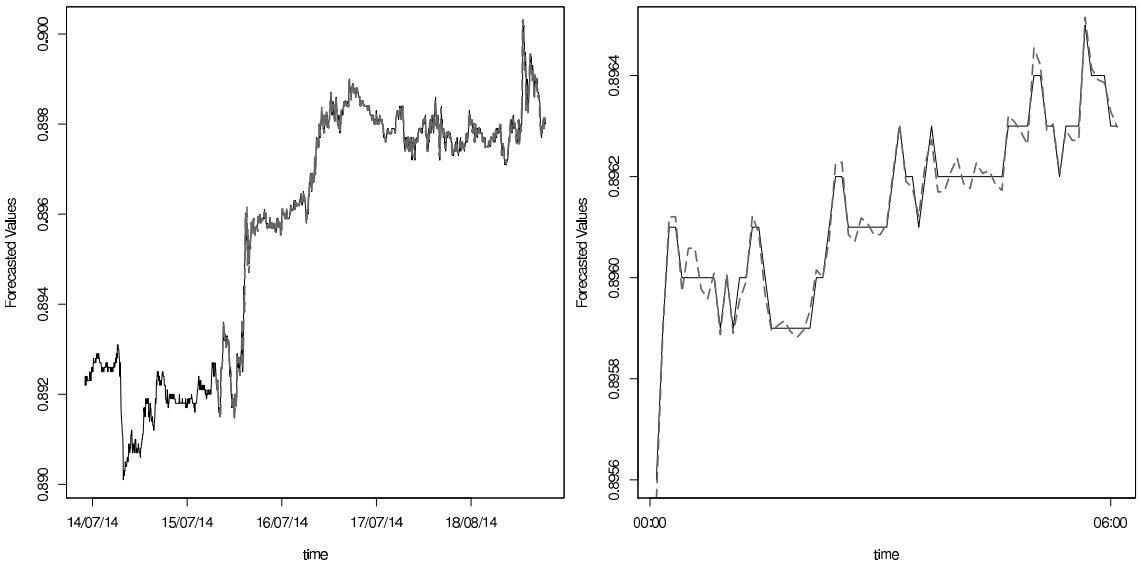
ARIMA(8,1,7) and this model are chosen.

5.3 Forecasting

When settling on a model, the future values of the currency pairs can be forecasted. The coefficients of the model are re-estimated every five minutes and after that the process are forecasted one step ahead. An example of the forecasted values of USDCHF week 29 are found in Figure 5.7a, a more detailed view of the first half of the 14 of July are shown in Figure 5.7b. In Table 5.6 and 5.7 the results for the root-mean-square deviation and the mean percentage error are presented. These all show low values which suggest a good accuracy of the forecast. Wk. 31 for the EURSEK pair has a bit higher values for the RMSD and the MPE the other weeks, a visual inspection of the plots of the data and the forecast are found in 5.8b. Here we can see that when a sharp jump in value of the currency the forecast method is not as good as one would have wanted, but this is an extreme event and dose not effect the overall fit of the model.

		USDCHF	EURUSD	EURSEK
wk. 26				
	ARIMA(10,1,6)	62.3%	58.4%	45.1%
	ARIMA(8,1,7)	61.8%	55.2%	52.3%
	ARIMA(10,1,7)	55.8%	57.4%	49.6%
wk. 27				
	ARIMA(10,1,6)	60.6%	83.3%	13.3%
	ARIMA(8,1,7)	77.3%	83.3%	19.8%
	ARIMA(10,1,7)	65.1%	81.6%	11.1%
wk. 28				
	ARIMA(10,1,6)	63.6%	65.3%	75.0%
	ARIMA(8,1,7)	56.7%	62.3%	71.7%
	ARIMA(10,1,7)	59.5%	62.9%	72.7%

Table 5.5: Precent of the Ljung-Box-Price statistic that fall in the 10% to 90% range of the χ^2-distribution



(a) The forecasted values, wk. 29 **(b)** Zoomed in forcasted values, 16/07/14

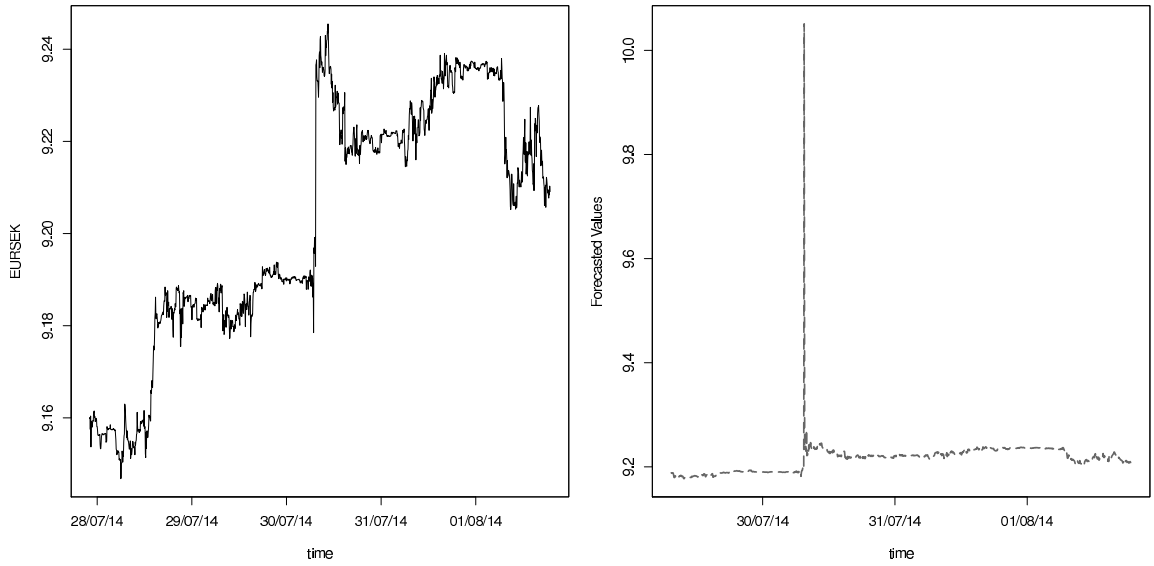
Figure 5.7: Forecasted values, in gray, for USDCHF, wk. 29

	USDCHF	EURUSD	EURSEK
wk. 29	$5.09 \cdot 10^{-5}$	$6.63 \cdot 10^{-5}$	$8.43 \cdot 10^{-4}$
wk. 30	$4.79 \cdot 10^{-5}$	$7.46 \cdot 10^{-5}$	$5.11 \cdot 10^{-4}$
wk. 31	$6.87 \cdot 10^{-5}$	$9.52 \cdot 10^{-5}$	0.0258
wk. 32	$5.20 \cdot 10^{-5}$	$6.34 \cdot 10^{-5}$	$6.95 \cdot 10^{-4}$
wk. 33	$5.70 \cdot 10^{-5}$	$6.78 \cdot 10^{-5}$	$4.95 \cdot 10^{-4}$
wk. 34	$4.56 \cdot 10^{-5}$	$6.62 \cdot 10^{-5}$	$5.51 \cdot 10^{-4}$
wk. 35	$8.55 \cdot 10^{-5}$	$7.90 \cdot 10^{-5}$	$7.38 \cdot 10^{-4}$
wk. 36	$1.91 \cdot 10^{-4}$	$3.39 \cdot 10^{-4}$	$8.68 \cdot 10^{-4}$
wk. 37	$7.01 \cdot 10^{-5}$	$8.62 \cdot 10^{-5}$	$6.28 \cdot 10^{-4}$

Table 5.6: Root-mean-square deviation of the forcasted values

	USDCHF	EURUSD	EURSEK
wk. 29	$5.38 \cdot 10^{-5}\%$	$-2.26 \cdot 10^{-5}\%$	$4.15 \cdot 10^{-4}\%$
wk. 30	$-1.26 \cdot 10^{-4}\%$	$2.63 \cdot 10^{-4}\%$	$1.42 \cdot 10^{-4}\%$
wk. 31	$1.28 \cdot 10^{-5}\%$	$2.64 \cdot 10^{-5}\%$	$8.79 \cdot 10^{-3}\%$
wk. 32	$2.72 \cdot 10^{-4}\%$	$2.88 \cdot 10^{-5}\%$	$-4.50 \cdot 10^{-5}\%$
wk. 33	$6.84 \cdot 10^{-5}\%$	$-1.44 \cdot 10^{-4}\%$	$2.80 \cdot 10^{-4}\%$
wk. 34	$-8.18 \cdot 10^{-5}\%$	$1.74 \cdot 10^{-4}\%$	$2.31 \cdot 10^{-4}\%$
wk. 35	$-4.73 \cdot 10^{-4}\%$	$1.09 \cdot 10^{-4}\%$	$-2.40 \cdot 10^{-4}\%$
wk. 36	$1.68 \cdot 10^{-4}\%$	$-7.15 \cdot 10^{-4}\%$	$-792 \cdot 10^{-6}\%$
wk. 37	$3.44 \cdot 10^{-4}\%$	$-1.43 \cdot 10^{-4}\%$	$-4.21 \cdot 10^{-4}\%$

Table 5.7: The mean percentage error of the forcasted values



(a) The value of EURSEK **(b)** The forecasted values of EURSEK

Figure 5.8: The values of EURSEK, week 31, and the forecasted values of that same week.

5.4 Buy and sell strategies

The buy and sell strategies outlined in Section 4.7 have been implemented using the ARIMA(8,1,7) model. For both of the strategies 1000 USD or 1000 EUR are "invested" at the start of the week after the initial estimation of the coefficients, which is after 400 data points.

5.4.1 Strategy 1

The results for the first strategy involving the forecasted values from the ARIMA(8,1,7) model are presented in Table 5.8. The method used are: for the USDCHF pair 1000 USD are used as starting value and when the strategy gives a signal for selling the USD and buying CHF this is done. The process are continued through out and evaluated at the end of the week. If the process ends with a value in CHF this is compared with the value of 1000 USD in CHF at the start of the buy and sell period.

5.4.2 Strategy 2

For strategy 2 the values used for calculating the *SMA* and the upper and lower band are $n = 20$, $k = 2$. These values are chosen to produce bands that are not to far away from the data values so that the strategy never is executed but not to close so that the effect of the strategy is to low. The bands and the data for USDCHF week 29 are plotted in Figure 5.9. The results, that is the closing values from an 1000 USD or 1000 EUR investment with strategy 2, are found in Table 5.9.

	USDCHF	EURUSD	EURSEK
wk. 29	1052.03 USD +5.20%	1425.60 USD +5.40%	9920.20 SEK +7.22%
wk. 30	1046.36 USD +4.64%	1407.85 USD +4.81%	1064.25 EUR +6.43%
wk. 31	1054.53 USD +5.45%	1046.65 EUR +4.66%	9785.03 SEK +6.25%
wk. 32	965.85 CHF +6.68%	1058.71 EUR +5.87%	1065.25 EUR +6.52%
wk. 33	1067.13 USD +6.71%	1413.82 USD +5.53%	1051.47 EUR +5.15%
wk. 34	1058.95 USD +5.90%	1401.32 USD +5.78%	9627.09 SEK +5.08%
wk. 35	1065.75 USD +6.58%	1065.71 EUR +6.57%	1046.39 EUR +4.64%
wk. 36	990.79 CHF +6.45%	1386.65 USD +7.00%	1073.73 EUR +7.37%
wk. 37	1099.00 USD +9.90%	1083.55 EUR +8.36%	1059.78 EUR +5.98%

Table 5.8: The gain from investing 1000 USD or 1000 EUR with buy and sell strategy 1.

5.5 Simulation

To preform the multivariate normality test, $\{\tilde{Y}_1, \ldots, \tilde{Y}_{17}\}$ are simulated with $n = 10000$. The result of the Mardia's test are shown in Table 5.10. On a 5% significance level none of the test statistics show a rejection of the null hypothesis, that $\{\tilde{Y}_1, \ldots, \tilde{Y}_{17}\}$ belong to the multivariate normal distribution. The Q-Q plots of the values of the skewness statistic, A, versus the χ^2_{680}-distribution are shown in Figure 5.10.

One can also see that the $\tilde{Y}_i$ are individually normal distributed in de histogram plots in appendix, Figures A.9, A.10 and A.11 for USDCHF, Figures A.12, A.13 and A.14 for EURUSD and Figures A.15, A.16 and A.17 for EURSEK. The result of the simulation of described in the method chapter, Section 4.6, are presented in Figure 5.11. This simulation is about one year, that is around 70 000 data point, to achieve this after the simulated coefficients who do not fulfil the stationary and invertible conditions, the

	USDCHF	EURUSD	EURSEK
wk. 29	895.00 CHF +0.30%	1000.37 EUR +0.04%	992.58 EUR -0.74%
wk. 30	992.16 USD -0.78%	1004.53 EUR +0.45%	999.45 EUR -0.06%
wk. 31	999.458 USD -0.054%	1339.29 USD -0.30%	9142.96 SEK -0.50%
wk. 32	998.02 USD -0.20%	1340.70 USD -0.11%	9227.57 SEK +0.08%
wk. 33	900.02 CHF -0.85	999.11 EUR -0.09%	1001.77 EUR +0.18%
wk. 34	910.58 CHF +0.37%	1008.23 EUR +0.82%	994.22 EUR -0.58%
wk. 35	996.84 USD -0.32%	1311.92 USD -0.66%	990.72 EUR -0.93%
wk. 36	996.97 USD -0.30%	1007.71 EUR +0.77%	9135.15 SEK -0.64%
wk. 37	1000.00 USD 0.00%	1293.99 USD +0.51%	991.77 EUR -0.82%

Table 5.9: Gains and losses from investing 1000 USD or 1000 EUR with strategy 2

	A	df	p-value for A	B	p-value for B
USDCHF	705.5740	680	0.2410	0.1416	0.8874
EURUSD	659.1540	680	0.7099	0.4762	0.6339
EURSEK	648.5997	680	0.8014	-0.9282	0.3533

Table 5.10: The result from the Mardia's test

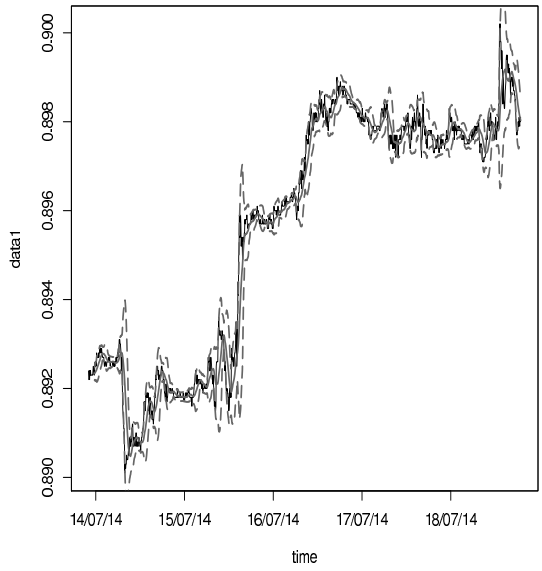


Figure 5.9: The *SMA*, the upper and lower band used in buy and sell strategy 2

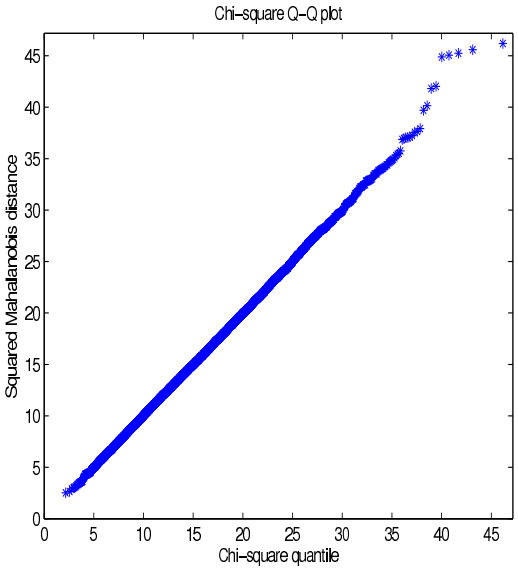
	USDCHF	EURUSD	EURSEK
Strategy 1	435724.8 CHF +46556.87%	137 514.00 EUR +13 651.40%	418 150.00 EUR +41 715%
Strategy 2	833.61 USD -16.64%	1 goodness-of-fit233.31 USD -4.76%	7 528.25 SEK -18.56%

Table 5.11: Results from buy and sell strategies 1 and 2 for the simulated values

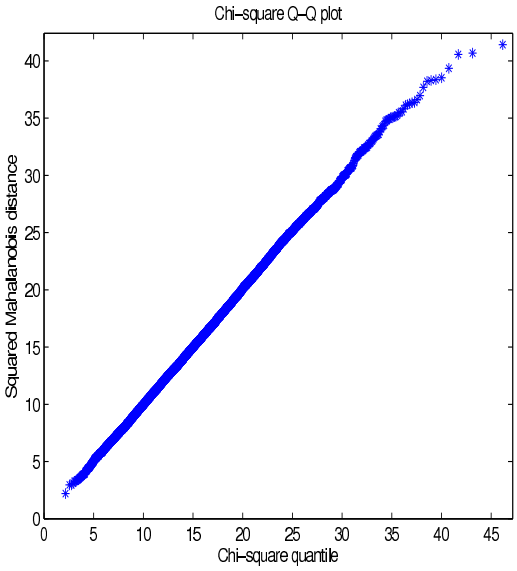
simulation starts with 500 000 points. The standard deviation used for the innovations in the simulation are: USDCHF; $\sigma = 1.614 \cdot 10^{-4}$, EURUSD; $\sigma = 1.977 \cdot 10^{-4}$, EURSEK; $\sigma = 1.811 \cdot 10^{-3}$.

5.5.1 Buy and sell strategies of the simulated data

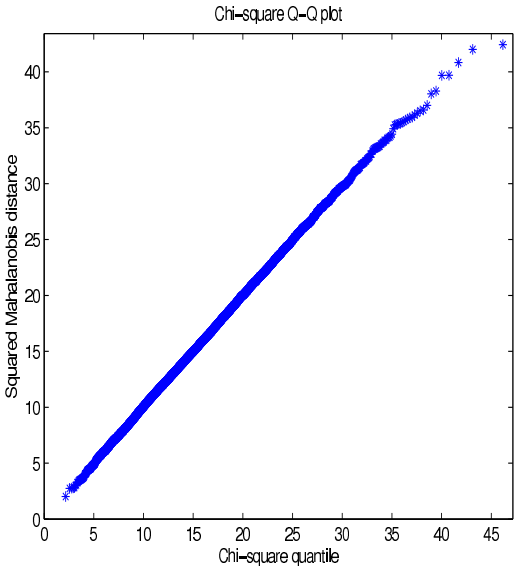
We preform the same buy and sell strategies on the simulated values as we did on the real data. The results are presented in Table 5.11.



(a) USDCHF

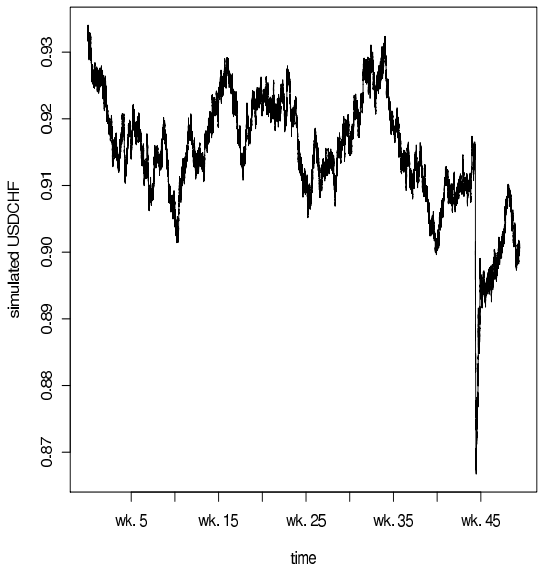


(b) EURUSD

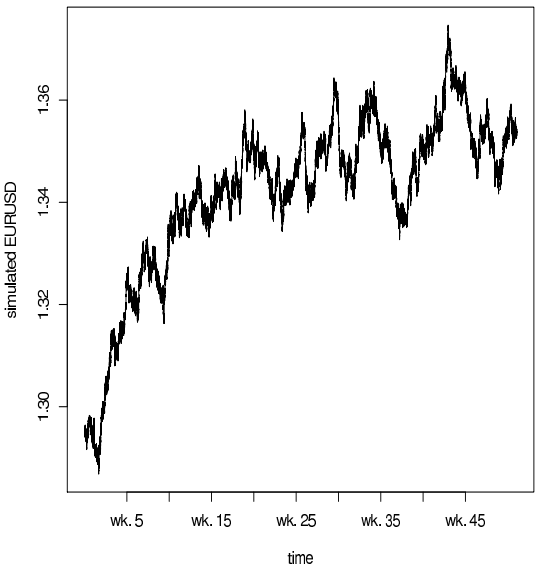


(c) EURSEK

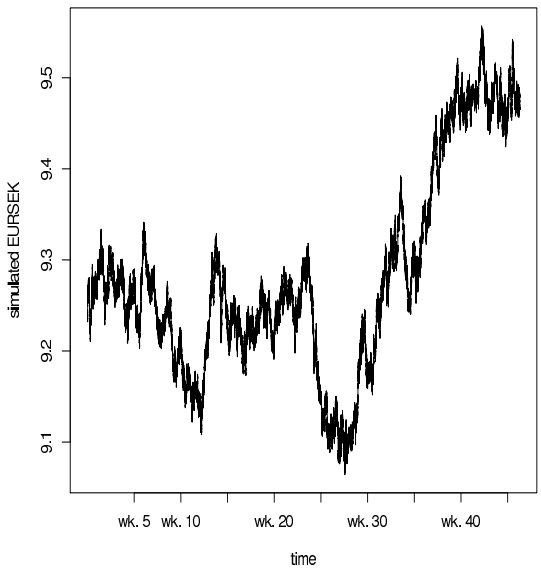
Figure 5.10: Q-Q plots of the Mardia test statisic, A, v.s. the χ^2_{680}-distribution.



(a) USDCHF



(b) EURUSD



(c) EURSEK

Figure 5.11: One year of simulated values of the tree different currency pairs

6

Conclusion

The general aim of this thesis that is, the search for a model that can describe the currency pairs, have been fruitful. The process shows clear signs of non-stationarity on a weekly basis so the ARIMA model is appropriate, and the results for the ARIMA(8,1,7) show a good fit.

The fact that the smaller model of the three that where tested for their goodness-of-fit had the best fit probably depends on the relative volatility of the currency pairs. A shorter "memory" of the model can better take care of the fickleness of the currency. There is one more autoregressive part then moving-average part in the model, so the current value of the currency is a bit more dependent on past values of the currency then the past values of the white noise. On the other hand, according to this model the value of a currency in one of these pairs only depends on the previous values of that currency back to 45 minutes before the current time.

The behaviour of the EURSEK pair is a bit different from the other pairs due to that it has much smaller trade volumes then the others. From looking at the plots of the data, Figure A.1 and A.2, one can see that the EURSEK pair has more sharp jumps and more calm periods then the USDCHF and EURUSD pairs. In the goodness-of-fit result, table 5.5, one can see a some what inferior result for EURSEK then for the others.

The forecasting of the currency pairs with the ARIMA model are quick enough to be able to preform the re-estimation and forecasting in the five minutes period of the data.

The result of the buy and sell strategy 1 are interesting, the result are very good. This must be an indication on the fact that the model represents the currency well. If this model can be implemented on a real world situation there is potential for real profit.

Buy and sell strategy 2 is not satisfactory, it dose not seem to be better then chance. There need to be more criteria then just the upper and lower bands for when to buy and sell.

The simulation of the currency values worked quite well. The test for normality where positive, which is a sign that the Copula simulation was successful. This result

can be used to improve the second buy and sell strategy. The result from strategy one, used on the simulated values, might not be entirely trusted, the data is generated with the same model which is used to forecast the values. The forecast is probably to good to be realistic.

6.1 Future Research

Investigating if the currency pairs show any signs of seasonality might improve the model. That is are there any seasonal trends that can be accounted for in the model. This can be the both in the short term, smaller currencies could be traded less during the night for that currency, and the long term, any trends over the year for example. This can be done with a seasonal ARIMA model, finding the trends will probably be a bit more difficult, but an interesting topic of study.

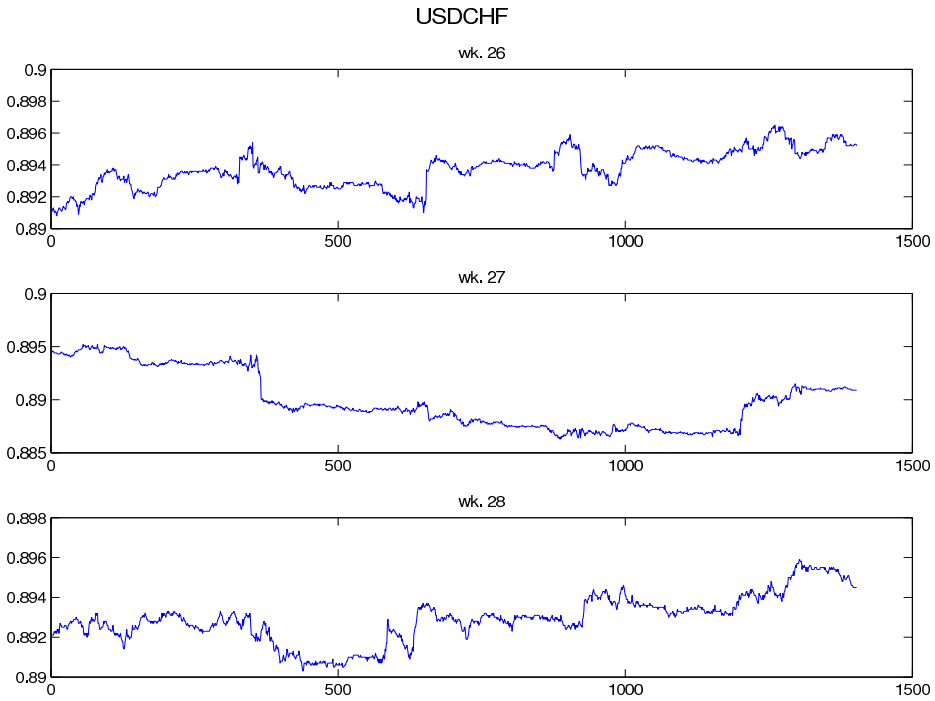
Another thing that will improve the model is to have separate models for the different currency pairs. One might also think of re-estimating the whole model, not just the coefficients, before each forecasting is done. If doing this the time might be an issue, it might take more time to re-estimate the entire model then five minutes. If that is the case an evaluation if the better model is worth the extra time.

The behaviour of the EURSEK should be interesting to study more, an interesting question to ask is if the EURSEK is representative for other ”small” currency pairs. Maybe there is one model that fits the *major* currency pairs and one for the rest.

Appendices

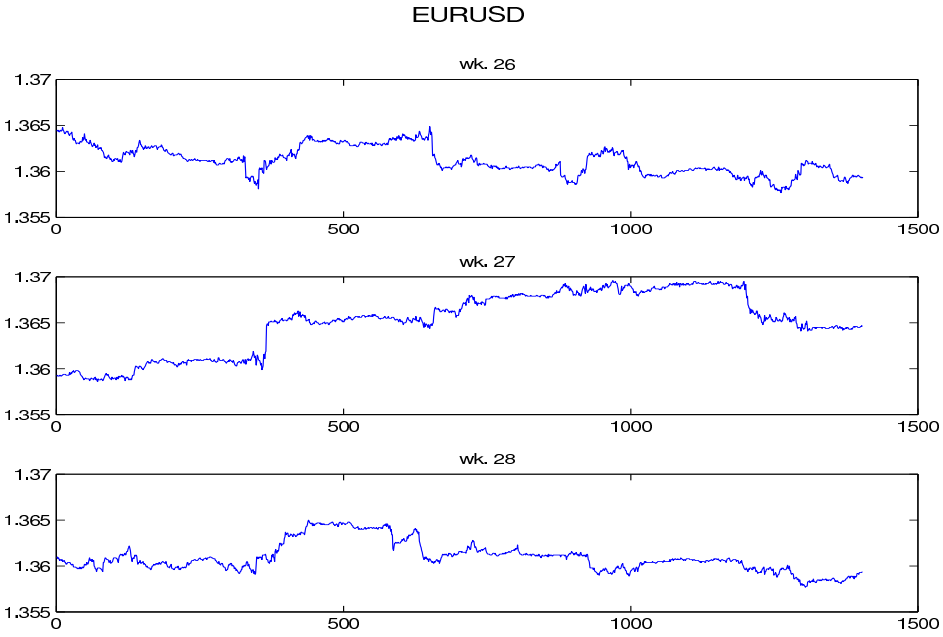
A

Results

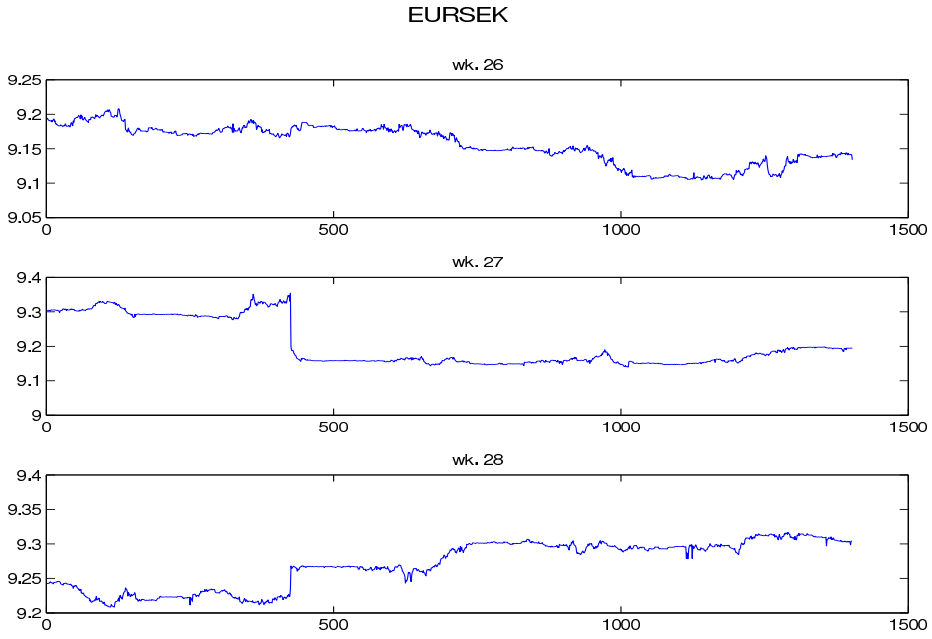


(a) The value of USDCHF in weeks 26, 27 and 28.

Figure A.1: Plots of the data

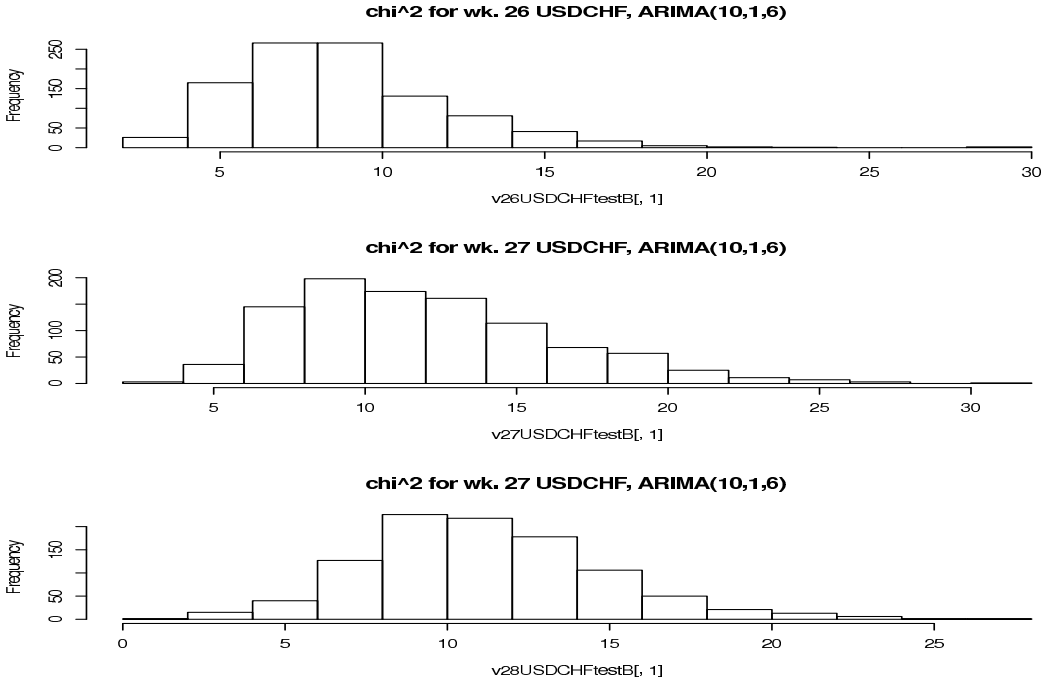


(a) The value of EURUSD in weeks 26, 27 and 28.

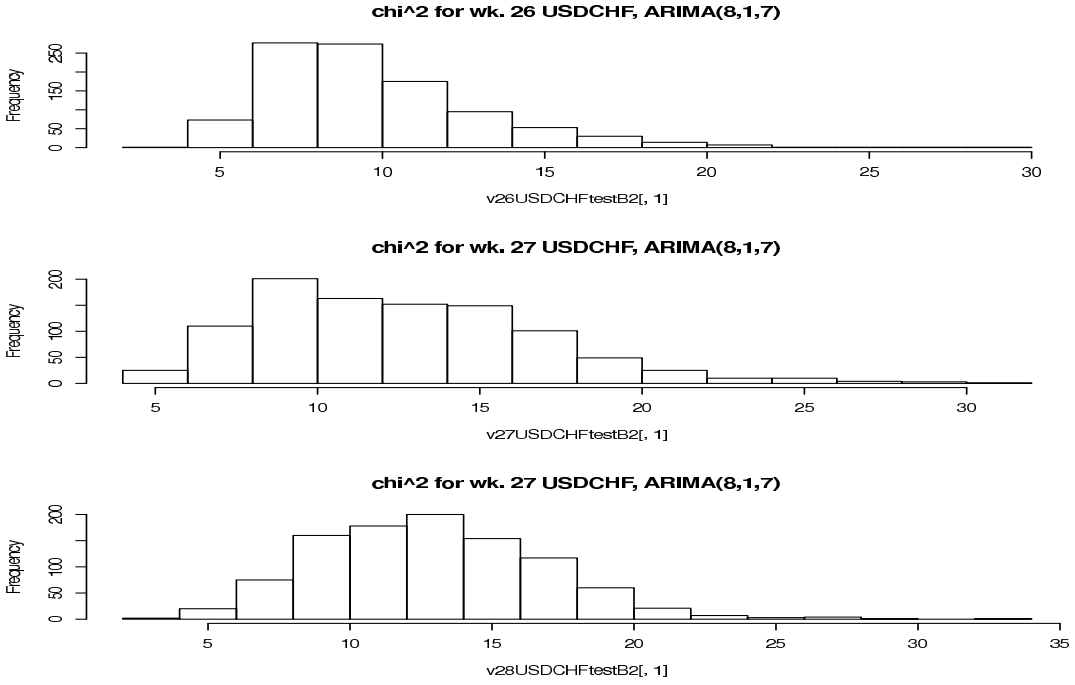


(b) The value of EURSEK in weeks 26, 27 and 28.

Figure A.2: Plots of the data

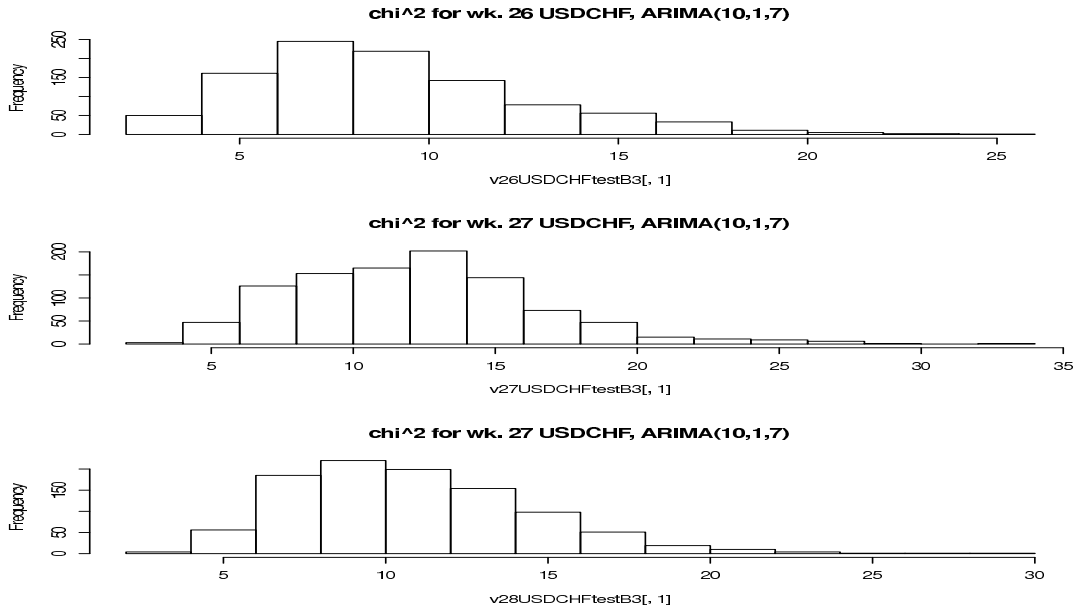


(a) Histogram of the Ljung-Box-Price statistic for $ARIMA(10,1,6)$



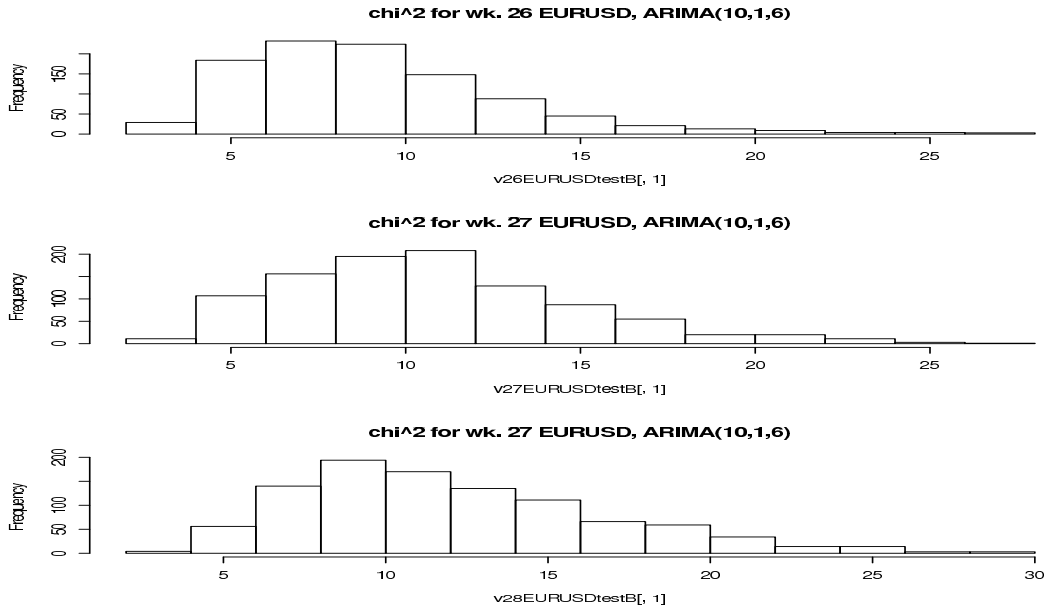
(b) Histogram of the Ljung-Box-Price statistic for $ARIMA(8,1,7)$

Figure A.3: Histograms of the Ljung-Box-Price statistic frot USDCHF



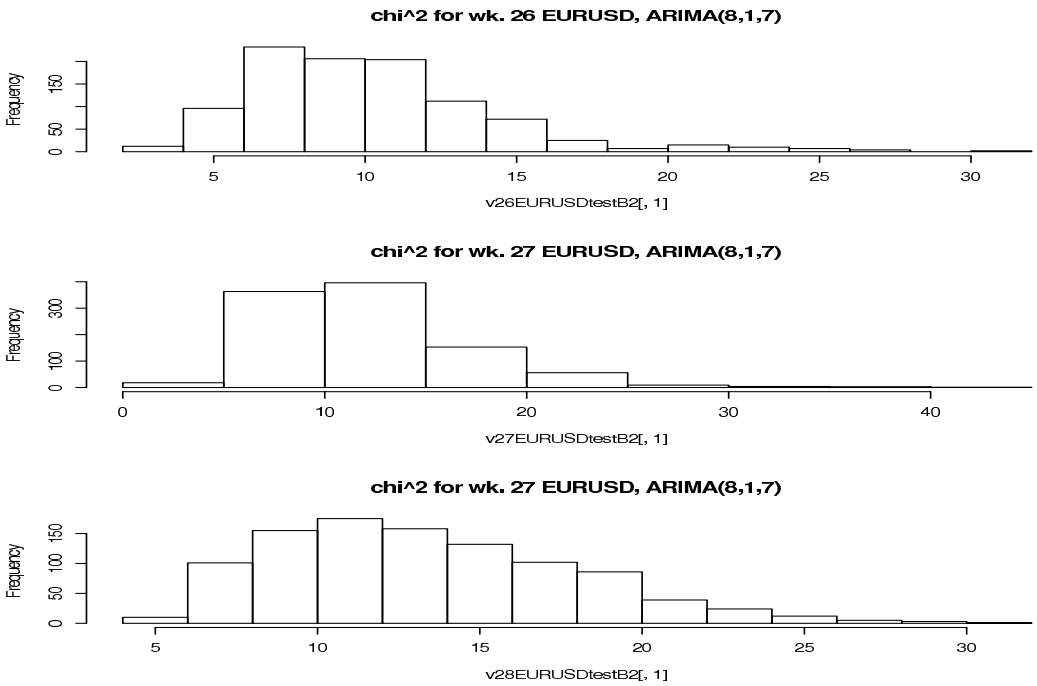
(a) Histogram of the Ljung-Box-Price statistic for *ARIMA*(10,1,7)

Figure A.4: Histograms of the Ljung-Box-Price statistic frot USDCHF

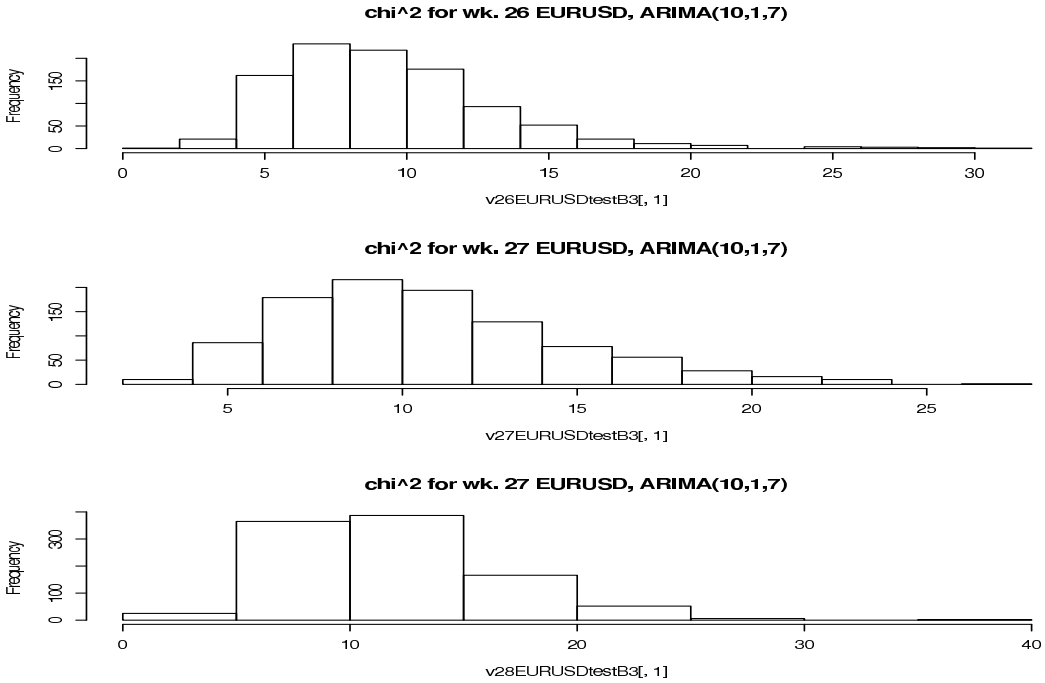


(a) Histogram of the Ljung-Box-Price statistic for *ARIMA*(10,1,6)

Figure A.5: Histograms of the Ljung-Box-Price statistic frot EURUSD

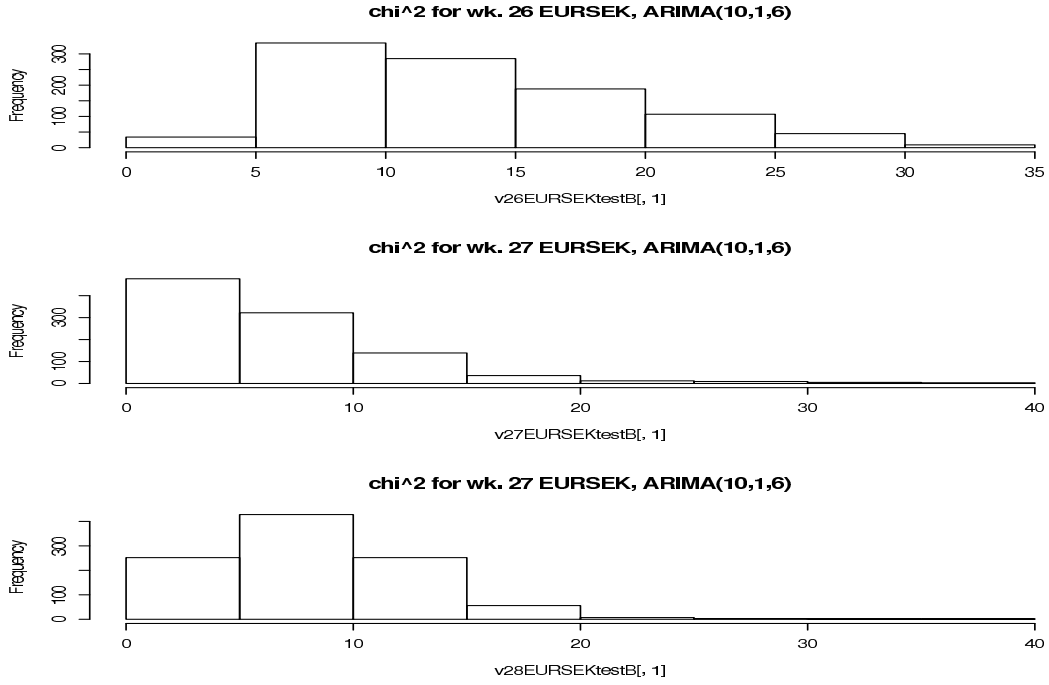


(a) Histogram of the Ljung-Box-Price statistic for *ARIMA*(8,1,7)

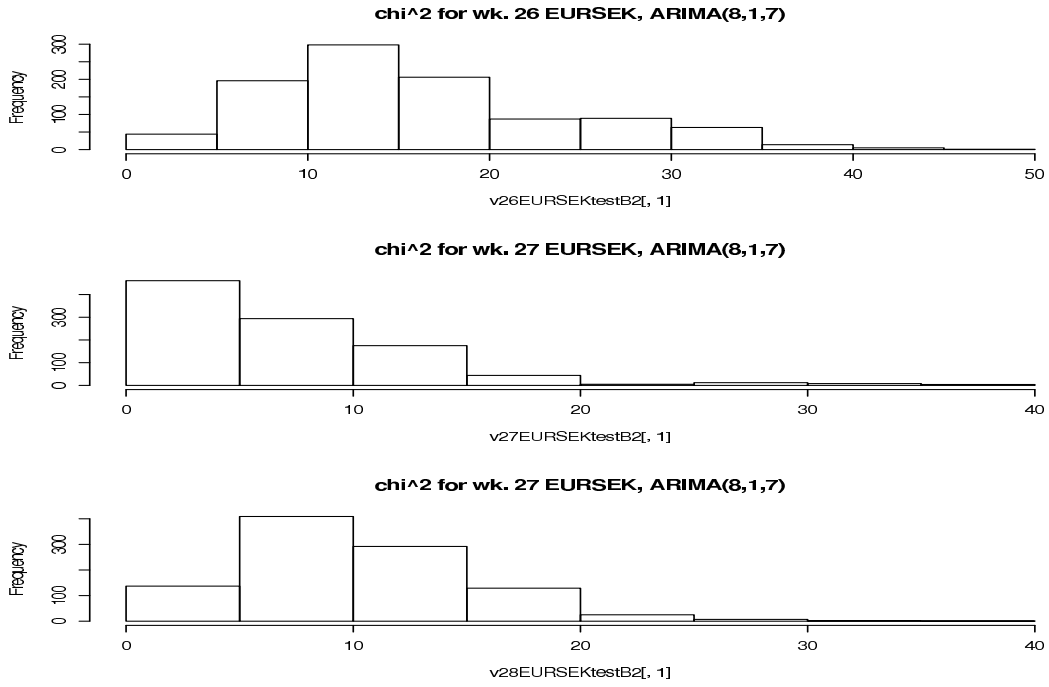


(b) Histogram of the Ljung-Box-Price statistic for *ARIMA*(10,1,7)

Figure A.6: Histograms of the Ljung-Box-Price statistic frot EURUSD

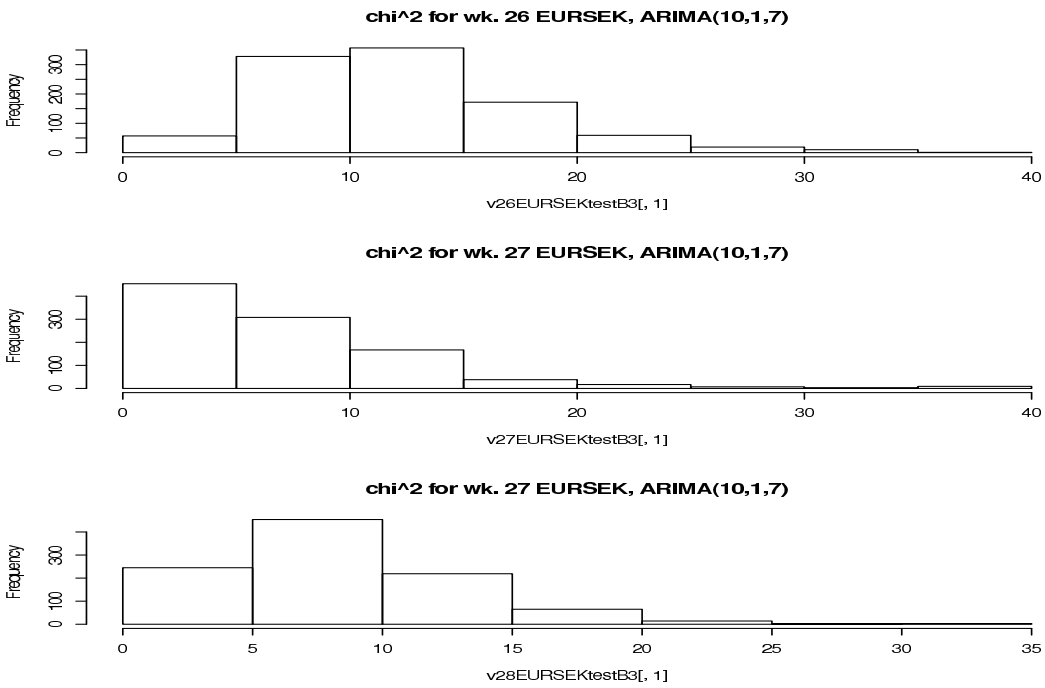


(a) Histogram of the Ljung-Box-Price statistic for *ARIMA*(10,1,6)



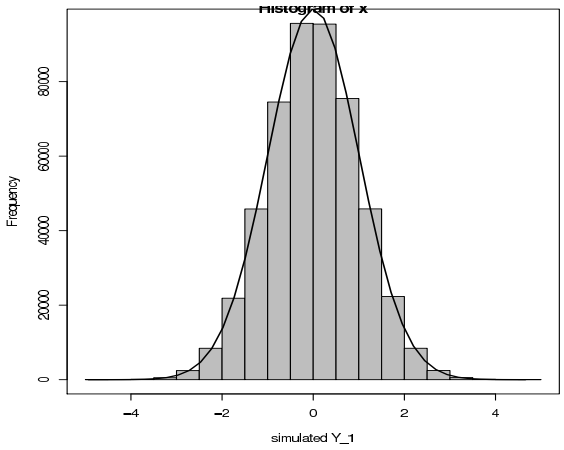
(b) Histogram of the Ljung-Box-Price statistic for *ARIMA*(8,1,7)

Figure A.7: Histograms of the Ljung-Box-Price statistic frot EURSEK

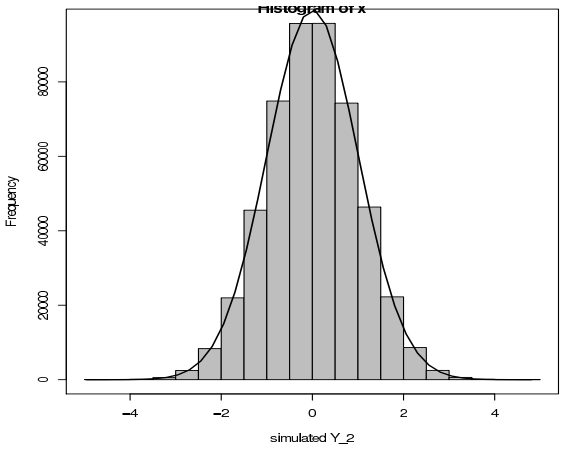


(a) Histogram of the Ljung-Box-Price statistic for $ARIMA(10,1,7)$

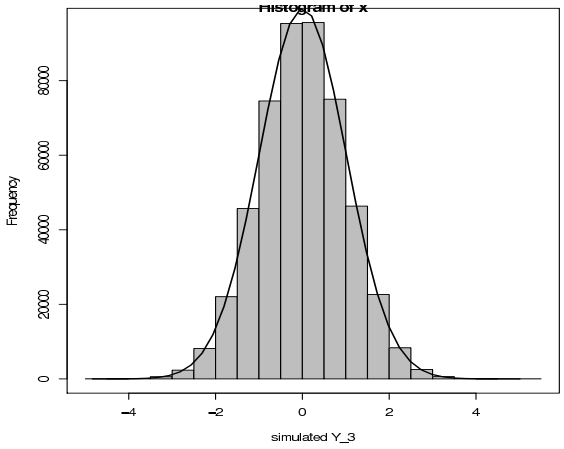
Figure A.8: Histograms of the Ljung-Box-Price statistic frot EURSEK



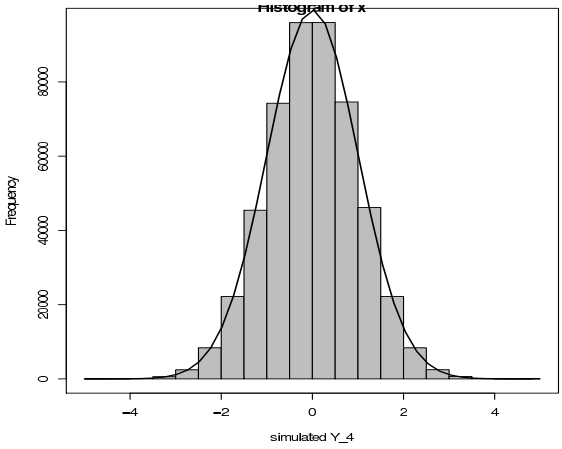
(a) Histogram of $\tilde{Y}_1$



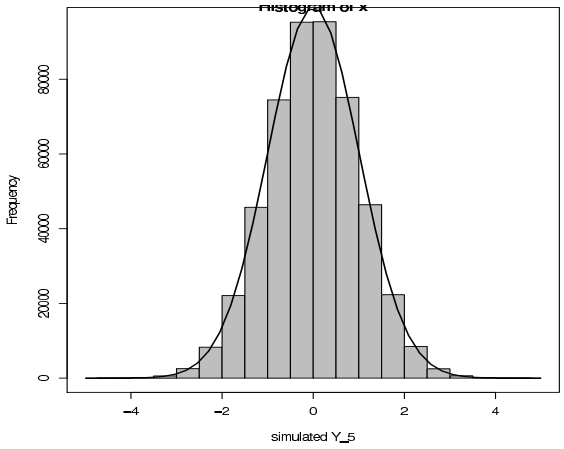
(b) Histogram of $\tilde{Y}_2$



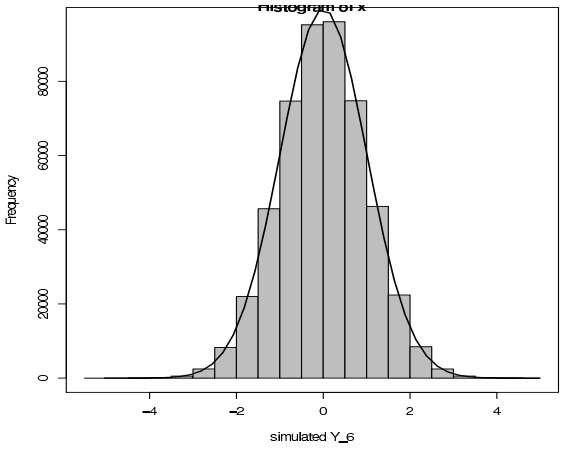
(c) Histogram of $\tilde{Y}_3$



(d) Histogram of $\tilde{Y}_4$

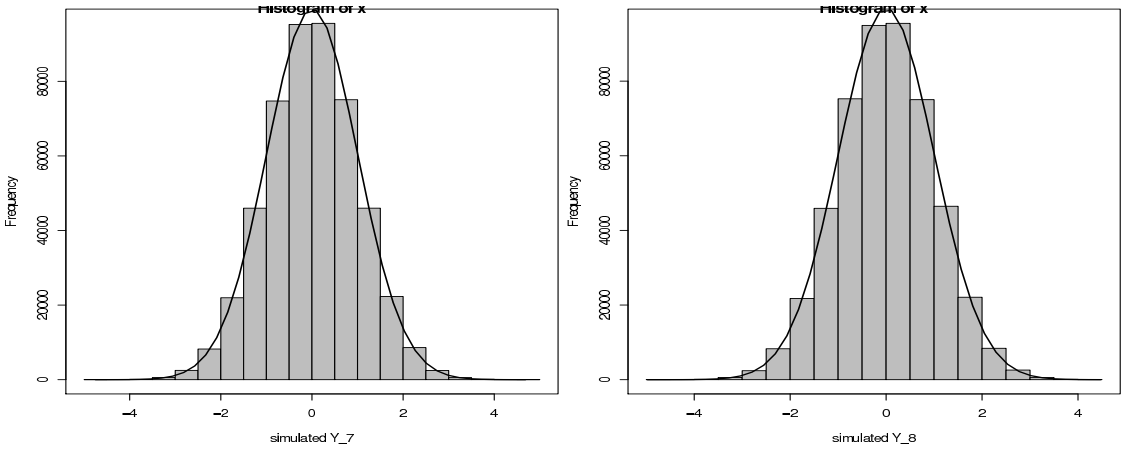


(e) Histogram of $\tilde{Y}_5$



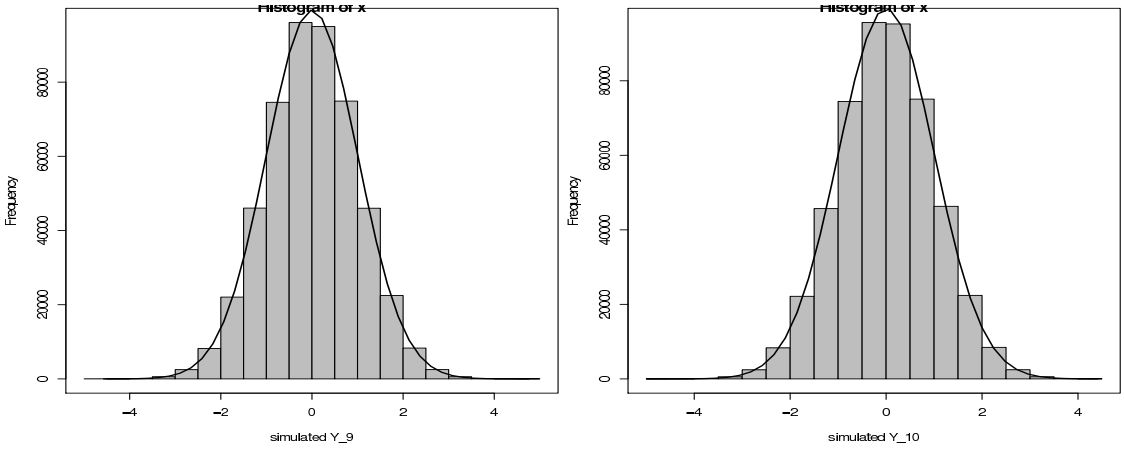
(f) Histogram of $\tilde{Y}_6$

Figure A.9: The histogram plots of $\tilde{Y}_1 - \tilde{Y}_6$ for USDCHF



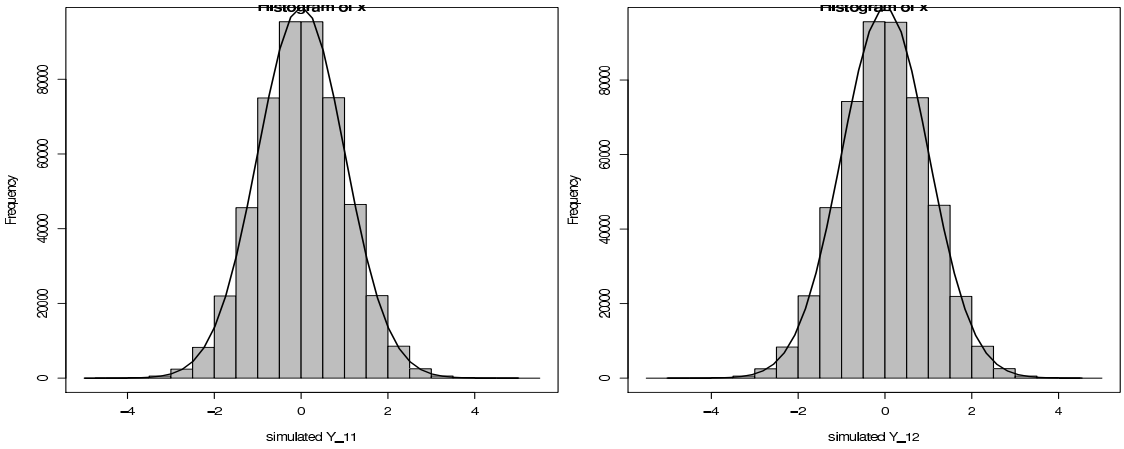
(a) Histogram of $\tilde{Y}_7$

(b) Histogram of $\tilde{Y}_8$



(c) Histogram of $\tilde{Y}_9$

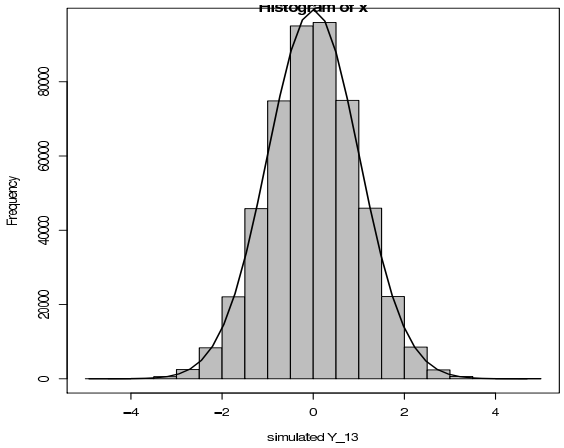
(d) Histogram of $\tilde{Y}_{10}$



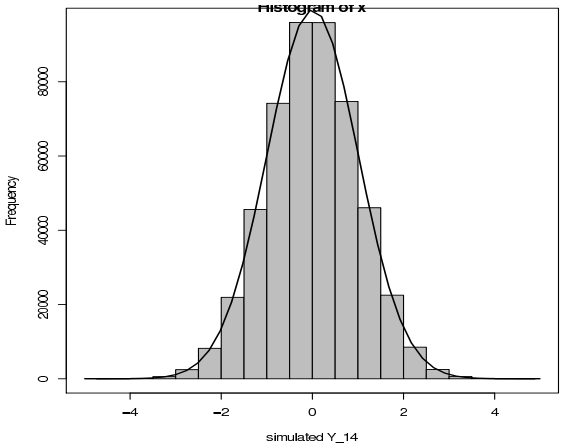
(e) Histogram of $\tilde{Y}_{11}$

(f) Histogram of $\tilde{Y}_{12}$

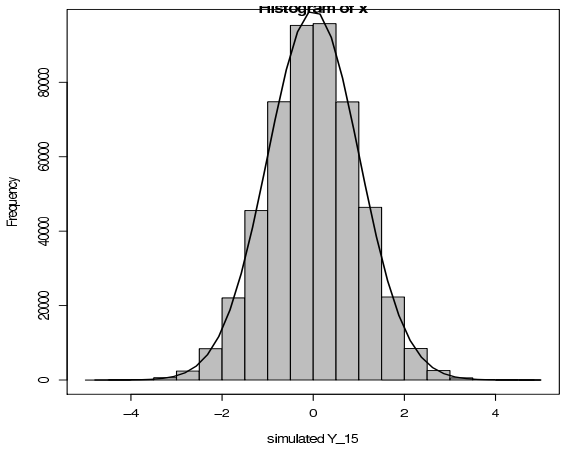
Figure A.10: The histogram plots of $\tilde{Y}_6 - \tilde{Y}_{12}$ for USDCHF



(a) Histogram of $\tilde{Y}_{13}$

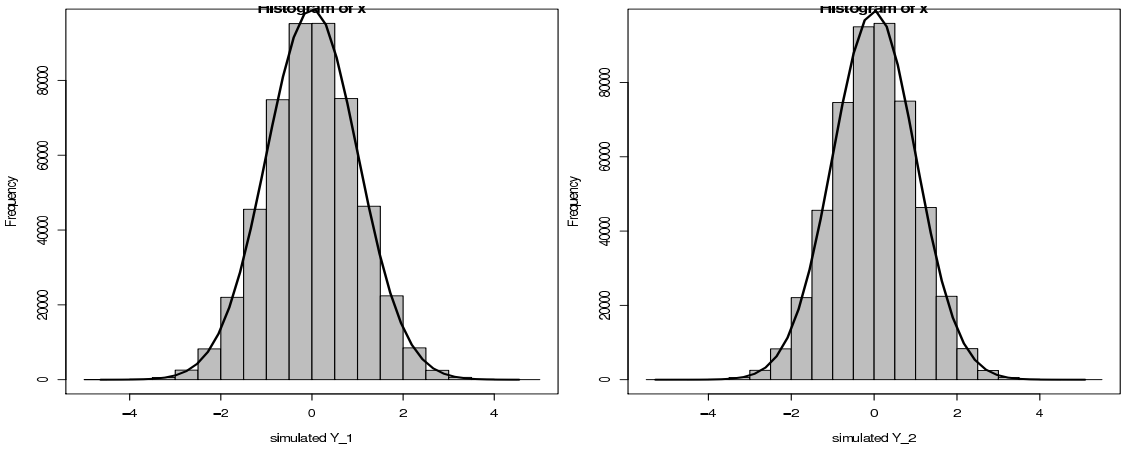


(b) Histogram of $\tilde{Y}_{14}$



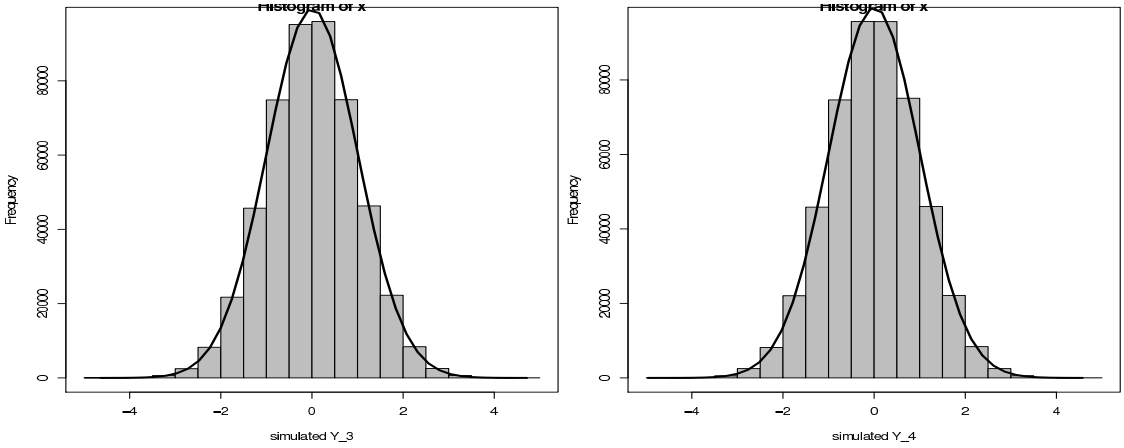
(c) Histogram of $\tilde{Y}_{15}$

Figure A.11: The histogram plots of $\tilde{Y}_{13} - \tilde{Y}_{15}$ for USDCHF



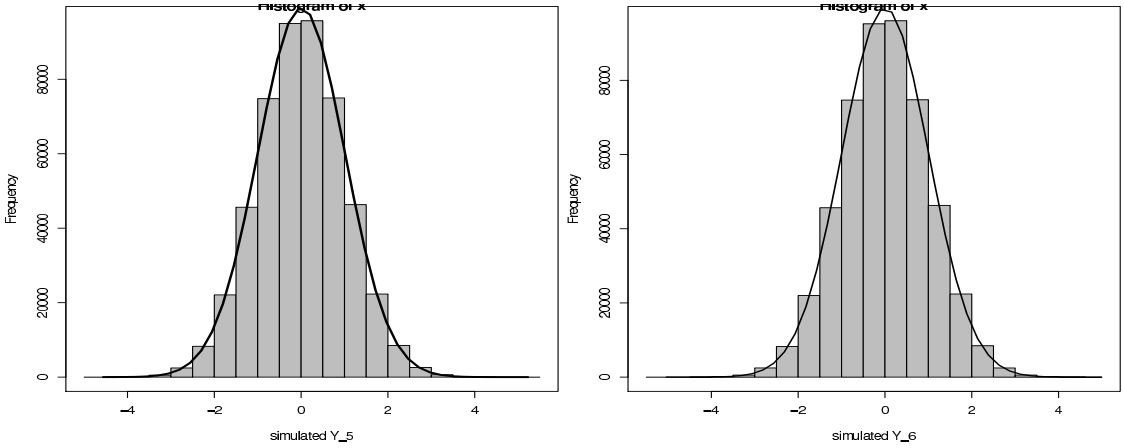
(a) Histogram of $\tilde{Y}_1$

(b) Histogram of $\tilde{Y}_2$



(c) Histogram of $\tilde{Y}_3$

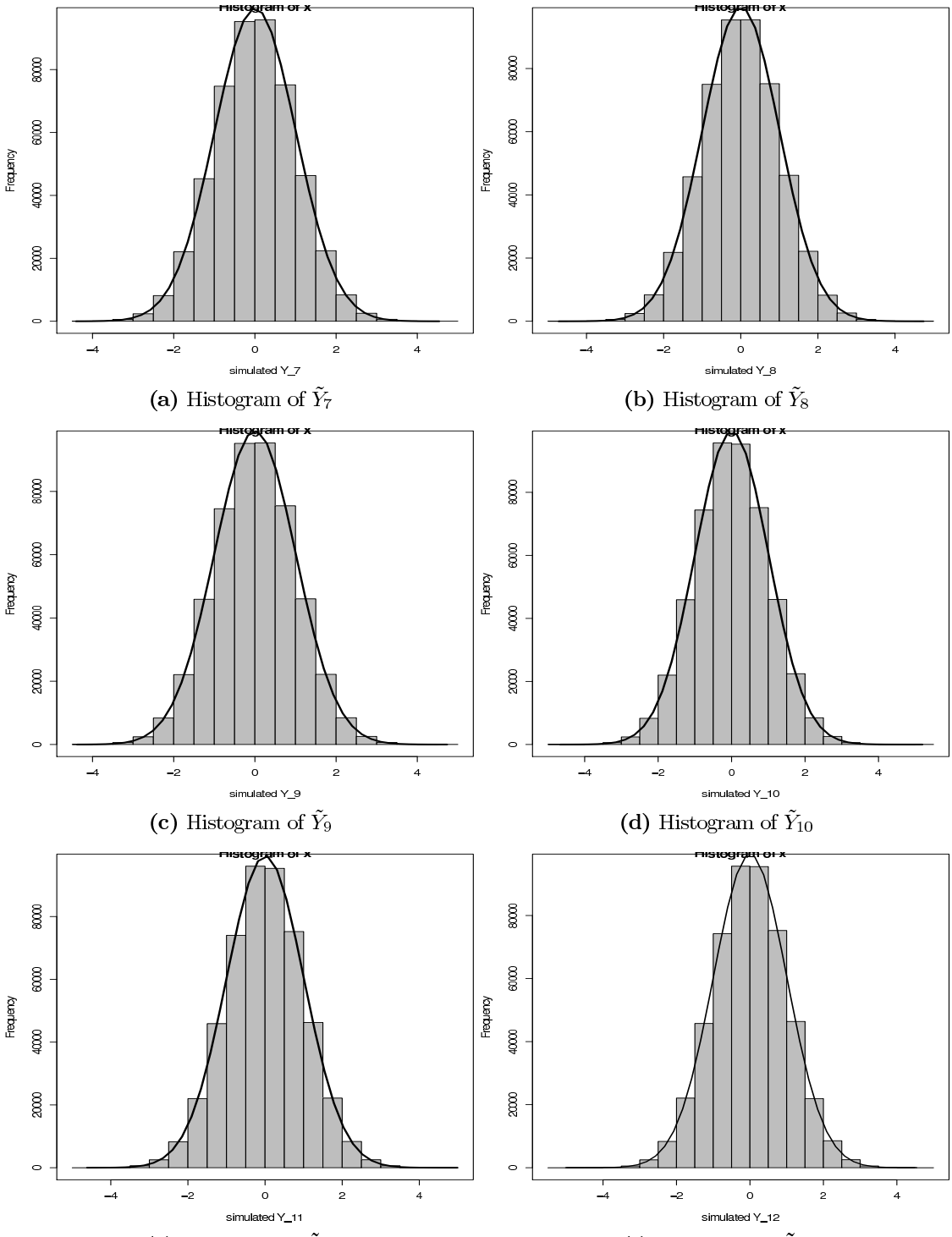
(d) Histogram of $\tilde{Y}_4$



(e) Histogram of $\tilde{Y}_5$

(f) Histogram of $\tilde{Y}_6$

Figure A.12: The histogram plots, together with the normal curve, of $\tilde{Y}_1 - \tilde{Y}_6$ for EURUSD



(a) Histogram of $\tilde{Y}_7$

(b) Histogram of $\tilde{Y}_8$

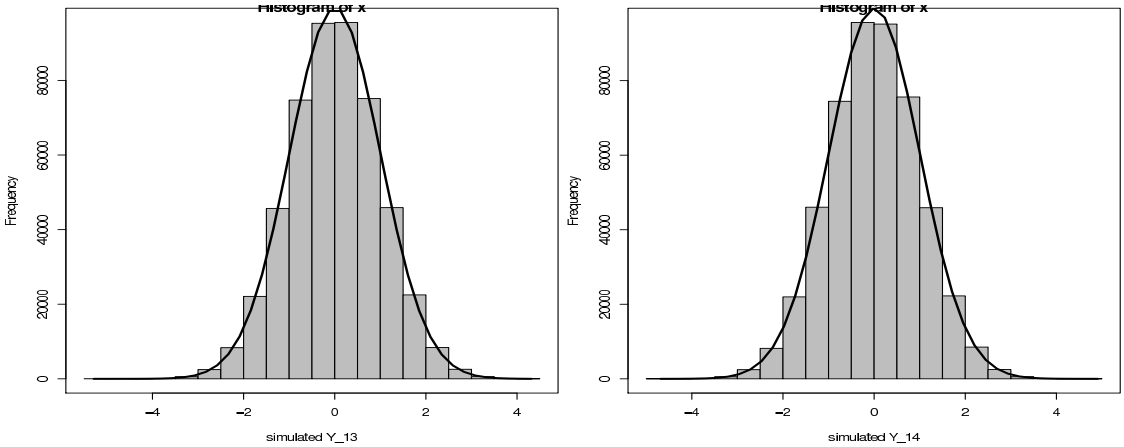
(c) Histogram of $\tilde{Y}_9$

(d) Histogram of $\tilde{Y}_{10}$

(e) Histogram of $\tilde{Y}_{11}$

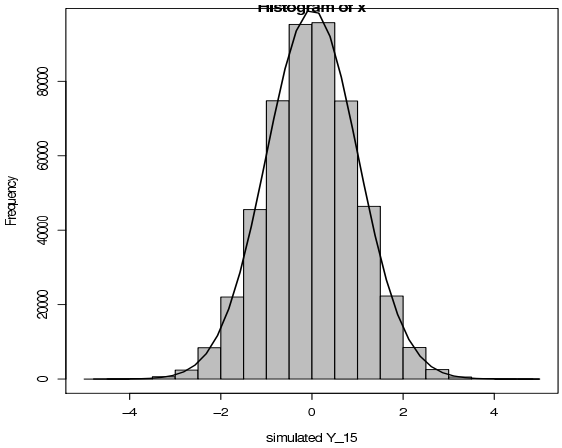
(f) Histogram of $\tilde{Y}_{12}$

Figure A.13: The histogram plots, together with the normal curve, of $\tilde{Y}_7 - \tilde{Y}_{12}$ for EURUSD



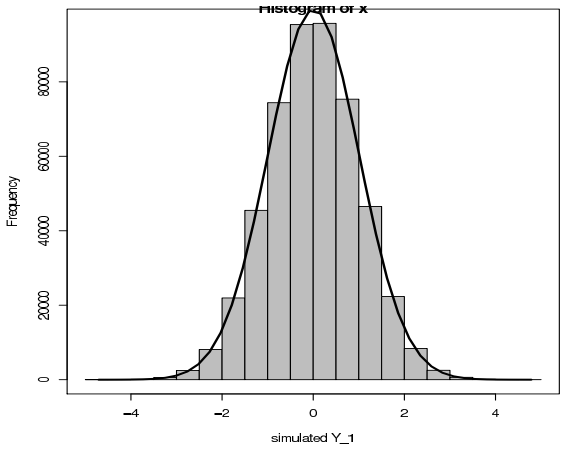
(a) Histogram of $\tilde{Y}_{13}$

(b) Histogram of $\tilde{Y}_{14}$

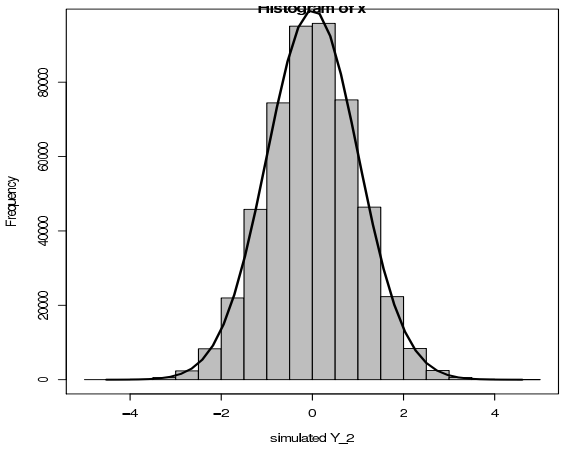


(c) Histogram of $\tilde{Y}_{15}$

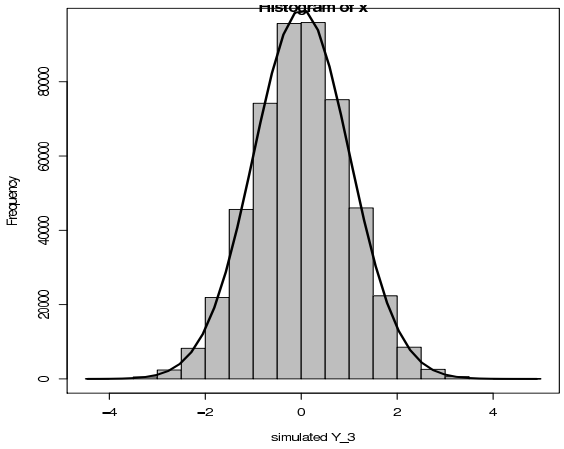
Figure A.14: The histogram plots, together with the normal curve, of $\tilde{Y}_{13} - \tilde{Y}_{15}$ for EURUSD



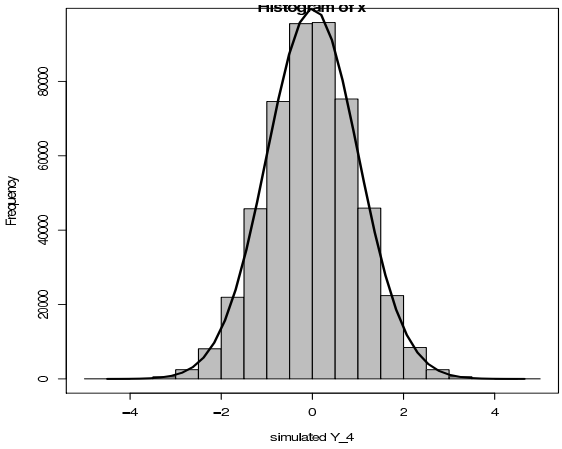
(a) Histogram of $\tilde{Y}_1$



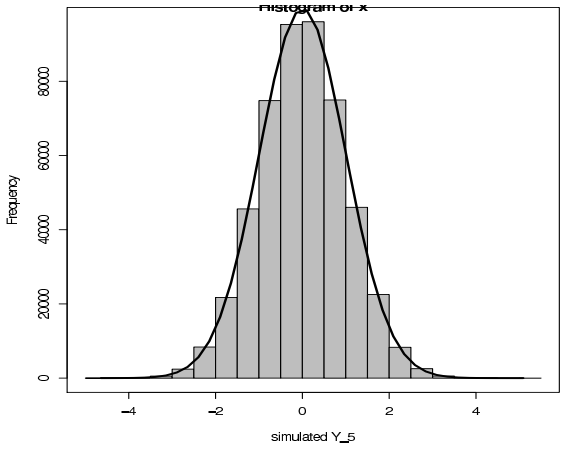
(b) Histogram of $\tilde{Y}_2$



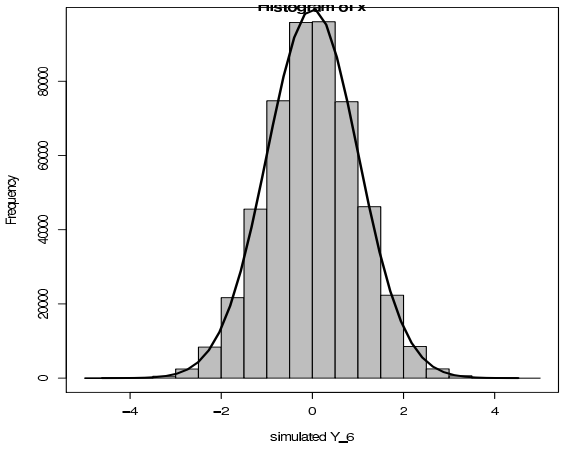
(c) Histogram of $\tilde{Y}_3$



(d) Histogram of $\tilde{Y}_4$

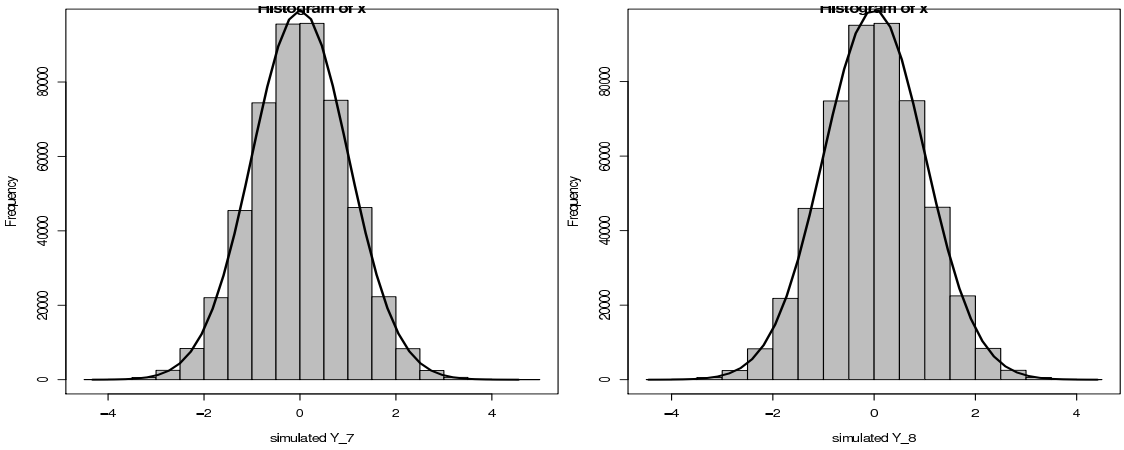


(e) Histogram of $\tilde{Y}_5$



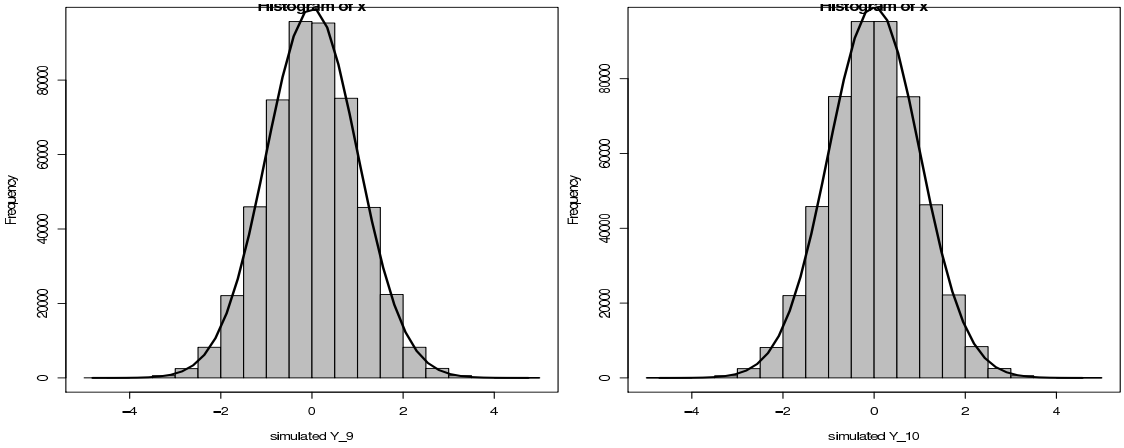
(f) Histogram of $\tilde{Y}_6$

Figure A.15: The histogram plots of $\tilde{Y}_1 - \tilde{Y}_6$ for EURSEK



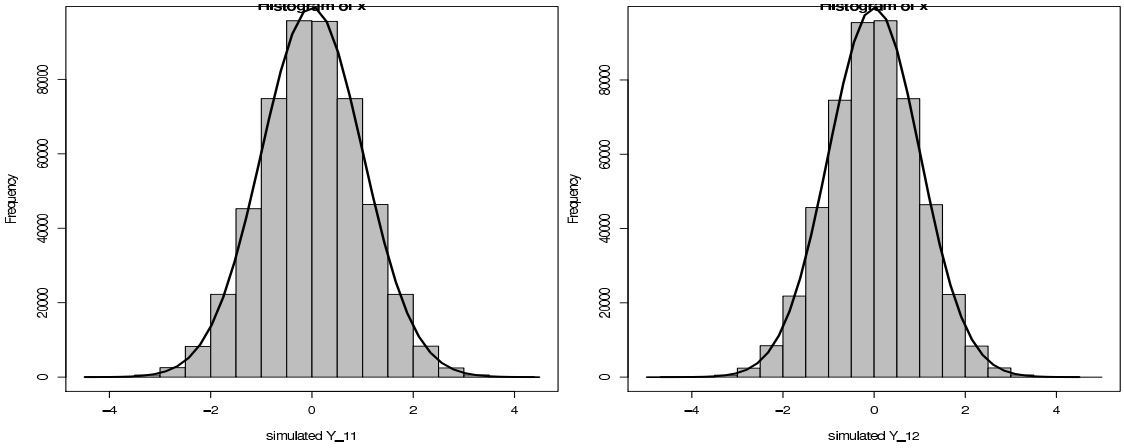
(a) Histogram of $\tilde{Y}_7$

(b) Histogram of $\tilde{Y}_8$



(c) Histogram of $\tilde{Y}_9$

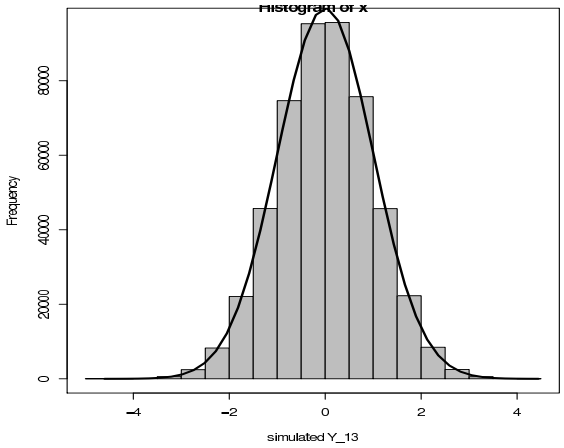
(d) Histogram of $\tilde{Y}_{10}$



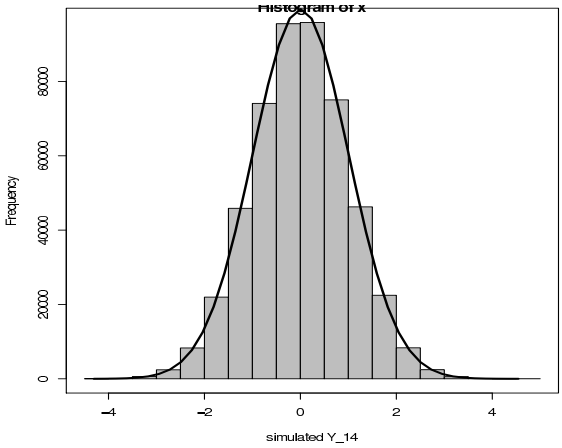
(e) Histogram of $\tilde{Y}_{11}$

(f) Histogram of $\tilde{Y}_{12}$

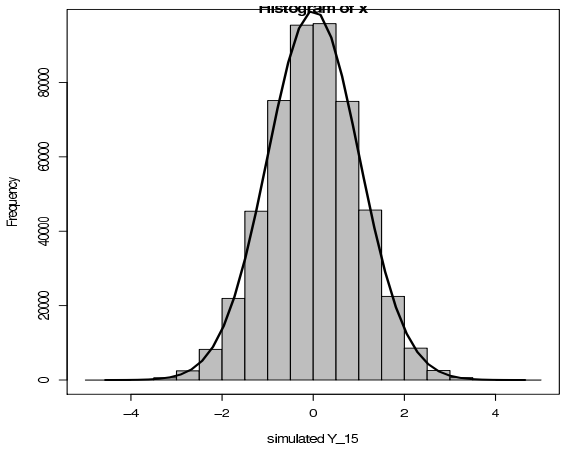
Figure A.16: The histogram plots of $\tilde{Y}_7 - \tilde{Y}_{12}$ for EURSEK



(a) Histogram of $\tilde{Y}_{13}$



(b) Histogram of $\tilde{Y}_{14}$



(c) Histogram of $\tilde{Y}_{15}$

Figure A.17: The histogram plots of $\tilde{Y}_{13} - \tilde{Y}_{15}$ for EURSEK

Bibliography

[1] Akaike, H. (1974) A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, **19**, 716–723.

[2] Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994) Time series analysis: forecasting and control. Prentice-Hall, Inc, Upper Saddle River, New Jersey, third edn.

[3] Box, G. E. P. and Pierce, D. A. (1970) Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, **65**, pp. 1509–1526.

[4] Brockwell, P. J. and Davis, R. A. (2006) Time Series: Theory and Methods. Springer Science + Business Media, LLC, New York, NY, second edn.

[5] FXStreet, FxStreet historical exchange rate. `http://www.fxstreet.com/forex-tools/rate-history-tools/`, accessed: 2014-11-15.

[6] Hannan, E. J. (1960) Time Series Analysis. Methuen & Co Ltd, London, Great Britain.

[7] Hyndman, R. J. and Khandakar, Y. (2008) Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software*, **27**, 1–22.

[8] Mardia, K. V. (1970) Measures of multivariate skewness and kurtosis with applications. *Biometrika*, **57**, 519–530.

[9] Nelsen, R. B. and (e-book collection), S. (2006) An introduction to copulas. Springer, New York.

[10] Pearlman, J. G. (1980) An algorithm for the exact likelihood of a high-order autoregressive- moving average process. *Biometrika*, **67**, pp. 232–233.

[11] Tsay, R. S. (2010) Analysis of Financial Time Series. John Wiley & Sons Inc, Hoboken, New Jersey, third edn.