Charles University in Prague
Faculty of Social Sciences
Institute of Economic Studies

BACHELOR THESIS

Switzerland as a Safe Haven:
Does the Foreign News Matter?

Author: David Kühnl
Supervisor: Mgr. Iuliia Brushko M.A.
Academic Year: 2014/2015
Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 14, 2015

__________________________
Signature
Acknowledgments

I am grateful especially to my supervisor Iuliia Brushko for her kind guidance and helpful advice.
Bibliographic Record


Character count: 66.028
Abstract

This thesis investigates the relationship between financial news and Swiss franc exchange rate in the context of Switzerland being safe haven for European investors. We employ the ARMA-GARCH econometric model extended by our custom component called “Floating Returns” to estimate the reaction of the investors to particular financial news. We find out that the bad news lead to significant short-term appreciation of the Swiss franc. Furthermore, we find out that not only the real macroeconomic data but also the investors’ expectations are important for exchange rate determination. Finally, our model quantify the reaction to the particular news depending on the expected values and the announced values.

Keywords
Swiss franc, financial news, exchange rate, FOREX, ARMA-GARCH

Author’s e-mail kuhnl.david@gmail.com
Supervisor’s e-mail iuliia.brushko@cerge-ei.cz

Abstrakt


Klíčová slova Švýcarský frank, finanční zprávy, devizový kurz, FOREX, ARMA-GARCH

E-mail autora kuhnl.david@gmail.com
E-mail vedoucího práce iuliia.brushko@cerge-ei.cz
Bachelor Thesis Proposal

<table>
<thead>
<tr>
<th>Author</th>
<th>David Kühnl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supervisor</td>
<td>Mgr. Iuliia Brushko M.A.</td>
</tr>
<tr>
<td>Proposed topic</td>
<td>Switzerland as a Safe Haven: Does the Foreign News Matter?</td>
</tr>
</tbody>
</table>

**Topic characteristics** Due to its stability Swiss Franc is commonly used as a safe haven currency among many investors. Therefore in times of decreased performance of economy or even crisis, it may be exposed to economic pressure which could lead to a significant appreciation as the investors may seek Swiss Franc based assets in order to avoid risk. The overall performance of economy can be approximated by the indicators as CPI, unemployment rate etc., which are periodically reported by the central banks and other institutions in the form of financial news. The purpose of this thesis is to examine the relationship between this kind of news and the Swiss Franc exchange rate volatility. In order to reach the goal, I will describe the forex market and the hypothetical behavior of the investors and traders after the news are released. I will also go through the major macro indicators and I will select those which should significantly affect the Swiss Franc exchange rate.

**Hypotheses** I will focus on the following questions: Can the appreciation or depreciation be forecasted by the foreign news? Does the expectations play any role? Can a linear relationship between the Swiss Franc exchange rate change and certain news be established?

**Methodology** To model the exchange rate we employ the ARMA-GARCH econometric model.

**Outline**
1. Introduction

2. Selecting crucial for forex market news

3. Description of the model and the data

4. Empirical analysis

5. Description of the results

6. Conclusion

Core bibliography


Contents

Thesis Proposal vi
List of Tables x
List of Figures xi
Acronyms xiii

1 Introduction 1

2 Switzerland as a Safe Haven 3
  2.1 Perceived Riskiness of the Country ......................... 3
    2.1.1 GDP ........................................ 3
    2.1.2 Unemployment .................................... 4
    2.1.3 Human Development Index .......................... 5
  2.2 Safe Haven and Carry Trade .............................. 6
  2.3 Risk Aversion Indicators ............................... 7
    2.3.1 Simple Indicators .............................. 7
    2.3.2 Aggregate Indicators ........................... 8

3 Hypothesis: Exchange Rate and the News 10

4 Data 12
  4.1 Exchange Rate ..................................... 12
  4.2 News .......................................... 12
    4.2.1 Euro Area .................................. 13
    4.2.2 Germany ................................... 17
  4.3 Data Summary .................................... 19

5 Model 21
  5.1 Primary Analysis and Transformation .................... 21
Contents

5.1.1 Stationarity ................................................. 21
5.1.2 Autocorrelation ............................................. 23
5.2 Modeling of Mean .............................................. 26
  5.2.1 Autoregressive Integrated Moving Average Model .... 26
  5.2.2 Floating Returns Extension .............................. 27
  5.2.3 Fitting the ARIMA Model ................................. 30
5.3 Modeling of Volatility ........................................ 38
  5.3.1 Autoregressive Conditional Heteroskedascity Model ... 40
  5.3.2 Generalized Autoregressive Conditional Heteroskedascity
      Model ....................................................... 41
5.4 Fitting the ARIMA Model with GARCH Disturbances ....... 41
  5.4.1 Identification ............................................ 42
  5.4.2 Estimation ................................................. 42
  5.4.3 Diagnostics ............................................... 42
5.5 External Variables ............................................ 44
5.6 Summary of the ARIMAX GARCH Fitting Procedure ....... 46

6 Results ......................................................... 48
  6.1 Autoregressive Part ......................................... 48
  6.2 Floating Returns ............................................ 49
  6.3 News ......................................................... 50

7 Conclusion ..................................................... 53

Bibliography ....................................................... 57

A Selected Figures ............................................... 1
List of Tables

2.1 Top 5 Countries by HDI ........................................ 5
2.2 LCVI Composition ............................................. 8
3.1 Intra-day Exchange Rate Determination ....................... 10
4.1 Included News Variables ...................................... 19
4.2 EURCHF Exchange Rate Time Series ......................... 20
5.1 EURCHF Returns Q-Test Results .............................. 26
5.2 ACF & PACF Patterns .......................................... 33
5.3 ARMA Specifications ........................................... 37
5.4 ARMA Results .................................................. 38
5.5 GARCH Specifications ........................................... 43
5.6 GARCH Results .................................................. 43
5.7 Final Model Summary .......................................... 45
5.8 Final Model Q-Test Results .................................... 45
6.1 Results of the Autoregressive Part ......................... 48
6.2 Floating Returns Results ..................................... 50
6.3 Euro Area News Results ....................................... 50
6.4 German News Results ........................................... 51
List of Figures

2.1 Switzerland’s Dominance in Terms of GDP .......................... 4
2.2 Low Unemployment in Switzerland ................................. 5
2.3 Switzerland: Strong International Position ....................... 7

4.1 CPI Differences .................................................. 15

5.1 Non-Stationary EURCHF Exchange Rate .......................... 22
5.2 Stationary EURCHF Returns ..................................... 23
5.3 Autocorrelation among the EURCHF Returns .................. 25
5.4 Floating Returns in Different Timeframes ........................ 29
5.5 Autocorrelation of AR(1) Process ................................. 32
5.6 Autocorrelation of MA(1) Process ................................. 32
5.7 Autocorrelation of ARMA(1,2) Process ............................ 33
5.8 ARMA(1,2) – Misleading ACF & PACF Pattern ................ 34
5.9 Autocorrelation among the EURCHF Returns ................. 35
5.10 Volatility Clustering among the Residuals ...................... 39
5.11 Autocorrelation among Squared Residuals ...................... 39
5.12 ARIMAX – GARCH Modelling Guide ............................. 47

6.1 Exchange Rate Fluctuation ...................................... 49
6.2 Tension Induced Irrationality .................................... 52

A.1 ARIMA Model #21 - No Significant Autocorrelation among the Residuals ................................................ I
A.2 ARIMA-GARCH Model #7 - No Significant Autocorrelation among the Residuals ........................................ II
A.3 ARIMA-GARCH Model #7 - No Significant Autocorrelation among the Squared Residuals ............................ III
A.4 Final Model - No Significant Autocorrelation among the Residuals IV
A.5 Final Model - No Significant Autocorrelation among the Squared Residuals
Acronyms

ACF  Autocorrelation Function
ARCH Autoregressive Conditional Heteroskedascity
ARIMA Autoregressive Integrated Moving Average
CBOE Chicago Board Options Exchange
CHF  Swiss franc
CPI  Consumer Price Index
EA   Euro area
EC   European Commission
ECB  European Central Bank
ESI  Economic Sentiment Indicator
EU   European Union
EURCHF Euro-Swiss franc exchange rate
FOREX Foreign exchange
GARCH Generalized Autoregressive Conditional Heteroskedascity
GDP  Gross Domestic Product
HDI  Human Development Index
LCVI Liquidity, Credit and Volatility Index
MoM  Month-over-Month
NAIRU Non-Accelerating Inflation Rate of Unemployment
NFA  Net Foreign Assets
PACF Partial Autocorrelation Function
PDF  Probability Distribution Function
PMI  Purchasing Managers Index
PPI  Producer Price Index
QE    Quantitative Easing
QoQ   Quarter-over-Quarter
SBD   Statistisches Bundesamt Deutschland
SNB   Swiss National Bank
UBS   Union des Banques Suisses
UIP   Uncovered Interest Rate Parity
VIX   Volatility Index
YoY   Year-over-Year
Chapter 1

Introduction

Switzerland is generally known for its trusted financial institutions. With one of the highest standards of living\(^1\) and not being a member of the European Union, Switzerland appears to be an independent island of stability in the Central Europe, which is often regarded as a safe haven — country sought-after by risk averse investors during the periods of increased financial risk.

We believe that the investors form expectations and make decisions based on the current state of economy. One way how to describe the state of economy is using various macroeconomic indicators such as rate of unemployment, gross domestic product, or rate of inflation. Majority of the indicators is usually published in the form of periodically released financial news. The Efficient Market Hypothesis (Fama 1970) suggests that financial market instantly adjusts to reflect all the available information. The investors immediately adjust their demand for the assets as the news are published. On the other hand, a few recent papers, for instance Bacchetta & van Wincoop (2009), Sarno & Valente (2009) and Nason & Rogers (2008), suggest random walk nature of exchange rates and speak about unstable relationship between the exchange rates and the macro fundamentals. Our goal is to find out whether such relationship can be established at least in the stable environment of safe haven country. We, therefore, examine the relationship between the Swiss franc exchange rate and the particular financial news. In other words, we would like to know how much the exchange rate changes when, for instance, the current unemployment figure is published. In case of success, the findings could be used in a profitable short-term trading strategy or for planning adjustments of short term monetary policy.

---

\(^1\)Source: United Nations - Human Development Index
1. Introduction

The thesis is structured as follows: Chapter 2 evaluates Switzerland as a safe haven country using three different approaches. Chapter 3 elaborates the hypothesis in detail and Chapter 4 describes the complete dataset we are going to analyze. Chapter 5, then, covers the methodology and all the tools used for building and estimating the econometric model. Moreover, simultaneously with the theory, we describe the process of modeling in practice. The results of the estimation are described in the Chapter 6. Finally, Chapter 7 concludes and provides suggestions for further improvements and future research possibilities.
Chapter 2

Switzerland as a Safe Haven

Before we proceed with the main analysis, we will examine the features which make the country safe haven and we will find out whether it is right to consider Switzerland to be a safe haven country.

The following section is dedicated to the three possible approaches one can use to study a given country for being a safe haven.

2.1 Perceived Riskiness of the Country

The first approach deals with the expected behavior of a risk averse investor during the period of increased risk. It is simply assumed that when the investors are looking for a safe place to invest money, they prefer countries which are sufficiently economically sound. The economic soundness is usually determined by such indicators as the Gross Domestic Product (GDP) growth, the level of unemployment, or the Human Development Index (HDI).

2.1.1 GDP

Figure 2.1 depicts the annual growth of GDP of Switzerland during the period from 2000 to 2013 and provides the comparison with such benchmarks as Germany and the Euro area (EA). Germany and the Euro area were taken as the benchmarks, since Germany is one of the best performing economies of the European Union and the Euro area is of interest for us since we are going to analyze the behavior of EURCHF currency pair.

Although Switzerland was falling slightly behind the Euro area, it performed better than Germany most of the time during the first three years and a half. Starting from the summer 2003, Switzerland has continuously over-performed
both benchmarks with only one exception when Germany managed to considerably recover from the global financial crisis 2007-2008. But the attempt to overtake Switzerland did not last long and recently Switzerland has been able to reclaim and hold its leading position.

**Figure 2.1:** Switzerland’s Dominance in Terms of GDP

![Figure 2.1: Switzerland’s Dominance in Terms of GDP](image)

*Source: World Bank - data of GDP*

Most important for this study is the period immediately after the current crisis started. Figure 2.1 clearly shows that while Germany plummeted almost to -6%, Switzerland, on the other hand, bottomed out only slightly below the -2% level. This may imply that Switzerland potentially could be deemed to possess some better investment opportunities.

### 2.1.2 Unemployment

Figure 2.2 provides the comparison of the unemployment rate of the same group of countries over the same period. The unemployment in Switzerland remains at a fairly moderate level below the 5% threshold during the whole period. At the same time, Germany has significantly higher rate of unemployment, however, it seems to converge to the 5% level as well. In contrast, the Euro
area has been recently heading quite the opposite direction towards the value of 13%.

**Figure 2.2: Low Unemployment in Switzerland**

![Graph showing low unemployment in Switzerland compared to Germany and the Euro Area.](image)

*Source: World Bank - data of unemployment*

### 2.1.3 Human Development Index

As the last measure, we chose the Human Development Index provided by the United Nations, which is widely used for comparison of the standards of living across different countries. Table 2.1 provides the list of countries with the five highest HDI from 2000 to 2013. As the table shows, Switzerland stays in the top five countries with the highest HDI during the whole period.

**Table 2.1: High HDI of Switzerland**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Norway</td>
<td>Norway</td>
<td>Norway</td>
<td>Norway</td>
<td>Norway</td>
<td>Norway</td>
<td>Norway</td>
</tr>
<tr>
<td>Australia</td>
<td>Australia</td>
<td>Australia</td>
<td>Australia</td>
<td>Australia</td>
<td>Australia</td>
<td>Australia</td>
</tr>
<tr>
<td>Sweden</td>
<td>Switzerland</td>
<td>United States</td>
<td>Switzerland</td>
<td>Switzerland</td>
<td>Switzerland</td>
<td>Switzerland</td>
</tr>
<tr>
<td>Switzerland</td>
<td>United States</td>
<td>Switzerland</td>
<td>United States</td>
<td>Netherlands</td>
<td>Netherlands</td>
<td>Netherlands</td>
</tr>
<tr>
<td>United States</td>
<td>New Zealand</td>
<td>Switzerland</td>
<td>Germany</td>
<td>Netherlands</td>
<td>United States</td>
<td>United States</td>
</tr>
</tbody>
</table>

*Source: United Nations - Human Development Index*
Assuming that continuous growth of GDP, low rate of unemployment and high standard of living is sufficient evidence, we can say that Switzerland appears to be a country with potential to provide stable and relatively low-risk investment environment. Therefore it can serve as a financial shelter for risk averse investors during the incidental downswings of the global economy.

2.2 Safe Haven and Carry Trade

The second approach is tightly linked to the concept of carry trade which has been addressed by a major body of literature. Carry trade occurs when one uses funds borrowed in the low-yielding currency (the funding currency) to obtain a higher-yielding asset denominated in a different currency. As the literature suggests, the safe haven currency is suitable to fund the carry trade for the risk-less nature of the safe haven country is usually reflected by low returns on its currency. For instance, Ranaldo et al. (2010) suggest to use explicitly Swiss franc with average annual returns -1.7%.

Furthermore, Habib & Stracca (2011) examine behavior of multiple currencies in order to find out the common fundamentals of safe haven currencies. They arrive at conclusion that the safe haven countries, currencies of which are used to fund carry trade, have often positive Net Foreign Assets (NFA) position. The NFA position is defined as difference between foreign assets and foreign liabilities. Positive NFA position suggests relatively high stability as the assets are distributed abroad, hence, the risk is diversified.

Figure 2.3 illustrates comparison of the Net Foreign Assets of the Euro area and Switzerland, proportional to their GDP. The NFA to GDP ratio of Switzerland is above the 20% level during almost the whole period and has been rallying recently. The NFA to GDP ratio of the Euro area, on the other hand, is significantly lower and stays in the 0-10% range all the time. The near-balanced position of the Euro area as a whole is due to the negative NFAs maintained in several previous years by such members as France, Spain, and Ireland.1 The dominance of Switzerland over the Euro area points out its plausible stability advantage.

2. Switzerland as a Safe Haven

Lastly, we should check whether Switzerland is actually used as a safe haven. One way how to do that is to look for any significant evidence of increased demand for the Swiss currency in the time of high financial stress. If we assume the simple Supply-Demand model, we know that an increase in demand will probably cause an upward pressure on price. Furthermore, for currencies, exchange rate is an equivalent of price, hence, we are simply looking for an appreciation of the Swiss franc during the periods of increased risk.

Generally, we distinguish two types of indicators assessing the overall risk aversion based on the method of their construction.

2.3 Risk Aversion Indicators

2.3.1 Simple Indicators

The changes in investors’ perception of the risk can be tracked by some economic indicators. For instance, the series of gold price can be used as such indicator since the elevated risk may lead the investors to reallocate wealth to assets perceived as relatively safe. (Illing & Meyer 2012)(Misina & Tkacz 2009)

Other method is to assess the risk by taking one of the volatility measures of the option market. One of such volatility measures is the Volatility Index (VIX)
created by the Chicago Board Options Exchange (CBOE) based on the S&P 500 index. (Coudert & Gex 2006)

To show risk-induced appreciation of the Swiss franc, we compare the correlation between the exchange rate and the price of gold between years 1997 and 2015. Indeed, there appears to be strong negative correlation -83% which implies that while the gold price rises (increased stress) the EURCHF exchange rate declines (appreciation).\(^2\)

### 2.3.2 Aggregate Indicators

Aggregate indicators are composite measures which are constructed by aggregating elementary series such as various yield spreads, commodity prices, or market volatility. The most often used are Merill Lynch’s financial stress index, JP Morgan’s Liquidity, Credit and Volatility Index (LCVI) or the Union des Banques Suisses (UBS) risk index. (Coudert & Gex 2006)

We have decided to use the LCVI for it contains neither the price of gold, used above, nor the Swiss franc exchange rate, subject of our interest. The detailed composition of the LCVI can be found in Table 2.2.

<table>
<thead>
<tr>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spreads on U.S. high-yield bonds</td>
</tr>
<tr>
<td>U.S. swap rates</td>
</tr>
<tr>
<td>U.S. Treasury bid/ask spreads</td>
</tr>
<tr>
<td>Spreads on emerging-market bonds</td>
</tr>
<tr>
<td>VIX</td>
</tr>
<tr>
<td>Implied currency volatilities</td>
</tr>
<tr>
<td>Global Risk Appetite Index</td>
</tr>
</tbody>
</table>

*Source: Illing & Meyer (2012)*

We again employ the measure of correlation. Not surprisingly, there appears to be considerable negative correlation, -78%, between the indicator and the EURCHF exchange rate. The strong correlation is another evidence of Swiss franc being safe haven currency.

\[\star\star\star\]

\(^2\)Note that in the case of EURCHF, Euro is the base currency and therefore the declining value means appreciation of the Swiss franc.
2. Switzerland as a Safe Haven

At this moment, we have shown the key assumption of this paper that Switzerland can be considered a safe haven country and having that done, we can proceed with the main analysis of this study.
Chapter 3

Hypothesis: Exchange Rate and the News

The Foreign exchange (FOREX) market is said to be one of the most volatile implying a lot of movement in the currency prices throughout the day. That inevitably raises the question about how one can exploit such a volatility to gain profits on a daily basis. But for doing that the successful forecast of exchange rate is required, which might be a tough thing to do. According to the recent survey published by Rosenberg (2003), more than 60% of the FX dealers were skeptical about the intra-day predictability of the exchange rates. Furthermore, almost nobody believed that the intra-day changes are driven by the economic fundamentals. This makes the idea of short term forecasting even less attractive. On the other hand, the dealers reported that the majority (approx. 85%) of the changes through the day is induced by overreaction to the news, by bandwagon effects, or by speculative forces. The overview of the survey results can be found in Table 3.1 below.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overreaction to News</td>
<td>32.8%</td>
</tr>
<tr>
<td>Bandwagon Effects</td>
<td>29.3%</td>
</tr>
<tr>
<td>Speculative Forces</td>
<td>25.3%</td>
</tr>
<tr>
<td>Technical Trading</td>
<td>10.3%</td>
</tr>
<tr>
<td>Economic Fundamentals</td>
<td>0.6%</td>
</tr>
<tr>
<td><em>Other</em></td>
<td>1.7%</td>
</tr>
</tbody>
</table>

Source: Rosenberg (2003)
In the light of the survey findings, the following scenario is likely to happen. At first, the news are announced, however, the numbers are different than expected. Consequently, the first investors start selling or buying the currency. The change in demand pushes the price up or down. The sudden price change makes the rest of the investors and speculators think that some players on the market are better informed or have some private information and therefore they “hop on the bandwagon”. The price movement then accelerates even more.

Although the initial change in the economic indicator could be insignificant, the amplifying chain-reaction can probably end as a rather significant appreciation or depreciation of the given currency.

That is why the goal of this paper is to analyze whether the information arrivals can lead to the overreaction on the FOREX market.
Chapter 4

Data

4.1 Exchange Rate

Since this work is trying to study whether the Switzerland is a safe haven for European investors, we choose EURCHF exchange rate as the main variable of our interest.

The reaction to the news mostly takes place within half an hour with the most significant change usually occurring within the first ten minutes after the announcement. We, therefore, use the one minute as the time interval for our analysis and our primary dataset contains EURCHF bid exchange rates from January 2001 to January 2015 reported on the basis of GMT0 time zone.

4.2 News

Having described the exchange rate time series, we can proceed with describing the news data. Almost every day there is an economic indicator, survey, or opinion released. Of course not all the indicators are equally important. Such indicators as employment, GDP, or industrial production attract attention of many investors and can cause substantial disruptions in the market. On the contrary, the other indicators such as some regional indexes, hardly ever affect the global market.

Selecting the news that are considered as important is a subjective process and it is based mostly on reasoning stemming from macroeconomics laws and common sense. The variables exploited in this research are described in the following section.
4.2.1 Euro Area

Since we assume Switzerland to be safe haven for Europe, we want to analyze the EURCHF currency pair. As the first group of indicators we are taking those, which describe the economic performance of the Euro area. These indicators can be divided in two groups:

1. **Primary indicators**—those, which are commonly used to evaluate the performance of the economy such as GDP, Unemployment, or Consumer Price Index (CPI).
2. **Secondary indicators**—those, which are not constructed on the real numbers but rather represent expectations and perceived state of the individual agents of the economy summarized by surveys or polls.

**Primary Indicators**

**Gross Domestic Product** The Gross Domestic Product is a measure broadly used to assess the overall performance of any economic unit. It sums up the value of all goods and services produced within the unit. One of the ways how to compute GDP is according to the following formula:

\[ Y = C + I + G + (X - M), \]

where \( C \) — consumption, consisting of durable goods, non-durable goods and services, represents all goods and services bought by households within the given country; \( I \) — investment are goods bought to be used in future to create wealth; \( G \) — government spendings reflects the value of the goods and services bought by the government, e.g. public services, military expenditure, etc.; \( X-M \) — net exports which stand for the exports net of the imports.

The Euro area GDP as an economic indicator is published by the Eurostat. It is reported in two forms: Quarter-over-Quarter (QoQ) and Year-over-Year (YoY). QoQ captures the change in the GDP between two consecutive financial quarters of the same year (e.g. Q2 2015 and Q1 2015). The YoY data, on the other hand, captures the change between the same quarters of two consecutive years (e.g. Q1 2015 and Q1 2014). Furthermore, it was empirically proven that the GDP exhibits certain seasonality.(see Boxall et al. 2009) It rises during the year and then it falls between the fourth and the first quarter of the following year. In order to remove the undesirable seasonal pattern, the figures reported by the Eurostat are seasonally adjusted preceding the release.
For more information on the seasonal adjustment of the financial time series see for example Boxall et al. (2009).

**Consumer Price Index** The Consumer Price Index is an indicator used to determine the rate of inflation and general cost of living in the given economic area. CPI is constructed as an aggregate price of goods and services contained in the predetermined consumption basket proportional to the price of the same consumption basket measured in particular base year. The items in the basked are assigned individual weights to reflect the consumers’ preferences for different goods and services. Generally, CPI can be expressed by the following formula:

\[
CPI = \frac{\sum_{i=1}^{n} w_i P_{Ai}}{\sum_{i=1}^{n} w_i P_{Bi}}
\]

where \( w_i \) are the respective weights, \( P_{Ai} \) are the actual prices and \( P_{Bi} \) are the prices recorded in the base year.

For us, the CPI indicator is important especially for one reason. During the recent past, increasing number of central banks (including the European Central Bank (ECB)) have been employing the framework of inflation targeting. They adjust the monetary policy in order to reach certain inflation target.\(^1\) This indicator is of interest for us since the comparison of target and real CPI levels may build the investors’ expectations about the central bank policies. In particular, we are interested in such central bank interventional technique as Quantitative Easing (QE), which can substantially affect the exchange rate.

Like the GDP, the Euro area CPI is published by the Eurostat and it is reported as YoY and MoM. Additionally, Eurostat measures also the core CPI. The core CPI aggregates the same basket excluding alcohol, tobacco, food, and energies, for prices of those goods are usually excessively volatile. Figure 4.1 depicts the comparison of the two measures in the period between years 2007 and 2015. Although the long term trend is more or less similar, in the short run, however, the differences can be up to 2 percentage points.

**ECB Interest Rate** Central bank interest rate is the rate the central bank provides short-term liquidity to the commercial banks at . Interest rate is one of the key tools the central bank can use to intervene in the economy. If the

\(^1\)More about inflation targeting can be found for example in Hammond (2012).
interest rate, the price of saving, is set high, saving becomes more attractive than spending. If, on the other hand, interest rate is low or even negative, things come the other way around. Consequently, by adjusting rates, central banks have the ability to some extent boost or slow down the economy.

Apart from affecting the domestic consumption, the interest rate is also important for us since it may lead to the exchange rate appreciation or depreciation. According to the Uncovered Interest Rate Parity (UIP), the difference between interest rates in two countries is balanced by the exchange rate of their currencies. The relationship can be written as:

\[ i_A - i_B = E(e), \]

where \( i_A \) is the interest rate of country A, \( i_B \) is the interest rate of country B, \( E(e) \) — the expected change in exchange rate.

Nowadays, the US FED is expected to raise interest rate over the few following months. If they do so, to make profit, investors start to borrow in low-rate countries, for instance in Japan, and they invest the cheap money into the US at the higher interest rate. However to do so, they have to exchange
the yens for US dollars. In other words buy dollars for Yens. According to the law of demand and supply, it inevitably leads to appreciation of the US dollar relatively to the yen.

Clearly, the interest rate is an important factor for the exchange rate determination. The decisions about its level are periodically announced by the European Central Bank.

**Rate of Unemployment** Another essential characteristics of any economy is the rate of unemployment. It is measured as the amount of unemployed people proportional to the total labor force. The Euro area rate of unemployment is monitored by the Eurostat and it is published on a monthly basis.

As most economist agree, in reality, there is a trade-off between unemployment and inflation. This belief is supported both by theory and empirical evidence. Study of the relationship of unemployment and inflation revolves around the Phillips curve and the NAIRU, an abbreviation for *Non-Accelerating Inflation Rate of Unemployment*. The theory behind NAIRU states the impossibility to keep the rate of unemployment below a certain threshold without accelerating the growth of inflation.

Crucial for this paper is that some low positive value of rate of unemployment is a sign of sound and growing economy. Significantly low or high figure, on the other hand, may be an incentive for central bank to pursue an interventional policy which, among other things, usually affects the exchange rate.

For more information about Phillips curve, NAIRU and the rate of unemployment see for example Turner *et al.* (2001).

**Retail Sales** One of the concerns of every central bank is the amount of consumer spending. Excessive consumption may create an inflationary pressure on the economy. Too low spending, on the other hand, could be a sign of economic recession or even depression, if it persists.

Retail Sales is an aggregate measure that reflects the amount of goods and services sold in the retail sector within the economy. The retail sales for the Euro area are monitored by the Eurostat as a Year-over-Year percentage change.

**Secondary Indicators**

**Consumer Confidence** The consume confidence indicator summarizes the general opinion of the consumers on the state of economy in the following 12
months and it is periodically published by the European Commission (EC). The data are collected by a survey consisting of several questions aimed to determine the expectation of individuals about various aspects of the economy. For instance, the respondents answer questions about the rate of unemployment, the financial situation of the households, etc.

**Economic Sentiment Indicator**  As the consumer confidence, the Economic Sentiment Indicator (ESI) depicts the expectations about the future state of economy. It also falls under the competence of the European commission. The range of respondents is, however, broader than in the case of the consumer confidence. The ESI is a composite indicator and it contains information from the following 5 sectors:

- industry
- services
- consumers
- construction
- retail trade

The details about the methods, construction and computation of the both indicators, including the survey questions, can be found in the EC (2014).

### 4.2.2 Germany

Germany is one of the most powerful economies of the EU and therefore, it is reasonable to think that it has a substantial influence on the whole EU. Providing almost 30bn EUR to the EU budget, Germany is the leading contributor and outperforms second France by nearly 6bn EUR.\(^2\) It is no doubt that downturn of German economy would substantially decrease the entire EU budget. Therefore, additionally to the data of the Euro area, we as well include a few indicators from the German economy.

Because of the similar nature of the German and the EA data, we present most of the indicators in a list without further explanation. Only those that did not appear in previous section are briefly commented.

- **GDP**—YoY, QoQ, seasonally adjusted, published by the Statistisches Bundesamt Deutschland (SBD).

\(^2\)Source: European Commission
4. Data

- **CPI**—YoY, MoM, published by the SBD.
- **Rate of Unemployment**—seasonally adjusted, released by the Bundesagentur für Arbeit.
- **Retail Sales**—YoY, MoM, seasonally adjusted, published by the SBD.

**Industrial Production & Manufacturing Purchasing Managers Index** Although the importance of the industry is declining while that of the service sector is increasing, it still employs over seven million people and its condition is, therefore, crucial for the whole country.\(^3\)

The industrial production indicator measures the performance and outputs of the German mines and factories. Both seasonally adjusted MoM and unadjusted YoY, the figures are published by the Statistisches Bundesamt Deutschland.

The second indicator, Manufacturing Purchasing Managers Index (PMI), collected by the non-governmental commercial organization *Markit Economics* evaluates the condition of the manufacturing sector. The PMI takes into account such data as output, the amount and volume of orders, stock balance, and employment, reported by the selected private enterprises. The result are published monthly and since Markit Economics operates internationally, the index offers comparison of multiple countries across the world.

**Trade Balance** The trade balance figure reflects the difference between imports and exports, i.e. a positive value is a sign of trade surplus, a negative one, in contrast, shows trade deficit. Generally speaking, a trade surplus is usually seen as a positive signal suggesting steady economic growth. On the other hand, a sudden change in the trade balance could be a precursor for increased market volatility.

The data on the seasonally adjusted trade balance are collected and released by the SBD.

**Producer Price Index** The Producer Price Index (PPI) is a similar indicator as the Consumer Price Index. But unlike the CPI, the PPI does not depict price of any consumption basket, but rather the price of goods sold in the primary markets. It can be considered as an indicator of commodity inflation and therefore, the figures, published by the SBD as MoM and YoY, provide valuable information when evaluating the overall rate of inflation.

---

\(^3\) Source: Statistisches Bundesamt Deutschland
4. Data

4.3 Data Summary

Conveniently, an organization FXStreet provided us with dataset containing all the variables of our interest with values since year 2007. For every observations the dataset contains not only the actual value, but also the previous value and the consensus value.

According to the FXStreet, the consensus value is determined based on multiple forecasts and expectation surveys carried out by major banks and other financial institutions worldwide. The figure is published preceding the announcement of the actual value and can be therefore considered as a benchmark for the actual value. It is of interest, however, that the difference between the announced value and the benchmark (forecast error) is usually deeply correlated with the after-news exchange rate development. Finally, Table 4.1 summarizes all the variables we are going to include into our research.

Table 4.1: Included News Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Form</th>
<th>Adjusted</th>
<th>Publisher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euro area</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>QoQ, YoY</td>
<td>seasonally</td>
<td>Eurostat</td>
</tr>
<tr>
<td>CPI</td>
<td>MoM, YoY</td>
<td>—</td>
<td>Eurostat</td>
</tr>
<tr>
<td>Core CPI</td>
<td>YoY</td>
<td>—</td>
<td>Eurostat</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>—</td>
<td>—</td>
<td>ECB</td>
</tr>
<tr>
<td>Unemployment</td>
<td>—</td>
<td>—</td>
<td>Eurostat</td>
</tr>
<tr>
<td>Retail Sales</td>
<td>YoY</td>
<td>—</td>
<td>Eurostat</td>
</tr>
<tr>
<td>Consumer Confidence</td>
<td>—</td>
<td>—</td>
<td>EC</td>
</tr>
<tr>
<td>ESI</td>
<td>—</td>
<td>—</td>
<td>EC</td>
</tr>
<tr>
<td>Germany</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>QoQ, YoY</td>
<td>seasonally</td>
<td>SBD</td>
</tr>
<tr>
<td>CPI</td>
<td>MoM, YoY</td>
<td>seasonally</td>
<td>SBD</td>
</tr>
<tr>
<td>Unemployment</td>
<td>—</td>
<td>seasonally</td>
<td>Bundesagentur für Arbeit</td>
</tr>
<tr>
<td>Retail Sales</td>
<td>MoM, YoY</td>
<td>seasonally</td>
<td>SBD</td>
</tr>
<tr>
<td>Industrial Production</td>
<td>MoM, YoY</td>
<td>seasonally, —</td>
<td>SBD</td>
</tr>
<tr>
<td>Manufacturing PMI</td>
<td>—</td>
<td>—</td>
<td>SBD</td>
</tr>
<tr>
<td>Trade Balance</td>
<td>—</td>
<td>seasonally</td>
<td>SBD</td>
</tr>
<tr>
<td>PPI</td>
<td>MoM, YoY</td>
<td>—</td>
<td>SBD</td>
</tr>
</tbody>
</table>
Since the news data starts in 2007, our research period starts also in 2007. Furthermore, in 2011 the Swiss National Bank (SNB) introduced an exchange rate peg that fixed the value of the Swiss franc against the Euro. During the period between 2011 and January 2015, when the peg was removed, the exchange rate oscillated in a narrow range around the price of 1.2 Francs per Euro. Employing the peg, the SNB focused on balancing the exchange rate and minimizing the fluctuations. If there were any over-reactions to the news during the period, they would not be, most probably, reflected by the exchange rate change and therefore, our research period ends in 2011.

Table 4.2 describes our final exchange rate time series.

**Table 4.2: EURCHF Exchange Rate Time Series**

<table>
<thead>
<tr>
<th>Observations</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 464 913</td>
<td>1.52993</td>
<td>0.1058848</td>
<td>1.2413</td>
<td>1.6825</td>
</tr>
</tbody>
</table>
Chapter 5

Model

Since the modeling of financial time series is usually done in multiple steps, for clarity, we first provide brief description of the whole procedure. Firstly, we analyze the data and if needed, we transform them as described in Section 5.1. Once the data are suitable for our model, we proceed with modeling of mean equation using the ARIMA model. The mean equation modeling is described in Section 5.2. Then, we model the volatility equation using the GARCH model as described in Section 5.3. After we have both the mean and the volatility equation, we estimate them together (Section 5.4) and if the model is valid, we add the external variables to the mean equation as described in Section 5.5.

5.1 Primary Analysis and Transformation

5.1.1 Stationarity

One of the crucial properties for time series analysis is the stationarity. We say that stochastic process is stationary when its Probability Distribution Function (PDF) is stable over time. In other words, we say that \( \{x_t\}_{t=1}^n \) is stationary when

\[
\text{PDF of } \{x_{t1}, x_{t2}, \ldots, x_{tj}\} = \text{PDF of } \{x_{t1+k}, x_{t2+k}, \ldots, x_{tj+k}\} \quad \forall k \geq 1.
\]

Stationarity is, however, a very strict assumption. It is often violated for the time series, and especially for the financial time series, and therefore, a weaker version called covariance stationarity is assumed. The covariance stationarity holds when

\[
\forall t, h \geq 1 : \text{cov}(x_t, x_{t+h}) \text{ depends only on } h.
\]
It implies that all the values $x_t$ fluctuates around some fixed level. For more about basic properties of the time series see for instance Tsay (2005).

To evaluate the extent of stationarity of our dependent variable EURCHF exchange rate, we first plot the data and inspect it visually.

Figure 5.1: Non-Stationary EURCHF Exchange Rate

As depicted in Figure 5.1, clearly, the process does not oscillate around any fixed value and there is visible downward trend during the whole period. Hence, we can speak neither about stationarity nor about covariance stationarity of the series.

To test the stationarity, one can employ the well-known Dickey-Fuller test introduced by Dickey & Fuller (1979) or its augmented form. The augmented D-F test is sometimes more suitable for financial time series because of their autoregressive tendencies. The null hypothesis of the augmented D-F test is $H_0 : \beta = 1$ against the alternative $H_a : \beta < 1$ using the regression

$$x_t = f(t) + \beta x_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta x_{t-i} + \epsilon_t,$$  \hspace{1cm} (5.1)

where $f(t)$ is a function of the time trend. The test statistics is then computed as ordinary t-statistics

$$t = \frac{\hat{\beta} - 1}{\text{std}(\hat{\beta})},$$

where $\hat{\beta}$ is the OLS estimate of the $\beta$ from Equation 5.1.

We applied the test on the EURCHF exchange rate and obtained p-value 0.35. High p-value means we cannot reject the null hypothesis at any conventional significance level. Our time series is not stationary.

One technique commonly used to transform non-stationary time series into stationary one is called differencing. We call the time series differenced of order one, when it was obtained by subtracting the consecutive values of the underlying series. Formally, we can write the transformation of a unit root time series $\{y_t\}$ into stationary time series $\{c_t\}$ as
5. Model

\[ c_t = y_t - y_{t-1} \quad \forall t \geq 2. \]

\( \{c_t\} \) is then referred as first differenced.

Other possibility how to deal with non-stationarity is to use log returns instead of prices. Transformation of the prices to the log returns can be expressed by the following formula:

\[ r_t = \log(p_t) - \log(p_{t-1}) = \log\left(\frac{p_t}{p_{t-1}}\right) \quad \forall t \geq 2. \]

Apart from solving the stationarity problem, another advantage of the log returns is that they are a scale-free. The relative terms are more flexible than the absolute values, especially in the field of finance.

After we applied the log returns transformation on the original time series, we checked the stationarity again. First we plotted the returns and inspected visually.

Figure 5.2: Stationary EURCHF Returns

As depicted in Figure 5.2, the new series does not exhibit any visible trend and the values fluctuate in limited range around zero. Furthermore, we again run the augmented Dickey-Fuller test and we obtained p-value much smaller than 0.01. There is enough evidence to reject \( H_0 \) at the 99% level and therefore it is confirmed that the log returns are covariance stationary, which implies the log returns are suitable for further analysis.

5.1.2 Autocorrelation

Next step is to look for correlation among the returns. If the correlation is present, the preceding return would have significant influence on the present one. In that case we speak of autocorrelation and extending the model by the past returns may improve its quality.
The correlation coefficient of two random variables \( X, Z \) is defined as
\[
\rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}
\]

Furthermore, in case of a random sample with \( N \) observations, the coefficient can be estimated across the whole sample as
\[
\hat{\rho}_{X,Y} = \frac{\sum_{t=1}^{N} (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^{N} (x_t - \bar{x})^2(y_t - \bar{y})^2}}, \tag{5.2}
\]
where
\[
\bar{x} = \frac{\sum_{t=1}^{N} x_t}{N} \quad \text{and} \quad \bar{y} = \frac{\sum_{t=1}^{N} y_t}{N},
\]
are the sample means of \( X \) and \( Y \).

In the context of time series, the correlation coefficient between two consecutive values is called the first lag autocorrelation. For the correlation between \( r_t \) and \( r_{t-l} \), the term can be generalized to the lag-\( l \) autocorrelation. Given a sample \( \{r_t\}_{1}^{N} \) of log returns, we can rewrite Equation 5.2 as
\[
\hat{\rho}_l = \frac{\sum_{t=l+1}^{N} (r_t - \bar{r})(r_{t-l} - \bar{r})}{\sum_{t=1}^{N} (r_t - \bar{r})^2}. \tag{5.3}
\]

If we use the covariance stationarity property that \( \forall t, l \geq 1 : \text{cov}(r_t, r_{t+l}) \) depends only on \( l \), we get the lag-\( l \) autocorrelation as a function of single variable \( l \). The function is called the Autocorrelation Function (ACF). For more detailed derivation of the ACF see for example Tsay (2005).

Additionally to the ACF, one can derive the Partial Autocorrelation Function (PACF). The Partial Autocorrelation Function (PACF) determines the autocorrelation of the \( l \)-th lag after it was accounted for the autocorrelation of the preceding lags. In other words, partial autocorrelation of the \( l \)-th lag is the conditional correlation defined as
\[
\text{Corr}(x_t, x_{t-l} | x_{t-1}, \ldots, x_{t-l+1})
\]

Together, ACF and PACF form a powerful tool that is often used to determine whether there are serial linear relationships in the time series. One way how to use the functions is to plot them and inspect the autocorrelation visually. The autocorrelation of our log return series is depicted in Figure 5.3.
5. Model

Figure 5.3: Autocorrelation among the EURCHF Returns

Both charts suggest certain relationship among the first few lags. More specifically, there is about 10% negative correlation on the first lag followed by clearly visible correlations on the second and third lag. The values then flows near the significance threshold up to the tenth lag. Further, the autocorrelation diminishes with only occasional spikes as for instance on the 27th lag. It is very probable that we need to take the autocorrelation into account.

Apart from the visual inspection, to be sure, one can test the joint significance of autocorrelation among multiple lags using the Ljung & Box (1978) test also known as the $Q$-test. The $Q$-test is defined the following way:

$$H_0 : \text{No autocorrelation},$$
$$H_a : \text{Data not independently distributed},$$

with the test statistics

$$Q(m) = N(N + 2) \sum_{l=1}^{m} \frac{\hat{\rho}_l^2}{N - l},$$
where $N$ is the size of the sample, $m$ denotes the number of lags, $\hat{\rho}_l^2$ is the squared estimate of the coefficient of correlation of the $l$-th lag from Equation 5.3. The $Q(m)$ follows the $\chi^2$ distribution with $m$ degrees of freedom.

We applied the $Q$-test on our data up to 40 lags. The respective tests statistics are in Table 5.1.

Table 5.1: EURCHF Returns Q-Test Results

<table>
<thead>
<tr>
<th>Number of lags</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4011.5</td>
</tr>
<tr>
<td>5</td>
<td>4206.5</td>
</tr>
<tr>
<td>10</td>
<td>4261.4</td>
</tr>
<tr>
<td>20</td>
<td>4281.9</td>
</tr>
<tr>
<td>40</td>
<td>4354.9</td>
</tr>
</tbody>
</table>

In all cases, the Q values are extremely high and all the respective p-values were close to zero. Consequently, we reject the null hypothesis at 99% confidence level in all the cases. Results of the Q-test confirm that we have a serial correlation in our data. Since we have serial correlation in our data, we are going to use ARIMA model, discussed in detail in the next section.

5.2 Modeling of Mean

5.2.1 Autoregressive Integrated Moving Average Model

The autocorrelation is sometimes understood as some kind of the flaw on the data needed to be removed. Yet, a better approach exists. The ARIMA methodology allows to work with the autocorrelation and use the extra information to improve the model.

The ARIMA model was introduced by Box & Tiao (1975) as a dynamic regression model. The ARIMA($p, d, q$) model combines the autoregressive process AR($p$) defined as

$$y_t = \alpha + \sum_{i=1}^{p} \delta_i y_{t-i} + \epsilon_t,$$

where $p$ is the order of the process with the moving average MA($q$)

$$y_t = \alpha + \epsilon_t - \sum_{i=1}^{q} \gamma_i \epsilon_{t-i},$$
where the parameter \( q \) is, again, the order of the process. The I in the ARIMA abbreviation stands for *Integration*. It means that the underlying time series is integrated of order \( d \) and should be differenced to achieve stationarity.

The general ARIMA\((p, d, q)\) model can be, then, using the back-shift operator, written as

\[
(1 - \sum_{i=1}^{p} \delta_i B_i) y_t = \alpha + (1 - \sum_{j=1}^{q} \gamma_j B_j) \epsilon_t,
\]

(5.4)

where

- \( y_t \) is the response variable or its differenced form (depending on the parameter \( d \))
- \( B_i \) stands for the back-shift operator such that \( B_i(y_t) = y_{t-i} \)
- \( \epsilon_t \) is the error term which is assumed to have zero mean and constant variance

Either of the components, AR or MA, can be excluded from the model by setting the parameter, \( p \) or \( q \) to 0. The similar goes for the differencing and the parameter \( d \).

In their paper, Box & Tiao (1975) mention the possibility to extend the dynamic model by one or more exogenous regressors. Mathematically speaking, we extend Equation 5.4 as follows:

\[
(1 - \sum_{i=1}^{p} \delta_i B_i) y_t = \alpha + (1 - \sum_{j=1}^{q} \gamma_j B_j) \epsilon_t + \sum_{k=1}^{n} \beta_k x_{kt}
\]

(5.5)

Model described by Equation 5.5 is then called ARIMAX model and its advantages are the following. Inclusion of the AR and the MA term allows for free, self-deterministic floating characteristic for financial series. The set of exogenous variables accounts for the occasional shocks induced by some external inputs, in our case, the news.

### 5.2.2 Floating Returns Extension

Although the ARIMA model seems very suitable for modeling of the FOREX data, there is one imperfection. The ARIMA model does not take into account the fact that the FOREX market runs in several different time frames. That mean that the time periods, within which the exchange rate changes are
recorded, can differ. The data can be reported in terms of minutes, hours, weeks, or even months and years.

All the subjects of the market, then, look at the same data but their scopes differs. The intra-day speculator will be, most probably, influenced by the most recent price movements and will choose the minute data, the long run investor operating in the monthly time frame, on the other hand, will make his decision based on the past months, maybe even years and he will, therefore, reach for a bigger timeframe. Their demand and supply, however, meet at the same time and affect the present exchange rate.

Very important property of the FOREX market is that the starting and ending times of the reporting periods are fixed. If we choose, for instance, 15-minute data and start at 12:00, our periods will be as follows 12:00 – 12:15, 12:15 – 12:30, 12:30 – 12:45 and so on. Similarly, if we choose the 5-minute data, the periods will be 12:00 - 12:05, 12:05 – 12:10…. Furthermore, since the investors, when making decision, usually consider only the periods which are already closed, their decision information can be lagged relatively to the current time.

Let us consider the following example as depicted in Figure 5.4. Suppose that current time is 0:26 and we have 3 investors, 1-minute, 5-minute, and 15-minute, who started to plan their trade at 0:00. At 0:26 each of the investors looks at the last closed period according to the timeframe he or she operates. As we specified the periods above and as depicted in Figure 5.4, each of the investors has different information disposable. The 1-minute investor’s last closed period is the one between 0:25 and 0:26. The 5-minute investor’s last closed period is the one between 0:20 and 0:25. Finally, the 15-minute investor’s last closed period is the one between 0:00 and 0:15.
As the demand and supply of all the investors meet at the same time, regardless the timeframe, in order to improve our model, we designed a new component, which reflect the return of the i-th closed period depending on the selected timeframe. We call it the floating return, as it floats lagged behind the current time. For clarity, if we once again consider the example depicted in Figure 5.4, the first 15-minute floating return at the time 0:26 would be the return of the period between 0:00 and 0:15. Mathematically speaking it would be

\[ L_{15_{0:26}} = \log \left( \frac{P_{0:15}}{P_{0:00}} \right) \]  

(5.6)

Generally, we define the i-th floating return as

\[ \text{FR}(i, \tau) = \log \left( \frac{P_a}{P_b} \right) \],  

(5.7)

where

\[ a = t - \text{mod}(t, \tau) - (i - 1)\tau \quad \text{and} \quad b = t - \text{mod}(t, \tau) - i\tau \]
where \( i \) is the lag order; \( \tau \) is the length of the period (the timeframe) and \( t \) is the number of minutes from the beginning of our research period. \( \text{Mod}(x, y) \) is the *modulo operation* which finds the remainder after division of \( x \) by \( y \).

According to the formula, the first 15-minute floating return at time 0:26 can be analyzed as follows:

\[
i = 1, \quad \tau = 15, \quad t = 26
\]

The index \( a \) in the numerator from Equation 5.7 would then be

\[
a = t - \text{mod}(t, \tau) - (i - 1)\tau = 26 - \text{mod}(26, 15) - (1 - 1) \times 15 = 26 - 9 = 15,
\]

which results in \( P_{0:15} \). Similarly, the index \( b \) in the denominator would be

\[
b = t - \text{mod}(t, \tau) - i\tau = 26 - \text{mod}(26, 15) - 1 \times 15 = 26 - 9 - 15 = 0,
\]

resulting in \( P_{0:00} \).

Clearly, the floating return \( FR(1, 15) \) at time 0:26, then, reflects the return of the period between 0:00 and 0:15, which is exactly as in Equation 5.6.

In our model, the floating returns are referred as \( L_{##} \) where ## denominates the timeframe, for instance the first 15-minute floating return is referred as \( L_{15} \).

Having described both the ARIMA model and the Floating Returns extension, we can proceed with the next section which describes how to fit the model in practice.

### 5.2.3 Fitting the ARIMA Model

Depending on the complexity of the model, especially when it contains all three components of high orders, the ARIMA estimation can be a difficult task. Box & Tiao (1975) suggest it to be done following an iterative algorithm consisting of the following three steps:

1. Identification
2. Estimation
3. Diagnostics

In practice, we try to identify the orders of the AR, the MA, and the I component. Then, we estimate the specified model and run a diagnostic checks.
If the diagnostic checks suggest any inadequacy in the model, we re-identify the model, estimate it, and check again. We repeat the procedure until we obtain a sufficient model that fits the data the best.

More than a set of strict rules, the procedure figures like a set of guidelines. The selection itself depends usually on the nature of the data and on the researcher’s intuition. It can vary, case to case, substantially and one may sometimes struggle when looking for an inspiration in the recent literature. Conveniently, Andrews et al. (2013) published a paper about the ARIMA methodology applied on the insurance sector. It contains detailed and precise description of the estimation in practice and it guided us through the process of fitting the ARIMA model on the log return series of the EURCHF currency pair.

**Identification**

Identification of the appropriate AR and MA orders relies heavily on the analysis of the autocorrelation as described in Subsection 5.1.2. In particular, we want to include all the lags that exhibit any significant autocorrelation. The question is, whether to include AR terms, MA terms or both. The ACF and the PACF turns out to be useful, as we can determine the particular orders according to the patterns in the plots of the functions.

The rule of thumb is that when the ACF values show sharp drop toward zero and the PACF is declining, we speak of the MA presence and including the MA terms is suggested. If, on the other hand, the PACF exhibits sharp drop and the ACF is declining, we speak of the AR presence and including the AR terms is suggested.

In order to illustrate the rule, we simulated the following ARMA processes and plotted their ACFs and PACFs.

\[
\text{AR}(1): \quad y_t = 0.7y_{t-1} + \epsilon_t, \quad n = 300
\]
Figure 5.5: Autocorrelation of AR(1) Process

MA(1): $y_t = 0.8 \epsilon_{t-1} + \epsilon_t$, $n = 300$

Figure 5.6: Autocorrelation of MA(1) Process

ARMA(1,2): $y_t = 0.3 y_{t-1} + 0.8 \epsilon_{t-1} + 0.7 \epsilon_{t-2} + \epsilon_t$, $n = 300$
Figure 5.7: Autocorrelation of ARMA(1,2) Process

The rules for valid identification can be summarized as it is in Table 5.2 which was originally published by Jaggia (2010) in a paper about forecasting with the ARIMA models.

Table 5.2: ACF & PACF Patterns – Rule Of Thumb

<table>
<thead>
<tr>
<th>Model</th>
<th>ACF</th>
<th>PACF</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Noise</td>
<td>No significant lags</td>
<td>No significant lags</td>
</tr>
<tr>
<td>MA(1)</td>
<td>Cut-off after first lag</td>
<td>Direct or oscillatory decay toward zero</td>
</tr>
<tr>
<td>MA(q)</td>
<td>Cut-off after q-th lag</td>
<td>Direct or oscillatory decay toward zero</td>
</tr>
<tr>
<td>AR(1)</td>
<td>Direct or oscillatory decay toward zero</td>
<td>Cut-off after first lag</td>
</tr>
<tr>
<td>AR(p)</td>
<td>Direct or oscillatory decay toward zero</td>
<td>Cut-off after p-th lag</td>
</tr>
<tr>
<td>ARMA(p,q)</td>
<td>Decay toward zero after q-th lag</td>
<td>Decay toward zero after p-th lag</td>
</tr>
</tbody>
</table>

Source: Jaggia (2010)

Although all of the three plots above display the ideal patterns and clearly suggest the right lag orders, in reality, one can easily obtain results as in Figure 5.8 which represents a different realization of the ARMA(1,2) process with slightly adjusted coefficients.

\[
ARMA(1,2): \quad y_t = 0.9y_{t-1} + 0.9\epsilon_{t-1} + 0.2\epsilon_{t-2} + \epsilon_t, \quad n = 300
\]
According to Table 5.2 and the plot in Figure 5.8, the original ARMA(1,2) process could be easily identified, for instance, as ARMA(3,10). That only emphasizes that the fitting can be often demanding and it may take several attempts to find the right model.
Finally, the ACF and the PACF of the EURCHF log returns are depicted in Figure 5.9. Both the plots display significant drop after the first lag which appears to be slightly bigger in the case of the ACF, suggesting the MA presence. We therefore decided to start with the MA(1) model.

**Figure 5.9: Autocorrelation among the EURCHF Returns**

![ACF and Partial ACF plots](image)

**Estimation**

We estimate the ARIMA model using the maximum likelihood estimation. We distinguish two approaches of the estimation, based on how we treat the initial shocks $\epsilon_t$ when $t < 0$. The *conditional likelihood* approach assumes the $\epsilon_t, t < 0$ to be zero. The *exact likelihood* approach, on the other hand, estimates the initial shocks as additional parameters of the model. Clearly, the second approach is more precise, but it is also more demanding regarding the computation. As the difference between the approaches diminishes with the increasing sample size, we use the conditional likelihood. For more details about the estimation, please refer to Box & Tiao (1975) or to Tsay (2005).
5. Model

Diagnostics

Once the model is identified and estimated, we proceed with the diagnostic phase. The main purpose of the ARIMA model is to deal with the autocorrelation, which will therefore be our primary interest. A simple way how to find out whether all the available information was used is to analyze the residuals of the estimation defined as

$$
\epsilon_t = r_t - \hat{r}_t,
$$

where $\hat{r}_t$ stands for the fitted values from the estimation. If the residuals do not exhibit any signs of autocorrelation\(^1\), the model efficiently uses all the information. In contrast, if there are any signs of the autocorrelation among the residuals, the model should be improved. To do so, we start again with the Identification phase (see Section 5.2.3), however, instead of the log returns, we analyze the residuals.

Although the zero residual autocorrelation is a necessary condition for a valid model, it does not automatically imply the best model. Actually, in practice, multiple different models may fulfill the zero autocorrelation condition. To select the most suitable model, one can use the Akaike (1973) (AIC) and Schwarz-Bayesian (BIC) information criterion defined as

$$
\text{AIC} = 2m - 2\mathcal{L},
$$
$$
\text{BIC} = m \cdot \log(n) - 2\mathcal{L},
$$

where $m$ is the number of parameters, $n$ is the number of observations and $\mathcal{L}$ is the maximized log likelihood from the estimation.

Although neither of the criterions measures quality of the particular model in absolute terms, they allow to compare different models by penalizing over-specification. In other words, the smaller is the criterion, the better is the model compared to the others.

\(^1\)For the analysis of the autocorrelation among the residuals, we use the tools described in Subsection 5.1.2
Altogether, we evaluated 21 different specifications of ARIMA models, all of which are summarized in Table 5.3 and in Table 5.4.

Table 5.3: ARMA Specifications

<table>
<thead>
<tr>
<th>#</th>
<th>Model specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MA(1)</td>
</tr>
<tr>
<td>2</td>
<td>AR(1) nc</td>
</tr>
<tr>
<td>3</td>
<td>AR(1) MA(1) nc</td>
</tr>
<tr>
<td>4</td>
<td>AR(2) nc</td>
</tr>
<tr>
<td>5</td>
<td>AR(2) MA(2) nc</td>
</tr>
<tr>
<td>6</td>
<td>MA(2) nc</td>
</tr>
<tr>
<td>7</td>
<td>AR(3) MA(2) nc</td>
</tr>
<tr>
<td>8</td>
<td>AR(1-3,5,6) nc</td>
</tr>
<tr>
<td>9</td>
<td>AR(3) MA(4) nc</td>
</tr>
<tr>
<td>10</td>
<td>AR(1-3,5,6) MA(2) nc</td>
</tr>
<tr>
<td>11</td>
<td>AR(6) MA(2) nc</td>
</tr>
<tr>
<td>12</td>
<td>AR(6) nc</td>
</tr>
<tr>
<td>13</td>
<td>AR(6) nc</td>
</tr>
<tr>
<td>14</td>
<td>ARMA(6) nc L5-L60</td>
</tr>
<tr>
<td>15</td>
<td>AR(6) nc L15-L60</td>
</tr>
<tr>
<td>16</td>
<td>AR(10) nc L15-L60</td>
</tr>
<tr>
<td>17</td>
<td>AR(8) nc L15-L60</td>
</tr>
<tr>
<td>18</td>
<td>MA(10) nc L15-L60</td>
</tr>
<tr>
<td>19</td>
<td>AR(10) MA(10) nc L15-L60</td>
</tr>
<tr>
<td>20</td>
<td>AR(9) nc L15-L60</td>
</tr>
<tr>
<td>21</td>
<td>AR(1-10,13,24,25,27) nc L15-L60</td>
</tr>
</tbody>
</table>

Four of the models sufficiently explained the autocorrelation and finally, based on the information criterion the model number 21 was selected as the most suitable candidate. From now on, the model number 21 will be referred as the candidate model. Additionally to the results in Table 5.4, in order to prove that there is no significant autocorrelation left among the residuals from the estimation of the candidate model, we present the ACF and the PACF plots of the residuals in Figure A.1 in the Appendix A.
5. Model

### Table 5.4: ARMA Results

<table>
<thead>
<tr>
<th>#</th>
<th>$L$</th>
<th>AIC</th>
<th>BIC</th>
<th>Q(1) p-val</th>
<th>Q(10) p-val</th>
<th>Q(20) p-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2722627</td>
<td>-5445208</td>
<td>-5445187</td>
<td>2.41</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>2722488</td>
<td>-5444971</td>
<td>-5444950</td>
<td>4.12</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>2724841</td>
<td>-5445438</td>
<td>-5445406</td>
<td>0.05</td>
<td>0.82</td>
<td>103.68</td>
</tr>
<tr>
<td>4</td>
<td>272676</td>
<td>-5445347</td>
<td>-5445314</td>
<td>0.08</td>
<td>0.78</td>
<td>188.5</td>
</tr>
<tr>
<td>5</td>
<td>272730</td>
<td>-5445449</td>
<td>-5445395</td>
<td>0.05</td>
<td>0.82</td>
<td>88.81</td>
</tr>
<tr>
<td>6</td>
<td>272705</td>
<td>-5445403</td>
<td>-5445371</td>
<td>0.02</td>
<td>0.89</td>
<td>135.89</td>
</tr>
<tr>
<td>7</td>
<td>272718</td>
<td>-5445424</td>
<td>-5445359</td>
<td>0.02</td>
<td>0.89</td>
<td>135.89</td>
</tr>
<tr>
<td>8</td>
<td>272742</td>
<td>-5445472</td>
<td>-5445407</td>
<td>0.01</td>
<td>0.92</td>
<td>63.95</td>
</tr>
<tr>
<td>9</td>
<td>272725</td>
<td>-5445434</td>
<td>-5445347</td>
<td>0.01</td>
<td>0.93</td>
<td>98.23</td>
</tr>
<tr>
<td>10</td>
<td>272747</td>
<td>-5445477</td>
<td>-5445391</td>
<td>0</td>
<td>1</td>
<td>54.34</td>
</tr>
<tr>
<td>11</td>
<td>272752</td>
<td>-5445486</td>
<td>-5445388</td>
<td>0</td>
<td>1</td>
<td>44.48</td>
</tr>
<tr>
<td>12</td>
<td>272750</td>
<td>-5445485</td>
<td>-5445409</td>
<td>0.96</td>
<td>48.81</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>272753</td>
<td>-5445486</td>
<td>-5445378</td>
<td>0.97</td>
<td>43.92</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>272754</td>
<td>-5445484</td>
<td>-5445354</td>
<td>0.97</td>
<td>43.67</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>272732</td>
<td>-5445444</td>
<td>-5445336</td>
<td>0.97</td>
<td>35.99</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>272750</td>
<td>-5445472</td>
<td>-5445321</td>
<td>0</td>
<td>1</td>
<td>54.34</td>
</tr>
<tr>
<td>17</td>
<td>272744</td>
<td>-5445463</td>
<td>-5445333</td>
<td>0</td>
<td>12.59</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>272749</td>
<td>-5445471</td>
<td>-5445320</td>
<td>0.99</td>
<td>12.59</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>272750</td>
<td>-5445452</td>
<td>-5445192</td>
<td>0.95</td>
<td>0.01</td>
<td>16.59</td>
</tr>
<tr>
<td>20</td>
<td>272746</td>
<td>-5445467</td>
<td>-5445326</td>
<td>0.99</td>
<td>7.26</td>
<td>0.01</td>
</tr>
<tr>
<td>21</td>
<td>272778</td>
<td>-5445520</td>
<td>-5445326</td>
<td>0</td>
<td>1</td>
<td>0.98</td>
</tr>
</tbody>
</table>

5.3 Modeling of Volatility

Up to this point, we have assumed homoskedascity in the error terms. That implies

$$\text{Var}(\epsilon_t|X) = \text{Var}(\epsilon_t) = \sigma^2$$

Yet, if the assumption is wrong and the $\{\epsilon_t\}$ is not an independent identically distributed sequence with constant variance, the $\hat{\rho}$ from Equation 5.3 will not be a consistent estimate of $\rho$. Consequently, neither the ACF and the PACF nor the Q-test, described in Subsection 5.1.2, yield a valid results. Obviously, the homoskedasticity assumption is crucial for the whole process of building the ARIMA model. Inconveniently, regarding the financial time series, the homoskedasticity is rather a rare property. And as the reader will see, our case does not differ.

First of all, we plotted the residuals from the estimation of our candidate model. They are depicted in Figure 5.10.

Clearly, the variance of the residuals is far from constant. Furthermore,
the plotted data exhibits switching between high volatility periods and low volatility periods. Such a behavior is called the volatility clustering and it is typical for the Autoregressive Conditional Heteroskedascity (ARCH) described by Engle (1982).

Figure 5.11: Autocorrelation among Squared Residuals

Although the necessary condition for the candidate model was to have serially uncorrelated residuals, it does not imply anything for the squared residuals. We therefore inspected the serial correlation among the squared residuals by
plotting their ACF and PACF and indeed, as can be seen in Figure 5.11, both the plots display significant autocorrelation.

One can also formally test for the ARCH errors using, for instance, the ARCH-LM test described by Engle (1982). The test uses auxiliary regression of squared residuals on their lagged values that can be written as

\[ e_t^2 = \alpha + \sum_{i=1}^{m} \beta_i e_{t-i}^2 + e_t, \quad t = m + 1, \ldots, N, \]

where \( m \) is the number of the lags we test up to and \( N \) is the sample size. The test is then equivalent to the common F-test for joint significance with

\[ H_0 : \beta_1 = \cdots = \beta_m = 0 \Rightarrow \text{no ARCH effects} \]

and the F-statistics is compared with the \( \chi^2_m \).

We run the ARCH-LM test and obtained high test statistics of 14699.61 with corresponding p-value close to zero. Such a low p-value suggests that there is enough evidence to reject the null hypothesis on all conventional significance levels and therefore, our candidate model exhibits ARCH effects in the error terms.

Although we said that violation of the homoskedascity assumption makes the Q-test and the ACF and the PACF results invalid, it does not mean that all the work done in the previous section is useless. We can still use the candidate model for further research even though it might not be the most appropriate specification. The only thing we have to do is to retest the autocorrelation and, if necessary, modify the model while taking the ARCH errors into account.

### 5.3.1 Autoregressive Conditional Heteroskedascity Model

One of the first frameworks for volatility modelling was the ARCH model published by Engle (1982) which allowed the conditional variance to be affected by its past realizations. Generally, the ARCH\((m)\) model could be specified as

\[ \epsilon_t = \sigma_t a_t \]
\[ \sigma_t^2 = \beta_0 + \sum_{i=1}^{m} \beta_i e_{t-i}^2, \]

with \( \{a\} \) being an iid sequence with zero mean and unit variance and often assumed to follow a standard normal distribution or a standardized t-distribution or a generalized error distribution. (Tsay 2005). Consequently, the shocks \( \epsilon \) do
not, then, suffer from autocorrelation and at the same time they depend on their past realizations.

### 5.3.2 Generalized Autoregressive Conditional Heteroskedascity Model

One of the limitations of the ARCH model is that sometimes it may be necessary to include high number of lags in order to sufficiently explain the volatility process which may result in inconveniently complicated model and therefore some alternatives may be demanded.

Bollerslev (1986) introduces a powerful modification of the ARCH model called the Generalized Autoregressive Conditional Heteroskedascity (GARCH) model defined the following way

$$
\epsilon_t = \sigma_t a_t \\
\sigma_t^2 = \beta_0 + \sum_{i=1}^{m} \beta_i \epsilon_{t-i}^2 + \sum_{j=1}^{n} \alpha_j \sigma_{t-j}^2 \\
\beta_0 > 0, \quad \sum_{i=1}^{\max(m,n)} (\beta_i + \alpha_i) < 1,
$$

(5.8)

where $m$ and $n$ are the lag orders, $\{a\}$ is again an iid sequence with zero mean and unit variance. And again, in practice the assumptions about its distribution may vary as in the case of ARCH. The constraint on the $\beta_0$ assures the non-zero unconditional variance of $\epsilon_t$. The constraint on the sum of the coefficients, then, assures that the unconditional variance is finite.

It is no doubt that the ARCH and the GARCH models play an important role in the field of financial modelling and many extensions and modifications of the original models have been introduced, especially in the recent past. Since there is already more than sufficient amount of literature on the topic, we do not provide any further description. For more details see for example Engle (1982), Bollerslev (1986) or Bera & Higgins (1993).

### 5.4 Fitting the ARIMA Model with GARCH Disturbances

Attentive readers may have already noticed the similarity of equations Equation 5.4 and Equation 5.8. Indeed, the GARCH volatility modelling resembles the ARIMA mean modeling with only one difference. The GARCH formula does
not include any error component. The model therefore describes the conditional volatility in deterministic manner, i.e. exactly\(^2\). In contrast, the ARIMA model has an error term and the estimated relationship is then only approximate.

For the process of identification, estimation and diagnostics is also similar, here, we provide only brief description how to fit ARIMA-GARCH model. From Section 5.2 we have our mean equation. From Section 5.3 we have our variance equation. Putting the equations together, we arrive at the ARIMA-GARCH model specified as:

\[
\text{Mean eq.:} \quad (1 - \sum_{i=1}^{p} \delta_i B_i) y_t = \gamma_0 + (1 - \sum_{j=1}^{q} \gamma_j B_j) \epsilon_t \\
\text{Variance eq.:} \quad \epsilon_t = \sigma_t a_t \\
\sigma_t^2 = \beta_0 + \sum_{i=1}^{m} \beta_i \epsilon_{t-i}^2 + \sum_{j=1}^{n} \alpha_j \sigma_{t-j}^2
\]

\[(5.9)\]

\(\beta_0 > 0, \sum_{l=1}^{\max(m,n)} (\beta_l + \alpha_l) < 1\)

5.4.1 Identification

In the phase of identification, we inspect ACF and PACF of the residuals to determine the orders of the AR and MA components of the mean equation. Additionally, we evaluate also ACF and PACF of the squared residuals for the ARCH and the GARCH terms in the variance equation.

5.4.2 Estimation

Again, we estimate the ARIMA model with GARCH errors using the maximum likelihood estimation and we can also choose either the conditional likelihood approach or the exact likelihood approach.

5.4.3 Diagnostics

Once the model is estimated, we run the Q-test to check whether the necessary condition of no autocorrelation among the residuals and also among their squares is not violated. If there is no significant autocorrelation, it means there is no information left in the residuals or in the squares. Finally, we compare all the suitable ARIMA-GARCH specifications using the Akaike and Bayesian information criterion.

\(^2\)Not meaning that estimated volatility equals the real one, but that the estimation is without the error term.
Altogether, we estimated seven ARIMA-GARCH specifications following from the candidate ARIMA model. Four of the models sufficiently explained the autocorrelation among both the residuals and their squared values. The particular specifications and estimations are summarized in Table 5.5 and Table 5.6.

### Table 5.5: GARCH Specifications

<table>
<thead>
<tr>
<th>#</th>
<th>Mean spec.</th>
<th>Variance spec.</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AR(1-10,13,24,25,27) nc L15-L60</td>
<td>GARCH(1,0)</td>
<td>SN</td>
</tr>
<tr>
<td>2</td>
<td>AR(1-10,13,24,25,27) nc L15-L60</td>
<td>GARCH(1,1)</td>
<td>SN</td>
</tr>
<tr>
<td>3</td>
<td>AR(1-10,13,24,25,27) nc L15</td>
<td>GARCH(1,0)</td>
<td>SN</td>
</tr>
<tr>
<td>4</td>
<td>AR(1-10,13,24,25,27) nc L15</td>
<td>GARCH(1,0)</td>
<td>STD</td>
</tr>
<tr>
<td>5</td>
<td>AR(1-10) nc L10,L15</td>
<td>GARCH(1,0)</td>
<td>STD</td>
</tr>
<tr>
<td>6</td>
<td>AR(1-10,20,23) nc L10, L15</td>
<td>GARCH(1,0)</td>
<td>STD</td>
</tr>
<tr>
<td>7</td>
<td>AR(1-10,20,23) nc L10, L15</td>
<td>GARCH(1,0)</td>
<td>SSTD</td>
</tr>
</tbody>
</table>

*SN=Standard Normal, STD=Standardized t-distribution, SSTD=Skewed Standardized t-distribution*

### Table 5.6: GARCH Results

<table>
<thead>
<tr>
<th>#</th>
<th>L</th>
<th>AIC</th>
<th>BIC</th>
<th>SQ</th>
<th>Q(1) p-val</th>
<th>Q(5) p-val</th>
<th>Q(10) p-val</th>
<th>Q(20) p-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>803962</td>
<td>-4.39</td>
<td>-4.39</td>
<td>536.37</td>
<td>0</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>2</td>
<td>2726208</td>
<td>-14.89</td>
<td>-14.89</td>
<td>2499.53</td>
<td>0</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>3</td>
<td>944916</td>
<td>-5.16</td>
<td>-5.16</td>
<td>0.06</td>
<td>0.8</td>
<td>7.51</td>
<td>0.19</td>
<td>8.62</td>
</tr>
<tr>
<td>4</td>
<td>2511999</td>
<td>-13.72</td>
<td>-13.72</td>
<td>0.01</td>
<td>0.94</td>
<td>0.17</td>
<td>0.68</td>
<td>5.88</td>
</tr>
<tr>
<td>5</td>
<td>2411621</td>
<td>-13.17</td>
<td>-13.17</td>
<td>0.09</td>
<td>0.77</td>
<td>0.51</td>
<td>0.47</td>
<td>2.19</td>
</tr>
<tr>
<td>6</td>
<td>2399633</td>
<td>-13.11</td>
<td>-13.11</td>
<td>0.03</td>
<td>0.87</td>
<td>0.12</td>
<td>0.73</td>
<td>0.20</td>
</tr>
<tr>
<td>7</td>
<td>2401692</td>
<td>-13.12</td>
<td>-13.12</td>
<td>0</td>
<td>0.97</td>
<td>2.91</td>
<td>0.09</td>
<td>8.12</td>
</tr>
</tbody>
</table>

As it is depicted in Table 5.5 we have experimented not only with different lag orders but also with the probability distribution of the iid sequence from the variance equation. If we use the information criterion to compare the
model number 3 with the model number 4, clearly, the standardized Student-t
distribution outperforms the standardized normal distribution. Furthermore,
the results of regressions 6 and 7 suggest that the skewed version of the stan-
dardized Student-t distribution seems to be even slightly better.

Finally, as the model number 7 explains the autocorrelation most efficiently
and obtained the highest IC scores, it becomes our new candidate model.
Again, additionally to the Q-test results depicted in Table 5.6, the ACF and
the PACF plots of the residuals and of the squared residuals can be found
in Figure A.2 and Figure A.3 in the Appendix A.

5.5 External Variables

We have successfully identified ARIMA model with GARCH errors that suffi-
ciently describes the log return time series of the EURCHF currency pair. We
have everything set for the final step of our research, i.e. incorporating the
news.

The framework for external variables was already established in Subsec-
tion 5.2.1 and the data was described in Section 4.2. Furthermore, we decided
to include the news in two forms. First, we include the actual announced value
that, after the first differencing, represents the difference between the actual
and the previous value. Second, as we assume the expectation of the general
public to substantially influence the market, we also include the difference be-
tween the consensus value, described in Section 4.3, and the actual value which
can be understood as a forecast error.

We extended the current candidate model by variables of the News and
we arrive at the ARIMAX model with GARCH errors. For clarity, the final
candidate model is specified as described by mean-variance Equation 5.10 and
by Table 5.7 of included regressors below.

\[
\begin{align*}
\text{Mean eq.}: & \quad (1 - \sum_{i=1}^{p} \alpha_i B_i) y_t = \sum_{j=1}^{n} \beta_j x_{jt} + \epsilon_t \\
\text{Variance eq.}: & \quad \epsilon_t = \sigma_t a_t \\
& \quad \sigma_t^2 = \gamma_0 + \gamma_1 \epsilon_{t-1}^2 & (5.10) \\
& \quad \gamma_0 > 0, \gamma_1 < 1
\end{align*}
\]

where the left-hand side of the mean equation represents the autoregressive
part of the model; the right-hand side represents the news variables and the
floating returns terms. The variance equation is rather simple as it contains only one ARCH term.

Table 5.7: Final Model Summary

<table>
<thead>
<tr>
<th>ARIMA</th>
<th>Floating Returns</th>
<th>Mean spec.</th>
<th>Variance spec.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Euro Area</td>
<td>Germany</td>
</tr>
<tr>
<td>Mean spec.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 L10</td>
<td>GDP QoQ</td>
<td>GDP QoQ</td>
<td>1</td>
</tr>
<tr>
<td>2 L15</td>
<td>GDP YoY</td>
<td>GDP YoY</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>CPI MoM</td>
<td>CPI MoM</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>CPI YoY</td>
<td>CPI YoY</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Core CPI YoY</td>
<td>Unemployment</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Interest Rate</td>
<td>Retail Sales MoM</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Unemployment</td>
<td>Retail Sales YoY</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Retail Sales MoM</td>
<td>Industrial Production YoY</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Retail Sales YoY</td>
<td>Manufacturing PMI</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Consumer Confidence</td>
<td>Trade Balance</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>ESI</td>
<td>PPI MoM</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>PPI YoY</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After the model is estimated, we verify whether it fulfills the condition of serially uncorrelated residuals and squared residuals. In order to do so, we run the Q-test on the residuals and their squared values. The results of the Q-test, depicted in Table 5.8, suggest that our model sufficiently explains the autocorrelation among the residuals and among the squared residuals. Furthermore, we also provide the ACFs and the PACFs plots depicted in Figure A.4 and Figure A.5 in Appendix A.

Table 5.8: Final Model Q-Test Results

<table>
<thead>
<tr>
<th>AR</th>
<th>Residuals</th>
<th>Squared Residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$m$</td>
<td>$Q(m)$ p-value</td>
</tr>
<tr>
<td>1</td>
<td>0.004</td>
<td>0.95 0.99</td>
</tr>
<tr>
<td>5</td>
<td>0.0087</td>
<td>0.99 0.99</td>
</tr>
<tr>
<td>10</td>
<td>2.0436</td>
<td>0.84 0.002</td>
</tr>
<tr>
<td>20</td>
<td>5.2357</td>
<td>0.81 0.0073</td>
</tr>
<tr>
<td>40</td>
<td>18.719</td>
<td>0.54 0.0334</td>
</tr>
</tbody>
</table>
5.6 Summary of the ARIMAX GARCH Fitting Procedure

As we reached the end of the procedure of identification, fitting and estimation of the autoregressive moving average model with external variables and conditional heteroskedastic errors, it is convenient to summarize the whole process. The following flowchart can be used as a compact guideline for ARIMAX – GARCH modelling.
Figure 5.12: ARIMAX – GARCH Modelling Guide

Initial TS

is it stationary?

yes

identify mean model (acf & pacf)

estimate, analyze residuals

residuals=new TS

are they autocorr. (Q-test)?

yes

ARCH-LM test

are there ARCH effects?

add external variables, estimate

still no autocorr. or ARCH effects

ok

describe results

identify variance model (acf & pacf of squared residuals)
Chapter 6

Results

Finally, our model is formally valid and therefore, we can discuss the results. First we describe the results of the autoregressive part of the model, then the floating returns and finally, we focus on the results of the news analysis.

6.1 Autoregressive Part

The results, for the AR terms, of the estimation are depicted in Table 6.1. As all the p-values are zero, we reject the null hypothesis of the t-test for individual significance in all cases. Therefore, we can say that all the included AR terms are significant. The significance of the AR terms is not surprising.

<table>
<thead>
<tr>
<th>AR lag</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.454</td>
<td>0.000188</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>-0.187</td>
<td>0.000122</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>-0.077</td>
<td>0.000076</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>-0.023</td>
<td>0.000025</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0.015</td>
<td>0.000022</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0.010</td>
<td>0.000019</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0.012</td>
<td>0.000021</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0.013</td>
<td>0.000024</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>-0.015</td>
<td>0.000022</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>-0.014</td>
<td>0.000018</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>-0.006</td>
<td>0.000017</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>0.005</td>
<td>0.000052</td>
<td>0</td>
</tr>
</tbody>
</table>
as all the insignificant terms were excluded during the primary identification of the ARIMA model.

Furthermore, we observe certain repeating pattern among the AR coefficients. As the four lags have negative coefficient, they steer the price to the opposite direction. The lags from 5 to 8, on the other hand, with the positive coefficients push the price in the same direction. The following lags, then, have a negative coefficients again. Therefore, we can assume that the pattern is repeating. It may seem odd that the most recent lags affect the price negatively, however, it corresponds to the usual behavior of the FOREX market. Usually, the price movement does not follow a straight line. More likely, the values fluctuates in certain range and we observe repeating pullbacks against the main direction of the movement. As depicted in Figure 6.1, the fluctuation occurs regardless the price following an up-trend, down-trend or not any trend at all.

Figure 6.1: Exchange Rate Fluctuation

As for the values of the coefficients, the biggest coefficient, in absolute terms, is at the first lag. Given that the return in t-1 is 100%, the current return is likely to be about 45 percentage points lower. The coefficients of the more distant lags are then decreasing suggesting a decreasing importance of the more distant lags.

6.2 Floating Returns

More interesting than the results of the ordinary AR terms are the results of the Floating Returns terms that we introduced in Subsection 5.2.2. The results are summarized in Table 6.2. Both the terms turn out to be significant and the coefficients are fairly high in absolute terms. The 100% last-10-minute return would increase the current return by 4.4 percentage points. Similarly, the 100% last-15-minute return would decrease the current return by 5.4 percentage
6. Results

The different signs again suggest the fluctuation within the main direction. Altogether, the results prove our assumption that the higher time-frame layers of the market substantially affect the underlying price developments.

Table 6.2: Floating Returns Results

<table>
<thead>
<tr>
<th>FR</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>L10</td>
<td>0,044</td>
<td>0,000027</td>
<td>0</td>
</tr>
<tr>
<td>L15</td>
<td>-0,054</td>
<td>0,000038</td>
<td>0</td>
</tr>
</tbody>
</table>

6.3 News

The results, regarding the News, are summarized Table 6.3 and Table 6.4. The first table depict the variables of the Euro area, the second depict the German ones. First of all, the estimation turned out to be more or less as we expected. There are no suspiciously high or low values. In absolute terms, the biggest effect has the EA Consumer Price Index MoM, when the change of the CPI by 1 percentage point would change the EURCHF return by 7.5 percentage points. The least effect, on the other hand, has the German Manufacturing PMI, which would if changed by one percentage point increase the EURCHF return by 0.18 percentage points.

Table 6.3: Euro Area News Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Previous Value Diff.</th>
<th>Consensus Value Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>SE</td>
</tr>
<tr>
<td>CPI MoM</td>
<td>0,0750</td>
<td>0,001292</td>
</tr>
<tr>
<td>CPI YoY</td>
<td>0,0491</td>
<td>0,000707</td>
</tr>
<tr>
<td>Retail Sales YoY</td>
<td>0,0345</td>
<td>0,000176</td>
</tr>
<tr>
<td>GDP YoY</td>
<td>0,0270</td>
<td>0,000692</td>
</tr>
<tr>
<td>Core CPI YoY</td>
<td>0,0100</td>
<td>0,000959</td>
</tr>
<tr>
<td>Unemployment</td>
<td>0,0076</td>
<td>0,000431</td>
</tr>
<tr>
<td>ESI</td>
<td>0,0033</td>
<td>0,000247</td>
</tr>
<tr>
<td>Consumer Conf.</td>
<td>-0,0218</td>
<td>0,000113</td>
</tr>
<tr>
<td>GDP QoQ</td>
<td>-0,0225</td>
<td>0,000097</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>-0,0280</td>
<td>0,000842</td>
</tr>
<tr>
<td>Retail Sales MoM</td>
<td>-0,0712</td>
<td>0,000042</td>
</tr>
</tbody>
</table>
Table 6.4: German News Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Previous Value Diff.</th>
<th>Consensus Value Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.    SE</td>
<td>p-value</td>
</tr>
<tr>
<td>CPI MoM</td>
<td>0.0571    0.001292</td>
<td>0</td>
</tr>
<tr>
<td>Industrial Prod. YoY</td>
<td>0.0346  0.000527</td>
<td>0</td>
</tr>
<tr>
<td>CPI YoY</td>
<td>0.0324    0.000267</td>
<td>0</td>
</tr>
<tr>
<td>Retail Sales YoY</td>
<td>0.0293    0.000785</td>
<td>0</td>
</tr>
<tr>
<td>PPI YoY</td>
<td>0.0284    0.000241</td>
<td>0</td>
</tr>
<tr>
<td>Trade Balance</td>
<td>0.0214    0.000761</td>
<td>0</td>
</tr>
<tr>
<td>GDP YoY</td>
<td>0.0114    0.000385</td>
<td>0</td>
</tr>
<tr>
<td>Manufacturing PMI</td>
<td>0.0018    0.000317</td>
<td>0</td>
</tr>
<tr>
<td>Unemployment</td>
<td>0.0006    0.000398</td>
<td>0.14</td>
</tr>
<tr>
<td>Retail Sales MoM</td>
<td>-0.0041   0.00035</td>
<td>0</td>
</tr>
<tr>
<td>GDP QoQ</td>
<td>-0.0227   0.000082</td>
<td>0</td>
</tr>
<tr>
<td>PPI MoM</td>
<td>-0.0346   0.000091</td>
<td>0</td>
</tr>
</tbody>
</table>

Considering the significance, based on the p-values, we find all the variables, except two, to be significant. The significance of the German unemployment was rejected with p-value of 0.14. The reason is probably not that the German unemployment is an unimportant variable, but more likely because the German rate of unemployment is rather steady. As the variation in the unemployment is very small during the whole period, its explanatory power is rather low. Similarly, the forecast error of the ECB interest rate is considered to be insignificant with p-value of 0.69. Again, the interest rate is an important variable, however, it does not change very often. Therefore the consensus value is most likely to reflect the real value, so the forecast error is minimal. Hence the insignificance as in the previous case.

As we expected, most of the coefficients was estimated with positive sign, reflecting the assumption that when the value of the indicator decreases, the EURCHF return decreases as well. The decreasing returns are, then, nothing else than an appreciation of the Swiss franc (CHF) against the Euro. A few of the variables, however, has a negative coefficient. But there may exist a rational explanation. Some of the news, for instance the QoQ and the YoY forms of the same indicator or the Consumer Confidence and the ESI, are released at the exactly same time. Consequently, although the joint effect of the variables may be estimated properly, it may be hard to distinguish the individual effects. The ambiguity between the individual effect results in invalid coefficients. Furthermore, the interest rate announcement is an important event which can cause...
a substantial tension in the market. When there is a tension, the investors are likely to make an irrational decisions. Consequently, it is not unlikely that right after the announcement, we observe a temporary short-term swing of the exchange rate. Figure 6.2 depicts a moment when an unchanged interest rate was published, yet the exchange rate exhibits large spike in only two following minutes.

Figure 6.2: Tension Induced Irrationality

There is also a few negative coefficient among the forecast errors, however, we do not have any rational explanation for it. It may be simply caused by some imperfection of the model.

Generally, the Euro area variables seems to have greater influence on the EURCHF exchange rate than the German ones. Moreover, the changes in the actual values are more important than the forecast errors, suggesting that the rational decisions, based on the real outputs of the economy, still outweighs the effect of the missed expectations.
In the first part, we used three different approaches to evaluate qualities of Switzerland and its currency as a safe haven for European investors. According to the major economic indicators, such as GDP, Unemployment or standard of living, and in the context of international trade, Switzerland proves to have potential to be a safe haven. Furthermore, empirical evidence suggests that the investors, indeed, use it that way.

Following from the safe haven status of Switzerland, we built an econometric model to analyze whether the values of the indicators of European economy, published as an economic news, affect the immediate demand for Swiss franc. We employed the ARIMA model with GARCH disturbances and external regressors and fitted it on the one-minute EURCHF exchange rate. Moreover we, extended the ordinary ARIMA model with our custom component: the Floating Returns. The Floating Returns project the influence of the higher-timeframe layers of the FOREX market on the underlying minute-data. After we described the process of fitting the ARIMA-GARCH model, we finally estimated the complete model including the News variables.

We arrive at the conclusion, that news have certain instant effect on the exchange rate and in most cases, the bad news lead to appreciation of the CHF. The results correspond to our hypothesis that the bad news motivate the investors to transfer wealth to the safe haven. Furthermore, we found out that, apart from the general macroeconomic indicators, the investors’ expectations also play some role in the exchange rate determination and if the actual values fall short of the expectations, it may induce appreciation of the CHF as well.

The most important results of our research are that the news have impact on the EURCHF exchange rate. These findings are important since might
be exploited in constructing a FOREX trading strategy, as well as they might be interesting for the policy makers. Moreover, we successfully employed the Floating Returns extension of our own design and it seems to be useful for modeling of the multi-layer market.

Nonetheless our model is not flawless and there are several possibilities for a further research. For instance, we employed simple GARCH model which assumes symmetric shocks. In practice, however, we usually observe asymmetric shocks as the markets reflect the biased perception of the investors. In other words, the negative change is often perceived more than a positive one, therefore model adjusted for asymmetric shocks may deliver better results. Furthermore, we were able to include only limited lag order of the Floating Returns and the news variables were not lagged at all. Although including lags of higher orders may improve our model, it would require considerably powerful computation technology, as the estimation of complex models fitted on high frequency data is extremely demanding.


Appendix A

Selected Figures

Figure A.1: ARIMA Model #21 - No Significant Autocorrelation among the Residuals

Source: author’s computations
Figure A.2: ARIMA-GARCH Model #7 - No Significant Autocorrelation among the Residuals

Source: author’s computations
Figure A.3: ARIMA-GARCH Model #7 - No Significant Autocorrelation among the Squared Residuals

Source: author’s computations
Figure A.4: Final Model - No Significant Autocorrelation among the Residuals

Source: author's computations
Figure A.5: Final Model - No Significant Autocorrelation among the Squared Residuals

Source: author's computations