Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Forecasting the Exchange Rate in the Czech Republic Using Non-linear Threshold Models

Author: **Bc. Petr Žák**

Supervisor: **PhDr. RNDr. Josef Stráský Ph.D.**

Academic Year: **2016/2017**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

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

Prague, January 6, 2017

Signature

Acknowledgments

I would like to express my sincere gratitude to my supervisor Josef Stráský for the idea to write the diploma thesis on the topic of the exchange rate forecasting. I would also like to thank him for his guidance, continuous support, all useful comments and remarks and infinite patience during the completion of this thesis.

I must also express my very profound gratitude to my girlfriend Pavlína Šmucerová for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis.

Abstract

The aim of this thesis is to analyze the performance of nonlinear threshold models in forecasting the exchange rate of Czech koruna against EUR.

Data for this study were obtained from Statistical Data Warehouse of European Central Bank (ECB) website, from Czech National Bank (CNB) Board decisions minutes and from the press releases of Governing Council of ECB. The data set was split into two periods - from 1999 until November, 2013 when CNB started to use interventions and from November, 2013 until April, 2016.

Models used in the thesis are Self-Exciting Threshold Auto Regressive (SETAR) models with one and two thresholds and two Threshold Auto Regressive (TAR) models with different threshold variables - meetings of CNB Board as dummy variable and average volatility over recent periods.

The forecasting results indicate that SETAR models did not outperform Random Walk in any period. TAR models offered promising results in the period before interventions and surprisingly failed in the period during interventions.

This study supports the general belief of exchange rates being difficult to forecast and that it holds in case of Czech koruna as well.

JEL Classification	F12, F21, F23 H25, H71, H87
Keywords	forecasting, exchange rate, time series, nonlinearity, SETAR, TAR

Author's e-mail	zaka.one@gmail.com
Supervisor's e-mail	josef.strasky@gmail.com

Abstrakt

Cílem této diplomové práce je analýza výkonnosti nelineárních prahových modelů při předpovídání směnného kurzu české koruny vůči euru.

Data byla získána z webových stránek Statistical Data Warehouse Evropské centrální banky (ECB), ze zápisů z jednání bankovní rady České národní banky (ČNB) a z tiskových zpráv Rady guvernérů ECB. Data set byl rozdělen na dvě období - od roku 1999 do listopadu 2013, kdy ČNB začala vyžívat devizových intervencí, a od listopadu 2013 do dubna 2016.

Modely, které byly použity v této diplomové práci, jsou SETAR s jednou a dvěma prahovými hodnotami a dva TAR modely s dvěma různými prahovými proměnnými - s mítinky bankovní rady jako dummy proměnnou a s průměrnou volatilitou v nedávných obdobích.

Výsledky předpovědí odhalují, že SETAR modely nepředčí model založený na náhodné procházce ani v jednom období. TAR modely projevily slibné výsledky v období před intervencemi, ale překvapivě selhaly v období v průběhu intervencí.

Tato studie je v souladu s obecným přesvědčením, že směnné kurzy je složité předpovídat, což platí i v případě české koruny.

Klasifikace JEL	F12, F21, F23 H25, H71, H87
Klíčová slova	předpovídání, směnný kurz, časové řady, nelinearita, SETAR, TAR

E-mail autora	`zaka.one@gmail.com`
E-mail vedoucího práce	`josef.strasky@gmail.com`

Contents

List of Tables ix

List of Figures xi

Acronyms xii

Thesis Proposal xiv

1 Introduction 1

2 Literature Overview 3

2.1 Back to the 80's . . . 3

2.2 Exchange Rate Forecasting in the Central and Eastern Europe . 4

2.3 Threshold Models . . . 6

2.4 Threshold Models in the Central and Eastern Europe . . . 7

2.5 Central Bank Communication and Exchange Rates . . . 7

3 Data and Tools 8

3.1 Data . . . 8

3.1.1 Descriptive Statistics . . . 9

3.2 Tools . . . 17

4 Methodology 18

4.1 Methodology Overview . . . 18

4.1.1 Descriptive Statistics . . . 18

4.1.2 Model Identification . . . 19

4.1.3 Forecasting . . . 19

4.1.4 Measuring Accuracy . . . 20

4.1.5 Univariate Time Series . . . 20

4.2 Testing for Nonlinearity . . . 21

4.2.1 Brock Dechert Scheinkman (BDS) Test 21
4.3 Measuring the Accuracy of Forecasts 23
4.3.1 Mean Square Error (MSE) and Root Mean Square Error (RMSE) 23
4.3.2 TheilU Statistics 24
4.3.3 Diebold-Mariano Test 24
4.3.4 Success Ratio 25
4.4 Random Walk and Univariate Models 26
4.4.1 Random Walk 26
4.4.2 Auto Regressive Integrated Moving Average (ARIMA) . . 26
4.4.3 Generalized Auto Regressive Conditional Heteroskedasticity (GARCH) 27
4.4.4 Auto Regressive Moving Average (ARMA)/GARCH 28
4.5 Univariate Threshold Models 29
4.5.1 Threshold Auto Regressive Model 29
4.5.2 Self-Exciting Threshold Auto Regressive Model 30
4.6 Multivariate Time Series 31
4.6.1 Vector Auto Regressive Model 32
4.6.2 Specification of Vector Auto Regressive (VAR) Model and its Estimation 32
4.6.3 Forecasting 33
4.6.4 Threshold Vector Auto Regressive model 33
4.6.5 Specification of Threshold Vector Auto Regressive (TVAR) and its Estimation 33
4.7 Summary of the Models Used in the Thesis 34

5 Results 35
5.1 Univariate Linear Models 36
5.1.1 ARIMA 36
5.1.2 Model Identification 36
5.1.3 BDS Test for Nonlinearity 38
5.1.4 Forecasting 39
5.1.5 ARMA/GARCH 41
5.2 Univariate Threshold Models 42
5.2.1 SETAR Models 42
5.2.2 Model Identification 42
5.2.3 Forecasting 45

5.2.4 TAR Models . . . 47
5.2.5 Model Identification . . . 49
5.2.6 Forecasting . . . 51

6 Discussion **54**
6.1 Data . . . 54
6.2 Random Walk (RW), ARIMA models and BDS test . . . 55
6.3 SETAR model . . . 56
6.4 TAR models . . . 57

7 Conclusion **59**

Bibliography **64**

A Data, source codes and supportive materials **I**

List of Tables

3.1 Descriptive statistics for the whole period, ie. 1999 - 2016 13
3.2 Descriptive statistics for the period before CNB interventions . . 13
3.3 Descriptive statistics for the period during recent CNB interventions, ie. from Nov 7th, 2013 till Apr 18, 2016 13
3.4 Descriptive statistics for weeks when CNB had a Board meeting 14
3.5 Descriptive statistics for weeks when CNB changed the repo rate 14

5.1 ARMA(1,1) and RW forecasting performance, rolling samples, before interventions 40
5.2 ARMA(1,1) and RW forecasting performance, recursive samples, before interventions 40
5.3 ARMA(1,3) and RW forecasting performance, rolling samples, during interventions 41
5.4 ARMA(1,3) and RW forecasting performance, recursive samples, during interventions 41
5.5 SETAR(3,1) and RW forecasting performance, rolling samples, before interventions 46
5.6 SETAR(3,1), SETAR(4,3,4) and RW forecasting performance, recursive samples, before interventions 46
5.7 SETAR(1,2), SETAR(1,3,1) and RW forecasting performance, rolling samples, during interventions 47
5.8 SETAR(1,2), SETAR(1,3,1) and RW forecasting performance, recursive samples, during interventions 47
5.9 TAR(1,1) with CNB meetings dummy as threshold variable and RW forecasting performance, rolling samples, before interventions 51
5.10 TAR(1,1) with CNB meetings dummy as threshold variable and RW forecasting performance, recursive samples, before interventions 52

5.11 TAR(3,4) with volatility over the last two weeks threshold variable and RW forecasting performance, rolling samples, before interventions . . . 52
5.12 TAR(3,4) with volatility over the last two weeks threshold variable and RW forecasting performance, recursive samples, before interventions . . . 52
5.13 TAR(2,1) with CNB meetings dummy as threshold variable and RW forecasting performance, rolling samples, during interventions 53
5.14 TAR(2,1) with CNB meetings dummy as threshold variable and RW forecasting performance, recursive samples, during interventions . . . 53
5.15 TAR(2,4) with volatility over the last two weeks threshold variable and RW forecasting performance, rolling samples, during interventions . . . 53
5.16 TAR(2,4) with volatility over the last two weeks threshold variable and RW forecasting performance, recursive samples, during interventions . . . 53

List of Figures

3.1 EUR/CZK weekly closing prices 9
3.2 Weekly returns 9
3.3 ACF - data 10
3.4 PACF - data 10
3.5 Leptokurtosis of EUR/CZK log returns 11
3.6 Histograms of Log Exchange Rate Changes 16

5.1 ARMA(1,1) residuals ACF 37
5.2 ARMA(1,1) fit residuals 37
5.3 ARMA(1,1) squared residuals ACF 37
5.4 ACF of ARMA(1,3) residuals 38
5.5 ARMA(1,3) fit residuals 38
5.6 ARMA(1,3) squared residuals ACF 38
5.7 ACF of Squared Standardized Residuals 42

Acronyms

ACF Auto Correlation Function

ADF Augmented Dickey-Fuller

AIC Akaike Information Criterion

AICc Akaike Information Criterion

AR Auto Regressive

ARCH Auto Regressive Conditional Heteroskedasticity

ARIMA Auto Regressive Integrated Moving Average

ARMA Auto Regressive Moving Average

BDS Brock Dechert Scheinkman

BIC Schwarz-Bayesian Information Criterion

BVAR Bayesian Vector Auto Regressive

BVEC Bayesian Vector Error Correction

CEE Central and Eastern European

CNB Czech National Bank

CZK Czech koruna

DM Diebold-Mariano

ECB European Central Bank

EU European Union

EUR Euro

GARCH Generalized Auto Regressive Conditional Heteroskedasticity

HQ Hannan-Quinn

MA Moving Average

MAD Mean Absolute Deviation

MAE Mean Absolute Error

MAPE Mean Absolute Percentage Error

ME Mean Error

MEDSE Median Squared Error

MSE Mean Square Error

MSFE Mean Square Forecasting Error

MSPE Mean Square Prediction Error

OLS Ordinary Least Squares

PACF Partial Auto Correlation Function

RMSE Root Mean Square Error

RVAR Restricted Vector Auto Regressive

RW Random Walk

SETAR Self-Exciting Threshold Auto Regressive

SSR Sum of Squared Residuals

SUR Seemingly Unrelated Regression

TAR Threshold Auto Regressive

TVAR Threshold Vector Auto Regressive

VAR Vector Auto Regressive

VEC Vector Error Correction

Master's Thesis Proposal

Author	Bc. Petr Žák
Supervisor	PhDr. RNDr. Josef Stráský Ph.D.
Proposed topic	Forecasting the Exchange Rate in the Czech Republic Using Non-linear Threshold Models

Topic characteristics The reasons to have reliable exchange rate forecasting tools are fundamental:

1. From the corporate point of view, as stated by Taušer & Buryan (2011), more and more companies are entering global markets nowadays, therefore the demand for applicable and efficient predictive tools that could give groundwork for international financial management decisions is growing.

2. From the government point of view, the Czech Republic is still on the way to accept Euro currency. Therefore it is, or it will be important to focus on one of the Maastricht criteria – exchange rate stability. The necessity of having a reliable forecasting tool is then obvious.

In 1983, Meese & Rogoff (1983) surprised academics by showing that monetary models cannot outperform a random walk in out-of-sample exchange rate forecasting. Since that time, many authors have contributed to this field with more or less controversial results.

The question whether the Czech koruna (CZK)/Euro (EUR) exchange rate is predictable has been addressed in a few studies, for example Cuaresma & Hlouskova (2004), Naszódi (2011) using a promising survey forecasts method or Mućk & Skrzypczyński (2012) using time series models.

Current research in this area could be more or less split into two groups:

1. Fundamental Approach (out of the scope of this thesis)

2. Technical Analysis

With regards to technical analysis - the common approach assumes linear relationships among variables (Ouliaris, S. (2012)), but for instance Franses & vn Dijk (2000) observed that exchange rate might be constrained to lie within a pre-defined target zone. Threshold models can accommodate this particular behavior.

My thesis is going to contribute to this topic by extending the current approach with nonlinear threshold models, which have not yet been covered sufficiently in the research on the Czech koruna exchange rate.

Hypotheses

1. Threshold models outperform random walk in forecasting exchange rates in the case of the Czech koruna.

2. There are non-linearities in exchange rate time series with regards to the Czech koruna.

3. Threshold models are efficient and accessible tools for exchange rate forecasting for practitioners.

Methodology For the purpose of my thesis I will use suitable threshold model to forecast exchange rates. The forecasting periods will be 3, 6 and 12 months. For comparison of the results and predictive ability I will use the Mean Square Error tests and more advanced tests evaluating model efficiency such as Theil statistics or Diebold-Mariano tests.

The Brock–Dechert–Scheinkman test will be used to find out whether there are non-linearities in the data.

All of the data that will be used in the thesis are provided by the CNB and are available online. The research will take into account the interventions on the exchange rate market made by CNB in November 2013 to get rid of possible biases.

Expected contribution Since the literature regarding exchange rate forecasting in CEE is rare (Mućk & Skrzypczyński 2012) and most of the research applied to Czech koruna encompasses linear models (VAR, RVAR, BVAR, etc.), which fail to prove their forecasting accuracy (Cuaresma & Hlouskova 2004) or do not outperform random walk (Mućk & Skrzypczyński 2012), there is a space or even necessity to shift to nonlinear models.

Non-linear models were covered by Cuaresma *et al.* (2005), but only for the period of 1999 – 2004, ie. before the Czech Republic joined the European Union. Their results confirmed presence of strong non-linearities, so there is certain evidence of non-linear models accuracy.

My thesis will introduce threshold models for exchange rate forecasting with regards to the Czech koruna. The contribution covers a methodology, which has rarely been used. Also the research period includes several milestones that have not been covered yet – joining the EU in 2004, financial crisis in 2008 and CNB interventions in 2013.

Critical assessment of the efficiency of threshold models will bring new insight to exchange rate forecasting, which might be valuable for financial agents such as the CNB, commercial banks, Forex traders or globally operating companies. Addressing the more general question about efficiency of linear and non-linear models shall result in theoretical knowledge about the behavior of CZK/EUR exchange rate, for instance.

Outline

1. Introduction
2. Literature review on exchange rate forecasting and non-linear threshold modelling
3. Data and Methodology
4. Results
5. Discussion, alternative approaches, suggestions for future research
6. Conclusion

Core bibliography

1. Bjorvatn, K. & C. Eckel (2006): "Policy Competition for Foreign Direct Investment Between Asymmetric Countries." *European Economic Review* **50(7)**: pp. 1891–1907.
2. Blomstrom, M. & A. Kokko (2003): "The Economics of Foreign Direct Investment Incentives." *NBER Working Papers 9489*, National Bureau of Economic Research, Inc.
3. Görg, H. & E. Strobl (2001): "Multinational Companies and Productivity Spillovers: A Meta-analysis." *Economic Journal* **111(475)**: pp. F723–39.
4. Haaparanta, P. (1996): "Competition for Foreign Direct Investment." *Journal of Public Economics* **63(1)**: pp. 141–53.

5. Haufler, A. & I. Wooton (2006): "The Effects of Regional Tax and Subsidy Coordination on Foreign Direct Investment." *European Economic Review* **50(2)**: pp. 285–305.

6. Stanley, T. D. (2001): "Wheat from Chaff: Meta-analysis as Quantitative Literature Review." *Journal of Economic Perspectives* **15(3)**: pp. 131–150.

7. Wells, L. T., N. Allen, J. Morisset, & N. Pirnia (2001): *"Using Tax Incentives to Compete for Foreign Investment: Are They Worth the Cost?"* Washington, DC: FIAS.

Author

Supervisor

Chapter 1

Introduction

According to Bank for International Settlements (Monetary and Economic Department 2013) the foreign exchange market reached \$5.3 trillion per day in its volume in April 2013.[1] The predictability of exchange rates is of interest for many market participants - investors, policymakers, exporters and importers and finally retail consumers as well. With the growth of the foreign exchange market over the past years and together with the 2013 CNB interventions and its recent announcements it provides a reasonable ground to have a reliable exchange rate forecasting tool.

The particular research interest of this thesis lies in the use of nonlinear threshold methods and their fitting and predicting capabilities in the exchange rate time series analysis. As we showed in the Chapter 2 such approach stressed on nonlinearity has not been sufficiently covered yet in the case of the Czech Republic nor in the case of Central and Eastern European (CEE) countries. In this paper we aim to bridge this gap.

We examined the data from January 1st, 1999 to April 18, 2016 using several univariate linear and nonlinear models. Except for Arima model - a representative of linear models family - we especially aimed to use SETAR with one and two thresholds and separately TAR models with two different threshold variables - meetings of CNB as a dummy variable and average volatility over recent weeks as the other one. The reason why we used these models is the anticipated nonlinearity discovered in our time series when we applied the BDS nonlinearity test.

As a comparison criteria we used several statistics common in the exchange rate forecasting, namely RMSE and TheilU statistics, as well as Diebold-

[1] \$4.0 trillion in April 2010, \$3.3 trilion in April 2007.

Mariano (DM) test. We provided a Success Ratio metric to extend previous measurements. It represents the percentage of successfully forecasted appreciations and depreciations of the CZK relative to EUR. However it does not handle how big the appreciation or depreciation are.

Linear and nonlinear approaches were compared to a common benchmark - random walk - to show advantages of more complex nonlinear models. However these were not confirmed and our analysis provides mixed results.

The organization of the thesis is as follows. Chapter 2 provides the overview of literature in relation to the topic of this thesis. Chapter 3 provides a descriptive statistics of the analyzed data and offers tools and R functions packages for all the computations. Chapter 4 presents methodology of our research together with theory overview. Results including graphics and tables are commented in Chapter 5. Last two chapters, discuss the results, suggest ideas for future research and conclude.

Chapter 2

Literature Overview

Since Meese & Rogoff (1983) seminal paper a lot of literature focusing on exchange rate forecasting in developed countries has been published with contradictory results (Ardic *et al.* 2008). In this overview we will focus mainly on the literature examining the CEE region and the use of threshold models in exchange rate forecasting. Additionaly we will mention two studies examining the effect of CNB communication as one of the determinants of exchange rate behavior. We will utilize these findings later on when dealing with suitable TAR model selection. Because of its importance we also provide a brief review of Meese & Rogoff (1983).

2.1 Back to the 80's

Meese & Rogoff (1983) provided a crucial paper which influenced the research in the field of exchange rate forecasting for the upcoming decades. To cover both fundamental and technical analysis they divided their research into two parts. In the first part they used three structural models (Flexible-price monetary model, Sticky-price monetary model and Hooper-Morton model), in the second part they used a variety of univariate time series models as well as an unconstrained vector autoregression. The purpose of their study was to find out whether these approaches can beat RW in exchange rate forecasting accuracy. To measure the accuracy they used three different statistics, namely Mean Error (ME), Mean Absolute Error (MAE) and RMSE as a principal criterion.

The results were surprising. None of the methods performed significantly better than RW without drift at one to twelve months horizons for any of the analyzed bilateral exchange rates (USD/GBP, USD/DM and USD/JPY).

They suggested possible reasons for such poor results, for instance simultaneous equation bias, sampling error, misspecification, etc. Among others one mainly serves as a motivation for this thesis:

"... we make no attempt to account for possible non-linearities in the underlying models."

Obviously, the work of Meese & Rogoff (1983) could not address all possible models and approaches. Nevertheless, it still even nowadays remains a benchmark for other exchange rate forecasting studies (Evans & Lyons 2005).

2.2 Exchange Rate Forecasting in the Central and Eastern Europe

Despite of the fact that Poland, Hungary and the Czech Republic are members of the European Union, they have not accepted the euro currency yet. They still keep their own currency as a monetary tool and as we saw recently in case of the Czech Republic to intervene on foreign exchange markets. It is therefore surprising that the literature related to exchange rate forecasting in CEE is rather scarce (Mućk & Skrzypczyński 2012). With regards to existing literature it is a common approach to use bilateral exchange rate against the Euro since it is simply the most important currency for this region.

In Cuaresma & Hlouskova (2004) and Cuaresma & Hlouskova (2005) authors examined several linear multivariate models including vector autoregressive models with their restricted and Bayesian versions and vector error correction and its Bayesian version. They aimed at five CEE region exchange rates against the Euro and US dollar. Their results support the conclusion made by Meese & Rogoff (1983), ie. that none of these models outperform naive RW for short term predictions. However, more sophisticated models show better forecasting accuracy compared to RW in case of long-term predictions (more than 6 months).

Mućk & Skrzypczyński (2012) research expands the data set up to 2012 applying the recursive samples method. It covers the period after Poland, Hungary and the Czech Republic (countries of their interest) joined the European Union in 2004. Mućk & Skrzypczyński (2012) employed fractionally integrated RW and few VAR models to eventually conclude that none of them outperforms naive RW as well.

Similar to the case of developed countries also the outcomes of studies related to emerging CEE countries are somehow contradictory. Ardic *et al.* (2008) shows that univariate as well as multivariate series accomplish better forecasting results than RW based on Mean Square Prediction Error (MSPE) criteria. This is not the case for structural models based on the same criteria as they perform better only in case of one-step-ahead forecast. However employing the hypothesis testing proposed by Diebold & Mariano (1995), and West (1996) (further referred to as DMW) does not provide these (unexpectedly) successful (in terms of previous research) results. As they claim: *"It is not possible to reject the null hypothesis that the MSPE from the RW and the MSPE from VAR do not differ at any conventional significance level."*. On the other hand they also declare DMW to be downward biased in favor of RW. Following the method of Clark & West (2006) and making this bias fixed, the results turned to be more favorable.

A different approach is used by Naszódi (2011). Aside from above mentioned studies the survey-based forecasts and their efficiency are compared to RW. The study confirms some of the former conclusions, ie. at some horizons there is a method which outperforms RW. In this case survey-based forecasts were shown to be remarkably better at horizons longer than 5, 6 or 7 months depending on the currency inspected.

Another study which evaluates the survey forecasts is provided by Baghestani & Danila (2014). It gives partly contradictory results to the previous one. At first they divided the analysts' to two groups - domestic and foreign. Then they examined the accuracy of their one-month-ahead and twelve-month-ahead forecasts and found them to be better than RW except for one-month-ahead domestic analysts' forecast. One more interesting outcome is that "unlike the foreign analysts' forecasts, the domestic analysts' forecasts are efficient."

Summary The research of exchange rate forecasting in CEE offers methods able to outperform RW, however all of them except Naszódi (2011) and Baghestani & Danila (2014) are based on various linear models and their performance is just slightly better than naive forecast. Also most of the research except Mućk & Skrzypczyński (2012) and Baghestani & Danila (2014) covers the period before financial crisis. Non-linear threshold methods have not been tested yet.

2.3 Threshold Models

Threshold autoregressive models were firstly introduced by Tong & Lim (1980). They have had numerous applications in many research areas. Comprehensive overview of their use related to economics and econometrics can be found in Hansen (2011).

With regard to exchange rate forecasting one of the first studies was published by Kräger & Kugler (1993). They investigated five currencies against US dollar over the period of June 1980 to January 1990 using weekly end-of-period data. The use of SETAR revealed strong threshold effects, literally *"Three regions with different autoregressive dependence were detected."*. The comparison of SETAR and GARCH models shows GARCH model to be worse since it led to completely mis-specified model in certain periods in contrast with SETAR model. Nevertheless both models were not stable over time. The BDS test was run on the residuals of both models showing that most but not all the non-linearities can be explained by either of above mentioned models.

Based on the previous study and same data, Clements & Smith (2001) confirms some predictability of SETAR models. However it depends on how the predictability itself is assessed. Working with evaluation of density forecasts and non-linear impulse response function analysis they unfold non-linear model, if judged by conventional criteria (Mean Square Forecasting Error (MSFE) for example), may reveal SETAR to be not significantly better than RW. On the other hand "non-linearities were apparent in the dependence of the shapes of the estimated densities on the regime".

Chappel *et al.* (1996) analyzed the behavior of French franc and deutschmark exchange rate on a daily data from May 1st, 1990 to March 30th, 1992. They employed two different SETAR models - with a single and two thresholds, "best" linear Auto Regressive (AR) model and RW to predict the exchange rate for one, two, three, five and ten days ahead. As an accuracy criteria MSE and Median Squared Error (MEDSE) were used. The results were remarkable. For both criteria (particularly for MSE) the SETAR model happened to have significantly better accuracy than RW, especially on five- and ten-steps-ahead forecasts.

2.4 Threshold Models in the Central and Eastern Europe

To our best knowledge there has been only one study which covered non-linear threshold methods and their use in the exchange rate modeling in CEE countries. Cuaresma *et al.* (2005) research was based on SETAR-GARCH model, more specifically (similar to Kräger & Kugler 1993) they proposed three-regime SETAR model with errors having GARCH(1,1) structure. Regarding the Czech Republic they provided an evidence of non-linear behavior on data covering the period of January 1st, 1999 to April 28th, 2004. However they did not use their resulting model to forecast the future exchange rates.

2.5 Central Bank Communication and Exchange Rates

The fact that CNB communication can have an effect on exchange rate behavior was analyzed in Fišer & Horvath (2010). Authors used extended GARCH(1,1) model to examine among others the effects of CNB communication, including Board members comments, meetings minutes and their timing on the exchange rate volatility. Throughout the period from January, 2005 to February, 2007 they discovered that central bank communication tends to decrease exchange rate volatility and that the timing seems to matter as well.

Alternative to the previous study is provided inHába (2016) where authors state that the verbal communication and its potential influence on exchange rates is limited and ambiguous. As they point out market participants sometimes reacts to the CNB Governor comments in an opposite way that would be expected based on the CNB Governor statements.

The contradiction of these studies is caused by the periods when they were performed. The first one analyses the period long before interventions, while the other was performed from 2014 to 2015 period, ie. during interventions.

Chapter 3

Data and Tools

3.1 Data

The exchange rates used in this analysis are the CZK relative to EUR. They were obtained from the Statistical Data Warehouse of ECB website.[1] Based on Mućk & Skrzypczyński (2012) we used weekly closing data for the purpose of the thesis. A total of 903 observations were gathered from January 1, 1999 (euro currency came into existence) to April 18, 2016. In this period Czech koruna had been under managed floating regime. The time span includes several important periods - period before Czech Republic joined the European Union (EU) in 2004, the financial crisis in 2008 and the interventions of CNB in November, 2013, to list a few.

The dataset also includes:

- Logarithmic exchange rates,
- logarithmic exchange rate changes,
- CNB Board decision dates (dummy),
- repo rates,
- repo rates changes (dummy),
- introduction of interventions (dummy),
- Governing Council of the ECB monetary meetings (dummy),

[1] http://sdw.ecb.europa.eu/

- interest rate on the main refinancing operations,
- interest rate on the main refinancing operations changes (dummy),
- asset purchase program introduction (dummy).

Since we did not have access to data provided by any traditional news agency we had to get them in a different way. To get the data regarding CNB Board decision dates and Governing Council of the ECB monetary meetings we manually went through the published CNB Board decision minutes[2] and through the press releases regarding monetary policy decisions of ECB.[3] During this data mining we also extracted the repo rates and interest rates on the main refinancing operations and afterwards processed a simple computations to get the corresponding dummy variables included in our data set.

3.1.1 Descriptive Statistics

First look at EUR/CZK exchange rate (Figure 3.1) shows obvious non-stationarity. The assumption is justified by Augmented Dickey-Fuller (ADF) test which yields the p-value = 0.7563 and therefore the null hypothesis of a unit root cannot be rejected.

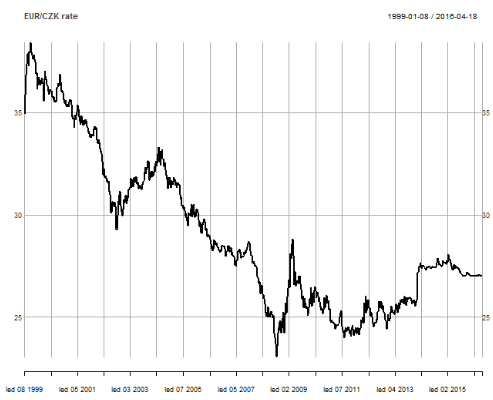

Figure 3.1: EUR/CZK weekly closing prices

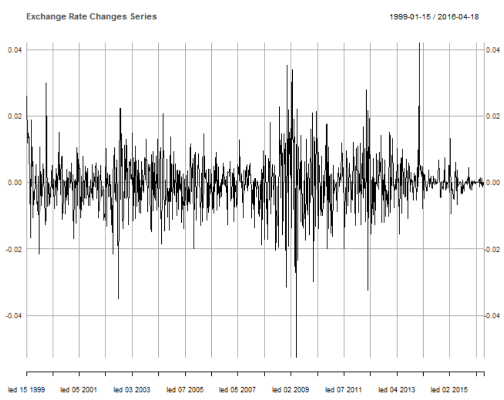

Figure 3.2: Weekly returns

Ljung-Box test up to 10 lags was performed to test the randomness of data. The null hypothesis was rejected with the p-value < 2.2e-16 for all lags, ie. data exhibit serial correlation. This is visually confirmed by the Auto Correlation Function (ACF) (Figure 3.3).

[2]`https://www.cnb.cz/en/monetary_policy/bank_board_minutes/`
[3]`https://www.ecb.europa.eu/press/pr/date/2016/html/index.en.html`

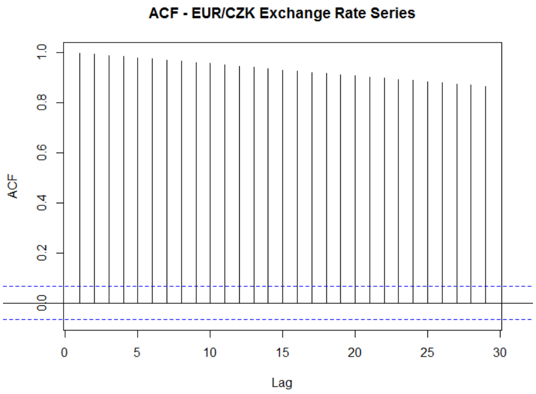

Figure 3.3: ACF - data

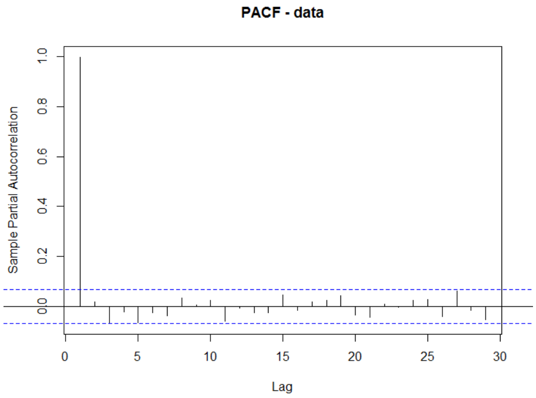

Figure 3.4: PACF - data

The weekly logarithmic exchange rate changes[4][5] exhibit volatility clustering (Figure 3.2). What does not surprise us is that the volatility was higher during the crisis in 2008. Similar to stock returns the logarithmic exchange rate changes series is also leptokurtic (Figure 3.5). The Jarque-Bera test gives us a p-value < 2.2e-16, ie. we strongly reject the null hypothesis of normality. The ADF test null hypothesis of a unit root is strongly rejected with the p-value < 0.01 for the time series of logarithmic returns.

We applied the Ljung-Box test also to diagnose potential serial correlations in the exchange rate changes and it failed to reject the null hypothesis of independence for all lags up to 10.

The following dates refer to local minima and maxima and the highest weekly logarithmic exchange rate change of the series and might be of interest for further investigation:

- February 4th, 1999
- May 7th, 2002
- January 30th, 2004
- June 18, 2008
- February 20, 2009
- November 11, 2013 - On November 7th, 2013 CNB decided to start using

[4]defined as $r_t = ln(P_t) - ln(P_{t-1})$, where P_t denotes price of 1 EUR in CZK at time t

[5]Without loss of generality we sometimes use a term "logarithmic returns" or simply "returns" for logarithmic exchange rate changes.

the exchange rate as an additional instrument for easing the monetary conditions.[6]

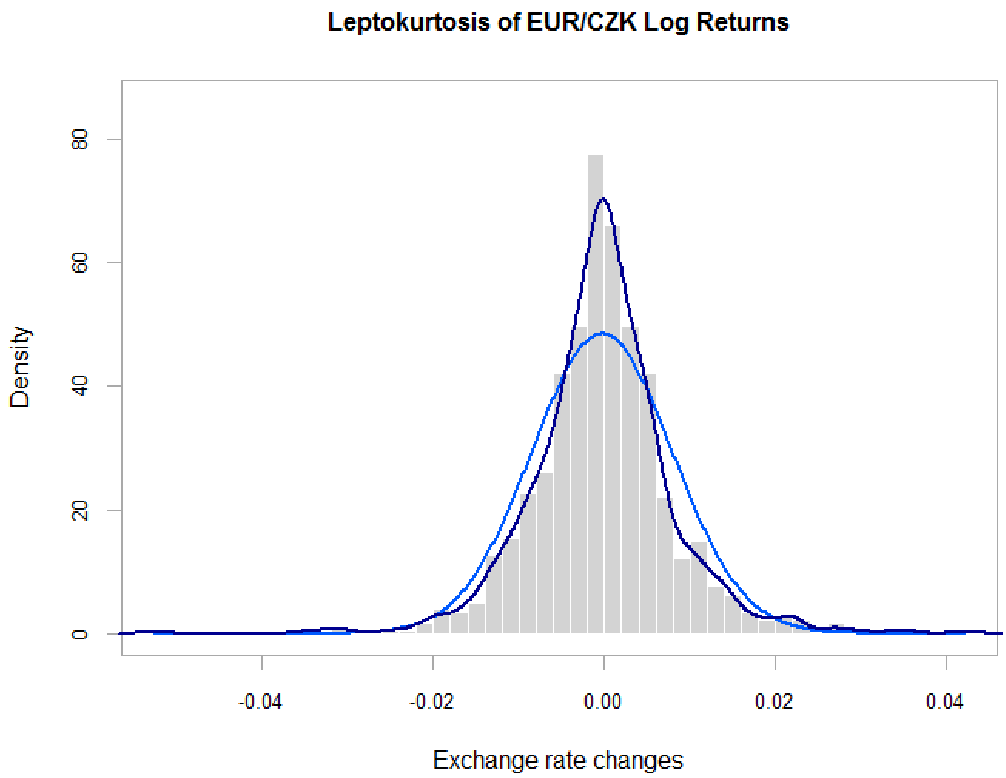

Figure 3.5: Leptokurtosis of EUR/CZK log returns

In our research we took a closer look at periods before and during the recent CNB interventions, as well as what happened when CNB had a Board meeting or even changed the repo rates. Therefore the descriptive statistics of both nominal exchange rates and logarithmic exchange rate changes is presented in the following tables:

- Table 3.1 - whole period statistics, ie. Jan 1st, 1999 to April 18th, 2016,
- table 3.2 - statistics of the period before CNB started to interventions, ie. Jan 1st, 1999 to November 7th, 2013,
- table 3.3 - statistics of the period during the recent interventions, ie. from November 7th, 2013 to April 18th, 2016,

[6] https://www.cnb.cz/en/monetary_policy/bank_board_minutes/2013/amom_131107.html

- table 3.4 - statistics of the weeks, when CNB had a Board meeting,
- table 3.5 - statistics of the weeks, when CNB changed the repo rate.

Table 3.1: Descriptive statistics for the whole period, ie. 1999 - 2016

Statistics	Exchange Rate	Logarithmic Exchange Rate Changes
Minimum	23.063	-0.05261192
Maximum	38.442	0.04211149
Mean	29.0276	-0.0002848378
Median	27.868	-0.0002265842
Std. Dev	3.761439	0.008226055
Variance	14.14842	6.766798e-05
Skewness	0.6843671	-0.02055613
Kurtosis	-0.6214113	4.25245

Table 3.2: Descriptive statistics for the period before CNB interventions

Statistics	Exchange Rate	Logarithmic Exchange Rate Changes
Minimum	23.063	-0.05261192
Maximum	38.442	0.03537033
Mean	29.30613	-0.0003895339
Median	28.478	-0.0003942311
Std. Dev	3.994335	0.008675058
Variance	15.95471	7.525663e-05
Skewness	0.4703632	-0.1412599
Kurtosis	-0.9833071	3.082396

Table 3.3: Descriptive statistics for the period during recent CNB interventions, ie. from Nov 7th, 2013 till Apr 18, 2016

Statistics	Exchange Rate	Logarithmic Exchange Rate Changes
Minimum	26.966	-0.0095602
Maximum	28.062	0.04211149
Mean	27.35642	0.0003425273
Median	27.423	-0.0001824917
Std. Dev	0.25861	0.004686656
Variance	0.06687912	2.196475e-05
Skewness	0.1485484	5.62053
Kurtosis	-0.9010903	47.63983

Table 3.4: Descriptive statistics for weeks when CNB had a Board meeting

Statistics	Exchange Rate	Logarithmic Exchange Rate Changes
Minimum	24.002	-0.03491811
Maximum	38.036	0.04211148
Mean	29.72062	0.001115133
Median	28.595	0.000259005
Std. Dev	3.862781	0.009571001
Variance	14.92108	9.160406e-05
Skewness	0.4543089	0.6002922
Kurtosis	-0.8932314	3.645361

Table 3.5: Descriptive statistics for weeks when CNB changed the repo rate

Statistics	Exchange Rate	Logarithmic Exchange Rate Changes
Minimum	24.211	-0.01878592
Maximum	38.036	0.03537033
Mean	30.86739	0.004740876
Median	30.5855	0.002484586
Std. Dev	4.258211	0.01104757
Variance	18.1323	0.0001220487
Skewness	0.221263	0.7282064
Kurtosis	-1.203615	0.4256797

We see some interesting facts in the data (see also (Figure 3.6)):

- Mean and variance of logarithmic exchange rate changes in the weeks when CNB changed the repo rate or had a Board meeting is much higher than in other weeks; we will use this observation when dealing with TAR models later on,
- the kurtosis of logarithmic exchange rate changes during the recent interventions period is extremely high compared to the period before interventions; it supports the idea to examine the periods before and during the recent interventions separately,
- logarithmic exchange rate changes before interventions are left skewed, data for the period during interventions shows the opposite.

The fact that mean and variance is much higher in the weeks when CNB presented a change in repo rates or had a Board meeting could be caused by at least two factors (Hába 2016): the decision rate of surprise (in other words market expectations and perception of the move) and the following CNB Bank Board or CNB Governor commentary. Hába (2016) summarized the verbal influence as very limited. Moreover they claimed the influence to be only short term and only in case of CNB Governor statements. Nevertheless they are mostly not statistically significant either for exchange rate moves. This conclusion leaves us the space to later discuss the expectations and their role in exchange rate forecasting.

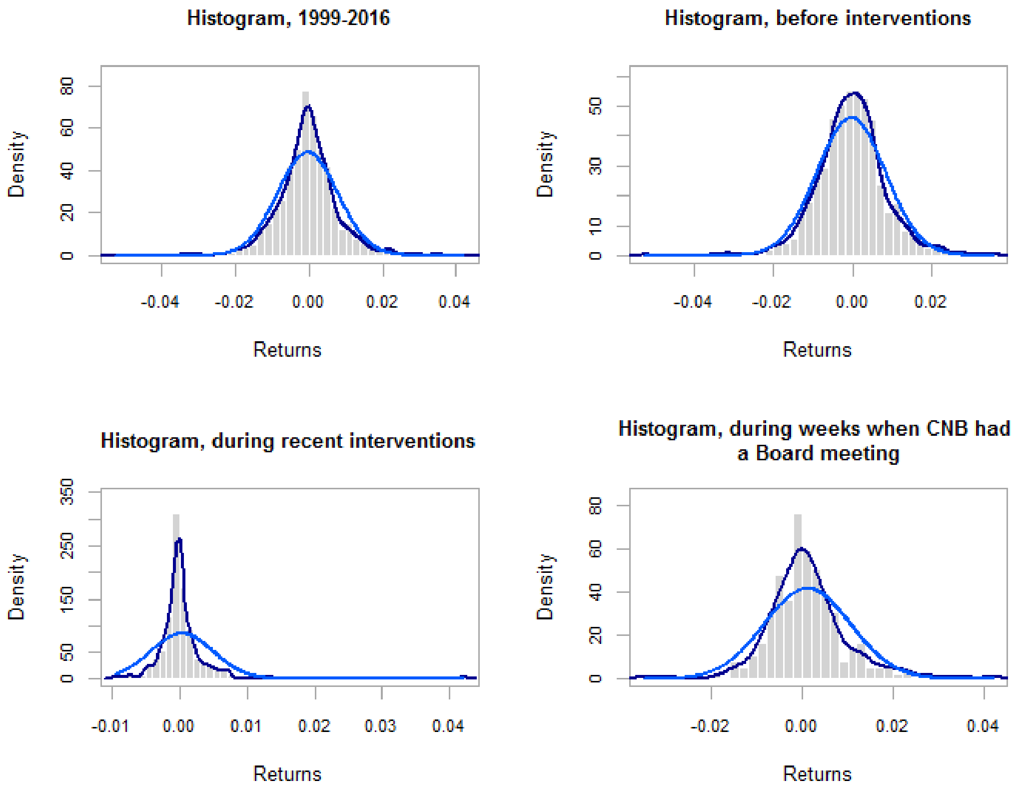

Figure 3.6: Histograms of Log Exchange Rate Changes

3.2 Tools

MS Excel was used to prepare the dataset. R programming language[7] and RStudio IDE[8] was used for all the computations in this thesis. Packages used to analyze the time series and for hypothesis testing were as follows:[9]

- *stats* - The R Stats Package,
- *tseries* - Time Series Analysis and Computational Finance,
- *xts* - eXtensible Time Series,
- *FinTS* - Tools for financial time series analysis,
- *DescTools* - Tools for Descriptive Statistics.
- *rugarch* - Univariate GARCH Models,
- *tsDyn* - Implements nonlinear time series models with regime switching,
- *fNonlinear* - Nonlinear and Chaotic Time Series Modeling,
- *forecast* - Forecasting Functions for Time Series and Linear Models,
- *hydroGOF* - Goodness-of-fit functions for comparison of simulated and observed time series

[7] `http://www.r-project.org/`
[8] `http://www.rstudio.com/`
[9] Source `http://cran.r-project.org/web/packages/available_packages_by_name.html`, more detailed information about packages and their use available ibidem

Chapter 4

Methodology

4.1 Methodology Overview

We have split the research into two parts. In the first one the univariate series of CZK/EUR before recent interventions was analysed, in the second one we switched to the period during recent interventions.

The organization of this chapter is as follows. In the first part we described our approach in general including related R functions we used. The detailed description of particular models, forecasting procedures and performance measurements is provided in following paragraphs.

4.1.1 Descriptive Statistics

To test non-stationarity and randomness of data we used ADF test and Ljung-Box test respectively. The first one is available as `adf.test` in *t-series* package and the other as `Box.test` with type parameter equal to "Ljung-Box" in *stats* package. Number of lags in Ljung-Box test was chosen to be up to 10 as suggested in Hyndman & Athanasopoulos (2013) for non-seasonal time series. From now on whenever we refer to Ljung-Box test we mean Ljung-Box test up to 10 lags. Jarque-Bera test for normality is available as `jarqueberaTest` in *fBasics* package. ACF and Partial Auto Correlation Function (PACF) are part of the *stats* package and we can find them under `acf` and `pacf` functions. The common statistics as mean, variance, etc. are all available in *base* package.

4.1.2 Model Identification

For model identification we applied the Information criteria procedure as follows (In our case we used the Akaike Information Criterion (AIC).):

1. Establish the stationarity of the series, identify seasonality if necessary.

2. Identification - Calculate AIC values for all models taken into account. Choose the model with the minimum AIC value.

3. Estimation - Estimate the parameters by conditional least squares or maximum likelihood using statistical software, check the significance of the parameters, eventually omit insignificant ones and refine the model.

4. Validation - Residuals of the fitted model should behave as a white noise. To check that we can use Ljung-Box test. If the assumption of white noise is not satisfied, we need to find more appropriate model.

We applied this procedure on our two samples, ie. on data before interventions and during interventions. If our estimated model fits the data well on the whole sample we can expect it to be good on all subsamples as well.

4.1.3 Forecasting

For the purpose of pseudo-out-of-sample forecast the two original samples were split into two subsamples each. We initially dedicated two years of observations for learning. For the estimations and predictions we used both rolling samples as well as recursive samples, what is suggested by different studies, for instance Mućk & Skrzypczyński (2012) and Kräger & Kugler (1993).

The rolling samples approach uses fixed length windows of data to re-estimate the models over set period of time, whereas the recursive approach makes use of an increasing window to re-estimate the models.

It brings up the question of appropriately chosen learning set in rolling samples. The minimum required is 50 observations, but at least 100 observations are recommended. Mućk & Skrzypczyński (2012) used 4 years of weekly data, ie. 208 observations for learning. We decided to use the minimum recommended two years of data for learning, ie. 104 observations. The reason is that we wanted to analyze performance of all the models also under CNB interventions.

In this thesis we were evaluating forecasts of three, six and twelve months ahead. Firstly we log-differenced the original series to obtain stationary series of logarithmic exchange rate changes (or shortly returns). Therefore we were forecasting multiple steps ahead using this transformed series and different models. Since comparison of predicted returns and original returns multiple steps ahead does not provide much valuable information we decided to accumulate the predicted returns from period 1 to n where n symbolizes the number of steps ahead to be predicted and reconstruct them back to the logarithmic series. We utilized the fact that in case of log returns the multiperiod log return is a sum of single period log returns. This reconstructed series was then compared with the benchmark series, ie. random walk.

In case of random walk forecast we simply moved the logarithmic series n steps ahead where n stands for number of periods to forecast. Similar approach as above would not provide the time series worthwhile for confrontation.

Both R functions we used for forecasting, `forecast` and `predict`, use the chain rule of forecasting.

4.1.4 Measuring Accuracy

We used RMSE and TheilU criteria to compare the forecast accuracy. The first one is available in `rmse` function in *hydroGOF* package, `TheilU` function had to be programmed manually with the use of the above mentiond `rmse` function. We also applied the Diebold-Mariano test which is ready to use via `dm.test` function in *forecast* package.

Another criterion we implemented and discussed in our research was based on the direction of the move of the logarithmic exchange rate. We were particularly interested whether our forecasts predicted correctly at least the direction of the move of the logarithmic exchange rate.

4.1.5 Univariate Time Series

We built up our research from the basics. Using the Box-Jenkins methodology or information criteria approach and fitting functions in R (`arma` function from *tseries* package and `ugarchspec, ugarchfit` functions from *rugarch* package) we shortly compared ARIMA model with the random walk to show this linear model is not adequate to forecast exchange rates. Another reason why we started with ARIMA model is to prove there is a nonlinear structure in the exchange rate series. The linear fit is required by the BDS test.

ARIMA/GARCH was used to show that we can sometimes model such nonlinearities using this hybrid model. In this case more detailed analysis was not performed as it is beyond the interest of this thesis.

The nonlinear models we introduced were the SETAR model (threshold variable is delayed series itself) and TAR models (threshold variable can be an exogenous variable).

To fit a SETAR model we used a `selectSETAR` and `setar` functions from *tsDyn* package. It is based on a Tong (1990) suggestions. The same function can be easily used for general threshold models since we can specify our own threshold variable. To forecast future values we used a `predict` function which allows us to use different forecasting methods, for instance Monte-Carlo method, bootstrap or block-bootstrap. `predict` is also available in the *tsDyn* package.

Nonlinearity Since nonlinearity plays a crucial role in our study we will test it with BDS test, which is ready to use as `bdsTest` function included in *fNonlinear* package. When applied to the residuals from a fitted linear time series model (ARMA(p,q) for example), the BDS test can be used to detect remaining dependence and the presence of omitted nonlinear structure (Zivot & Wang 2007).

4.2 Testing for Nonlinearity

4.2.1 BDS Test

As we mentioned above we used a BDS test to detect nonlinearity in the time series. It was firstly published by Brock *et al.* (1996) and even though its original purpose was to detect non-random chaotic dynamics, it was proven to be powerful against wide range of nonlinear alternatives (Zivot & Wang 2007). One of the advantages is that the BDS does not require any distributional assumptions on the tested data.

Let us define m-history of a time series as $x_t^m = (x_t, x_{t-1}, \dots, x_{t-m+1})$, $t = 1, 2, \dots, T$. The BDS statistic is based on the correlation integral[1] and it is defined as follows:

$$V_{m,\epsilon} = \sqrt{T}\frac{C_{m,\epsilon} - C_{1,\epsilon}^m}{s_{m,\epsilon}} \tag{4.1}$$

[1]For further details see Brock *et al.* (1996) or Zivot & Wang (2007)

where

- $C_{m,\epsilon}$ is the correlation integral at embedding dimension m, see (4.2)
- $C^m_{1,\epsilon}$ is a special case of correlation integral when x_t are considered independently and identically distributed; the exact formula is derived below with the use of equations (4.2), (4.3) and (4.4)
- $s_{m,\epsilon}$ is the standard deviation of $\sqrt{T}(C_{m,s} - C^m_{1,\epsilon})$,

The correlation integral at embedding dimension m can be estimated by (Zivot & Wang 2007):

$$C_{m,\epsilon} = \frac{2}{T_m(T_m-1)} \sum_{m \le s <} \sum_{t \le T} I(x^m_t, x^m_s; \epsilon) \tag{4.2}$$

where

- $T_m = T - m + 1$, T is the length of the time series,
- $I(x^m_t, x^m_s; \epsilon)$ is an indicator function equal to one when $|x_{t-i} - x_{s-i}| < \epsilon$ for $i = 0, 1, \ldots, m-1$ and zero otherwise.

In other words, the correlation integral estimates the probability that any two m-dimensional points are within a distance of ϵ of each other (Zivot & Wang 2007), thus the joint probability is estimated as follows:

$$Pr(|x_t - x_s| < \epsilon, (|x_{t-1} - x_{s-1}| < \epsilon, \ldots, (|x_{t-m+1} - x_{s-m+1}| < \epsilon \tag{4.3}$$

In practice the distance threshold ϵ is specified in units of sample standard deviations. In case of x_t being independently and identically distributed this probability formula shrinks in the limiting case to:

$$C^m_{1,\epsilon} = Pr(|x_t - x_s| < \epsilon)^m \tag{4.4}$$

Under reasonable regularity conditions it then applies that BDS statistic converges in distribution to $N(0,1)$:

$$V_{m,\epsilon} \xrightarrow{d} N(0,1) \tag{4.5}$$

For correct testing we have to fit our (first-differenced) time series to a linear model and then apply the BDS test to the residuals. The null hypothesis is then:

$$H_0 : \textit{Remaining residuals are iid.} \tag{4.6}$$

If we reject the null hypothesis it implies the mis-specification of the original linear model and so we can speak about hidden nonlinearity. BDS test is a two-tailed test, we reject the null if $|V_{m,\epsilon}| > 1.96$.

4.3 Measuring the Accuracy of Forecasts

Throughout the thesis we used several criteria to compare the accuracy of forecasts of the selected models.

4.3.1 MSE and RMSE

MSE is one of the most common measures of forecasting accuracy. It is defined as follows:

$$MSE = \frac{1}{T}\sum_{t=1}^{T}(X_t - F_t)^2 = \frac{1}{T}\sum_{t=1}^{T} e_t^2 \tag{4.7}$$

where

- X_t is the actual data series,
- F_t is forecasted series produced by some model or method,
- e_t is the forecast error,
- T is the number of observations used to compare both series.

Frequently used alternative representation of MSE is RMSE defined as:

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(X_t - F_t)^2} = \sqrt{\frac{1}{T}\sum_{t=1}^{T} e_t^2} \tag{4.8}$$

In both cases the smaller the value is the better forecast we have. The advantage of both approaches is that they include quadratic loss function which gives more weight to larger prediction errors. On the other hand this could be considered also as disadvantage in case of extreme values of prediction errors.

4.3.2 TheilU Statistics

TheilU statistics allows comparison among models. It is defined in the following way:

$$TheilU = \sqrt{\sum_{t=1}^{T} \frac{\frac{1}{T}\left(\frac{X_t - F_t}{X_t}\right)^2}{\frac{1}{T}\left(\frac{X_t - FN_t}{X_t}\right)^2}} = \sqrt{\sum_{t=1}^{T} \frac{\left(\frac{X_t - F_t}{X_t}\right)^2}{\left(\frac{X_t - FN_t}{X_t}\right)^2}} \tag{4.9}$$

where

- FN_t is a benchmark forecast, in this thesis random walk was used,
- other notations refer to the same variables as in the case of MSE and RMSE,
- we can easily rewrite the equation using familiar RMSE, where nominator and denominator are RMSE of the proposed forecasting and benchmark model respectively.

Depending on TheilU value we distinguish three cases:

$$TheilU \in \begin{cases} \langle 0, 1) & \text{the forecast is better than benchmark} \\ \langle 1 \rangle & \text{the forecast is as accurate as benchmark} \\ (1, \infty) & \text{the forecast is worse than benchmark} \end{cases}$$

If $TheilU = 0$ we have the exact forecast. Since it involves RMSE it is also hugely influenced by outliers. Moreover, it provides somehow incomplete information. Having two models and their TheilU's (for instance $TheilU_1 = 0.7$ and $TheilU_2 = 0.8$) reveals what model is better, but it is not apparent what it says about how much better the first model actually is.

4.3.3 Diebold-Mariano Test

Diebold & Mariano (1995) presented a hypothesis testing concept with regard to forecasting accuracy of distinct models. It was originally motivated by "cheap" vs sophisticated model comparison and its significance, ie. whether the sophisticated model has enough benefits over the "cheap" one.[2] They developed the following statistic to use it in equal accuracy hypothesis testing:

[2]It is actually questionable, the cost of more sophisticated model are rather small. Moreover "... directors of forecasting institutions may actually prefer a sophisticated model over a simple one, as such choice will improve the reputation of the institution." (Kunst 2003, page 1)

$$S = \frac{\frac{1}{T}\sum_{t=1}^{T}\left(g(e_{1t}) - g(e_{2t})\right)}{\sqrt{\frac{2\pi \hat{f}_d(0)}{T}}} = \frac{\frac{1}{T}\sum_{t=1}^{T} d_t}{\sqrt{\frac{2\pi \hat{f}_d(0)}{T}}} = \frac{\bar{d}}{\sqrt{\frac{2\pi \hat{f}_d(0)}{T}}} \tag{4.10}$$

where

- $e_{it}, i = 1, 2$, refer to forecast errors from two different models
- $g(\cdot)$ is a loss function
- $\hat{f}_d$ is a consistent estimator of the spectral density of d_t at frequency 0[3]

The null hypothesis tested is:

$$H_0 : E[g(e_{1t})] = E[g(e_{2t})] \tag{4.11}$$

or using the above notation:

$$H_0 : E[d_t] = 0 \tag{4.12}$$

Ie. the null hypothesis is that the two methods have the same forecast accuracy. The advantage of this approach is that it allows the loss function as well as the errors to be of a variety of types (see Diebold & Mariano (1995) for more details). One of the weak points of Diebold-Mariano test is the hypothesis testing itself as it is subject to type I and type II errors, ie. with regard to model selection we cannot consider it appropriate if compared with the information criteria for example (Kunst 2003).

4.3.4 Success Ratio

This function analyses how successful were corresponding models in forecasting the direction of the move of the exchange rate. In other words we checked only whether the CZK appreciations and depreciations were forecasted correctly, but we threw out the exact values of particular forecasts.

We created a success ratio measurement which is defined as follows:

- Let DR denotes number of forecasts when we correctly predicted that CZK appreciated or depreciated; we were interested just in the direction of this forecast and not its value,

[3]For further details see Diebold & Mariano (1995) or Kunst (2003)

- let L denotes the overall number of forecasts we made while forecasting n steps ahead.

Success ratio is then:

$$successRatio = \frac{DR}{L} \tag{4.13}$$

The higher the ratio the more successful particular model was in forecasting the direction. The reasoning behind this is based on the idea that while any model could overestimate or underestimate its predictions (and overall can be worse than random walk), it could on the other hand successfully predicts the directions and therefore provide better forecasts than random walk in this manner, because random walk as it is defined says that the exchange rate remains the same and thus do not appreciate or depreciate. The disadvantage of this approach is that we cannot evaluate its significance based on our results.

4.4 Random Walk and Univariate Models

The purpose of this section is by no means to introduce these models to a reader but rather give a brief repetition of the basic concepts further used in the thesis.

4.4.1 Random Walk

Random walk without drift is considered and widely used as a benchmark in the exchange rate forecasting comparisons (see Chapter 2). A time series r_t is a random walk process if:

$$r_t = r_{t-1} + \epsilon_t \tag{4.14}$$

where r_t refers to logarithm of exchange rate and ϵ_t is a white noise series. The k-steps ahead forecast is defined as

$$r_{T+k} = r_T \tag{4.15}$$

4.4.2 ARIMA

Let us consider a standard ARMA(p,q) model:

$$r_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} \tag{4.16}$$

where p denotes the order of AR model and q denotes the order of Moving Average (MA) model. ARIMA(p,d,q) model is a generalization of ARMA(p,q) to handle non-stationary time series. A common approach to control the non-stationarity is the use of differencing ($r'_t = r_t - r_{t-1}$). Parameter d in ARIMA(p,d,q) stands for its degree. ARIMA is unit-root non-stationary model with strong memory. From now on whenever we use ARMA notation we refer to differenced series, therefore we refer to ARIMA(p,1,q).

In the case of ARMA models we have an alternative to Information criteria procedure for estimation and model selection. We can use Box-Jenkins methodology. The main difference is the use of ACF and PACF functions to determine the ARMA(p,q) orders.

Let h be the forecast origin. Then we can derive the 1-step ahead forecast r_{h+1} from the model as

$$r_{h+1} = \phi_0 + \sum_{i=1}^{p} \phi_i r_{h+1-i} + \sum_{j=1}^{q} \theta_j \epsilon_{h+1-j}. \tag{4.17}$$

The multistep ahead forecasts of an ARMA model can be computed recursively (Tsay 2002).

4.4.3 GARCH

Bollerslev (1986) extended the Auto Regressive Conditional Heteroskedasticity (ARCH) model, first systematic framework to specify volatility, to deal with the empirically found long memory in ARCH models and allow more flexible lag structure.[4] We say that $a_t = r_t - \mu_t$ (mean-corrected log return) follows a GARCH(m, s) model if (Tsay 2002):

$$\begin{aligned} a_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2, \end{aligned} \tag{4.18}$$

[4] Empirically, relatively long lags are called for with the use of ARCH model (Bollerslev 1986). For instance, ARCH(9) is necessary to model S&P 500 monthly returns (Tsay 2002). The readers are recommended to examine this phenomena by themselves for different financial time series.

where

- $\{\epsilon_t\}$ is iid N(0, 1)
- $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0,$
- $\sum_{i=1}^{max(m,s)}(\alpha_i + \beta_i) < 1$, ie. unconditional variance of a_t is finite and conditional variance σ_t^2 evolves over time.

If we allow $s = 0$ we get ARCH(m) model. Similarly to ARMA model the forecasts of GARCH model can be obtined recursively. Let us again denote h the forecast origin, then 1-step ahead forecast is defined as

$$\sigma_{h+1}^2 = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2. \tag{4.19}$$

For multistep ahead forecast we use $a_t^2 = \sigma_t^2 \epsilon_t^2$ and since $E(\epsilon_{h+1}^2 | F_h) = 1$ we can rewrite

$$\begin{aligned} \sigma_{h+1}^2 &= \alpha_0 + (\alpha_1 + \beta_1)\sigma_h^2, \\ \sigma_{h+2}^2 &= \alpha_0 + (\alpha_1 + \beta_1)\sigma_{h+1}^2, \\ &\dots \\ &\dots \\ \sigma_{h+l}^2 \rightarrow & \frac{\alpha_0}{1 - \alpha_1 - \beta_1}, \ for\ l \rightarrow \infty. \end{aligned}$$

The forecast converges to unconditional variance for forecasting horizon going to infinity.

4.4.4 ARMA/GARCH

This model is a combination of two models defined above. Tsay (2002) describes a procedure to build this type of model for asset return series which we can use for exchange rate changes series as well. It comprises of four steps:

- Specify a mean equation and remove any linear dependence,
- test the residuals for ARCH effects,
- specify a volatility model if ARCH effects are statistically significant and perform a joint estimation with the mean equation,

- check the resulting model and refine if necessary.

We used this model as a short excursion to a different approach to threshold models to carry out nonlinearities in time series.

4.5 Univariate Threshold Models

4.5.1 Threshold Auto Regressive Model

TAR models were firstly introduced in Tong & Lim (1980). The general form is described by the following equation:

$$X_t = a_0^{(J_t)} + \sum_{i=1}^{p} a_i^{(J_t)} X_{t-i} + b^{(J_t)} \epsilon_t, \tag{4.20}$$

with

- ϵ_t being iid $(0, \sigma_2)$ and
- J_t is a switching mechanism.

The switching mechanism divides the model to pieces with different linear model representations. For instance, the series can follow AR(1) when the indicator is < 0 and AR(2) for indicator > 0 or the model representation could remain the same with a change in parameter values. The switching property can be easily applied to extend other models.

To correctly estimate our TAR models we used an exhaustive grid search algorithm over all possible combinations of values of specified parameters, hence threshold delay, threshold value and number of lags in each regime are computed. We will not get into details of this algorithm, but we refer the reader to Gonzalo & Pitarakis (2002) were the comprehensive description of model selection framework in the context of multiple threshold models is provided. The best model is chosen according to the AIC information criteria.

Parametric bootstrap method is used in practice for forecasting (Tsay 2002). Let h denote the forecast origin and l denote the forecast horizon. The forecast l-step ahead forecast is computed sequentially in two steps:

- draw a new innovation from the specified innovational distribution of the model,

- given the model, data and previous predictions $X_{h+1}, ..., X_{h+i-1}$ compute X_{h+l}.

By repeating this procedure M times we get M realizations of X_{h+l}. The point forecast is then the sample average of all M realizations of X_{h+l}. Tsay (2002) recommends $M = 3000$.

4.5.2 Self-Exciting Threshold Auto Regressive Model

SETAR model, firstly introduced by Tong & Lim (1980), is one of the most popular models from TAR models family. A time series X_t follows a k-regime SETAR with itself delayed as a threshold variable if it satisfies (Tsay 2002):

$$X_t = \phi_0^{(j)} + \sum_{i=1}^{p} \phi_i^{(j)} X_{t-p} + \epsilon_t^{(j)}, \qquad \text{if } \gamma_{j-1} \leq X_{t-d} < \gamma_j, \tag{4.21}$$

where

- $k, d > 0$, d is a delay parameter,
- $j = 1, \ldots, k$ refers to a particular regime,
- $\gamma_i \in R$, $-\infty = \gamma_0 < \gamma_1 < \ldots < \gamma_{k-1} < \gamma_k = \infty$, γ_i are thresholds,
- $\epsilon_t^{(j)}$ is each strictly white noise

Now we face two problems before parameters estimation itself - how to choose a delay parameter d and how to find out the right thresholds. For the delay parameter d there are two different approaches. Tong & Lim (1980) suggests to use Akaike or Schwarz information criteria to select it after all other parameters have already been chosen, whereas Tsay (1989) provides his own method to select d before even having thresholds. Regarding thresholds Chan (1993) suggested that they are elements of the series itself and proposed a search algorithm.

In this thesis we decided to use an approach the same approach as in the case of SETAR models that deals with both problems mentioned above at once - an exhaustive grid search algorithm over all possible combinations of values of specified parameters. The best model is again chosen according to the AIC information criteria.

There are several methods applicable for forecasting, for instance Bootstrap method, Monte Carlo method or Dynamic Estimation method to name a few

(Clements & Smith 1997). We decided to use the bootstrap method to stay consistent with TAR models forecasting.

We analyzed SETAR models with one and two thresholds. We believed that the log differenced exchange rates time series behaves differently when there are appreciations and when there are depreciations of CZK against EUR. The other approach, ie. the use of two thresholds was used in Kräger & Kugler (1993). The authors expected the time series to behave differently in case of strong appreciation and depreciations and in case of moderate changes. We ran this analysis as well.

The notation we further used in the thesis is as follows. For SETAR model with one threshold the SETAR(m, n) was used, where m and n stand for the AR order in the lower and higher regime respectively. When talking about models with two thresholds the final model is labeled SETAR(m, n, o) where analogically m, n and o stand for the AR order in low, middle and high regime.

4.6 Multivariate Time Series

One of the original ideas of this thesis was also to introduce multivariate threshold models and compare them with univariate ones. But the analysis of univariate models was so voluminous and interesting that we decided to focus this thesis only on advanced univariate models. We eventually decided to keep the methodology parts related to multivariate threshold models for prospective future extension of this thesis even though they are not used in this thesis.

We based the multivariate analysis on VAR and TVAR models. These models used several variables, namely money supply, short term interest rate, price level and output as suggested by Cuaresma & Hlouskova (2004).

Fitting and forecasting functions for VAR model, namely `VAR`, `VARselect` and `predict`, are available in *vars* package. *tsDyn* package offers all necessary functions to fit and forecast TVAR model, namely `TVAR` and `predict`. The latter function is masked based on a package used in the computations.

Finally we compared univariate models to multivariate using the Mean Absolute Deviation (MAD)/mean ratio. Mean Absolute Percentage Error (MAPE) could be used as well unless there are zeros in the time series.

Before introducing TVAR let us define (linear) VAR model which we will use as a basis for (nonlinear) TVAR model definition and also for comparison of forecasts.

4.6.1 Vector Auto Regressive Model

The VAR model is a multivariate extension of the univariate AR model. As such it has several advantages. Principally the biggest advantage steps from the nature of the model - it describes and captures the dynamics in multiple time series, ie. it allows the investigated variable to be influenced by other endogenous variables in the model. It also (often) provides superior forecasts to univariate series. In addition it can be used for policy analysis and structural inference. However it posses also disadvantages. The major one is the number of parameters to be estimated. Others include dealing with overparametrization, overfitting or identification problem.

4.6.2 Specification of VAR Model and its Estimation

The VAR is a model of n equations and n variables where each variable is explained by its own lagged values and current and past values of the remaining n-1 variables. Formally VAR(p) where p denotes lag, can be written as

$$Y_t = \mu + \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + \ldots + \Pi_p Y_{t-p} + \epsilon_t = \mu + \sum_{i=1}^{p} \Pi_i Y_{t-i} + \epsilon_t, \quad (4.22)$$

with

- $t = 1, \ldots, T$,
- Π_i are $(n \times n)$ coefficient matrices,
- ϵ_t is $(n \times 1)$ unobservable zero mean white noise vector process (serially uncorrelated or independent) with time invariant covariance matrix $\sum$

We considered no deterministic terms nor exogenous variables. Let us assume that VAR(p) model is covariance stationary with no restrictions on parameters. We can rewrite the model to the form of Seemingly Unrelated Regression (SUR) what leads to each equation having the same explanatory variables. Now we can estimate each equation using Ordinary Least Squares (OLS) (Zivot & Wang 2007).

To choose optimal lag of the model we can use several criteria. Three most common are AIC), Schwarz-Bayesian Information Criterion (BIC) and Hannan-Quinn (HQ) (Zivot & Wang 2007).

4.6.3 Forecasting

Forecasting using VAR model is analogous to the forecasting from AR model, ie. 1-step ahead forecast of VAR(p) is obtained as follows:

$$Y_{T+1|T} = \mu + \Pi_1 Y_T + \ldots + \Pi_p Y_{T-p+1}. \tag{4.23}$$

For longer horizons we can use the so called chain rule of forecasting. It means that values predicted for T+1 are used on the right hand side of the prediction equation for T+2, etc.

$$Y_{T+h|T} = \mu + \Pi_1 Y_{T+h-1|T} + \ldots + \Pi_p Y_{T+h-p|T}. \tag{4.24}$$

where $Y_{T+j|T} = Y_{T+j}$ for $j \leq 0$

4.6.4 Threshold Vector Auto Regressive model

Similar to the case of univariate TAR and AR models, TVAR models are generalized VAR models to capture possible nonlinearities. Besides its simplicity the other advantage is that the variable we use for model switching could itself be an endogenous variable included in the model. This means we can switch regimes after shocks happened to each variable.

4.6.5 Specification of TVAR and its Estimation

Based on (Tsay 1998) we say that Y_t follows a multivariate TVAR with threshold variable z_t and delay d if it satisfies

$$Y_t = \mu_j + \sum_{i=1}^{p} \Pi_i^{(j)} Y_{t-i} + \sum_{i=1}^{q} \Phi_i^{(j)} X_{t-i} + \epsilon_t^{(j)}, \text{if } r_{j-1} < z_{t-d} \leq r_j, \tag{4.25}$$

where

- $-\infty = r_0 < r_1 < \ldots < r_{s-1} < r_s = \infty$, r_i are thresholds, s is number of regimes,
- X_t are v-dimensional exogenous variables,
- μ_j are constant vectors,
- p and q are non-negative integers,

- the threshold variable z_t is assumed to be stationary with continuous distribution.

The obvious problem which steps out is the identification of the threshold variable z_t. The most difficult according to Tsay (1998) is the specification of number of regimes. Based on the analyzed data it is often restricted to 2 or 3 regimes, what is again supported by Tsay (1998) or Kräger & Kugler (1993). When given s and z_t the selection of the model is then based on modified AIC:

$$AIC(p,q,d,s) = \sum_{j=1}^{s}[2ln(L_j(p,q,d,s)) + 2k(kp+vq+1)], \tag{4.26}$$

where $L_j(p,q,d,s)$ is the likelihood function of regime j evaluated at the maximum likelihood estimates of $\mu_j, \Pi_i^{(j)}$ and $\Phi_j^{(j)}$ (Tsay 1998).

4.7 Summary of the Models Used in the Thesis

Throughout the thesis we investigated the following models:

- ARIMA as a representative of linear models family and a model on whose residuals we applied the BDS test to check the nonlinearity,
- SETAR with one threshold since we believe that the exchange rate time series behaves differently when we observe appreciations and when depreciations,
- SETAR with two thresholds based on Kräger & Kugler (1993) suggestion that the time series behaves differently in case of strong appreciations and depreciations and in case of moderate changes,
- TAR model with CNB Board meeting as a dummy threshold variable, because we identified higher mean and volatility of our time series during weeks of CNB Board meetings,
- TAR model with volatility in the last periods as a threshold variable what is based on revealed volatility clusters in return series.

The data set was split to the periods before CNB started interventions in 2013 and during the interventions. The analysis was processed using both rolling and recursive samples. For measurement of forecasting accuracy we used RMSE, TheilU, Diebold-Mariano test and success ratio.

Chapter 5

Results

To remind readers the analysis of the forecasting performance of all models used in the thesis is split into two periods - we analyse the period before CNB decided to use interventions on November 7th, 2013 and the period during the interventions, ie. from November 7th, 2013 till April 18, 2016.[1] The "interventions" jump, which followed the CNB decision, could not be predicted by any conventional method. We explicitly mentioned the time period under examination in all cases.

For forecasting we used rolling and recursive samples. Since we are interested in the forecasting performance rather than the stability of our models we used a common approach of re-fitting the data in each rolling and recursive step. All forecasts were compared in means of RMSE and TheilU statistics. We also performed the Diebold-Mariano test based on hypothesis testing concept. The last measurement we employed was success ratio representing the ratio of correctly predicted directions of exchange rate move.

Results were compared with random walk forecast what is widely used benchmark in analyzing exchange rates models. In all tables we repeat the following notation - 3m, 6m and 12m - they stand for three months, six months and twelve month ahead forecasts.

In the period during interventions we did not have enough pseudo-out-of-sample data to check the forecasting performance of six and twelve months ahead. Hence we examined only the three months ahead forecasts.

[1] To be clear April 18, 2016 is the day when we downloaded our data, but it is not the day when CNB stopped to use interventions. CNB still uses this monetary tool as of December 26th, 2016.

5.1 Univariate Linear Models

5.1.1 ARIMA

Let us at first recall the Data and Tools section (Chapter 3), where we found the original series to be non-stationary. Before proceeding to estimation and rolling and recursive forecasts we used the log transformation of the original series as it provides several advantageous properties. Then we differentiated the original series once. The final series is stationary and integrated of order 1 (I(1)). We will refer to this series as a return series.

5.1.2 Model Identification

Before interventions

Since neither ACF nor PACF provided any useful information when applied on our data we used the information criteria approach. We checked all ARMA(p,q) models where p,q = 0,...,5 and while taking a closer look at the parameters values we chose ARMA(1,1) without intercept as the best model. It was the one with the lowest AIC value. All parameters were significant at 5% level.

$$r_t = 0.68629 r_{t-1} - 0.71950 \epsilon_{t-1} + \epsilon_t$$

The adequacy of the model was checked on residuals by Ljung-Box test. It showed we cannot reject the null hypothesis of no residual serial correlation since the p-values up to 10th lag are higher than 0.05. ACF of model residuals confirmed our conclusion, see Figure 5.1. There are some significant peaks at lag 18 and 20, but we decided not to take them into account in order to keep simplicity. We confirmed the residuals to be serially uncorrelated, but they were still dependent. Volatility clusters (Figure 5.2) suggested dependency of squared residuals. The dependency is confirmed by the ACF of squared residuals, see Figure 5.3.

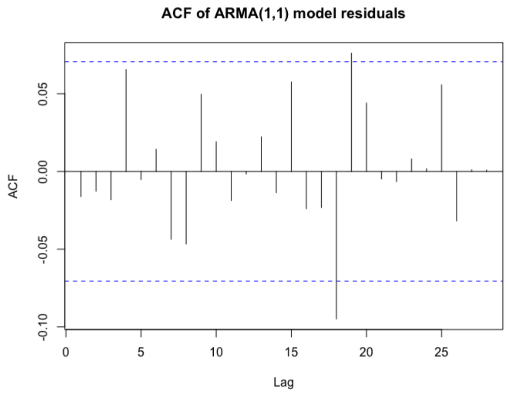

Figure 5.1:
ARMA(1,1) residuals ACF

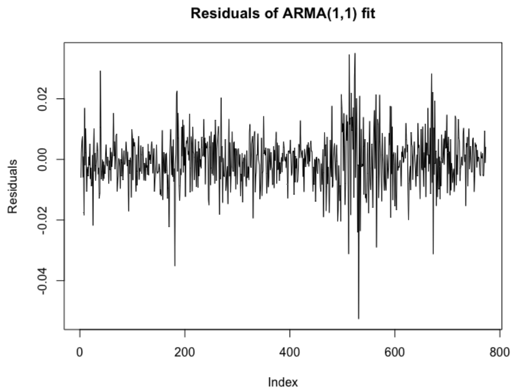

Figure 5.2:
ARMA(1,1) fit residuals

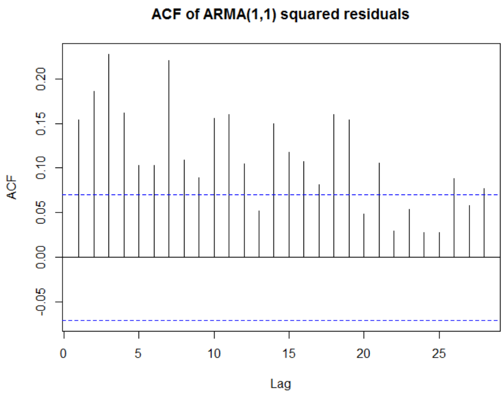

Figure 5.3:
ARMA(1,1) squared residuals ACF

During interventions

We used the same approach to find that in this case ARMA(1,3) provides the best results in fitting the data. All parameters were again significant at 5% level.

$$r_t = 0.50078 r_{t-1} - 0.62797 \epsilon_{t-1} + 0.29237 \epsilon_{t-2} - 0.34764 \epsilon_{t-3} + \epsilon_t$$

Ljung-Box test does not reject the null hypothesis of no residual serial correlation. ACF plot and residuals of fit plot are available in Figure 5.4 and Figure 5.5 respectively. The ACF plot of squared residuals in this case shows only the first lag to be significant compared to the period before interventions.

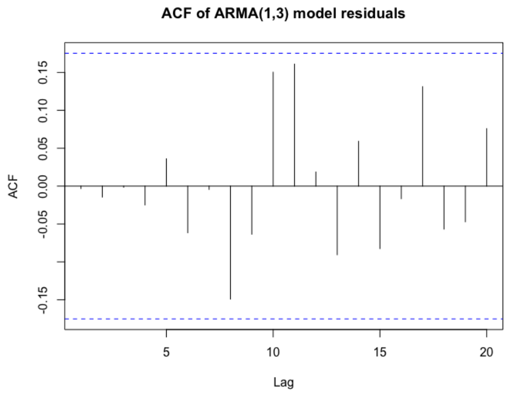

Figure 5.4:
ACF of ARMA(1,3) residuals

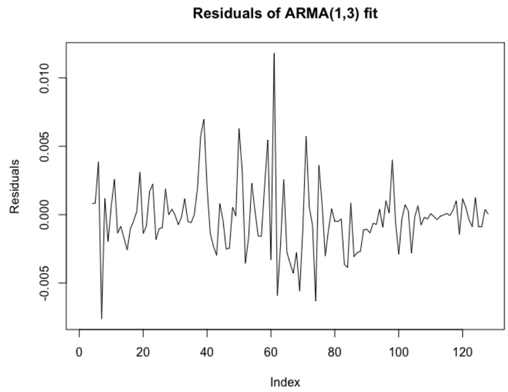

Figure 5.5:
ARMA(1,3) fit residuals

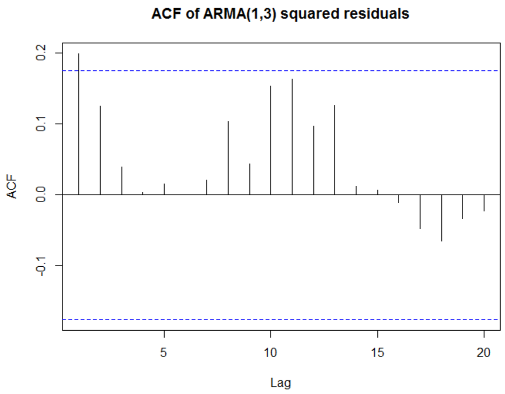

Figure 5.6:
ARMA(1,3) squared residuals ACF

5.1.3 BDS Test for Nonlinearity

Before proceeding to forecasting we checked the nonlinearity in our data series. We applied the BDS test to ARIMA models' residuals given the models above. The null hypothesis is that the residuals are independently and identically distributed.

Before interventions

For a better idea we just once in this thesis present the output of BDS test below. We can see that the null hypothesis is rejected for all combinations of m (*embedding dimension*) and ϵ. The result therefore suggests that the linear model is mis-specified and we can speak about a nonlinear structure in the data.

```
Embedding dimension =  2 3 4 5
Epsilon for close points =  0.0043 0.0086 0.0129 0.0171
p-value =
[ m ] [ 0.0043 ] [ 0.0086 ] [ 0.0129 ] [ 0.0171 ]
[ 2 ]          0          0          0          0
[ 3 ]          0          0          0          0
[ 4 ]          0          0          0          0
[ 5 ]          0          0          0          0
```

During interventions

Residuals from our fitted model of data during interventions provides the same conclusion as in the previous case. Null hypothesis is rejected for all combinations of m and ϵ as well and therefore again suggesting a nonlinear structure in the data.

```
Embedding dimension =  2 3 4 5
Epsilon for close points =  0.0013 0.0026 0.0039 0.0052
p-value =
[ m ] [ 0.0013 ] [ 0.0026 ] [ 0.0039 ] [ 0.0052 ]
[ 2 ]          0          0     0.0011     0.0286
[ 3 ]          0          0     0.0001     0.0061
[ 4 ]          0          0     0.0000     0.0019
[ 5 ]          0          0     0.0000     0.0004
```

5.1.4 Forecasting

Both models in the previous sections are quite simple, easy to understand and providing promising basis for the exchange rate forecasting.

Before interventions

ARMA(1,1) rolling and recursive forecasting performance results are summarized in Table 5.1 and Table 5.2 respectively.

Regarding rolling samples it gave us results that we somehow expected based on Meese & Rogoff (1983). In all three cases the performance of ARMA(1,1) was worse than RW in terms of RMSE and TheilU. On the other hand the Diebold-Mariano test could not reject the null hypothesis of both methods having the same forecast accuracy and so it could not favor neither ARMA(1,1) nor RW.

ARMA(1,1) was in all three cases able to successfully forecast the right direction of exchange rate move only in a little more than 50% cases.

The recursive samples method provided better results than the rolling one. In three and six months ahead forecasts the ARMA(1,1) was even better than RW in terms of RMSE and TheilU, but only a little bit. Diebold-Mariano test revealed we could not reject its null hypothesis what signalizes that both ARMA(1,1) and RW have comparable forecasting accuracy. The success ratio is close to or above 60%. This is rather surprising.

Table 5.1: ARMA(1,1) and RW forecasting performance, rolling samples, before interventions

Model	RMSE	TheilU	DM p-value	Success ratio
ARMA(1,1) 3m	0.03230934	1.068957	0.2597	0.5570776
ARMA(1,1) 6m	0.04928561	1.113638	0.3182	0.5326087
ARMA(1,1) 12m	0.07458868	1.201788	0.1753	0.5113269
RW 3m	0.03022511	1	N/A	N/A
RW 6m	0.04425642	1	N/A	N/A
RW 12m	0.06206476	1	N/A	N/A

Table 5.2: ARMA(1,1) and RW forecasting performance, recursive samples, before interventions

Model	RMSE	TheilU	DM p-value	Success ratio
ARMA(1,1) 3m	0.02981656	0.9864831	0.7001	0.5920852
ARMA(1,1) 6m	0.04415803	0.9977768	0.9716	0.6195652
ARMA(1,1) 12m	0.06270562	1.010326	0.9111	0.6148867
RW 3m	0.03022511	1	N/A	N/A
RW 6m	0.04425642	1	N/A	N/A
RW 12m	0.06206476	1	N/A	N/A

During interventions

The results in the period during the recent interventions favor RW over ARMA(1,3) in all accuracy measurements as we can see in Table 5.3 and Table 5.3. In case of rolling samples Diebold-Mariano test actually rejects the null hypothesis of the same forecasting performance. We believed that RW was performing significantly better and our belief was confirmed by one sided Diebold-Mariano test with alternative hypothesis of RW method being more accurate. The null hypothesis was rejected in favor of the alternative one with the p-value = 0.000254.

Table 5.3: ARMA(1,3) and RW forecasting performance, rolling samples, during interventions

Model	RMSE	TheilU	DM p-value	Success ratio
ARMA(1,3) 3m	0.002512549	3.04563	0.000508	0.25
RW 3m	0.000824969	1	N/A	N/A

Table 5.4: ARMA(1,3) and RW forecasting performance, recursive samples, during interventions

Model	RMSE	TheilU	DM p-value	Success ratio
ARMA(1,3) 3m	0.000837945	1.015729	1	0.4166667
RW 3m	0.000824969	1	N/A	N/A

5.1.5 ARMA/GARCH

The nonlinearities in the financial time series suggested by the BDS test are quite often handled using hybrid ARMA/GARCH model. In our case we at first used the Lagrange multiplier test on our ARMA fits residuals to test for ARCH effects. The null hypothesis of no ARCH effects is strongly rejected for the period before interventions with the p-value = 2.687e-14. On the other hand it could not be rejected for the period during interventions with the p-value = 0.5121.

Since the ARCH effects were present in the first period we used the popular GARCH(1,1) model to take care of it and then performed a joint estimation of the mean and volatility equations. The resulting model ARMA(1,1)/GARCH(1,1) had all the parameters significant at 5% significance level. The adequacy of the model to handle the ARCH effect was checked by Ljung-Box test on squared standardized residuals and the null hypothesis could not be rejected. This was confirmed by ACF of squared standardized residuals, see Figure 5.7.

As mentioned earlier in Chapter 4 we introduced this model as a short excursion to a different approach to threshold models to carry out nonlinearities in time series.

The BDS test on standardized residuals fails to reject the null hypothesis of remaining residuals to be independently and identically distributed. In this case the BDS test reveals no evidence of nonlinear dependence left in residuals.

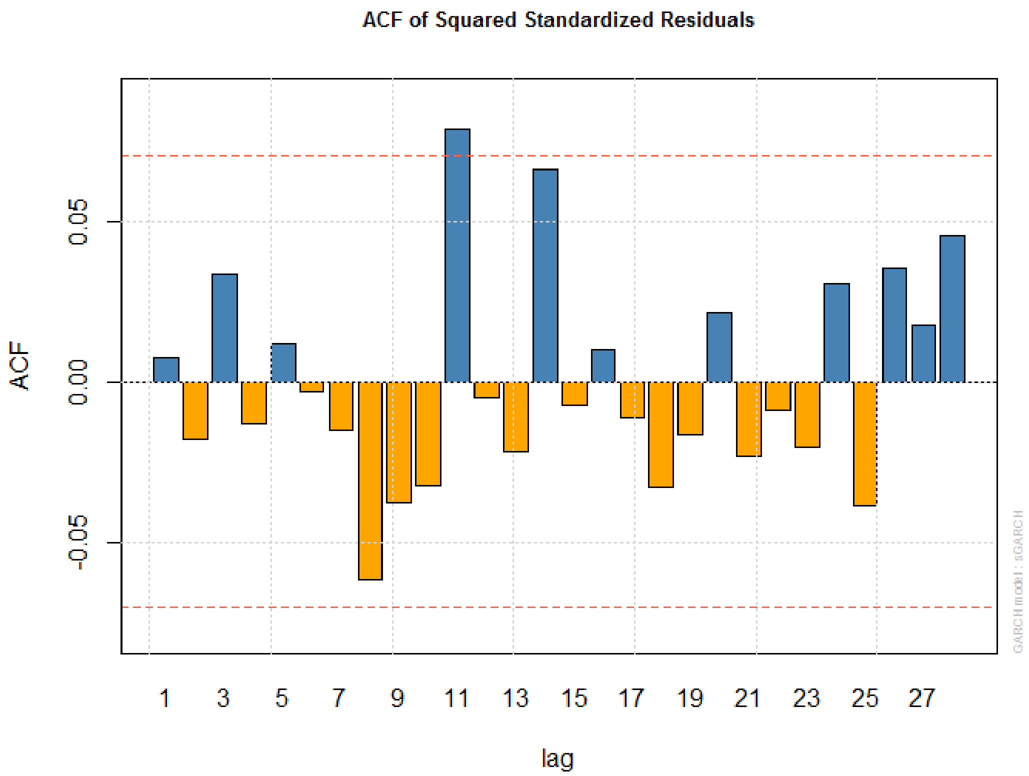

Figure 5.7: ACF of Squared Standardized Residuals

5.2 Univariate Threshold Models

5.2.1 SETAR Models

Let us bring to mind that we examined two SETAR models. The first one with one threshold value and the second one with two thresholds. See Chapter 4 for this motivation.

5.2.2 Model Identification

In all cases the exhaustive grid search over all possible thresholds values, threshold delays and number of lags was performed. We restricted the parameters as follows. Embedding dimension $m = 4$, threshold delay must be obviously smaller than embedding dimension, therefore $thDelay = 0, \ldots, 3$, number of thresholds was set based on the number of regimes of particular model under investigation and threshold values were chosen based on number of unique (in terms of their value) observations used for model identification.

Before Interventions

We used the AIC information criteria for model selection. The chosen model with one threshold and lowest AIC was SETAR(3,1). The threshold was 0.006530889 and the threshold delay was determined as 2. Formally:

$$r(t) = \begin{cases} -0.00087520 - 0.032273 r_{t-1} - 0.072006 r_{t-2} - 0.093281 r_{t-3} + \epsilon_t \\ \quad \text{if } r_{t-2} \leq 0.006530889 \\ \\ 0.000068904 - 0.14634 r_{t-1} + \epsilon_t \\ \quad \text{if } r_{t-2} > 0.006530889 \end{cases}$$

Not all the parameters were significant at 5% level and omitting some of them would improve the model a little bit. On the other hand we had to take into account the restriction in R functions where if some of the middle lags were omitted the `predict` function would not work.

Model with two thresholds and lowest AIC was SETAR(4,3,4). The computed thresholds were -0.01024 and 0.01184. Threshold delay was determined to be 1. Formally:

$$r(t) = \begin{cases} 0.0021276520 + 0.3044095899 r_{t-1} - 0.0003476093 r_{t-2} + \\ +0.1186155566 r_{t-3} - 0.4453011422 r_{t-4} + \epsilon_t \\ \quad \text{if } r_{t-1} \leq -0.01024 \\ \\ -0.001237363 - 0.083162209 r_{t-1} - 0.060026822 r_{t-2} - 0.082521586 r_{t-3} + \epsilon_t \\ \quad \text{if } -0.01024 \leq r_{t-1} \leq 0.01184 \\ \\ -0.004391336 - 0.225405437 r_{t-1} - 0.414868951 r_{t-2} + \\ +0.126937715 r_{t-3} + 0.252345896 r_{t-4} + \epsilon_t \\ \quad \text{if } r_{t-1} > 0.01184 \end{cases}$$

As in previous case not all the parameters were significant at 5% level. We handled them in the same manner. [2] The adequacy of both models was checked by Ljung-Box test and we could not reject its null hypothesis. The Ljung-Box test ran on squared residuals showed remaining dependency. The BDS test performed on both models' residuals and embedding dimension up to 5 revealed the nonlinear structure is not captured well in any case. The Lagrange multiplier test disclosed ARCH effects in residuals by rejecting its null hypothesis with the p-value = 4.552e-15.

During Interventions

During interventions the preferred model with one threshold based on AIC is SETAR(1,2) with all the parameters significant at 5% level. The threshold was 0.001516 and threshold delay was equal to 2:

$$r(t) = \begin{cases} 0.1941159 r_{t-1} + \epsilon_t & \text{if } r_{t-2} \leq 0.001516 \\ -0.4257522 r_{t-1} + 0.2857692 r_{t-2} + \epsilon_t & \text{if } r_{t-2} > 0.001516 \end{cases}$$

Two thresholds model chosen based on AIC was SETAR(1,3,1) with most of the parameters significant at 5% level, thresholds were 0.001626 and 0.002046, threshold delay was equal to 2:

[2] If not stated otherwise we always treated the nonsignificant paramaters in the same manner for reasons stated earlier.

$$r(t) = \begin{cases} 0.00009005083 + 0.1229938880 r_{t-1} + \epsilon_t \\ \quad \text{if } r_{t-2} \leq 0.001626 \\ \\ 0.06871248 - 0.60300860 r_{t-1} + 0.68019361 r_{t-2} - 37.10976160 r_{t-3} + \epsilon_t \\ \quad \text{if } 0.001626 \leq r_{t-2} \leq 0.002046 \\ \\ -0.001123077 - 0.525041311 r_{t-1} + \epsilon_t \\ \quad \text{if } r_{t-2} > 0.002046 \end{cases}$$

Ljung-Box test ran on both models' residuals could not reject its null hypothesis, in case of squared residuals Ljung-Box test rejected it only for the first two lags in case of model with two thresholds. It could not reject it in other cases. The BDS test disclosed remaining nonlinear structure in residuals after fitting data with both models described above. Nevertheless our suspicion of ARCH effects were not confirmed by Lagrange multiplier test.

5.2.3 Forecasting

Forecasting was performed using bootstrap nethod. Hence when replicating the computations the results might slightly differ. The number of replications was set to 3000 in each forecasting cycle based on Tsay (2002). We decided to keep RW forecasts in all tables just for the purpose of immediate comparison.

Before Interventions

Results of forecasting using SETAR(3,1) and SETAR(4,3,4) are sumed up in Table 5.5 and Table 5.6.

SETAR models offered worse results than RW in terms of RMSE and TheilU in all cases. Nevertheless Diebold-Mariano test rejected the possibility of RW having better forecasting accuracy.

In comparison with ARMA models the SETAR models offer mixed results using RMSE and TheilU. However they beat ARMA models in all the predictions of appreciations or depreciations (success ratio statistic).

Table 5.5: SETAR(3,1) and RW forecasting performance, rolling samples, before interventions

Model	RMSE	TheilU	DM p-value	Success ratio
SETAR(3,1) 3m	0.03164654	1.047028	0.3201	0.56621
SETAR(3,1) 6m	0.04936115	1.115344	0.2845	0.5465839
SETAR(3,1) 12m	0.07468392	1.203322	0.2017	0.5323625
RW 3m	0.03022511	1	N/A	N/A
RW 6m	0.04425642	1	N/A	N/A
RW 12m	0.06206476	1	N/A	N/A

Table 5.6: SETAR(3,1), SETAR(4,3,4) and RW forecasting performance, recursive samples, before interventions

Model	RMSE	TheilU	DM p-value	Success ratio
SETAR(3,1) 3m	0.03057484	1.011571	0.8073	0.608828
SETAR(3,1) 6m	0.04426897	1.000284	0.9975	0.6335404
SETAR(3,1) 12m	0.06217421	1.001763	0.9908	0.631068
SETAR(4,3,4) 3m	0.03188281	1.054845	0.4607	0.6073059
SETAR(4,3,4) 6m	0.04849542	1.095783	0.4972	0.6350932
SETAR(4,3,4) 12m	0.07330917	1.181172	0.4517	0.631068
RW 3m	0.03022511	1	N/A	N/A
RW 6m	0.04425642	1	N/A	N/A
RW 12m	0.06206476	1	N/A	N/A

During Interventions

Table 5.7 and Table 5.8 provide results in period during interventions. We came to the same conclusion that neither of SETAR models provided better results than RW. Again the exception is in case of success ratio where, except for the prediction by SETAR(1,2) using rolling samples, all the other models forecasted the direction of move with higher than 50% success. Anyway we have to keep in mind that we were predicting only twelve future data points.

Table 5.7: SETAR(1,2), SETAR(1,3,1) and RW forecasting performance, rolling samples, during interventions

Model	RMSE	TheilU	DM p-value	Success ratio
SETAR(1,2) 3m	0.001203081	1.458335	1	0.4166667
SETAR(1,3,1) 3m	0.001562299	1.893768	0.07789	0.75
RW 3m	0.000824969	1	N/A	N/A

Table 5.8: SETAR(1,2), SETAR(1,3,1) and RW forecasting performance, recursive samples, during interventions

Model	RMSE	TheilU	DM p-value	Success ratio
SETAR(1,2) 3m	0.00097173	1.177899	0.5325	0.5833333
SETAR(1,3,1) 3m	0.001257648	1.524479	1	0.75
RW 3m	0.000824969	1	N/A	N/A

5.2.4 TAR Models

Since R functions related to SETAR model offer the extension in a form of specifying own external threshold variable we can easily use them for TAR models analysis.

The question that immediately arises is how to choose such external threshold variable. In Chapter 2 we found some interesting facts about the exchange rate changes data. Let us remind that the mean of returns as well as their variance in the weeks where there was a CNB Board meeting were higher than in other cases. This brings us to an idea that one of the possible candidates for the external threshold variable is a dummy variable representing weeks of Boards meetings.

Few other time series in our data set, specifically repo rates, repo rates changes, interest rate on the main refinancing operations and its changes, were investigated for their influence on the exchange rates as well. We considered and examined all of them as possible threshold variable. In case of latter two, ie. time series related to ECB, their effect on EUR/CZK exchange rate was scunt. The repo rates changes time series provided the most attractive results with regards to their influence on the exchange rate time series but only before interventions. That is because the last change of repo rate was recorded on November 2nd, 2012. Therefore we stuck to the series of CNB meetings. The supporting argument is mentioned in Fišer & Horvath (2010) where authors showed the importance of CNB communication. We must add that the CNB Board meetings are immediately followed with a press conference.

Another idea is based on volatility clusters detected in the plot of returns. We consider the volatility as a good candidate for a threshold variable. Since volatility is not directly observable in financial time series we decided to use one of the most common volatility proxies - squared returns. We calculated the volatility based on average squared returns over the last few weeks:

$$Z_t = \frac{1}{n} \sum_{j=1}^{n} r_{t-j}^2 \tag{5.1}$$

where

- Z_t represents volatility and our threshold variable,
- $n = 1 \dots 26$, represents number of last periods over which we calculated the average volatility,
- t is the current time spot of the time series,
- r_{t-j}^2 are squared returns at time $t - j$.

In our research we among other things chose such n that minimizes the AIC value of the fit of our TAR models. The observed fact was that the higher the n the worse AIC ratio we got, for $n > 4$.

To sum up we investigated two TAR models with different threshold variables. The first model used series of CNB Board meetings and the second one average volatility.

We used the same notation principle as in case of SETAR models, ie. TAR(1,2) represents a model with AR(1) in low regime and AR(2) in high regime.

5.2.5 Model Identification

Similarly to SETAR models in all cases the exhaustive grid search over all possible thresholds values, threshold delays and number of lags was performed. We restricted the parameters as follows. Embedding dimension $m = 4$, threshold delay must be obviously smaller than embedding dimension, therefore $thDelay = 0, \ldots, 3$, number of thresholds was set to one based on our threshold variables specification and threshold values were chosen based on number of unique (in terms of their value) observations used for model identification.

Before interventions

Following the same approach based on AIC we used while analyzing ARIMA and SETAR models we selected TAR(1,1) to be the best model when we applied CNB Board meetings dummy as a threshold variable:

$$r(t) = \begin{cases} -0.00022992 - 0.07793668 r_{t-1} + \epsilon_t & \text{if } Z_{t-1} \leq 0.5 \\ -0.001587467 + 0.089540724 r_{t-1} + \epsilon_t & \text{if } Z_{t-1} > 0.5 \end{cases}$$

where Z_{t-1} represents the CNB Board meetings dummy.

When we used average volatility in recent weeks as threshold variable the best AIC results were provided by TAR(3,4) model together with the average volatility over the last two weeks:

$$r(t) = \begin{cases} -0.0007797951 + 0.0242006344 r_{t-1} - 0.0386308663 r_{t-2} - \\ -0.0950754556 r_{t-3} + \epsilon_t \\ \quad \text{if } Z_t \leq 0.0001272 \\ \\ -0.0006316862 - 0.2077773043 r_{t-1} - 0.0548459174 r_{t-2} + \\ +0.1095609837 r_{t-3} + 0.2073444521 r_{t-4} + \epsilon_t \\ \quad \text{if } Z_t > 0.0001272 \end{cases}$$

with Z_t representing the average volatility over the last two weeks.

As usual the adequacy of our models was tested by Ljung-Box test on residuals and squared residuals, their ability to handle nonlinearities was examined by BDS test. Regarding both models Ljung-Box test did not reject the null hypothesis of no residual serial correlation with p-values > 0.05. In case of squared residuals Ljung-Box revealed that the residuals were still dependent. BDS test rejected the null hypothesis of independently and identically distributed residuals and proved neither model could handle the nonlinearities in residuals. For completion we diagnosed the residuals using the Langrange multiplier test and it rejected the null hypothesis in favor of the alternative, ie. there are ARCH effects in residuals.

During interventions

In the period during interventions we selected TAR(2,1) to be the best model when we applied CNB Board meetings dummy as a threshold variable:

$$r(t) = \begin{cases} 0.01074175 r_{t-1} + 0.16448600 r_{t-2} + \epsilon_t & \text{if } Z_{t-1} \leq 0.5 \\ -0.5292964 r_{t-1} + \epsilon_t & \text{if } Z_{t-1} > 0.5 \end{cases}$$

where Z_{t-1} represents the CNB Board meetings dummy.

Using volatility as threshold variable the best model was TAR(2,4) with volatility in the last week as threshold variable.

$$r(t) = \begin{cases} 0.0003878168 + 0.0356212967 r_{t-1} + 0.4103271765 r_{t-2} + \epsilon_t \\ \quad \text{if } Z_{t-1} \leq 0.000002107 \\ \\ -0.0007762333 - 0.2162858775 r_{t-1} + 0.1100322624 r_{t-2} - \\ -0.1567837409 r_{t-3} - 0.2412991667 r_{t-4} + \epsilon_t \\ \quad \text{if } Z_{t-1} > 0.000002107 \end{cases}$$

with Z_{t-1} representing the average volatility over the last week.

Regarding the first model Ljung-Box test on residuals and squared residuals did not reject its null hypothesis. This was not the case for the second model where Ljung-Box test rejected its null in case of squared residuals up to the fifth lag. BDS test again rejected its null hypothesis and therefore we can speak about remaining nonlinearities in residuals, yet no ARCH effects were detected in any case.

5.2.6 Forecasting

Bootstrap method was used for forecasting with 3000 replications in every iteration as suggested by Tsay (2002).

Before interventions

Rolling and recursive forecasts using CNB meetings as threshold variable are summarized in Table 5.9 and Table 5.10. Values in RMSE and TheilU columns in all cases exhibit surprising fact of better performance of TAR over RW. Success ratio provides better results over ARIMA and SETAR models when using rolling samples and comparable results with SETAR when recursive samples took place. Only the Diebold-Mariano test mitigates our results not favoring any model over RW.

When using average volatility as threshold variable the results are comparable to the ones we just described, see Table 5.11 and Table 5.12.

Table 5.9: TAR(1,1) with CNB meetings dummy as threshold variable and RW forecasting performance, rolling samples, before interventions

Model	RMSE	TheilU	DM p-value	Success ratio
TAR(1,1) 3m	0.02973523	0.9837921	0.6825	0.6057839
TAR(1,1) 6m	0.04305763	0.9729126	0.7088	0.6350932
TAR(1,1) 12m	0.05774956	0.9304725	0.5453	0.631068
RW 3m	0.03022511	1	N/A	N/A
RW 6m	0.04425642	1	N/A	N/A
RW 12m	0.06206476	1	N/A	N/A

During interventions

Both our TAR models provided exceptionally poor results in this period, see Table 5.13, Table 5.14, Table 5.15 and Table 5.16.[3] They failed to beat RW in any statistics. Moreover one sided Diebold-Mariano test preferred RW in all

[3]We stayed consistent with our previous way to preview the results even though in this case it might have been better to make it brief and just mark it correctly in one table.

Table 5.10: TAR(1,1) with CNB meetings dummy as threshold variable and RW forecasting performance, recursive samples, before interventions

Model	RMSE	TheilU	DM p-value	Success ratio
TAR(1,1) 3m	0.02997812	0.9918282	0.8383	0.6073059
TAR(1,1) 6m	0.04327548	0.9778351	0.7679	0.6350932
TAR(1,1) 12m	0.05783566	0.9318598	0.5505	0.631068
RW 3m	0.03022511	1	N/A	N/A
RW 6m	0.04425642	1	N/A	N/A
RW 12m	0.06206476	1	N/A	N/A

Table 5.11: TAR(3,4) with volatility over the last two weeks threshold variable and RW forecasting performance, rolling samples, before interventions

Model	RMSE	TheilU	DM p-value	Success ratio
TAR(3,4) 3m	0.02972962	0.9836067	0.5489	0.6027397
TAR(3,4) 6m	0.04328494	0.9780488	0.7363	0.6350932
TAR(3,4) 12m	0.05836947	0.9404606	0.635	0.631068
RW 3m	0.03022511	1	N/A	N/A
RW 6m	0.04425642	1	N/A	N/A
RW 12m	0.06206476	1	N/A	N/A

Table 5.12: TAR(3,4) with volatility over the last two weeks threshold variable and RW forecasting performance, recursive samples, before interventions

Model	RMSE	TheilU	DM p-value	Success ratio
TAR(3,4) 3m	0.0300074	0.992797	0.8051	0.6073059
TAR(3,4) 6m	0.04298157	0.971194	0.6589	0.6350932
TAR(3,4) 12m	0.05822023	0.9380561	0.6249	0.631068
RW 3m	0.03022511	1	N/A	N/A
RW 6m	0.04425642	1	N/A	N/A
RW 12m	0.06206476	1	N/A	N/A

cases. This is rather surprising when we take into account good results in the previous period.

Table 5.13: TAR(2,1) with CNB meetings dummy as threshold variable and RW forecasting performance, rolling samples, during interventions

Model	RMSE	TheilU	DM p-value	Success ratio
TAR(2,1) 3m	0.003084047	3.738381	0.001814	0.1666667
RW 3m	0.000824969	1	N/A	N/A

Table 5.14: TAR(2,1) with CNB meetings dummy as threshold variable and RW forecasting performance, recursive samples, during interventions

Model	RMSE	TheilU	DM p-value	Success ratio
TAR(2,1) 3m	0.00257131	3.116858	0.00008429	0.3333333
RW 3m	0.000824969	1	N/A	N/A

Table 5.15: TAR(2,4) with volatility over the last two weeks threshold variable and RW forecasting performance, rolling samples, during interventions

Model	RMSE	TheilU	DM p-value	Success ratio
TAR(2,4) 3m	0.004480661	5.43131	0.002845	0.25
RW 3m	0.000824969	1	N/A	N/A

Table 5.16: TAR(2,4) with volatility over the last two weeks threshold variable and RW forecasting performance, recursive samples, during interventions

Model	RMSE	TheilU	DM p-value	Success ratio
TAR(2,4) 3m	0.007707147	9.342351	$2.992e-07$	0.1666667
RW 3m	0.000824969	1	N/A	N/A

Chapter 6

Discussion

6.1 Data

We decided to investigate only the exchange rate of CZK relative to EUR. There are two reasons. The first one is that EUR currency is simply the most important currency in this region. The other is based on the 2013 CNB decision to intervene on foreign exchange market and to keep the lower bound of 27 CZK per 1 EUR exchange rate. Upper bound remained free from interventions.

The exchange rates data are easily available for download at Statistical Data Warehouse of ECB website. Data related to CNB and ECB meetings and corresponding changes in repo rates and interest rates on the main refinancing operations were manually mined from CNB Board meetings minutes and from the press releases of Governing Council of ECB. They are free to download at institutions' websites respectively.

Chapter 2 provided descriptive statistics of our data. The common stylized facts regarding financial time series were confirmed in the case of the exchange rate time series as well. The original time series was non-stationary and exhibited serial correlation. We used log-differencing to obtain time series of exchange rate changes (or shortly returns). Returns data exposed volatility clusters, their distribution showed leptokurtosis and we rejected the null hypothesis of normality.

There was one fact that stood out of this descriptive analysis - the difference in mean and variance of returns in the periods when CNB had a Board meeting. This is in line with Fišer & Horvath (2010) who found that CNB communication (for instance press conferences after the meetings were held) affects the exchange rate. On the other hand it is a little contradictory to them in the sense of type of the effect. They showed that CNB communication aims to decrease the noise in the financial market what is obviously not our case. We attributed this to the fact that CNB meetings are typically held on Thursdays and our data represent Friday closing rates. Therefore the news and communication might not be fully absorbed and moreover the initial market reaction could be an overreaction or a reverse reaction due to the unfulfilled expectations no matter what was the Board communication. The latter fact is in line with Hába (2016) findings.

6.2 RW, ARIMA models and BDS test

Random walk is widely used as a benchmark for exchange rate forecasting. It was used in this thesis as well. Two of our forecasting measurement statistics to compare different models were TheilU and Diebold-Mariano test. In both the random walk is used as a benchmark series.

BDS test for testing nonlinearity was performed on residuals when we fitted the returns series using ARMA models. It supported one of our initial hypothesis that there are nonlinearities in the CZK/EUR time series. This finding opened the space for the use of threshold models. It complies with Chappel *et al.* (1996). They applied a different testing procedure using McLeod-Li test. The nonlinerity was present in the exchange rate time series of French Franc/Deutschmark, two developed countries, which at that time were even under Exchange Rate Mechanism. The nonlinearity is found also in some other studies in emerging countries such as Kadilar *et al.* (2009) or Fahimifard *et al.* (2009) to name a few. They also pointed out that it is quite common to observe it in many exchange rate time series as there are many determinants making the behavior of exchange rate nonlinear and volatile.

We found out that in case of data before CNB decided to use interventions in 2013 the discovered nonlinearity could be well handled using hybrid ARMA/GARCH model. This was not the case regarding data during the recent interventions as there were no ARCH effects detected in residuals based on Lagrange multiplier test.

The forecasting performance of ARMA(1,1) and ARMA(1,3) in periods before and during interventions was in all cases poor and comparable with RW in terms of RMSE and TheilU. Diebold-Mariano test rejected the null hypothesis of same forecasting performance only in case of rolling samples and ARMA(1,3) model during interventions. In this case one sided alternative of Diebold-Mariano test even preferred RW over ARMA(1,3). All these results comply with former findings, for instance Meese & Rogoff (1983) and Chappel *et al.* (1996).

The last measurement - the success ratio - exposed that the best result were achieved in the period before interventions using recursive samples when ARMA(1,1) model could successfully forecast the appreciation or depreciation in about 60% of cases. Using rolling samples the achieved success ratio was slightly above 50%. At first sight it might seem satisfactory, on the other hand tossing a coin would return comparative results and in this manner this is not satisfactory at all.

The recursive samples offered better results over rolling samples in all instances.

6.3 SETAR model

In compliance with Kräger & Kugler (1993) and Chappel *et al.* (1996) formulation we analyzed models with one and two thresholds. When fitting procedure was finished we also performed corresponding tests to check whether our models captured the nonlinear structure. Neither before interventions nor during them any of our SETAR models performed well. BDS test revealed remaining nonlinear structure in all our models' residuals.

In case of data before interventions we found ARCH effects in residuals using Lagrange multiplier test. This was not the case of data during interventions and therefore the nonlinear structure there had a different nature.

The forecasting performance of all SETAR models in both periods was very bad. This is in contrast to Chappel *et al.* (1996) who reported superiority of SETAR forecasts over the linear models representative and random walk in their study. Nevertheless we must admit that these results are complicatedly comparable. Firstly we used weekly data compared to daily data in previous study and secondly they used just the logarithmic exchange rate series itself, which was de facto stationary in their case. We had to use the series of returns to fulfill stationarity condition.

On the other hand our findings are supported by Boero & Marrocu (2002). Similarly to our study they used weekly data and proved that Diebold-Mariano test did not show significant forecast gains for the nonlinear models over the linear benchmark. Their study was performed on the exchange rate data of developed countries.

There was not a case when any of our SETAR models performed better than

RW in terms of RMSE and TheilU. However Diebold-Mariano test did not reject its null hypothesis of same forecasting performance of SETAR against RW in any case.

We reached higher than 50% success in predicting the exchange rate moves in all cases except for one threshold SETAR and rolling samples during interventions. The success ratio of SETAR models was always more successful than success ratio of ARMA models. More satisfactory results we accomplished using the success ratio and recursive samples.

Overall our initial hypothesis that threshold models outperform RW in forecasting exchange rate was not confirmed in case of SETAR models. They do not even outperform a linear ARMA model in most of the measures.

6.4 TAR models

The analysis of TAR models starts with the specification of the threshold variable. In Chapter 3 we discovered the unusual mean and variance of the exchange rate changes in weeks when CNB Board had a meeting. We decided to implement this finding in our first model where the threshold variable was a dummy variable representing CNB Board meetings. In our second model we worked with a volatility as a threshold variable. Volatility itself is a subject of study regarding many financial time series including exchange rate time series as well (for instance Stancık (2007) or Frömmel *et al.* (2007)). We employed one of the common proxies of volatility - squared returns. We used them in a form of average squared returns over the specified last periods.

Similarly to SETAR models BDS test rejected its null hypothesis suggesting remaining nonlinear structure in residuals in both periods. Lagrange multiplier test unveiled ARCH effects in residuals in the period before interventions but not then in the period during interventions.

The performance results of our TAR models were mixed especially when focusing on periods before and during interventions. Surprisingly good results were achieved in the period before interventions. Both TAR models, no matter what threshold variable we used, outperformed RW in all cases in terms of RMSE and TheilU.

Despite the performance of the TAR models was better than that of RW, the improvement was not significant according to Diebold-Mariano test. Hence we cannot conclude TAR models had better forecasting accuracy using this measurement. The success ratio showed better results in case of rolling samples.

It reachad success in more than 60% of cases and provided better results than ARIMA and SETAR models. When using recursive samples success ratio had comparable results with SETAR models.

During interventions the performance of both TAR models was also surprising but in other sense. Their performance was in all cases distinctly worse than RW. Moreover Diebold-Mariano test favored RW in all cases.

For possible reasons let us at first recall (in case of CNB Board meetings dummy variable) the findings of Fišer & Horvath (2010) and Hába (2016). They showed a limited influence of CNB Board meetings on exchange rate during interventions compared to the period before them. We can assume based on CNB Board decision minutes and market behavior that during interventions CNB Board meetings did not produce no new information - the repo rate changes were unlikely and the change of currency pledge rather unrealistic. Therefore we could suppose that when TAR switching mechanism was activated the parameters would be only slightly different to the regime in weeks without CNB Board meeting. But that was not our case. This complies with Ahmad (2008) who pointed out that in regimes with limited number of observations small sample biases exist and lead to failure in identification of the true value of parameters in the data.

Obviously there might be several other reasons why we achieved such poor performance - badly chosen threshold variable or absence of threshold effects and thus misspecification of our models to name a few. The nonlinearity could be of different nature where threshold models do not provide the required improvement.

Chapter 7

Conclusion

This thesis explores different univariate threshold methods used for estimation and forecasting the exchange rate of CZK relative to EUR. The methods include SETAR models with one and two threshold values and TAR models with two different threshold variables - dummy of CNB Board meetings and average volatility over recent periods. We divided the data to two subsamples - 1999 to October, 2013 and November, 2013 to April, 2016 - to reflect the interventions which CNB started to use at the beginning of November, 2013.

The forecasting results were compared using RMSE, TheilU and Diebold Mariano test. In latter two we used a random walk as benchmark. Additional measurement we implemented was a success ratio to assess the predictability of appreciations and depreciations of our models.

In terms of RMSE, TheilU and Diebold Mariano test neither of SETAR models provided better results than random walk in any period.

TAR models provided surprisingly good results in the period before interventions where they outperformed random walk in all cases in terms of RMSE and TheilU. Nevertheless Diebold-Mariano test did not confirm their forecasting superiority. During interventions TAR completely failed in their forecasting ability. Success ratio results of SETAR and TAR models was higher in all cases before interventions, but provided mixed results during them.

Overall, the conclusion drawn from our analysis supports the general belief of exchange rates being difficult to forecast and that it holds in case of Czech koruna as well.

The thesis contributes to the area of exchange rate forecasting. Specifically it introduces the idea of TAR regime switching based on dummy threshold variable representing the Czech National Bank Board meetings. The second contribution lies in the use of volatility as possible threshold variable with the influence on exchange rates. It complements to traditional GARCH models.

Possible extension to this thesis regarding univariate threshold models lies for instance in choosing different threshold variables with possible influence on our original series. The analysis of type of nonlinearity could be also beneficial. Natural proposition is the use of multivariate threshold models therefore we provided the possible methodology in Chapter 4.

Bibliography

Ahmad, Y. S. (2008): "The effects of small sample bias in threshold autoregressive models." *Economics Letters* **101(1)**: pp. 6–8.

Ardic, O. P., O. Ergin, & B. G. Senol (2008): "Exchange rate forecasting: Evidence from the emerging central and eastern european economies." *Technical Report 7505*, Munich Personal RePEc Archive.

Baghestani, H. & L. Danila (2014): "Interest rate and exchange rate forecasting in the czech republic: Do analysts know better than a random walk?" *Czech Journal of Economics and Finance* **64(4)**: pp. 282–295.

Boero, G. & E. Marrocu (2002): "The performance of non-linear exchange rate models: a forecasting comparison." *Journal of Forecasting* **21(7)**: pp. 513–542.

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity." *Journal of Econo* **31**: pp. 307–327.

Brock, W., J. A. Scheinkman, W. D. Dechert, & B. LeBaron (1996): "A test for independence based on the correlation dimension." *Econometric reviews* **15(3)**: pp. 197–235.

Chan, K. S. (1993): "Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model." *The Annals of Statistics* **21(1)**: pp. 520–533.

Chappel, D., J. Padmore, P. Mistry, & C. Ellis (1996): "A threshold model for the french franc/deutschmark exchange rate." *Journal of Fore* **15**: pp. 155–164.

Clark, T. E. & K. D. West (2006): "Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis." *Journal of Econometrics* **135(1-2)**: pp. 155–186.

Clements, M. P. & J. Smith (1997): "The performance of alternative forecasting methods for setar models." *International Journal of Forecasting* **13(4)**: pp. 463–475.

Clements, M. P. & J. Smith (2001): "Evaluating forecasts from setar models of exchange rates." *Journal of International Money and Finance* **20**: pp. 133–148.

Cuaresma, J. C., B. Égert, & R. MacDonald (2005): "Non-linear exchange rate dynamics in target zones: A bumpy road towards a honeymoon." *Working Paper 771*, William Davidson Institute.

Cuaresma, J. C. & J. Hlouskova (2004): "Forecasting exchange rates in transition economies: A comparison of multivariate time series models." *Empirical Economics* **29**: pp. 787–801.

Cuaresma, J. C. & J. Hlouskova (2005): "Beating the random walk in central and eastern europe." *Journal of Forecasting* **24**: pp. 189–201.

Diebold, F. X. & R. S. Mariano (1995): "Comparing predictive accuracy." *Journal of Business & Economic Statistics* **13(3)**: pp. 253–263.

Evans, M. D. & R. K. Lyons (2005): "Meese-Rogoff Redux: Micro-based exchange rate forecasting." *Technical Report 11042*, National Bureau of Economic Research.

Fahimifard, S., M. Homayounifar, M. Sabouhi, & A. Moghaddamnia (2009): "Comparison of anfis, ann, garch and arima techniques to exchange rate forecasting." *Journal of Applied Sciences* **9(20)**: pp. 3641–3651.

Fišer, R. & R. Horvath (2010): "Central bank communication and exchange rate volatility: a garch analysis." *Macroeconomics and Finance in Emerging Market Economies* **3(1)**: pp. 25–31.

Franses, H. P. & D. vn Dijk (2000): *Nonlinear Time Series Models in Empirical Finance*. Cambridge University Press.

Frömmel, M. *et al.* (2007): "Volatility regimes in central and eastern european countries' exchange rates." *Faculteit Economie en Bedrijfskunde Working Paper* **487**.

Gonzalo, J. & J.-Y. Pitarakis (2002): "Estimation and model selection based inference in single and multiple threshold models." *Journal of Econometrics* **110(2)**: pp. 319–352.

Hansen, B. E. (2011): "Threshold autoregression in economic." *Statistics and Its Interface* **4**: pp. 123–127.

Hába, S. (2016): "Verbální intervence Čnb: Reaguje devizový kurz na slova bankovní rady?" *Politická ekonomie* **04**: pp. 405–419.

Hyndman, R. J. & G. Athanasopoulos (2013): *Forecasting: principles and practice*. OTexts.

Kadilar, C., M. Simsek, C. H. Aladag *et al.* (2009): "Forecasting the exchange rate series with ann: the case of turkey." *Istanbul University Econometrics and Statistics e-Journal* **9(1)**: pp. 17–29.

Kräger, H. & P. Kugler (1993): "Non-linearities in foreign exchange markets : a different perspective." *Journal of International Money and Finance* **12**: pp. 195–208.

Kunst, R. M. (2003): "Testing for relative predictive accuracy: A critical viewpoint." *Technical report*, Reihe Ökonomie/Economics Series, Institute for Advanced Studies, Vienna.

Meese, R. A. & K. Rogoff (1983): "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?" *Journal of International Economics* **14**: pp. 3–24.

Monetary and Economic Department (2013): "Foreign exchange turnover in april 2013: preliminary global results." *Technical report*, Bank for International Settlements.

Mućk, J. & P. Skrzypczyński (2012): "Can we beat the random walk in forecasting cee exchange rates?" *Working paper 127*, National Bank of Poland.

Naszódi, A. (2011): "Beating the random walk in central and eastern europe by survey forecasts." *Working Paper 2011/3*, Magyar Nemzeti Bank, The Central Bank of Hungary, Budapest.

Stancik, J. (2007): "Determinants of exchange rate volatility: The case of the new eu members." *Czech Journal of Economics and Finance* **57**.

Taušer, J. & P. Buryan (2011): "Exchange rate predictions in international financial management by enhanced gmdh algorithm." *Prague Economic Papers* **3**: pp. 232–249.

Tong, H. (1990): *Non-linear time series: a dynamical system approach*. Oxford University Press.

Tong, H. & K. S. Lim (1980): "Threshold autoregression, limit cycles and cyclical data." *Journal of the Royal Statistical Society* **42(3)**: pp. 245–292.

Tsay (1998): "Testing and modeling multivariate threshold models." *Journal of the American Statistical Association* **93(443)**: pp. 1188–1202.

Tsay, R. (1989): "Testing and modeling threshold autoregressive processes." *Journal of the American Statistical Association* **84(405)**: pp. 231–240.

Tsay, R. S. (2002): *Analysis of Financial Time Series*. John Wiley & Sons, Inc.

West, K. D. (1996): "Asymptotic inference about predictive ability." *Econometrica* **64(5)**: pp. 1067–1084.

Zivot, E. & J. Wang (2007): *Modeling Financial Time Series with S-PLUS, Second Edition*, volume 191. Springer Science & Business Media.

Appendix A

Data, source codes and supportive materials

All materials used while working on the thesis are available to view and download at the following link: `https://goo.gl/5HVPYo`