Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Exchange Rate Forecasting: An Application with Model Averaging Techniques

Author: **Bc. Jaroslav Mida**

Supervisor: **doc. Roman Horváth Ph.D.**

Academic Year: **2014/2015**

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, April 30, 2015

Signature

Acknowledgments

I would like to give my thanks to my supervisor, doc. Roman Horváth Ph.D., who helped in the process of choosing the topic for my master's thesis as well as during the writing. His remarks and advice were invaluable for me. I would also like to offer my gratitude to my family and friends, who always cheered me up and supported me all the time.

Abstract

The exchange rate forecasting has been an interesting topic for a long time. Beating the random walk model has been the goal of many researchers, who applied various techniques and used various datasets. We tried to beat it using bayesian model averaging technique, which pools a large amount of models and the final forecast is the average of forecasts of these models. We used quarterly data from 1980 to 2013 and attempted to predict the value of exchange rate return of five currency pairs. The novelty was the fact that none of these currency pairs included U.S. Dollar. The forecasting horizon was one, two, four and eight quarters. In addition to random walk, we also compared our results to historical average return model using several benchmarks, such as root mean squared error, mean absolute error or direction of change statistic. We found out that bayesian model averaging can not generally outperform random walk or historical average return, but in specific setting it can produce forecasts with low error and with high percentage of correctly predicted signs of change.

JEL Classification	C5, C11, C12, E5, F31
Keywords	exchange rate forecasting, Bayesian model averaging, random walk, historical average return

Author's e-mail	`jaro.mida@gmail.com`
Supervisor's e-mail	`roman.horvath@fsv.cuni.cz`

Abstrakt

Predpovedanie výmenných kurzov vždy bolo zaujímavou témou. Cieľom mnohých akademikov bolo predpovedať vývoj hodnoty výmenného kurzu s menšou chybou ako náhodná predpoveď. Títo akadamici využili vo svojich prácach rozmanité techniky a datasety. V tejto práci sme použili techniku Bayesovho priemerovania modelu, kde konečná predpoveď je priemer predpovedí všetkých modelov. Aplikovali sme štvrťročné dáta od roku 1980 do 2013 a pokúsili sme sa odhadnúť hodnoty výnosov výmenných kurzov piatich menových párov, ktoré neobsahujú americký dolár. Predikcia bola vykonaná na horizonte jedného, dvoch, štyroch a ôsmych kvartálov. Výsledné predpovede sme okrem náhodnej predpovedi porovnali aj s priemerným historickým výnosom pomocou niekoľkých kriterií, ako napríklad stredná kvadratická chyba, stredná absolútna odchýlka alebo hodnota smeru zmeny. Zistili sme, že Bayesove priemerovanie modelov zvyčajne neporazí náhodnú predpoveď alebo priemerný historický výnos. Na druhej strane, v niektorých špeciálnych situáciach táto metóda dokáže predpovedať s menšou chybou a s vyšším percentom správne predpokladaných zmien znamienka.

Klasifikace JEL	C5, C11, C12, E5, F31
Klíčová slova	predikcia výmenného kurzu, Bayesovo priemerovanie modelov, náhodná predpoveď, priemerný historický výnos

E-mail autora	jaro.mida@gmail.com
E-mail vedoucího práce	roman.horvath@fsv.cuni.cz

Contents

List of Tables

List of Figures

Acronyms

BMA Bayesian Model Averaging

RW Random Walk

HAR Historical Average Return

BMW Bayesian Model Winner

VAR Vector Autoregressive Model

BMAP Bayesian Model Averaging Using Predictive Likelihood

RMSE Root Mean Squared Error

MAE Mean Absolute Error

Q1 One Quarter

Q2 Two Quarters

Q4 Four Quarters

Q8 Eight Quarters

Master's Thesis Proposal

Author	Bc. Jaroslav Mida
Supervisor	doc. Roman Horváth Ph.D.
Proposed topic	Exchange Rate Forecasting: An Application with Model Averaging Techniques

Topic characteristics Exchange rate predictability is a very important issue, especially for policymakers and central banks. It is monitored on regular basis for macroeconomic reasons and market surveillance purposes. Wieland & Wolters (2012) presented the usage of forecasts in assessing the outcome and effect of a particular policy measure on the targets the policymakers want to achieve. Then the decisions are taken based on what the policymakers believe is the most likely scenario. Exchange rate forecasts are especially important for countries whose economies heavily depend on imports and exports, because exchange rate plays a huge role in trade environment.

The collapse of Bretton Woods system marked a new era in exchange rate regimes. Fixed exchange rates among major industrial countries were abandoned and they started to use floating exchange rate regime. Since then there has been a considerable effort in forecasting the exchange rate movements. Mussa (1979) concludes that the spot exchange rate approximately follows random walk, meaning that the changes are in fact unpredictable. Meese & Rogoff (1983) come to conclusion that random walk without drift outperforms several structural models, suggesting unpredictability of fluctuations. And although few years later, Mark (1995) or Mark & Sul (2001) provide some empirical evidence that exchange rates can be forecasted, at least on longer horizons, the prevailing opinion on this topic has been that exchange rate fluctuations are not predictable.

Exchange rates are not the only data that have been difficult to predict, for example, Atkeson and Ohanian (2001) show that forecasts of inflation based on Phillips curve give high prediction errors. Stock and Watson (2001, 2004) examine forecast of inflation and output growth and come to conclusion that naďve time series forecast provide smaller root mean squared error than the models they considered. Therefore, new methodologies are being invented and used in order to try to solve this problem. According to Liu & Mahieu (2010), the method which has been recently applied in out-of-sample forecasts of growth, stock return or exchange rate with partial success, is the Bayesian Model Averaging (BMA) technique. For example, Wright (2008) shows that BMA forecasts of chosen exchange rates do sometimes better than driftless random walk (RW) and never much worse. Liu & Mahieu (2010) present results consistent with Wright (2008) that BMA slightly outperforms RW.

Hypotheses 1. Hypothesis: The BMA model predicts the future exchange rate value better than driftless RW in medium to long term horizons.

2. Hypothesis: The BMA model predicts the future exchange rate value worse than driftless RW in short term horizons.

3. Hypothesis: The proportion of times, the BMA model predicts the sign of the change correctly, is more than 0.5.

Methodology The first necessary thing to do will be to choose the currency pairs to be forecasted and the predictors to be used in the analysis. I will focus on currency crosses, such as EUR/YEN, GBP/YEN or EUR/GBP. The choice of variables will follow Wright (2008), but I will only consider quarterly data for all the variables. To collect the data, I will use various datasources, e.g. Statistical Data Warehouse of ECB, Office for National Statistics of UK, OECD database etc. In addition, I will transform the data according to previous literature, split my dataset into two groups and use the out-of-sample forecasting technique.

The choice of the method used is very straightforward. BMA technique, originally presented by Leamer (1978), addresses the problem of model uncertainty.

According to Hoeting et al. (1999), it provides better out-of-sample performance than a method with only one single model. Madigan & Raftery (1994) state that BMA shows better average predictive ability than any single model. The basic idea behind the model is to consider a set of possible models (I will consider only linear models) and of those one is the true model, but we do not know which one. We have some prior beliefs about the probability that the i-th model is the true one. Firstly, I will set the probability equal to 0.5, which equally supports all the models, as stated by Wright (2008). Then, I will experiment and set the prior probabilities to be smaller, which suggests prior support to lower dimensional models, as explained by Tortora (2010). I will then compute the posterior probabilities and use them to weight each of the forecasts. My baseline model, against which I will compare the BMA, will be driftless RW model.

The general view on the exchange rate predictability is that it is very hard, if not impossible to predict on short horizons and less difficult on longer horizons. Therefore, I expect BMA to do better than RW on longer and worse on shorter horizons. To test it, I will compute the Root Mean Squared Forecast Errors and test the null that the difference is zero against alternative that it is negative, using t-test, as suggested by Rossi (2013). To assess the third hypothesis, I will follow Wright (2008) and construct the percentage of times BMA predicts the sign correctly for different time horizons.

Outline

1. Introduction
2. Literature review
3. Dataset
4. Methodology
5. Results
6. Conclusion

After introducing the topic, we will have a look at the previous research done on this issue. We will continue with the description of variables and currency pairs, which will be used in the thesis. Next part we shall describe the theoretical background behind the model, benchmarks etc. We will present and discuss the results before we conclude.

Core bibliography

1. Atkeson, A. & Lee E. Ohanian (2001): "Are Phillips curves useful for forecasting inflation?" *Federal Reserve Bank of Minneapolis Quarterly Review.*

2. Della Cortea, P. & I. Tsiakasa (2007): "An economic evaluation of empirical exchange rate models: Robust evidence of predictability and volatility timing." *CEPR Working Paper.*

3. Hoeting, Jennifer A. & et al. (1999): "Bayesian model averaging: A tutorial." *Statistical science* : pp. 382–401.

4. Leamer, Edward E. (1978): "Specification searches: Ad hoc inference with nonexperimental data." *New York: Wiley.*

5. Liu, Y. & R. J. Mahieu (2010): "Exchange Rate Predictability: Bayesian Model Selection."

6. Madigan, D. & Adrian E. Raftery (1994) : "Model selection and accounting for model uncertainty in graphical models using Occam's window." *Journal of the American Statistical Association* pp. 1535–1546.

7. Mark, Nelson C. (1995): *"Exchange rates and fundamentals: Evidence on long-horizon predictability." The American Economic Review* pp. 201–218.

8. Mark, Nelson C. & D. Sul (2001) : "Nominal exchange rates and monetary fundamentals: evidence from a small post-Bretton Woods panel." *Journal of International Economics* **53(1)**: pp. 29–52.

9. Meese, Richard A. & K. Rogoff (1983) : "Empirical exchange rate models of the seventies: do they fit out of sample?" *Journal of International Economics* **14(1)**: pp. 3–24.

10. Mussa, M. (1979) : "Empirical regularities in the behavior of exchange rates and theories of the foreign exchange market." *Carnegie-Rochester Conference Series on Public Policy, North-Holland* pp. 9–57.

11. Rossi, B. (2003) : "Exchange rate predictability." *Journal of Economic Literature* **51(4)**: pp. 1063–1119.

12. Stock, James H. & Mark W. Watson (2001) : "Forecasting output and inflation: the role of asset prices." *National Bureau of Economic Research.*

13. Stock, James H. & Mark W. Watson (2004) : "Combination forecasts of output growth in a seven?country data set." *National Bureau of Economic Research* **23(6)**: pp. 405–430.

14. Tortora, Andrea D. (2010) : "Exchange Rate Forecasting: Bayesian Model Averaging and Structural Instability."

15. Wieland, V. & Maik H. Wolters (2012) : "Forecasting and policy making." *Handbook of economic forecasting.*

16. Wright, Jonathan H. (2008) : "Bayesian model averaging and exchange rate forecasts." *Journal of Econometrics* **146(2)**: pp. 329–341.

Author

Supervisor

Chapter 1

Introduction

Exchange rate is a very important concept in today's world. As the time passed by, the countries started to become more open to the world. Various barriers were abolished and international trade and foreign investment began to play the main role in countries' economies. Nowadays, countries trade heavily with each other, manufacturers look for suppliers, who can deliver necessary goods/services for a cheaper price, investors look for profitable opportunities even outside their domestic country and so on. As a consequence, the exchange rate and exchange rate risk, which is a risk of large depreciation of currency, rose to prominence, because to acquire a foreign good or invest into foreign assets, one has to usually exchange his domestic currency for foreign currency. Therefore, monitoring the exchange rate development became crucial in achieving favorable prices or profit.

The floating regime, which has been adopted in many developed countries and is one of the most common currency regimes out there, was not always so popular. Its popularity started to gain momentum around the time, when Bretton Woods system collapsed. This was the agreement, which stated that currencies were pegged against the U.S. Dollar, which was backed by gold. It had some success, as it was able to increase the amount of international trade and stabilize some of the economies. Eventually, it collapsed and created space for free floating regime. That was the moment the exchange rate predictability became a hot topic among researchers.

The core paper at that time, which addressed the issue of exchange rate predictability, was written by Meese & Rogoff (1983). They showed how difficult it actually is to forecast exchange rates and, in process, discouraged many others from trying to do so, because they showed that exchange rates can be approximated by Random Walk. However, as already mentioned, exchange rate is very important for a wide range of institutions and people, for whom it was and still is crucial to predict the exchange rate changes. Therefore, new methods and techniques were developed, which in specific setting, were able to outperform Random Walk.

One of the newest method applied to the problem of forecasting in general was the Bayesian Model Averaging. At the beginning it was mainly used to carry out the forecast of, e.g. growth rates, but later on, some researchers realized its potential in exchange rate forecasting as well. The issue in any forecasting exercise is to correctly choose the model, right number of variables etc. The BMA method tries to solve this issue of uncertainty, because it allows all possible models to be the true ones and then averages over the forecasts of these models based on model prior probability. The first one to actually apply this methodology on exchange rates was Wright (2008) and his promising findings encouraged others to use BMA with/without some smaller modifications to predict the movements.

However, what most of the researchers did, was that they attempted to forecast currency pairs vis-*à*-vis U.S. Dollar, i.e. the most traded and liquid currency pairs. But there are so many more currency pairs, which are also liquid and often traded, although not in such a large amount as U.S. Dollar. Therefore, in this study, the goal is to examine the predictive power of BMA method as well as slightly modified BMA called Bayesian Model Winner method, for the currency crosses, i.e. pairs not involving U.S. Dollar, such as EUR/GBP or AUD/JPY. They have some specific features, which can either be helpful in predicting the changes or might worsen our chances to forecast with low error. We will use a large amount of variables, with quarterly data ranging from 1980 to 2013, to perform the forecasts.

The forecasts will be done using several different settings of BMA. Firstly, we will predict the changes on four different time horizons, namely one, two, four and eight quarter ahead, so that we can see the predictive power on short, medium and long horizons. We will always keep some part of the data aside, so that we can conduct out-of-sample forecast and compare predictions with reality. Due to the fact that previous evidence shows that models with smaller amount of variables have better predictive power, we will implement this feature into our estimation as well.

The most important part of this study is to compare forecasts from BMA and slightly modified BMA called Bayesian Model Winner method with Random Walk and, ideally, outperform it. In addition, we will also follow previous literature and use Historical Average Return model as another baseline statistic to evaluate forecast performance. To carry out the comparison one must choose benchmarks. There is quite a wide range of benchmarks previously applied, and we chose three of them. Root Mean Squared Error and Mean Absolute Value benchmarks look at the forecasting error of model versus baseline model, whereas Direction of Change looks at the percentage of times the model is able to correctly predict the sign of the change. Using more benchmarks is helpful, as then we can come to conclusions with more certainty.

The thesis is structured as follows: Chapter 1 introduces the reader into the problem of exchange rate forecasting. Chapter 2 is the literature review, where reader can read about past studies conducted on this topic and which is divided into three parts. Chapter 3 explains the development of chosen currency pairs over the period 1980-2013 and provides description of variables used. Chapter 4 explains the theory behind BMA technique and benchmarks as well as explains how they are applied in this particular study. Chapter 5 includes the discussion of results and tables with actual results. Chapter 6 concludes.

Chapter 2

Literature review

The topic of exchange rates predictability has been discussed and analysed in many studies. Various researchers used a wide range of different models combined with different datasets and explanatory variables. The general opinion was at the beginning that it is very difficult to beat the random walk without drift model, which basically means that exchange rates are unpredictable. Therefore, many tried to prove otherwise. However, the results of these studies and attempts are very mixed, some confirming the unpredictability, whereas the others usually reporting partial success. In this chapter we will have a look at past studies, which have been carried out and what evidence they brought into this study field.

2.1 Exchange rate is not predictable

One of the first researchers to look into this topic was Mussa (1979). In his study he examined the empirical regularities in the behaviour of exchange rates and theories of the forex. Analysing the exchange rates of U.S. Dollar against major currencies, he stated that under condition that exchange rates are not controlled by interventions, the natural logarithm of the spot exchange rate follows random walk without drift model. In addition to this regularity, also described as the stochastic behaviour, he also examined the standard flow market model, which was very popular in theory. However, his findings were that this model was not useful in explaining the exchange rates behaviour as well as the behaviour of other related variables.

The core paper at that time written on the topic of exchange rate forecasting was the one by Meese & Rogoff (1983). They collected monthly, seasonally unadjusted data over period 1973 to 1981 and compared the out-of-sample forecast of structural models, time series models and random walk without drift model. The chosen structural models were flexible-price model by Frenkel-Bilson, sticky-price monetary models by Dornbusch-Frankel and stick-price model, which included current account, by Hooper-Morton. They had three currency pairs, U.S. dollar against British pound, Japanese yen, German mark and, in addition, the trade-weighted dollar exchange rate. The forecasting horizon was from one up to twelve months. The conclusion was that RW model performed no worse than the time series and structural models. The poor performance of structural models was especially surprising, because they based the forecasts of these models on actual realised values of future explanatory variables. These results showed that exchange rates could be approximated using the RW without drift model, i.e. it is very difficult to forecast them.

The results from Meese & Rogoff (1983) discouraged others from analysing this area of economics for a while. As Frankel & Rose (1995) pointed out in their survey, which was included as a chapter in Handbook of International Economics edited by Jones *et al.* (1997), these negative results had "pessimistic effect on the field of empirical exchange rate modeling in particular and international finance in general" (p. 1704). This negative view was further confirmed by Berkowitz & Giorgianni (1996). They decided to use the methodology according to the Monte Carlo study and applied it on historical data with the goal to try to predict the exchange rate movements in U.S. dollar exchange rates. What they found out was that fundamentals did not help in out-of- or in-sample prediction as the forecasting horizon widened.

Later on, in 2000s, there already were some studies, which we will discuss later on in this chapter, which showed that some models could outperform RW model, at least in specific conditions. However, very good point was made by Sarno & Taylor (2002) in their book. They emphasized that although there already existed several models, there still did not exist models, which would be sufficiently reliable, satisfactory and robust when applied on different exchange rates, explanatory variables, datasets etc. They said that some models performed quite well in in-sample forecast, but failed terribly in out-of-sample forecast. On the other hand, there were some which had good out-of-sample forecasting accuracy, but when applied on different currencies/horizons, the satisfactory results could not be replicated.

An interesting study, which partially confirmed the unpredictability of the exchange rates, only on short horizons up to one year, was conducted by Kilian & Taylor (2003). The difference to previous studies was that they considered empirical evidence of nonlinear relationship between exchange rates and fundamentals, incorporated it into their model and analysed whether this nonlinearity can explain why it had been so difficult in the past to predict with low prediction error. They used exponential smooth transition autoregressive models and applied it to the dataset consisting of seven countries with quarterly data during the period after the collapse of Bretton Woods system. They found out that real exchange rate can be approximated by RW close to the equilibrium. This could help explain the success of RW when predicting the nominal exchange rate, especially at shorter horizons.

Very similar results were also presented by Cuaresma & Hlouskova (2005). They decided to compare the forecasting accuracy of vector autoregressive model, restricted VAR, Bayesian VAR, vector error correction and Bayesian vector correction models compared to RW. They chose exchange rates of five countries from Central and Eastern Europe, namely Slovakia, Czech republic, Hungary, Poland and Slovenia, against U.S. dollar and the Euro. The findings were that none of the above mentioned models was able to outperform RW model at shorter horizons. Moreover, there was another paper written by Muck & Skrzypczynski (2012), which also examined the exchange rate predictability of three countries, Czech republic, Hungary and Poland, against the Euro. Using the fractionally integrated RW and a variety of VAR-type models, they concluded that it is very hard to beat the RW model, as none of the chosen, more complex models, were able to outperform RW consistently.

Van Wincoop & Bacchetta (2003) introduced another new element into the exchange rate forecasting framework. Considering studies, which showed that exchange rate volatility is related to order flows at shorter horizons, they decided to introduce investor heterogeneity into the basic monetary model of exchange rate determination. There were two types of heterogeneity: dispersed information about fundamentals and non-fundamentals based heterogeneity. The results of their model were in line with the past evidence on this topic.

It confirmed that fundamentals play no significant role in forecasting exchange rates movements at short to medium horizons.

Another researcher, who extended previous literature was Yuan (2011). He proposed a different model compared to previous studies. This model was combining the multi-state Markov-switching model with smoothing techniques. He based his paper on the fact that exchange rates are likely to follow highly persistent trends. He found out two important things and although he actually presented results that exchange rates can be forecastable, which will be discussed in the next section, he also confirmed that fundamentals-based linear models in most cases are not able to capture the persistence in exchange rates. According to him, this is the reason why RW outperforms these models.

2.2 Exchange rate is predictable

The previous section offered only one side of the coin. Although it is clearly very difficult to forecast exchange rates with high accuracy, it is not impossible. There exists numerous studies and papers since the famous Meese & Rogoff (1983) paper, which showed that exchange rates can be predicted, especially at longer horizons. On of the first researchers to look into this topic was Hakkio (1986). Based on claims from several authors that nominal and real exchange rates are unpredictable, he decided to look into the conditions under which they actually do follow RW. The results, however, presented him with a puzzle. The evidence was mixed, some of it confirmed the theory of exchange rates following RW, some of it, however, showed otherwise. What he did next was the analysis of four different tests for RW. The findings were that these tests have low power. As a consequence, even if we can not reject the hypothesis that exchange rate follows RW, it is not correct to make conclusions, because of the evidence he presented.

Hakkio (1986) showed some uncertainty and helped to heat up interest in this topic. Mark (1995) in his paper provided the evidence that there exists an economically significant predictable component in changes in log exchange rates at longer horizons. The important thing was that there is a lot of noise at short horizons. However, at longer horizons, it is averaged out and the movements of exchange rates start to be systematic, and therefore forecastable using macroeconomic fundamentals. In his forecast, he was able to outperform the driftless RW in three out of four examined exchange rates at longer horizons, which was a breakout in this economic field, because the general view still was that it is impossible to predict exchange rates.

The study by Mark (1995) was an encouragement for further studies. One, carried out by Mark & Sul (2001), analysed the relationship between nominal exchange rates and monetary fundamentals. The data used were from 1973 to 1997, collected quarterly and from 19 different countries. Using the panel regression and panel-based forecasts, they showed that nominal exchange rate is co-integrated with fundamentals and they have significant forecasting power for future movements of exchange rates.

Many studies about the predictability of exchange rates came up with mixed results. On one hand, they showed that it is very difficult to beat RW at short horizons, e.g. up to one year. On the other hand, their results provided evidence of predictability at longer horizons. This also is the case of a few papers and studies mentioned in the Section 2.1. First of them is the paper by Kilian & Taylor (2003). In addition to what we already explained, he also recommended a new regression test of RW model - bootstrap long horizon test. His choice was also influenced by the fact that it was showed that tests for RW have low power. However, this test proved to be reliable and quite powerful. In the end, it provided evidence against RW at longer horizons (2-3 years), which meant that exchange rates can be predicted at longer horizons.

Next in this group of papers is the study by Cuaresma & Hlouskova (2005). They used a variety of multivariate time series models (different versions of vector autoregressive and vector error correction models) and in most cases, they found evidence that these models outperform the RW at longer horizons, in this case it was six months and more. Another one is the paper by Van Wincoop & Bacchetta (2003). Their special contribution was the addition of investors heterogeneity, which could be, according to them, a crucial element in understanding the dynamics of exchange rates. They presented three core results of their paper. For this section, the second and third results are important. Their model was able to show that exchange rate is actually influenced by fundamentals over longer horizons. However, changes in exchange rates have a weak predictive power in forecasting future fundamentals.

A researcher to propose a different approach to this problematic was Rossi (2006). She based her study on the fact already noticed by Meese & Rogoff (1988) that parameters are instable. She believed that this could be a reason why monetary models of exchange rate determination fail to outperform RW. She decided to incorporate a new tests for nested models that are robust to the issue of instability. The results were that she was able to, at least for some countries, reject the hypothesis of exchange rates being random walks. In addition, she estimated RW time-varying parameter model and a forecast combination method, which should increase the accuracy of forecasts in case we observe structural breaks. The findings were encouraging as the latter method was in fact able to improve forecasts relative to RW.

Engel *et al.* (2007) presented some very interesting points in their study. Firstly, due to the fact that many models actually imply that exchange rates can be approximated by RW, they think that we should not expect high forecasting power of various models. Secondly, they point out that the inability of these models to beat RW does not necessarily mean that wrong exchange rates models had been used. To support this, they applied panel techniques on monetary models and the resulting evidence was that these models generally performed better than RW when forecasting exchange rates, they had lower mean squared prediction error.

Another paper presenting positive findings about exchange rates predictability was written by Carriero *et al.* (2009). They incorporated a new approach into the forecasting framework. Due to the success of RW, it is safe to construct a model, in which exchange rates a priori follow RW. Moreover, it is important that this model takes into account information provided by a panel of exchange rates, in this case 33 of them vis-a-vis U.S. Dollar, because exchange rates are likely to co-move. Finally, they constructed a Bayesian vector autoregressive model and adopted RW prior. The results of the forecasts showed that their model was able to outperform RW for most of the countries, and even at shorter horizons, where large majority of models fails.

A different approach was applied by López-Suárez & Rodríguez-López (2011). They wanted to explain the failure to predict exchange rates by incorporating the nonlinear behaviour of them into the model. Using a Smooth transition error-correction model, panel dataset with 19 countries and three numéraires (the United States, Japan, Switzerland), they came up with evidence of out-of-sample nonlinear predictability of exchange rates. The forecast accuracy of this model was higher than RW model, even if applied on different horizons (shorter/longer) and using different numéraires. The problem, which they found was, that the robustness was limited - the model dominated, on average, RW only for specific forecast horizons and the horizons were different for different numéraires. However, despite this fact, they still obtained significant predictability gains.

We already discussed the study by Yuan (2011). He provided evidence of fundamentals having no predictive power for exchange rates. However, in addition, he also presented other important findings. His model, described in Section 2.1, was able to outperform RW at shorter horizons. Moreover, the results were also robust across different sample spans. The crucial fact he used was that RW usually fails to capture trends and this is its main weakness. As a result, he believed that if we identify these highly persistent trends, we should be capable of beating RW with relative ease.

2.3 Literature on Bayesian Model Averaging predictability power

In this thesis the focus will be put on the BMA technique and its forecasting accuracy. This technique is relatively new and we will have a closer look at the overall model and the origin of this method in the later chapters. The purpose of this section is to present a few studies that used BMA to predict exchange rates and have a look at their findings.

A core paper, which incorporated BMA into the exchange rates forecasting framework, was written by Wright (2008). He based his study on the fact that in recent times researchers tried to use methods which take into account large amount of information from large amount of time series and then simple average the forecasts of different models. This approach brought encouraging results, because it provided better out-of-sample prediction than a single model. As a consequence, he decided to apply BMA technique to pseudo-out-of-sample forecast of U.S. Dollar vis-a-vis Canadian Dollar, Yen, Euro/Mark and Pound, over ten years. His results confirmed that this method could be a suitable tool, because BMA did in some cases slightly better than RW (lower mean square prediction error), but never much worse. Another thing he presented was that BMA forecasts were very similar to RW. This is, however, not a bad thing, according to him, because there exists general scepticism about models whose predictions are not flatline/near flatline. This comes from the fact that exchange rates are actually very close to RW. He stated that such models, e.g. BMA, which pool information from large amount of indicators, could in some cases produce slightly better forecasts than flatline prediction.

Partly motivated by the work of Wright (2008), Lam *et al.* (2008) decided to re-examine not only the predictive power of BMA, but also three other models, namely Purchasing Power Parity model, Uncovered Interest Rate Parity model and Sticky Price Monetary model. In addition, they included in their analysis the combined forecast based on these four models. They used two baseline models - usual RW model and historical average return, and undertook the forecast on three major exchange rates (Euro/Mark, Pound, Yen) against U.S. Dollar from 1973 to 2007. The forecast horizons were one through four, and eight quarters. To assess the accuracy carefully, they decided to use various measures, such as root mean squared error, direction of change or t-statistic. However, BMA model did well only in some cases. It outperformed other models and combined forecast for EUR/USD at shorter horizons as well as in predicting the correct sign of the change (direction of change statistic). For GBP/USD currency pair, BMA was superior only for eight-quarters ahead forecast and for USD/JPY, BMA was not the best model at any horizons. This only confirmed findings by Cheung *et al.* (2005), who concluded that a model might do well for a certain currency pair, but not for other currency pairs. At the end, the combined forecast was the winner among the models as the root mean squared error ratios were lower relative to other models and it also outperformed them when looking at predicting the sign of the movement correctly.

Another study inspired by Wright (2008) was conducted by Tortora (2010). He carried out the analysis for two exchange rates, Yen and Canadian Dollar vs. U.S. Dollar and predicted the development for one-quarter ahead, i.e. for a short horizon. The new element added to BMA was the consideration of parameter instability, which he incorporated by using a mixture innovation approach.

This method should in theory reduce the uncertainty of a model and provide more flexibility. However, his findings suggested that this approach hardly improved the forecast compared to the prediction under constant parameters. This can be contributed, according to him, to the fact that parameters might contain extra estimation error or uncertainty. Eventually, this approach worked quite nicely for the USD/JPY, whereas the evidence for other currency pair was mixed - model outperformed benchmarks only in the subsample, which was studied by Wright (2008), i.e. finishing in 2005. At the end, Tortora (2010) concluded that we can get the best results by having constant parameters or by limiting number of breaks, meaning that the mixture innovation approach should be better than time-varying parameter model.

A little bit overall different approach was taken by Liu (2010). He decided to use BMA model, together with RW model and historical average return model, as baseline models. He considered two specific models for forecasting exchange rates, Bayesian Model Winner and Bayesian Model Averaging Using Predictive Likelihood. In the case of first model, BMA was able to outperform it in out-of-sample forecast, as model combination incorporated by BMA proved to be a better method than a single model, BMW. In addition, BMA was able to produce better predictions than RW when looking at forecasting errors. In case of the BMAP, BMA was not able to produce better results and was outperformed. BMAP was even able to beat RW in most cases, having lower mean squared prediction error. According to Liu (2010), the success of BMAP can be contributed to the fact that it gets rid of two problems BMA model faces - in-sample overfitting of data and the condition that true model is included among considered models.

Let us finish this chapter by introducing a very interesting study by Rossi (2013). In her article she considered a large variety of methods and fundamentals for forecasting exchange rates, e.g. Single equation linear models, Error correction model, Nonlinear models, BMA, Vector autoregressive models etc. She collected data for several currencies vis-a-vis U.S. Dollar as well as various economic fundamentals, such as overnight interest rates, 3-month Treasury Bills etc. The evidence in favour of BMA in her study, however, was very weak as BMA did not have significantly higher predictive power than RW for any chosen horizon or test statistic. However, during this complex analysis she came across numerous stylised facts, which then lead her to make five major conclusions. Firstly, traditional fundamentals have lower predictive more in out-of-sample forecasts than Taylor-rule and net foreign assets fundamentals. Secondly, the most successful model specifications were the linear ones. Thirdly, data transformation, e.g. seasonal adjustment, can have a lagre influence on the accuracy of prediction. Next, the choice of benchmarks, sample periods and evaluation method of prediction accuracy, is really important and can have a strong effect on final results. And lastly, conducted analysis confirmed some findings in the previous literature, but also rejected others, e.g. evidence showed that some models and fundamentals used in previous studies had actually lower out-of-sample predictive power.

This chapter provided us with various views on the exchange rate forecasting framework. This topic is still very hot and no consensus, except maybe the fact that exchange rates can be closely approximated by RW, has been found yet. Therefore, it is interesting to look at this topic from slightly different angle, by examining currency crosses, because as we have read, most of the studies focused on major U.S. Dollar currency pairs and it was already suggested that some methods work differently for different exchange rates.

Chapter 3

Data description

3.1 Currency pairs

Most of the studies, which analysed the predictability of the exchange rates, focused on major currency pairs, i.e. currency pairs vis-a-vis U.S. dollar. That was a pretty obvious choice as these currency pairs belong to the most traded and used ones in the foreign exchange market. However, there exists many currency pairs, so called crosses, which are not vis-a-vis U.S. dollar, but some other currency. To those belong for example, EUR/JPY, GBP/JPY, EUR/GBP, EUR/AUD or AUD/JPY, which will be the centre of our interest in this thesis, because these crosses are especially important when countries are performing trade or some financial transactions between each other. Then, they use these crosses to exchange money. Therefore, being able to predict the future development is crucial for them.

But why should the predictive performance differ from major currency pairs? There are several facts suggesting that it really might be different. Firstly, these pairs are not so strongly affected by development in the largest economy in the world, the United States. Situation in the U.S. definitely influences movements in these pairs, but the movements are not so strong compared to U.S. dollar pairs. Secondly, the trade volumes are smaller as less of these currencies is bought or sold in the foreign exchange market. In addition, some pairs, such as EUR/JPY and especially GBP/JPY, are quite volatile, not trending and popular among private forex traders. This could potentially mean that it is harder to predict them due to these strong, sometimes unexpected movements. On the other hand, we have EUR/GBP and AUD/JPY, pairs of countries whose economies are interlinked. This means that events happening in one country, either bad or good will have weaker effect on the overall exchange rate, because both countries will be affected.

Thanks to the reasons stated above, it is interesting for us to evaluate the potential for crosses to be forecasted. In this section we will have a closer look at the pairs we have chosen. Following sections in this chapter will address the choice of variables, their potential transformation and data sources.

3.1.1 EUR/JPY

EUR/JPY is the currency exchange rate of a Euro to Japanese Yen. This currency pair is very popular among private forex traders, because it is a very active and volatile pair that could produce large profits (but also losses) for traders. It is usually used for short-term trading strategies due to its volatility. It represents around 3% of all transactions completed daily on forex and as a result is ranked as the seventh most traded currency pair. Both currencies are largely affected by the monetary policy of European Central Bank and Bank of Japan respectively.

The Euro is the official currency of Eurozone countries. The trading of Euro started as soon as the virtual currency was established, i.e. in 1999. Since its introduction, it has become the world's second most popular reserve currency after the U.S. dollar. Recent crisis in the Euro area and political turmoil in the Ukraine could potentially weaken Euro's position in the financial world.

Yen is the official currency of Japan. Compared to the Euro, it is an old currency, being adopted in 1871 and used ever since. Yen is a very popular and often traded currency, accounting for around 17% of trading volume. Moreover, similar to Euro, it is a reserve currency. One specialty of Yen is that it is a safe haven currency, into which traders invest in worse times. In addition, Japanese Yen is heavily influenced by commodity prices, mainly oil prices. The reason is that Japan is a net importer of oil and as a consequence, oil prices have a strong affect on Yen value.

Looking at the historical graph, Figure 3.1, of development of EUR/JPY exchange rate, we can see that before the adoption of the Euro, the DM/JPY (Deutsche Mark vis-a-vis Yen) exchange rate was declining (the values in this graph are already recalculated using the exchange rate 1 Euro = 1.96 DM and DM is used as a predecessor of Euro). Both countries appreciated their currencies during this time, Japan slightly more, after the Plaza Accord, which is described in the next subsection. Right after the adoption, Euro started to appreciate against Yen, which is no surprise, because Euro area was doing well and Euro was a hot prospect at that time. The crisis hit Eurozone strongly, investors moved their funds into safe haven currencies, such as Yen, and Euro depreciated heavily, almost reaching its minimum since its introduction. The future is uncertain, as Eurozone fights with present crisis and political turmoil, and Japan with deflation worries.

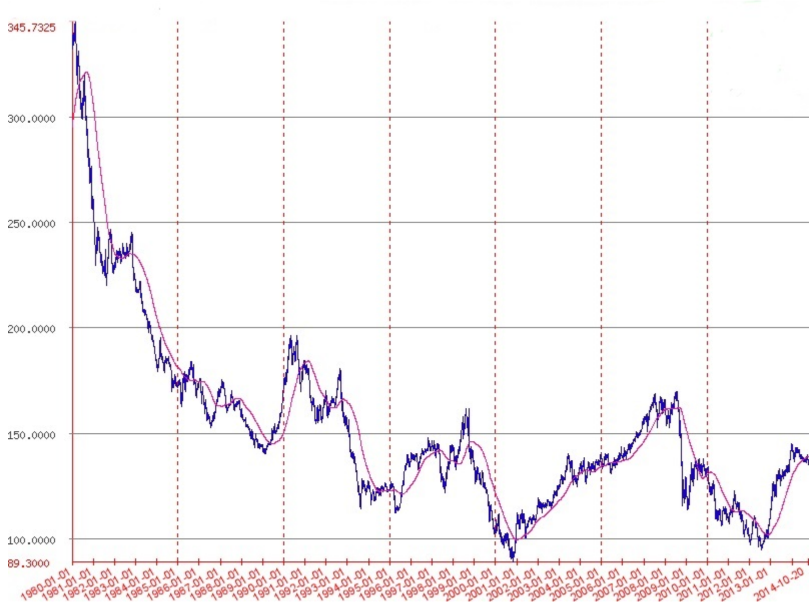

Figure 3.1: EUR/JPY historical development
Source: Fxtop.com - historical exchange rates graphs

3.1.2 GBP/JPY

GBP/JPY is the currency exchange rate of a British Pound to Japanese Yen. Similarly to EUR/JPY, this pair is very popular among private traders. It is a very volatile currency pair, mainly during Asian trading session, which creates opportunities for traders to cash on it. As a result, the value of this pair is often influenced by the traders' sentiment, whether they believe it can rise or decline, i.e. whether it is bull or bear market. It has a quite large volume of transactions, it ranks approximately fourth among major crosses. In addition to traders' perception about the market, GBP/JPY reacts strongly on interest rate or quantitative easing decisions made by respective central banks. Another interesting thing about it is that before the crisis, it was heavily used by traders for carry trade, because the United Kingdom had much higher interest rates, so it was profitable to hold Pound and sell Yen.

British Pound or Pound Sterling is the official currency of the United Kingdom of Great Britain and has been used since 1707. It is the fourth most traded currency, accounting for about 15% of trading volume. In addition, it is heavily used as a reserve currency, currently ranking third after U.S. dollar and Euro.

The historical graph, Figure 3.2, tells us that the development of this pair has been similar to the EUR/JPY. The exchange rate had been declining rapidly since 1980 until the middle of 1996. This could be attributed to the Plaza Accord. It was a meeting, where it was declared that U.S. dollar is overvalued. The solution to this problem was for Japan and Germany to boost domestic demand and appreciate their currencies, according to International Monetary Fund (2011) report. As a result, Japanese Yen appreciated 46% against U.S. dollar and eventually against other currencies as well. This resulted in a macroeconomic stimulus in terms of interest rates cut, slowly generating stock prices bubble, as pointed out in International Monetary Fund (2011) report. After the collapse, there was another appreciation up until mid 1996, according to Obstfeld (2009). In addition, during this period, UK's economy was going through hard times, which is explained in more detail in next subsection. Looking at more recent history, the exchange rate had been appreciating since 2001 up until the financial crisis due to U.K.'s economy being strong as well as the interest rate differential, which was quite large and boosting Sterling's value. When the crisis hit, investors moved their funds to safe havens, such as Yen, which caused Yen to appreciate. In addition, Bank of England decided to cut the interest rates, causing further decline in Sterling. Future of this currency pair is uncertain. Although Bank of England decided to hike the rates, which helped the value of Pound to rise, recent weak economic results can cause the hiking to stop or even the rates cut. Together with Bank of Japan's love for manipulation of Yen's value to boost its economy and love of private traders for this pair, it could be challenging to forecast the value of the exchange rate.

Figure 3.2: GBP/JPY historical development
Source: Fxtop.com - historical exchange rates graphs

3.1.3 EUR/GBP

EUR/GBP denotes the exchange rate of a Euro to Pound Sterling. As both of the currencies have large volumes of trade, this pair has quite a large amount of transactions, ranking second among major currency crosses. Compared to the previous two pairs, however, it has much lower volatility, which we can see from the Figure 3.3. The reason behind is that economies of Eurozone and the United Kingdom are interlinked and so happenings in one country move the currency of other country usually in the same direction a bit, causing the overall volatility to be lower. In addition, the exchange rate usually develops in strong trends, compared to strong and rapid movements in previous two currency pairs. As with other currency pairs, interest rate, inflation, unemployment, GDP levels etc., are important factors influencing the value of the exchange rate.

Looking at the graph of development of EUR/GBP we can notice the strong trends we talked about. Since 1980, the Deutsche Mark had been appreciating against the British Pound until mid 1996. This massive decline of value of Sterling was caused mainly due to massive inflation in the UK, when it reached

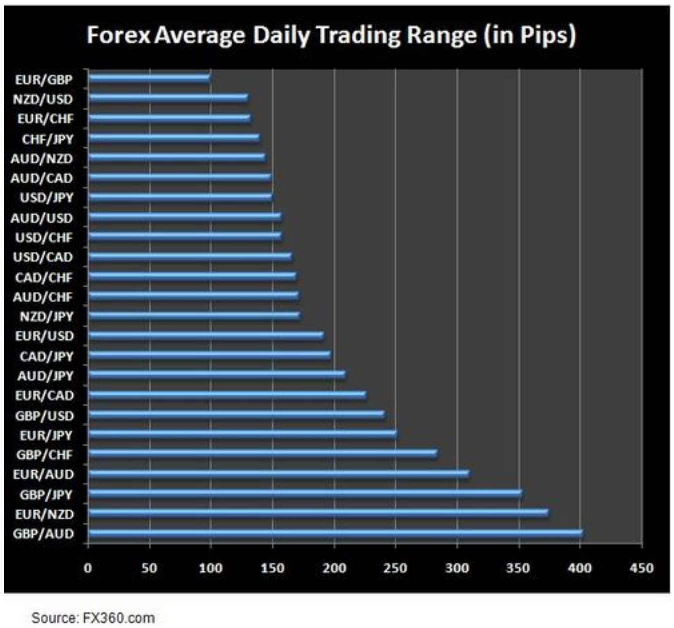

Figure 3.3: Forex average daily trading range - volatility
Source: FX360.com

over 20%. This pushed UK government to act on it, increasing interest rates and taxes, but in the process causing recession with very high unemployment, as pointed out by Pissarides (2006). Afterwards, the economy began its slow recovery. The next period is the downfall of DM (Euro later), which started mid 1995/1996, when, as stated by Ecfin (2002), Germany started to experience very lacklustre economic growth, averaging only 1.6% between 1995-2001, 1% below average in EMU/EU. In the recent history, as mentioned before, introduction of Euro together with strong economic performance was a hot prospect for investors and Euro appreciated against Pound, even throughout the financial crisis. Both Eurozone countries and UK had to cut the interest rates and introduce quantitative easing programs during this period. However, recent data showed that UK got over the crisis little bit quicker, hiking the interest rates, which caused the appreciation of Pound, which can be seen from the last part of the graph, which is declining.

3.1.4 EUR/AUD

EUR/AUD is the exchange rate of a Euro to the Australian Dollar. It is one of the most volatile currency pairs traded on forex, as illustrated from the Figure 3.3 and it could also be seen from the historical development graph, which will

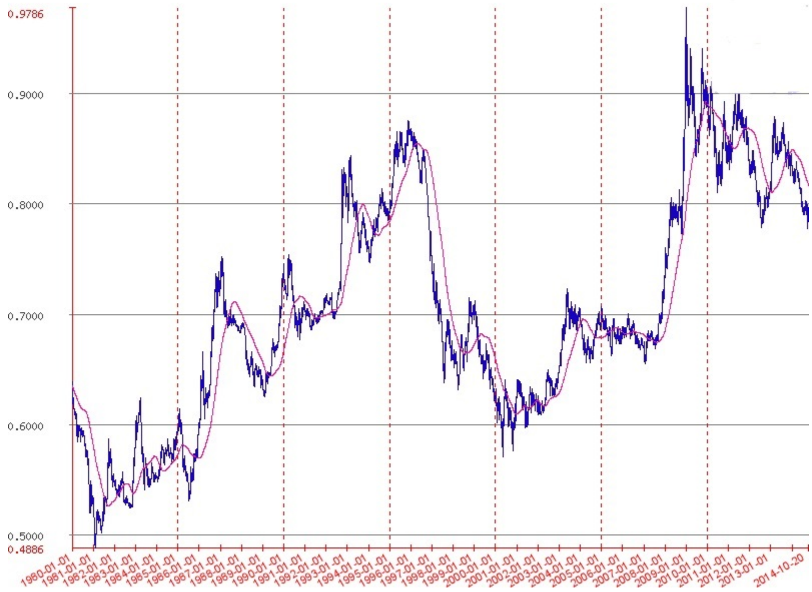

Figure 3.4: EUR/GBP historical development
Source: Fxtop.com - historical exchange rates graphs

be provided later. Although it represented only a 1% of all transactions on forex in 2010, it is an interesting currency pair due to its large volatility and the fact that Australian Dollar is significantly correlated with gold prices due to Australia being the second largest gold producer after China in the world, according to the U.S. Geological Survey published in 2014. As gold is considered to be a safe investment during bad times, Australian Dollar appreciates against most of the currencies during economic crisis. Moreover, Australia's economy is heavily linked to the economy of New Zealand, so positive happenings there can affect Australian Dollar in a positive way and vice versa.

Australian Dollar is the official tender in Commonwealth of Australia and is the fifth most traded currency on forex, according to Bank for International Settlements survey published in 2013. It is sometimes nicknames Aussie. It was introduced in the form we know today in 1966, replacing the Australian Pound. It is quite often used with currencies of countries, which have low interest rates, for a carry trade due to Australia having higher interest rates. It is sometimes referred to as commodity currency because of Australia's large exports of raw materials.

Early 1980s can be characterised as a period of economic boom for Australia. People were spending money, companies were able to generate large profits.

This changed in October 1987, when the stock markets crashed and Australia was sent to recession with high inflation and unemployment. Combined with strong economic growth of West Germany, DM/AUD as predecessor of EUR/AUD, appreciated rapidly. It wasn't until mid 1990s, when Australia nicely recovered from the recession and Germany's growth began to slow down, which resulted in appreciation of AUD against DM and later Euro. Introduction of Euro helped to appreciate EUR/AUD, but only for a short while before the pair got into the range and stayed there between 2001-2008. At that time, there was no clear trend of this pair. However, when the crisis hit the world, investors began to purchase a lot of gold, perceived to be a safe investment. This resulted in the appreciation of Australian Dollar against Euro, which we can see from the massive plunge in value of the exchange rate in the last part of the graph.

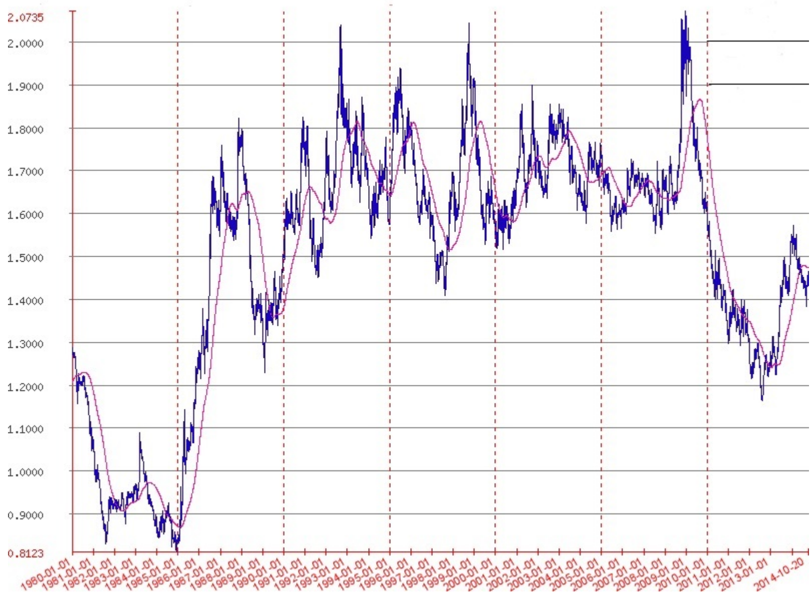

Figure 3.5: EUR/AUD historical development
Source: Fxtop.com - historical exchange rates graphs

3.1.5 AUD/JPY

AUD/JPY is the exchange rate of an Australian Dollar to Japanese Yen. It is considered as a pair with average level of volatility, as seen from Figure 3.3. It is a perfect example of a currency pair, which is used for carry trade due to Australia having much larger interest rates than Japanese Yen. The economies of the two countries are strongly interlinked, mainly due to the geographical closeness of them and the fact that Japan imports many of the raw materials Australia produces. The interesting fact is that both of them should in theory appreciate during the worse economic times, Yen because of its status as safe haven currency and Australian Dollar due to increased gold investment. As we will see from the upcoming graph, during the recent financial crisis the use of carry trade declined due to liquidity shortages and Yen appreciated compared to Aussie.

From 1980s up until 1995/96 Yen was appreciating rapidly against the Aussie. This can be contributed to the reasons already discussed, i.e. the artificial appreciation of Yen after Plaza Accord and the recession Australia went through during this period. Afterwards, the Japanese economy went into recession, which was even worsened by the government policy to decrease spendings in public sector, increase taxes and strengthen the monitoring of bank loans. Together with the Asian financial crisis, it led to the period of increased value of the AUD/JPY exchange rate. In the upcoming period, Japan started to slowly recover, according to Hoshi & Kashyap (2011), whereas the growth of Australia slowed down, as stated by Battellino (2010). Afterwards, investors were using the carry trade a lot, as Japan kept its interest rate at very low levels. This contributed to the steady appreciation of AUD/JPY until the financial crisis, when the already mentioned liquidity shortages occurred and investors decided to put their money into safe haven currency, Yen.

Figure 3.6: AUD/JPY historical development
Source: Fxtop.com - historical exchange rates graphs

3.2 Variables

There are many economic, financial or even political variables that can potentially have an effect on the forecast of exchange rates. However, in this study, we will not have a look at each variable and describe its possible effect. This has already been done in previous literature. A good example to look at is the bachelor thesis of Mida (2013), where one can find a detailed description of how various variables can influence the exchange rate movement and forecast.

In choosing the specific variables, their transformation or time horizon, we followed previous studies on this subject, mostly papers by Wright (2008), Tortora (2010) or Liu (2010), as all of them used BMA method to forecast exchange rates using a wide range of variables. After some consideration, we decided to use quarterly data, ranging from 1980 to 2013. We needed to collect data for five countries, namely Japan, Australia, the United Kingdom, Germany and Euro area. Some data for the Euro area were not available for the whole period. Therefore, we decided to use the Germany's data as a replacement where needed. In addition, not all of the variables were available for all countries, resulting in different exchange rate pairs having different amount of explanatory variables to be used in the forecast. In addition, all variables except from exchange rates, exchange rate signs and oil price index, were constructed as base currency relative to quote currency, i.e. for example for EUR/GBP pair, the relative GDP was constructed as GDP of Euro area divided by GDP of the United Kingdom. Simple summary statistics are provided in the Appendix A.

The final chosen variables were:

- **Exchange rate** - data were downloaded from three different sources, namely Oanda.com, Fxtop.com (contains historical exchange rates from 1953) and from Statistical Data Warehouse. We used log transformation and calculated returns.

- **Long term interest rate** - data came from OECD database (stats.oecd.org) and from Federal Reserve Bank of St. Louis (FRED database). We used 10 year government bonds as an equivalent for long term interest rate. Japan was a special case as data were only available from 1989 onwards. As a consequence, we decided to use a different series as proxy for the first nine years, namely yield on government bonds.

- **Short term interest rate** - data downloaded from OECD and FRED database. We used three month interbank offered rate as proxy. Japan, again, was a special case as data on interbank rates were available only from 2002 onwards. As a result, it was a combination of Treasury Bills and interbank rates.

- **Gross domestic product** - data from OECD database. GDP calculated using expenditure approach, seasonally adjusted, nominal value. We took logs of it.

- **Consumer price index** - data from OECD database. We collected data on relative CPI using 2010=100 and conducted log transformation.

- **Money supply** - source was OECD database. We used narrow money (M1), seasonally adjusted with 2010=100 as proxy for money supply and took logs of it. For the United Kingdom we were unable to gather the data for the whole studied period and therefore, currency pairs with British Pound will not have this explanatory variable.

- **Share price** - data came from OECD. We used share prices index with 2010=100 and took logs.

- **Index of industrial production** - data from OECD. They are seasonally adjusted and we used Production of total industry index with 2010=100. Again, we carried out log transformation.
- **Oil price** - data came from Quandl.com. We used Crude oil (petroleum) price index with 2005=100, which is a simple average of three spot prices, namely of Dated Brent, West Texas Intermediate and the Dubai Fateh. We also performed log transformation.
- **Yield spread** - this is a simple difference between our long term interest rate represented by 10 year government bonds and short term interest rates represented by three month interbank rate.
- **Labour productivity** - data are from OECD. We proxied productivity by GDP per person employed index with 2010=100, seasonally adjusted, log-transformed. We were not able to collect data for Germany/Euro area and as a result this variable is used only in non-Euro pairs.
- **Current account** - data came from OECD. We used current account as % of GDP, seasonally adjusted. In this case, we could not obtain data for Japan and current account was omitted in Yen pairs.
- **Sign of exchange rate returns** - computed from exchange rate returns over the previous quarter.

Chapter 4

Bayesian Model Averaging

For a long period of time researches have been trying to forecast exchange rate (or to estimate/forecast other financial and economic variables in general) using a single model, which they believed is a true model and, therefore would be able to outperform simple random walk. However, this might not have been a completely appropriate approach. The reason is that one normally uses a large variety of variables, or at least has a wide range of choices. It is very difficult to correctly choose the right amount and the most suitable variables. This is first problem. The second issue is the choice of model. There have been many models proposed, many of which were applied and nicely summarized in Rossi (2013). This creates model uncertainty when a researcher can not be sure that his chosen model is the most suitable one. The Bayesian Model Averaging (BMA) technique tries to cope with these two issues.

BMA has been slowly getting attention in the statistics and economics area. The idea itself was first proposed by Leamer (1978), but it wasn't until later that BMA rose up to prominence. At the beginning, it was mainly used in statistics, for example by Hoeting *et al.* (1999), where they provide a comprehensive tutorial into BMA technique. But as time passed, researches started to experiment with this method and use it in econometrics and economics as well, for example Min & Zellner (1993) to forecast international growth rates, to investigate the issue of model uncertainty in cross-country growth regressions (Fernandez *et al.* (2001b)) or in stock return predictability (Avramov (2002)).

4.1 Description of the model

In this section we will have a look at how this model works and why it is able to, in some cases, resolve the aforementioned problems. The greatest strength of BMA is the fact that it does not take any model as the true model at the beginning of estimation. Rather, it considers a set of n models $M_1, M_2, ..., M_n$. We know that one of them is the true one, but do not know exactly which one. However, we have some expectations about the probability that a model from this set, denoted as M_i, is the true model, as stated by Wright (2008). We can denote this probability as $P(M_i)$. The next step in the procedure is to collect and thereby observe data, and, as a consequence, update our expectations. Then we can compute the posterior probability of M_i:

$$P(M_i|D) = \frac{P(D|M_i)P(M_i)}{\sum_{j=1}^{n} P(D|M_j)P(M_j)}$$

where:

$$P(D|M_i) = \int P(D|\theta, M_i)P(\theta|M_i)d\theta$$

is the marginal likelihood of M_i, θ includes all parameters, which are used to explain dependent variable, $P(\theta|M_i)$ is the prior density of parameter vector and $P(D|\theta, M_i)$ is the likelihood, according to Wright (2008).

When conducting the forecast, each model produces forecast density. As a result, we have $f_1, f_2, ..., f_n$ forecasts to choose from. In an ideal world, we would know which model is the true one, and we would use its single density for the forecast. In reality, we cope with uncertainty and BMA methodology weights each of the point forecast by the model's posterior probability to get the final point prediction:

$$f = \sum_{i=1}^{n} P(M_i|D)f_i$$

In this study, our focus will be put on linear regression models of the form:

$$y = X\beta + \epsilon$$

where y is our dependent variable, we want to forecast, X is a matrix of explanatory variables, i.e. predictors, β is a vector of unknown coefficients and ϵ is an independent and identically distributed disturbance vector such that $\epsilon \sim N(0, \sigma^2)$, as previously defined by Liu (2010).

Next step in our estimation and forecast is to choose the estimation framework, as stated by Zeugner (2011), because specific expressions for posterior probabilities and marginal likelihoods are dependent on it. In this study, we will set our priors on constant and error variance to be improper, which means, according to Zeugner (2011), that there is complete prior uncertainty where the prior is located.

The most important prior is the one set on coefficients β. In many cases researchers form expectations about parameters into normal distribution, with certain mean and variance. Although one can choose the values according to his opinion, many studies implement a conservative prior zero mean, which, as pointed out by Zeugner (2011), reflects the fact that a researcher does not know a lot about these coefficients. For their variance, we apply Zellner's g to come to:

$$\beta|g \sim N\left(0, \sigma^2(\frac{1}{g}X'X)^{-1}\right)$$

where we will set $g = max(N, K^2)$, where N is total number of observations and K is the number of covariates. This prior was suggested by Fernandez *et al.* (2001a).

What does however hyper-parameter g imply ? According to Zeugner (2011), it shows how certain we are that the coefficients are zero. Low g means that a researcher is confident that they are indeed equal to zero, whereas high g means exactly the opposite. In this study we use a fair amount of variables (10 or 11 explanatory variables depending on the currency pair) and work with 132 observations for each variable (period from 1980 to 2013 with quarterly observations). Therefore, our g is fairly high, which is in line with theory, because we chose the variables we expect to influence our dependent variable and so the coefficients should not in most cases be zero.

We have already defined the basic linear regression model we will be using. Although a researcher can choose to shrink the number of models, e.g. by estimating models including only one variable and constant, we will use all possible permutations of explanatory variables and constant here. Altogether, we will have 2^K models with K being the number of covariates, each with a constant. This presents us with a new issue, because we need to think about model priors. Wright (2008) wrote that giving equal prior probability to all models would likely in this setting give too little prior weight to models, which contain only a few predictors.

As a result, instead of uniform prior for models, we will use binomial prior, which is helpful in setting different prior probabilities for different models. The probability that i-th model is the true one can then be defined as:

$$P(M_i) = \rho^{\kappa}(1-\rho)^{K-\kappa}$$

where κ reflects the model size. By changing the value of hyper parameter ρ we can change model's prior probability depending on its size. Setting ρ to 0.5 is the same as using the uniform prior. In addition, we will implement ρ equal to $0.25, 0.1, 0.05$, because smaller models were shown to have better predictive ability than models with large amount of variables when forecasting the exchange rate. Then the expected model size is equal to ρK and the probability of no-predictors model being the true one as $(1-\rho)^{\kappa}$, as stated by Wright (2008).

Moreover, inspired by the work of Liu (2010), we decided to estimate a slightly different version of BMA called Bayesian Model Winner (BMW). Author used only the prediction from the single best model and then compared the results to normal BMA plus an additional model, which we do not consider here. However, we will slightly change BMW. Choosing a single model is very restricting, as the whole point of BMA is the idea of combination of results from various models. Therefore, we chose to use the information from 10 best models and then compare the results with standard BMA.

4.2 Forecast application and benchmark description

In the forecasting part, we will follow the procedure applied in the past by researchers, who tried to predict changes in exchange rates. Instead of forecasting exchange rate, we will predict exchange rate returns, so that our model looks as follows:

$$e_{t+h} - e_t = \beta' X_t + \epsilon_t$$

where e_t is the logarithm of exchange rate, X_t denotes the vector of explanatory variables, in other words regressors, h reflects the forecasting horizon and ϵ_t is the error term. Here, we conduct the forecast on four different time horizons, namely one, two, four and eight quarters, representing short, medium and long horizon.

To better understand the forecasting procedure, we shall briefly describe, how exactly it works. We will be using out-of sample prediction, meaning that we will be keeping some observations aside to compare the prediction with reality. Data from Q1 1980 to Q4 2000, i.e. first 84 observations, will always be kept as a source data for forecast and we will be predicting the exchange rate returns for Q1 2001 to Q4 2013. However, after each forecast, this subsample will be updated with newer data, in other words, after using data from 1980 to 2000 to predict the exchange rate return in Q1 2001, the data from Q1 2001 will be added and the forecast will be re-estimated with updated subsample for Q2 2001. The same procedure will be applied for all time horizons. This is called the expanding window.

A very important part in the forecasting exercise is to choose the baseline statistic (model) as well as benchmarks, which will be used to assess and compare the accuracy of Bayesian Model Averaging against the baseline model. As Random Walk without drift is very difficult to beat, which has been shown in the literature review section, we will implement it as well. It looks exactly like our general prediction model, but there are no regressors. Therefore, it is in the form:

$$e_{t+h} - e_t = \epsilon_t$$

In addition to RW, we decided to follow Lam *et al.* (2008) or Liu (2010) and implement Historical Average Return baseline model into our analysis. Acronym HAR is usually used with regards to Heterogenous Autoregressive model, but we will use it in this study for Historical Average Return model. This model takes the average of the returns up until now and uses it as forecast prediction for the next period. Formally, it takes form:

$$r_{t+h} = \frac{1}{t}\sum_{t} r_t$$

where r_t is the exchange rate return in period t.

There are quite a few benchmarks, which a researcher can apply when comparing performance of various models. It is better to use more than just one, as we can not say which benchmark is the most appropriate. As a consequence, by using more types of benchmarks, we can more clearly interpret our results.

First method applied will be Root Mean Square Error (RMSE). It takes the form:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{f} - f)^2}{N}}$$

where $\hat{f}$ is our predicted value and f is the real value of exchange rate return. After summing the differences between forecast and reality, one must take square root of the average of this sum to get the final statistics. As we want to compare performance of BMA against RW, we will construct the ratio:

$$\text{Ratio BMA to RW} = \frac{\text{RMSE of BMA}}{\text{RMSE of RW}}$$

We can interpret this ratio quite easily then: If this ratio is greater than 1, then BMA method produced forecast with larger error then RW. On the other side, if it is lower than 1, BMA outperformed RW. More information about this benchmark as well as for following one can be found in Wooldridge (2009).

In addition, some researchers prefer a different measure to RMSE, called Mean Absolute Error (MAE). One of the reasons behind it, as pointed out by Tortora (2010), is that RMSE sometimes drastically penalizes single large error. According to him, a suitable benchmark should balance true costs and benefits, but this is not always the case for RMSE. As a result, he is one of those, who implemented more than one measure, including MAE, which can be defined as:

$$\text{MAE} = \frac{\sum_{i=1}^{N} |\hat{f} - f|}{N}$$

where, similarly to RMSE, $\hat{f}$ is the forecast and f is the reality. The difference here is that to get the error, we take only average of absolute values of differences. We again will proceed to create the ratio, in the same manner as for RMSE, and again, ratio with value higher than 1 indicates RW dominance and vice versa.

One more benchmark will be used in this study. We can call it the Direction of Change, also called Hit Rate, statistics. The concept is different to RMSE or MAE, because the final value of the statistics is not an error. It does not compare the forecasted values as whole, but rather it only looks at the sign of the change and whether the model predicted it correctly. We will assign number 1, if the model predicts the sign correctly and zero otherwise. We can then compute the statistics as the proportion of "ones" over total number of changes, as said by Liu (2010). He also stated that due to the fact that RW is unable to predict the sign of future change, we take the value 0.5 as a limit. Below this value we assume that RW beats our model and vice versa. We will present this statistics for each time horizon, currency pair and model prior separately as well as for the model as a whole (including all horizons and all model priors).

In upcoming section where results are discussed, we will have a look and compare the ratios and assess the accuracy. In addition, we will follow Rossi (2013) and perform t-test to test:

$$H_0 : RMSE_{forecast} - RMSE_{RW} = 0$$

against:

$$H_A : RMSE_{forecast} - RMSE_{RW} < 0$$

In other words, we will try to reject the null hypothesis of no difference between our model and RW in favor of the alternative that BMA predicts with smaller error. We will do it separately for every time horizon and currency pair and the same procedure will be applied to MAE benchmark.

One last thing to mention is the computer software we will be using. Several choices were available and we decided to go with the R Studio statistical software, where all the forecasts and t-tests were conducted. In addition, we used Excel to prepare the datasets and to compute the benchmark statistics.

Chapter 5

Discussion of results

So far we have covered the theory behind our forecast and estimation procedure. In this chapter we will focus solely on the empirical part of the thesis. We have used several benchmarks to compare the forecast with Random walk model as well as different time horizons and prior model probabilities. As a consequence, we have a lot of results for each currency pair and a lot to explain. We will divide this chapter into five sections corresponding to five currency pairs and present one table with results for each pair.

Before moving on to the actual results, we should describe the tables so that it is clear for the reader what he can see in them. Each table contains four parts, corresponding to the three benchmarks used, i.e. RMSE, MAE and Direction of Change (DoC), and one part stating the P-values of t-tests for RMSE and MAE differences. The test was described in the previous chapter. Each part includes results for BMA and BMW model and an additional column, where we compare the average of the values between BMA and BMW for each time horizon, i.e. for forecast of one, two, four and eight quarter ahead. In addition, for DoC and P-values part, we included one more row, where the overall results for all horizons and model priors are stated. The values in the tables are computed as ratios for MAE and RMSE part, percentages for DoC part and probabilities for P-values part.

Last thing before we proceed is to state our hypotheses/expectations. From the theory and previous literature point of view, we would expect the RW model to be superior to our model at short horizons and if inferior, then at medium to long horizons. However, in the core paper for exchange rate forecasting using

BMA from Wright (2008), he obtained smaller errors in most cases at short horizons rather than long horizons. However, we will stick to the general point of view and examine the hypothesis that BMA/BMW should be superior at medium/long horizons and inferior at short one. Regarding the direction of change, we expect our models to beat RW, meaning that they should forecast the sign of the change correctly in more than 50% cases.

Looking at the specifics of the currency pairs studied, we established some expectations about them. In chapter three, we presented a graph with volatilities of our pairs. The more volatile ones, i.e. EUR/JPY, GBP/JPY and EUR/AUD should be more difficult to forecast due to the fact they are quite popular among private forex traders, who can change the sentiment and unexpectedly influence the exchange rates with large amount of transactions, meaning it is hard to predict any movement in them. On the other side, less volatile EUR/GBP and AUD/JPY, i.e. exchange rates of countries which are economically interconnected, should be easier to predict, as one currency moves, when the other does, leading to lower spikes and not so rapid movements.

5.1 EUR/GBP results

We will start our discussion of results with the least volatile currency pair in our sample. Looking at the RMSE ratio we can see that BMA and BMW were able to outperform our benchmark RW model at the shortest horizon, which is something we would not expect according to the theory, However, it is not so surprising if we consider results of Wright (2008). Our results are confirmed by running the t-test as we strongly reject the null hypothesis of no difference between RMSE of BMA/BMW and RW in favor of the alternative that errors from our models are lower than from RW. On longer horizons RW was much more successful and we failed to reject the null hypothesis, Comparing performance of BMA and BMW, on average for different horizons, BMA produced consistently better forecasts.

MAE benchmark was mainly used so that we could confirm our results from RMSE and eventually, we did. Again, we were able to outperform RW only at the shortest interval of one quarter with BMA being, on average, more successful than BMW model. Another thing we can observe, which is line with previous literature, was that smaller models in terms of variables predicted with smaller error. We started with uniform distribution of prior model probabilities $P(M_i)$ and later forecasted with priors, which favored models with less variables. As one can see, in most cases, the errors were decreasing as we assigned higher prior to smaller models (0.50 = uniform prior, smaller than 0.50 prior = smaller models prioritised).

Regarding the percentage of times our models were able to correctly predict the sign of the change, we were quite successful. On most of the time horizons and model priors, we were able to predict the sign with higher than 50% accuracy, with the overall accuracy of over 57%, meaning that our models produced better forecasts in terms of DoC, with BMA being better than BMW.

Looking at overall statistics, BMA performed much better than BMW, as BMW was only able to outperform BMA in 3 out of 32 forecasts (in terms of RMSE and MAE). More interestingly, in terms of DoC, BMW produced equally good results in 5 and better in 4 cases than BMA, suggesting BMW can be quite useful for forecasting the sign of change. The most important thing, however, is the failure to reject our null hypothesis in case we tested the difference between all errors at all horizons and model priors, meaning the superiority of RW.

EUR/GBP

RMSE	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.79	0.75	0.78	0.80	BMA	0.92	0.75	0.78	0.80
Q2	1.96	1.66	1.34	1.12	BMA	2.08	1.70	1.36	1.12
Q4	2.06	1.87	1.53	1.36	BMA	2.15	1.96	1.54	1.37
Q8	1.29	1.16	1.07	1.01	BMA	1.35	1.18	1.08	1.02

MAE	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.81	0.76	0.77	0.79	BMA	0.86	0.75	0.77	0.79
Q2	2.02	1.67	1.34	1.19	BMA	2.14	1.70	1.34	1.19
Q4	2.13	1.89	1.44	1.22	BMA	2.21	1.95	1.43	1.22
Q8	1.48	1.34	1.22	1.15	BMA	1.54	1.36	1.24	1.16

DoC	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.71	0.75	0.75	0.75	BMA	0.73	0.71	0.75	0.75
Q2	0.55	0.57	0.53	0.55	BMA	0.53	0.57	0.51	0.55
Q4	0.48	0.48	0.50	0.46	BMW	0.46	0.48	0.54	0.50
Q8	0.54	0.49	0.49	0.49	BMA	0.51	0.51	0.46	0.46
Overall	0.577				BMA	0.574			

P-values	BMA (RMSE)	BMA (MAE)	BMW (RMSE)	BMW (MAE)
Q1	0.0001	0.0001	0.0084	0.0014
Q2	0.9669	0.9706	0.9640	0.9657
Q4	0.9896	0.9756	0.9875	0.8728
Q8	0.9416	0.9872	0.9393	0.9851
Whole	0.8557	0.8996	0.8804	0.9129

Table 5.1: Forecast results for EUR/GBP currency pair

The table consists of four parts. The first numeric row in each of first three subtables includes model prior probabilities. The "Whole" column then shows the comparison between the average of errors/DoC in particular quarter for all model priors between BMA and BMW. If it says "BMA", it means that in this quarter the BMA outperformed BMW. In the RMSE and MAE part, the number larger than 1 indicates that RW outperformed BMA/BMW and vice versa. In DoC part, the number indicates the percentage of correct predictions of signs of changes. The row "Overall" states the total percentage of correctly predicted signs of changes throughout all quarters and model priors. In P-values part the number close to 0 means that we are able to reject the null hypothesis and vice versa. The row "Whole" shows the p-values when running the test of differences between all errors across all quarters and model priors.

5.2 AUD/JPY results

We will continues with the second least volatile pair. Starting with RMSE benchmark, again, we produced better forecast at the shortest horizon. However, in this case, our BMA matched RW much more closely. You can see that on the medium horizon of two quarters, BMA forecasted only slightly worse than RW and on the interval of four quarters, our model outperformed RW in all cases, but uniform model prior. Same can be said about BMW. These quite promising results were, unfortunately, then rejected as only for one quarter we rejected the null hypothesis.

Examining MAE benchmark shows us reason why a researcher should, if possible, use more than one benchmark. Every benchmark uses a different loss function, which can then produce different results. We confirmed better performance of our models at Q1 and Q4 interval, but, in addition, we slightly outperformed RW even in case of Q2 horizon. This all sounded, once more, very promising, but results of t-test proved us otherwise, because for Q2, Q4 and Q8, we were unable to reject the null.

Regarding the DoC, our models showed some stellar performance. In all cases, for all horizons and model priors, the percentage of correctly predicted sign of change was higher than 50%, thus models comfortably outperformed RW. BMA was able to beat BMW on all intervals, on average and produced a quite staggering accuracy figure of around 67%, which as we will see later, was the highest number among all currency pairs.

Although BMA had solid performance for AUD/JPY, at the end we again failed to reject the null. BMW model was no competition for BMA, as it forecasted with lower error only in 5 out of 32 cases and had higher accuracy of change in only 3 cases out of 16 (with 3 equally good). We did expect some better forecasting results for this and the previous currency pair due to their lower volatility and interconnected economies, but we were right only in case of the DoC.

AUD/JPY

RMSE	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.68	0.71	0.74	0.76	BMA	0.68	0.72	0.75	0.76
Q2	1.01	1.00	1.01	1.02	BMA	1.01	1.03	1.01	1.02
Q4	1.56	0.96	0.82	0.79	BMA	1.31	0.93	0.82	0.80
Q8	1.31	1.27	1.26	1.22	BMA	1.32	1.28	1.28	1.24

MAE	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.58	0.64	0.70	0.73	BMA	0.58	0.65	0.70	0.73
Q2	0.97	0.97	1.03	1.06	BMA	0.96	0.99	1.04	1.07
Q4	1.09	0.82	0.76	0.78	BMA	1.07	0.83	0.77	0.79
Q8	1.14	1.13	1.15	1.11	BMA	1.14	1.15	1.19	1.13

DoC	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.86	0.88	0.82	0.82	BMA	0.90	0.86	0.80	0.80
Q2	0.67	0.67	0.59	0.59	BMA	0.63	0.57	0.59	0.59
Q4	0.60	0.58	0.60	0.53	BMW	0.53	0.56	0.53	0.56
Q8	0.59	0.57	0.59	0.57	BMA	0.62	0.57	0.57	0.54
Overall	0.669				BMA	0.648			

P-values	BMA (RMSE)	BMA (MAE)	BMW (RMSE)	BMW (MAE)
Q1	0.0024	0.0003	0.0009	0.0009
Q2	0.9713	0.9958	0.6049	0.6817
Q4	0.5611	0.3943	0.0877	0.0767
Q8	0.9996	0.9997	0.9996	0.9994
Whole	0.6198	0.4146	0.5973	0.4264

Table 5.2: Forecast results for AUD/JPY currency pair

The table consists of four parts. The first numeric row in each of first three subtables includes model prior probabilities. The "Whole" column then shows the comparison between the average of errors/DoC in particular quarter for all model priors between BMA and BMW. If it says "BMA", it means that in this quarter the BMA outperformed BMW. In the RMSE and MAE part, the number larger than 1 indicates that RW outperformed BMA/BMW and vice versa. In DoC part, the number indicates the percentage of correct predictions of signs of changes. The row "Overall" states the total percentage of correctly predicted signs of changes throughout all quarters and model priors. In P-values part the number close to 0 means that we are able to reject the null hypothesis and vice versa. The row "Whole" shows the p-values when running the test of differences between all errors across all quarters and model priors.

5.3 EUR/JPY results

In the following three sections we will comment on the results from the three pairs, which are volatile and quite popular among private traders. We will start with the EUR/JPY. which was the only pair for which BMA outperformed RW model at the longest horizon of eight quarters. Moreover, our models also produced better forecasts for Q1. The results were confirmed by rejection of the null. On the other side, results for medium horizons of Q2 and Q4 were quite bad. As you can see, the ratios of errors are very high, for both BMA and BMW, leading us to a conclusion that RW had a much lower forecasting RMSE than chosen models.

Very similar results were computed using MAE as benchmark. For Q1 and Q8, BMA beat RW, which was again confirmed by the t-test. At medium horizons, the RW predicted with smaller errors, especially for Q2, where the ratio is again quite high in favor of RW. In this case, BMW was able to outperform BMA at Q4 horizon, for which it had three values of errors out of four lower than BMA. However, as BMW found no success for other horizons, its success could be attributed to a specific combination of currency pair, forecasting horizon and model prior.

A very interesting thing can be seen from values of DoC statistic. Although our models outperformed RW at both Q1 and Q8 horizon, in terms of DoC, they failed largely at Q8 horizon, They achieved average accuracy of only around 28%, which was quite surprising due to the fact, how heavily BMA beat RW (RMSE and MAE ratios quite low). After seeing the results of RMSE and MAE comparison, we expected at least 50% overall predictability of sign of change, but the models achieved only around 48%, which was a disappointing number.

Although there were some promising results, eventually, we failed to reject the null of no difference. BMW model showed little promise regarding the RMSE and MAE, as only in 6 out of 32 cases it produced forecast with lower error (both times at Q4 interval). More interestingly, BMW was quite successful in terms of DoC. It produced equally good prediction in 7 and better in 4 cases out of 16, which could suggest its viability for this type of forecast.

EUR/JPY

RMSE	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.93	0.90	0.85	0.85	BMA	0.98	0.92	0.85	0.85
Q2	2.13	2.12	2.00	1.84	BMA	2.26	2.21	2.05	1.85
Q4	2.96	2.21	1.61	1.40	BMA	3.26	2.20	1.53	1.24
Q8	0.74	0.71	0.70	0.69	BMA	0.76	0.73	0.70	0.69

MAE	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.88	0.89	0.88	0.88	BMA	0.90	0.91	0.88	0.88
Q2	1.81	1.81	1.76	1.70	BMA	1.87	1.86	1.79	1.71
Q4	1.53	1.38	1.27	1.21	BMW	1.56	1.37	1.25	1.17
Q8	0.62	0.62	0.62	0.61	BMA	0.63	0.63	0.63	0.62

DoC	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.71	0.71	0.75	0.73	BMA	0.71	0.71	0.73	0.71
Q2	0.57	0.51	0.39	0.35	BMW	0.59	0.53	0.41	0.35
Q4	0.51	0.44	0.40	0.36	DRAW	0.51	0.47	0.38	0.36
Q8	0.30	0.30	0.27	0.27	BMA	0.24	0.27	0.27	0.27
Overall	0.488				BMA	0.485			

P-values	BMA (RMSE)	BMA (MAE)	BMW (RMSE)	BMW (MAE)
Q1	0.0042	0.02093	0.00004	0.0003
Q2	0.9997	0.9994	1.0000	0.9999
Q4	0.9706	0.9505	0.9924	0.9860
Q8	0.00005	0.0001	$1.52E-07$	$1.41E-06$
Whole	0.9206	0.4600	0.9250	0.4771

Table 5.3: Forecast results for EUR/JPY currency pair

The table consists of four parts. The first numeric row in each of first three subtables includes model prior probabilities. The "Whole" column then shows the comparison between the average of errors/DoC in particular quarter for all model priors between BMA and BMW. If it says "BMA", it means that in this quarter the BMA outperformed BMW. In the RMSE and MAE part, the number larger than 1 indicates that RW outperformed BMA/BMW and vice versa. In DoC part, the number indicates the percentage of correct predictions of signs of changes. The row "Overall" states the total percentage of correctly predicted signs of changes throughout all quarters and model priors. In P-values part the number close to 0 means that we are able to reject the null hypothesis and vice versa. The row "Whole" shows the p-values when running the test of differences between all errors across all quarters and model priors.

5.4 GBP/JPY results

Moving on, we have a look at the most volatile pair in our sample. As in previous cases, chosen models forecasted with smaller errors at Q1 interval, which was confirmed by very small p-values of t-tests, both for BMA and BMW. However, on different horizons we did not find much success. We outperformed RW at Q4 interval for model prior of 0.10 and 0.05, but as we could not reject the null, we can attribute this finding to the specific setting rather than general better performance. Assessing BMW statistics we can see that BMW, on average, did not outperform BMA in any setting.

MAE confirmed our findings. It showed exactly the same patent as RMSE, i.e. better than RW performance at shortest horizon and worse at longer horizons, with the exception at Q4. BMW was again unable to beat BMA in any forecasting interval.

Compared to the previous studied currency pairs, the DoC statistics was very low, even for horizons for which BMA outperformed RW in terms of RMSE and MAE. Throughout Q2, Q4 and Q8, the average accuracy is only 36% and 63% for Q1.This is also confirmed by the overall accuracy, which is only 43% for BMA and 40% for BMW, i.e. the lowest we have had so far.

Unsurprisingly, we could not reject the null hypothesis for the whole sample of errors. These results are in line with what we expected due to the nature of this currency pair, which often experiences unexpected movements, which are hard to predict. BMW produced better forecast than BMA in terms of our error benchmarks in 7 cases out of 32, again confirming that averaging over large amount of models should produce better forecast. In terms of DoC, BMW had three better performances and five equally good compared to BMA. However, overall percentage was around 3% lower.

GBP/JPY

RMSE	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.83	0.83	0.84	0.85	BMA	0.83	0.84	0.84	0.85
Q2	1.24	1.10	1.08	1.07	BMA	1.25	1.09	1.08	1.07
Q4	2.00	1.50	0.87	0.75	BMA	2.33	1.57	0.85	0.74
Q8	1.41	1.38	1.33	1.26	BMA	1.43	1.40	1.33	1.27

MAE	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.80	0.82	0.84	0.84	BMA	0.81	0.83	0.84	0.85
Q2	1.42	1.18	1.15	1.15	BMA	1.41	1.15	1.15	1.15
Q4	1.80	1.37	0.86	0.72	BMA	1.95	1.40	0.84	0.69
Q8	1.61	1.56	1.48	1.40	BMA	1.62	1.58	1.49	1.41

DoC	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.67	0.63	0.61	0.63	BMA	0.59	0.59	0.61	0.63
Q2	0.37	0.41	0.33	0.33	BMA	0.39	0.35	0.24	0.27
Q4	0.38	0.31	0.33	0.31	DRAW	0.36	0.33	0.33	0.31
Q8	0.41	0.30	0.49	0.35	BMA	0.35	0.32	0.32	0.35
Overall	0.434				BMA	0.404			

P-values	BMA (RMSE)	BMA (MAE)	BMW (RMSE)	BMW (MAE)
Q1	$2.06E-05$	$9.57E-06$	0.0001	$9.62E-05$
Q2	0.9708	0.9687	0.9795	0.9770
Q4	0.7979	0.8070	0.7492	0.7519
Q8	0.9991	0.9990	0.9993	0.9992
Whole	0.8060	0.8277	0.8329	0.8405

Table 5.4: Forecast results for GBP/JPY currency pair

The table consists of four parts. The first numeric row in each of first three subtables includes model prior probabilities. The "Whole" column then shows the comparison between the average of errors/DoC in particular quarter for all model priors between BMA and BMW. If it says "BMA", it means that in this quarter the BMA outperformed BMW. In the RMSE and MAE part, the number larger than 1 indicates that RW outperformed BMA/BMW and vice versa. In DoC part, the number indicates the percentage of correct predictions of signs of changes. The row "Overall" states the total percentage of correctly predicted signs of changes throughout all quarters and model priors. In P-values part the number close to 0 means that we are able to reject the null hypothesis and vice versa. The row "Whole" shows the p-values when running the test of differences between all errors across all quarters and model priors.

5.5 EUR/AUD results

We shall finish the explanation of results with the currency pair, which is somewhere in the middle in terms of volatility in our sample. Evaluating the results of RMSE benchmark we can confirm the trend that our chosen models are able to beat RW at short horizon. This is further acknowledged by the t-test, which vert strongly reject the null hypothesis. Apart from this, we did not find any evidence of BMA having more accurate forecasting power than RW, as ratios are all larger than one, except from the Q8 forecast with model prior set to 0.05. This again, can be attributed to specific setting.

Using MAE as benchmark, we came to the same conclusions as for RMSE. In both cases, RW produced much smaller forecasting error, which can be seen from ratios being higher than one by quite a high margin. BMW, unlike in RMSE, where it, on average, produced slightly better results than BMA, was not able to outperform BMA on any horizon.

The results of DoC statistic are quite similar to the previous currency pair. The values are quite low, except the Q1 horizon, and the average is only 35%. In this case, however, the interesting point is the performance of BMW model. In four cases it predicted the sign with better and in nine cases with equally good accuracy. For the whole forecast, it had the same percentage of correct predictions as BMA. This would be a very promising result, because BMW forecast is more simple and faster to produce, but the overall percentage was only around 45% percent, meaning that RW outperformed both methods.

As for all previous currency pairs, we fail to reject the null hypothesis using all the errors. Although BMW produced better forecasts in terms of errors in 10 out of 32 cases, which is the highest number in our sample, it is still a low portion to come to any conclusions. Furthermore, the inability of chosen models to outperform RW in terms of DoC disregards also the fact that BMW showed good performance compared to BMA.

EUR/AUD

RMSE	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.82	0.81	0.80	0.80	BMA	0.85	0.82	0.81	0.80
Q2	1.61	1.58	1.47	1.27	BMA	1.61	1.59	1.49	1.28
Q4	1.69	1.71	1.75	1.78	BMW	1.69	1.71	1.75	1.77
Q8	1.89	1.78	1.13	0.76	BMA	1.92	1.90	1.15	0.76

MAE	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.81	0.81	0.79	0.78	BMA	0.83	0.82	0.79	0.78
Q2	1.67	1.66	1.52	1.28	BMA	1.67	1.67	1.55	1.29
Q4	1.73	1.77	1.80	1.83	BMA	1.74	1.76	1.80	1.83
Q8	2.07	1.91	1.22	0.90	BMA	2.12	2.03	1.24	0.89

DoC	BMA				Comparison	BMW			
	0.50	0.25	0.10	0.05	Whole	0.50	0.25	0.10	0.05
Q1	0.55	0.65	0.78	0.76	BMA	0.49	0.65	0.78	0.76
Q2	0.37	0.39	0.37	0.37	BMW	0.39	0.39	0.37	0.39
Q4	0.40	0.31	0.33	0.36	BMW	0.38	0.36	0.33	0.36
Q8	0.32	0.35	0.35	0.32	DRAW	0.32	0.35	0.38	0.30
Overall	0.448				DRAW	0.448			

P-values	BMA (RMSE)	BMA (MAE)	BMW (RMSE)	BMW (MAE)
Q1	$2.23E-05$	0.0002	$6.83E-05$	0.0003
Q2	0.9958	0.9963	0.9952	0.9957
Q4	1.0000	1.0000	1.0000	1.0000
Q8	0.8796	0.8861	0.9217	0.9228
Whole	0.9411	0.9655	0.9445	0.9674

Table 5.5: Forecast results for EUR/AUD currency pair

The table consists of four parts. The first numeric row in each of first three subtables includes model prior probabilities. The "Whole" column then shows the comparison between the average of errors/DoC in particular quarter for all model priors between BMA and BMW. If it says "BMA", it means that in this quarter the BMA outperformed BMW. In the RMSE and MAE part, the number larger than 1 indicates that RW outperformed BMA/BMW and vice versa. In DoC part, the number indicates the percentage of correct predictions of signs of changes. The row "Overall" states the total percentage of correctly predicted signs of changes throughout all quarters and model priors. In P-values part the number close to 0 means that we are able to reject the null hypothesis and vice versa. The row "Whole" shows the p-values when running the test of differences between all errors across all quarters and model priors.

5.6 Comparison with Historical Average Return

As we already mentioned in Chapter 4, in addition to RW, we also compared the performance of BMA and BMW against the HAR model. It is another baseline statistic (model), which is hard to beat and we expect the results to be similar in terms of errors. Based on the results of the two studies from Lam *et al.* (2008) and Liu (2010), we believe it could be even more difficult for our chosen models to outperform HAR. In these papers, the ratios of errors were actually larger for HAR than for RW.

We will not focus on DoC statistics for HAR model as it can not be interpreted due to the way the HAR forecasts are computed. They are the average of past returns. In next period, the prediction is the average of the same values as before and one additional value and so on. As a result, the sign of exchange rate change does not change its value very often, because the averages are quite similar in each period.

Due to the fact that we already found out that BMW predicts with larger error than BMA in most cases and after running the estimation, we found almost the same pattern, we will not present BMW results here. However, they will be included in the Appendix B. Moreover, as we already discussed the results in previous sections in this chapter separately for each currency pair and as we do not expect large differences, we will focus on the comparison of BMA/RW and BMA/HAR error ratios.

We first looked at the RMSE statistic. As expected, we mostly found the same pattern as when using RW as baseline model. However, there were some interesting differences. The only pair, which had totally identical pattern was EUR/GBP. On the other hand, for EUR/JPY, HAR proved to be a tough benchmark to beat as BMA did not outperform HAR at Q8 horizon, unlike when RW was used. In addition, for Q1 horizon, BMA was unable to beat HAR when the model prior was equal to 0.5 and 0.25.

For GPB/JPY and EUR/AUD we found only small differences. For GBP/JPY the only change was inability to outperform HAR at Q2 horizon with model prior equal to 0.1 and 0.05, for EUR/AUD at Q8 for model prior 0.05. Looking at the last currency pair, we were able to outperform HAR at the whole

Q2 horizon, but not at Q4 horizon (for all priors but 0.5). This is the exact opposite as for RW and probably was the most interesting finding, although it is not possible to explain why it happened in such way.

Basically the same as for RMSE happened for MAE benchmark. The only difference was for EUR/JPY pair, where HAR outperformed BMA for Q1 horizon with model prior 0.25. Otherwise the pattern was the same as for RMSE. At the end HAR proved to be a tough baseline model to beat, we weren't able to outperform it with our model in some cases, where we, on the other hand, succeeded when the baseline statistic was RW. In addition, the error ratios were higher than when RW was used, which is in line with the results of the two studies we used as an inspiration.

	RMSE				P-values	MAE				P-values
EUR/GBP	0.50	0.25	0.10	0.05		0.50	0.25	0.10	0.05	
Q1	0.95	0.90	0.93	0.96	0.007	0.98	0.92	0.93	0.95	0.013
Q2	1.94	1.64	1.33	1.11	0.965	1.95	1.61	1.29	1.14	0.965
Q4	3.23	2.94	2.41	2.14	0.997	3.29	2.93	2.22	1.88	0.992
Q8	4.57	4.11	3.77	3.57	0.999	5.04	4.58	4.17	3.92	0.999
EUR/JPY										
Q1	1.07	1.03	0.98	0.98	0.755	0.99	1.01	0.99	0.99	0.256
Q2	2.06	2.05	1.93	1.78	0.999	1.68	1.68	1.63	1.57	0.999
Q4	5.00	3.74	2.71	2.36	0.958	2.72	2.46	2.27	2.16	0.999
Q8	2.05	1.99	1.94	1.92	1.000	2.19	2.19	2.18	2.16	1.000
GBP/JPY										
Q1	0.97	0.98	0.99	0.99	0.014	0.94	0.96	0.98	0.99	0.026
Q2	1.22	1.10	1.06	1.05	0.958	1.34	1.11	1.08	1.09	0.957
Q4	3.72	2.79	1.61	1.40	0.958	3.61	2.75	1.72	1.44	0.966
Q8	4.54	4.43	4.26	4.06	1.000	6.27	6.08	5.78	5.48	1.000
EUR/AUD										
Q1	0.99	0.97	0.96	0.96	0.008	0.96	0.96	0.94	0.92	0.006
Q2	1.68	1.64	1.53	1.32	0.997	1.71	1.70	1.55	1.31	0.996
Q4	3.12	3.16	3.23	3.27	1.000	3.37	3.43	3.51	3.55	1.000
Q8	5.90	5.56	3.54	2.38	0.986	6.25	5.76	3.69	2.70	0.988
AUD/JPY										
Q1	0.87	0.91	0.95	0.94	0.019	0.97	0.83	0.90	0.94	0.020
Q2	0.99	0.98	0.98	0.98	0.008	0.99	0.89	0.94	0.98	0.199
Q4	2.43	1.50	1.27	1.26	0.941	1.23	1.34	1.23	1.26	0.975
Q8	2.74	2.66	2.63	2.55	1.000	2.55	3.21	3.27	3.14	1.000

Table 5.6: Comparison of BMA model with HAR

The first numeric row includes model prior probabilities. For RMSE and MAE columns, the number larger than 1 indicates that HAR outperformed BMA/BMW and vice versa. For P-values columns the number close to 0 means that we are able to reject the null hypothesis for RMSE/MAE for particular quarter and vice versa.

5.7 Summary

In this section we shall sum up the results, which were separately presented for each currency pair in previous sections:

- **RMSE** - in all cases we were able to outperform RW with our chosen models at the shortest interval of one quarter, which was also confirmed by the ability to reject the null hypothesis in all cases. For all other horizons, we did not find much evidence of BMA or BMW being consistently able to beat RW, as we failed to reject the null in all cases, but EUR/JPY Q8 interval. Moreover, BMW was able to produce forecasts with smaller RMSE only in 15 out of 80 cases in total, suggesting it is not a better method than BMA.

- **MAE** - we found very similar evidence as when using RMSE as benchmark. Basically, it produced the same findings, i.e. strong performance at Q1 and weak at other horizons. The only difference was found for AUD/JPY Q2 interval for model prior of 0.50 and 0.25. BMW performance was also very similar, beating BMA in only 17 out of 80 cases. MAE in some cases can produce different results than RMSE due to different loss function, but in this study, it was not the case.

- **DoC** - our models found the most success in correctly forecasting the sign of change. For two currency pairs we comfortably beat RW, achieving accuracy well over 50%. In two cases, however, the percentage was well smaller than 50%. On average, the forecasting accuracy was 52%, suggesting BMA could be a viable tool for this type for forecast, but the results depend heavily on the currency pair, because two best results were achieved for the less volatile pairs. Even BMW showed some promise, as altogether it produced equally good forecast in 29 cases and better in 18 cases out of 80. However, we will still be better off using BMA as, on average, neither did BMW beat BMA for the whole sample, reaching the value of 51%, nor for whole sample for individual currency pairs.

- **HAR baseline model** - as expected, we found very similar pattern to RW. However, in some specific cases, HAR proved to be even tougher to beat than RW. This again proves the point that choosing different baseline model (or benchmark) could potentially lead to different results.

As we expected, BMW was unable to outperform BMA in most cases, the pattern was almost the same as for RW.

5.8 Comparison with previous literature

In the last part of this chapter we will briefly have a look at the results of past studies and compare them with ours. We will first have a look at the core study by Wright (2008). He used RMSE benchmark to compare BMA forecasts with RW. For each of the four currency pairs, BMA outperformed RW at the shortest horizon of Q1. We were able to achieve the same in this study. However, he also succeeded at longer horizons and for all model priors, except for GBP/USD pair, where he did not find any success at Q2, Q4 or Q8. We, on the other hand, could not consistently beat RW at any of the longer horizons. In terms of DoC, BMA was able to predict most of the times with accuracy over 50%. We showed the same thing for less volatile pairs, but performed worse with more volatile pairs.

Next, we move on to Lam *et al.* (2008). He implemented RW and HAR baseline models and RMSE as benchmark for three currency pairs against U.S. Dollar. For EUR/USD and USD/JPY he found much success, with BMA constantly outperforming both baseline models in terms of errors. However, for GBP/USD, baseline models performed better in most cases than BMA except for the Q1 horizon. This again confirmed the fact that at Q1 horizon, BMA method performs very strongly and GBP pairs are often hard to predict, as in our case. Regarding DoC, the results were mixed, with strong performance for GBP/USD and EUR/USD and weak for USD/JPY.

A very thorough analysis was performed by Liu (2010). He used rolling and expanding window, BMA and two other modifications of BMA, RMSE and DoC benchmark and RW and HAR as baseline models. BMA method was in this case quite unsuccessful, outperforming neither RW or HAR consistently. BMW was even worse, unable to beat baseline statistics for any horizon, model prior or currency pair except four cases (out of 48). His results are more similar to ours, but we were slightly more successful. In this study, chosen models could not outperform benchmarks event at shortest horizon, which so far, has been the most successful interval. Looking at DoC, BMA and BMW performed nicely, with only few setting producing accuracy lower than 50% In this regard, he was more successful.

Although we were not so successful with beating baseline models, we definitely did not perform the worst. We have seen that even two authors using the same currency pairs, but slightly different settings, can come up with opposite results. This shows how hard it is to find the correct setting even when one uses method, which should decrease uncertainty thanks to model averaging. We have also learned that GBP pairs are quite hard to predict. On the other hand, BMA can quite consistently produce good results in terms of correctly predicted sign of changes.

Chapter 6

Conclusion

The objective of this thesis was to examine whether BMA and BMW methods, which have shown promising performance in the previous studies, can be applied to exchange rate forecasting problem of currency pairs, which do not include U.S. Dollar. We chose five of them, namely Euro/Japanese Yen, Euro/British Pound, Euro/Australian Dollar, British Pound/Japanese Yen and Australian Dollar/Japanese Yen. Although they do not include U.S. Dollar, they are still traded in large amounts.

Based on the previous literature we decided to use RW as the primary baseline model. The reason is that since the first studies it has been very difficult and in many cases researchers were unsuccessful to outperform RW. Eventually, we also decided to implement HAR model and compare the results with the setting when RW was used. Over the time, researchers came up with various methods, but still were usually only partially able to, in some specific setting, produce more accurate forecasts than RW. We had some prior expectations about the performance of BMA versus RW. More volatile pairs were expected to be harder to predict than less volatile pairs of countries, which in addition, are economically interconnected. We implemented four different forecasting horizons as well as four different model prior probabilities.

We used three different benchmarks to examine the performance of our models against RW. RMSE and MAE are statistics, which measure the predictive error and DoC looks at the sign of change of exchange rate and the accuracy with which models are able to predict it. We computed ratios for the first two benchmarks of model error against RW error. If the ratio is greater than one, RW forecasts with smaller error than chosen model and vice versa. Regarding the DoC, we assume RW has a statistic equal to 0.5 or 50%. As a result, a value higher than 0.5 indicates higher accuracy. In addition to typical BMA method, we also employed slightly adjusted BMW technique, which takes into account only the performance from the ten best models. We expected it to perform worse in terms of errors than BMA, as it is more restrictive. However, we were hopeful it could be a viable tool for forecasting sign changes.

The final results found little evidence in favor of BMA or BMW method in terms of RMSE and MAE. We did consistently outperform RW at the shortest horizon of one quarter, but failed to do the same at other horizons, where we failed to reject the null hypothesis of no difference in favor of the alternative that errors from BMA/BMW are smaller than errors from RW forecast. Testing the null for all the errors, throughout all horizons and model priors, we failed to reject it in all cases. Examining results for BMW we came to conclusion that BMA is much more accurate, because only in a very few cases BMW outperformed BMA. Comparing the results with HAR, we found only a few interesting differences. At the end, the pattern was very similar to RW, however, in some cases, HAR ended up as a model, which is even tougher to outperform.

Slightly different story was the performance in terms of DoC. In case of less volatile currency pairs, namely EUR/GBP and AUD/JPY, we succeeded in achieving high overall accuracy of 58% and 67% respectively, which is quite a high percentage. For more volatile pairs, we failed to beat the benchmark of 50%. This was quite disappointing for us, as in previous literature, BMA almost always predicted the sign of change at least half the time. The performance of BMW was mediocre. It produced better forecasts in only around 25% cases, so we are better off using BMA.

Although we have used a number of currency pairs, benchmarks and model prior probabilities, there is still a possible room for further research. First of all, we examined only five crosses, but there are many more, which can be more suited to BMA method. Secondly, we used only linear models, but it is possible to implement non-linear models and apply BMA afterwards. In addition, one can also divide the dataset into more subsets, because our dataset contains data from the financial crisis, during which it would be extremely difficult to predict any changes as most of them were unexpected. We implemented here the expanding window forecast, but one can also use the rolling window, where we would not use all previous data for forecast, but only the most recent ones. We have seen that in some specific setting, BMA can beat RW, so finding this setting is crucial for the higher predictive power of this model.

Bibliography

AVRAMOV, D. (2002): "Stock return predictability and model uncertainty." *Journal of Financial Economics* **64(3)**: pp. 423–458.

BATTELLINO, R. (2010): "Twenty years of economic growth." *Structural Change in the Australian Economy 1 Durable Goods and the Business Cycle 11 Economic Change in India 19 Ownership of Australian Equities and Corporate Bonds 25 Interpreting Market Responses to Economic Data 35* p. 103.

BERKOWITZ, J. & L. GIORGIANNI (1996): "Long-horizon exchange rate predictability?" *Review of Economics and Statistics* **83(1)**: pp. 81–91.

CARRIERO, A., G. KAPETANIOS, & M. MARCELLINO (2009): "Forecasting exchange rates with a large bayesian var." *International Journal of Forecasting* **25(2)**: pp. 400–417.

CHEUNG, Y.-W., M. D. CHINN, & A. G. PASCUAL (2005): "Empirical exchange rate models of the nineties: Are any fit to survive?" *Journal of International Money and Finance* **24(7)**: pp. 1150–1175.

CUARESMA, J. C. & J. HLOUSKOVA (2005): "Beating the random walk in central and eastern europe." *Journal of Forecasting* **24(3)**: pp. 189–201.

ECFIN, D. (2002): "Germany's growth performance in the 1990's." *Technical report*, Directorate General Economic and Monetary Affairs (DG ECFIN), European Commission.

ENGEL, C., N. C. MARK, & K. D. WEST (2007): "Exchange rate models are not as bad as you think." *Technical report*, National Bureau of Economic Research.

FERNANDEZ, C., E. LEY, & M. F. STEEL (2001a): "Benchmark priors for bayesian model averaging." *Journal of Econometrics* **100(2)**: pp. 381–427.

Fernandez, C., E. Ley, & M. F. Steel (2001b): "Model uncertainty in cross-country growth regressions." *Journal of applied Econometrics* **16(5)**: pp. 563–576.

Frankel, J. A. & A. K. Rose (1995): "A survey of empirical research on nominal exchange rates." *Technical report*, University of California at Berkeley.

Hakkio, C. (1986): "Does the exchange rate follow a random walk? a monte carlo study of four tests for a random walk." *Journal of International Money and Finance* **5(2)**: pp. 221–229.

Hoeting, J. A., D. Madigan, A. E. Raftery, & C. T. Volinsky (1999): "Bayesian model averaging: a tutorial." *Statistical science* pp. 382–401.

Hoshi, T. & A. Kashyap (2011): "Why did japan stop growing?" *NIRA Report* .

International Monetary Fund, I. (2011): "World economic outlook–tensions from the two-speed recovery: Unemployment, commodities and capital flows." *World Economic and Financial Surveys* .

Jones, R. W., G. M. Grossman, & K. S. Rogoff (1997): *Handbook of international economics*, volume 3. Elsevier.

Kilian, L. & M. P. Taylor (2003): "Why is it so difficult to beat the random walk forecast of exchange rates?" *Journal of International Economics* **60(1)**: pp. 85–107.

Lam, L., L. Fung, & I.-w. Yu (2008): "Comparing forecast performance of exchange rate models." *Technical report*.

Leamer, E. E. (1978): *Specification searches: Ad hoc inference with nonexperimental data*. John Wiley & Sons Inc.

Liu, Y. (2010): "Exchange rate predictability: Bayesian model selection." .

López-Suárez, C. F. & J. A. Rodríguez-López (2011): "Nonlinear exchange rate predictability." *Journal of International Money and Finance* **30(5)**: pp. 877–895.

Mark, N. C. (1995): "Exchange rates and fundamentals: Evidence on long-horizon predictability." *The American Economic Review* pp. 201–218.

MARK, N. C. & D. SUL (2001): "Nominal exchange rates and monetary fundamentals: evidence from a small post-bretton woods panel." *Journal of International Economics* **53(1)**: pp. 29–52.

MEESE, R. & K. ROGOFF (1988): "Was it real? the exchange rate-interest differential relation over the modern floating-rate period." *The Journal of Finance* **43(4)**: pp. 933–948.

MEESE, R. A. & K. ROGOFF (1983): "Empirical exchange rate models of the seventies: Do they fit out of sample?" *Journal of international economics* **14(1)**: pp. 3–24.

MIDA, J. (2013): "Forecasting exchange rates: A var analysis." .

MIN, C.-k. & A. ZELLNER (1993): "Bayesian and non-bayesian methods for combining models and forecasts with applications to forecasting international growth rates." *Journal of Econometrics* **56(1)**: pp. 89–118.

MUCK, J. & P. SKRZYPCZYNSKI (2012): "Can we beat the random walk in forecasting cee exchange rates?" *Available at SSRN 2163518* .

MUSSA, M. (1979): "Empirical regularities in the behavior of exchange rates and theories of the foreign exchange market." In "Carnegie-Rochester Conference Series on Public Policy," volume 11, pp. 9–57. Elsevier.

OBSTFELD, M. (2009): "Time of troubles: the yen and japan's economy, 1985-2008." *Technical report*, National Bureau of Economic Research.

PISSARIDES, C. A. (2006): "Unemployment in britain: a european success story." *Structural unemployment in Western Europe: Reasons and remedies* pp. 209–235.

ROSSI, B. (2006): "Are exchange rates really random walks? some evidence robust to parameter instability." *Macroeconomic dynamics* **10(01)**: pp. 20–38.

ROSSI, B. (2013): "Exchange rate predictability." *Journal of Economic Literature* **51(4)**: pp. 1063–1119.

SARNO, L. & M. P. TAYLOR (2002): *The economics of exchange rates.* Cambridge University Press.

Tortora, A. D. (2010): "Exchange rate forecasting: Bayesian model averaging and structural instability." .

Van Wincoop, E. & P. Bacchetta (2003): "Can information heterogeneity explain the exchange rate determination puzzle?" *Technical report*, National Bureau of Economic Research.

Wooldridge, J. M. (2009): *Introductory econometrics: a modern approach.* South-Western Pub.

Wright, J. H. (2008): "Bayesian model averaging and exchange rate forecasts." *Journal of Econometrics* **146(2)**: pp. 329–341.

Yuan, C. (2011): "Forecasting exchange rates: The multi-state markov-switching model with smoothing." *International Review of Economics & Finance* **20(2)**: pp. 342–362.

Zeugner, S. (2011): "Bayesian model averaging with bms." *Technical report*, mimeo, Available at http://cran. rproject. org/web/packages/BMS/vignettes/bms. pdf.

Appendix A

Summary statistics

EUR/GBP	Maximum	Minimum	Mean	Standard Deviation
Exchange rate return	0.056	-0.033	0.009	0.015
Long-term int. rate	1.798	0.762	1.009	0.178
Short-term int. rate	1.733	0.371	0.736	0.284
GDP	0.231	0.099	0.164	0.039
CPI	0.056	-0.185	-0.067	0.057
Share price	0.240	-0.103	0.033	0.075
IIP	0.022	-0.099	-0.049	0.037
Oil index	2.357	1.339	1.764	0.281
Yield spread	68.750	-38.667	0.571	8.656
Current account	14.000	-36.000	-0.207	5.242

Table A.1: Summary statistics for EUR/GBP currency pair

The table shows the maximum, minimum, mean values and standard deviation for variables used for EUR/GBP forecast.

EUR/JPY	Maximum	Minimum	Mean	Standard Deviation
Exchange rate return	0.063	-0.107	-0.003	0.022
Long-term int. rate	6.712	1.304	2.342	0.924
Short-term int. rate	158.800	0.625	3.154	40.783
GDP	-0.083	-0.194	-0.156	0.023
CPI	0.255	-0.165	-0.015	0.078
Share price	0.111	-0.870	-0.351	0.264
IIP	0.063	-0.091	-0.027	0.035
Oil index	2.357	1.339	1.764	0.281
Yield spread	7.268	-0.635	1.771	8.656
M1	0.003	-0.248	-0.126	0.062

Table A.2: Summary statistics for EUR/JPY currency pair

The table shows the maximum, minimum, mean values and standard deviation for variables used for EUR/JPY forecast.

GBP/JPY	Maximum	Minimum	Mean	Standard Deviation
Exchange rate return	0.050	-0.128	-0.003	0.025
Long-term int. rate	7.237	1.373	2.562	0.904
Short-term int. rate	268.812	1.180	15.82	35.754
GDP	-0.2471	-0.401	-0.3194	0.045
CPI	0.319	-0.143	-0.082	0.098
Share price	0.096	-0.878	-0.384	0.295
IIP	0.085	-0.039	0.021	0.027
Oil index	2.357	1.339	1.764	0.281
Yield spread	5.683	-3.224	0.285	1.565
Labour productivity	0.022	-0.086	-0.022	0.027

Table A.3: Summary statistics for GBP/JPY currency pair

The table shows the maximum, minimum, mean values and standard deviation for variables used for GBP/JPY forecast.

EUR/AUD	Maximum	Minimum	Mean	Standard Deviation
Exchange rate return	0.087	-0.052	0.001	0.023
Long-term int. rate	1.192	0.598	0.846	0.132
Short-term int. rate	1.685	0.063	0.605	0.333
GDP	0.749	0.532	0.652	0.066
CPI	0.205	-0.074	0.087	0.073
Share price	0.339	-0.069	0.108	0.098
IIP	0.167	-0.029	0.079	0.047
Oil index	2.357	1.339	1.764	0.281
Yield spread	77.400	-208.000	-0.074	20.526
M1	0.207	-0.149	0.025	0.102
Current account	1.333	-1.306	-0.085	0.448

Table A.4: Summary statistics for EUR/AUD currency pair

The table shows the maximum, minimum, mean values and standard deviation for variables used for EUR/AUD forecast.

AUD/JPY	Maximum	Minimum	Mean	Standard Deviation
Exchange rate return	0.075	-0.170	-0.003	0.029
Long-term int. rate	8.576	1.157	3.120	1.231
Short-term int. rate	253.965	1.648	22.432	40.761
GDP	-0.671	-0.912	-0.807	0.071
CPI	0.175	-0.332	-0.072	0.129
Share price	0.049	-0.944	-0.459	0.301
IIP	0.070	-0.208	-0.106	0.072
Oil index	2.357	1.339	1.764	0.281
Yield spread	3.707	-5.400	0.117	1.180
M1	0.016	-0.438	-0.150	0.136
Labour productivity	0.056	-0.060	0.00001	0.025

Table A.5: Summary statistics for AUD/JPY currency pair

The table shows the maximum, minimum, mean values and standard deviation for variables used for AUD/JPY forecast.

Appendix B

Results from comparing BMW with HAR

	RMSE				P-values	MAE				P-values
EUR/GBP	0.50	0.25	0.10	0.05		0.50	0.25	0.10	0.05	
Q1	1.11	0.90	0.94	0.96	0.319	1.04	0.91	0.93	0.96	0.113
Q2	2.06	1.68	1.35	1.11	0.961	2.07	1.64	1.29	1.15	0.960
Q4	3.38	3.07	2.42	2.15	0.996	3.42	3.02	2.21	1.89	0.990
Q8	4.77	4.15	3.80	3.60	0.999	5.26	4.64	4.22	3.96	0.999
EUR/JPY										
Q1	1.12	1.06	0.98	0.98	0.820	1.02	1.03	0.99	0.99	0.779
Q2	2.18	2.13	1.98	1.79	0.999	1.73	1.72	1.66	1.58	0.999
Q4	5.50	3.72	2.58	2.10	0.977	2.78	2.45	2.24	2.08	0.999
Q8	2.10	2.02	1.96	1.92	0.999	2.23	2.23	2.20	2.17	1.000
GBP/JPY										
Q1	0.97	0.98	0.99	0.99	0.010	0.95	0.97	0.98	0.99	0.024
Q2	1.23	1.08	1.06	1.05	0.957	1.33	1.08	1.09	1.09	0.951
Q4	4.31	2.91	1.58	1.38	0.946	3.90	2.81	1.68	1.39	0.957
Q8	4.60	4.48	4.28	4.08	1.000	6.33	6.15	5.82	5.51	0.999
EUR/AUD										
Q1	1.02	0.98	0.96	0.96	0.110	0.99	0.98	0.94	0.92	0.042
Q2	1.67	1.65	1.55	1.33	0.997	1.70	1.71	1.59	1.32	0.996
Q4	3.12	3.15	3.23	3.27	1.000	3.38	3.43	3.50	3.55	1.000
Q8	6.01	5.95	3.61	2.37	0.985	6.40	6.14	3.75	2.68	0.987
AUD/JPY										
Q1	0.87	0.92	0.95	0.97	0.023	0.76	0.84	0.91	0.95	0.023
Q2	0.99	1.00	0.99	0.99	0.058	0.88	0.91	0.95	0.98	0.029
Q4	2.05	1.45	1.28	1.25	0.964	1.74	1.35	1.25	1.28	0.981
Q8	2.76	2.68	2.68	2.59	1.000	3.22	3.26	3.37	3.20	1.000

Table B.1: Comparison of BMW model with HAR

The first numeric row includes model prior probabilities. For RMSE and MAE columns, the number larger than 1 indicates that HAR outperformed BMA/BMW and vice versa. For P-values columns the number close to 0 means that we are able to reject the null hypothesis for RMSE/MAE for particular quarter and vice versa.