PREDICTIVE POWER OF THE TERM STRUCTURE

Department of Applied Finance and Actuarial Studies, Faculty of Business and Economics, Macquarie University Sydney, Australia

Dennis Wellmann

MACQUARIE
University

Thesis submitted to the Faculty of Business and Economics, Macquarie University, in fulfilment of requirement for the degree of Doctor in Philosophy

Sydney, October 2016

Declaration

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed.

D. Wellmann

Dennis Wellmann, Sydney 14 October 2016

Contents

List of Figures

List of Tables

1 Abstract

This PhD thesis analyzes the predictive power entailed in the term structure of interest rates and interest rate differentials. It consists of three key chapters based on three research papers.

The first research paper titled '*Forecasting the Term Structure of Interest Rates near the Zero Bound - a New Era?*' investigates the forecasting performance of popular dynamic factor models of the yield curve after the global financial crisis (GFC). This time period is characterized by a low and non-volatile interest rate environment in most major economies with short rates close to the zero bound. We focus on two popular factor models which exploit the information contained in the cross-section of the term structure – the dynamic Nelson-Siegel model and regressions on principal components – to show that subsequent to the GFC both models are significantly outperformed by a random walk no-change forecast. Especially for short and medium term yields, the random walk is up to ten times more accurate. Interestingly, these results are not picked up by traditional global forecast evaluation metrics. We further show that combining forecasts mitigates model uncertainty and improves the disappointing forecasting accuracy especially after the GFC.

The research work '*Factors of the Term Structure of Sovereign Yield Spreads*' investigates the term structure of sovereign yield spreads for five advanced economies against the US and provides novel insights into the key drivers of the spread term structure. We show that the spread term structure dynamics are driven by three latent factors which can be labeled as spread level, slope and curvature, similar to common interpretations found in the yield curve literature. We further show that these estimated spread factors have predictive power for exchange rate movements and excess returns above the predictability of an uncovered interest rate parity approach. As the yield curve contains information about expected future economic conditions, we

conjecture that these yield spread factors reflect expected macroeconomic differentials which in turn drive exchange rates. Using the information content of yield spread curves may thus be a promising approach to improve the forecasting accuracy of exchange rate models.

The third research paper titled '*Exchange Rates and Unobservable Fundamentals: A New Approach to Out-of Sample Forecasting*' builds on the key findings of the previous paper and suggests applying the empirical sovereign yield spread level and slope to forecast exchange rates out-of-sample. Traditional exchange rate models are usually based on differences in observable macroeconomic fundamentals such as output and inflation. However, while being well grounded in economic theory, these fundamental models have a rather poor out-of sample forecasting record. This empirical failure may be a result of the overly restrictive choice of macroeconomic fundamentals. We thus apply the empirical sovereign yield spread level and slope as unobservable proxies of the market's expectations for current and future fundamentals. Our approach outperforms traditional exchange rate models in forecasting accuracy and profitability for all applied forecasting evaluation metrics. It is also superior to a random walk in terms of direction of change forecasts and profitability of the forecasts.

2 Introduction

This thesis deals with potentials and limitations of the predictive power entailed in the term structure of interest rates and interest rate differentials. This introduction describes the concept of the term structure and discusses the different approaches applied in the literature to investigate the predictive content comprised in the term structure. It concludes with the describing the structure of this thesis and the contributions made by each of the research papers to the discipline.

2.1 The Term Structure of Interest Rates and Interest Rate Differentials

The term structure of interest rates denotes the relation between an interest rate and the time to maturity of the underlying instrument. While a term structure theoretically exists for any financial instrument that is available for different maturities, it is often equated with the yields of government bonds ('*yield curve*') which serve as a benchmark for the interest rate level in an economy.[1]

The term structure of interest rates is an important financial and macroeconomic variable. The shape of the yield curve is a central input factor for monetary policy decisions and denotes the cost of funding for individuals and corporations. Accurate forecasts of the term structure of interest rates are crucial in bond portfolio management, derivative pricing and risk management. The yield curve is therefore subject to an enormous body of literature that deals with modeling (Ang and Piazzesi, 2003; Diebold et al., 2005; Gürkaynak et al., 2007; Chen and Niu, 2014) and forecasting (Duf-

[1] In line with existing research, this thesis also uses government bonds zero yields, which are generally referred to when discussing the term structure of interest rates, to proxy interest rates and uses the terms '*yield*' and '*interest rate*' interchangeably.

fee, 2002; Diebold and Li, 2006; Christensen et al., 2011; Laurini and Hotta, 2014) the term structure.

The difference between two government yields of equal maturity is denoted the '*sovereign yield spread*' or '*sovereign spread*'. As sovereign spreads can be calculated for any maturity, they exhibit a term structure or '*sovereign spread curve*' of their own.

Sovereign yield spreads reflect a government's creditworthiness and are key indicators for expected exchange rate movements within the uncovered interest rate parity (UIRP) approach, see, e.g. Sarno (2005); Engel (2013) for recent surveys of the enormous UIRP literature. As they offer a clear picture of the underlying trade-offs for investors, research on sovereign spreads has grown significantly in the past years. These studies mainly focus on the determinants of spreads for emerging economies against the US (Rocha and Garcia, 2005; Hilscher and Nosbusch, 2010) and, recently, spreads within the Eurozone following the advent of the recent European debt crisis (Bernoth and Erdogan, 2012; Maltritz, 2012; Oliveira et al., 2012). However, studies investigating the term structure of sovereign spreads are relatively scarce as the above mentioned research on UIRP and sovereign spreads usually focuses on certain maturities and disregards the dynamics of the term structure.

For both variables – yield curves and yield spread curves – the predictive content entailed in the term structure is of particular importance to academics, market participants and policy makers alike. As the yield curve summarizes expectations about future paths of short interest rates ('*expectations hypothesis*'), it contains information about expected future economic fluctuations (Chinn and Kucko, 2010). The shape and movements of the yield curve have thus long been used to provide readings of market expectations about future macroeconomic fundamentals and the development of financial market. In turn, the term structure of sovereign spreads – being the difference between individual yield curves – naturally contains valuable long term information about expected cross-country differentials which play an important role in exchange rate determination.

This has naturally led to three streams of research investigating the predictive, forward-looking content entailed in the term structure. First, the yield curve forecasting literature uses the information embodied in the cross-section of the yield curve to forecast individual maturities (Duffee, 2002; Diebold and Li, 2006; Pooter et al., 2010; Exterkate et al., 2013). As the different maturities are often influenced by the same factors, exploiting the cross-sectional information entailed in the term structure with dynamic factor models is by now the most popular approach to forecast interest rates (Christensen et al., 2011; Favero et al., 2012; Xiang and Zhu, 2013).

In the second stream, macroeconomic forecasting studies explore the predictive power of the yield curve for macroeconomic variables, in particular output (Ang et al., 2006), recessions (Erdogan et al., 2015), inflation (Rudebusch and Wu, 2008), monetary policy (Heidari and Wu, 2010) and financial crises (Guidolin and Tam, 2013).

The third, relatively young, stream of research investigates the relation between the term structure interest rate differentials and exchange rates. This relation is traditionally expressed in the uncovered interest rate parity (UIRP) condition (Sarno, 2005; Backus et al., 2010; Engel, 2013), which only uses the information content up to a certain maturity. However, the exchange rate is now commonly modeled as an asset price (Mark, 1995; Engel and West, 2005), and thus depends heavily on expected long term macroeconomic fundamentals. Recent studies (see in particular Chen and Tsang (2013) and Bui and Fisher (2016)), argue that these future fundamentals may be reflected more accurately in the entire term structure of interest rate differentials.

This thesis focuses on the first stream of literature – interest rate forecasting – in Chapter 3 as well as the third stream – relation to exchange rates – in Chapters 4 and 5 to investigate the predictive power of the term structure from different perspectives with a special focus on the impact of the recent global financial crisis (GFC) in 2007-2009.

The outbreak of the global financial crisis in 2007 has shaken not only financial institutions, but also long-held beliefs and theories on the behaviour of financial markets (Caprio et al., 2014). It has deteriorated bond markets, both in the US and internationally, where persistently high, often historically abnormal yields and yield spreads have been observed (Contessi et al., 2014; Choudhry, 2015). The GFC thus has had a significant effect on the term structures of interest rates and yield spreads around the globe. It has also triggered sharp and unexpected currency movements, with the US dollar appreciating significantly against virtually all currencies (Fratzscher, 2009). Any research conducted with financial time series encompassing the crises period should therefore carefully take into account the impact of the crisis on results and conclusions. We consider this in particular in Chapters 3 and 5.

2.2 Investigating the Predictive Power of the Term Structure

Predictive power of financial and macroeconomic variables in academic research is usually assessed through in-sample analysis or out-of-sample forecasting exercises.

In-sample analysis means to estimate a model using the entire sample and then comparing the model's fitted values to the actual realizations. It is commonly applied to identify useful predictors with tests of statistical significance which require the common assumptions and restrictions of regression analysis for external validity.

For out-of sample forecasting the sample is split into an estimation period and a forecasting period. The forecasted values of the latter are then compared to the actual realizations and the resulting forecasting errors are used to get an estimate of the model's forecasting accuracy. Out-of sample forecasting exercises are usually implemented to assess whether predictors provide accurate forecasts in environments that mimic the one faced by forecasters in practice as closely as possible and demand a careful selection of appropriate forecasting evaluation methods (Rossi, 2013).

Forecast evaluation generally requires two main choices: which loss function to use, and which test statistic to apply in order to assess the significance of the results. Regarding the choice of loss function, researchers typically evaluate models out-of-sample forecasting performance according to their mean squared forecast error (MSE), the root mean square forecast error (RMSE) or, less frequently, the mean absolute error (MAE), all of which give different weight to the deviations of the forecast from the target. For example, the MSE gives equal weight to over- and underpredictions of the same magnitude. The significance of superior forecast performance is then typically assessed via out-of-sample predictive ability tests such as the tests proposed by Diebold and Mariano (1995), West (1996), Giacomini and White (2006) or Clark and West (2006). Recent studies (Cheung et al., 2005; Moosa and Burns, 2014) have also suggested that the use of criteria based solely on the loss function may not always be appropriate to measure forecasting accuracy. Minimizing the forecasting error based on the global loss function is not necessarily required from an economic standpoint and can miss out on important aspects of forecasts. These aspects may be revealed, for instance, with dynamic forecasting evaluation measures, which expose time-varying forecasting accuracy such as the fluctuations test developed by Giacomini and Rossi (2010). Depending on the predicted variables, alternative approaches to forecast evaluation may also target (Clark and McCracken, 2013): (i) the direction accuracy of the forecast, (ii) the predictive density or interval of the forecast as an indication for the uncertainty associated with the prediction, (iii) utility-based measures or (iv) the profitability of the forecast.

Considering the different underlying assumptions and objectives, the choice between in-sample and out-of-sample forecasting remains a debatable topic. The empirical literature points towards a tendency to find significant evidence of in-sample predictability but not so in out-of sample predictability.[2] Out-of-sample forecasting is thus considered to be a more challenging exercise as predictors which pass in-sample test may still not have predictive ability in a truly out-of-sample forecasting exercise.

Nevertheless, both approaches provide important insights and are used for different objectives depending on the field of research. The yield curve forecasting literature, for example, predominantly uses out-of sample forecasting

[2]This outcome is often caused by overfitting a model – including irrelevant regressors, which improve the in-sample fit of the model but penalize the model in an out-of-sample forecasting exercise.

as external predictors are rarely used and models usually derive the term structure based on the short rate by eliminating arbitrage opportunities or utilize the cross-sectional dependence of the term structure. The exchange rate forecasting literature uses in-sample as well as out-of-sample forecasting techniques alike, as the objective may be either to identify useful explanatory variables or to assess the forecasting accuracy of the predictors.
In this thesis, both in-sample (Chapter 4) and out-of sample (Chapters 3 and 5) forecasting methods are applied to investigate the potential and limitations of the predictive power of yield curves and spread curves. For Chapters 3 and 5 this thesis also carefully considers the impact of different evaluation metrics.

2.3 Content, Structure and Contributions of the Thesis

As indicated, this thesis consists of three key chapters (Chapters 3–5) based on three research papers. These chapters investigate the predictive content of the term structure from different perspectives and contribute to the literature as follows:

Chapter 3 titled '*Forecasting the Term Structure of Interest Rates near the Zero Bound - a New Era?*' investigates the forecasting performance of popular dynamic factor models of the yield curve after the global financial crisis (GFC) – a time period, which is characterized by a low and non-volatile interest rate environment. The results illustrate that subsequent to the GFC period these models are significantly outperformed by a random walk no-change forecast for short and medium term yields. We further demonstrate that these results are not picked up by traditional global forecast evaluation metrics and that this disappointing forecasting performance can be mitigated by combining the forecasts.

The paper thus contributes to the literature in several dimensions. To begin with, it is the first paper to systematically document and explain the poor relative forecasting performance for medium and short term US yields associated with the popular class of dynamic factor yield curve models in the current low interest rate environment. Second, it shows how sensitive the forecasting performance is to the choice of evaluation metrics. This is an important point to consider for future yield curve forecasting studies that will most likely include the unique period after the GFC. Finally, it provides further evidence that combining different models can significantly improve the forecasting accuracy, especially in the current low interest rate environment, where many of the individual models perform rather poorly.

Chapter 4 titled '*Factors of the Term Structure of Sovereign Yield Spreads*' investigates the term structure of sovereign yield spreads which has not been thoroughly studied in the literature yet. We therefore apply principal component analysis (PCA) to examine the latent factors driving the term structure. We then proceed with investigating the predictive power of these factors for exchange rates in-sample. We show that the factors extracted from the term structure of sovereign spreads have predictive power for movements in the exchange rate and excess returns in line with economic intuition and further illustrate that the extracted factors provide additional predictive power in comparison to the traditional UIRP approach.
This research paper contributes to the literature as follows: To start with, this is the first study to thoroughly explore the dynamics of the term structure of yield spreads for advanced economies which has received only limited attention in previous research. Second, it provides novel insights on latent key factors driving the term structure of sovereign spreads. Third, it provides further evidence that the spread term structure reflects macroeconomic fundamentals applied in exchange rate determination and confirms the view that the exchange rate can be modeled as an asset price.

Chapter 5 titled '*Exchange Rates and Unobservable Fundamentals: A New Approach to Out-of-Sample Forecasting*' is built on two main insights of the

previous Chapter 4. First, factors driving the term structure of sovereign yield spreads have predictive power for exchange rates. Second, these factors are highly correlated with the empirical level and slope of the spread term structure. We therefore suggest using the empirical sovereign yield spread level and slope as proxies of the market's expectations for current and future fundamentals in exchange rate out-of-sample forecasting. We find promising results when we investigate the forecasting accuracy of our approach. Applying the yield spread level and slope as a set of proxies for unobservable fundamentals, our model outperforms traditional exchange rate models based on macroeconomic fundamentals such as output and inflation for all considered evaluation metrics. It is also superior to a random walk in terms of direction of change forecasts and profitability.
With this study we contribute to the literature of exchange rate forecasting in the following dimensions: First, we present an innovative, parsimonious, market-driven approach to exchange rate forecasting based on readily and easily available data. This makes it a promising proposition in particular for market practitioners. Second, we provide further evidence that financial variables are useful indicators to be considered in exchange rate forecasting as they are naturally forward looking variables and susceptible to the same macroeconomic risk as exchange rates. Finally, we confirm that the random walk is beatable by models using observable and unobservable models if appropriate statistical and economic evaluation measures are applied.

To conclude, Chapter 6 summarizes the main results, describes the major overall contributions of the thesis to the discipline and identifies a number of directions for future research.

Bibliography

Ang, A., Piazzesi, M., 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. Journal of Monetary Economics 50 (4), 745–787.

Ang, A., Piazzesi, M., Wei, M., 2006. What does the yield curve tell us about GDP growth? Journal of Econometrics 131 (1-2), 359–403.

Backus, D., Gavazzoni, F., Telmer, C., Zin, S. E., 2010. Monetary Policy and the Uncovered Interest Parity Puzzle. NBER Working Paper 16218.

Bernoth, K., Erdogan, B., 2012. Sovereign bond yield spreads: A time-varying coefficient approach. Journal of International Money and Finance 31 (3), 639–656.

Bui, A. T., Fisher, L. A., 2016. The relative term structure and the Australian–US exchange rate. Studies in Economics and Finance (Forthcoming).

Caprio, G., D'Apice, V., Ferri, G., Puopolo, G. W., 2014. Macro Financial Determinants of the Great Financial Crisis: Implications for Financial Regulation. Journal of Banking & Finance.

Chen, Y., Niu, L., 2014. Adaptive dynamic Nelson–Siegel term structure model with applications. Journal of Econometrics 180 (1), 98–115.

Chen, Y.-C., Tsang, K. P., 2013. What Does the Yield Curve Tell Us about Exchange Rate Predictability? Review of Economics and Statistics 95 (1), 185–205.

Cheung, Y.-W., Chinn, M. D., Pascual, A. G., 2005. Empirical exchange rate models of the nineties: Are any fit to survive? Journal of International Money and Finance 24 (7), 1150–1175.

Chinn, M. D., Kucko, K. J., 2010. The Predictive Power of the Yield Curve across Countries and Time. NBER Working Paper 16398.

Choudhry, T., 2015. Time-varying risk premium yield spread effect in term structure and global financial crisis: Evidence from Europe. International Review of Financial Analysis , in press.

Christensen, J. H. E., Diebold, F. X., Rudebusch, G. D., 2011. The affine arbitrage-free class of Nelson-Siegel term structure models. Journal of Econometrics 164 (1), 4–20.

Clark, T. E., McCracken, M. W., 2013. Advances in Forecast Evaluation. In: Elliott, G., Granger, C. W. J., Timmermann, A. (Eds.), Handbook of economic forecasting. Vol. 2. Elsevier, Amsterdam and Boston, pp. 1107–1201.

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

Contessi, S., Pace, P. d., Guidolin, M., 2014. How did the financial crisis alter the correlations of U.S. yield spreads? Journal of Empirical Finance 28, 362–385.

Diebold, F. X., Li, C., 2006. Forecasting the term structure of government bond yields. Journal of Econometrics 130 (2), 337–364.

Diebold, F. X., Mariano, R. S., 1995. Comparing Predictive Accuracy. Journal of Business & Economic Statistics 13 (3), 253–263.

Diebold, F. X., Piazzesi, M., Rudebusch, G. D., 2005. Modeling Bond Yields in Finance and Macroeconomics. American Economic Review 95 (2), 415–420.

Duffee, G. R., 2002. Term Premia and Interest Rate Forecasts in Affine Models. Journal of Finance 57 (1), 405–443.

Engel, C., 2013. Exchange Rates and Interest Parity. NBER Working Paper 19336.

Engel, C., West, K. D., 2005. Exchange Rates and Fundamentals. Journal of Political Economy 113 (3), 485–517.

Erdogan, O., Bennett, P., Ozyildirim, C., 2015. Recession Prediction Using Yield Curve and Stock Market Liquidity Deviation Measures. Review of Finance 19 (1), 407–422.

Exterkate, P., van Dijk, D., Heij, C., Groenen, P. J. F., 2013. Forecasting the Yield Curve in a Data-Rich Environment Using the Factor-Augmented Nelson-Siegel Model. Journal of Forecasting 32 (3), 193–214.

Favero, C. A., Niu, L., Sala, L., 2012. Term Structure Forecasting: No-Arbitrage Restrictions versus Large Information Set. Journal of Forecasting 31 (2), 124–156.

Fratzscher, M., 2009. What explains global exchange rate movements during the financial crisis? Journal of International Money and Finance 28 (8), 1390–1407.

Giacomini, R., Rossi, B., 2010. Forecast comparisons in unstable environments. Journal of Applied Econometrics 25 (4), 595–620.

Giacomini, R., White, H., 2006. Tests of Conditional Predictive Ability. Econometrica 74 (6), 1545–1578.

Guidolin, M., Tam, Y. M., 2013. A yield spread perspective on the great financial crisis: Break-point test evidence. International Review of Financial Analysis 26, 18–39.

Gürkaynak, R. S., Sack, B. P., Wright, J. H., 2007. The U.S. Treasury yield curve: 1961 to the present. Journal of Monetary Economics 54 (8), 2291–2304.

Heidari, M., Wu, L., 2010. Market Anticipation of Fed Policy Changes and the Term Structure of Interest Rates. Review of Finance 14 (2), 313–342.

Hilscher, J., Nosbusch, Y., 2010. Determinants of Sovereign Risk: Macroeconomic Fundamentals and the Pricing of Sovereign Debt. Review of Finance 14 (2), 235–262.

Laurini, M. P., Hotta, L. K., 2014. Forecasting the Term Structure of Interest Rates Using Integrated Nested Laplace Approximations. Journal of Forecasting 33 (3), 214–230.

Maltritz, D., 2012. Determinants of sovereign yield spreads in the Eurozone: A Bayesian approach. Journal of International Money and Finance 31 (3), 657–672.

Mark, N. C., 1995. Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability. American Economic Review 85 (1), 201–218.

Moosa, I. A., Burns, K., 2014. The unbeatable random walk in exchange rate forecasting: Reality or myth? Journal of Macroeconomics 40, 69–81.

Oliveira, L., Curto, J. D., Nunes, J. P., 2012. The determinants of sovereign credit spread changes in the Euro-zone. Journal of International Financial Markets, Institutions and Money 22 (2), 278–304.

Pooter, M. D., Ravazzolo, F., van Dijk, D., 2010. Term structure forecasting using macro factors and forecast combination. Board of Governors of the Federal Reserve System, International Finance Discussion Papers 993.

Rocha, K., Garcia, F. A. A., 2005. The Term Structure of Sovereign Spreads in Emerging Markets. Journal of Fixed Income 14 (4), 45–56.

Rossi, B., 2013. Exchange Rate Predictability. Journal of Economic Literature 51 (4), 1063–1119.

Rudebusch, G. D., Wu, T., 2008. A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy. Economic Journal 118 (530), 906–926.

Sarno, L., 2005. Viewpoint: Towards a solution to the puzzles in exchange rate economics: where do we stand? Canadian Journal of Economics 38 (3), 673–708.

West, K. D., 1996. Asymptotic Inference about Predictive Ability. Econometrica 64 (5), 1067.

Xiang, J., Zhu, X., 2013. A Regime-Switching Nelson-Siegel Term Structure Model and Interest Rate Forecasts. Journal of Financial Econometrics 11 (3), 522–555.

3 Forecasting the Term Structure of Interest Rates near the Zero Bound - a New Era?

Dennis Wellmann (contribution: 70%), Stefan Trück (contribution: 30%)

This paper was presented at:

- 'Quantitative Methods in Finance' (QMF), Sydney, December 2013
- 'AFAANZ Doctoral Symposium' in Auckland, NZ, July 2014
- 'AFAS PhD Day' at Macquarie University, Sydney, October 2014

Abstract

We investigate the forecasting performance of popular dynamic factor models of the yield curve after the global financial crisis (GFC). This time period is characterized by an low and non-volatile interest rate environment in the US and most major economies. We focus on the dynamic Nelson-Siegel model and regressions on principal components and use a dataset of monthly US treasury bond yields to show that subsequent to the GFC both models are significantly outperformed by the random walk no-change forecast. Especially for short and medium term yields the random walk is up to ten times more accurate. Interestingly, these results are not picked up by traditional global forecast evaluation metrics. We show that combining forecasts mitigates the model uncertainty and improves the disappointing forecasting accuracy especially after the GFC.

3.1 Introduction

The global financial crisis (GFC) in 2007-2009 has caused major eruptions in bond and interest rate markets, rendering many traditional yield and bond pricing models useless (Bianchetti, 2010; Walker and McCormick, 2014). The GFC has also led to an prolonged period of low interest rates in several advanced economies subsequent to the crisis. The US is a prime example of this development. Following the expansive monetary policy of the Federal Reserve during the GFC and thereafter, US short and medium term yields have been close to zero since 2009. We test the forecasting accuracy of popular dynamic factor models in this unique interest rate environment and show how interest rates close to the zero bound severely challenge the forecasting accuracy for short and medium yields.

Forecasting the yield curve generally is a strenuous task. Despite major advances in yield curve modeling (Vasicek, 1977; Cox et al., 1985; Nelson and Siegel, 1987; Dai and Singleton, 2000) and forecasting (Diebold and Li, 2006; Exterkate et al., 2013), the high persistence of yields makes it typically hard for any model to outperform a simple random walk no-change forecast (Ang and Piazzesi, 2003; Moench, 2008; Carriero et al., 2012; Xiang and Zhu, 2013). We illustrate in this study how the current interest rate environment further aggravates this challenge.

After the GFC, short and medium yield forecasts of popular dynamic factor models not only fail to beat the random walk but are clearly outperformed in relative terms, with a random walk model being up to ten times more accurate. We also show that the poor forecasting performance in this time period is not reflected in traditional, global forecast evaluation metrics such as the root mean squared error (RMSE) that are typically applied to measure the forecasting performance (Diebold and Li, 2006; Carriero et al., 2012). Hence, this outcome may not be perceived in comprehensive forecasting studies and potentially distort future results and interpretations. Finally, we suggest forecast combination strategies as a mitigating measure and show that combining forecasts significantly improves the forecasting accuracy when individual models perform poorly in the near zero environment.

The unique interest rate environment has already led to an increased interest in modeling the term structure of interest rates near the zero bound. Kim and Singleton (2012) have shown that a quadratic yield curve model better fits the yield curve close to the zero bound than affine term structure models. Krippner (2013) has developed an adjustment to Gaussian models for use in near-zero environments. Christensen et al. (2015) introduce shadow-rate arbitrage-free Nelson-Siegel models designed for nonlinearities near the zero lower bound. Christensen and Rudebusch (2015) apply a shadow-rate arbitrage-free model to U.S. Treasury yields since 1985 and study its performance in both normal times and near the lower bound. Monfort et al. (2015) introduce a novel class of affine term structure models that can generate prolonged spells with the short rate stuck at its lower bound. Filipović et al. (2016) introduce a class of linear-rational term structure models that respect a lower bound and allow for unspanned stochastic volatility.

While this growing stream of literature focuses predominantly on yield curve modeling, the impact of the current interest rate environment on the forecasting accuracy of term structure models has, surprisingly, not thoroughly been investigated and documented in previous research yet. To the best of our knowledge, the only paper focusing on the out-of-sample forecasting accuracy in the recent low interest rate environment thus far is Steeley (2014) who examines several term structure models for UK yields. He finds that a slope regression and the Nelson-Siegel model did rather well before the crises and a simple AR(1) model performs best when short-term rates are near zero, especially for longer horizons. Our study thus aims to contribute to a deeper understanding of the forecasting performance of term structure models in the new interest rate environment near the zero bound.

To investigate the forecasting performance before and after the crisis, we study a dataset of monthly US Treasury bond zero-coupon yields for the time period from January 1995 to December 2013, focusing on the class of dynamic factor models. While they may lack the theoretical foundations of no-arbitrage models, they promise to deliver the most accurate forecasting results as suggested by Duffee (2002, 2011); Chen and Niu (2014)), just to name a few. They are thus also the class predominantly used in recent forecasting studies (Yu and Zivot, 2011; Hautsch and Yang, 2012; Xiang and Zhu, 2013; Laurini and Hotta, 2014; Chen and Niu, 2014). We focus on different variations of the Nelson-Siegel model which imposes a parametric structure on factor loadings, as well as regressions on principal components which extract factors and factor loadings directly from the data. We benchmark the forecasting performance of these models against a random walk no-change forecast. We also include a simple autoregressive AR(1) model as an additional benchmark, since this approach has been reported to forecast the yield curve surprisingly well (Diebold and Li, 2006; Pooter et al., 2010; Steeley, 2014).

The forecasting accuracy is measured with the commonly used root mean squared error (RMSE) and Diebold-Mariano statistics (DM). We also apply innovative dynamic forecast evaluation metrics, such as dynamic relative RSMEs or Giacomini and Rossi's (2010) fluctuation test, which reveal the development of the forecasting accuracy throughout the entire forecasting period.

In our analysis, we use an in-sample period from 1995:1 to 2003:12 to forecast the period from 2004:1 to 2013:12. As we are mainly interested in the forecasting performance when interest rates are close to the zero bound, we investigate the forecasting accuracy for two subsamples. The first one includes the pre-crisis and crises periods (2004:01-2008:12) and the second one comprises the crucial post crises period (2009:01-2013:12).

For the first subsample we find results similar to other pre-crisis forecasting studies such as Pooter et al. (2010) or Yu and Zivot (2011). The selected factor models perform relatively well in terms of the RMSE, especially for short maturities and long forecast horizons. Nevertheless, all models fail to consistently beat the random walk.

However, in the low interest rate environment subsequent to the GFC, the forecasting accuracy for short and medium yields in the second subsample worsens dramatically relative to the random walk. For nearly all maturities below five years the random walk is many times more accurate across all forecasting horizons. For six and twelve-months ahead forecasts the random walk is even up to ten times more accurate. Similar results are also found when comparing the performance of the applied dynamic factor models relative to an AR(1) model.

In addition, Diebold-Mariano statistics show that these results are also statistically significant. In other words, since the end of the GFC, when interest rates dropped to the zero bound, the random walk and a simple AR(1) process have significantly outperformed popular yield curve forecasting models in predicting short- and medium term yields. In the presence of exceptional central bank interference, this crucial time period dominated by low and non-volatile yields obviously favors a random walk model. Nevertheless, the extent of our results is still striking over such a prolonged period.

We find several possible explanations for these results. First, we argue that the cross-sectional structure of popular dynamic factor models, which includes information of the entire term structure to forecast individual yields, fails in an environment with short rates close to the zero bound. As the short yields become more segmented from the rest of the yield curve, the forecasting accuracy for all maturities declines. Furthermore, the applied dynamic models are typically calibrated over a period that also includes significant changes in interest rates as well as in the term structure of the yield curve. As a result, these models are outperformed by a random walk no-change forecast in a low yield environment with hardly any fluctuations for the observed interest rates. Moreover, the models were also estimated during periods when interest rates were significantly higher than during the post GFC period so that forecasts created by the applied models may not only overstate the dynamics of the interest rate term structure but also interest rate levels.

Interestingly, the poor forecasting performance of the applied term structure models in the low interest rate environment is not picked up by commonly used global forecast evaluation metrics, when these are calculated for the entire out-of-sample period. This is a surprising finding, since the critical time period subsequent to the GFC makes up half of the entire forecasting period. Nevertheless, the poor forecasting performance is not reflected in the full sample RMSE's. Further analysis reveals that the forecasting errors during the critical period become relatively small in absolute terms, especially for short and medium term yields and therefore contribute relatively little to the global average. Thus, without a thorough subsample analysis or dynamic forecast evaluation, we may have arrived at entirely different conclusions about the predictive abilities of factor models in the current interest rate environment. This is an important issue to consider for future yield curve forecasting studies and highlights one of the most crucial points of this paper: investigating the global average forecasting performance may fail to reveal important information about the relative forecasting performance over time.

Considering these results, a natural question to be asked is how to approach the unique yield curve dynamics subsequent to the GFC in future forecasting exercises. Different approaches have been developed to account for structural instability in yield curve forecasting. Ang and Bekaert (2002) or Xiang and Zhu (2013), for example, suggest applying regime-switching models that may capture the different interest rate environment. Exterkate et al. (2013) have also shown that including macroeconomic factors may improve the forecasting performance especially in volatile time periods, while gains in the forecasting performance are clearly less significant when volatility is low. Also, the recently developed term structure models for interest rates near the zero bound such as Filipović et al. (2016); Christensen et al. (2015); Christensen and Rudebusch (2015); Monfort et al. (2015) may help to achieve better results for the current low yield environment when applied in future yield curve forecasting studies.

In this study, we suggest the use of forecast combination techniques (Timmermann (2006); Guidolin and Timmermann (2009); Pooter et al. (2010)) as a possible strategy to mitigate the model uncertainty and improve the disappointing forecasting accuracy, especially for the crucial time period after the GFC. We find that simply combining all applied factor models already significantly improves the forecasting accuracy compared to the individual models, albeit this strategy is still outperformed by a random walk. We also combine two diametrically biased variations of the Nelson-Siegel model and a principal component model with an AR(1) model and find that this strategy is able to further improve the poor forecasting performance for shorter maturities after the GFC. It is also interesting to note, that a random walk forecast can be further improved, by combining it with forecasts from selected dynamic factor models. Generally, our results also indicate that performance weighted forecast combination schemes may lead to more accurate forecasts than the equally weighted performance schemes in particular for the more recent low interest rate environment.

With these results, we contribute to the literature on yield curve forecasting in several dimensions. To begin with, this is the first paper to systematically document and explain the poor forecasting performance for medium and short term US yields associated with the popular class of dynamic factor yield curve models in the current low interest rate environment. While we focus on the most popular variations of these models, further research may be required to fully understand how other models perform in this time period.

Second, we show how sensitive the forecasting performance is to the choice of the evaluation metrics. It is still common to select the model with the best global forecasting performance, which in practice amounts to selecting the model that forecasts best on average over the entire out-of-sample period. However, in the presence of time-varying yield curve dynamics, averaging the results over time will result in a significant loss of information. This is an important point to be considered for future yield curve forecasting studies including the unique period after the GFC.

Finally, we provide further evidence that combining different models can significantly improve the forecasting accuracy, especially in the current low interest rate environment, where many of the individual models perform rather poorly. While the forecasting accuracy of the selected models varies heavily over time, forecast combinations are less affected by structural instability than either of the individual models.

The remainder of this chapter is organized as follows: Section 2 provides a review of the relevant yield curve forecasting literature. Section 3 reports descriptive statistics and illustrates the dynamic behavior of yields during the considered sample period. In Section 4 we introduce the selected models, while Section 5 describes the forecasting framework and discusses the out-of-sample forecasting results. In Section 6 we apply different forecast combination strategies and examine whether results can be improved in comparison to using individual models only. Finally, Section 7 concludes and provides suggestion for future work in the area of research.

3.2 Related Literature

The numerous existing term structure models can typically be divided into two streams of literature (Chen and Niu, 2014). The first stream consists of models deriving the term structure based on the short rate, by eliminating arbitrage opportunities between current and future interest rates under various assumptions about the risk premium. Building on the work of Vasicek (1977) and Cox et al. (1981), seminal contributions to the development of these no-arbitrage and affine equilibrium models include Hull and White (1990); Duffie and Kan (1996) and Dai and Singleton (2000). More recent contributions to this stream of literature also relate the short rate to macroeconomic variables (Ang and Piazzesi, 2003; Dewachter and Lyrio, 2006; Rudebusch and Wu, 2008; Moench, 2008). Unfortunately, no-arbitrage and affine-equilibrium models often exhibit poor empirical forecasting performance as pointed out by Duffee (2002, 2011).

The second stream of literature consists of reduced-form models based on more data-driven statistical approaches. This stream has evolved from univariate to multivariate time series models to the class of empirical factor models predominantly used today. Popular univariate models are, for example, the slope regression model, the Fama-Bliss forward rate regression model (Fama and Bliss, 1987) or simple autoregressive processes. The multivariate class includes in particular vector autoregressive (VAR) models and error correction models (ECMs). Unlike univariate models, these models are also able to utilize the cross-sectional dependence structure and cointegration of observed yields at different maturities.

In this study we mainly focus on the class of empirical dynamic factor models that recently have been extensively applied to the modeling and prediction of the yield curve (Christensen et al., 2011; Favero et al., 2012; Exterkate et al., 2013; Xiang and Zhu, 2013). Dynamic factor models allow to model and forecast the term structure based on low-dimensional, latent factors which are extracted from the entire yield curve while retaining the dependence structure of different maturities. The latent factors are usually either estimated by imposing a parametric structure on the factor loadings or extracted directly from the term structure, e.g., by means of a principal component analysis (PCA). While these models may lack the theoretical foundation of the first stream, the empirical literature suggests that they typically provide more accurate forecasts of the yields (Duffee, 2002; Diebold and Li, 2006; Pooter et al., 2010).

Most of the parametric factor models build on the ground-breaking work of Nelson and Siegel (1987) and Diebold and Li (2006). Nelson and Siegel (1987) introduced a parsimonious three-factor model to fit the term structure by using flexible, smooth parametric functions. They demonstrate that their model is capable of capturing most of the typically observed shapes assumed by the yield curve over time. Among the various extensions that have been proposed to incorporate additional flexibility, the most popular one is probably the Svensson (1994) four-factor model. Both, the Nelson-Siegel as well as the Svensson model are heavily used by market practitioners and central banks to construct zero-coupon yield curves, see, for example, Gürkaynak et al. (2007); Coroneo et al. (2011).

Diebold and Li (2006) have extended Nelson-Siegel's initial approach into a dynamic framework enabling the Nelson-Siegel model to be successfully applied to term structure forecasting. Since their seminal study, the literature on forecasting yield curves has grown significantly and in particular their dynamic Nelson-Siegel model has been extended in various ways. Diebold et al. (2006) integrate the initial Diebold and Li (2006) two-step forecasting approach into a single dynamic factor model by specifying the Nelson-Siegel weights as an unobserved vector autoregressive process. Diebold et al. (2008) further extend the initial dynamic Nelson-Siegel model to a global context in which modeling a large set of yield curves allows for global and country specific factors. Almeida and Vicente (2008) explore the role of no-arbitrage restrictions for the forecasting performance and Christensen et al. (2011) develop an arbitrage-free version of the Nelson-Siegel model. Yu and Zivot (2011) include the evaluation of a state-space approach and nine different ratings for corporate bonds. Hautsch and Yang (2012) allow for stochastic volatility of the estimated yield factors, while Xiang and Zhu (2013) develop a regime-switching Nelson-Siegel model. Most recently, Laurini and Hotta (2014) and Chen and Niu (2014) integrate Bayesian estimation methods and adaptive forecasting techniques into the dynamic factor framework.

An alternative non-parametric forecasting approach based on factor dynamics is to apply a PCA to extract the factors directly from the term structure. PCA works best with correlated time series (Duffee, 2012) and is therefore a natural and popular choice to reduce the dimensions of highly correlated yield curve datsets. A small number of orthogonal and uncorrelated factors or principal components can usually already account for a high fraction of variability in relatively high-dimensional datsets. Following Litterman and Scheinkman (1991), several studies apply PCA and find that the variation in interest rates can already be explained by the first three principal components, see, e.g., Bikbov and Chernov (2010); Leite et al. (2010). These three common factors also have an intuitive interpretation as level, slope and curvature[1] based on their effect on the yield curve (Afonso and Martins, 2012) and can be successfully applied in forecasting exercises. Reisman and Zohar (2004), for example, use their forecasting results in bond portfolio selection and suggest that frequent rebalancing leads to substantially higher returns. Blaskowitz and Herwartz (2009) successfully apply PCA to the prediction of the term structure of Euribor swap rates, while Carcano and Dall'O (2011) use error-adjusted principal component analysis to hedge yield curve risk.

While some of the studies described above report superior forecasting results for particular datsets and dynamic factor models are considered to be the most promising class of yield curve forecasting models, the near unit root

[1]The correlations between the first, second and third component and the empirical level, slope and curvature of the US yield data applied in this paper are .99, .77 and .45 respectively and thus confirm these interpretations.

behavior of the yields generally makes it hard for any model to consistently outperform the random walk. As Pooter et al. (2010) show in an extensive forecasting study of US yields, no model clearly performs well across all maturities or different sample periods. Moreover, the forecasting ability of individual models considerably varies over time.

Recent studies have shown that combining the forecasts of different models may mitigate this model uncertainty (Guidolin and Timmermann, 2009). A different approach may also be to include macroeconomic variables into the forecasting procedure. Amongst others, Koopman and van der Wel (2013) and Exterkate et al. (2013) have demonstrated that including macroeconomic variables can significantly improve the forecasting performance for yield term structures, especially during periods of poor forecasting accuracy.

Nevertheless, despite these recent advances, forecasting the yield curve remains a challenging task. In this study we show that forecasting short and medium yields becomes even more arduous in the current low-interest rate environment after the GFC. Modelling the yield curve in this environment has recently been addressed by Kim and Singleton (2012) who analyze the in-sample fit of affine and quadratic yield curve models for Japanese yields and Krippner (2013) who adjusts Gaussian models to be applied in near zero-environments. More recently, Christensen et al. (2015); Christensen and Rudebusch (2015); Monfort et al. (2015); Filipović et al. (2016) have also suggested term structure models specifically designed to capture interest rate dynamics near the zero bound. However, none of these studies considers the out-of-sample forecasting performance of the proposed models. So far, only Steeley (2014) investigates the forecasting ability of different term structure models for UK government yields before and after the crises. He finds that there is little difference in the forecasting accuracy prior to the GFC, especially for short horizons, while a simple AR(1) process is found to be the most accurate after the crisis.

3.3 Data

For our analysis, we use the end-of-month zero-coupon rates of US Treasury bonds obtained from Bloomberg for the sample period January 1995 to December 2013.[2] Selecting US yields is an obvious choice as they have predominantly been used in the literature due to their supreme data quality and availability. The US also is a prime example of an extended period of low and non-volatile interest rates after the GFC. Using monthly observations (n=228), we construct the term structure with 12 maturities ranging from 3, 6, 12, 24, 36, 48, 60, 72, 84, 96, 108 up to 120 months.

Table 3.1 provides the descriptive statistics of the considered dataset. The

Maturity (months)	Mean	St Dev	Min	Max	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(30)$	$\hat{\alpha}(2)$	$\hat{\alpha}(12)$	ADF
3	2.85	2.26	0.02	6.47	0.99	0.74	0.25	-0.27	-0.05	-1.25
6	3.00	2.32	0.04	6.74	0.99	0.74	0.25	-0.35	-0.05	-1.36
12	3.12	2.32	0.11	6.88	0.99	0.75	0.27	-0.27	-0.10	-1.29
24	3.33	2.22	0.22	7.48	0.98	0.76	0.33	-0.18	-0.07	-1.28
36	3.55	2.11	0.28	7.59	0.98	0.77	0.37	-0.16	-0.06	-1.42
48	3.77	2.00	0.44	7.68	0.98	0.77	0.40	-0.12	-0.06	-1.66
60	3.95	1.85	0.62	7.72	0.97	0.76	0.40	-0.10	-0.05	-1.80
72	4.12	1.76	0.79	7.79	0.97	0.75	0.41	-0.10	-0.05	-1.90
84	4.31	1.67	0.97	7.86	0.97	0.74	0.42	-0.10	-0.05	-2.02
96	4.44	1.58	1.17	7.87	0.97	0.74	0.42	-0.10	-0.06	-2.07
108	4.54	1.50	1.37	7.89	0.97	0.73	0.41	-0.10	-0.08	-1.89
120	4.60	1.42	1.60	7.90	0.97	0.70	0.39	-0.08	-0.07	-2.30

Table 3.1. Descriptive statistics for the term structure of US yields for the time period 2000:01 to 2013:12. For each maturity we report (from left to right) mean, standard deviation, minimum, maximum, autocorrelations ($\hat{\rho}$) at displacements of 1, 12, and 30 months, partial autocorrelations ($\hat{\alpha}$) at displacements of two and 12 months and augmented Dickey-Fuller (ADF) test statistics. For the ADF, the critical values for a rejection of the unit root hypothesis are -3.45 at the 1% level (indicated by ***), -2.87 at the 5% level (**) and -2.57 at the 10% level (*). SIC is applied to determine the lag length.

reported characteristics are in line with the stylized facts commonly found in yield curve data, see e.g Diebold and Li (2006); Pooter et al. (2010) or Koopman and van der Wel (2013). The average yield curve during the sample period is upward sloping and concave, volatility is decreasing with maturity

[2] All Bloomberg zero yield curves are constructed daily with government bonds that have Bloomberg Generic (BGN) and/or supplemental proprietary contributor prices. The bonds are adjusted for embedded options and the curve is estimated using a piecewise linear function.

and autocorrelations are very close to unity. The ADF statistics confirm that yields are indeed all but non-stationary. The partial autocorrelation function suggests that autoregressive processes of limited lag order may fit the data well. Correlations between yields of different maturities are not reported here but are typically high, especially for adjacent maturities.

In Figure 3.1, we plot the dynamic behavior for yields of selected maturities. The plot confirms that the yield curve is mostly upward sloping (as indicated by long yields being consistently higher than short yields) with only two short periods of inverted yield curves preceding the two recessions (March - November 2001 and December 2007 - June 2009) after the bursting of the dotcom bubble and the GFC period. These periods also reveal that short and long maturities react quite differently to economic shocks, as both recessions are characterized by a sharp decline in short yields and, thus, an increase in the spread between short and long yields. While this

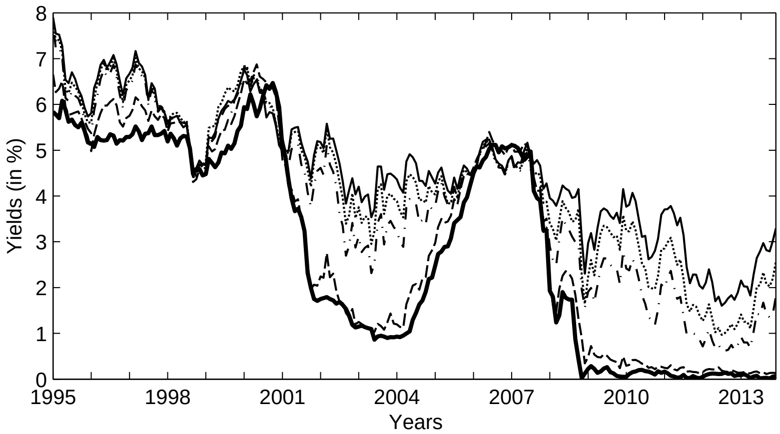

Figure 3.1. Time series of selected US yields. We plot yields for three-month (bold —), 12-month (– –), 36-month (– ·), 60-month (· · ·) and 120-month maturities (—) for the time period 1995:01 – 2013:12.

term spread is generally known to remain rather large for quite some time after a recession, the behavior of short and medium term interest rates, e.g., three months up to 36 months, after the GFC is unprecedented for the US. Following the Federal Reserve Bank's expansive monetary policy in response to the crisis, short yields have remain flat and non-volatile for more than five years up until the end of the sample period. Medium-term yields behave similar, reflecting the Fed's strong commitment to maintaining the expansionary policy as long as required for economic recovery.[3] Assisted by several programs of 'quantitative easing'[4] this has led to an unprecedented, prolonged period of low, non-volatile short and medium US yields. This unique interest rate environment is expected to favor a random walk no-change forecast and we expect that it will pose a peculiar challenge for the forecasting models introduced in the subsequent section.

3.4 Models

In Section 3.2 we have described the numerous empirical factor models that have been developed to model and forecast the yield curve in recent decades. To keep the number of models tractable, we focus on a representative subset of basic models which are commonly used in the academic literature and by practitioners.

In particular we include the dynamic Nelson-Siegel model as one model imposing a parametric structure on factor loadings and regressions on principal components as a model that extracts the loadings and factors directly from the observed term structure. For both models, we apply AR(1) and VAR(1) factor dynamics. While the jointly specified VAR(1) process has the advantage of capturing the interdependence between the derived factors, both approaches have been reported to work well in forecasting exercises, see, e.g. Diebold and Li (2006); Pooter et al. (2010). Furthermore, we include an AR(1) model directly applied on yield levels. AR(1) models can be considered as simple workhorse models and have been reported to fit and forecast yield levels quite well. Thus, the models applied in our empirical analysis are specified as follows:

[3]See, for example, chairman Bernanke's famous quote "The Federal Reserve has done, and will continue to do, everything possible within the limits of its authority to assist in restoring our nation to financial stability", when speaking at the National Press Club in 2009.

[4]The acquisition of financial assets from commercial banks to lower longer yields while simultaneously increasing the monetary base.

Dynamic Nelson-Siegel Model

The Nelson and Siegel (1987) model is a parsimonious three-factor model, based on the three parametric loadings $[1, (1-e^{-\lambda_t \tau})/\lambda_t \tau, ((1-e^{-\lambda_t \tau})/\lambda_t \tau) - e^{-\lambda_t \tau}]$. In the dynamic Nelson-Siegel model (Diebold and Li, 2006) the yield y_t for maturity τ is thus modeled as

$$y_{t,\tau} = \beta_{1,t} + \beta_{2,t} \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} \right) + \beta_{3,t} \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau} \right), \tag{3.1}$$

where $\beta_{1,t}, \beta_{2,t}, \beta_{3,t}$ denote the three latent factors, and the parameter λ controls the exponential decay rate of the second and third loading. In line with Diebold and Li (2006), Diebold et al. (2008) and Chen and Tsang (2013) we fix λ at 0.0609.

To forecast the term structure, we follow Diebold and Li's (2006) two-step approach.[5] First, the Nelson-Siegel factors $\beta_{1,t}, \beta_{2,t}, \beta_{3,t}$ are estimated for the in-sample period applying ordinary least squares. Then the factors are forecasted as autoregressive processes, i.e. for the AR(1) approach each $\hat{\beta}_{k,t+h/t}$[6] is forecasted as

$$\hat{\beta}_{k,t+h/t} = \hat{c}_{k,h} + \hat{\phi}_{k,h} \hat{\beta}_{k,t}, \tag{3.2}$$

where $\hat{c}_{k,h}$ and $\hat{\phi}_{k,h}$ are obtained by regressing $\hat{\beta}_{k,t}$ on an intercept and $\hat{\beta}_{k,t-h}$. For VAR(1) factor dynamics, the three $\hat{\beta}_{t+h/t}$ are forecasted accordingly as

$$\hat{\beta}_{t+h/t} = \hat{c}_h + \hat{\Phi}_h \hat{\beta}_t. \tag{3.3}$$

For both approaches, each individual yield forecast for maturity τ is then given by

$$\hat{y}_{t+h/t,\tau} = \hat{\beta}_{1,t+h/t} + \hat{\beta}_{2,t+h/t} \left(\frac{1 - e^{-\lambda \tau}}{\lambda \tau} \right) + \hat{\beta}_{3,t+h/t} \left(\frac{1 - e^{-\lambda \tau}}{\lambda \tau} - e^{-\lambda \tau} \right). \tag{3.4}$$

[5]Note that we do not apply the one-step state-space framework developed by Diebold et al. (2006) since previous studies, see, e.g., Pooter et al. (2010); Steeley (2014), report no substantial improvement in forecasting accuracy. As the initialization of the one-step state-space estimation further reduces the in-sample size, the two-step approach may even offer a gain in efficiency.

[6]The notation $t + h/t$ denotes that the forecast is made in period t to forecast the value in period $t + h$

In the following we will denote the two approaches by **NSAR** and **NSVAR**.
Regression on principal components
For the PCA approach, each yield is given by the following dynamic latent factor model:

$$y_{t,\tau} = \gamma_{1,\tau}\beta_{1,t} + ... + \gamma_{K,\tau}\beta_{K,t} + \epsilon_{t,\tau}, \tag{3.5}$$

where the $\gamma_{K,\tau}$ describe the K factor loadings and $\beta_{K,t}$ represent K vectors of latent factors.[7] The factors and loadings are estimated with a principal component analysis on the full set of yields for every forecasting iteration.[8] The PCA effectively decomposes the yield covariance matrix as $\Sigma = \Gamma\Lambda\Gamma'$, where the diagonal elements of $\Lambda = diag(\lambda_1, ..., \lambda_K)$ are the eigenvalues and the columns of Γ are the associated eigenvectors. The eigenvectors are arranged in decreasing order of the eigenvalues and the first K eigenvectors of Γ denote the factor loadings $[\gamma_1, ..., \gamma_K]$. The K latent factors $[\beta_1, ..., \beta_K]$ are then constructed by $\beta_{k,t} = \gamma_k' Y_t$, where Y_t is a vector of yields for all maturities at time t.[9] In line with previous research, we decide to use the first three latent factors ($K = 3$), which have been found to be already sufficient to explain a high fraction of the variance in yields (Litterman and Scheinkman, 1991; Bikbov and Chernov, 2010).[10] We apply the two-step forecasting procedure outlined above, forecasting the latent factors $\hat{\beta}_{[1,2,3],t}$ as AR(1) and VAR(1) processes. Thus, h-step ahead yield forecasts for maturity τ are given by

$$\hat{y}_{t+h/t,\tau} = \gamma_{1,\tau,t}\hat{\beta}_{1,t+h/t} + \gamma_{2,\tau,t}\hat{\beta}_{2,t+h/t} + \gamma_{3,\tau,t}\hat{\beta}_{3,t+h/t}. \tag{3.6}$$

In the following, we will refer to these models as **PCAAR** and **PCAVAR**.
Autoregressive (AR(1)) model on yield levels
We also apply an AR(1) model to individual yields of maturity τ directly,

[7] Please note that we use the terms '*factor*' and '*principal component*' interchangeably throughout this analysis.
[8] Note that we standardize the yields with zero mean and unit variance for the PCA.
[9] See also Chapter 4, Section 4.4.1, for a detailed description of the PCA and derivation of the factors.
[10] We also find that for the applied dataset the first three components explain more than 99% of the variation in the term structure.

determining h-step ahead forecasts as

$$\hat{y}_{t+h/t,\tau} = \hat{c}_{\tau,h} + \hat{\phi}_h y_{t,\tau}, \tag{3.7}$$

where $\hat{c}_k$ and $\hat{\phi}_k$ are obtained by regressing $y_{t,\tau}$ on an intercept and $y_{t-h,\tau}$. We denote this model as **AR1**.

Random Walk

As the main benchmark model throughout the forecasting exercise we use a random walk model. In this model any h-step ahead forecast is simply equal to the value observed at time t. Hence the forecast is always 'no change' and can be denoted by

$$\hat{y}_{t+h/t,\tau} = y_{t,\tau}. \tag{3.8}$$

In the following we will refer to the random walk benchmark model as **RW**.

3.5 Out-of-Sample Forecasting

3.5.1 Forecasting Framework and Evaluation

In the following we thoroughly investigate the performance of the applied econometric models in forecasting the US yield curve against a random walk benchmark. For the forecasting exercise, the sample of size T is divided into an in-sample period of length R and an out-of-sample period of length P. We use an initial in-sample period from 1995:1 to 2003:12 to forecast the period from 2004:1 to 2013:12. Thus, the in-sample period includes the bursting of the dotcom bubble and the subsequent recession and recovery, while the out-of-sample period includes the GFC as well as pre- and post crisis periods. The considered sample period allows us to have enough observations to estimate the parameters of the models with sufficient accuracy and still evaluate the forecasting performance over sufficiently long (sub)periods with different yield curve dynamics. As we are mainly interested in the forecast-

ing accuracy in the low interest rate environment after the GFC, we also cut the forecasting period in half for the forecast evaluation. This provides us with two subsamples, 2004:01-2008:12 (including the pre-crisis and the crises period) and 2009:01-2013:12 (the crucial post crises period).
We apply a recursive window such that in each time step the in-sample period is extended by one observation to calculate the forecasts for t+h. In particular, we create one-month ($h = 1$), six-months ($h = 6$) and twelve-months ($h = 12$) ahead forecasts whereas all models are forecasted iteratively.[11]
To assess the forecasting accuracy, we report the commonly applied root mean squared error (RMSE) and Diebold-Mariano (DM) statistic.[12] The RMSE is a measure of global forecasting performance and summarizes the forecasting errors over the entire forecasting period. For each considered model m, maturity τ and forecasting horizon h the RMSE for the forecasting period P is calculated as

$$RMSE^{m}_{\tau,h} = \sqrt{\frac{1}{P}\sum_{t=1}^{P}(\hat{y}^{m}_{t+h/t,\tau} - y_{t+h,\tau})^2}. \tag{3.9}$$

The lower the RMSE the more accurate the forecast. However, a smaller RMSE in a particular sample of forecasts does not necessarily mean that the corresponding model is truly better in population. Diebold and Mariano (1995) address this concern and propose a test to assess the statistical significance of predictive superiority. The Diebold-Mariano (DM) test statistic

[11]It is still being debated whether iterated or direct forecasts are more accurate. Carriero et al. (2012), for example, find that the iterated approach produces more accurate forecasts in yield curve forecasting. Comparing both approaches we also find better results for the iterated approach, albeit not by much, and henceforth apply it throughout the analysis. We also note that our main results and conclusions do not change when a rolling instead of a recursive window approach is applied.

[12]While theoretically the mean average error (MAE) or mean error (ME) could also be applied, the RMSE is the main forecasting evaluation metric used in the yield curve forecasting literature (see, e.g., Diebold and Li (2006); Diebold et al. (2008); Chen and Tsang (2013); Koopman and van der Wel (2013). We thus follow the literature and focus on the RMSE to make the results comparable with past and future forecasting studies.

is calculated as

$$DM^m_{\tau,h} = \frac{\bar{d}}{\sqrt{\widehat{LRV}_{\bar{d}}/P}}, \tag{3.10}$$

where $\bar{d}$ is the average difference d between the loss functions of two competing forecast models $m = 1, 2$ given as

$$\bar{d} = \frac{1}{P}\sum_{t=1}^{P} d_t = \frac{1}{P}\sum_{t=1}^{P} L(\varepsilon_{1,t}) - L(\varepsilon_{2,t}), \tag{3.11}$$

with an applied loss function $L(\varepsilon_{m,t}) = (\varepsilon_{m,t})^2 = (\hat{y}^m_{t+h/t,\tau} - y_{t+h,\tau})^2$ for model m, maturity τ and forecasting horizon h.[13] $\widehat{LRV}_{\bar{d}}$ denotes the HAC estimator of the asymptotic (long-run) variance of $\bar{d}$ given by

$$\widehat{LRV}_{\bar{d}} = \gamma_0 + 2\sum_{j=1}^{\infty} \gamma_j, \tag{3.12}$$

where $\gamma_0 = var(\bar{d})$ and $\gamma_j = cov(d_t, d_{t-j})$ and the lag length being $h - 1$. The null hypothesis is equal predictive accuracy of the considered models. Note that we will conduct two-sided tests, since we are interested in both, statistically significant superior and inferior forecasting performance of the selected models against a random walk benchmark.

3.5.2 Forecasting Results

Table 3.2 presents the forecasting results for the two subsample periods. In the first line of each table, we report the RMSE of the random walk expressed in basis points. We then report the RMSEs of all models relative to the random walk. Hence, numbers smaller than one (**reported in bold**) indicate that a model outperforms the random walk. The significantly better forecasting performance of a model against the random walk benchmark based on conducted DM tests is indicated by ("), while we indicate the significantly

[13]Note that in our analysis we apply the commonly used quadratic loss function. However, theoretically Diebold and Mariano (1995) do not limit the loss functions that could be used.

inferior performance of a model against the random walk by (∗).[14]
For the first subsample from 2004:01-2009:12 results are reported in the upper panel of Table 3.2. We find roughly similar outcomes to previous comprehensive forecasting studies, see, e.g., Pooter et al. (2010); Yu and Zivot (2011): in relative terms, the applied factor models perform relatively well, especially for shorter maturities. Nevertheless, all models fail to consistently beat the random walk - not a single model clearly performs well across all maturities and forecast horizons. The Diebold-Mariano statistics reported in Appendix A.3 confirm that, despite two exceptions, no model is able to significantly outperform the random walk at any maturity. In absolute terms the RMSEs are usually smaller for longer term maturities and the forecasting performance worsens with longer forecasting horizons. Generally, the absolute RMSEs of the subsample are relatively high, as all models and the random walk struggle to predict the sudden drop in yields during the GFC. Comparing the different models, we find no compelling difference in forecasting accuracy between the factor models. The AR(1) process performs surprisingly well and is on par with the factor models for most maturities and forecast horizons.
Results for the crucial subsample period after the GFC (2009:01-2013:10) are reported in the lower panel of Table 3.2 and indicate striking differences between the forecasts generated by the examined models. In absolute terms, the RMSE drops notably for all models and maturity horizons. In relative terms, the forecasting accuracy of the selected dynamic factor models for short and medium term yields worsens significantly compared to the random walk. For some of the models, calculated RMSEs are even more than ten times higher than for the random walk forecasts. The poor forecasting performance of the considered models relative to the random walk is particularly pronounced for shorter and medium maturities, i.e. three-month, six-month and twelve-month yields.
Moreover, the conducted Diebold-Mariano tests show that the outperformance by the RW are statistically significant for many maturities and forecasting horizons, often even at the 1% level. Further analysis confirms that

[14]Tests are conducted at the 5% level of significance. Detailed results and test statistics for the conducted DM tests are reported in Appendix A.3.

similar results are obtained when the forecasting performance of the factor models is compared to the performance of the simple AR(1) models applied directly to the yield levels.[15] In other words, after the GFC the random walk and a simple first order autoregressive process are able to significantly outperform all considered dynamic factor model variations.

We further explore this outcome by investigating the intertemporal development of the forecasting accuracy throughout the forecasting period. Therefore, we construct sequences of local relative RMSEs. We define a dynamic relative RSME as the sequence of local relative RMSEs over centered rolling windows of size r (we choose $r = 24$) for $p* = R + r/2, ..., T - r/2 + 1$. The intention of this measure is to examine the evolution of the models' relative forecasting performance through time.

For each model, the local RMSE for the respective rolling window is given by

$$RMSE^{m,local}_{p*,\tau,h} = \sqrt{\frac{1}{r} \sum_{j=p*-r/2}^{p*+r/2-1} (\hat{y}^m_{j+h/t,\tau} - y_{j+h,\tau})^2}. \tag{3.13}$$

We then express the sequence of local $RMSE^{m,local}_{p*,\tau,h}$ for all models relative to the random walk local $RMSE^{RW,local}_{p*,\tau,h}$ sequence. As indicated above, values smaller than one indicate that a model outperforms the random walk, while values larger than one indicate inferior forecasting performance against the random walk. In Figure 3.2 we plot the series of local relative RMSEs for a 12-months forecasting horizon and selected short (3-months; 12-months), medium (60-months) and long term (120-months) maturities. The dynamic forecast evaluation confirms that prior and during the GFC all models compete relatively closely, with the random walk for all maturities with some periods of superior and some periods of inferior performance. For the medium and long term yield, the dynamic relative RMSE also remain relatively close together throughout the entire forecasting period. However, for the short and medium term yields, things change dramatically subsequent to the GFC period (also note the different scale of the y-axis for the top and bottom plots). The forecasting accuracy worsens significantly in relative terms for

[15]Detailed results are not reported here, but are available upon request to the authors.

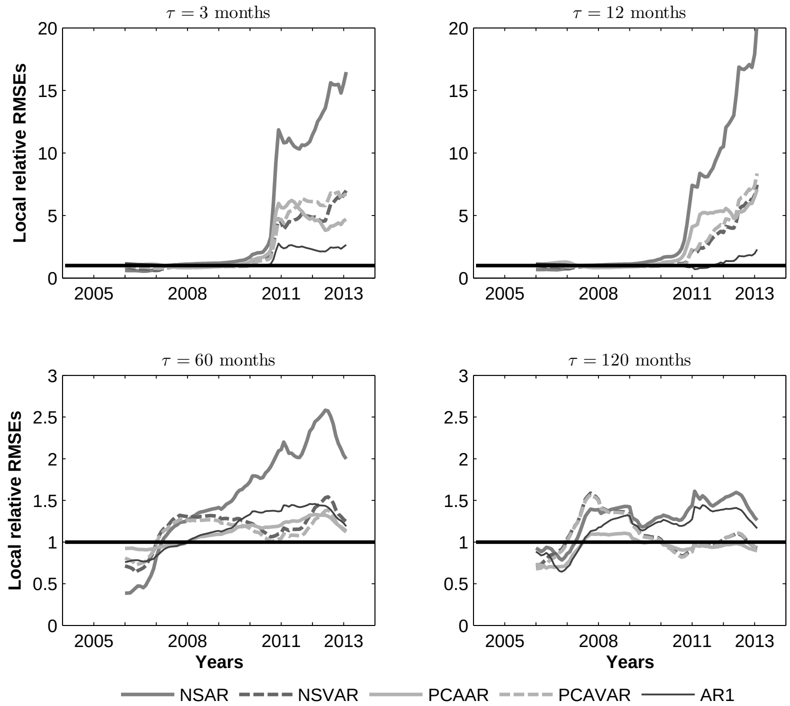

Figure 3.2. Dynamic relative three-month, twelve-month, 60-month and 120-month yield RMSEs for all models against the random walk for a $h = 12$ months forecast horizon. The dynamic RMSEs are calculated as follows: For each model and the random walk we calculate sequences ($p* = R + r/2, ..., T - r/2 + 1$) of local RMSEs for rolling windows of size r=24 throughout the the forecasting period from 2004:01 - 2013:12. We then calculate the dynamic relative RMSE by expressing the sequence of local $RMSE^{m}_{p*,\tau,h}$ for each model relative to the random walk local $RMSE^{RW}_{p*,\tau,h}$ sequence. Hence, values smaller than one indicate that a model outperforms the random walk, while values larger then one indicate inferior forecasting performance against the random walk. **Note the different scale of the y-axis for top and bottom plots**.

all factor models. The Nelson-Siegel model with $AR(1)$ factor dynamics performs particularly badly. While the $AR(1)$ process performs better than the factor models, it is still consistently dominated by the random walk from 2010 onwards.

We thus take a closer look at a time series of short term yields forecasts in Figure 3.3, where we provide an illustrative plot of the $h = 12$ months ahead forecasts for 3-months yields.[16] For ease of presentation we focus on the random walk, the NSAR, PCAAR and the AR1 model. The plot confirms that

[16]We note that plots and conclusions are similar for other short and medium term yields up to 36 months. We also present additional plots for twelve-month, 60-month and 120-month yield forecasts in Appendix A.1.

the selected dynamic factor models produce rather poor forecasts for the period after January 2010. While the AR(1) model and the random walk adapt rather quickly to the changed environment, both factor models, in particular the parametric Nelson-Siegel model with $AR(1)$ factor dynamics, continuously over- or under-predict the actual yields. Only the PCAAR model picks up the new interest rate environment towards the end of the period. It is also important to note that at times all models predict negative yields when the actual yield is close to the zero bound. This is a highly undesired effect for many pricing and hedging purposes and confirms that models of interest rate dynamics may need to be revised in the current environment as suggested by, e.g., Krippner (2013); Christensen et al. (2015); Monfort et al. (2015); Filipović et al. (2016). Overall, we obtain quite different results for the two

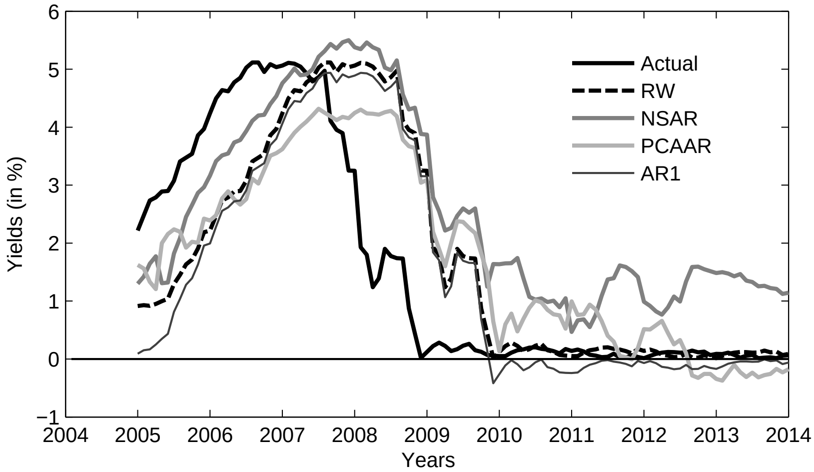

Figure 3.3. 3-months yield forecasts for random walk, NSAR, PCAAR and AR1 model. We plot the actual three-months yields together with the forecasts of the selected models for a forecasting horizon of $h = 12$ months.

subsample periods. While the forecasting accuracy for the first subsample period is in line with results reported in earlier studies, factor models perform rather poorly, when short rates are close to the zero bound.

Given the different interest rate regimes during the two subsample periods this outcome is not entirely surprising. Obviously, periods with hardly any volatility in interest rates will favor random walk no change forecasts. However, the dimension of relative outperformance of a random walk model over

the applied factor models over such an extended period is still striking.
Our results also point towards methodological problems of the applied econometric factor models, when short and medium yields are relatively constant and close to the zero bound. Due to the cross-sectional structure of dynamic factor models, these models consider information based on the entire yield curve for forecasting individual yields. However, during the post GFC low yield environment, short yields become more segmented from the rest of the curve and this will worsen the predictive accuracy of such models.
Another reason for the striking difference in forecasting performance may be that the models are calibrated over a time period that also includes a dynamic behavior of the term structure of the yield curve as well as significant changes in interest rates for given maturities. The estimated models may then overstate the dynamics of individual yields as well as for the entire yield curve during the unique low interest rate period from 2009 to 2013. Further, since the models are estimated during periods when short-term interest rates were significantly higher than after the GFC, created forecasts may not only overstate the dynamics of the interest rate term structure but possibly also the level of short-term interest rates.[17]
Finally, the poor forecasting performance in the second subsample may also be due to increased government interventions during this period of market turmoil. This would obviously be difficult to be picked up by any forecasting model.

[17]As a robustness check we also implemented the forecasting exercise with rolling windows, However, we do not find a significant change in results and conclusions.

		3m	6m	12m	2y	3y	5y	7y	10y
		2004:01 - 2008:12							
	RW	30.8	30.1	30.5	31.4	32.3	30.8	32.7	27.9
h=1	NSAR	1.26*	**0.92"**	**0.91**	1.04	1.08	1.11	1.07	1.01
	NSVAR	1.06	**0.83**	**0.93**	1.04	1.05	1.09*	1.05	1.00
	PCAAR	1.14	**0.95**	**0.99**	1.09*	1.05	1.02	1.04	1.01
	PCAVAR	**0.98**	**0.82"**	**0.96**	1.07	1.03	1.02	1.06	1.02
	AR1	1.01	1.01	1.01	1.01	1.01	1.01	1.02	1.02
	RW	113.0	114.0	108.5	97.8	92.3	72.9	65.9	53.7
h=6	NSAR	1.02	**0.95**	**0.93**	**1.00**	1.01	1.09	1.10	1.08
	NSVAR	**0.92**	**0.92**	**0.96**	1.04	1.05	1.11	1.10	1.07
	PCAAR	**0.97**	**0.98**	1.00	1.03	1.00	1.00	1.00	**0.97**
	PCAVAR	**0.92**	**0.97**	1.02	1.07	1.05	1.07	1.10	1.06
	AR1	1.04	1.04	1.03	1.00	**0.99**	**0.98**	1.02	1.04
	RW	207.0	204.3	186.6	156.8	136.5	100.0	83.9	63.6
h=12	NSAR	**0.97**	**0.93**	**0.92**	**0.99**	1.01	1.10	1.15	1.21
	NSVAR	**0.90**	**0.90**	**0.95**	1.05	1.08	1.16	1.17	1.21
	PCAAR	**0.87**	**0.90**	**0.93**	**0.98**	**0.98**	**0.99**	1.00	**0.98**
	PCAVAR	**0.92**	**0.96**	1.01	1.08	1.10	1.14	1.18	1.19
	AR1	1.05	1.05	1.03	**0.98**	**0.96**	**0.96**	1.02	1.08
		2009:01 - 2013:12							
	RW	4.1	4.7	7.4	12.9	18.0	24.6	28.1	28.7
h=1	NSAR	5.83*	2.22*	1.16	1.06	1.31*	1.35*	1.09	1.04
	NSVAR	2.95*	2.34*	2.45*	1.30*	1.06*	1.19*	1.03	1.07
	PCAAR	1.93*	1.25	1.28*	1.30*	1.10*	1.03	**0.99**	**0.99**
	PCAVAR	1.82*	2.22*	1.69*	1.21*	1.03	1.05	1.01	**0.98**
	AR1	1.11	1.08	1.02	1.01	1.01	1.01	1.00	1.00
	RW	7.9	9.4	9.9	22.4	38.9	63.7	75.9	76.2
h=6	NSAR	9.19*	6.74*	5.87*	2.91*	2.10*	1.47	1.18	1.03
	NSVAR	4.35*	3.40*	2.67*	**0.92**	1.04	1.04	**0.95**	**0.89**
	PCAAR	3.57*	3.07*	3.19*	1.84*	1.34*	1.04	**0.95**	**0.91"**
	PCAVAR	4.32*	3.59*	2.87*	1.05	**0.98**	**0.96**	**0.94**	**0.89**
	AR1	1.59	1.33	1.04	1.07	1.09	1.07	1.06	1.04
	RW	8.9	10.5	11.9	26.2	43.8	74.5	92.7	93.6
h=12	NSAR	13.15*	10.65	9.41	4.76	3.24*	2.03*	1.55*	1.31
	NSVAR	5.43	4.27	3.14*	1.37	1.36*	1.18	1.02	**0.92**
	PCAAR	5.28*	4.71*	4.47*	2.33*	1.65*	1.17	**1.00**	**0.91**
	PCAVAR	5.46*	4.52*	3.43*	1.42	1.24	1.08	**1.00**	**0.91**
	AR1	2.50*	2.00	1.10	1.19	1.28*	1.27*	1.24	1.20

Table 3.2. Subsample forecasting results for h=1, h=6 and h=12 months ahead forecasting horizons and three-month, six-month, twelve-month, two-year, three-year, five-year, seven-year and ten-year maturities. We report the root mean squared error (RMSE) for the subsample periods from **2004:01-2008:12** and **2009:01-2013:12**. The first line reports the RMSE for the random walk (expressed in basis points). The RMSEs of all other models are expressed relative to the random walk. Hence, numbers smaller than one (**reported in bold**) indicate that models outperform the random walk. Numbers larger than one indicate inferior performance. (") indicates statistical significant forecasting superiority of the respective models against the random walk measured by the DM-statistic on a 5% or smaller significance level. (*) indicates statistical significant forecasting inferiority against the random walk. The DM-statistics are reported in Appendix A.3. See Section 3.4 for a description of the selected models.

3.5.3 Sensivity of Results Towards Forcast Evaluation Metrics

We are also interested in how the dismal forecasting performance of the second subsample period is reflected in the forecasting accuracy calculated over the entire out-of-sample period. As the time period of the GFC and its aftermath make up nearly half of the forecasting period, we would also expect a rather poor overall forecasting performance of the applied dynamic Nelson-Siegel and PCA models.

Table 3.3 presents the forecasting results for the entire out-of-sample forecasting period from 2004:01 to 2013:12. To our surprise, we find results broadly in line with the first subsample period and previous research. The selected factor models perform relatively well in terms of the RMSE, especially for short maturities and long forecast horizons. Nevertheless, all models fail to consistently beat the random walk.

These results obviously raise the question why the poor relative forecasting performance for short and medium term yields subsequent to the GFC is not fully reflected in the results reported for the entire forecasting period. After all, the crucial time period makes up nearly half of the out-of-sample period. This will also be highly important for future yield curve forecasting studies which will most likely comprehend the time period subsequent to the GFC.

The answer can be found in the decreasing magnitude of forecasting errors caused by the low interest rate environment after the GFC. Not surprisingly, with low short and medium yields close to the zero bound, forecasting errors and the RMSE drop significantly in absolute terms. This is illustrated in Figure 3.4 where we plot forecasting errors for three-month yields for a $h = 12$ months forecast horizon for the random walk as well as for the NSAR and PCAAR model. Usually, forecast errors for shorter maturities are relatively large as interest rates for shorter maturities are rather volatile. Figure 3.4 confirms that this is also observable in the forecasting errors at the beginning of the forecasting period. The plot also illustrates how the forecasting errors for short and medium yields become rather small in absolute terms subse-

		3m	6m	12m	2y	3y	5y	7y	10y
	RW	21.9	21.5	22.1	24.0	26.1	27.7	30.0	27.9
	NSAR	1.47*	**0.97**	**0.92**	1.04	1.14*	1.21*	1.07	1.03
	NSVAR	1.12	**0.90**	1.07	1.08*	1.06	1.13*	1.04	1.04
h=1	PCAAR	1.16*	**0.95**	1.01	1.13*	1.06*	1.02	1.01	1.00
	PCAVAR	1.00	**0.88**	1.02	1.09*	1.03	1.03	1.04	**1.00**
	AR1	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01
	RW	81.1	82.6	79.8	72.7	72.4	70.4	72.5	66.9
	NSAR	1.23	1.11	1.08	1.20	1.26*	1.28*	1.16	1.06
	NSVAR	**0.99**	**0.97**	**0.98**	1.05	1.06	1.09	1.02	**0.97**
h=6	PCAAR	1.03	1.03	1.04	1.09	1.07	1.03	**0.98**	**0.94"**
	PCAVAR	**0.98**	1.01	1.03	1.07	1.05	1.04	1.02	**0.96**
	AR1	1.04	1.03	1.03	1.01	1.01	1.04	1.05	1.05
	RW	145.0	144.0	133.4	112.5	100.7	86.7	85.7	78.1
	NSAR	1.17	1.12	1.14	1.30	1.43	1.55*	1.42	1.31
	NSVAR	**0.95**	**0.95**	**0.98**	1.08	1.14	1.20	1.12	1.04
h=12	PCAAR	**0.95**	**0.97**	**1.00**	1.06	1.09	1.09	1.02	**0.94**
	PCAVAR	**0.96**	**0.99**	1.03	1.10*	1.13	1.15	1.11	1.03
	AR1	1.05	1.04	1.03	1.00	1.02	1.12	1.18	1.20

Table 3.3. Forecasting results for $h = 1$, $h = 6$ and $h = 12$ months ahead forecasting horizons and three-month, six-month, twelve-month, two-year, three-year, five-year, seven-year and ten-year maturities. We report the root mean squared error (RMSE) for the out-of-sample period **2004:1 - 2013:12** (N = 96). The first line reports the RMSE for the random walk (expressed in basis points). The RMSE of all other models is expressed relative to the random walk. Hence, numbers smaller than one (**reported in bold**) indicate that models outperform the random walk. Numbers larger than one indicate inferior performance. (") indicates statistical significant forecasting superiority of the respective models against the random walk measured by the DM-statistic at a 5% level of significance, while (*) indicates statistical significant forecasting inferiority against the random walk. The DM-statistics are reported in Appendix A.3. See Section 3.4 for a description of the selected models.

quent to the GFC. Thus, the poor relative forecasting performance after the GFC vanishes in global forecast evaluation measures averaged over the entire forecasting period. The RMSE being based on a quadratic loss function further aggravates this effect.

Interestingly, the unique behaviour of yields after the GFC also poses a challenge for dynamic forecasting measures relying on absolute differences in forecasting errors. As an additional evaluation metric we apply the innovative fluctuation test developed by Giacomini and Rossi (2010). This test allows to examine the statistical significance of the forecasting performance

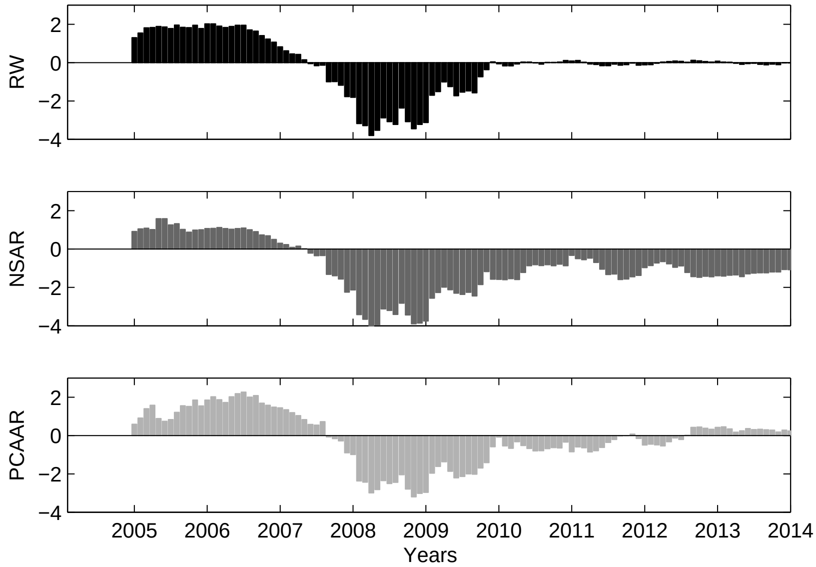

Figure 3.4. Forecasting errors (3-months actual yield - 3-months forecasted yield) for the random walk, NSAR and PCAAR model. We plot the difference between actual three-month yields and $h = 12$ months horizon forecasts for the time period from 2004:01-2013:12.

over time. The sequence of local test statistics is calculated based on a local loss function differential ΔL_j computed over centered rolling windows of size r and given as

$$F^{m}_{p*,\tau,h} = \hat{\sigma}^{-1} r^{-1/2} \sum_{j=p*-r/2}^{p*+r/2-1} \Delta L_j(\hat{y}^{RW}_{j+h/j,\tau}, \hat{y}^{m}_{j+h/j,\tau}), \tag{3.14}$$

where $\hat{\sigma}^2$ is a HAC estimator of the asymptotic (long-run) variance. The test statistic $F^{m}_{p*,\tau,h}$ is equivalent to the one proposed by Diebold and Mariano (1995) computed over rolling windows. Giacomini and Rossi (2010) also provide critical values to test the null hypothesis of equal predictive accuracy. In Figure 3.5 we plot the fluctuation test statistics for the three-month, twelve-month, five-year and ten-year yields and a forecast horizon of $h = 12$ months based on rolling windows of size $r = 24$ with corresponding two-sided critical values.[18] The fluctuation test correctly reflects the direction of the

[18]Note that unlike the forecasting exercises conducted in previous sections that were based on a recursive window estimation, fluctuations tests were conducted based on a rolling

superior and inferior performance of the models relative to the random walk. However, none of the local test statistics for the short term maturities $\tau = 3$ and $\tau = 12$ months post the GFC indicates a statistically significant outperformance of the other models by the random walk. This is surprising given the statistically significant superior performance of the random walk for the second subsample period reported in the lower panel of Table 3.2 and the illustration of these results in Figure 3.2. Further analysis reveals that the decreasing loss functions post the GFC distort the local test statistics calculated based on the global estimator of the asymptotic (long-run) variance $\widehat{LRV}_{\bar{d}}$. This confirms the observation of Martins and Perron (2012) who find power problems of the fluctuation test in the presence of instabilities in the differences of the loss functions.

Overall, these findings indicate that typically applied performance measures such as RMSE or Diebold-Mariano tests may entirely fail to reflect the superior or inferior forecasting performance of a model during a certain subsample period. In our study, the RMSE does not provide any valuable insights for which particular time periods the models perform well, since it only measures the global forecasting performance over the entire out-of-sample period. Thus, information about the dynamic performance of the models throughout the forecasting period is lost. Our findings also illustrate that even tests designed to examine the dynamic forecasting performance of competing models over time, e.g. the fluctuation test proposed by Giacomini and Rossi (2010), will not necessary detect the extent of the inferior performance of a subset of the applied models.

This highlights one important point of this paper: In the presence of time-varying yield curve dynamics, averaging the results over time in global forecasting measures such as the RMSE may result in a significant loss of information and may lead to false conclusions with regards to the true forecasting capabilities of a model. This is an important point to consider for future yield curve forecasting studies most likely including the unique period after the GFC.

---

window estimation as proposed by Giacomini and Rossi (2010). As noted above, the forecasting results do not differ much for rolling and recursive windows.

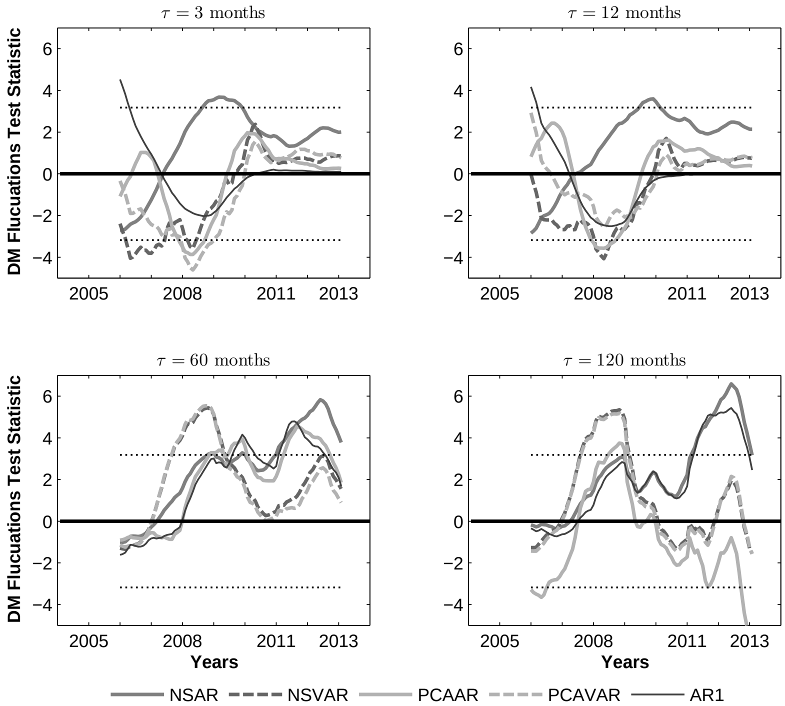

Figure 3.5. Fluctuations test statistics for all models against the random walk for three-month, twelve-month, five-year and ten-year maturities and $h = 12$ month forecast horizon. The $p* = R+r/2...T-r/2+1$ sequence of fluctuations test statistics is calculated based on rolling windows of size r=24 throughout the the forecasting period from **2004:01 - 2013:12**. Values smaller than zero zero indicate that models outperform the random walk. Values larger then zero indicate inferior forecasting performance against the random walk. Values larger/smaller than the critical values indicated statistically significance on a 5% level. The critical values [3.01; −3.01] are obtained from Giacomini and Rossi (2010).

3.6 Forecast Combinations

3.6.1 Methodology and Strategies

A natural question to ask is how to best approach the instability in the relative performance of the selected models. Previous research discusses several interesting measures to approach unstable forecasting environments, for example adaptive forecasting techniques (Chen and Niu, 2014) or regime switching models (Xiang and Zhu, 2013). A promising approach advocated in the recent literature is also to combine the forecasts of individual models. Several studies (Guidolin and Timmermann, 2009; Pooter et al., 2010) have shown that combining multiple forecasts may increase the forecasting accuracy for interest rates. This approach is particularly promising in our case, since the forecasting accuracy of our selected models heavily varies over time, often diametrically. Thus, combined forecasts are likely to be more robust to structural instability than either of the individual models.

In this section, we therefore investigate different combinations of individual forecasts in order to improve the forecasting accuracy particularly for the crucial time period after the GFC. We consider three different forecast combination strategies. The first simply includes all four factor models (NSAR, NSVAR, PCAAR, PCAVAR) and is denoted as '**factor**'. The second includes the NSAR, PCAAR as well as the AR1 forecasting models and is labeled '**far1**'. This strategy is motivated by the graphical analysis in Figure 3.3, which illustrates that the NSAR and PCAAR models seem to be diametrically biased in their forecasts of short and medium term yields after the GFC. Combining both models should thus improve the individual forecasts. Including the AR1 forecasts is an obvious choice as the AR(1) process performs rather well during the post GFC period in comparison to the applied factor models. For the third strategy we combine forecasts generated by a random walk with the forecasts of the second combination approach ('**far1RW**'). Given the superior forecasting performance of the random walk in particular after the GFC, it will be interesting to see, whether the combination with other forecasts will be able to improve the results for the random walk.

For each strategy we consider two forecasting combination schemes: equal weights (**FCEW**) and performance weights (**FCPW**). With M models and hence M individual forecasts for a τ-maturity yield at time t a linear combination of the forecasts based on weights $w_{t,m}^{\tau}$ is generally given by

$$\hat{y}_{t+h/t,\tau}^{FC} = \sum_{m=1}^{M} w_{t,m}^{\tau} \hat{y}_{t+h/t,\tau}^{m}, \tag{3.15}$$

where the $M \times 1$ vector of weights w_m^{τ} is time-varying. For equal weights, the weight for each model is then simply given by

$$w_{t,m}^{\tau} = \frac{1}{M}. \tag{3.16}$$

For performance weights each forecast is weighed by the inverse of its MSE (Mean squared error)[19] over the previous $v = 24$ months. [20] The MSE for each model m and maturity τ at time t is calculated as

$$MSE_{t,m}^{\tau} = \frac{1}{v} \sum_{j=t-v}^{t} e_{j+h/j,m}^{2}, \tag{3.17}$$

where $e_{j+h/j,m}^{2}$ is the squared forecast error of model m at point in time j. Each weight is then given by

$$w_{t,m}^{\tau} = \frac{1/MSE_{t,m}^{\tau}}{\sum_{m=1}^{M} 1/MSE_{t,m}^{\tau}}. \tag{3.18}$$

This way, a model with a previously lower MSE is given a relatively larger weight than a model with a previously higher MSE.

Combining the three forecast combination strategies with the two combina-

[19]Note that we follow Timmermann (2006) in using the MSE to construct the weights instead of the RMSE that has been previously reported in Table 3.2 and 3.3.

[20]We choose 24 months as a compromise between increasing forecast accuracy (which would require a long training period) and limiting forecast uncertainty over time (which would require a shorter time period.

tion schemes leaves us with six forecast combination strategies. We denote these strategies by **FCEWfactor**, **FCPWfactor**, **FCEWfar1**, **FCPW-far1**, **FCEWfar1RW** and **FCPWfar1RW**.

3.6.2 Forecasting Performance of Combined Forecasts

Table 3.4 presents the results of these six forecasting strategies for the two subsample periods.[21] For the first subsample period from 2004:01 to 2008:12, results are reported in the upper panel of Table 3.4. We observe that combining individual forecasts slightly improves the forecasting accuracy, as most strategies perform better than the individual dynamic factor models. For several short term yields and in particular for a $h = 12$-month forecasting horizon the superior performance of the models over a random walk is even statistically significant. All three strategies fare comparably well while including the random walk into the combination strategy does not significantly improve the performance. Interestingly, for the first subsample there is also no notable difference between equally weighted and performance weighted combination schemes.

For the crucial second subsample period after the GFC (2009:01 to 2013:12), combining different models significantly improves the forecasting performance, albeit most of the strategies are still being dominated by the random walk. For example, the relative RMSE for a three-month yield forecast over a six-month horizon decreases to 1.88 relative to the forecasting error of a random walk for the performance weighed combination of all factor models (FCPW-factor). Recall that the initial relative RMSEs for the individual models in Table 3.2 range from 3.57 to 9.19 for the same maturity and forecasting horizon. In particular the FCPWfar1 strategy performs comparably well with relative RMSEs being significantly smaller than the individual RMSEs for this period. Obviously, this is partly due to the relatively good performance of the simple AR(1) approach.

[21] We also report the RMSE calculated over the entire forecasting period in Appndix A.2 while the all corresponding Diebold-Mariano statistics are reported in Appendix A.4.

		3m	6m	12m	2y	3y	5y	7y	10y
				2004:01 - 2008:12					
	RW	30.8	30.1	30.5	31.4	32.3	30.8	32.7	27.9
h=1	FCEWfactor	1.08	**0.86"**	**0.93**	1.05	1.04	1.05	1.05	1.01
	FCPWfactor	1.05	**0.85"**	**0.92**	1.04	1.04	1.04	1.05	1.00
	FCEWfar1	1.08	**0.95**	**0.95**	1.03*	1.03	1.03	1.04	1.01
	FCPWfar1	1.05	**0.94"**	**0.94**	1.02*	1.02	1.03	1.03	1.00
	FCEWfar1RW	1.05	**0.96**	**0.96**	1.02*	1.02	1.02	1.02	1.00
	FCPWfar1RW	1.01	**0.95**	**0.95**	1.02*	1.01	1.02	1.02	**1.00**
	RW	113.0	114.0	108.5	97.8	92.3	72.9	65.9	53.7
h=6	FCEWfactor	**0.92**	**0.92**	**0.95**	1.01	1.01	1.05	1.06	1.03
	FCPWfactor	**0.88**	**0.89**	**0.91"**	**0.98**	**0.98**	1.03	1.05	1.02
	FCEWfar1	**0.98**	**0.97**	**0.97**	**0.99**	**0.98**	1.01	1.03	1.01
	FCPWfar1	**0.92**	**0.93**	**0.93"**	**0.95**	**0.95**	**0.99**	1.02	1.00
	FCEWfar1RW	**0.98**	**0.98**	**0.98**	**0.99**	**0.98**	**1.00**	1.01	**1.00**
	FCPWfar1RW	**0.93**	**0.94**	**0.94**	**0.95**	**0.95**	**0.98**	1.01	**0.99**
	RW	207.0	204.3	186.6	156.8	136.5	100.0	83.9	63.6
h=12	FCEWfactor	**0.89"**	**0.90"**	**0.93"**	1.00	1.02	1.08	1.11	1.13
	FCPWfactor	**0.86"**	**0.87"**	**0.90"**	**0.96**	**0.97**	1.04	1.09	1.10
	FCEWfar1	**0.94"**	**0.94**	**0.95**	**0.96**	**0.97**	**1.00**	1.04	1.07
	FCPWfar1	**0.89"**	**0.90"**	**0.90"**	**0.92**	**0.92**	**0.96**	1.03	1.04
	FCEWfar1RW	**0.95**	**0.95**	**0.96**	**0.97**	**0.97**	**0.99**	1.02	1.03
	FCPWfar1RW	**0.91**	**0.92**	**0.92**	**0.93**	**0.92**	**0.94**	1.00	1.02
				2009:01 - 2013:12					
	RW	4.1	4.7	7.4	12.9	18.0	24.6	28.1	28.7
h=1	FCEWfactor	2.19*	1.33*	1.43*	1.14*	1.07	1.12	1.02	1.01
	FCPWfactor	1.27	1.05	1.22	1.10	1.05	1.09	1.01	1.00
	FCEWfar1	2.26*	1.20*	1.07	1.06	1.07	1.07	1.00	**0.99**
	FCPWfar1	1.01	**0.98**	1.00	1.03	1.05	1.04	**0.99**	**0.99**
	FCEWfar1RW	1.82*	1.11	1.03	1.04	1.05	1.04	**0.99**	**0.99**
	FCPWfar1RW	**0.97**	**0.94**	**0.96**	1.01	1.03	1.02	**0.97**	**0.99**
	RW	7.9	9.4	9.9	22.4	38.9	63.7	75.9	76.2
h=6	FCEWfactor	2.78*	2.18*	1.99*	1.30*	1.24*	1.08	**0.98**	**0.91**
	FCPWfactor	1.88*	1.69*	1.69*	**0.94**	1.05	1.02	**0.95**	**0.90**
	FCEWfar1	3.13*	2.49*	2.43*	1.62*	1.35*	1.12	1.03	**0.97**
	FCPWfar1	1.19	1.09	1.07	1.27*	1.22	1.09	1.00	**0.96**
	FCEWfar1RW	2.49*	2.04*	2.02*	1.43*	1.23*	1.07	1.00	**0.97**
	FCPWfar1RW	**0.98**	**0.95**	**1.00**	1.14	1.11	1.03	**0.97**	**0.96**
	RW	8.9	10.5	11.9	26.2	43.8	74.5	92.7	93.6
h=12	FCEWfactor	4.12*	3.50*	3.34*	2.09*	1.74*	1.32*	1.11	**0.99**
	FCPWfactor	2.73*	2.47*	2.35*	1.41*	1.37	1.17	1.03	**0.94"**
	FCEWfar1	4.34*	3.73*	3.72*	2.44*	1.90*	1.42*	1.22	1.11
	FCPWfar1	1.42	1.29	1.01	1.55*	1.55*	1.30	1.15	1.04
	FCEWfar1RW	3.39*	2.98*	3.00*	2.05*	1.64*	1.28	1.14	1.06
	FCPWfar1RW	**0.93**	**0.80**	**0.82**	1.27*	1.27*	1.15	1.06	1.00

Table 3.4. Subsample forecasting combination results for h=1, h=6 and h=12 month-ahead forecasting horizons and three-month, six-month, twelve-month, two-year, three-year, five-year, seven-year and ten-year maturities. We report the root mean squared error (RMSE) for the out-of-sample periods **2004:1 - 2008:12** and **2009:1 - 2013:12**. The first line reports the RMSE for the random walk (expressed in basis points). The RMSEs of all other models are expressed relative to the random walk. Hence, numbers smaller than one (**reported in bold**) indicate that models outperform the random walk. Numbers larger than one indicate inferior performance. (") indicates statistical significant forecasting superiority of the respective models against the random walk measured by the DM-statistic on a 5% or smaller significance level. (*) indicates statistical significant forecasting inferiority against the random walk. The DM-statistics are reported in Appendix A.4. See text for a description of the selected forecast combination strategies.

Not surprisingly, the most promising strategy turns out to be the performance weighted forecast combination of the NSAR and PCAAR model with both the AR(1) model and the random walk (FCPWfar1RW). This strategy even outperforms the simple random walk forecast for most forecast horizons and maturities, in particular for three-month, six-month and 12-month yields as well as for yields with longer maturities such as seven or 10 years.
For the second subsample, the results of Table 3.4 also indicate that weighing the individual models based on their previous performance generally makes a remarkable difference compared to the equally weighted forecast combinations. We have a closer look at this finding by investigating the weights allocated to each of the included models, when the performance based weighting technique is applied to create forecast combinations. Figure 3.6 displays the development of the weights for the two most promising strategies, FCPWfar1 and FCPWfar1RW. The figure illustrates how the AR(1) process (for FCPWfar1) and the random walk (for FCPWfar1RW) receive increasing weights in the combined forecasting scheme due to their superior relative forecasting performance. While forecasts created by the PCA and Nelson-Siegel based factor models still obtain relatively high weights at the beginning of the forecasting period, from 2010 onwards the AR(1) process and the random walk respectively become the dominant models and crowd out the factor models.

Overall, our results clearly illustrate that forecast combinations are able to improve predictions for the term structure of interest rates made by individual econometric models. As especially during separate regimes of yield curve behavior, different models will provide the most appropriate forecasts and weights allocated to the individual models change dramatically, forecast combination strategies may adapt quicker to changes in the interest rated environment. In particular during the transition from a more volatile behavior of the yield curve to the current low interest rate regime with only minor fluctuations, our results thus strongly encourage the use of forecast combination schemes.

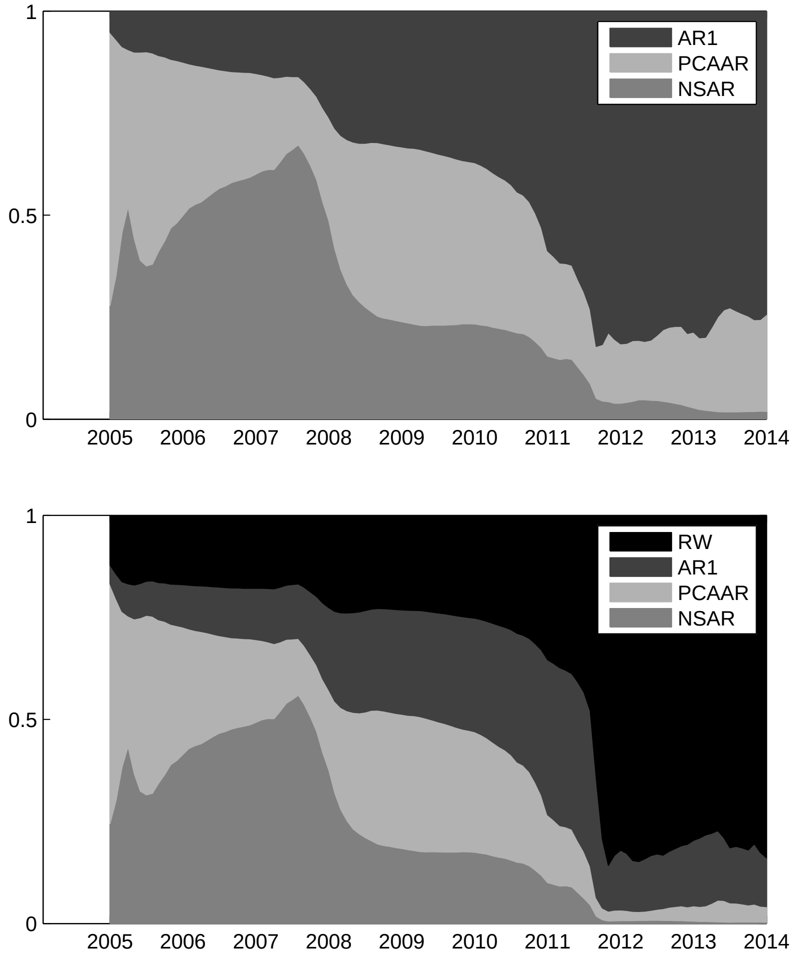

Figure 3.6. Development of forecast combination weights. We plot the weight dynamics of the FCPWfar1 (top plot) and FCPWfar1RW (bottom plot) strategy for the three-months yield forecasts and h=12 forecast horizon. The FCPWfar1 strategy encompasses the NSAR, PCAAR and AR1 model. The FCPWfar1RW additionally includes the random walk. The weights are calculated based on the inverse MSE of the previous v=24 months. See Section Appendix3.6.1 for a more detailed description of the selected forecast combination strategies.

3.7 Conclusion

This paper provides a pioneer study in documenting the challenge which the current low interest rate environment poses to popular dynamic factor yield curve forecasting models. To examine the forecasting accuracy during this unique time period, we apply a dataset of monthly US Treasury yields with maturities ranging from three up to 120 months for the time period January 1995 to December 2013. We focus on the popular class of dynamic factor models and investigate variations of the parametric dynamic Nelson-Siegel model and regressions on principal components (PCA).

We are particularly interested in the forecasting performance subsequent to the GFC that is dominated by relatively low and non-volatile short and medium term interest rates. While results for the pre-crisis and crisis periods are in line with findings from other comprehensive forecasting studies, the forecasting accuracy of the applied econometric models worsens dramatically with short rates close to the zero bound. The investigated dynamic factor models not only fail to beat the random walk but are clearly outperformed in relative terms. Diebold-Mariano statistics show that the outperformance of the models by a simple random walk no-change forecast is also statistically significant, often at the 1% level. Dynamic forecasting metrics and graphical analysis of the forecasted time series further reveal that the AR(1) model and random walk adapt rather quickly to the changed environment while factor models continuously over- or under-predict the actual level of the yields. While one would naturally expect the low, non-volatile interest rate environment to favor a random walk no-change forecast, the extent of the outperformance by the random walk in comparison to the applied dynamic factor models is still striking, also compared to previous studies which report differences in forecasting results for different subsample periods (Moench, 2008; Pooter et al., 2010; Steeley, 2014).

We identify several potential reasons for these results. First, we suspect that the cross-sectional structure of the dynamic factor models, which also includes additional information of other maturities, worsens the forecasting accuracy as the short end of the yield curve becomes more segmented from the rest of the curve. Further, the applied dynamic factor models are typically calibrated over a sample period that also includes significant changes in interest rates as well as volatile periods for the term structure of the yield curve. Therefore, these models may overstate the dynamics of individual yields and the entire yield curve during the unique low interest rate period from 2009 to 2013. Moreover, the models were also estimated during periods when interest rates were significantly higher than during the post GFC period, such that forecasts created by the applied models will not only overstate the dynamics of the interest rate term structure, but possibly also interest rate levels, which is also evidenced by our results. Finally, the crucial time period has been subject to significant market interventions by the Fed ('Quantitative Easing') which may further hamper the performance of traditional yield currve forecasting models.

With regards to modeling, our results indicate a need for approaches that are specifically designed to also capture term structure dynamics for interest rates near the zero bound. Therefore, recently developed models by, e.g. Kim and Singleton (2012); Krippner (2013); Christensen et al. (2015); Christensen and Rudebusch (2015); Monfort et al. (2015); Filipović et al. (2016) are more likely to appropriately describe and potentially forecast the behavior of interest rates in the current low yield environment.

Another important finding of this study is that the poor forecasting performance of the applied term structure models in the low interest rate environment is not reflected in commonly used global forecast evaluation metrics such as the RMSE when calculated over the entire forecasting period. When we investigate the forecasting performance for the entire forecasting period from January 2004 to December 2013, we surprisingly find that all models perform rather well and in line with results reported in previous research. Additional analysis reveals, that the forecasting errors during the critical period after the GFC become relatively small in absolute terms, especially for short and medium term yields, and therefore contribute relatively little to the global average. Therefore, investigating only the global forecasting performance may fail to detect important information about the relative forecasting performance of competing models through time and may lead to entirely different conclusions about the predictive accuracy of econometric models. As the current interest rate environment may well last for some more time into the future,[22] this finding has important implications for current and future

[22]Fed chair Janet Yellen only recently confirmed there will be 'considerable time' before the central bank may raise its benchmark rate. See the transcript of Chair Yellen's Press Conference on 19 March, 2014.

yield curve modeling and forecasting exercises. Not considering the unique behavior of short and medium yields in this time period may distort future results and interpretations.

With regards to forecast evaluation, it is thus crucial to carefully examine the dynamic behavior of the term structure and conduct subsample analysis accordingly. It is still common to measure forecasting accuracy predominantly with RMSEs computed over the entire sample period and select the model with the best global forecasting performance. However, a thorough subsample analysis and dynamic forecasting measures are crucial to truly expose a model's predictive abilities. Dynamic forecast evaluation measures should thus be applied to identify periods of superior or inferior forecasting accuracy. However, as illustrated in our study, even tests specifically designed to evaluate the dynamic performance of forecasting models, such as, e.g. the fluctuation test suggested by Giacomini and Rossi (2010), may have difficulties in detecting significant performance differences in the presence of structural instabilities of the loss functions (Martins and Perron, 2012).

Finally, it is important to develop mitigating measures to improve the relative forecasting accuracy in periods of low and non-volatile interest rates. As a potential approach we identify forecast combination strategies in this study. Simply combining all factor models already notably improves the inferior performance relative to the random walk. Strategically combining forecasts from a Nelson-Siegel, a PCA model, an AR(1) model further improves the results. This is particularly true when the forecast combination weights are based on the recent forecasting performance of the individual models. It is also interesting to note, that combining the random walk with other models further improves the random walk no change forecast. Overall, these results show that combining forecasts has the potential to significantly improve the forecasting accuracy, especially for a time period where many of the individual models perform rather poorly.

Our results also point towards the benefits of using adaptive forecasting techniques or regime switching models to predict the yield curve in different economic environments as they have recently been suggested by Xiang and Zhu (2013); Chen and Niu (2014). Such models may be more suitable to identify different phases of interest rate and yield curve behavior and may capture the change between volatile or quiet regimes also in their forecasts. Recent results, see, e.g., Koopman and van der Wel (2013); Exterkate et al. (2013) have also shown that including macroeconomic variables can significantly improve the forecasting performance of yield curve models. Finally, the recently developed models by Kim and Singleton (2012); Krippner (2013); Filipović et al. (2016); Christensen et al. (2015); Christensen and Rudebusch (2015); Monfort et al. (2015) for interest rate dynamics near the zero bound should also help to achieve better term structure forecasts in the current environment. We recommend to thoroughly investigate the forecasting performance of these models in future research.

Appendix A

A.1 Yield forecasts plots for additional maturities

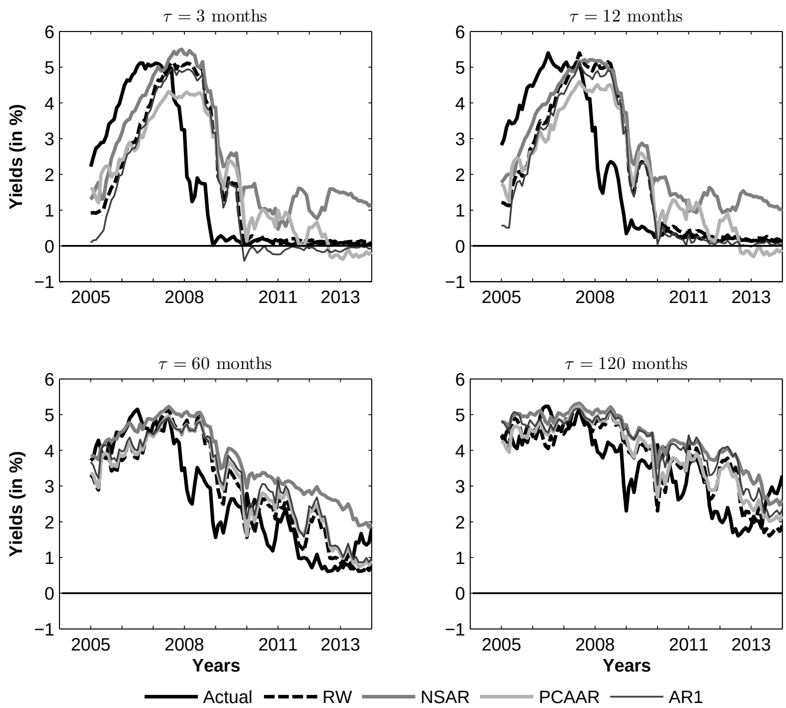

Figure A.1. Yield forecasts for random walk, NSAR, PCAAR and AR1 model. In addition to the three-month maturity (on which we have also focused in Figure 3.3 above), we plot the twelve-month, five-year and ten-year actual yields together with the forecasts of the selected models for a forecasting horizon of $h = 12$ months.

A.2 Forecasting combination performance - Full sample

		3m	6m	12m	2y	3y	5y	7y	10y
		2004:01 - 2013:12							
	RW	21.9	21.5	22.1	24.0	26.1	27.7	30.0	27.9
h=1	FCEWfactor	1.08	**0.86”**	**0.93**	1.05*	1.04	1.05*	1.05	1.01
	FCPWfactor	1.05	**0.85”**	**0.92**	1.04*	1.04	1.04*	1.05	1.00
	FCEWfar1	1.11*	**0.95**	**0.96**	1.04*	1.04	1.05	1.02	**1.00**
	FCPWfar1	1.05	**0.94**	**0.95**	1.03*	1.03	1.03	1.01	**0.99**
	FCEWfar1RW	1.07	**0.96**	**0.96**	1.03*	1.03	1.03	1.01	**1.00**
	FCPWfar1RW	1.02	**0.95**	**0.95**	1.02*	1.02	1.02	1.00	**0.99**
	RW	81.1	82.6	79.8	72.7	72.4	70.4	72.5	66.9
h=6	FCEWfactor	**0.92**	**0.92**	**0.95**	1.01	1.01	1.05	1.06	1.03
	FCPWfactor	**0.88**	**0.89**	**0.91**	**0.98**	**0.98**	1.03	1.05	1.02
	FCEWfar1	1.02	**1.00**	**1.00**	1.05	1.06	1.08	1.04	**1.00**
	FCPWfar1	**0.95**	**0.95**	**0.95**	1.00	1.02	1.06	1.03	**0.99**
	FCEWfar1RW	1.01	**1.00**	**0.99**	1.03	1.04	1.04	1.02	**0.99**
	FCPWfar1RW	**0.95**	**0.96**	**0.96**	**0.99**	**1.00**	1.02	1.01	**0.98**
	RW	145.0	144.0	133.4	112.5	100.7	86.7	85.7	78.1
h=12	FCEWfactor	**0.89**	**0.90**	**0.93**	1.00	1.02	1.08	1.11	1.13
	FCPWfactor	**0.86**	**0.87**	**0.90**	**0.96**	**0.97**	1.04	1.09	1.10
	FCEWfar1	**0.99**	**0.98**	**1.00**	1.07	1.13	1.21	1.18	1.13
	FCPWfar1	**0.93**	**0.94**	**0.95**	**1.00**	1.04	1.14	1.13	1.07
	FCEWfar1RW	**0.99**	**0.98**	**0.99**	1.04	1.08	1.14	1.11	1.07
	FCPWfar1RW	**0.94**	**0.95**	**0.95**	**0.98**	**1.00**	1.06	1.06	1.03

Table A.1. Forecasting combination results for $h = 1$, $h = 6$ and $h = 12$ month-ahead forecast horizons and three-month, six-month, twelve-month, two-year, three-year, five-year, seven-year and ten-year maturities. We report the root mean squared error (RMSE) for the out-of-sample period **2004:1 - 2013:12** (N = 96). The first line reports the RMSE for the random walk (expressed in basis points). The RMSEs of all other models are expressed relative to the random walk. Hence, numbers smaller than one (**reported in bold**) indicate that models outperform the random walk. Numbers larger than one indicate inferior performance. (") indicates statistical significant forecasting superiority of the respective models against the random walk measured by the DM-statistic at the 5% level of significance, (*) indicates statistical significant forecasting inferiority against the random walk. The DM-statistics are reported in Appendix A.4. See text for a description of the selected forecast combination strategies.

A.3 Diebold-Mariano test statistics - Individual Models

		3m	6m	12m	2y	3y	5y	7y	10y
					2004:01 - 2013:12				
h=1	NSAR	4.97*	**-0.76**	**-1.45**	1.82	2.57*	3.77*	1.54	0.99
	NSVAR	1.84	**-1.11**	0.80	2.70*	1.66	3.20*	0.98	1.25
	PCAAR	2.13*	**-1.06**	0.23	4.20*	2.13*	1.24	0.47	0.12
	PCAVAR	0.00	**-1.51**	0.23	2.30*	1.42	1.48	1.06	**-0.11**
	AR1	1.93	1.80	1.49	1.45	1.23	0.85	0.71	0.62
h=6	NSAR	1.74	1.26	0.99	1.94	2.39*	2.15*	1.31	0.78
	NSVAR	**-0.10**	**-0.39**	**-0.32**	0.93	1.01	1.06	0.29	**-0.60**
	PCAAR	0.30	0.27	0.45	1.66	1.81	1.06	**-0.47**	**-2.17"**
	PCAVAR	**-0.24**	0.07	0.53	1.52	1.00	0.57	0.25	**-0.58**
	AR1	0.88	0.80	0.66	0.39	0.53	0.94	0.93	0.78
h=12	NSAR	1.08	0.91	1.01	1.64	1.94	2.15*	1.89	1.93
	NSVAR	**-0.61**	**-0.68**	**-0.36**	1.43	1.45	1.52	0.97	0.50
	PCAAR	**-0.48**	**-0.31**	**-0.04**	0.82	1.42	1.42	0.35	**-1.74**
	PCAVAR	**-0.54**	**-0.11**	0.47	2.10*	1.71	1.35	0.89	0.37
	AR1	0.63	0.54	0.38	0.08	0.49	1.30	1.52	1.53

Table A.2. Diebold-Mariano forecast accuracy test statistics of all investigated models against the random walk. We report the results of the period from **2004:01 to 2013:12** for one-month, six-month and twelve-month forecast horizons and three-month, six-month, twelve-month, two-year, three-year, five-year, seven-year and ten-year maturities. Note that **negative values indicate superiority of the investigated models against the random walk**. (") denotes a significantly superior performance against a random walk relative to the asymptotic null distribution at the 5% level of significance. (*) denotes significance of the inferior performance against the random walk relative to the asymptotic null distribution at the 5% level. See section 3.4 for a description of the selected models.

		3m	6m	12m	2y	3y	5y	7y	10y
		2004:01 - 2008:12							
h=1	NSAR	2.57*	**-2.07"**	**-1.68**	1.54	1.24	1.92	1.05	0.48
	NSVAR	0.85	**-1.78**	**-0.79**	1.30	1.24	2.01*	0.89	0.15
	PCAAR	1.85	**-1.23**	**-0.12**	2.81*	1.38	1.03	0.80	0.39
	PCAVAR	**-0.35**	**-2.24"**	**-0.43**	1.56	1.11	1.02	1.20	0.82
	AR1	1.67	1.55	1.35	1.28	1.08	0.81	0.99	1.03
h=6	NSAR	0.19	**-0.77**	**-1.31**	**-0.05**	0.15	0.83	1.05	0.97
	NSVAR	**-1.07**	**-0.98**	**-0.64**	0.80	0.74	0.99	1.00	0.88
	PCAAR	**-0.33**	**-0.15**	0.05	0.50	0.15	0.19	0.03	**-0.91**
	PCAVAR	**-1.12**	**-0.43**	0.27	1.43	0.98	0.82	0.99	0.72
	AR1	0.86	0.83	0.69	0.03	**-0.67**	**-0.69**	0.37	0.56
h=12	NSAR	**-0.22**	**-0.75**	**-0.75**	**-0.09**	0.06	0.41	0.64	1.05
	NSVAR	**-1.80**	**-1.29**	**-0.84**	0.80	0.71	0.75	0.72	0.85
	PCAAR	**-1.24**	**-0.86**	**-0.58**	**-0.29**	**-0.45**	**-0.19**	0.01	**-0.21**
	PCAVAR	**-1.12**	**-0.46**	0.16	1.62	1.12	0.81	0.74	0.77
	AR1	0.64	0.57	0.40	**-0.32**	**-1.65**	**-0.77**	0.19	0.60
		2009:01 - 2013:12							
h=1	NSAR	11.53*	6.46*	1.43	1.15	3.32*	3.33*	1.14	0.86
	NSVAR	6.11*	4.81*	6.06*	2.95*	2.08*	2.51*	0.54	1.24
	PCAAR	4.51*	1.85	2.60*	4.04*	2.25*	0.78	**-0.25**	**-0.17**
	PCAVAR	4.02*	5.06*	3.51*	2.31*	1.47	1.10	0.18	**-0.77**
	AR1	1.51	1.20	0.70	1.08	0.65	0.42	0.11	0.02
h=6	NSAR	6.65*	6.67*	6.38*	7.89*	4.51*	1.90	0.85	0.25
	NSVAR	4.81*	3.31*	2.70*	**-0.43**	0.27	0.22	**-0.37**	**-1.49**
	PCAAR	3.12*	2.83*	3.35*	3.99*	2.59*	0.71	**-0.85**	**-2.37"**
	PCAVAR	2.78*	3.01*	3.26*	0.27	**-0.18**	**-0.27**	-0.47	**-1.13**
	AR1	1.90	1.21	0.44	1.68	1.49	0.86	0.61	0.39
h=12	NSAR	17.73*	0.00	0.00	0.00	8.85*	3.62*	2.09*	1.71
	NSVAR	0.00	0.00	9.92*	1.37	4.68*	1.42	0.16	0.00
	PCAAR	2.50*	2.34*	2.38*	3.13*	3.45*	1.64	0.00	0.00
	PCAVAR	4.18*	6.90*	7.11*	1.62	0.00	1.19	**-0.05**	0.00
	AR1	3.94*	0.00	0.50	1.84	2.10*	2.08*	1.63	1.25

Table A.3. Diebold-Mariano forecast accuracy test statistics of the random walk against all selected models and the AR(1) model against all selected models. We report the results for the subsample periods **2004:01-2009:12** (upper table) and **2009:01-2013:12** (lower table) for one-month, six-month and twelve-month forecast horizons and three-month, six-month, twelve-month, two-year, three-year, five-year, seven-year and ten-year maturities. Note that **negative values indicate superiority of the random walk**. (") denotes a significantly superior performance of the models against a random walk relative to the asymptotic null distribution at the 5% level. (*) denotes significance of the inferior performance against the random walk relative to the asymptotic null distribution at the 5% level. See Section 3.4 for a description of the selected models.

A.4 Diebold-Mariano test statistics - Forecast Combination

		3m	6m	12m	2y	3y	5y	7y	10y
		2004:01 - 2013:12							
h=1	FCEWfactor	1.56	**-2.02”**	**-0.51**	2.76*	1.47	2.29*	0.91	0.35
	FCPWfactor	0.94	**-2.23”**	**-0.85**	2.31*	1.28	1.98*	0.84	0.19
	FCEWfar1	2.12*	**-1.57**	**-0.99**	3.41*	1.48	1.72	0.63	**-0.15**
	FCPWfar1	1.06	**-1.89**	**-1.29**	2.67*	1.20	1.35	0.53	**-0.42**
	FCEWfar1RW	1.79	**-1.75**	**-1.21**	3.16*	1.30	1.29	0.31	**-0.31**
	FCPWfar1RW	0.50	**-2.10**	**-1.55**	2.45*	0.99	0.92	0.17	**-0.61**
h=6	FCEWfactor	**-0.41**	**-0.76**	**-0.79**	1.14	1.30	1.06	0.35	**-0.67**
	FCPWfactor	**-0.85**	**-1.18**	**-1.36**	0.18	0.47	0.79	0.20	**-0.85**
	FCEWfar1	0.34	**-0.03**	**-0.08**	1.28	1.40	1.19	0.59	**-0.03**
	FCPWfar1	**-0.70**	**-0.96**	**-1.16**	0.00	0.39	0.98	0.51	**-0.22**
	FCEWfar1RW	0.19	**-0.16**	**-0.21**	1.13	1.16	0.89	0.31	**-0.29**
	FCPWfar1RW	**-0.91**	**-1.09**	**-1.28**	**-0.30**	**-0.10**	0.56	0.17	**-0.53**
h=12	FCEWfactor	**-0.84**	**-0.93**	**-0.51**	1.35	1.54	1.63	1.17	0.80
	FCPWfactor	**-1.21**	**-1.33**	**-1.10**	0.30	0.62	1.08	0.89	0.31
	FCEWfar1	**-0.24**	**-0.33**	**-0.03**	0.98	1.29	1.54	1.36	1.27
	FCPWfar1	**-1.02**	**-1.10**	**-0.95**	**-0.04**	0.41	1.13	1.20	0.97
	FCEWfar1RW	**-0.36**	**-0.44**	**-0.16**	0.86	1.16	1.36	1.12	0.97
	FCPWfar1RW	**-1.11**	**-1.18**	**-1.04**	**-0.28**	**0.00**	0.60	0.80	0.53

Table A.4. Diebold-Mariano forecast accuracy test statistics of the forecast combination strategies against the random walk. We report the results of the forecasting period **2004:01-2013:12** for one-month, six-month and twelve-month forecast horizons and three-month, six-month, twelve-month, two-year, three-year, five-year, seven-year and ten-year maturities. Note that **negative values indicate superiority of the investigated models against the the random walk**. (") denotes a significantly superior performance of the models against a random walk relative to the asymptotic null distribution at the 5% level. (*) denotes significance of the inferior performance against the random walk relative to the asymptotic null distribution at the 5% level. See Section 3.6 for a description of the selected combination strategies.

		3m	6m	12m	2y	3y	5y	7y	10y
					2004:01 - 2008:12				
h=1	FCEWfactor	1.10	**-2.22"**	**-0.97**	1.93	1.00	1.39	0.97	0.32
	FCPWfactor	0.81	**-2.35"**	**-1.12**	1.57	0.90	1.26	0.95	0.17
	FCEWfar1	1.50	**-1.78**	**-1.12**	2.88*	0.94	1.12	0.90	0.36
	FCPWfar1	0.93	**-2.00"**	**-1.38**	2.27*	0.75	0.94	0.86	0.04
	FCEWfar1RW	1.31	**-1.91**	**-1.28**	2.78*	0.86	0.96	0.77	0.31
	FCPWfar1RW	0.42	**-2.20**	**-1.59**	2.19*	0.63	0.73		
h=6	FCEWfactor	**-1.52**	**-1.39**	**-1.17**	0.33	0.19	0.58	0.76	0.44
	FCPWfactor	**-1.92**	**-1.86**	**-1.97"**	**-0.58**	**-0.30**	0.41	0.65	0.24
	FCEWfar1	**-0.63**	**-0.82**	**-0.80**	**-0.42**	**-0.83**	0.12	0.45	0.19
	FCPWfar1	**-1.37**	**-1.72**	**-2.00"**	**-1.63**	**-1.55**	**-0.18**	0.40	0.08
	FCEWfar1RW	**-0.69**	**-0.88**	**-0.85**	**-0.52**	**-1.10**	**-0.06**	0.26	**-0.01**
	FCPWfar1RW	**-1.50**	**-1.77**	**-2.04**	**-1.73**	**-1.76**	**-0.54**	0.12	**-0.17**
h=12	FCEWfactor	**-5.33"**	**-2.73"**	**-2.58"**	0.02	0.23	0.44	0.53	0.62
	FCPWfactor	**-4.08"**	**-3.39"**	**-3.31"**	**-0.60**	**-0.22**	0.19	0.45	0.52
	FCEWfar1	**-4.27"**	**-1.89**	**-1.63**	**-1.42**	**-0.64**	**-0.03**	0.27	0.45
	FCPWfar1	**-2.85"**	**-2.73"**	**-2.59"**	**-1.73**	**-1.06**	**-0.33**	0.20	0.30
	FCEWfar1RW	**-4.30**	**-1.94**	**-1.68**	**-1.47**	**-0.69**	**-0.10**	0.17	0.28
	FCPWfar1RW	**-2.66**	**-2.67**	**-2.53**	**-1.68**	**-1.13**	**-0.53**	0.02	0.14
					2009:01 - 2013:12				
h=1	FCEWfactor	4.84*	2.86*	3.14*	2.54*	1.86	1.84	0.30	0.21
	FCPWfactor	1.54	0.39	1.77	1.95	1.29	1.48	0.16	0.09
	FCEWfar1	5.72*	2.19*	1.00	1.86	1.61	1.33	0.03	**-0.42**
	FCPWfar1	0.18	**-0.14**	0.04	0.97	1.10	0.89	**-0.28**	**-0.56**
	FCEWfar1RW	4.65*	1.52	0.51	1.51	1.33	0.91	**-0.26**	**-0.57**
	FCPWfar1RW	**-0.58**	**-0.59**	**-0.57**	0.60	0.78	0.46	**-0.67**	**-0.73**
h=6	FCEWfactor	2.32*	2.23*	2.19*	4.04*	2.14*	0.54	**-0.17**	**-1.17**
	FCPWfactor	2.56*	2.24*	2.42*	**-0.47**	0.47	0.13	**-0.37**	**-1.35**
	FCEWfar1	3.63*	2.89*	2.65*	2.82*	2.32*	0.95	0.23	**-0.35**
	FCPWfar1	0.94	0.54	1.04	2.00*	1.78	0.77	0.04	**-0.55**
	FCEWfar1RW	3.23*	2.59*	2.59*	2.64*	1.99*	0.67	0.02	-0.55
	FCPWfar1RW	**-0.39**	**-0.73**	**-0.05**	1.76	1.34	0.36	**-0.27**	**-0.81**
h=12	FCEWfactor	2.26*	2.14*	2.22*	4.18*	6.14*	2.00*	0.74	**-0.15**
	FCPWfactor	3.02*	3.44*	4.21*	2.27*	0.00	1.32	0.24	**-2.85"**
	FCEWfar1	4.40*	3.18*	2.55*	2.91*	3.07*	2.15*	1.28	0.93
	FCPWfar1	1.40	1.39	0.03	2.13*	2.46*	1.86	0.98	0.50
	FCEWfar1RW	4.42*	3.08*	2.48*	2.78*	2.79*	1.81	1.01	0.63
	FCPWfar1RW	**-3.09**	**-2.47**	**-1.60**	2.12*	2.37*	1.48	0.52	0.00

Table A.5. Diebold-Mariano forecast accuracy test statistics of the forecast combination strategies against the random walk. We report the results for the subsample periods **2004:01 to 2009:12** (upper table) and **2009:01-2013:12** (lower table) for one-month, six-month and twelve-month forecast horizons and three-month, six-month, twelve-month, two-year, three-year, five-year, seven-year and ten-year maturities. Note that **negative values indicate superiority of the investigated models against the the random walk**. (") denotes a significantly superior performance of the models against a random walk relative to the asymptotic null distribution at the 5% level. (*) denotes significance of the inferior performance against the random walk relative to the asymptotic null distribution at the 5% level. See Section 3.6 for a description of the selected combination strategies.

Bibliography

Afonso, A., Martins, M. M., 2012. Level, slope, curvature of the sovereign yield curve, and fiscal behaviour. Journal of Banking & Finance 36 (6), 1789–1807.

Almeida, C., Vicente, J., 2008. The role of no-arbitrage on forecasting: Lessons from a parametric term structure model. Journal of Banking & Finance 32 (12), 2695–2705.

Ang, A., Bekaert, G., 2002. Regime Switches in Interest Rates. Journal of Business & Economic Statistics 20 (2), 163–182.

Ang, A., Piazzesi, M., 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. Journal of Monetary Economics 50 (4), 745–787.

Bianchetti, M., 2010. Two curves, one price. Risk (August), 74–80.

Bikbov, R., Chernov, M., 2010. Yield Curve and Volatility: Lessons from Eurodollar Futures and Options. Journal of Financial Econometrics 9 (1), 66–105.

Blaskowitz, O., Herwartz, H., 2009. PCA-based ex-ante Forecasting of Swap Term Structures. International Journal of Theoretical and Applied Finance 12 (04), 465–489.

Carcano, N., Dall'O, H., 2011. Alternative models for hedging yield curve risk: An empirical comparison. Journal of Banking & Finance 35 (11), 2991–3000.

Carriero, A., Kapetanios, G., Marcellino, M., 2012. Forecasting government bond yields with large Bayesian vector autoregressions. Journal of Banking & Finance 36 (7), 2026–2047.

Chen, Y., Niu, L., 2014. Adaptive dynamic Nelson–Siegel term structure model with applications. Journal of Econometrics 180 (1), 98–115.

Chen, Y.-C., Tsang, K. P., 2013. What Does the Yield Curve Tell Us about Exchange Rate Predictability? Review of Economics and Statistics 95 (1), 185–205.

Christensen, J. H. E., Diebold, F. X., Rudebusch, G. D., 2011. The affine arbitrage-free class of Nelson-Siegel term structure models. Journal of Econometrics 164 (1), 4–20.

Christensen, J. H. E., Lopez, J. A., Rudebusch, G. D., 2015. How Efficient is the Kalman Filter at Estimating Affine Term Structure Models? Manuscript.

Christensen, J. H. E., Rudebusch, G. D., 2015. Estimating Shadow-Rate Term Structure Models with Near-Zero Yields. Journal of Financial Econometrics 13 (2), 226–259.

Coroneo, L., Nyholm, K., Vidova-Koleva, R., 2011. How arbitrage-free is the Nelson-Siegel model? Journal of Empirical Finance 18 (3), 393–407.

Cox, J. C., Ingersoll, J. E., Ross, S. A., 1981. A Re-Examination of Traditional Hypotheses about the Term Structure of Interest Rates. Journal of Finance 36 (4), 769.

Cox, J. C., Ingersoll, J. E., Ross, S. A., 1985. A Theory of the Term Structure of Interest Rates. Econometrica 53 (2), 385–407.

Dai, Q., Singleton, K. J., 2000. Specification Analysis of Affine Term Structure Models. Journal of Finance 55 (5), 1943–1978.

Dewachter, H., Lyrio, M., 2006. Macro Factors and the Term Structure of Interest Rates. Journal of Money, Credit and Banking 38 (1), 119–140.

Diebold, F. X., Li, C., 2006. Forecasting the term structure of government bond yields. Journal of Econometrics 130 (2), 337–364.

Diebold, F. X., Li, C., Yue, V. Z., 2008. Global yield curve dynamics and interactions: A dynamic Nelson–Siegel approach. Journal of Econometrics 146 (2), 351–363.

Diebold, F. X., Mariano, R. S., 1995. Comparing Predictive Accuracy. Journal of Business & Economic Statistics 13 (3), 253–263.

Diebold, F. X., Rudebusch, G. D., Aruoba, S. B., 2006. The macroeconomy and the yield curve: a dynamic latent factor approach. Journal of Econometrics 131 (1-2), 309–338.

Duffee, G. R., 2002. Term Premia and Interest Rate Forecasts in Affine Models. Journal of Finance 57 (1), 405–443.

Duffee, G. R., 2011. Forecasting with the term structure: The role of no-arbitrage restrictions. Economics Working Paper Archive, Johns Hopkins University.

Duffee, G. R., 2012. Forecasting interest rates. Economics Working Paper Archive, Johns Hopkins University.

Duffie, D., Kan, R., 1996. A Yield-Factor Model Of Interest Rates. Mathematical Finance 6 (4), 379–406.

Exterkate, P., van Dijk, D., Heij, C., Groenen, P. J. F., 2013. Forecasting the Yield Curve in a Data-Rich Environment Using the Factor-Augmented Nelson-Siegel Model. Journal of Forecasting 32 (3), 193–214.

Fama, E. F., Bliss, R. R., 1987. The Information in Long-Maturity Forward Rates. American Economic Review 77 (4), 680–692.

Favero, C. A., Niu, L., Sala, L., 2012. Term Structure Forecasting: No-Arbitrage Restrictions versus Large Information Set. Journal of Forecasting 31 (2), 124–156.

Filipović, D., Larssi, M., Trolle, A. B., 2016. Linear-Rational Term Structure Models. Journal of Finance , forthcoming.

Giacomini, R., Rossi, B., 2010. Forecast comparisons in unstable environments. Journal of Applied Econometrics 25 (4), 595–620.

Guidolin, M., Timmermann, A., 2009. Forecasts of US short-term interest rates: A flexible forecast combination approach. Journal of Econometrics 150 (2), 297–311.

Gürkaynak, R. S., Sack, B. P., Wright, J. H., 2007. The U.S. Treasury yield curve: 1961 to the present. Journal of Monetary Economics 54 (8), 2291–2304.

Hautsch, N., Yang, F., 2012. Bayesian inference in a Stochastic Volatility Nelson–Siegel model. Computational Statistics & Data Analysis 56 (11), 3774–3792.

Hull, J., White, A., 1990. Pricing Interest-Rate-Derivative Securities. Review of Financial Studies 3 (4), 573–592.

Kim, D. H., Singleton, K. J., 2012. Term structure models and the zero bound: An empirical investigation of Japanese yields. Journal of Econometrics 170 (1), 32–49.

Koopman, S. J., van der Wel, M., 2013. Forecasting the US term structure of interest rates using a macroeconomic smooth dynamic factor model. International Journal of Forecasting 29 (4), 676–694.

Krippner, L., 2013. A theoretical foundation for the Nelson-Siegel class of yield curve models. Journal of Applied Econometrics 30 (1), 97–118.

Laurini, M. P., Hotta, L. K., 2014. Forecasting the Term Structure of Interest Rates Using Integrated Nested Laplace Approximations. Journal of Forecasting 33 (3), 214–230.

Leite, A. L., Filho, R. B. P. G., Vicente, J., 2010. Forecasting the yield curve: A statistical model with market survey data. International Review of Financial Analysis 19 (2), 108–112.

Litterman, R., Scheinkman, J., 1991. Common Factors Affecting Bond Returns. Journal of Fixed Income 1 (1), 54–61.

Martins, L. F., Perron, P., 2012. Forecast Comparisons in Unstable Environments: Comments and Improvements. Working Paper.

Moench, E., 2008. Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach. Journal of Econometrics 146 (1), 26–43.

Monfort, A., Pegaro, F., Renne, J.-P., Rousselet, J.-P., 2015. Staying at Zero with Affine Processes: A New Dynamic Term Structure Model. Working Paper.

Nelson, C. R., Siegel, A. F., 1987. Parsimonious Modeling of Yield Curves. Journal of Business 60 (4), 473.

Pooter, M. D., Ravazzolo, F., van Dijk, D., 2010. Term structure forecasting using macro factors and forecast combination. Board of Governors of the Federal Reserve System, International Finance Discussion Papers 993.

Reisman, H., Zohar, G., 2004. Short-Term Predictability of the Term Structure. Journal of Fixed Income 14 (3), 7–14.

Rudebusch, G. D., Wu, T., 2008. A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy. Economic Journal 118 (530), 906–926.

Steeley, J. M., 2014. Forecasting the Term Structure when Short-Term Rates are Near Zero. Journal of Forecasting 33 (5), 350–363.

Svensson, L. E. O., 1994. Estimating and Interpreting Forward Interest Rates: Sweden 1992 - 1994. NBER Working Paper 4871.

Timmermann, A., 2006. Forecast Combinations. In: Elliott, G., Granger, C. W. J., Timmermann, A. (Eds.), Handbook of economic forecasting. Vol. 1 of Handbook of Economic Forecasting. Elsevier, Amsterdam and Boston, pp. 135–196.

Vasicek, O., 1977. An equilibrium characterization of the term structure. Journal of Financial Economics 5 (2), 177–188.

Walker, S., McCormick, L., 2014. Unstoppable 100 Trillion Bond Market Renders Models Useless. Bloomberg.com, 3 June 2014.

Xiang, J., Zhu, X., 2013. A Regime-Switching Nelson-Siegel Term Structure Model and Interest Rate Forecasts. Journal of Financial Econometrics 11 (3), 522–555.

Yu, W.-C., Zivot, E., 2011. Forecasting the term structures of Treasury and corporate yields using dynamic Nelson-Siegel models. International Journal of Forecasting 27 (2), 579–591.

4 Factors of the Term Structure of Sovereign Yield Spreads

Dennis Wellmann (contribution: 80%), Stefan Trück (contribution: 20%)

This paper was presented at:

- 'Postgrad Research Seminar' at Humboldt-University, Berlin, May 2015
- 13th 'INFINITI Conference', Ljubljana, June 2015
- 5th 'International Conference of the Financial Engineering and Banking Society' (FEBS), Nantes, June 2015
- 'Finance Research Seminar' at University of Goettingen, July 2015
- 22nd 'Annual Meeting of the German Finance Association' (DGF), Leipzig, September 2015

Abstract

We investigate the term structure of sovereign yield spreads for five advanced economies against the US and provide novel insights on the key drivers of the term structure. We show that the spread term structure dynamics are driven by three latent factors, which can be labeled as spread level, slope and curvature similar to common interpretations found in the yield curve literature. We further show that these estimated spread factors have predictive power for exchange rate movements and excess returns, above the predictability of an uncovered interest rate parity approach. As the yield curve contains information about expected future economic conditions we conjecture that these yield spread factors reflect expected macroeconomic differentials which in turn drive exchange rates. Using the information content of yield spread curves may thus be a promising approach to improve the forecasting accuracy of exchange rate models.

4.1 Introduction

Sovereign yield spreads denote the difference between two government yields of equal maturity. They are important variables for investment practitioners, risk management and policy makers as they reflect the relative economic position against other economies and are key input factors for exchange rate forecasts and carry trade strategies. Commonly used as proxies for the difference in interest rate levels between economies, they play a crucial role in one of the cornerstones of the academic finance literature – the uncovered interest rate parity (UIRP) hypothesis.

Interestingly, the dynamic behavior of yield spreads, in particular the term structure of these spreads has received only limited attention in previous research. The class of literature decomposing the credit-risk driven determinants of sovereign spreads for emerging economies usually focuses on selected long term maturities (Rocha and Garcia, 2005; Hilscher and Nosbusch, 2010). The same holds for a recent stream of literature exploring the drivers of sovereign spreads within the Eurozone area (Bernoth and Erdogan, 2012; Maltritz, 2012; Eichler and Maltritz, 2013; Monfort and Renne, 2014). The large number of studies testing the UIRP, see, e.g., McCallum (1994); Chinn and Meredith (2004); Backus et al. (2010) or Sarno (2005); Engel (2013) for recent surveys, usually apply spreads between short term interest rates and naturally disregard the dynamics of the entire spread term structure. However, these term structure dynamics may be worth exploring as the information provided by yield spreads at various maturities may be different.

We investigate the term structure of sovereign yield spreads and provide novel insights on the latent factors driving the term structure. We further show that these latent factors can predict exchange rate changes and excess returns up to 24 months ahead. We conclude that the spread factors proxy expected fundamental differentials in price levels, output and monetary policy.

For our analysis we investigate the term structure of yield spreads of six advanced economies (Australia, Canada, Switzerland, Japan, UK and the US) with highly liquid markets of bonds issued in their own, free floating currency and little to no credit or default risk. The spreads are calculated as the difference between end-of-months zero-bond yields of equal maturity and the term structure is constructed with 12 maturities ranging from three-months, six-months, 12-months and 24-months up to 120-months for the time period from January 1995 to December 2013. In line with previous research (Dungey et al., 2000; Boudoukh et al., 2005; Sarno et al., 2012) we calculate all spreads against US yields.[1] This leaves us with five datasets of sovereign spread curves - US-AU, US-CA, US-CH, US-JP and US-UK.

To identify the drivers of the sovereign spreads term structure we derive latent factors by means of principal component analysis (PCA). This allows us to extract market expectations directly from the data. PCA is a common technique to describe dynamic term structure behaviour in a parsimonious manner and has been applied successfully to various financial assets, including, for example, the term structure of interest rates (Barber and Copper, 2012), swap spreads (Cortes, 2006) or CDS spreads (Longstaff et al., 2007). Note that we apply PCA directly to the yield spreads and thus differ from the class of studies modeling the co-movement of government yields to derive common global factors (Driessen et al., 2003; Martell, 2007; Pérignon and Smith, 2007; Afonso and Martins, 2012; Juneja, 2012).

Our analysis of the yield spread term structure dynamics shows that the variation in the entire term structure can be explained through a relatively small number of three factors. For the considered economies, the three estimated factors explain approximately 99% of the entire variation in the term structure of spreads between US interest rates and yields in Australia, Canada, Japan, Switzerland and United Kingdom. Interestingly, similar to interpretations found in yield curve models (Litterman and Scheinkman, 1991; Leite et al., 2010), the identified latent factors can be labeled as spread level, spread slope and spread curvature. These findings are stable across all investigated sovereign spreads.

We further test the ability of the extracted latent spread factors to predict exchange rate movements and excess returns. The yield curve is well known

[1]We note that the results and conclusions presented in this paper also hold for other combinations of these economies not including the US.

to provide valuable information about future macroeconomic conditions in particular output, inflation and monetary policy. Differences between these same fundamentals are commonly used to determine and forecast exchange rates (Rossi, 2013). We thus argue that latent factors derived from the term structure of yield spreads – the cross-country differences between yield curves – may serve as a natural measure to the fundamental aspects of exchange rate determination.

Results for in-sample predictive regressions confirm that the spread factors can explain and predict bilateral exchange rate movements and excess returns three month up to two years ahead. The predictive content seems to be highest for the '*safe haven currencies*'[2] Swiss franc and Japanese yen as well as the Australian dollar and to a lesser extent for the Canadian dollar and the British pound. We find that in particular the derived spread level and the spread slope factor are significant.

The negative coefficients for these two factors indicate that an increase in the yield spread level or slope factor (i.e. the entire home yield curve shifts up or becomes flatter relative to the foreign yield curve) predicts an appreciation of the home currency and increase in excess returns. These negative coefficients are consistent with economic intuition and findings in previous literature, in particular Chen et al. (2012); Chen and Tsang (2013).[3]

We further find additional explanatory power for the extracted factors compared to the UIRP approach for most currencies and horizons. These results make intuitive sense when the exchange rate is understood as an asset price and relies more on long term expectations than on current fundamentals. Other than the UIRP approach - which only uses the information content up to a certain maturity - the spread factors take advantage of the information embodied in the entire spread term structure. The estimated latent spread

[2]A safe haven currency provides hedging benefits to investors in particular in times of market turbulence (Ranaldo and Soderlind, 2010)

[3]Note that while the Dornbusch (1976) overshooting model would rather suggest the opposite pattern, i.e. an immediate appreciation and subsequent depreciation in response to a higher interest rate, more recent empirical studies (Eichenbaum and Evans, 1995; Gourinchas and Tornell, 2004; Clarida and Waldmann, 2008; Chen and Tsang, 2013) find that higher interest rates in a country may lead to a persistent appreciation of its currency.

factors can thus be interpreted as augmenting the single horizon UIRP relation with the information in spreads of additional maturities.
Overall, these results indicate that the obtained spread factors can be interpreted as an alternative set of latent fundamentals incorporating expected differences in observed macroeconomic fundamentals. This confirms the conclusions of Engel and West (2005), Bacchetta and van Wincoop (2013) and Balke et al. (2013) who find an important role of unobserved and expected fundamentals to explain exchange rate fluctuations. Considering the widespread forecasting failure of empirical exchange rate models based on observed macroeconomic fundamentals (Meese and Rogoff, 1983; Molodtsova and Papell, 2009; Rossi, 2013), the estimated spread factors may thus be helpful in future forecasting studies.
With these findings, we contribute to the macro-finance literature in several dimensions. First, this is the first study to thoroughly explore the dynamics of the entire term structure of yield spreads for advanced economies. Second, we provide key insights on latent key factors driving the term structure of sovereign spreads between advanced economies. Third, we show that the factors extracted from the term structure of sovereign spreads have predictive power for movements in the exchange rate and excess returns in line with economic intuition. We also illustrate that the extracted factors provide additional predictive information in comparison to the traditional UIRP approach.
The remainder of this chapter is organized as follows. The subsequent section describes the relation between yield spread curves, macroeconomic fundamentals and exchange rates as found in the previous literature. Section 4.3 reports descriptive statistics and investigates the dynamic behavior of the applied yield spread data. In Section 4.4 we estimate and interpret the latent yield spread factors, while Section 4.5 investigates the predictive ability of these factors for exchange rate movements and excess returns. Section 4.6 concludes and provides suggestions for future work in this area of research.

4.2 The Relation between Yield Curves, Yield Spreads, Macroeconomic Fundamentals and Foreign Exchange Rates

4.2.1 Yield Curves and the Term Structure of Yield Spreads

The yield curve or the term structure of interest rates describes the relationship between yields and their time to maturity. Yield curves exist for numerous instruments and securities, theoretically for any interest bearing instrument that is available for different maturities. However, the term *yield curve* or *term structure of interest rates* in an academic sense is mainly used to describe the term structure of government bond yields which are commonly considered to reflect a benchmark for the level of interest rates in an economy.

The yield curve summarizes expectations about future paths of short interest rates ('Expectations Hypothesis') and perceived future uncertainty expressed in the term premiums (Bulkley et al., 2011). Consequently it also contains information about expected future economic conditions such as output, recessions inflation and monetary policy (Ang et al., 2006; Rudebusch and Wu, 2008; Favero et al., 2012; Erdogan et al., 2015). The shape and movements of the yield curve have thus long been used to provide readings of market expectations and they are common indicators for central banks to receive timely feedback on their policy actions.

Research on the term structure of interest rates suggests that the variation in the term structure can be explained by a small number of underlying factors, see, e.g. Litterman and Scheinkman (1991) or more recently Bikbov and Chernov (2010). Typically, three factors already capture more than 99% of the variation in yields and are reflected in the entire term structure of interest rates and its dynamic behavior over time. They have an intuitive interpretation as level, slope and curvature related to the economically meaningful shift, rotation and butterfly moves of the yield curve which describe how the yield curve changes in response to macroeconomic changes and monetary policy.

These yield curve factors also have the power to predict fluctuations in future economic conditions. Diebold et al. (2006) find that an increase in the US level factor raises capacity utilization, the US fund rate and inflation. Dewachter and Lyrio (2006) estimate an affine model for the yield curve with macroeconomic variables and suggest that the level factor reflects long run inflation expectations, the slope factor captures the business cycle, and the curvature represents the monetary stance of the central bank. Rudebusch and Wu (2007) also contend that the level factor incorporates long-term inflation expectations, and the slope factor captures the central bank's dual mandate of stabilizing the real economy and keeping inflation close to its target. They show that when the central bank tightens monetary policy, the slope factor rises, forecasting lower growth in the future. Moench (2012) analyzes the economic underpinnings of level, slope, and curvature and finds that a rising slope factor is associated with a future decline of output while surprise surges of the yield curve level are followed by a strong and persistent increase of inflation rates.

The difference between two government yields of equal maturity – the '*sovereign yield spread*' or '*sovereign spread*' [4] - is also of particular importance to market participants and policy makers. Sovereign yield spreads reflect a government's creditworthiness and are heavily used by investment practitioners in exchange rate forecasting and carry trade strategies as these yield spreads reflect interest rate differentials which are key indicators for expected exchange rate movements.

As sovereign spreads can be calculated for any maturity, they exhibit a term structure or spread curve of their own. This term structure of sovereign spreads naturally contains valuable long term information about expected

[4] Note that the academic literature uses various terms to denote the difference between government yields. Sovereign yield spreads are also commonly referred to as '*government bond spreads*' (Dungey et al., 2000), '*sovereign credit spreads*' (Sueppel, 2005; Oliveira et al., 2012), '*sovereign risk premia*' (Haugh et al., 2009) or '*relative yield curves*' (Chen and Tsang, 2013). We use the terms '*sovereign yield spreads*' and '*yield spreads*' or just '*spreads*' throughout the course of the analysis.

cross-country differentials in the economic conditions reflected in the individual yield curves. These are the same macroeconomic differentials which play an important role in exchange rate determination.

4.2.2 Yield Spreads, Macroeconomic Fundamentals and Exchange Rate Determination

The relation between differences in interest rates and exchange rates is traditionally expressed in the uncovered interest rate parity (UIRP) condition. Under the assumptions of risk neutral and rational market participants, the UIRP links expected changes in the exchange rate to interest rate differences over the same horizon:

$$\Delta s_{t+h} = i_t^{\tau} - i_t^{*,\tau} + \rho_t, \tag{4.1}$$

with Δs_{t+h} being the change in the logarithm of the nominal spot exchange rate (home currency price per unit of foreign currency) between time t and $t+h$, i_t^{τ} and $i_t^{*,\tau}$ the monthly domestic and foreign interest rates at maturity τ with $\tau = h$ and ρ_t being the risk premium of holding foreign relative to home currency investments.[5]

This relationship naturally builds on interest rate differentials or sovereign spreads of a certain maturity. A τ-month spread accordingly only embodies information up until the underlying instrument matures. However, the literature predominantly suggests to consider the exchange rate as an asset price (Mark, 1995; Engel and West, 2005), where the nominal exchange rate is determined as the present value of the discounted sum of current fundamentals f_t and expectations about fundamentals in future periods $E_t(f_{t+j})$,

[5] Empirically, this hypothesis has mostly been rejected, see Sarno (2005); Engel (2013) for recent surveys.

given information available at time t:

$$s_t = (1-\omega)f_t + (1-\omega)\sum_{j=1}^{\infty}\omega^j E_t(f_{t+j}), \tag{4.2}$$

with ω being a discount factor less than one.

Based on this approach the exchange rate not only relies on information up to a certain maturity but depends heavily on expected long term fundamentals. These may be reflected more accurately in the entire term structure of sovereign spreads, capturing market expectations of future cross-country differentials, such as output, inflation, money supply and monetary policy, which are commonly used in traditional exchange rate models.

The traditional monetary model, for example, describes exchange rate behavior in terms of relative demand for and supply of money in the two economies. Assuming purchase price parity (PPP) only holds in the long run (i.e. a sticky price version of the monetary model), the fundamental f_t then becomes:[6]

$$f_t^{mon} = (m_t - m_t^*) - \eta(y_t - y_t^*). \tag{4.3}$$

where m_t and m_t^* are the domestic and foreign money market supply, y_t and y_t^* denote the domestic and foreign income and η the income elasticity.

The frequently used model based on Taylor rule fundamentals builds on the view that two central banks set interest rates in response to changes in the output gap and deviations from target inflation rates, and the bilateral exchange rate will reflect their relative interest rates through UIRP. For a symmetric Taylor rule with homogeneous coefficients, the fundamentals f_t become:

$$f_t^{TR} = (1+\phi)(\pi_t - \pi_t^*) + \gamma(y_t^{gap} - y_t^{gap*}). \tag{4.4}$$

where π_t and π_t^* are the domestic and foreign inflation rates and y_t^{gap} and y_t^{gap*} are the domestic and foreign output gaps.

However, these standard fundamental models have a rather poor forecasting

[6]See Molodtsova and Papell (2009) and Rossi (2013) for a more detailed derivation of the most common exchange rate models based on observable macroeconomic fundamentals.

performance. In a large body of empirical forecasting studies, the random walk model has proven to be almost unbeatable by models with traditional economic predictors.[7] This empirical failure may also be a result of using inappropriate proxies for market expectations of future fundamentals, rather than a failure of the models themselves. Engel and West (2005), for example, find that the exchange rate is not explained only by observable fundamentals. Balke et al. (2013) also show that it is difficult to obtain sharp inferences about the relative contribution of fundamentals, using only data on observed fundamentals. Bacchetta and van Wincoop (2013) conclude that the reduced form relationship between exchange rates and fundamentals is driven not by the structural parameters themselves, but rather by expectations of these parameters. Properly measuring expectations thus becomes especially important in empirical testing. Latent or unobservable factors, summarizing market expectations contained in the term structure of yield spreads, may therefore serve as a natural measure to the fundamental aspects of exchange rate determination. Ang and Chen (2010) and Chen and Tsang (2013), for example, show that yield curve factors based on portfolio strategies and a Nelson-Siegel model can explain foreign exchange rate returns.

These insights highlight two important points. First, the term structure of sovereign spreads is a highly relevant economic variable and it is crucial to understand the dynamics of sovereign spread curves. We will further investigate these dynamics as well as driving factors of the dynamics in the subsequent sections. Second, the latent factors driving the term structure of sovereign spreads may be helpful to predict exchange rates as they summarize information about long term differences in macroeconomic fundamentals. We further explore this assertion in Section 4.5.

[7] Rossi (2013) provides an excellent overview of the empirical exchange rate forecasting literature since the seminal paper by Meese and Rogoff (1983). See also Section 5.2.2.

4.3 Yield Spread Data

4.3.1 Source and Calculation

For our analysis of sovereign yield spreads we choose the five most advanced markets of government bonds issued in their own currency[8] and little to no credit risk (US, UK, Japan, Canada and Switzerland). We also consider yields in Australia that have gained particular interest in the recent literature on carry-trades and foreign exchange risk premiums (Darvas, 2009; Christiansen et al., 2011; Lustig et al., 2011; Sarno et al., 2012). The government bond zero yield data is directly obtained from Bloomberg for the time period from January 1995 (the first availability of the time series) up to December 2013.[9]

The monthly sovereign yield spreads Δsy_t^τ are calculated as the difference $sy_t^\tau - sy_t^{*,\tau}$ between sovereign yields of equal maturity τ at the end of each month. The term structure is constructed with 12 maturities ranging from three-months, six-months, 12-months, 24-months, 36-months up to 120-months. Following the existing literature, all spreads are calculated against US yields. This leaves us with five datasets of sovereign spreads: US-AU, US-CA, US-CH, US-JP and US-UK.[10]

[8] We thus have to exclude Euro countries, because these bonds are issued in a currency which reflects the macroeconomic prospects of the entire Euro area instead of the individual economies. This will become crucial for our analysis in Section 4.5. Using e.g. German yields as a proxy does not seem reasonable for our approach, especially as our sample includes the recent Euro crisis.

[9] Bloomberg yields have the advantage that they are consistently available for all considered economies in this study. All Bloomberg zero yield curves are constructed daily with government bonds that have Bloomberg Generic (BGN) and/or supplemental proprietary contributor prices.

[10] We note that the results and conclusions presented in this paper also hold for other combinations of the economies not including the US.

4.3.2 Statistical Properties

We summarize selected statistical properties of all sovereign spread data sets in Table 4.1. Note that sovereign spreads can be either positive or negative depending on the respective yield being lower or higher than US yields. The negative mean for the US-Australian, US-British and US-Canadian sovereign yield spreads, for example, indicates that the respective yields have on average been higher than US yields. The opposite holds for Switzerland and Japan, where yields have been significantly lower than US yields throughout the sample period. It is also interesting to note that calculated against US yields all average spread curves, indicated by the mean spreads of the different maturities, are upward sloping. This implies that throughout our sample period from 1995-2013, for economies with yields mostly higher than US yields (Australia, UK and Canada) the average spread narrows with longer maturities, while for economies with yields mostly lower than US yields (the two safe haven currencies Japan and Switzerland), the average spread widens with longer maturities. The mean spreads for different maturities also indicate that the slope of the average spread curve is relatively small.

The standard deviations point towards a difference in volatility between short-term and long-term maturities. We observe that yield spreads for longer maturities have far lower standard deviations for all currency pairs. For example, the shortest maturity (three-months) has a standard deviation as high as 2.16 for the US-JP series and 1.66 for US-AU. In contrast, the 120-months series for the same pair of countries is characterized by a standard deviation of just 0.90 for US-JP and 0.72 for US-AU. Calculated coefficients of correlation reveal that sovereign spreads at different maturities are highly correlated. As could be expected, correlation coefficients are typically the highest for adjacent maturities. The correlation between the 3-months and 60-months spread ranges from 0.85 to 0.95 for example. For the US-CH and US-JP series, even correlations between the 3-months and 120-months spreads are still relatively high (around 0.8), while they are significantly lower for the US-UK and US-CA series, ranging from 0.3 to 0.46.

Maturity (months)	Mean	St Dev	Min	Max	Skew	Kurt	Corr(3)	Corr(12)	Corr(60)	Corr(120)
US-AU Spread										
3	-2.17	1.66	-5.50	0.94	0.11	1.96	1.00			
12	-2.02	1.66	-5.19	1.15	0.16	1.92	0.98	1.00		
60	-1.68	1.06	-3.89	0.47	0.11	1.99	0.94	0.98	1.00	
120	-1.33	0.72	-3.20	0.14	-0.36	2.61	0.67	0.74	0.81	1.00
US-CA Spread										
3	-0.24	0.98	-2.59	2.32	0.48	3.11	1.00			
12	-0.25	0.95	-2.43	2.34	0.38	2.85	0.97	1.00		
60	-0.31	0.61	-1.83	0.80	-0.22	2.52	0.92	0.96	1.00	
120	-0.22	0.58	-2.02	0.66	-1.07	3.78	0.46	0.49	0.62	1.00
US-CH Spread										
3	1.64	1.59	-1.09	4.52	0.18	1.44	1.00			
12	1.72	1.61	-1.03	4.66	0.19	1.42	0.98	1.00		
60	1.81	1.00	-0.13	4.09	0.28	2.09	0.95	0.98	1.00	
120	1.85	0.58	0.15	3.37	0.22	2.71	0.80	0.84	0.90	1.00
US-JP Spread										
3	2.63	2.16	-0.46	6.28	0.00	1.38	1.00			
12	2.83	2.17	0.01	6.75	-0.01	1.44	0.99	1.00		
60	3.00	1.40	0.36	5.77	-0.28	2.03	0.95	0.97	1.00	
120	2.91	0.90	0.76	4.79	-0.46	2.76	0.78	0.81	0.90	1.00
US-UK Spread										
3	-1.03	1.06	-3.40	0.89	-0.59	2.18	1.00			
12	-0.88	0.98	-3.23	0.65	-0.66	2.31	0.95	1.00		
60	-0.60	0.64	-2.47	0.88	-0.08	2.52	0.83	0.90	1.00	
120	-0.29	0.67	-2.39	1.39	-0.45	4.04	0.30	0.33	0.59	1.00

Table 4.1. Descriptive statistics of sovereign yield spreads at monthly frequency for the time period from 1995:01 - 2013:12. For each spread and selected maturities (3-months, 12-months, 60-months and 120-months) we report (from left to right) mean, standard deviation, minimum, maximum, skewness, kurtosis and correlations between the reported maturities.

Overall we find that the different yield spread curve data sets share some common characteristics:

- The short end of the spread curve is more volatile than the long end;
- Average spread curves are upward sloping;
- Yield spreads at different maturities are highly correlated.

These properties are relatively similar to characteristics also commonly found in yield curve datasets, see, for example, Pooter et al. (2010) or Koopman and van der Wel (2013). However, yield curves usually exhibit a steeper and more concave sloping average curve and higher volatility at different maturities. Furthermore, yields are usually not negative, while the sovereign spreads considered in this study are typically negative for three of the currency pairs (US-AU, US-CA, US-UK).

4.3.3 Dynamic Behaviour

The aggregated descriptive statistics should also not hide the fact that spread curves may differ significantly from characteristic yield curve shapes. We provide illustrative plots of sovereign spread curves on selected days during the sample period in Figure 4.1. Apparently they can take on a wide range of shapes through time, including upward and downward sloping, but other than yield curves, they are generally rather flat. This reflects the fact that in a highly connected global economy, advanced economies often face similar economic conditions and consequently often simultaneously experience an upward or downward sloping yield curve. Rather uncommon for yield curves, spread curves may also regularly contain several bumps. In Figure 4.2 we plot the dynamics of spreads for selected short (3-months), medium (36-months) and long (120-months) maturities for the considered sample period. The difference in the level of the different spreads is obvious: the US-JP spread is typically the highest, while the US-AU spread is usually the most negative, indicating the substantially higher interest rates in Australia. Nevertheless the different spreads move surprisingly coherent throughout time.

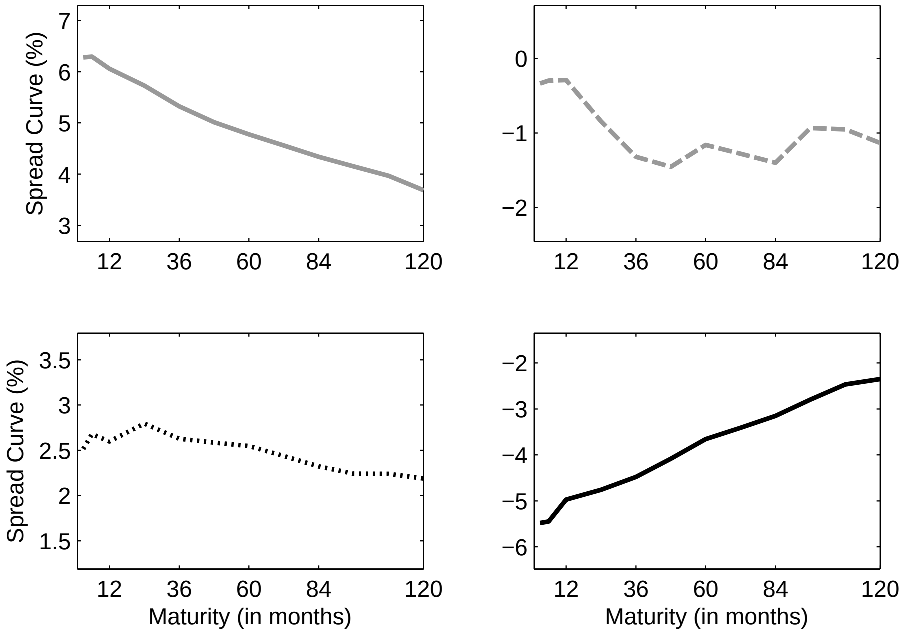

Figure 4.1. Representative sovereign spread curves for selected spreads and dates thoughout the sample period. The curves present the US-Japanese spread curve on 31 Aug 2000 (upper left panel), the US-British spread curve on 31 Dec 2008 (upper right panel), US-Swiss spread curve on 30 April 2010 (lower left panel) and the US-Australian spread curve on 31 Jan 2007 (lower right panel).

All spreads decrease for example after the bursting of the dotcom bubble in 2001 when US yields dropped more in relative terms than other advanced economies' yields. In particular spreads for short-term but also for medium-term maturities exhibit a characteristic drop during earlier periods of the global financial crisis (GFC) in 2007-2008, when the US significantly reduced short term interest rates to nearly zero. Towards the end of the crisis we observe an upwards shift of spreads for Australia, Canada and the UK, since also these countries started to significantly reduce interest rates.

After the GFC, especially short term spreads exhibit a striking behavior. While prior to the crisis period, there is a large temporal variation in the short term spreads, in its aftermath all spreads - except the US-Australian - narrow and remain flat until the end of the sample period in December 2013. This is obviously a direct consequence of the unprecedented expansive monetary policy of the major central banks. During the financial crisis the major central banks (except the Australian RBA, as Australia had been impacted less by the GFC) decreased their policy rate close to the zero bound and also directly intervened in the markets to bring yields down.[11] The central banks' long term commitment to these policies and the rather dire economic prospects have led to a prolonged period of low and non-volatile short and medium yields in most advanced economies. This unique interest rate environment is naturally reflected in short and less distinctively in medium term sovereign spreads as well.

Comparing short, medium and long term maturities, the difference in volatility mentioned above is clearly noticeable. While short term spreads are quite volatile, long term spreads remain rather stable throughout the sample period as the underlying long term structural differences between economies do not change as quickly as short term economic fluctuations.

[11]The most prominent example is, of course, the controversial quantitative easing of the US Federal Reserve.

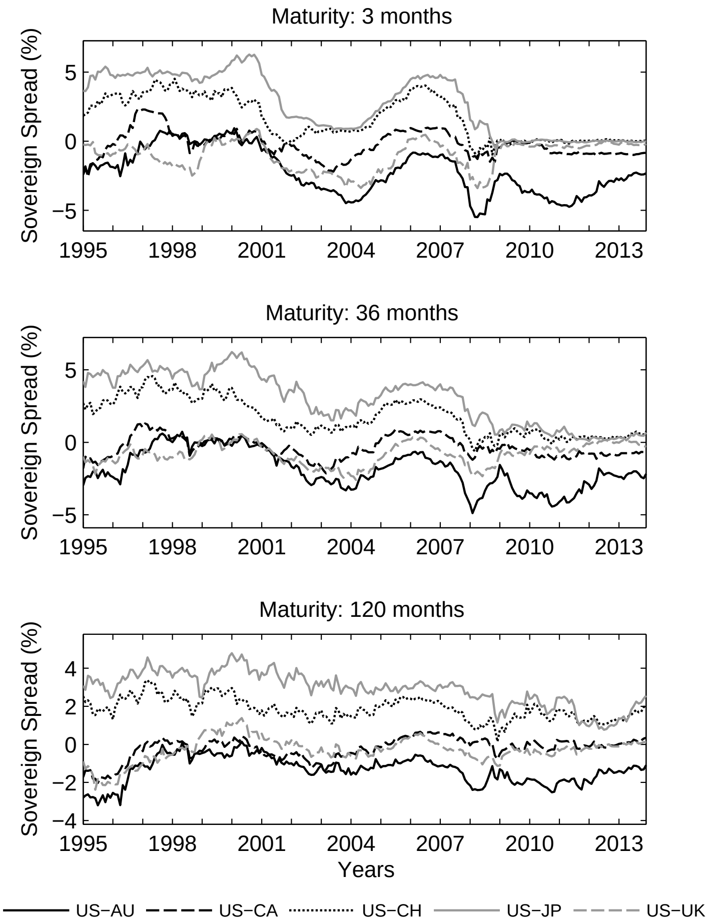

Figure 4.2. Time series for sovereign spreads of selected short term (three-month), medium term (36-month) and long term (120-month) maturities. We plot US-Australian, US-Canadian, US-Swiss, US-Japanese and US-British sovereign yield spreads for the sample period 1995:01 – 2013:12.

4.4 Latent Yield Spread Curve Factors

4.4.1 Estimating the Latent Factors

Our main objective is to identify and investigate the underlying factors driving the term structure of sovereign spreads. To derive these factors we conduct a principal component analysis (PCA). Principal component analysis is a statistical method that reduces the dimensionality of a data set by compressing the information it contains into a limited number of components or factors.[12] These factors thus summarize the main features of the original term structure of sovereign yield spreads in parsimonious form. PCA is a common approach applied to term structure dynamics (Longstaff et al., 2007; Blaskowitz and Herwartz, 2009; Barber and Copper, 2012) and works best with correlated time series, see, e.g., Duffee (2011). The approach therefore seems a natural choice to reduce the dimension of the highly correlated sovereign spreads. Practitioners may also benefit from the flexibility of factors that are not postulated a priori, but are rather derived from actual market data. Note that we apply PCA directly to the sovereign spreads and use standardized spreads with zero mean and unit variance.

To derive the orthogonal factors $F_{1,...,K}$ that can account for the variability in the term structure, assume S to be a TxN matrix of standardized sovereign spreads Δsy_t^τ, where T is the number of maturities and N is the number of observation dates. To extract loadings γ_K and factors F_K a PCA seeks an orthogonal KxT matrix Γ which yields a linear transformation

$$\Gamma S = \Phi, \tag{4.5}$$

where Φ is a KxN-dimensional matrix of latent factors F.

The matrix Γ is constructed using an eigenvector decomposition. Let Σ denote the TxT covariance matrix of S that can be decomposed as

$$\Sigma = \Gamma \Lambda \Gamma', \tag{4.6}$$

where the diagonal elements of $\Lambda = diag(\lambda_1, ..., \lambda_T)$ are the eigenvalues and the columns of Γ are the eigenvectors. Arranging the eigenvectors in decreasing order of the eigenvalues, the first K eigenvectors of Γ denote the factor loadings $[\gamma_1, ..., \gamma_K]$. Then the K latent factors $[F_1, ..., F_K]$ are defined by $F_{k,t} = \gamma_k' S_t$, where S_t is a T-dimensional vector of the term structure of sovereign yield spreads at time t. Given the wealth of literature detailing the use of PCA for examining term structure dynamics, we refer to Jolliffe

[12] Note that in the following we will use the terms factors and components interchangeable throughout the analysis.

(2002); Lardic et al. (2003) or Barber and Copper (2012) for further details.

4.4.2 Factor Dynamics

Applying PCA allows for a data-driven selection of the K most important latent factors. Table 4.2 displays the variance explained by the first three principal components extracted by the applied PCA.

	US-AU	US-CA	US-CH	US-JP	US-UK
F_1	91.8	84.2	94.3	93.4	77.5
F_2	7.1	13.9	4.6	6.0	19.9
F_3	0.7	1.1	0.6	0.4	1.7
Total	99.6	99.2	99.5	99.8	99.1

Table 4.2. Explained variance of first three principal components (F_1, F_2, F_3) in percent extracted by principal component analysis (PCA) for US-Australian, US-Canadian, US-Swiss, US-Japanese and US-British sovereign yield spread curves over the time period 1995:01 – 2013:12.

For all spreads the three leading principal components already account for about 99% of the variance in the term structure, with the first principal component playing the most dominant role. The first factor already explains more than 90% of the variance for three out of the five spreads. For US-UK and US-CA sovereign spreads the explanatory power of the first component is slightly lower, but still explains 84.2%, respectively 77.5%, of the variance. The second factor explains a further 4.6% up to 19.9% and the third factor an additional 0.4% to 1.7%. Including the first two factors in our analysis is an obvious choice. Note that we decided to also include the third factor, as this allows us to interpret the factors in line with common yield curve models in the subsequent sections.

We plot the time series of the first three estimated factors F_1, F_2, F_3 in Figure 4.3, which reveals that the three factors behave quite differently.[13] The first factor is the most volatile and seems to be relatively persistent. It also seems to mirror the dynamic behavior of the yield spreads relatively closely. Most

[13]Note that, as the PCA extracts orthogonal factors the correlation between the first, second and third factor is zero. All factor series have a zero mean.

prominent is the characteristic drop at the beginning of the GFC. The second factor is relatively noisy. It also exhibits a distinctive spike towards the end of the GFC. The third factor is the least volatile and also relatively small in magnitude. Comparing the factors of the different spread pairs, it is also interesting to note that the respective time series of the different spread pairs move together relatively closely through time. This holds especially for the first factor.

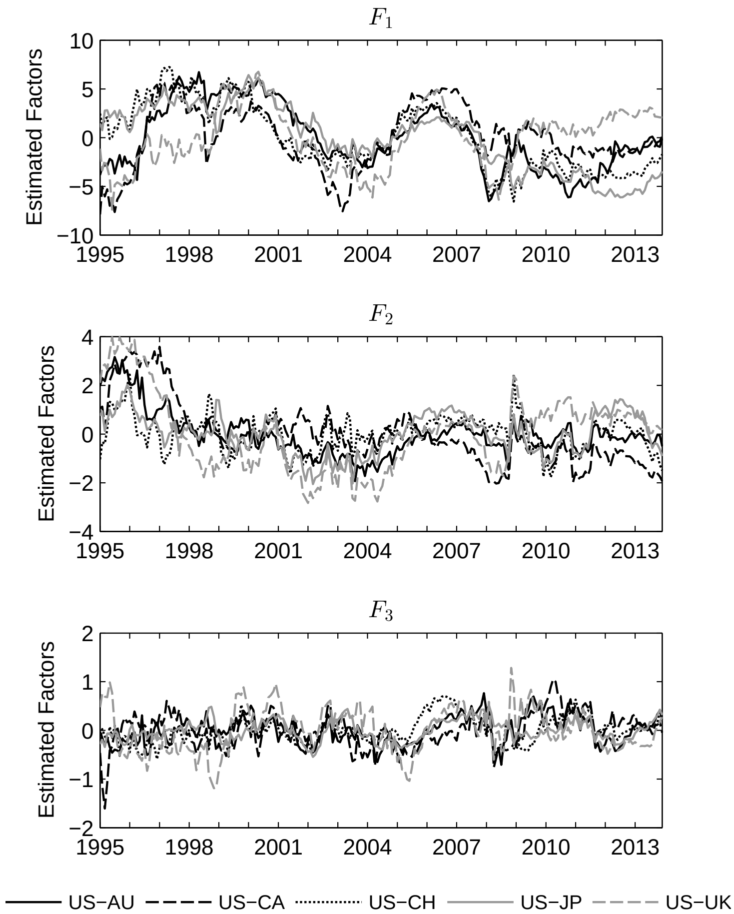

Figure 4.3. Time series of the first three Factors (F_1, F_2, F_3) estimated by principal component analysis (PCA) for US-Australian, US-Canadian, US-Swiss, US-Japanese and US-British sovereign yield spread curves for the time period 1995:01 – 2013:12.

4.4.3 Interpreting the Factors as Level, Slope and Curvature

Our methodological framework allows us to further analyze and interpret the estimated latent factors. To start with, we present the shape of the loadings $\gamma_1, \gamma_2, \gamma_3$ on the first three estimated latent factors as a function of maturity in Figure 4.4. The loadings are surprisingly similar for all sovereign spread

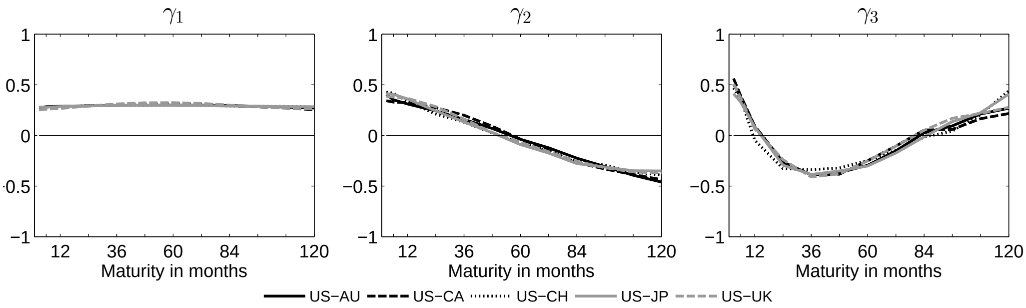

Figure 4.4. Loadings as a function of maturity for all spreads against US yields. We plot the loadings derived by PCA for the first three factors for US-Australian, US-Canadian, US-Swiss, US-Japanese and US-British sovereign spreads. Note that principal components are not unique up to sign, i.e. multiplying a principal component by (-1) has no effect on the explanatory power of the component.

pairs. Given the differences in sign and magnitude of the different spreads, one could have expected different shapes in the loadings. Interestingly, the shape of the factor loadings is also quite similar to those found in other works where PCA has been applied the term structure of interest rates (Litterman and Scheinkman, 1991; Dai and Singleton, 2000; Afonso and Martins, 2012). The first loading is almost constant across all maturities and does not decay with longer maturities - hence the first component can be interpreted as a *level factor*. The loading is responsible for parallel shifts of the spread curve. The second loading has opposite signs at both ends of the spread term structure so it affects short-term and long-term spreads differently. Thus, it can be interpreted as a *slope factor*, determining variations in the slope of the spread curve. The third loading has equal signs at both ends of the maturity spectrum, but an opposite sign for medium-term maturities mainly affecting changes in the *curvature* of sovereign spreads. Therefore, the observed shape of the loadings allows us to interpret the components as spread level, slope and curvature factors.

We verify this interpretation by investigating the relationship between the

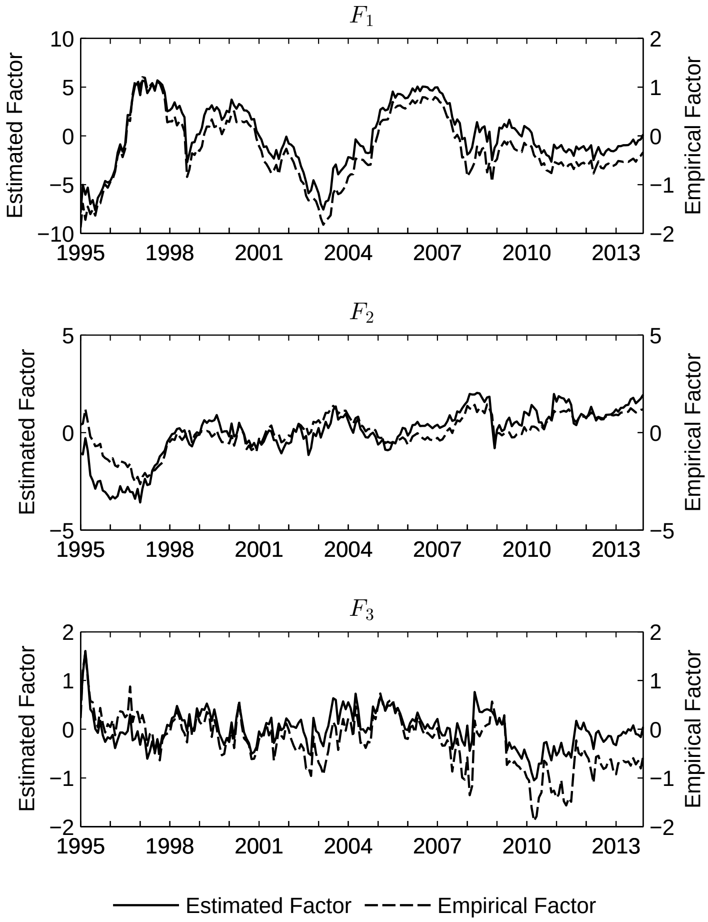

Figure 4.5. Time series of the first three estimated latent factors (F_1, F_2, F_3 —) against the empirical level, slope and curvature (- - -) for US-Canadian sovereign spreads. In line with existing literature, we calculate the empirical level as the average of the longest (120-months), the shortest (3-months) and a medium-term maturity (we chose 36-months); the empirical slope as the difference between the longest and shortest maturity and the curvature as twice the medium-term maturity minus the sum of the shortest and longest maturity. Note that the second and third factors are negatively correlated, thus the estimated factor series have been multiplied with (-1) for illustrative purposes.

estimated factor time series and the empirical spread level, slope and curvature. In line with the existing literature, see, e.g, Diebold et al. (2006); Afonso and Martins (2012), we calculate the empirical spread level as the average of the longest (120-months), the shortest (three-months) and a medium-term maturity (we chose 36-months), the spread slope as the difference between the longest and shortest maturity and the spread curvature as twice the medium-term maturity minus the sum of the shortest and longest maturity. Figure 4.5 provides an illustrative plot of the relationship between estimated and empirical factors for the US-Canadian spread.[14] For all three factor series the relation between the latent and empirical factors is visually apparent confirming a close relationship between estimated factors and their empirical proxies.

This also holds for all other investigated spread pairs. Table 4.3 summarizes the correlations between the empirical and estimated factor series for all spreads. The correlations are compellingly high especially for the first factor and confirm our interpretation that the three estimated latent factors correspond to level, slope and curvature of the yield spread curves. These

	US-AU	US-CA	US-CH	US-JP	US-UK
F_1 / level	0.99	0.99	0.99	0.99	0.97
F_2 / slope	-0.71	-0.87	-0.57	-0.66	-0.94
F_3 / curv.	-0.76	-0.83	-0.69	-0.62	-0.96

Table 4.3. Correlation between the time series of the first three estimated factors (F_1, F_2, F_3) and the empirical level, slope and curvature for US-Australian, US-Canadian, US-Swiss, US-Japanese and US-British sovereign spreads over the time period 1995:01-2013:12.

interpretations also make intuitive sense with regards to the shape of the spread curves and the percentage of the variation in yield spreads they explain. As spread curves are often rather flat, the slope factor plays a relatively smaller role than the level factor in explaining the variance. Interpreting the first factor as a level factor also helps to understand, why the first factors of the US-CA and especially the US-UK spread explain a relatively smaller fraction in the variance compared to the other spreads. As indicated, all datasets exhibit high correlation between the maturities - they are driven by the same '*level*' factor. As illustrated in Table 4.1, for the US-UK spread and to a lesser extent for the US-CA spread, the correlation between yield spreads at different maturities is less pronounced than for the other spread datasets. The correlation between the three-months and 120-months US-UK

[14] Corresponding plots for the remaining spread pairs are reported in Appendix B.1.

spread, for example, is only 0.3. Thus, a level factor explains less variance for the spreads between US and UK yields than for the other spreads.

Overall, the conducted analysis suggests that the dynamics of the entire term structure for sovereign spreads can be decomposed by using a small number of three latent factors. Further, the three factors can be suitably labelled as spread '*level* (F_L)', '*slope* (F_S)' and '*curvature* (F_C)' and are highly correlated with empirical measures of the factors. Our results are robust across sovereign spreads between advanced economies, namely spreads between the US and Australia, Canada, Japan, Switzerland and United Kingdom sovereign yields.

4.5 Exchange Rate Predictability

4.5.1 Rationale and Previous Findings

Section 4.2 concluded that the latent factors estimated from the term structure of yield spreads may serve as natural measures of exchange rate determination. The term structure of sovereign yield spreads contains information about expected cross-country differentials in economic fundamentals that are known to drive exchange rates. As the latent spread factors summarize this information they may also possess predictive power for the expected path of exchange rates.

In this regard, previous literature has produced some encouraging results. Clarida et al. (2003) provide evidence that the term structure of forward premiums contains valuable information for forecasting future spot exchange rates using a regime-switching vector equilibrium correction model. Bekaert et al. (2007) advocate that risk factors driving the premiums in the term structure of interest rates may drive the risk premiums in currency returns. Ang and Chen (2010) use the domestic empirical level and slope factors of the term structure together with interest rate volatility to predict FX rate returns in a cross-sectional setting based on portfolio strategies. They find an economically and statistically significant ability of changes in interest rates and slopes of the yield curve to predict foreign exchange returns, above the predictability of carry. Chen and Tsang (2013) examine the predictive power of cross-country yield curve factors constructed with the parametric Nelson-Siegel Model for exchange rates. They find that Nelson-Siegel factors extracted from two countries' relative yield curves can predict future exchange rate movements and excess currency returns up to 24 months ahead. Bui and Fisher (2016) confirm their findings for the relative yield curves of the US and Australia.

These results suggest that the macroeconomic information contained in the latent spread factors may be useful in exchange rate determination. To verify this assumption, we examine the ability of the extracted yield spread factors to capture the variation in exchange rate changes and excess returns in the subsequent sections.

4.5.2 Exchange Rate Data

For the analysis we retrieve the AUD, CAD, CHF, JPY and GBP end-of-the-month exchange rates against the USD from Bloomberg. We consider the US as the home country, thus, the exchange rate is measured as the USD price per unit of foreign currency. Therefore, a rise in the exchange rate represents a depreciation of the USD and a lower value an appreciation of the USD.

Figure 4.6 plots the time series of the log exchange rates for the considered sample period. The majority of exchange rates are relatively erratic and volatile with only the JPY and GBP being relatively stable throughout the entire time period and mainly fluctuating around a long-term equilibrium value. The bursting of the dotcom bubble in 2001 does not seem to be clearly reflected, but all currencies except the JPY experience a characteristic depreciation against the USD throughout the GFC period in 2008. This is followed by a relatively quick and sharp recovery.

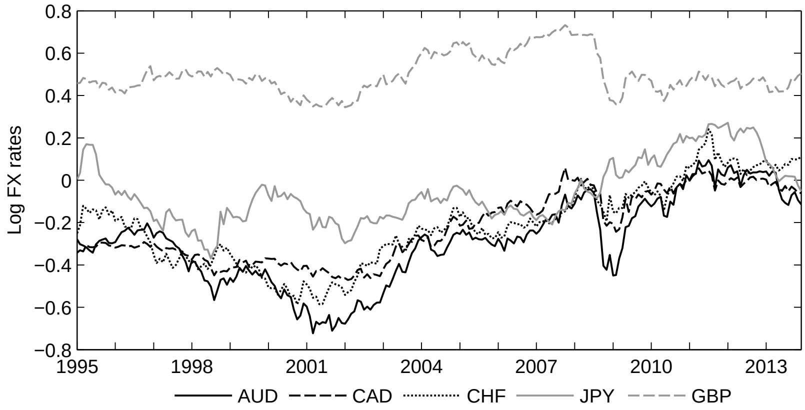

Figure 4.6. Time Series of log exchange rates for AUD, CAD, CHF, JPY (×100) and GBP against the USD for the sample period 2005:01 - 2013:12. The exchange rate is measured as the US dollar price per unit of foreign currency. Thus, an increase in the exchange rate represents a depreciation of the USD.

4.5.3 Estimation Specifications

To test the predictive ability of the latent yield spread factors we regress exchange rate movements and excess returns on the extracted factors for horizons $h = 3, 6, 12, 24$[15] months:

$$\Delta s_{t+h} = \alpha_h^{\Delta s} + \beta_{h,L}^{\Delta s} F_{L,t} + \beta_{h,S}^{\Delta s} F_{S,t} + \beta_{h,C}^{\Delta s} F_{C,t} + u_{t+h}; \tag{4.7}$$

$$xs_{t+h} = \alpha_h^{xs} + \beta_{h,L}^{xs} F_{L,t} + \beta_{h,S}^{xs} F_{S,t} + \beta_{h,C}^{xs} F_{C,t} + v_{t+h}. \tag{4.8}$$

The exchange rate return Δs_{t+h} for horizon h is defined as the annualized change of the log exchange rate s. Annualized excess returns xs_{t+h} are calculated by adjusting the exchange rate change of horizon h with the corresponding yield spread of equal maturity τ, see, e.g., Christiansen et al. (2011),

$$xs_{t+h} = \Delta s_{t+h} + i_t^{*,\tau} - i_t^{\tau}, \tag{4.9}$$

where i_t^{τ} and $i_t^{*,\tau}$ are the monthly domestic and foreign interest rates at maturity $h = \tau$.

[15]We start with $h = 3$ as the the shortest available maturity to calculate the excess returns xs is $m = 3 - months$. We note, that for Equation 4.7 the results for $h = 1$ months are consistent with the forecasting horizons presented in this analysis.

F_L, F_S, and F_C are the three latent spread factors extracted from the PCA above.[16] The predictive ability is evaluated by estimating $\beta^{\Delta s}_{h,[L,S,C]}$ and $\beta^{xs}_{h,[L,S,C]}$ over the entire sample. If the yield spread factors contain relevant information, these coefficients should be different from zero.

Regressions using longer time horizons need to address an inference bias due to overlapping observations (Harri and Brorsen, 2009). Since the horizon h for exchange rate movements and excess returns is longer than the frequency of data (one month in this case), the left hand side variable overlaps across observations and the error terms u_{t+h} and v_{t+h} will be moving average processes of order $h-1$. In this case OLS parameter estimates would be inefficient and hypothesis tests biased (Hansen and Hodrick, 1980). One way to deal with this problem is to only use non-overlapping observations. This would eliminate the autocorrelation problem, but is obviously highly inefficient as it dramatically reduces the number of observations and dismisses valuable information. An alternative and more efficient approach is to account for the moving average error term in hypothesis testing. Thus, we use heteroskedasticity and autocovariance consistent (HAC) estimators developed by Newey and West (1987) for the OLS estimation.[17]

Further, it is well known that the finacial crisis period in 2007-2009 has caused major eruptions in bond and foreign exchange markets. Bianchetti (2010), for example, finds that standard yield curve no-arbitrage relations are no longer valid. Fratzscher (2009) reports that the GFC has caused sharp movements in global exchange rate markets.

Thus, we first run equations (4.7) and (4.8) with the latent spread factors and their interaction with a GFC dummy (2007:08 to 2009:05).[18] We find significant coefficients[19] on the interaction terms for most of the spread pairs and horizons and conclude that the crisis period differs significantly from the rest of the sample. We thus drop the time period from 2007:08 to 2009:05

[16] We note that the null of a unit root is generally rejected for exchange rate movements, excess returns and the factor time series.

[17] Following Schwert (2002)'s method, we determine the number of lags of the residual autocorrelations as 12.

[18] Guidolin and Tam (2013) provide an extensive overview of the crisis dating literature and provide a conservative consensus dating centered around August 2007 - May 2009.

[19] Results for these regressions are reported in Appendix B.2.

for the predictive regressions.

4.5.4 Regression Results

The results of the predictive regressions are reported in Tables 4.4 – 4.8. We find predominantly negative coefficients on the yield spread level and slope factors for nearly all investigated horizons and currencies. For example, a 1% increase in the US-Australian yield spread level factor (i.e. the entire yield curve of the US shifts up by 1% relative to the Australian yield curve) predicts a 1.82% (annualized) depreciation of the Australian dollar against the US dollar and a 2.28% (annualized) drop in AUD excess return in the next quarter. Likewise, an 1% increase in the US-Swiss yield spread slope factor (i.e. the US yield curve becomes steeper relative to the Swiss one) predicts a 5.16% (annualized) depreciation of the Swiss franc over the next three months. The same 1% increase in the relative slope factor predicts an 5.84% drop in CHF (annualized) excess returns for the next three months.
There is no clear pattern for the sign and magnitude of the curvature factor and its estimated coefficients are often insignificant. Therefore it does not seem to play an important role in determining exchange rate movements and excess returns.[20] This is not entirely surprising, as it also only explains a relatively small amount of the variation in the term structure of yield spreads and has not clearly been linked to economic variables in previous literature.[21]
The negative coefficients on spread level and slope factors are in line with economic intuition and previous findings in the literature, see, in particular Chen and Tsang (2013). The yield curve literature described in Section 4.2 suggests that the level factor can be seen as a long-run inflation expectation factor, while the slope factor reflects business cycle and output growth dynamics. A higher yield curve level in a country thus indicates that the market

[20] Omitting the spread curvature factor from the regressions (results are not reported here) confirms that the additional explanatory power of including the third factor is indeed rather limited.
[21] Moench (2012) for example suggests a more indirect interpretation with innovations of the curvature factor announcing changes in the slope factor.

expects rising inflation and a flat yield curve points to the market expecting a forthcoming economic downturn. In these situations, a country's currency may be less desirable and will potentially face depreciation pressure. As Bui and Fisher (2016) describe, the currency will appreciate and recover towards its long-run equilibrium value. An increase in the yield spread level or slope factor (the home yield curve shifts up or the foreign yield curve becomes steeper) thus indicates a decrease in the nominal exchange rate equivalent to an appreciation of the USD or a depreciation of the foreign currency.

A similar logic applies to the spread level and slope factor coefficients in the excess return regressions. As noted above, excess foreign currency returns can be considered as risk premiums associated with holding a currency. An increase in the yield spread level and slope factors indicates higher expected foreign growth and lower expected foreign inflation. Thus, investors may demand smaller risk premiums for holding the foreign currency.

Note that these findings are contrary to the classic Dornbusch (1976) overshooting model but in line with recent empirical studies, see, e.g. Clarida and Waldmann (2008). The Dornbusch (1976) model would suggest an initial appreciation and subsequent depreciation of the home currency in response to a higher interest rate. However, there is empirical evidence indicating that quite regularly currencies of high interest rate countries tend to appreciate subsequently, rather than depreciate. Eichenbaum and Evans (1995) find that a rise in the U.S. federal funds rate can lead to persistent appreciation of the dollar for two years or longer. Gourinchas and Tornell (2004) also demonstrate that when investors systematically underestimate the persistence in the interest rate process, high interest rates in a country may lead to the subsequent appreciation of its currency. This is in line with our finding that an upward shift in the spread level or a flatter spread slope predict subsequent home currency appreciation and a high home risk premium. Since these movements are typically considered a signal for an economic slow-down or rising inflation, Chen and Tsang (2013) convincingly argue that in accordance with the present value relation, the home currency consequently faces depreciation pressure as investors pull out, and ceteris paribus, appreciates back up over time towards its long-term equilibrium value.

We obtain mixed results with regards to the significance of the coefficients for

US-Australian Spread

	FX Rate Change (Δs_{t+h}) Regression				Excess Return (xs_{t+h}) Regression			
	h=3	h=6	h=12	h=24	h=3	h=6	h=12	h=24
F_L	**-1.82*** **	**-1.71*** **	**-1.52*** **	**-0.84** **	**-2.28*** **	**-2.19*** **	**-2.00*** **	**-1.27*** **
	(-2.94)	(-3.03)	(-3.02)	(-2.17)	(-3.69)	(-3.88)	(-3.95)	(-3.28)
F_S	-2.23	-2.21	**-2.86***	**-3.63*** **	**-2.90***	**-2.81***	**-3.38** **	**-3.99*** **
	(-1.35)	(-1.41)	(-1.75)	(-2.66)	(-1.76)	(-1.78)	(-2.05)	(-2.90)
F_C	-3.23	-3.04	-1.80	-1.69	-4.10	-3.66	-1.89	-1.33
	(-0.48)	(-0.68)	(-0.40)	(-0.39)	(-0.61)	(-0.81)	(-0.42)	(-0.30)
nob	200	194	182	158	200	194	182	158
adj R^2	0.08	0.17	0.26	0.29	0.13	0.25	0.37	0.40

Table 4.4. Results for regressing USD/AUD exchange rate changes Δs_{t+h} (equation (4.7) and excess returns xs_{t+h} (equation 4.8) on latent US-AU spread factors F_L, F_S, F_C over the sample period 1995:01 - 2013:12. Newey-West robust t-statistics are reported in parantheses. *,** , *** indicate significance of the coefficients on a 10%, 5%, 1% level, respectively. Nob denotes number of observations, h denotes the forecasting horizon in months. Note that we omit the estimates of the constant. The time period from 2007:08 to 2009:05 (GFC) has been dropped from the sample.

US-Canadian Spread

	FX Rate Change (Δs_{t+h}) Regression				Excess Return (xs_{t+h}) Regression			
	h=3	h=6	h=12	h=24	h=3	h=6	h=12	h=24
F_L	-0.60	-0.63	**-0.67** **	**-0.63** **	**-0.87** **	**-0.90** **	**-0.94*** **	**-0.88*** **
	(-1.37)	(-1.57)	(-2.12)	(-2.44)	(-1.98)	(-2.27)	(-2.98)	(-3.36)
F_S	-0.40	-0.38	-0.87	**-1.12***	-0.73	-0.71	**-1.19** **	**-1.36** **
	(-0.55)	(-0.58)	(-1.58)	(-1.86)	(-1.01)	(-1.08)	(-2.15)	(-2.24)
F_C	-1.72	-2.58	-2.79	-1.81	-2.27	-2.97	-2.89	-1.57
	(-0.73)	(-1.10)	(-1.12)	(-1.10)	(-0.95)	(-1.27)	(-1.16)	(-0.95)
nob	200	194	182	158	200	194	182	158
adj R^2	0.01	0.06	0.20	0.30	0.04	0.12	0.32	0.43

Table 4.5. Results for regressing USD/CAD exchange rate changes Δs_{t+h} (equation (4.7) and excess returns xs_{t+h} (equation 4.8) on latent US-CA spread factors F_L, F_S, F_C over the sample period 1995:01 - 2013:12. Newey-West robust t-statistics are reported in parantheses. *,** , *** indicate significance of the coefficients on a 10%, 5%, 1% level, respectively. Nob denotes number of observations, h denotes the forecasting horizon in months. Note that we omit the estimates of the constant. The time period from 2007:08 to 2009:05 (GFC) has been dropped from the sample.

the examined currency pairs. The derived latent yield spread factors seem to exhibit the most significant predictive power for the safe haven currencies Swiss franc and Japanese yen.[22] For the Swiss franc both spread level as well as spread slope factor are statistically significant in predicting exchange rate movements and excess returns for three to 24 months, while for the Japanese

[22] Amongst others Ranaldo and Soderlind (2010) find that the Swiss franc and to a smaller extent the Japanese yen have significant safe-haven characteristics.

US-Swiss Spread

	FX Rate Change (Δs_{t+h}) Regression				Excess Return (xs_{t+h}) Regression			
	h=3	h=6	h=12	h=24	h=3	h=6	h=12	h=24
F_L	**-1.21***	**-1.18***	**-1.20****	**-1.18*****	**-1.64*****	**-1.64*****	**-1.67*****	**-1.60*****
	(-1.93)	(-1.94)	(-2.35)	(-4.02)	(-2.63)	(-2.69)	(-3.23)	(-5.42)
F_S	**-5.16*****	**-4.91*****	**-4.67*****	**-4.11*****	**-5.84*****	**-5.58*****	**-5.20*****	**-4.41*****
	(-2.59)	(-3.50)	(-3.19)	(-3.79)	(-2.94)	(-3.99)	(-3.51)	(-4.03)
F_C	3.91	5.60	2.62	-4.64	3.18	5.01	2.63	-4.23
	(0.58)	(0.97)	(0.41)	(-0.99)	(0.47)	(0.87)	(0.41)	(-0.89)
nob	200	194	182	158	200	194	182	158
adj R^2	0.07	0.15	0.28	0.42	0.11	0.22	0.39	0.53

Table 4.6. Results for regressing USD/CHF exchange rate changes Δs_{t+h} (equation (4.7) and excess returns xs_{t+h} (equation 4.8) on latent US-CH spread factors F_L, F_S, F_C over the sample period 1995:01 - 2013:12. Newey-West robust t-statistics are reported in paranthesis. *,** , *** indicate significance of the coefficients on a 10%, 5%, 1% level,respectively. Nob denotes number of observations, h denotes the forecasting horizon in months. Note that we omit the estimates of the constant. The time period from 2007:08 to 2009:05 (GFC) has been dropped from the sample.

US-Japanese Spread

	FX Rate Change (Δs_{t+h}) Regression				Excess Return (xs_{t+h}) Regression			
	h=3	h=6	h=12	h=24	h=3	h=6	h=12	h=24
F_L	-0.05	-0.16	-0.20	0.16	-0.65	-0.78	-0.82	-0.42
	(-0.09)	(-0.28)	(-0.39)	(0.31)	(-1.02)	(-1.33)	(-1.61)	(-0.82)
F_S	**-8.01*****	**-8.07*****	**-6.96*****	**-5.19*****	**-8.89*****	**-8.96*****	**-7.74*****	**-5.68*****
	(-5.32)	(-5.90)	(-5.87)	(-5.49)	(-5.88)	(-6.57)	(-6.53)	(-5.98)
F_C	2.24	5.95	9.58	5.07	1.32	5.24	9.42	5.63
	(0.23)	(0.88)	(1.56)	(1.09)	(0.14)	(0.77)	(1.54)	(1.21)
nob	200	194	182	158	200	194	182	158
adj R^2	0.08	0.19	0.34	0.30	0.11	0.24	0.42	0.39

Table 4.7. Results for regressing USD/JPY exchange rate changes Δs_{t+h} (equation (4.7) and excess returns xs_{t+h} (equation 4.8) on latent US-JP spread factors F_L, F_S, F_C over the sample period 1995:01 - 2013:12. Newey-West robust t-statistics are reported in parantheses. *,** , *** indicate significance of the coefficients on a 10%, 5%, 1% level, respectively. Nob denotes number of observations, h denotes the forecasting horizon in months. Note that we omit the estimates of the constant. The time period from 2007:08 to 2009:05 (GFC) has been dropped from the sample.

yen the slope factor seems to play the most dominant role.

We also find equally promising results for the Australian dollar. The US-AU spread level factor has significant explanatory power for exchange rate movements and excess returns across all horizons, while the US-AU spread slope factor is mostly significant for long horizons.

For the Canadian dollar we find significant results only for longer horizons. We assume that these results are mainly due to the Canadian dollar's commodity currency status, as characterized by Chen and Rogoff (2003). The currency responds mainly to the world price of the country's primary commodity exports and thus appears to be dominated by factors that are not

US-British Spread

	FX Rate Change (Δs_{t+h}) Regression				Excess Return (xs_{t+h}) Regression			
	h=3	h=6	h=12	h=24	h=3	h=6	h=12	h=24
F_L	-0.50	-0.49	-0.61	**-0.57****	-0.76	**-0.76***	**-0.88****	**-0.82*****
	(-1.03)	(-1.08)	(-1.50)	(-2.00)	(-1.57)	(-1.68)	(-2.16)	(-2.87)
F_S	-0.82	-0.75	-0.42	-0.38	-1.22	-1.15	-0.79	-0.62
	(-0.97)	(-0.93)	(-0.64)	(-0.70)	(-1.46)	(-1.43)	(-1.18)	(-1.13)
F_C	-0.11	-1.44	-1.33	0.92	-0.67	-1.82	-1.35	1.12
	(-0.03)	(-0.48)	(-0.69)	(0.52)	(-0.21)	(-0.61)	(-0.70)	(0.64)
nob	200	194	182	158	200	194	182	158
adj R^2	0.01	0.03	0.09	0.14	0.03	0.08	0.18	0.26

Table 4.8. Results for regressing USD/GBP exchange rate changes Δs_{t+h} (equation (4.7) and excess returns xs_{t+h} (equation 4.8) on latent US-UK spread factors F_L, F_S, F_C over the sample period 1995:01 - 2013:12. Newey-West robust t-statistics are reported in parantheses. *,** , *** indicate significance of the coefficients on a 10%, 5%, 1% level, respectively. *Nob* denotes number of observations, *h* denotes the forecasting horizon in months. Note that we omit the estimates of the constant. The time period from 2007:08 to 2009:05 (GFC) has been dropped from the sample.

directly related to its macroeconomic fundamentals in the short term. Krippner (2006) also found that the USD/CAD exchange rate is rather unrelated to the difference in interest rates due to the cyclical component of Canadian interest rates.

The results for the British pound appear to be the weakest among all currency pairs. We do not find consistently significant predictive power in any of the yield curve factors for the GBP. As indicated by the low correlation between distant maturities in Section 4.3, the US-UK spread term structure seems to be somewhat disconnected. Thus, the different yield spread maturities are presumably not driven by common factors related to fundamentals and exchange rates to the same extent as for the other currency pairs examined in this study.

For all currencies, the explanatory power indicated by the adjusted coefficient of determination increases with lengthening horizon. This also agrees with previous findings that the relation between differences in interest rates and exchange rate dynamics seems to be more pronounced for longer horizons (Chinn and Meredith, 2004; Rossi, 2013).

One may argue that in the longer term exchange rate movements seem to be less affected by risk and more affected by the fundamental differentials incorporated in the yield spreads. Note, however, that as pointed out by

Boudoukh et al. (2008) higher levels of predictability with widening horizons are to be expected in longer term horizon regressions. As the sampling error that is almost surely present in small samples shows up in each regression, both the estimator and R^2 are proportional to the horizon. Therefore, better results for long horizons in the form of higher βs and increasing R^2s generally provide little if any evidence for a better forecasting performance over and above the one month horizon results. From this perspective, the increasing explanatory power of the applied models for longer horizons should be interpreted with care.
Overall, our predictive regression results show that the extracted yield spread factors can help to explain and forecast bilateral exchange rate movements and excess currency returns three month to two years ahead.[23] Most dominant are the spread level and slope factor, while the spread curvature factor seems to have no consistent explanatory power. The negative signs of the coefficients for the level and slope factors are consistent with economic theory and previous findings. These results provide strong evidence for the linkage between the term structure of yield spreads, macroeconomic fundamentals and exchange rates.

4.5.5 Comparison with the UIRP

Based on the insights described in Section 4.2, one way to understand the estimated latent spread factors in relation to exchange rate movements, is to interpret them as augmenting the traditional h-horizon UIRP approach with the information in spreads of additional maturities. Both latent spread factors and UIRP relate differences in interest rates between economies to changes in exchange rates. However, while the UIRP relation only uses the information content up to a certain maturity, the latent spread factors summarize the information embodied in the entire term structure of sovereign

[23]We also examine the assertion that US factors may drive the spread (with a negative sign) as an alternative explanation. However, except for the US-Japanese spread we mainly find small correlations between US yield and yield spread factors. US factors are also inferior in their explanatory power compared to the yield spread factors. See Appendix B.3.

yield spreads.

To formally test whether the latent spread factors provide additional, valuable information for predicting exchange rate movements, we apply a likelihood ratio (LR) test[24] between the traditional UIRP regression model

$$\Delta s_{t+h} = \alpha^{\Delta s}_{h,UIRP} + \beta^{\Delta s}_{h,UIRP}(i^h_t - i^{h*}_t) + \epsilon_{t+h} \tag{4.10}$$

based on the UIRP relation in equation (4.1) and an extended UIRP regression model which also includes the three latent spread factors F_L, F_S, and F_C:

$$\Delta s_{t+h} = \alpha^{\Delta s}_{h,UIRP} + \beta^{\Delta s}_{h,UIRP}(i^h_t - i^{*,h}_t) + \beta^{\Delta s}_{h,L}F_{L,t} + \beta^{\Delta s}_{h,S}F_{S,t} + \beta^{\Delta s}_{h,C}F_{C,t} + \epsilon_{t+h}. \tag{4.11}$$

The LR test is commonly used to evaluate the difference between two nested models, e.g. when the simpler model is a special case of the more complex model. It is based on a comparison of the maximum likelihood of the two models.[25] If L_{UIRP} is the likelihood of the simple UIRP model in equation (4.10) and $L_{UIRPext}$ is the likelihood of the more complex model in equation (4.11) the LR test statistic (LRT) is calculated as

$$LRT = -2\,log\,\frac{L_{UIRP}}{L_{UIRPext}}. \tag{4.12}$$

Asymptotically, the test statistic follows a chi-square distribution, with the degrees of freedom equal to the difference in the number of parameters between the two models.

We present the results of the conducted LR tests in Table 4.9. The LRT-values indicate that using the information of the entire yield spread term structure summarized in the latent spread factors clearly provides additional explanatory power in comparison to a simple UIRP regression. We find significant LR test statistics, often at the 1% level, for nearly all currencies and

[24] We note that applying an F-Test leads to similar conclusions.

[25] Adding additional parameters will always result in a higher likelihood score. However, the LR test provides an objective criterion whether the difference in likelihood scores among the two models is statistically significant considering the loss of degrees of freedom for the more complex model.

horizons except for the British pound ($h = 3$ and $h = 24$) and the Canadian dollar ($h = 3$ and $h = 6$). Not surprisingly, we find the strongest results for the two safe haven currencies Swiss franc and Japanese yen.

These results make intuitive sense when the exchange rate is understood as

	LR Test Statistic			
	Horizon			
Spread	h=3	h=6	h=12	h=24
US-AU	**7.00***	0.57	**26.50*** **	**51.53*** **
US-CA	0.26	2.26	**6.85***	**21.36*** **
US-CH	**12.99*** **	**24.94*** **	**41.34*** **	**61.52*** **
US-JP	**21.59*** **	**42.69*** **	**70.55*** **	**64.48*** **
US-UK	2.50	**9.87*** *	**7.46***	1.12
nob	198	192	180	156

Table 4.9. Results of a likelihood ratio (LR) test between the simple UIRP model based on equation (4.10) and an extended UIRP model also including the three latent spread factors F_L, F_S, F_C in equation (4.11). We present the LR test statistics and p-values for all considered sovereign spread pairs and horizons h=3, h=6, h=12, h=24. *, ** , *** indicates a significantly superior performance of the extended model at the 10%, 5%, 1% level, respectively.

an asset price. A specific maturity, such as a 12-month yield only embodies information for the time period until the underlying instrument matures. However the exchange rate as an asset price is determined to a large extent by the expected long term future values of the fundamentals. While these can obviously not be reflected in a short or medium term yield, they are reflected in the yield spread factors which summarize the information embodied in the entire term structure up to 120 months. Thus, our results strongly support the additional predictive power of the extracted latent factors from the term structure of sovereign yield spreads over a standard UIRP relationship. These findings also suggest that it may be useful to apply these factors in future empirical exchange rate forecasting studies.

4.6 Conclusion

This paper provides a novel analysis of the term structure of sovereign yield spreads. We focus on the yield spreads of advanced economies with highly liquid markets of bonds issued in their own currency, little to no credit or default risk and free floating exchange rates (Australia, Canada, Switzerland, Japan, UK and US). Using a monthly frequency, we investigate all yield spreads against the US for the time period from January 1995 to December 2013.

Our main objective is to derive and examine latent factors driving the term structure of sovereign yield spreads that, to the best of our knowledge, has not been thoroughly studied in the literature yet. We apply principal component analysis on each of the five sovereign spread data sets. Our analysis shows that the term structure of all sovereign spreads is driven by similar factors and the first three estimated factors are already sufficient to explain more than 99% of the variation in the entire spread term structure. Interestingly, the identified factors show a very similar shape to those reported in studies analysing the term structure of interest rates, see, e.g., Litterman and Scheinkman (1991); Bikbov and Chernov (2010), and can be labeled as spread level, spread slope and spread curvature.

We further find that the extracted yield spread factors can explain and predict bilateral exchange rate movements and excess returns three months to two years ahead. Most dominant are the spread level and spread slope factor. The negative signs of the predictive regression coefficients on these factors indicate that an increase in the spread level or spread slope factor, i.e. when the foreign yield curve shifts down or becomes steeper relative to the US, predicts a depreciation and smaller excess returns of the foreign currency against the US dollar. Our results are also consistent with recent empirical findings in related studies (Chen and Tsang, 2013; Bui and Fisher, 2016).

With regards to economic theory, in contrast to the pattern suggested by Dornbusch (1976)'s overshooting model, our results rather suggest an immediate depreciation and subsequent appreciation of a currency in response to a higher interest rate. However, these results are consistent with more recent empirical studies, see, e.g., Eichenbaum and Evans (1995); Gourinchas and Tornell (2004); Clarida and Waldmann (2008), where it is proposed that an increase in interest rates in a country may lead to a persistent appreciation of its currency.

When we test the additional explanatory power of the extracted spread factors in comparison to the traditional UIRP approach, we find significant results for most currencies and horizons. We therefore infer that using the information of the entire spread curve summarized in the spread factors clearly adds valuable information.

This finding make intuitive sense when the exchange rate is understood as an asset price. Within a present value framework, exchange rates rely more on future than on current fundamentals. In other words, the exchange rate as an asset price is determined to a large extent by the long term expected values of fundamentals. While the UIRP relation only reflects information up to a limited horizon, these long term fundamentals are reflected in the yield spread factors which summarize the information of the entire term structure up to 120 months ahead. Thus, the term structure of yield spreads may provide more accurate information for expected fundamentals.

Our results have several important implications for future studies in this area. To start with, our results highlight the importance of the term structure of yield spreads as a factor containing valuable macro-financial information. While the spread between certain maturities is subject to the enormous body of UIRP-literature, the term structure of spreads has been widely neglected so far. We have identified and successfully labeled the latent driving forces of the spread term structure, but more research is required to fully understand the fundamental information it contains. Furthermore, we provide additional evidence of the link between interest rates, macroeconomic fundamentals and exchange rates and confirm the view that the exchange rate can be modeled as an asset price. Considering the widespread forecasting failure of empirical exchange rate models based on observable macroeconomic fundamentals (Meese and Rogoff, 1983; Rossi, 2013), the estimated latent spread factors may be particularly helpful in future forecasting studies. Recent results point out that this empirical failure may also be a result of using inappropriate proxies for the market expectations of future fundamentals rather than the failure of the models themselves (Bacchetta and van Wincoop, 2013; Balke et al., 2013). Including the fundamental information embodied in the latent factors of sovereign yield spreads may thus be a promising approach to improve the forecasting accuracy of traditional exchange rate models.

Appendix B

B.1 Additional Plots of Estimated and Empirical Factors

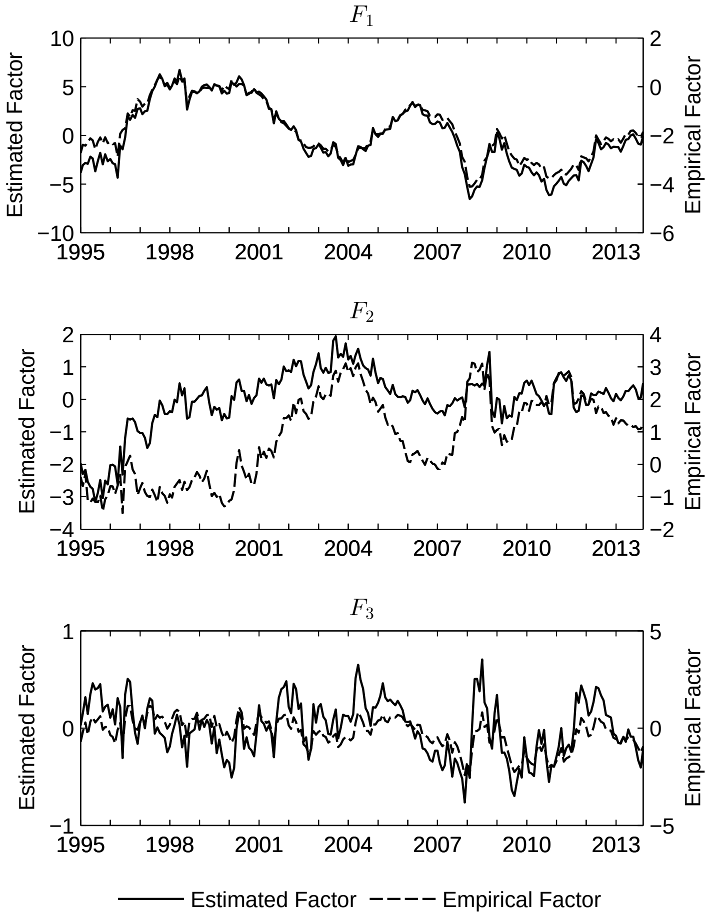

Figure B.1. Time series of the first three estimated latent factors (F_1, F_2, F_3 —) against the empirical level, slope and curvature (- - -) for US-Australian sovereign spreads. In line with existing literature, we calculate the empirical level as the average of the longest (120-months), the shortest (3-months) and a medium-term maturity (we chose 36-months); the empirical slope as the difference between the longest and shortest maturity and the curvature as twice the medium-term maturity minus the sum of the shortest and longest maturity. Note that the second and third factors are negatively correlated, thus the estimated factor series have been multiplied with (-1) for illustrative purposes.

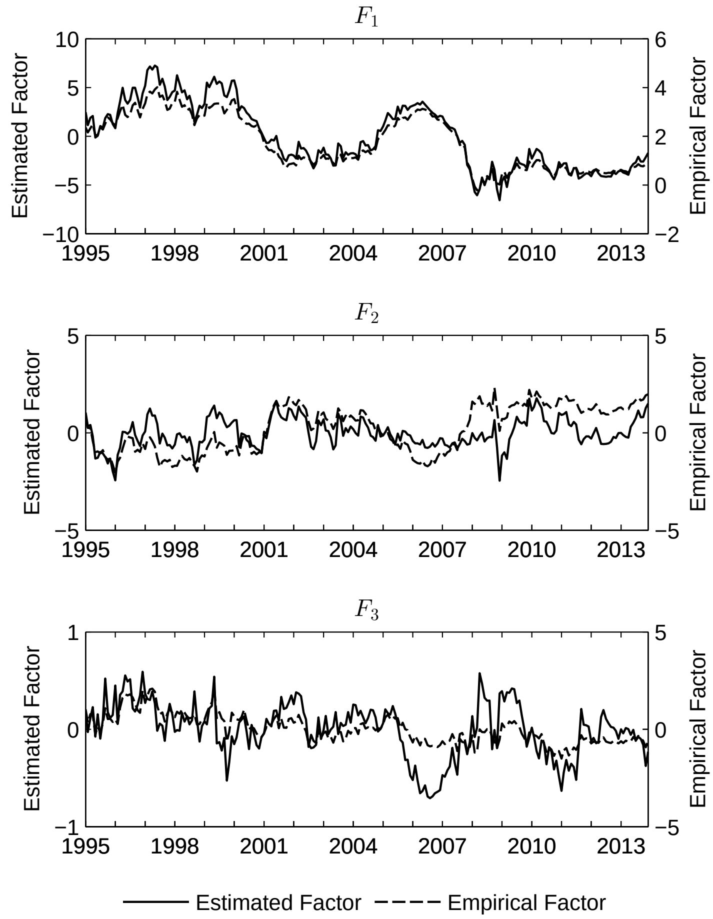

Figure B.2. Time series of the first three estimated latent factors (F_1, F_2, F_3 —) against the empirical level, slope and curvature (- - -) for US-Swiss sovereign spreads. In line with existing literature, we calculate the empirical level as the average of the longest (120-months), the shortest (3-months) and a medium-term maturity (we chose 36-months); the empirical slope as the difference between the longest and shortest maturity and the curvature as twice the medium-term maturity minus the sum of the shortest and longest maturity. Note that the second and third factors are negatively correlated, thus the estimated factor series have been multiplied with (-1) for illustrative purposes.

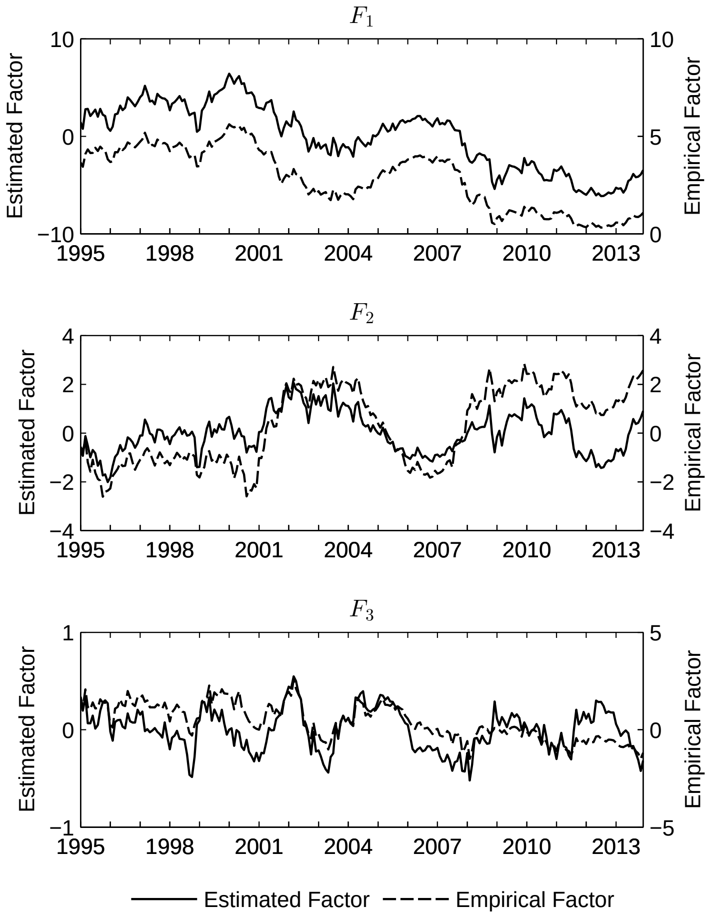

Figure B.3. Time series of the first three estimated latent factors (F_1, F_2, F_3 —) against the empirical level, slope and curvature (- - -) for US-Japanese sovereign spreads. In line with existing literature, we calculate the empirical level as the average of the longest (120-months), the shortest (3-months) and a medium-term maturity (we chose 36-months); the empirical slope as the difference between the longest and shortest maturity and the curvature as twice the medium-term maturity minus the sum of the shortest and longest maturity. Note that the second and third factors are negatively correlated, thus the estimated factor series have been multiplied with (-1) for illustrative purposes.

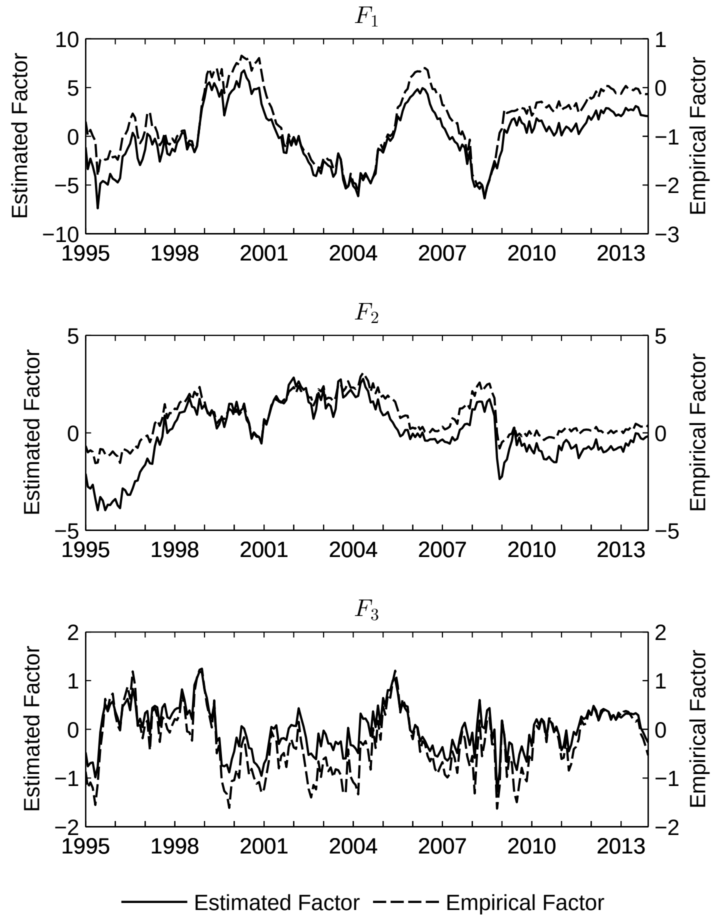

Figure B.4. Time series of the first three estimated latent factors (F_1, F_2, F_3 —) against the empirical level, slope and curvature (- - -) for US-British sovereign spreads. In line with existing literature, we calculate the empirical level as the average of the longest (120-months), the shortest (3-months) and a medium-term maturity (we chose 36-months); the empirical slope as the difference between the longest and shortest maturity and the curvature as twice the medium-term maturity minus the sum of the shortest and longest maturity. Note that the second and third factors are negatively correlated, thus the estimated factor series have been multiplied with (-1) for illustrative purposes.

B.2 Regression Results with GFC Interaction Terms

US-Australian Spread

	FX Rate Change (Δs_{t+h}) Regression				Excess Return (xs_{t+h}) Regression			
	h=3	h=6	h=12	h=24	h=3	h=6	h=12	h=24
F_L	**-1.84*** **	**-1.73*** **	**-1.48*** **	**-0.75* **	**-2.29*** **	**-2.21*** **	**-1.96*** **	**-1.18*** **
	(-2.84)	(-2.86)	(-2.61)	(-1.74)	(-3.56)	(-3.64)	(-3.44)	(-2.75)
F_S	-2.20	-2.09	**-2.65* **	**-3.83*** **	**-2.86* **	**-2.68* **	**-3.17** **	**-4.18*** **
	(-1.37)	(-1.41)	(-1.69)	(-2.81)	(-1.79)	(-1.81)	(-2.00)	(-3.06)
F_C	-1.50	-0.55	2.39	-2.88	-2.37	-1.17	2.28	-2.50
	(-0.23)	(-0.11)	(0.45)	(-0.69)	(-0.37)	(-0.24)	(0.43)	(-0.59)
F_L x GFCd	**5.33*** **	**7.03*** **	**5.61*** **	0.71	**5.30*** **	**7.03*** **	**5.65*** **	0.71
	(2.71)	(5.30)	(3.80)	(0.81)	(2.70)	(5.31)	(3.83)	(0.81)
F_S x GFCd	**25.48*** **	**18.64** **	7.28	6.51	**25.44*** **	**18.69** **	7.32	6.42
	(4.40)	(2.01)	(0.76)	(1.12)	(4.37)	(2.02)	(0.76)	(1.12)
F_C x GFCd	**62.91*** **	**59.29*** **	**-15.64* **	5.50	**63.07*** **	**59.17*** **	**-15.76* **	5.55
	(3.36)	(6.01)	(-1.66)	(1.14)	(3.37)	(5.99)	(-1.67)	(1.15)
nob	225	222	216	204	225	222	216	204
adj R^2	0.20	0.38	0.27	0.21	0.22	0.40	0.32	0.30

Table B.1. Results for regressing USD/AUD exchange rate changes Δs_{t+h} (equation 4.7) and excess returns xs_{t+h} (equation 4.8) on latent US-AU spread factors F_L, F_S, F_C and their interactions with a GFC dummy (from 2007:08-2009:05) over the sample period 1995:01 - 2013:12. Newey-West robust t-statistics are reported in parantheses. *,** , *** indicate significance of the coefficients at the 10%, 5%, 1% level, respectively. *Nob* denotes number of observations, h denotes the forecasting horizon in months. Note that we omit the estimates of the constant.

US-Canadian Spread

	FX Rate Change (Δs_{t+h}) Regression				Excess Return (xs_{t+h}) Regression			
	h=3	h=6	h=12	h=24	h=3	h=6	h=12	h=24
F_L	-0.51	-0.42	-0.39	**-0.51***	**-0.78***	-0.70	**-0.66***	**-0.75*****
	(-1.15)	(-0.96)	(-1.07)	(-1.94)	(-1.76)	(-1.59)	(-1.81)	(-2.86)
F_S	-0.65	-0.70	**-1.18***	**-1.22****	-0.98	-1.02	**-1.50****	**-1.45****
	(-0.81)	(-0.92)	(-1.96)	(-2.06)	(-1.22)	(-1.34)	(-2.48)	(-2.44)
F_C	-1.14	-2.26	-3.01	-2.55	-1.68	-2.65	-3.10	-2.31
	(-0.46)	(-1.06)	(-1.36)	(-1.37)	(-0.68)	(-1.24)	(-1.41)	(-1.23)
F_L x GFCd	2.79	-2.10	1.18	-0.14	2.78	-2.11	1.21	-0.10
	(0.62)	(-0.85)	(0.93)	(-0.15)	(0.61)	(-0.85)	(0.96)	(-0.11)
F_S x GFCd	**10.43*****	**11.88*****	**11.18*****	**3.51*****	**10.39*****	**11.89*****	**11.25*****	**3.53*****
	(3.75)	(6.67)	(6.34)	(3.68)	(3.74)	(6.68)	(6.39)	(3.67)
F_C x GFCd	**21.85****	15.97	**-17.47*****	-4.57	**21.84****	16.01	**-17.47*****	-4.67
	(2.32)	(1.58)	(-3.93)	(-1.57)	(2.32)	(1.59)	(-3.95)	(-1.60)
nob	225	222	216	204	225	222	216	204
adj R^2	0.10	0.22	0.31	0.23	0.11	0.24	0.35	0.33

Table B.2. Results for regressing USD/CAD exchange rate changes Δs_{t+h} (equation 4.7) and excess returns xs_{t+h} (equation 4.8) on latent US-CA spread factors F_L, F_S, F_C and their interactions with a GFC dummy (from 2007:08-2009:05) over the sample period 1995:01 - 2013:12. Newey-West robust t-statistics are reported in parantheses. *,** , *** indicate significance of the coefficients at the 10%, 5%, 1% level, respectively. Nob denotes number of observations, h denotes the forecasting horizon in months. Note that we omit the estimates of the constant.

US-Swiss Spread

	FX Rate Change (Δs_{t+h}) Regression				Excess Return (xs_{t+h}) Regression			
	h=3	h=6	h=12	h=24	h=3	h=6	h=12	h=24
F_L	**-1.22***	**-1.19***	**-1.12****	**-0.85****	**-1.66*****	**-1.65*****	**-1.59*****	**-1.27*****
	(-1.92)	(-1.93)	(-2.14)	(-2.14)	(-2.61)	(-2.68)	(-2.99)	(-3.17)
F_S	**-4.89****	**-4.44*****	**-3.87****	**-3.08****	**-5.57*****	**-5.12*****	**-4.40*****	**-3.37****
	(-2.43)	(-3.07)	(-2.33)	(-2.37)	(-2.78)	(-3.54)	(-2.61)	(-2.58)
F_C	4.79	7.22	7.50	5.27	4.05	6.64	7.56	5.74
	(0.73)	(1.35)	(1.31)	(1.14)	(0.62)	(1.25)	(1.31)	(1.23)
F_L x GFCd	0.78	2.19	1.86	0.72	0.74	2.20	1.94	0.73
	(0.61)	(1.45)	(1.31)	(0.87)	(0.58)	(1.46)	(1.39)	(0.87)
F_S x GFCd	**14.53****	**19.31*****	**8.91*****	**6.99*****	**14.48****	**19.33*****	**9.06*****	**6.98*****
	(2.53)	(2.85)	(3.36)	(2.87)	(2.51)	(2.86)	(3.47)	(2.85)
F_C x GFCd	**41.36****	18.63	-2.73	-3.72	**41.66****	18.50	-3.20	-3.71
	(2.14)	(1.38)	(-0.25)	(-0.57)	(2.16)	(1.37)	(-0.29)	(-0.57)
nob	225	222	216	204	225	222	216	204
adj R^2	0.08	0.16	0.23	0.28	0.11	0.22	0.32	0.40

Table B.3. Results for regressing USD/CHF exchange rate changes Δs_{t+h} (equation 4.7) and excess returns xs_{t+h} (equation 4.8) on latent US-CH spread factors F_L, F_S, F_C and their interactions with a GFC dummy (from 2007:08-2009:05) over the sample period 1995:01 - 2013:12. Newey-West robust t-statistics are reported in parantheses. *,** , *** indicate significance of the coefficients at the 10%, 5%, 1% level, respectively. Nob denotes number of observations, h denotes the forecasting horizon in months. Note that we omit the estimates of the constant.

US-Japanese Spread

	FX Rate Change (Δs_{t+h}) Regression				Excess Return (xs_{t+h}) Regression			
	h=3	h=6	h=12	h=24	h=3	h=6	h=12	h=24
F_L	-0.03	-0.14	-0.16	0.12	-0.63	-0.76	-0.78	-0.46
	(-0.05)	(-0.24)	(-0.30)	(0.20)	(-0.97)	(-1.26)	(-1.42)	(-0.75)
F_S	**-7.72*** **	**-7.57*** **	**-5.97*** **	**-3.23** **	**-8.61*** **	**-8.47*** **	**-6.75*** **	**-3.71** **
	(-5.02)	(-5.41)	(-4.52)	(-2.20)	(-5.57)	(-6.06)	(-5.11)	(-2.52)
F_C	5.41	10.81	**15.62** **	**11.02** **	4.48	10.10	**15.47** **	**11.62** **
	(0.56)	(1.42)	(2.28)	(1.96)	(0.47)	(1.32)	(2.26)	(2.06)
F_L x GFCd	-0.27	-1.77	**-1.70** **	**-2.19** **	-0.30	-1.76	**-1.67** **	**-2.18** **
	(-0.16)	(-1.59)	(-2.51)	(-2.49)	(-0.18)	(-1.59)	(-2.47)	(-2.46)
F_S x GFCd	-7.32	8.83	**5.96***	4.45	-7.47	8.86	**6.09***	4.50
	(-0.49)	(1.19)	(1.71)	(1.32)	(-0.50)	(1.20)	(1.77)	(1.33)
F_C x GFCd	**44.66** **	15.35	14.01	10.65	**44.78** **	15.23	13.92	10.69
	(2.54)	(1.14)	(1.48)	(1.40)	(2.54)	(1.13)	(1.47)	(1.40)
nob	225	222	216	204	225	222	216	204
adj R^2	0.09	0.17	0.33	0.23	0.12	0.22	0.40	0.33

Table B.4. Results for regressing USD/JPY exchange rate changes Δs_{t+h} (equation 4.7) and excess returns xs_{t+h} (equation 4.8) on latent US-JP spread factors F_L, F_S, F_C and their interactions with a GFC dummy (from 2007:08-2009:05) over the sample period 1995:01 - 2013:12. Newey-West robust t-statistics are reported in parantheses. *,** , *** indicate significance of the coefficients at the 10%, 5%, 1% level, respectively. Nob denotes number of observations, h denotes the forecasting horizon in months. Note that we omit the estimates of the constant.

US-British Spread

	FX Rate Change (Δs_{t+h}) Regression				Excess Return (xs_{t+h}) Regression			
	h=3	h=6	h=12	h=24	h=3	h=6	h=12	h=24
F_L	-0.51	-0.49	-0.55	-0.42	-0.77	-0.76	**-0.82***	**-0.66** **
	(-1.05)	(-1.06)	(-1.30)	(-1.38)	(-1.58)	(-1.65)	(-1.92)	(-2.19)
F_S	-0.81	-0.73	-0.39	-0.47	-1.21	-1.13	-0.75	-0.71
	(-0.96)	(-0.90)	(-0.57)	(-0.77)	(-1.45)	(-1.39)	(-1.09)	(-1.16)
F_C	0.10	-1.17	-1.00	-1.08	-0.46	-1.55	-1.04	-0.87
	(0.03)	(-0.38)	(-0.54)	(-0.47)	(-0.14)	(-0.51)	(-0.56)	(-0.38)
F_L x GFCd	**6.05*** **	**4.99*** **	**3.47*** **	**2.23*** **	**6.02*** **	**4.98*** **	**3.51*** **	**2.25*** **
	(3.12)	(4.91)	(4.17)	(5.34)	(3.10)	(4.90)	(4.21)	(5.37)
F_S x GFCd	**9.57** **	**12.45*** **	**7.63*** **	**3.64*** **	**9.58** **	**12.47*** **	**7.62*** **	**3.58*** **
	(2.40)	(4.32)	(4.23)	(5.39)	(2.41)	(4.32)	(4.25)	(5.29)
F_C x GFCd	3.55	11.28	-4.68	0.13	3.71	11.21	-4.86	0.15
	(0.34)	(1.40)	(-0.82)	(0.05)	(0.36)	(1.39)	(-0.86)	(0.06)
nob	225	222	216	204	225	222	216	204
adj R^2	0.23	0.47	0.39	0.26	0.22	0.46	0.39	0.27

Table B.5. Results of regressing USD/GBP exchange rate changes Δs_{t+h} (equation 4.7) and excess returns xs_{t+h} (equation 4.8) on latent US-UK spread factors F_L, F_S, F_C and their interactions with a GFC dummy (from 2007:08-2009:05) over the sample period 1995:01 - 2013:12. Newey-West robust t-statistics are reported in parantheses. *,** , *** indicate significance of the coefficients at the 10%, 5%, 1% level, respectively. Nob denotes number of observations, h denotes the forecasting horizon in months. Note that we omit the estimates of the constant.

B.3 Correlation of Yield Spread Factors with US Factors

	US-AU	US-CA	US-CH	US-JP	US-GB
Level	0.63	0.33	0.86	0.96	0.01
Slope	0.19	-0.07	0.59	0.78	0.20
Curvature	0.48	0.46	0.35	0.73	0.42

Table B.6. Correlation of yield spread factors with US yield factors over the sample period 1995:01 - 2013:12.

Bibliography

Afonso, A., Martins, M. M., 2012. Level, slope, curvature of the sovereign yield curve, and fiscal behaviour. Journal of Banking & Finance 36 (6), 1789–1807.

Ang, A., Chen, J., 2010. Yield Curve Predictors of Foreign Exchange Returns. AFA 2011 Denver Meetings Paper.

Ang, A., Piazzesi, M., Wei, M., 2006. What does the yield curve tell us about GDP growth? Journal of Econometrics 131 (1-2), 359–403.

Bacchetta, P., van Wincoop, E., 2013. On the unstable relationship between exchange rates and macroeconomic fundamentals. Journal of International Economics 91 (1), 18–26.

Backus, D., Gavazzoni, F., Telmer, C., Zin, S. E., 2010. Monetary Policy and the Uncovered Interest Parity Puzzle. NBER Working Paper 16218.

Balke, N. S., Ma, J., Wohar, M. E., 2013. The contribution of economic fundamentals to movements in exchange rates. Journal of International Economics 90 (1), 1–16.

Barber, J. R., Copper, M. L., 2012. Principal component analysis of yield curve movements. Journal of Economics and Finance 36 (3), 750–765.

Bekaert, G., Wei, M., Xing, Y., 2007. Uncovered interest rate parity and the term structure. Journal of International Money and Finance 26 (6), 1038–1069.

Bernoth, K., Erdogan, B., 2012. Sovereign bond yield spreads: A time-varying coefficient approach. Journal of International Money and Finance 31 (3), 639–656.

Bianchetti, M., 2010. Two curves, one price. Risk (August), 74–80.

Bikbov, R., Chernov, M., 2010. Yield Curve and Volatility: Lessons from Eurodollar Futures and Options. Journal of Financial Econometrics 9 (1), 66–105.

Blaskowitz, O., Herwartz, H., 2009. Adaptive forecasting of the EURIBOR swap term structure. Journal of Forecasting 28 (7), 575–594.

Boudoukh, J., Richardson, M., Whitelaw, R., 2005. The Information in Long-Maturity Forward Rates: Implications for Exchange Rates and the Forward Premium Anomaly. NBER Working Paper 11840.

Boudoukh, J., Richardson, M., Whitelaw, R. F., 2008. The Myth of Long-Horizon Predictability. Review of Financial Studies 21 (4), 1577–1605.

Bui, A. T., Fisher, L. A., 2016. The relative term structure and the Australian–US exchange rate. Studies in Economics and Finance (Forthcoming).

Bulkley, G., Harris, R. D., Nawosah, V., 2011. Revisiting the expectations hypothesis of the term structure of interest rates. Journal of Banking & Finance 35 (5), 1202–1212.

Chen, R.-R., Cheng, X., Wu, L., 2012. Dynamic Interactions Between Interest-Rate and Credit Risk: Theory and Evidence on the Credit Default Swap Term Structure. Review of Finance 17 (1), 403–441.

Chen, Y.-C., Rogoff, K., 2003. Commodity currencies. Journal of International Economics 60 (1), 133–160.

Chen, Y.-C., Tsang, K. P., 2013. What Does the Yield Curve Tell Us about Exchange Rate Predictability? Review of Economics and Statistics 95 (1), 185–205.

Chinn, M. D., Meredith, G., 2004. Monetary Policy and Long-Horizon Uncovered Interest Parity. IMF Staff Papers 51 (3), p 409–430.

Christiansen, C., Ranaldo, A., Söderlind, P., 2011. The Time-Varying Systematic Risk of Carry Trade Strategies. Journal of Financial and Quantitative Analysis 46 (04), 1107–1125.

Clarida, R. H., Sarno, L., Taylor, M. P., Valente, G., 2003. The out-of-sample success of term structure models as exchange rate predictors: a step beyond. Journal of International Economics 60 (1), 61–83.

Clarida, R. H., Waldmann, D., 2008. Is Bad News About Inflation Good News for the Exchange Rate? And, If So, Can That Tell Us Anything about the Conduct of Monetary Policy. NBER Chapters.

Cortes, F., 2006. Understanding the Term Structure of Swap Spreads. Bank of England Quarterly Bulletin 46 (1), 45–56.

Dai, Q., Singleton, K. J., 2000. Specification Analysis of Affine Term Structure Models. Journal of Finance 55 (5), 1943–1978.

Darvas, Z., 2009. Leveraged carry trade portfolios. Journal of Banking & Finance 33 (5), 944–957.

Dewachter, H., Lyrio, M., 2006. Macro Factors and the Term Structure of Interest Rates. Journal of Money, Credit and Banking 38 (1), 119–140.

Diebold, F. X., Rudebusch, G. D., Aruoba, S. B., 2006. The macroeconomy and the yield curve: a dynamic latent factor approach. Journal of Econometrics 131 (1-2), 309–338.

Dornbusch, R., 1976. Expectations and Exchange Rate Dynamics. Journal of Political Economy 84 (6), 1161–1176.

Driessen, J., Melenberg, B., Nijman, T., 2003. Common factors in international bond returns. Journal of International Money and Finance 22 (5), 629–656.

Duffee, G. R., 2011. Forecasting with the term structure: The role of no-arbitrage restrictions. Economics Working Paper Archive, Johns Hopkins University.

Dungey, M., Martin, V. L., Pagan, A. R., 2000. A multivariate latent factor decomposition of international bond yield spreads. Journal of Applied Econometrics 15 (6), 697–715.

Eichenbaum, M., Evans, C. L., 1995. Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates. Quarterly Journal of Economics 110 (4), 975–1009.

Eichler, S., Maltritz, D., 2013. The term structure of sovereign default risk in EMU member countries and its determinants. Journal of Banking & Finance 37 (6), 1810–1816.

Engel, C., 2013. Exchange Rates and Interest Parity. NBER Working Paper 19336.

Engel, C., West, K. D., 2005. Exchange Rates and Fundamentals. Journal of Political Economy 113 (3), 485–517.

Erdogan, O., Bennett, P., Ozyildirim, C., 2015. Recession Prediction Using Yield Curve and Stock Market Liquidity Deviation Measures. Review of Finance 19 (1), 407–422.

Favero, C. A., Niu, L., Sala, L., 2012. Term Structure Forecasting: No-Arbitrage Restrictions versus Large Information Set. Journal of Forecasting 31 (2), 124–156.

Fratzscher, M., 2009. What explains global exchange rate movements during the financial crisis? Journal of International Money and Finance 28 (8), 1390–1407.

Gourinchas, P.-O., Tornell, A., 2004. Exchange rate puzzles and distorted beliefs. Journal of International Economics 64 (2), 303–333.

Guidolin, M., Tam, Y. M., 2013. A yield spread perspective on the great financial crisis: Break-point test evidence. International Review of Financial Analysis 26, 18–39.

Hansen, L. P., Hodrick, R. J., 1980. Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis. Journal of Political Economy 88 (5), 829–853.

Harri, A., Brorsen, W. B., 2009. The Overlapping Data Problem. Quantitative and Qualitative Analysis in Social Sciences 3 (3), 78–115.

Haugh, D., Ollivaud, P., Turner, D., 2009. What Drives Sovereign Risk Premiums?: An Analysis of Recent Evidence from the Euro Area. OECD Economics Department Working Papers 718.

Hilscher, J., Nosbusch, Y., 2010. Determinants of Sovereign Risk: Macroeconomic Fundamentals and the Pricing of Sovereign Debt. Review of Finance 14 (2), 235–262.

Jolliffe, I. T., 2002. Principal component analysis, 2nd Edition. Springer series in statistics. Springer, New York.

Juneja, J., 2012. Common factors, principal components analysis, and the term structure of interest rates. International Review of Financial Analysis 24, 48–56.

Koopman, S. J., van der Wel, M., 2013. Forecasting the US term structure of interest rates using a macroeconomic smooth dynamic factor model. International Journal of Forecasting 29 (4), 676–694.

Krippner, L., 2006. A Yield Curve Perspective on Uncovered Interest Parity. Working Paper.

Lardic, S., Priaulet, P., Priaulet, S., 2003. PCA of the yield curve dynamics: questions of methodologies. Journal of Bond Trading and Management 1 (4), 327–349.

Leite, A. L., Filho, R. B. P. G., Vicente, J., 2010. Forecasting the yield curve: A statistical model with market survey data. International Review of Financial Analysis 19 (2), 108–112.

Litterman, R., Scheinkman, J., 1991. Common Factors Affecting Bond Returns. Journal of Fixed Income 1 (1), 54–61.

Longstaff, F. A., Pan, J., Pedersen, L. H., Singleton, K. J., 2007. How Sovereign is Sovereign Credit Risk? NBER Working Paper 13658.

Lustig, H., Roussanov, N., Verdelhan, A., 2011. Common Risk Factors in Currency Markets. Review of Financial Studies 24 (11), 3731–3777.

Maltritz, D., 2012. Determinants of sovereign yield spreads in the Eurozone: A Bayesian approach. Journal of International Money and Finance 31 (3), 657–672.

Mark, N. C., 1995. Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability. American Economic Review 85 (1), 201–218.

Martell, R., 2007. Understanding Common Factors in Domestic and International Bond Spreads. Review of Finance 12 (2), 365–389.

McCallum, B. T., 1994. A reconsideration of the uncovered interest parity relationship. Journal of Monetary Economics 33 (1), 105–132.

Meese, R. A., Rogoff, K., 1983. The out-of-sample failure of empirical exchange rate models: sampling error or misspecification? In: Frenkel, J. A. (Ed.), Exchange rates and international macroeconomics. University of Chicago Press, Chicago, pp. 67–112.

Moench, E., 2012. Term structure surprises: the predictive content of curvature, level, and slope. Journal of Applied Econometrics 27 (4), 574–602.

Molodtsova, T., Papell, D. H., 2009. Out-of-sample exchange rate predictability with Taylor rule fundamentals. Journal of International Economics 77 (2), 167–180.

Monfort, A., Renne, J.-P., 2014. Decomposing Euro-Area Sovereign Spreads: Credit and Liquidity Risks. Review of Finance 18 (6), 2103–2151.

Newey, W. K., West, K. D., 1987. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica 55 (3), 703–708.

Oliveira, L., Curto, J. D., Nunes, J. P., 2012. The determinants of sovereign credit spread changes in the Euro-zone. Journal of International Financial Markets, Institutions and Money 22 (2), 278–304.

Pérignon, C., Smith, D. R., 2007. Yield-factor volatility models. Journal of Banking & Finance 31 (10), 3125–3144.

Pooter, M. D., Ravazzolo, F., van Dijk, D., 2010. Term structure forecasting using macro factors and forecast combination. Board of Governors of the Federal Reserve System, International Finance Discussion Papers 993.

Ranaldo, A., Soderlind, P., 2010. Safe Haven Currencies. Review of Finance 14 (3), 385–407.

Rocha, K., Garcia, F. A. A., 2005. The Term Structure of Sovereign Spreads in Emerging Markets. Journal of Fixed Income 14 (4), 45–56.

Rossi, B., 2013. Exchange Rate Predictability. Journal of Economic Literature 51 (4), 1063–1119.

Rudebusch, G. D., Wu, T., 2007. Accounting for a Shift in Term Structure Behavior with No-Arbitrage and Macro-Finance Models. Journal of Money, Credit and Banking 39 (2-3), 395–422.

Rudebusch, G. D., Wu, T., 2008. A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy. Economic Journal 118 (530), 906–926.

Sarno, L., 2005. Viewpoint: Towards a solution to the puzzles in exchange rate economics: where do we stand? Canadian Journal of Economics 38 (3), 673–708.

Sarno, L., Schneider, P., Wagner, C., 2012. Properties of foreign exchange risk premiums. Journal of Financial Economics 105 (2), 279–310.

Schwert, G. W., 2002. Tests for Unit Roots. Journal of Business & Economic Statistics 20 (1), 5–17.

Sueppel, R., 2005. Forecasting Sovereign Credit Spreads. Journal of Fixed Income 15 (1), 54–67.

5 Exchange Rates and Unobservable Fundamentals: A New Approach to Out-of-Sample Forecasting

Dennis Wellmann (contribution: 80%), Stefan Trück (contribution: 20%)

This paper was presented at:

- 28th 'Australasian Finance & Banking Conference' (AFBC), Sydney, December 2015
- 'Research Seminar' at German Institute for Economic Research (DIW), Berlin, May 2016

Abstract

Traditional exchange rate models are based on differences in macroeconomic fundamentals. However, despite being well grounded in economic theory they have a rather poor out-of-sample forecasting record. This empirical failure may be a result of the overly restrictive choice of macroeconomic fundamentals. We suggest using the sovereign yield spread level and slope as proxies of the market's expectations for current and future fundamentals and find promising results when we investigate the out-of-sample forecasting accuracy of these variables. Using the yield spread level and slope as a set of unobservable fundamentals, our approach outperforms traditional exchange rate models for most considered currencies and horizons. It is also superior to a random walk in terms of direction of change forecasts and profitability.

5.1 Introduction

Traditional exchange rate models, e.g. the monetary model or the purchasing power parity approach, are based on differences in macroeconomic fundamentals such as monetary supply, inflation and output. However, these standard fundamental-based models have a rather poor out-of-sample forecasting performance. Starting with the seminal contribution of Meese and Rogoff (1983), a vast body of empirical research finds that models based on macroeconomic fundamentals cannot outperform a naive random walk model in terms of the root mean squared error (RMSE), see e.g., Cheung et al. (2005); Molodtsova and Papell (2009); Rossi (2013).

The literature has put forth several reasons for this dismal record. Existing structural models may, for instance, be overly restrictive in their choice of macroeconomic fundamentals (Engel and West, 2005; Balke et al., 2013). The empirical failure may also be a result of using inappropriate proxies for the market expectations of future fundamentals which become highly important when the exchange rate is understood as an asset price (Mark, 1995; Bacchetta and van Wincoop, 2013).

Instead of applying traditional observable fundamentals, we therefore suggest using the level and slope of sovereign yield spread curves between economies[1] as market-based proxies for current and future macroeconomic fundamentals to forecast exchange rates. We find that this approach delivers promising forecasting results based on statistical and economic evaluation measures when compared against the random walk and commonly used fundamental exchange rate models.

The motivations for our innovative approach are twofold. First, interpreted as an asset price, exchange rates are now commonly considered to equal the sum of discounted future macroeconomic fundamentals. The yield spread level and slope are forward-looking financial indicators which summarize

[1] Sovereign yield spreads are the difference between two sovereign bond yields of equal maturity. The sovereign yield spread level $L^{\Delta sy}$ is defined as an average of short, medium and long term spreads and the spread slope $S^{\Delta sy}$ denotes the difference between long and short term yield spreads.

the long-term macroeconomic information contained in the term structure of yield spreads – the difference between the yield curves of two different economies. Thus, these variables naturally contain unobservable information about the same expected macroeconomic differentials that drive exchange rates (Chen and Tsang, 2013; Bui and Fisher, 2016). Chapter 4 has further confirmed that the yield spread level and slope factors have predictive power for exchange rate changes and excess returns in an in-sample analysis. Second, because bond yields and foreign exchanges are susceptible to the same macroeconomic risk, the expected risk premiums that investors require for holding these assets might closely relate to each other.

Our approach has several further advantages compared to traditional fundamental models. The yield spread approach is market based, as the expectations about future economic fundamentals reflect the view of a large number of market participants in highly liquid sovereign bond markets. Yield data is also readily and easily available on a daily basis as opposed to monthly and quarterly macroeconomic data which is often published with a considerable time lag and revised afterwards. Finally, our parsimonious models are straightforward to implement and therefore an appealing approach for investment practitioners.

To assess the out-of-sample forecasting accuracy of our approach we conduct an extensive forecasting exercise for forecasting horizons of one, three and six months against the random walk and several traditional fundamental exchange rate models based on interest rate, price, monetary and Taylor rule fundamentals. We use major currencies of advanced countries with free floating exchange rates and highly liquid bond markets with little to no credit risk (Australian Dollar, the Canadian Dollar, the Swiss Franc, the Japanese Yen and the British Pound) against the US Dollar.[2]

We assess the forecasting accuracy of the investigated models based on several different forecasting evaluation methods. Previous research has shown that the sole focus on the traditional RMSE metric may not be entirely appropriate for exchange rates (Cheung et al., 2005; Moosa and Burns, 2014). In addition to the RMSE, we therefore apply further evaluation measures

[2]We note that our findings also hold for other cross exchange rates.

to provide a multifaceted assessment of the forecasting performance of our approach and the benchmark models. In particular, we also apply a measure of direction accuracy and assess the forecasting uncertainty based on density forecasts.

Since statistical evidence of superior exchange rate forecasting accuracy does not necessarily guarantee an investor to make a profit when exploiting this predictability, the ultimate test of forecasting power is the economic viability (Abhyankar et al., 2005; Corte et al., 2009; Moosa and Burns, 2014). We thus also implement a period-by-period trading strategy to assess the profitability of the forecasts produced by the implemented models.

Considering all of the applied statistical and economic evaluation metrics, we find promising results for our yield spread approach. Using the spread level or slope is generally superior in terms of the RMSE and direction accuracy, when being compared to traditional fundamental models. The approach typically also provides better results in terms of its density forecasts. While neither our approach nor the benchmark models are able to consistently beat the random walk in terms of the RMSE – which should hardly be surprising given the findings in previous literature – the suggested yield spread approach clearly outperforms the random walk in forecasting the direction of exchange rate changes. We also find that our approach consistently yields higher (lower) risk-adjusted profits (losses) than the considered fundamental benchmark models and also outperforms the random walk in terms of profitability for several currencies.

As the global financial crisis (GFC) from 2007-2009 lies within our forecasting period, we also investigate the impact of this period on exchange rate forecasting accuracy and trading profitability. While we generally find a strong impact on foreign exchange markets, the impact on RMSE, direction accuracy and profitability seems rather limited. However, as could be expected we do find that the GFC significantly increases the uncertainty of all model's exchange rate forecasts.

Taken together, these results drawn from different statistical and economical evaluation measures provide an encouraging view with regards to the forecasting ability of our approach. The promising out-of-sample results also confirm previous studies which have investigated the predictive power of financial variables for exchange rates. Guo and Savickas (2008), for example, show that financial variables that have been commonly used as predictors of stock returns, or bond returns also have the ability to forecast exchange rates. Evans and Lyons (2007) and Rime et al. (2010) show that order flow forecasts exchange rates because it contains information about future fundamentals. Overall, our results support the view that financial variables may be an intuitive and promising forecasting approach when exchange rates are understood as an asset price and equal the sum of expected future fundamentals.

It is important to note that our results do not imply that the macroeconomic fundamentals applied in traditional models cannot forecast exchange rates. Quite the opposite, our results are consistent with the view that the principal drivers of exchange rates are standard macro fundamentals. The difference between our approach and traditional fundamental models is that we apply the spread level and spread slope as proxies for *unobservable* macro fundamentals instead of using selected, often restricted *observable* macroeconomic variables directly in the forecasting equation.

With this study, we thus contribute to the literature of exchange rate forecasting in several dimensions. First, we present an innovative, parsimonious, market driven approach to exchange rate forecasting based on readily and easily available data. This makes it a promising proposition in particular for market practitioners. Second, we provide further evidence that financial variables are useful indicators to be considered in exchange rate forecasting and thus hope that the results inspire a renewed interest in exchange rate forecasting models based on financial variables. Third, we thoroughly investigate the impact of the GFC on foreign exchange markets and the forecasting accuracy of exchange rate models. Nevertheless, further research is required to fully understand the impact of the GFC on foreign exchange markets and models. Finally, we confirm that the random walk is beatable by models using observable and unobservable models if appropriate evaluation measures and trading profitability are applied. The difference in conclusions for the implemented evaluation metrics also further highlights the importance of applying different measures to provide a conclusive assessment of a model's forecasting ability. As Rossi (2013) has put it, "the choice of the evaluation method matters, and matters a lot."

The remainder of this chapter is structured as follows. The next section provides an overview of traditional exchange rate models and their empirical forecasting performance. Section 5.3 introduces our forecasting approach based on the empirical sovereign yield spread level and slope. In Sections 5.4 and 5.5 we conduct the out of sample forecasting exercise against several popular benchmark models and discuss the results. In Section 5.6 we implement the trading strategy to investigate the profitability of our forecasts. Section 5.7 concludes.

5.2 Exchange Rate Determination and Forecasting

5.2.1 Traditional Fundamental Exchange Rate Models

Economic theory states that the exchange rate is determined by differences between macroeconomic fundamentals such as money supply, inflation, output and interest rates. This relationship between the exchange rate and its fundamentals can be described by different models based on varying economic variables and econometric techniques such as error correction models (ECM), time-varying parameter (TVP) models and – still most commonly applied – linear models, see Rossi (2013) for an excellent recent overview.

For expositional[3] purposes, let the basic model be linear with a constant term. Assume that s_t denotes the log of the nominal exchange rate (home currency price per unit of foreign currency) and f_t the (potentially multivariate) fundamental(s) of the exchange rate. The general relationship can then be expressed as:

$$s_t = \alpha + \beta f_t. \tag{5.1}$$

[3]Theoretically, also non-linear models could be used.

This framework gives way for the most commonly used models tying floating exchange rates to differences in interest rates and macroeconomic fundamentals:

Interest Rate Differentials

Traditionally, the relation between differences in interest rates and exchange rates is expressed in the uncovered interest rate parity (UIRP) condition. The UIRP relates exchange rate changes to interest rate differentials between two economies over the same horizon:

$$\Delta s_{t+h} = \alpha + \beta(i_t^{\tau} - i_t^{\tau,*}), \tag{5.2}$$

where Δs_{t+h} is the h-horizon exchange rate change and i_t^{τ} and $i_t^{\tau,*}$ are the domestic and foreign interest rates of maturity τ where $\tau = h$. If uncovered interest rate parity holds, α and β should be 0 and 1 respectively.

Price Level Fundamentals

According to Purchasing Power Parity (PPP), the real price of comparable commodity baskets in two countries should be the same. Thus, the price level in the home country should equal the price level of the foreign country converted to the currency of the foreign country. It follows that a unit of currency in the home country will have the same purchasing power in the foreign country. Accordingly, PPP implies that

$$s_t = \alpha + \beta(p_t - p_t^*), \tag{5.3}$$

where p_t and p_t^* denote the logarithm of the price index in the home and foreign country, respectively.[4]

Monetary and Output Fundamentals

The frequently used monetary model builds upon PPP and UIP but assumes additional restrictions. It models exchange rate behavior in terms of relative demand for and supply of money in the two economies. To start with, real

[4] PPP does imply $\alpha = 0$ and $\beta = 1$. However in empirical forecasting this relationship is usually estimated.

money demand is viewed as a function of income and interest rates:

$$m_t - p_t = \eta i_t + \phi y_t, \tag{5.4}$$

where m_t is the log of nominal money demand, i_t denotes the interest rate, y_t is the logarithm of real output and η and ϕ are coefficients. Assuming that a similar equation holds for the foreign country with symmetric (equal) coefficients and taking the difference between the two gives the relative money demand equation:

$$m_t - m_t^* - (p_t - p_t^*) = \eta(i_t - i_t^*) + \phi(y_t - y_t^*). \tag{5.5}$$

The 'flexible price version' of the monetary model (valid if prices and exchange rates are completely flexible) assumes that PPP holds at every point in time. Substituting the PPP relation into the relative money demand equation we get

$$s_t = \eta(i_t - i_t^*) - \phi(y_t - y_t^*) + (m_t - m_t^*). \tag{5.6}$$

In the presence of sticky price adjustment, either the relative price level or inflation differentials are included to obtain the 'sticky price version' of the monetary model:

$$s_t = \eta(i_t - i_t^*) - \phi(y_t - y_t^*) + (m_t - m_t^*) + (p_t - p_t^*). \tag{5.7}$$

In this case it is assumed that PPP holds in the long run but does not hold in the short run.

Taylor Rule Fundamentals

Recently, studies have proposed fundamentals based on a Taylor rule for monetary policy (Engel and West, 2005; Molodtsova and Papell, 2009). At the core of models using Taylor rule fundamentals is the idea that if two economies set interest rates based on a Taylor rule, their bilateral exchange rate will reflect their relative interest rates through UIRP.

Consequently, this approach assumes that both central banks adjust the target rate i_t^T according to a Taylor rule in response to changes in the output gap and deviation from target inflation:

$$i_t^T = \pi_t + \phi(\pi_t - \pi^T) + \gamma y_t^{gap} + r, \tag{5.8}$$

where π_t is the inflation rate, π^T is the target level of inflation, y_t^{gap} is the output gap[5] and r is the equilibrium level of the real interest rate.
Assuming that a similar condition applies to the foreign country with equal coefficients ϕ and γ (symmetric Taylor rule with homogeneous coefficients) and further assuming that UIRP and PPP hold leads to:[6]

$$\Delta s_{t+h} = (1 + \phi)(\pi_t - \pi_t^*) + \gamma(y_t^{gap} - y_t^{gap*}). \tag{5.9}$$

Hence, under this basic Taylor rule approach, the fundamentals that determine the exchange rate are the country differentials in inflation and output gap.[7]

5.2.2 Empirical Evidence

The empirical validation for these theoretical frameworks remains rather elusive. In a large body of empirical out-of-sample forecasting studies the random walk model has proven almost unbeatable by models with traditional macroeconomic predictors. Meese and Rogoff (1983) first established this result in their seminal paper. They evaluated the out-of-sample fit of several exchange rate models in the short run and concluded that a random walk predicts exchange rates better than macroeconomic models in terms of the RMSE.
Many studies have subsequently claimed to find success for various versions

[5] The output gap is the difference between actual output and potential output $y_t^{gap} = y_t - \bar{y}_t$ at time t, where y_t is the logarithm of real output and $\bar{y}_t$ is the logarithm of potential output measured e.g. by a linear time trend.

[6] See Giacomini and Rossi (2010) for a more detailed derivation.

[7] Under different assumptions, e.g heterogenous coefficients or central banks also considering the real exchange rate, other fundamentals such as the country differentials in interest rates and price levels may be included as well. Molodtsova and Papell (2009) provide a comprehensive overview of different approaches applying Taylor rule fundamentals.

of fundamentals-based models. Kilian and Taylor (2003), for example, find that exchange rates can be predicted from economic models at horizons of two to three years after taking into account the possibility of nonlinear exchange rate dynamics. Bjørnland and Hungnes (2006) combine the purchasing power parity condition with the interest rate differential in the long run and show that their approach outperforms a random walk in an out-of-sample forecasting exercise for several horizons. Molodtsova and Papell (2009) investigate the predictability of models that incorporate Taylor rule fundamentals and provide evidence of short-run exchange rate predictability.

However, the success of these models has not proven to be universally reliable and robust. Models that work well in one period or for one currency do not necessarily work well in another period or for other currencies. The study by Cheung et al. (2005) examines the out-of-sample performance of several popular fundamental based models and finds that none of the models consistently outperforms the random walk. More recently, Rossi (2013) also concludes in a recent comprehensive survey that forecasting success largely depends on the choice of predictor, forecast horizon, sample period, model, and forecast evaluation method. Thus, even after more than 30 years, the Meese and Rogoff (1983) results have not yet been convincingly overturned.

Several reasons have been put forward for the empirical out-of-sample forecasting failure of traditional exchange rate models. The poor forecasting performance may for example reflect, at least in part, econometric issues. In their original paper, Meese and Rogoff (1983) attribute the failure to underlying econometrics such as a simultaneous equations bias, sampling errors, stochastic movements in the true underlying parameters, misspecification and nonlinearities. Moosa (2013) also demonstrates that failure to outperform the random walk should be the rule rather than the exception due to the characteristics of the underlying processes.

The empirical failure may further be a result of using inappropriate proxies for the market expectations of future fundamentals rather than the failure of the models themselves. It has long been suggested, see, e.g., Frenkel (1983) for an early survey, that exchange rates should be viewed as an asset price determined in financial markets, similar to stock, bond and commodity markets, in which current prices reflect the market's expectations about present and future. Following Mark (1995) and Engel and West (2005), the exchange rate is now commonly modeled as an asset price, where the nominal exchange rate is determined as the present value of the discounted sum of current and expected fundamentals:

$$s_t = (1 - \omega) f_t + \omega E_t(s_{t+1}), \tag{5.10}$$

where ω is a discount factor less than one. Iterating this equation forward then leads to

$$s_t = (1 - \omega) f_t + (1 - \omega) \sum_{j=1}^{\infty} \omega^j E_t(f_{t+j}). \tag{5.11}$$

This approach implies that the exchange rate is determined by the weighted average of fundamentals such as economic growth, inflation or money supply which are determined by the chosen model. It also follows that within the present value framework exchange rates rely more on expectations about the future than on current fundamentals. Properly measuring expectations thus becomes especially important in empirical studies (Bacchetta and van Wincoop, 2013). Standard empirical approaches, however, often reduce the sum of expected future fundamentals to equal current fundamentals (Chen and Gwati, 2014).

Existing structural models grounded in economic theory may also be overly restrictive in their choice of observable macroeconomic fundamentals. Engel and West (2005), for example, argue that exchange rates are not only affected by observable fundamentals. Balke et al. (2013) also show that it is difficult to obtain sharp inferences about the relative contribution of fundamentals using only data on observed fundamentals.

Finally, it has been suggested that the use of the RMSE and similar statistical criteria solely based on minimizing the loss function may not be entirely appropriate to measure exchange rate forecasting accuracy. A correct prediction of the direction of change can often be more important than the magnitude of the error (Cheung et al., 2005), for example when it comes to hedging decisions. Researchers have also suggested that the ultimate test

of forecasting power is the ability to make profits based on the predicted exchange rate changes (Corte et al., 2009). Moosa and Burns (2014), for example, demonstrate that the conventional monetary model can outperform the random walk in out-of-sample forecasting if forecasting power is measured by direction accuracy and profitability. Several studies also argue that it is important to asses the uncertainty of exchange rate forecasts (Diebold et al., 1999; Rapach and Wohar, 2006). Wang and Wu (2012), for example, find that Taylor rule models can outperform the random walk, especially at long horizons, based on interval forecasting criteria.
We take these arguments into account and judge the forecasting accuracy in this analysis based on several different statistical measures. We further implement a trading strategy to asses the profitability of the derived forecasts.

5.2.3 Alternative Approaches

The perceived failure of traditional fundamentals-based exchange rate models in empirical out-of-sample forecasting has motivated numerous alternative approaches to model and forecast the exchange rate. Engel et al. (2007), for example, include expectations of fundamentals drawn from survey data and demonstrate that the predictive power of the models can be greatly increased by using panel techniques. Engel et al. (2012) construct factors from a cross-section of exchange rates and use the idiosyncratic deviations from the factors to beat the random walk benchmark.
One recent stream of literature also investigates financial variables as predictors for exchange rates. Evans and Lyons (2007) and Rime et al. (2010) for example show that order flow forecasts exchange rates because it contains information about future fundamentals. Christiansen (2011) use a smooth transition model to show that typical FX carry trade strategies have a high exposure to the stock market. Molodtsova and Papell (2012) incorporate indicators of financial stress to improve the forecasting performance of models based on Taylor rule fundamentals. Ferraro et al. (2015) further document the relationship between commodity prices and exchange rates.
We follow a similar path by applying the level and slope of cross-country

yield spread curves as forward-looking financial variables which reflect expectations of future and unobservable macroeconomic fundamentals. This approach is further described in the subsequent section.

5.3 A Market Driven Approach using the Sovereign Spread Level and Slope

5.3.1 Financial Variables and Exchange Rates

One of the main findings of the previous Section 5.2 suggests that the failure of empirical exchange rate forecasting models may be due to using inappropriate proxies for market expectations of future and non-observable fundamentals. The fact that plausible models now consider the exchange rate as an asset price means that short-run movements in exchange rates are primarily determined by changes in expectations and that unobservable fundamentals play a significant part in this process. However, future expectations and unobservable fundamentals both are difficult to capture with traditional empirical models which commonly reduce the sum of expected future fundamentals to equal current fundamentals (Chen and Gwati, 2014) and are too stylized to be successfully applied to forecast exchange rates (Rossi, 2013).

In this context, financial variables may be an intuitive, promising resolution. Through their forward-looking character many financial variables incorporate market expectations of future economic conditions (Stock and Watson, 2003). Share and bond prices, for instance, reflect discounted future cash flows based on expectations about the firm level and macroeconomic environment. When exchange rates are understood as an asset price and equal the sum of expected future fundamentals, financial variables may thus naturally have predictive power for exchange rates.

Furthermore, financial variables such as stock or bond returns might also be related to exchange rates because the expected risk premiums that investors require for holding stocks, bonds, and foreign currencies might closely relate to each other. Guo and Savickas (2008) investigate whether financial variables that have been commonly used as predictors of stock or bond returns also forecast exchange rates. They document in particular a strong relation between idiosyncratic stock volatility and exchange rates.

5.3.2 Yield Curves and Macroeconomic Fundamentals

While the exchange rate literature has so far focused more on the relation between stock prices and exchange rates (Evans and Lyons, 2007; Rime et al., 2010; Cenedese et al., 2015), bond yields are another obvious choice. Yield curves are well known to summarize expectations about future paths of short interest rates and thus contain information about expected future economic conditions such as output, inflation, recessions and monetary policy (Stock and Watson, 2003; Ang et al., 2006; Rudebusch and Wu, 2008; Favero et al., 2012; Erdogan et al., 2015).

Findings in previous studies suggest that this macroeconomic information entailed in the yield curve is summarized in the level, slope and curvature of the term structure. Estrella and Mishkin (1998), for example, argue that the yield curve slope is a serious candidate as predictor of output growth and recessions. Diebold et al. (2006) find that an increase in the US yield curve level factor raises capacity utilization, the US fund rate and inflation. Dewachter and Lyrio (2006) suggest that the level factor reflects long run inflation expectations. Rudebusch and Wu (2007) also contend that the level factor incorporates long-term inflation expectations and the slope factor captures the business cycle. Moench (2012) finds that a rising yield curve slope factor is associated with a future decline of output while surprise surges of the yield curve level are followed by a strong and persistent increase of inflation rates.

The shape and movements of the yield curve have therefore long been used to provide readings of market expectations about the same fundamentals whose differentials are commonly used to model and forecast exchange rates (see Section 5.2).

5.3.3 Macroeconomic Fundamentals and Sovereign Spread Factors

We thus argue that the term structure of sovereign yield spreads – the difference between two economies respective yield curves – can be considered as a natural candidate for exchange rate forecasting.

Sovereign yield spreads are the difference between two government bond yieldsof equal maturity. The τ-maturity sovereign yield spread Δsy_t^m is thus calculated as:

$$\Delta sy_t^\tau = sy_t^\tau - sy_t^{\tau,*}, \tag{5.12}$$

where sy_t^τ and $sy_t^{\tau,*}$ are τ-maturity home and foreign country sovereign yields respectively.

As sovereign spreads can be calculated for any maturity, they exhibit a term structure – or spread curve – of their own. We conjecture that this spread curve naturally contains valuable information about the market expectation of the differences in macroeconomic conditions that determine exchange rates. The findings in the yield curve literature described above further suggest that the information about macroeconomic differentials entailed in the sovereign spread curves will be reflected in the spread level, spread slope and spread curvature.

Recent research comprising the term structure of sovereign yield spreads confirms this suspected link between spread curve factors and exchange rates. In a cross-country setting based on portfolio strategies, Ang and Chen (2010) find an economically and statistically significant ability of the yield level and slope factors of the term structure to predict exchange rate profitability. Chen and Tsang (2013) find in-sample that cross-country Nelson-Siegel factors which are related to the sovereign spread level, slope and curvature can predict future exchange rate changes and excess currency returns. Bui and Fisher (2016) support their findings for the relative yield curves of the US and Australia. The results of Chapter 4 have also confirmed that in particular the spread level and slope are capable of predicting exchange rate changes and excess returns in-sample.

5.3.4 Using the Sovereign Spread Level and Slope to Forecast Exchange Rates

Encouraged by these promising results we propose to exploit the fundamental information contained in the sovereign yield spread curve between economies to forecast exchange rates out-of-sample. To make this information applicable within a parsimonious forecasting model, we build on the two main insights of the previous Chapter 4. First, latent factors derived from the term structure of sovereign spreads have predictive power for exchange rates. Second, these factors are highly associated with the empirical sovereign spread level and slope. We thus suggest using the empirical sovereign yield spread curve level $L_t^{\Delta sy}$ and slope $S_t^{\Delta sy}$ from Chapter 4 as a set of financial proxies which summarize the information of the spread curve and reflect the market expectations of future and unobservable fundamentals.[8]

Following the common approach in the yield curve literature (Diebold et al., 2006; Afonso and Martins, 2012), the sovereign spread level $L_t^{\Delta sy}$ is defined as an average of short, medium and long term spreads:

$$L_t^{\Delta sy} = \frac{\Delta sy_t^{short} + \Delta sy_t^{medium} + \Delta sy_t^{long}}{3}, \tag{5.13}$$

and the spread slope $S_t^{\Delta sy}$ denotes the difference between long and short term spreads:

$$S_t^{\Delta sy} = \Delta sy_t^{long} - \Delta sy_t^{short}. \tag{5.14}$$

Note, that we do not use the spread curvature which is commonly identified as a third factor in yield curve literature (Diebold and Li, 2006; Moench, 2012). We opt not to include it in our approach because the in-sample results

[8] We decide to use the empirical rather than the factors estimated in Chapter 4 because applying the empirical factors in an out-of-sample forecasting framework is intuitive and less computationally extensive. It is thus straightforward to apply in practice. Robustness tests comparing the forecasting accuracy between empirical and estimated factors (results available upon request) also indicate that there is no considerable difference in the overall forecasting accuracy.

in Chapter 4 indicate that the additional predictive power of the curvature factor for exchange rates is rather limited. Robustness tests with the spread curvature (results available upon request) also confirm that the overall forecasting accuracy is not very promising.

Our innovative approach has several advantages compared to traditional fundamental models. First, the sovereign yield spread level and slope are driven by the sentiment of highly liquid financial markets. Sovereign bond markets are amongst the largest and most liquid financial markets in the world and therefore summarize the expectations of a large number of market participants. This also means that changes in expectations about future economic fluctuations are quickly incorporated into the variables. Second, yield data is readily and easily available on a daily basis as opposed to monthly or quarterly macroeconomic data which is often published with a significant delay and revised in hindsight. Finally, our approach is parsimonious, which makes it straightforward to apply in practice.

Naturally, our approach is somewhat related to the UIRP inspired fundamental model based on interest rate differentials described in Section 5.2.1. However, while the interest rate differential model only uses yield spreads of one specific maturity, the spread level and slope naturally exploit forward-looking information contained in the entire spread curve and thus seem to better reflect the idea that exchange rates are now considered as asset prices. Purists may further criticize a lack of a clear theoretical foundation. The results from the yield curve literature described above suggest that the level factor can be seen as a long-run inflation expectation factor while the slope factor reflects business cycle and output growth dynamics, see also Chen and Tsang (2013). Still, the factors cannot clearly be tied to specific macroeconomic variables as they may reflect a range of latent, unobservable fundamentals. However, in empirical forecasting, this should rather be considered as an advantage as it allows for parsimonious yet flexible modeling compared to often overly restrictive structural models and is therefore less prone to the omitted variable bias.

5.4 Out-of-sample Forecasting Framework

5.4.1 Forecating Specifications

To investigate the forecasting accuracy of our approach, we conduct an extensive out-of-sample forecasting exercise using the major currencies of advanced countries with free floating exchange rates and highly liquid bond markets with little to no credit risk. We thus include the Australian Dollar, the Canadian Dollar, the Swiss Franc, the Japanese Yen and the British Pound (all measured against the US Dollar following the convention in the exchange rate literature, see, e.g., Molodtsova and Papell (2009); Giacomini and Rossi (2010); Rossi (2013)).[9] We do not include the Euro because there is no Euro yield curve which reflects the macroeconomic prospects of the entire Euro area. Using, e.g. German yields as a proxy does not appear to be reasonable for our approach, especially as our sample includes the recent Euro crisis.

Our analysis covers the time period from 1995:01 to 2014:12. To evaluate the out-of-sample forecasting ability, the sample of size $T = 240$ monthly observations is split into an in-sample period, consisting of observations from $t = 1$ to R, and an out-of-sample portion of size $P = T - R$. We adopt the convention, see, e.g., Wang and Wu (2015), in the empirical exchange rate forecasting literature of implementing 'rolling windows'[10] and use a rolling window of size $R = 60$. This means that the models are estimated over an initial in-sample window from 1995:01-1999:12 to produce h-months ahead forecasts and then the in-sample window is moved up or 'rolled' forward one observation before the procedure is repeated. We thus produce h-months ahead forecasts for the period $R+h-1, ..., T$ for the forecast horizons $h = 1$, $h = 3$ and $h = 6$ months.[11]

---

[9] We note, that our findings also hold for other cross exchange rates.

[10] While the rolling regressions do not incorporate the possible efficiency gains of 'recursive windows' as the sample moves forward through time, the procedure has the potential benefit of alleviating parameter instability effects over time, which is a commonly conceived phenomenon in exchange rate forecasting.

[11] Note that, different to the in-sample analysis in Chapter 4, we focus on forecasting horizons up to six months which are commonly applied in exchange rate out-of-sample

5.4.2 Forecasting Models

Empirically, the most common approach to evaluating exchange rate models out of sample (following Mark (1995)) is to represent a change in the log of the nominal exchange rate as a function of its deviation from its fundamental value (see also Molodtsova and Papell (2009); Giacomini and Rossi (2010)). Thus, the h-period-ahead change in the log exchange rate can be denoted as:

$$\widehat{s_{t+h}} - s_t = \alpha + \beta z_t + \epsilon_{t+h}, \tag{5.15}$$

where $\widehat{s_{t+h}}$ is the h-period forecast of the log exchange rate s and $z_t = f_t - s_t$ with f_t being the long-run equilibrium level of the nominal exchange rate determined by its fundamentals. The choice of fundamentals f_t is then determined by the respective model.

Yield Spread Model

As described in the previous section, our approach considers the empirical yield spread level $L_t^{\Delta sy}$ and spread slope $S_t^{\Delta sy}$ as market driven indicators reflecting unobservable fundamentals. Treating these variables as unobservable fundamentals, we can state the long-run equilibrium level of the nominal exchange rate as

$$f_t^{SPRDLEV} = (L_t^i - L_t^{i*}) + s_t = L_t^{\Delta sy} + s_t \tag{5.16}$$

for the yield spread level factor, and as

$$f_t^{SPRDSL} = (S_t^i - S_t^{i*}) + s_t = S_t^{\Delta sy} + s_t, \tag{5.17}$$

for the slope factor, where $(L_t^i - L_t^{i*})$ and $(S_t^i - S_t^{i*})$ are the differences in yield curve level and slope between home and foreign country. Following the convention in the yield curve literature (see, e.g., Diebold et al. (2006); Afonso and Martins (2012)), we calculate the empirical spread level $L_t^{\Delta sy}$ as

forecasting, see e.g. Moosa and Burns (2014).

the average of the 3-months, the 36-months and the 120-months yield spread:

$$L_t^{\Delta sy} = \frac{\Delta sy_t^{\tau=3} + \Delta sy_t^{\tau=36} + \Delta sy_t^{\tau=120}}{3}, \tag{5.18}$$

and the empirical spread slope $S_t^{\Delta sy}$ as the difference between the 120-months and 3-months yield spread:

$$S_t^{\Delta sy} = \Delta sy_t^{\tau=120} - \Delta sy_t^{\tau=3}. \tag{5.19}$$

We denote these models as $L^{\Delta sy}$ and $S^{\Delta sy}$ or **SPRDLEV** and **SPRDSL** respectively.

Macroeconomic Benchmark Models

We are particularly interested in how the yield spread level and slope models performs against traditional exchange rate models and thus include one prominent version for any of the major fundamental models described in Section 5.2:[12]

Interest Rate Differentials: To apply differences in interest rates, the UIRP relation is often directly used in forecast equations. However, empirical evidence indicates that, while exchange rate movements may be consistent with UIRP in the long-run, it clearly does not hold in the short-run, see, e.g., Sarno (2005); Engel (2013). Following Clark and West (2006), we thus implement a more flexible specification with:

$$f_t^{IRD} = (i_t^h - i_t^{h*}) + s_t, \tag{5.20}$$

and do not restrict $\alpha = 0$ and $\beta = 1$ in the forecasting equation (5.15). Note, that this approach yields the standard h-period ahead forecasting equation

$$\Delta s_{t+h} = \alpha + \beta(i_t^{\tau} - i_t^{\tau,*}) + \epsilon_{t+h} \tag{5.21}$$

[12] With the wide variety of these models being applied in the literature one necessarily has to be selective with respect to model choice in order to keep the results manageable. Molodtsova and Papell (2009) for example test 48 variations of the model based on Taylor rule fundamentals.

for maturity $\tau = h$. We denote the interest rate differential model as **IRD**.
Price Fundamentals: If market participants believe that the future exchange rate is formed in line with PPP, the fundamental f_t is specified as:

$$f_t^{PPP} = (p_t - p_t^*). \tag{5.22}$$

This model is denoted as **PPP**.
Monetary Model: We follow Molodtsova and Papell (2009) and apply a version of the flexible-price monetary model. The fundamental f_t thus becomes:

$$f_t^{MON} = (m_t - m_t^*) - \eta(y_t - y_t^*). \tag{5.23}$$

We fix $\eta = 3$ which is successfully applied by Molodtsova and Papell (2009) and denote the model as **MON**.
Taylor Rule Fundamentals: We consider a symmetric Taylor rule with homogeneous coefficients, see Giacomini and Rossi (2010). Assuming that UIRP and PPP hold leads to

$$f_t^{TR} = (1 + \phi)(\pi_t - \pi_t^*) + \gamma(y_t^{gap} - y_t^{gap*}) + s_t. \tag{5.24}$$

The model based on Taylor rule fundamentals is indicated as **TR**.
Random Walk
The traditional benchmark model for exchange rate forecasts is the random walk. We therefore also include the commonly used random walk without drift which stipulates that the best predictor of next period's exchange rate is the current exchange rate. Thus, the random walk always predicts 'no change' in the h-months horizon exchange rate:

$$\widehat{s_{t+h}} - s_t = 0. \tag{5.25}$$

The random walk is denoted as **RW**.

5.4.3 Data

To construct the yield spread level $L_t^{\Delta sy}$ and slope $S_t^{\Delta sy}$ we use $\tau = 3$, $\tau = 36$ and $\tau = 120$ months sovereign bonds zero-coupon yields available from Bloomberg. The sovereign yield spreads Δsy_t^{τ} are then calculated as the difference between yields of equal maturity τ. Bloomberg yields are available from January 1995 onwards, so we use the time period from 1995:01-2014:12 for our analysis.

The corresponding nominal spot exchange rates are also taken from Bloomberg. We consider the US as the home country and thus the exchange rate is measured as the USD price per unit of foreign currency (USD/foreign currency). Therefore, a rise in the nominal exchange rate represents a depreciation of the USD and a lower value an appreciation of the USD.

The primary source of data used to construct the macroeconomic fundamentals for the benchmark models is the IMF's International Financial Statistics (IFS) database. We follow Molodtsova and Papell (2009) in selecting the data and calculating the fundamental differentials. We use the seasonally adjusted industrial production index as a proxy for output since GDP data is available only at the quarterly frequency.[13] The output gap is calculated as a percentage deviation of actual output from a linear trend. We use the money market rate as a measure of the short-term ($\tau = 1$ months) interest rate required to implement the 1- month ahead forecast for the IRD approach. The price level is measured by the consumer price index (CPI).[14] The inflation rate is the annual inflation rate, measured as the 12-month difference of the CPI. Finally, we use M1 to measure the money supply for all countries.[15]

We provide descriptive statistics of the yield spreads, foreign exchange rates

[13] Industrial production data for Australia and Switzerland are also only available at quarterly frequency and hence are transformed from quarterly to monthly observations using the quadratic-match average method as applied by Molodtsova and Papell (2009).

[14] Australian CPI data is also only available at quarterly frequency, and hence transformed from quarterly to monthly observations applying the same quadratic-match average interpolation.

[15] M1 data for the UK is not provided by the IMF so we use M1 data provided by Datastream.

and macroeconomic variables applied in this analysis in Appendix C.1.
The time period for our analysis also entails the global financial crisis from 2007 - 2009 which is well known to have caused major eruptions in financial markets (Guidolin and Tam, 2013; Fratzscher, 2009). To illustrate the impact on foreign exchange markets, we first provide an illustrative plot of the USD/AUD nominal exchange rate changes for all considered forecasting horizons in Figure 5.1.[16] The strong impact of the GFC on the exchange rate

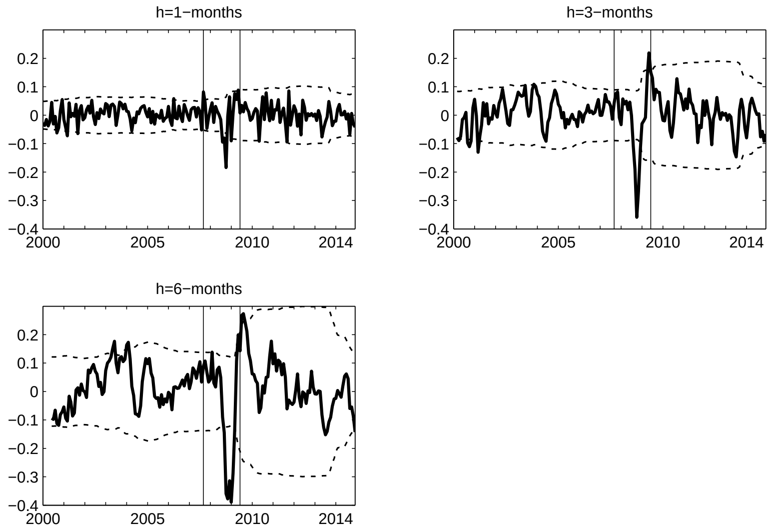

Figure 5.1. Time series of h=1, h=3 and h=6 months USD/AUD exchange rate changes with two standard deviation bands for the time period 1995:01 – 2014:12. The vertical lines indicate the GFC period from 2007:08-2009:05.

is clearly obvious, with abnormally large exchange rate changes during the crises period from August 2007 - May 2009.[17]
In order to provide more comprehensive insights to which extent all currencies have been affected by the GFC, we further calculate the standard deviation of the exchange rate changes during the GFC period against the standard deviation of the remaining sample. Table 5.1 summarizes the results. Val-

[16]For illustrative purposes and to save space, we focus on one exchange rate (USD/AUD) here. Plots for other currencies provide similar conclusions and are available upon request.
[17]Guidolin and Tam (2013) provide an extensive overview of the crisis dating literature and provide a conservative consensus period from August 2007 - May 2009.

horizon	USD/ AUD	USD/ CAD	USD/ CHF	USD/ JPY	USD/ GBP
1	2.06	2.22	1.64	1.29	1.74
3	2.53	2.12	1.33	1.43	2.70
6	3.05	2.41	1.47	1.04	2.89

Table 5.1. Ratios of the standard deviation of h=1, h=3 and h=6 months exchange rate changes during the GFC period (2007:08-2009:05) against the standard deviation of the remaining sample. Values larger than one indicate that the standard deviation of exchange rate changes during the GFC is higher than in the remaining sample.

ues larger than one indicate that the standard deviation of exchange rate changes during the GFC is higher than in the remaining sample. The results confirm that all currencies have been highly affected by the GFC with all of the ratios being significantly larger than one. Severely impacted have been in particular the USD/AUD, USD/CAD and USD/GBP currency pairs, where the standard deviation of exchange rate changes has been more than twice as high during the GFC as otherwise. We will therefore also consider the impact of the GFC on the forecasting accuracy in our analysis.

5.4.4 Forecasting Evaluation Measures

As indicated in Section 5.2, we examine the predictive power of the various models along different criteria. Each of the selected evaluation metrics has a different focus and we thus consider the use of these criteria as complementary. Taken together, they provide a multifaceted picture of the forecasting performance of the tested models. Naturally, depending on the purpose of a specific exercise one may favour one metric over the other.

RMSE The most commonly used measure of predictive ability in the out-of-sample exchange rate forecasting literature is the root-mean-square error (RMSE). The RMSE is a measure of global forecasting performance and summarizes the forecasting errors of a specific model M over the entire forecasting period P:

$$RMSE^M = \sqrt{\frac{1}{T-R-h} \sum_{t=R+h}^{T} (\widehat{\Delta s}^M_{t+h/t} - \Delta s^M_{t+h})^2}. \tag{5.26}$$

The lower the RMSE, the more accurate the forecast. To facilitate comparison between the yield spread models and the benchmarks models, we report the relative forecasting accuracy with the ratio of the RMSE from the respective yield spread model and the benchmark model M:[18]

$$RMSE\,ratio^{M} = \frac{RMSE^{YLDSPRD}}{RMSE^{M}}. \tag{5.27}$$

Accordingly, if $RMSE\,ratio^{M} < 1$, forecasts from the yield spread model are more accurate than the benchmark model.

Direction Accuracy From a market timing perspective it is often more important to correctly predict the direction of the exchange rate change. We therefore also apply a measure of direction accuracy (DA). The DA is computed as the number of correct predictions of the direction of change over the total number of predictions:

$$DA^{M} = \frac{1}{T-R-h} \sum_{t=R+h}^{T} a_t, \tag{5.28}$$

where $a_t = 1$ if the direction of change in period t is forecasted correctly and $a_t = 0$ otherwise. The higher the DA the better the model correctly predicts the direction of change.[19]

Density Forecasts Researchers have also started to realize that it is important to asses the uncertainty around point forecasts (Sarno et al., 2006; Rapach and Wohar, 2006; Inoue and Rossi, 2008; Hong et al., 2007). One way to achieve this is to use density forecasts. A density forecast is an estimate of the probability distribution of the point forecast, conditional on the information available at time t and thus represents a complete characterization of the uncertainty associated with the forecast (Rossi, 2015).

To evaluate density forecasts, Diebold et al. (1999) have pioneered the use of probability integral transforms (PIT). A PIT is the cumulative probability

[18] Note that this ratio is often reported reciprocal as the ratio of a model against the random walk.

[19] Compared to a random walk, a value above (below) 0.5 indicates a better (worse) forecasting performance than the naive RW model which has an equal chance of going up or down.

evaluated at the actual, realized value of the forecasted variable (Rosenblatt, 1952). Diebold et al. (1999) demonstrate that the PIT is uniform[20] and i.i.d. if the density forecast is correctly specified. In practice, density evaluation is thus implemented with formal tests measuring whether an observed PIT is U(0,1). Assume, we are interested in the distribution of the exchange rate change Δs^M_{t+h} which is being forecasted at time t. If the probability density of Δs^M_{t+h} is $f(\Delta s^M_{t+h})$ then the associated distribution function is

$$F(\Delta s^M_{t+h}) = \int_{-\infty}^{\Delta s^M_{t+h}} f(x)dx. \tag{5.29}$$

Following Rossi (2015), we determine the unknown variance of the forecast error with the estimated variance of the in-sample fitted errors and then test for violations of independence and uniformity of $\widehat{F}(\Delta s^M_{t+h})$ with a Berkowitz test.[21]

5.5 Out-of-sample Forecasting Results

5.5.1 RMSE

Table 5.2 reports the RMSEs for all investigated currencies and forecast horizons. The first line shows the results of the yield spread model in absolute terms. As we are mainly interested in the forecasting performance relative

[20] The uniformity property means that the probability that the realized value is higher (lower) than the forecasted value is the same (on average over time) no matter whether we consider high realizations or low realizations of the variable we are forecasting.

[21] Berkowitz (2001) suggests transforming the PIT series using the inverse of the standard normal cumulative distribution function. If the PITs are i.i.d. uniformly distributed, the transformed PITs then are normally distributed and i.i.d. The test for normality and independence of this transformed PIT sequence is then achieved through a likelihood-ratio test based on the estimated coefficients of an AR(1) process for the transformed PITs. The test examines the null hypothesis of the mean of the transformed PITs being equal to zero, the variance being equal to one, and the autocorrelation coefficient being equal to zero. Note that the test statistic is approximately χ^2 distributed with three degrees of freedom. See Berkowitz (2001) for further details.

to our yield spread models, we present the RMSEs of the other models below as a ratio against both approaches.[22] Hence, numbers smaller than one (reported in bold) indicate a smaller RMSE and accordingly superior forecasting performance of the yield spread model.

The table shows that our approach predominantly outperforms the fun-

	Spread Level $L^{\Delta sy}$					**Spread Slope** $S^{\Delta sy}$				
	AUD	CAD	CHF	JPY	GBP	AUD	CAD	CHF	JPY	GBP
					h = 1					
$L^{\Delta sy}/S^{\Delta sy}$	0.038	0.026	0.032	0.028	0.025	0.039	0.026	0.032	0.029	0.026
IRD	**0.983**	1.006	**0.982**	**0.988**	1.010	**0.991**	**0.988**	**0.985**	1.004	1.034
PPP	**0.997**	1.004	**1.000**	**0.992**	**0.954**	1.004	**0.986**	1.003	1.008	**0.977**
Mon	**0.999**	1.007	**0.991**	**0.959**	**0.973**	1.007	**0.989**	**0.994**	**0.975**	**0.996**
TR	**0.973**	**0.994**	**0.974**	**0.988**	**0.955**	**0.980**	**0.976**	**0.977**	1.004	**0.978**
RW	1.014	1.016	1.019	1.015	1.005	1.021	**0.998**	1.022	1.032	1.029
					h = 3					
$L^{\Delta sy}/S^{\Delta sy}$	0.074	0.045	0.055	0.053	0.048	0.074	0.043	0.057	0.056	0.050
IRD	**0.992**	1.020	**0.978**	**0.955**	1.001	**0.986**	**0.970**	1.001	1.019	1.044
PPP	1.009	1.009	1.015	**0.965**	**0.901**	1.003	**0.959**	1.039	1.030	**0.940**
Mon	1.021	1.032	**0.991**	**0.837**	**0.951**	1.015	**0.981**	1.015	**0.893**	**0.992**
TR	0.920	1.021	**0.968**	**0.927**	**0.862**	**0.915**	**0.971**	**0.991**	**0.989**	**0.899**
RW	1.069	1.039	1.077	1.046	1.025	1.063	**0.988**	1.103	1.117	1.070
					h = 6					
$L^{\Delta sy}/S^{\Delta sy}$	0.122	0.071	0.085	0.074	0.078	0.116	0.065	0.082	0.087	0.081
IRD	1.007	1.025	**0.949**	**0.908**	**0.998**	**0.950**	**0.946**	**0.908**	1.070	1.048
PPP	1.033	1.032	1.103	**0.868**	**0.986**	**0.975**	**0.953**	1.055	1.023	1.036
Mon	1.055	1.040	1.039	**0.634**	**0.943**	**0.996**	**0.959**	**0.994**	**0.747**	**0.990**
TR	**0.844**	1.025	**0.909**	**0.719**	**0.887**	**0.797**	**0.946**	**0.869**	**0.847**	**0.931**
RW	1.186	1.080	1.190	1.016	1.067	1.120	**0.997**	1.138	1.198	1.121

Table 5.2. Root mean squared errors (RMSEs) for the time period from 1995:01 – 2014:12 and h=1, h=3 and h=6 months-ahead forecasting horizons. The first line reports the RMSE for the yield spread models $L^{\Delta sy}$ and $S^{\Delta sy}$. The RMSEs of all other models and the random walk are expressed as the ratio against the yield spread model. Hence, numbers smaller than one (**reported in bold**) indicate a smaller RMSE and accordingly superior forecasting performance of the yield spread model. Numbers larger than one indicate inferior forecasting performance in terms of the RMSE. See Section 5.4.1 for a detailed description of the models.

damental models with the majority of the RMSE ratios being smaller than one. These results hold across all considered forecasting horizons and currencies, except for the USD/CAD pair for the spread level approach and the USD/JPY for the spread slope approach. Our approach seems to work

[22]Note that this is reciprocal to studies that focus only on the forecasting performance compared to the random walk and thus express the RMSE relative to the random walk. It is thus also reciprocal to Chapter 4.

particularly well for the Australian Dollar and British Pound. It is also interesting to note that both models seem to work well for different currencies, e.g. the spread level approach works rather well for the Japanese Yen and the British Pound and spread slope approach for the Australian and Canadian Dollar.

However, neither our yield spread model nor the traditional models are able to consistently beat the random walk in terms of the RMSE. Except for a few occasions, the random walk consistently yields the lowest RMSE (reflected in RMSE ratios larger than one). These results are not entirely surprising based on the findings of previous literature described in Section 5.2 and provide further evidence to the well documented failure of exchange rate models to outperform the random walk in terms of the RMSE (Cheung et al., 2005; Rossi, 2013). However, this conclusion changes when we turn to additional statistical evaluation measures and profitability.

5.5.2 Direction Accuracy

Our first alternative evaluation metric is the direction of change statistic. Table 5.3 reports the proportion of forecasts that correctly predict the direction of the exchange rate movement over horizon h. The first line reports the direction accuracy of the yield spread models. Below, we report the results for the benchmark models. The higher the proportion of correctly forecasted directions of change, the better. Superior direction accuracy of the yield spread models is indicated in bold.

In terms of direction accuracy, we find that our model is able to consistently beat the random walk with the DA statistics being predominantly larger than 0.50 except for the USD/CAD (spread level model) and the USD/GBP exchange rate (spread slope model). These results hold across all considered forecasting horizons.

Comparing the yield spread model to the traditional fundamental models, we also find promising results. Our model is consistently among the models with the highest proportion of forecasts that correctly predicted the direction of

	Spread Level $L^{\Delta sy}$					Spread Slope $S^{\Delta sy}$				
	AUD	CAD	CHF	JPY	GBP	AUD	CAD	CHF	JPY	GBP
					h = 1					
$L^{\Delta sy}/S^{\Delta sy}$	0.54	0.51	0.52	0.56	0.56	0.61	0.66	0.55	0.56	0.45
IRD	0.54	0.51	**0.51**	**0.52**	**0.55**	**0.58**	**0.51**	0.58	0.62	0.54
PPP	**0.51**	**0.47**	**0.49**	**0.51**	**0.43**	**0.45**	**0.44**	0.55	**0.53**	**0.43**
MON	**0.53**	0.56	**0.49**	**0.51**	**0.49**	**0.49**	**0.58**	0.58	**0.40**	0.45
TR	**0.51**	**0.48**	0.54	**0.53**	**0.54**	**0.51**	**0.48**	0.54	**0.53**	0.54
RW	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	0.50
					h = 3					
$L^{\Delta sy}/S^{\Delta sy}$	0.54	0.47	0.61	0.57	0.56	0.59	0.57	0.56	0.57	0.53
IRD	0.56	0.53	**0.57**	**0.54**	0.56	**0.56**	**0.53**	0.57	**0.54**	0.56
PPP	**0.44**	**0.43**	**0.47**	**0.55**	**0.43**	**0.44**	**0.43**	**0.47**	**0.55**	**0.43**
MON	0.55	0.52	**0.51**	**0.40**	**0.47**	**0.55**	**0.52**	**0.51**	**0.40**	**0.47**
TR	**0.51**	0.58	**0.50**	**0.49**	**0.53**	**0.51**	0.58	**0.50**	**0.49**	**0.53**
RW	**0.50**	0.50	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**
					h = 6					
$L^{\Delta sy}/S^{\Delta sy}$	0.56	0.45	0.57	0.58	0.53	0.61	0.66	0.55	0.56	0.45
IRD	0.58	0.51	0.58	0.62	0.54	**0.58**	**0.51**	0.58	0.62	0.54
PPP	**0.45**	**0.44**	**0.55**	**0.53**	**0.43**	**0.45**	**0.44**	0.55	**0.53**	**0.43**
MON	**0.49**	0.58	0.58	**0.40**	**0.45**	**0.49**	**0.58**	0.58	**0.40**	0.45
TR	**0.49**	0.55	**0.46**	**0.49**	0.57	**0.49**	**0.55**	**0.46**	**0.49**	0.57
RW	**0.50**	0.50	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	0.50

Table 5.3. Direction accuracy (DA) for the time period from 1995:01 – 2014:12 and h=1, h=3 and h=6 months-ahead forecasting horizons. The DA-statistic reports the proportion of forecasts that correctly predict the direction of the exchange rate movement over horizon h. The higher the proportion the better the direction accuracy. The first line reports the direction accuracy of the yield spread models. Direction accuracy smaller than the yield spread model's indicating a superior forecasting performance is **indicated in bold**. A value above (below) 0.5 indicates a better (worse) forecasting accuracy than a random walk. See Section 5.4.1 for a detailed description of the models.

change for all currencies except for the USD/CAD (spread level model) and the USD/GBP (spread slope model) currency. Similar to the RMSE, the direction accuracy seems to be particularly high for the Australian Dollar and Japanese Yen. The z-scores of a conventional test of the significance of proportions reported in Apendix C.2 also indicate that the superior direction accuracy of our approach is statistically significant for many cases, especially for the longer forecasting horizons.

5.5.3 Density Forecasts

We present the results of the evaluation of the density forecasts in Tables 5.4 and 5.5. Table 5.4 summarizes the results of the Berkowitz test for h=1 months forecasting horizon. The first line reports the Berkowitz test statistic for the yield spread models. The results for the benchmark models are reported below. Recall that a greater value of the $\chi^2(3)$ test statistic indicates a more likely rejection of the null hypothesis of an appropriate density forecasts for the model. Test statistics that are larger for the benchmark models than for SPRDLEV and SPRDSL therefore indicate a superior forecasting performance of the spread level and slope models and are highlighted in bold. Note that we focus on fundamental models, as for the random walk we cannot proxy the unknown variance of the forecast errors with the estimated variance of the in-sample fitted errors.

We find strong evidence that our approach yields more appropriate density

	Spread Level $L^{\Delta sy}$					**Spread Slope** $S^{\Delta sy}$				
	AUD	CAD	CHF	JPY	GBP	AUD	CAD	CHF	JPY	GBP
					h = 1					
$L^{\Delta sy}/S^{\Delta sy}$	10.91*	7.52	3.28	2.29	6.50	11.76*	6.57	3.41	2.02	6.75
IRD	**11.54***	6.98	**3.89**	1.03	6.04	11.54*	**6.98**	**3.89**	1.03	6.04
PPP	**13.37***	**9.66***	**5.01**	**4.17**	**13.73***	**13.37***	**9.66***	**5.01**	**4.17**	**13.73***
MON	**11.78***	6.41	2.86	**2.97**	**6.67**	**11.78***	6.41	2.86	**2.97**	6.67
TR	**19.39***	**11.39***	**7.52**	0.54	**9.48***	**19.39***	**11.39***	**7.52**	0.54	**9.48***

Table 5.4. Berkowitz test statistics for the time period from 1995:01 – 2014:12, all considered currencies against the USD and h=1 months-ahead forecasting horizons. The larger the test statistic, the more likely it is that the density forecasts are not correctly specified. The first line reports the Berkowitz test statistic for the yield spread models $L^{\Delta sy}$ and $S^{\Delta sy}$. The results for the benchmark models are reported below. Test statistics larger than for the yield spread models are **indicated in bold**. Rejection of the null of uniformity of the PITs at the 5% level is indicated with (*). See Section 5.4.1 for a detailed description of all models.

forecasts than the traditional fundamental models. Overall, the spread factor models deliver smaller Berkowitz statistics than the benchmark models for most of the currencies and forecasting horizons. In absolute terms, the null of uniformity of the PITs based on our model forecasts is only rejected for the USD/AUD exchange rate. Note, however, that for this currency pair the null of an appropriately specified density forecast is also rejected for all other models.

For longer forecasting horizons the validation techniques of the one period ahead density forecasts are no longer reliable. While one-period forecasts are i.i.d. under the null hypothesis of model adequacy and can therefore be validated using standard statistical techniques, multiple-period density forecasts and PITs are subject to common shocks which create temporal dependence and undermine the i.i.d. assumption. We therefore follow the bootstrap approach described in Dowd (2007) and construct i.i.d. resamples from the original PITs. For each PIT we create 10,000 resamples and then examine how often we end up with p-values below the adjusted significance level of 5%.

The results are summarized in Table 5.5. This table shows the number of rejections of the null of uniformity indicated by a Berkowitz test at a 5% level of significance based on the 10,000 resamples. Thus the smaller the number of rejections, the more appropriately are the density forecasts specified. Superior density forecasts of the respective yield spread model is indicated in bold.

Similar to the short-term results ($h = 1$) we find that for most currencies the

	Spread Level $L^{\Delta sy}$					**Spread Slope $S^{\Delta sy}$**				
	AUD	CAD	CHF	JPY	GBP	AUD	CAD	CHF	JPY	GBP
					h = 3					
$L^{\Delta sy}/S^{\Delta sy}$	8,634	8,153	4,805	1,815	9,336	8,568	6,910	6,522	4,817	9,657
IRD	**8,818**	7,440	**6,862**	**2,246**	**9,465**	**8,868**	**7,453**	**6,828**	2,262	9,469
PPP	**9,644**	**9,297**	**8,342**	**5,985**	**9,964**	**9,623**	**9,308**	**8,316**	**5,944**	**9,950**
MON	**8,952**	7,430	4,077	**8,899**	**9,778**	**8,929**	**7,569**	4,076	**8,961**	**9,775**
TR	**10,000**	**9,571**	**6,804**	**6,283**	**9,886**	**10,000**	**9,571**	**6,804**	**6,283**	**9,886**
					h = 6					
$L^{\Delta sy}/S^{\Delta sy}$	9,998	9,977	9,999	7,893	9,996	9,993	9,918	9,992	9,974	10,000
IRD	**10,000**	9,956	**10,000**	**9,149**	**9,999**	**9,999**	**9,962**	**10,000**	9,183	10,000
PPP	**10,000**	**10,000**	9,999	**9,998**	**10,000**	**10,000**	**9,999**	**9,998**	**9,999**	10,000
MON	**10,000**	**9,988**	9,960	**10,000**	**10,000**	**10,000**	**9,986**	9,951	**10,000**	10,000
TR	10,000	10,000	10,000	**10,000**	**10,000**	**10,000**	**10,000**	**10,000**	**10,000**	10,000

Table 5.5. Number of rejections at the 5% level of significance for the null hypothesis of uniformity of the PITs based on 10,000 created i.i.d. bootstrap resamples and application of the Berkowitz test (with p-values adjusted following Dowd (2007)). Tests are applied for the time period 1995:01 – 2014:12 and h=3 and h=6 months-ahead forecasting horizons. The smaller the number of rejections, the more likely it is that the density forecasts of the model are correctly specified. The first line reports the number of rejections for the yield spread models $L^{\Delta sy}$ and $S^{\Delta sy}$. Results for the benchmark models with more rejections than for the yield spread models are highlighted in bold. See Section5.4.1 for a detailed description of the models.

yield spread models yield more appropriate density forecasts than traditional models also for longer forecasting horizons. For a $h = 3$ months forecasting

horizon the number of rejections of the null hypothesis at the 5% level is the smallest for nearly all currencies. In relative terms, our model also does well for a $h = 6$ months forecasting horizon. In absolute terms however, the generally high number of rejections indicate that the density forecasts are not correctly specified for all assessed models.
We advocate two potential explanations. First, as described in Section 5.2, empirical exchange rate forecasting generally is a cumbersome task, especially over long horizons. It is not entirely surprising that the forecasting performance of the applied models deteriorates with the length of the forecasting horizon. Furthermore, as described in section 5.4.3, our forecasting period encompasses the global financial crisis (GFC) which had significant impacts on foreign exchange markets. We suspect that this will also impact the forecasting uncertainty as measured by tests of uniformity, as the exchange rates with the highest forecasting uncertainty (AUD, CAD and GBP) have been previously identified as severely impacted by the GFC in Section 5.4.3.

5.5.4 Impact of the GFC

To provide an intuitive indication of the impact of the GFC on the general forecasting accuracy, we first provide an illustrative plot of the forecasted 6-months USD/AUD exchange rate changes against the corresponding forecasting errors in Figure 5.2.[23] Observations from the GFC period (2007:08-2009:05) are highlighted with (*). This illustrative plot confirms, that the GFC period accounts for a high number of relatively large USD/AUD exchange rate changes throughout the forecasting period. It also becomes obvious that all models have problems to pick up the extend of these changes as the forecast errors are relatively large compared to the remaining sample. To thoroughly investigate the impact of the crises on the forecasting accuracy, we further re-calculate all statistical evaluation measures excluding the

[23]Due to the high number of currency/horizon combinations, we focus on one illustrative forecasting horizon ($h = 6$ months) currency which has been identified in Section 5.4.3 as being highly impacted by the GFC (USD/AUD). Further plots of other currencies and forecasting horizons are available upon request.

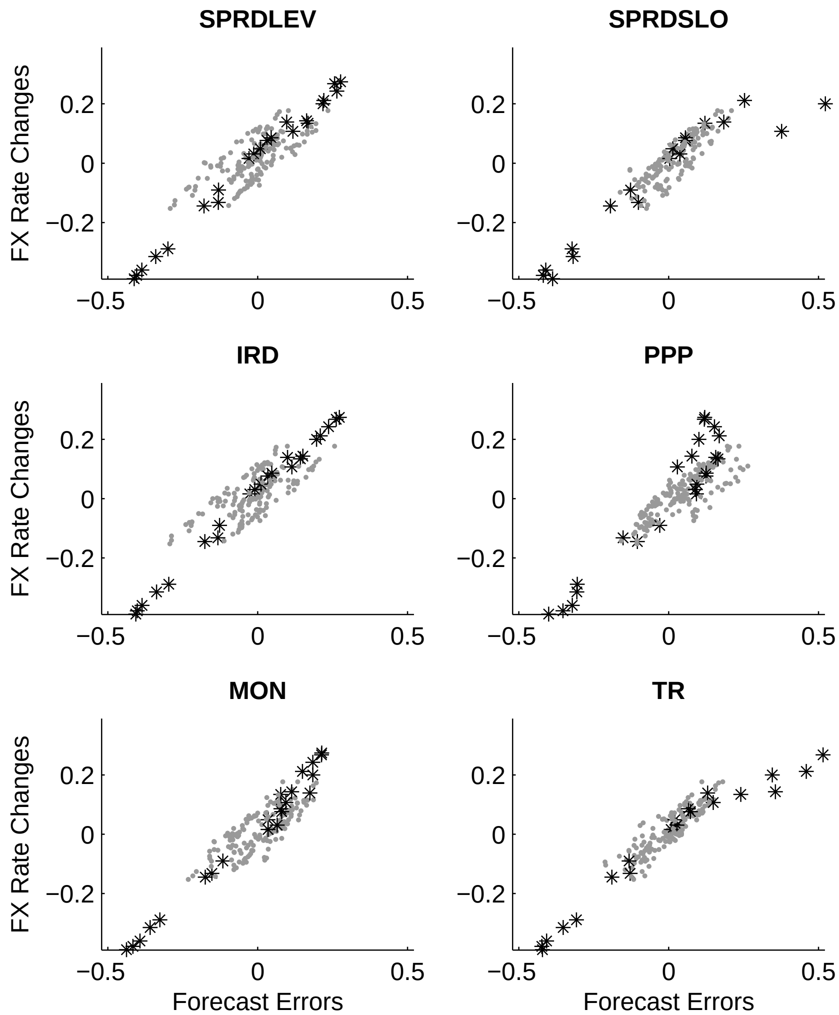

Figure 5.2. USD/AUD exchange rate changes vs. forecasting errors for h=6 months and all considered forecasting models. Observations from the GFC period (2007:08-2009:05) are indicated with (*). See Section 5.4.2 for a detailed description of the selected models.

forecasts of the crises period (2007:08-2009:05) from the sample. The respective tables are reported in Appendix C.3.

The results show, that Berkowitz statistics for h=1 and h=3 months forecasting horizons improve significantly, indicating that the forecasting uncertainty is much lower prior to and after the GFC. Not surprisingly, this effect is the strongest for the exchange rates identified as severely impacted by the GFC in

Section 5.4.3. While the forecasting performance of the models also slightly improves for the forecasting horizon $h = 6$ months, it still remains rather poor, indicating that longer term exchange rate forecasts are generally associated with high forecasting uncertainty.

With regards to other forecast evaluation measures, we find the impact to be somewhat limited. The RMSEs reported in Table C.5 generally decrease as the large forecasting errors during the crisis period are excluded, but in relative terms the results remain rather stable. The overall impact on the direction accuracy reported in Table C.6 is – not surprisingly – rather small, as the crisis only comprises a relatively small part (20 observations) of the entire forecasting period and the magnitude of forecasting errors does not affect the DA metric.

Overall, we do not find much difference in terms of relative performance between the models for all evaluation metrics when the crises period is excluded. It thus seems that, while the GFC generally had some impact on the forecasting accuracy in terms of the RMSE and a significant impact on the density forecasts, no model stands out as a particular good or bad performer during the crisis.

5.6 Trading Strategy

5.6.1 Trading Rule Implementation

The ultimate test of predictive power is the profitability of the forecasts (Abhyankar et al., 2005; Corte et al., 2009; Moosa and Burns, 2014). After all, statistical evidence of exchange rate predictability does not guarantee an investor to make profits with a strategy exploiting this predictive power.

To asses the profitability of the different models, we thus implement a simple trading strategy that utilizes the respective exchange rate forecasts $\widehat{\Delta s}^{M}_{t+h}$. Following Moosa and Burns (2014) we apply an intuitive approach that involves period-by-period trading based on the forecasted h-month horizon excess returns $\widehat{xs}^M_{t/t+h}$ predicted by model M:

$$\widehat{xs}^M_{t/t+h} = (i^\tau_t - i^{\tau,*}_t) - \widehat{\Delta s}^M_{t+h}. \tag{5.30}$$

Note that the maturity τ of the interest rate differential $(i^\tau_t - i^{\tau,*}_t)$ equals the forecasting horizon h.

The decision rule for trading is then based on whether the forecasted excess return $\widehat{xs}^M_{t/t+h}$ derived from the model forecast $\widehat{\Delta s}^M_{t+h}$ is positive or negative. A negative excess return, for example, indicates that the model forecasts an appreciation of the foreign currency which outweighs the interest rate differential[24] and thus suggests an investment in the foreign currency. The trading rule can therefore be defined as:

$$\begin{array}{lll} if \quad \widehat{xs}^M_{t/t+h} > 0 & \rightarrow & invest\ in\ home\ currency, \\ if \quad \widehat{xs}^M_{t/t+h} < 0 & \rightarrow & invest\ in\ foreign\ currency. \end{array}$$

We then calculate the actual return xs^M_{t+h} for every trade based on the actual exchange rate changes over the corresponding horizon as:

$$xs^M_{t+h} = \begin{cases} (i^\tau_t - i^{\tau,*}_t) - \Delta s_{t+h} & for\ investments\ in\ home\ currency \\ (i^{\tau,*}_t - i^\tau_t) + \Delta s_{t+h} & for\ investments\ in\ foreign\ currency. \end{cases} \tag{5.31}$$

Note that we ignore transaction costs as the main purpose of our analysis is to compare the profitability of the considered models and the same number of trades are executed for all models.

We implement the trading strategy for every month of the forecasting period $P = T - R - h$ and summarize the annualized returns in the mean return $\overline{xs}^M$ across the entire forecasting period. We then calculate the risk adjusted profitability xs^M_{ra} for every model M as a ratio of the mean return $\overline{xs}^M$ and

[24] Recall, that a rise in the nominal exchange rate s represents a depreciation of the home currency (USD) and a lower value an appreciation.

the standard deviation σ_{xs}^M of the returns:

$$xs_{ra}^M = \frac{\overline{xs}^M}{\sigma_{xs}^M}. \tag{5.32}$$

This can be illustrated with a simple example. Let us assume that s is the log of the USD/CAD exchange rate and the $\tau = 6$ months interest rate differential $i_t^{\tau=6} - i_t^{\tau=6,*}$ between the US (home country) and Canada (foreign country) is $5\% - 2\% = 3\%$. Let us further assume that model M predicts a depreciation of the US dollar over the next $h = 6$ months period of $\widehat{\Delta s}_{t+6}^M = 4\%$.[25] The predicted negative semi-annual excess return $\widehat{xs}_{t/t+6}^M$ of $\frac{3\%}{2} - 4\% = -2.5\%$[26] would indicate an investment in the Canadian dollar as the predicted CAD appreciation outweighs the interest rate differential. Further assuming that the actual appreciation Δs_{t+6} of the Canadian dollar over the next 6-months horizon is only 1%, the investment in the Canadian dollar would yield a negative actual semi-annual return of $xs_{t/t+6}^M = -\frac{3\%}{2} + 1\% = -0.5\%$ and thus an annualized return of $xs_{t/t+6,ann.}^M = -0.5\% \cdot 2 = -1.0\%$. This strategy is now implemented for every month of the forecasting period. If we assume that the mean return of all these trades is $\overline{xs}^M = 2.8\%$ with a standard deviation of $\sigma_{xs}^M = 14.6\%$ the annualized return-risk ratio or risk adjusted return is $xs_{ra}^M = \frac{2.8\%}{14.6\%} = 0.19$.

5.6.2 Risk-adjusted Trading Returns

We report the annualized risk adjusted returns or return-risk ratio xs_{ra}^M of the implemented trading strategy for all considered models including the random walk in Table 5.6. The first line reports the return-risk ratios for the yield spread models. Risk adjusted returns smaller than those of the yield spread models then indicate a superior risk-return relationship for our model and are highlighted in bold.

[25] Again, recall that a rise in the nominal exchange rate s represents a depreciation of the home curreny (USD).

[26] Note, that the difference in annual interest rates has to be adjusted to match the semi-annual ($h = 6$-month) horizon of the exchange rate change.

Focusing on the yield spread models first, we find positive risk-adjusted

	Spread Level $L^{\Delta sy}$					**Spread Slope** $S^{\Delta sy}$				
	AUD	CAD	CHF	JPY	GBP	AUD	CAD	CHF	JPY	GBP
					h = 1					
$L^{\Delta sy}/S^{\Delta sy}$	0.06	0.03	0.02	0.04	0.10	0.09	0.15	0.01	0.14	-0.01
IRD	**0.04**	0.07	**-0.02**	0.06	**0.04**	**0.04**	**0.07**	**-0.02**	**0.06**	0.04
PPP	**0.04**	0.10	**-0.02**	0.07	**-0.08**	**0.04**	**0.10**	**-0.02**	**0.07**	**-0.08**
MON	**0.02**	0.04	**-0.04**	**0.01**	**0.03**	**0.02**	**0.04**	**-0.04**	**0.01**	0.03
TR	**0.05**	0.04	0.12	0.11	**0.07**	**0.05**	**0.04**	0.12	**0.11**	0.07
RW	**-0.02**	0.05	0.11	0.04	**0.03**	**-0.02**	0.05	0.11	**0.04**	0.03
					h = 3					
$L^{\Delta sy}/S^{\Delta sy}$	0.00	0.01	0.10	0.09	0.06	0.09	0.22	0.01	0.16	0.04
IRD	0.05	0.10	**0.08**	**0.05**	0.13	**0.05**	**0.10**	0.08	**0.05**	0.13
PPP	**0.00**	**-0.01**	**-0.03**	0.10	**-0.14**	**0.00**	**-0.01**	**-0.03**	**0.10**	**-0.14**
MON	**-0.02**	**0.01**	**0.01**	**-0.08**	**0.03**	**-0.02**	**0.01**	0.01	**-0.08**	**0.03**
TR	0.06	0.26	**0.02**	**0.05**	**-0.08**	**0.06**	0.26	0.02	**0.05**	**-0.08**
RW	**-0.02**	0.08	0.25	0.17	0.08	**-0.02**	**0.08**	0.25	0.17	0.08
					h = 6					
$L^{\Delta sy}/S^{\Delta sy}$	0.13	-0.07	0.11	0.14	0.05	0.21	0.28	0.07	0.10	-0.06
IRD	0.15	-0.01	0.13	0.15	0.08	**0.15**	**-0.01**	0.13	0.15	0.08
PPP	**0.06**	**-0.08**	0.12	**0.07**	**0.03**	**0.06**	**-0.08**	0.12	**0.07**	0.03
MON	**-0.07**	-0.04	0.17	**-0.13**	0.08	**-0.07**	**-0.04**	0.17	**-0.13**	0.08
TR	**-0.07**	0.22	**0.00**	**-0.04**	**-0.14**	**-0.07**	**0.22**	**0.00**	**-0.04**	**-0.14**
RW	**-0.09**	0.13	0.39	0.17	0.09	**-0.09**	**0.13**	0.39	0.17	0.09

Table 5.6. Annualized risk adjusted returns or return-risk ratios of a monthly trading strategy based on model exchange rate predictions for the time period from 1995:01 – 2014:12 and h=1, h=3 and h=6 months-ahead forecasting horizons. The risk adjusted returns are calculated as a ratio of the mean return $\overline{xs}^M$ and the standard deviation σ^M_{xs} of the returns: $xs^M_{ra} = \overline{xs}^M/\sigma^M_{xs}$. The higher the risk adjusted return the better is the risk-return relation of the model's forecasts. Risk adjusted returns's smaller than the yield spread models indicate a superior risk-return relationship of the yield spread models, and are **reported in bold**. See Section 5.4.1 for a detailed description of the models and Section 5.6.1 for a description of the trading strategy.

returns for nearly all currencies and forecast horizons. The spread slope approach seems to work particularly well for the Canadian Dollar, where the returns yield annualized return-risk ratios of up to 0.28.

Overall, we also find promising results for our approach relative to the fundamental models and the random walk when the forecasting accuracy is assessed in terms of trading profitability. Our models predominantly generate higher risk-adjusted returns than the traditional fundamental models.[27] This holds in particular for the Australian Dollar for both models and the Canadian Dollar and Japanese Yen for the spread slope approach. Note, that the ran-

[27] We note that this difference in risk-adjusted trading returns is not statistically significant

dom walk does surprisingly well, which is further validation for the success of the carry trade.[28] Nevertheless, our yield spread model is able to beat the random walk for some of the currency/forecast horizon combinations.[29]

In general, profits and losses seem to become larger with a widening forecasting horizon. Trading profitability also seems to be more currency specific than the previously applied statistical evaluation metrics, with some currencies working exceptionally well (especially the Australian Dollar) across all forecasting horizons while the profitability for others (e.g. the Swiss Francs) is somewhat disappointing.

It is also important to note, that simply trading based on exchange rate forecasts without additional discretionary trading rules or the creation of portfolios does generally not produce promising results. The majority of the annualized return-risk ratios is not very high and several trading strategies based on forecasts created by the benchmark models even yield negative risk-adjusted returns. Taking into account reasonable transaction costs would probably see more of the remaining profits diminish.

5.7 Conclusion

This paper investigates whether two selected financial variables — the level and slope of the yield spread curve between two economies — can successfully forecast exchange rates. We apply these two variables as proxies reflecting the market's unobservable expectations of current and future macroeconomic fundamentals and investigate their forecasting accuracy in an extensive out-

[28]Note that under the random walk without drift, the forecast change in the exchange rate is always zero. This means that the decision rule leads to going short on the low-interest currency and long on the high-interest currency which represents the common carry trade.

[29]We also investigate the impact of the GFC on profitability by re-calculating the risk adjusted returns excluding the GFC period (Tables reported in Appendix C.3). While some of the individual results for the different forecasting model/horizon/currency combinations change quite substantially, the overall relative forecasting performance when comparing the models remains rather stable compared the results presented in this section.

of-sample forecasting exercise against traditional models based on interest rate, price, monetary and Taylor rule fundamentals as well as the random walk.

We find that our approach is able to consistently beat traditional fundamental exchange rate models in terms of all considered forecasting evaluation metrics such as the RMSE, direction accuracy and density forecasts. While it fails to beat the random walk in terms of the RMSE – which should hardly be surprising based on previous findings in the literature – it is also superior to a random walk in forecasting the direction of exchange rate changes.

We also assess the profitability of our approach by implementing a period-by-period trading strategy and find that trading based on the implemented yield spread level and spread slope models consistently generates higher risk-adjusted returns compared to traditional fundamental models and is also able to beat the random walk for several currencies.

As our sample includes the GFC period from 2007-2009, we further investigate the impact of this time period on the forecasting performance. We find that, while the GFC generally had a strong impact on foreign exchange markets, the effect on point forecasts is somewhat limited in terms of the RMSE and direction accuracy as well as profitability. However, not surprisingly, we find that the GFC period has significantly increased the forecasting uncertainty as measured by density forecasts.

Interestingly, depending on the forecasting measures the results seem to be at least partly currency specific. While the superior performance of our approach compared to the fundamental models is comprehensive in terms of direction accuracy and density forecasts, it is less distinct for some currencies in terms of the RMSE. The investigated density forecasts also reveal that some currencies can generally be forecasted with higher degress of certainty than others. Trading on exchange rate forecasts also seems to work generally well for some currencies while it fails to deliver positive returns for others.

This difference in conclusions depending on the choice of statistical and economic forecasting metrics highlights the importance of applying a variety of measures to provide a conclusive assessment of a model's forecasting ability. Simply minimizing the mean squared error is not always adequate from an economic standpoint and may miss out on important aspects of exchange rate forecasts such as the direction of change and profitability.

Overall, our promising results provide further evidence that there is an important place for models based on financial variables in exchange rate forecasting. Evans and Lyons (2007), Guo and Savickas (2008) and Rime et al. (2010) have previously shown that financial variables related to stock returns and order flow have the ability to improve exchange rate forecasts. Recently, Chen and Tsang (2013) and Bui and Fisher (2016) also found predictive power of cross-country Nelson-Siegel factors for exchange rates conducting an in-sample analysis. These factors are closely associated with the empirical spread level, slope and curvature suggested in this study. We complement their findings and the conclusions of the previous Chapter 4 and show that the empirical yield spread level and slope also successfully forecast exchange rates out-of-sample.

There are several possible explanations why the yield spread level and slope work so well in forecasting exchange rates. When the exchange rate is understood as an asset price and determined by the sum of expected future fundamentals these indicators act as natural proxies whose forward-looking character summarizes the market expectations for these fundamentals. Further, because sovereign yield spreads and foreign exchanges are susceptible to the same macroeconomic risk the expected risk premia that investors require for holding these assets might closely relate to each other.

From an empirical forecasting perspective it is also favorable to focus on a small number of variables which reflect a broad range of unobservable macroeconomic information and business conditions. First, this allows for parsimonious modeling and previous research has shown, that simple specifications often deliver accurate forecasts, see, for example, Clark and McCracken (2013). Second, our approach is more flexible to changes in business conditions over time and less vulnerable to the omitted variables bias than traditional models based on selected observable fundamentals.

It is important to note that exchange rates are notoriously difficult to forecast empirically and the forecasting success often depends on the choice of currency, forecast horizon, in-sample window length, sample period and forecast evaluation method. We also find that our approach works better for some currencies and horizons than for others. But considering all applied evaluation metrics as well as the trading profitability, we provide a promising, multifaceted picture of the forecast performance of the proposed yield spread models. In addition, our approach is parsimonious and based on readily available, market-driven data, which makes it straightforward to apply in practice.

We thus hope that our study sparks a renewed interest in the empirical assessment of exchange rate forecasting models based on financial variables, as they are an intuitive and promising resolution, in particular when exchange rates are understood as an asset price. Further research is required, for instance, to fully understand the dynamics of the term structure of yield spreads and its relation to macroeconomic fundamentals and exchange rates. It may also be worthwhile to combine our approach with other financial variables, e.g. stock returns, or factors derived from a range of financial variables which may reflect other aspects of the business cycle and exchange rate determination, to further increase the forecasting accuracy.

Appendix C

C.1 Descriptive Statistics

Currency	Mean	Std	Min	Max	Skewness	Kurtosis
			h = 1			
USD/AUD	0.000	0.035	-0.18	0.09	-0.73	5.92
USD/CAD	0.001	0.024	-0.14	0.09	-0.64	7.94
USD/CHF	0.001	0.031	-0.12	0.13	0.15	4.38
USD/JPY	-0.001	0.032	-0.10	0.16	0.52	5.78
USD/GBP	0.000	0.024	-0.10	0.09	-0.34	4.59
			h = 3			
USD/AUD	0.002	0.064	-0.36	0.22	-1.07	8.24
USD/CAD	0.003	0.040	-0.17	0.15	-0.19	6.01
USD/CHF	0.003	0.051	-0.12	0.13	0.05	2.64
USD/JPY	-0.003	0.058	-0.16	0.22	0.47	4.06
USD/GBP	0.000	0.043	-0.20	0.14	-0.83	7.70
			h = 6			
USD/AUD	0.005	0.095	-0.39	0.27	-0.77	6.05
USD/CAD	0.005	0.059	-0.22	0.16	-0.32	5.10
USD/CHF	0.006	0.072	-0.20	0.18	-0.10	2.69
USD/JPY	-0.006	0.082	-0.20	0.22	0.04	2.99
USD/GBP	0.000	0.065	-0.32	0.16	-1.46	8.48

Table C.1. Descriptive statistics for h=1, h=3 and h=6 months nominal exchange rate changes (home currency price per unit of foreign currency) for the time period from 1995:01 – 2014:12. Source: Bloomberg.

Country	Mean	Std	Min	Max	Skewness	Kurtosis
			log(p)			
US	4.478	0.140	4.23	4.69	-0.08	1.68
AU	4.449	0.160	4.19	4.71	0.03	1.66
CA	4.501	0.115	4.31	4.68	-0.09	1.68
CH	4.553	0.044	4.47	4.62	-0.20	1.57
JP	4.618	0.013	4.60	4.65	0.62	2.33
UK	4.493	0.122	4.30	4.72	0.47	1.96
			π			
US	0.023	0.011	-0.02	0.05	-0.67	5.04
AU	0.027	0.013	0.00	0.06	0.26	3.64
CA	0.019	0.009	-0.01	0.05	-0.04	3.89
CH	0.007	0.008	-0.01	0.03	0.30	3.14
JP	0.001	0.011	-0.03	0.04	1.04	4.60
UK	0.021	0.010	0.01	0.05	0.86	3.55
			log(m1)			
US	7.272	0.294	6.97	7.97	1.00	2.80
AU	5.123	0.399	4.32	5.74	-0.34	2.00
CA	5.812	0.478	5.02	6.62	0.09	1.83
CH	5.631	0.402	4.93	6.36	0.31	2.03
JP	5.871	0.456	4.95	6.41	-0.61	1.77
UK	6.453	0.532	5.40	7.22	-0.34	1.84
			i			
US	2.854	2.350	0.07	6.54	0.06	1.30
AU	5.054	1.361	2.50	7.52	-0.06	2.54
CA	3.017	1.843	0.24	8.06	0.34	2.38
CH	0.967	1.055	-2.00	3.50	0.56	2.66
JP	0.217	0.329	0.00	2.25	3.51	19.76
UK	3.818	2.396	0.40	7.50	-0.39	1.67
			log(y)			
US	4.601	0.095	4.36	4.77	-0.90	3.39
AU	4.531	0.098	4.32	4.73	-0.31	2.57
CA	4.490	0.148	4.21	4.71	-0.49	2.07
CH	4.638	0.155	4.36	4.86	-0.07	1.60
JP	4.623	0.069	4.34	4.76	-0.70	5.18
UK	4.657	0.053	4.55	4.73	-0.68	1.96
			y^{gap}			
US	-0.048	0.062	-0.21	0.06	-0.33	2.79
AU	-0.011	0.032	-0.08	0.07	0.07	2.35
CA	-0.018	0.036	-0.09	0.07	-0.09	2.08
CH	0.020	0.047	-0.07	0.15	0.70	2.89
JP	0.005	0.067	-0.28	0.11	-1.33	5.90
UK	-0.043	0.040	-0.14	0.03	-0.23	2.51

Table C.2. Descriptive statistics of the macroeconomic time series for the US, Australia, Canada, Switzerland, Japan and the United Kingdom for the time period from 1995:01 – 2014:12. Sources: IMF's International Financial Statistics, Datastream. See Section 5.4.3 for a detailed description of the variables.

Spread	Mean	Std	Min	Max	Skewness	Kurtosis
			US - AU			
3m	-2.180	1.619	-5.50	0.94	0.14	2.06
60m	-1.864	1.314	-4.89	0.72	0.05	2.16
120m	-1.313	0.705	-3.20	0.14	-0.41	2.71
Spread Level	-1.786	1.146	-4.11	0.45	0.19	2.10
Spread Slope	0.867	1.269	-1.51	3.13	-0.09	1.85
			US - CA			
3m	-0.273	0.965	-2.59	2.32	0.56	3.21
60m	-0.352	0.721	-2.14	1.30	0.20	2.42
120m	-0.189	0.578	-2.02	0.66	-1.12	3.86
Spread Level	-0.271	0.667	-1.88	1.20	0.07	2.72
Spread Slope	0.084	0.895	-2.65	1.44	-0.76	3.33
			US - CH			
3m	1.560	1.586	-1.09	4.52	0.27	1.46
60m	1.712	1.271	-0.61	4.56	0.37	1.88
120m	1.859	0.562	0.15	3.37	0.20	2.83
Spread Level	1.710	1.100	-0.22	4.01	0.30	1.68
Spread Slope	0.299	1.205	-2.08	2.32	-0.21	1.62
			US - JP			
3m	2.502	2.181	-0.46	6.28	0.10	1.36
60m	2.868	1.824	0.09	6.23	-0.04	1.60
120m	2.874	0.896	0.76	4.79	-0.35	2.65
Spread Level	2.748	1.582	0.32	5.62	-0.02	1.59
Spread Slope	0.372	1.568	-2.61	2.84	-0.14	1.49
			US - UK			
3m	-0.987	1.048	-3.40	0.89	-0.68	2.28
60m	-0.759	0.745	-2.59	0.56	-0.40	2.27
120m	-0.271	0.656	-2.39	1.39	-0.52	4.17
Spread Level	-0.672	0.700	-2.25	0.65	-0.21	2.16
Spread Slope	0.716	1.049	-1.56	3.08	0.18	2.15

Table C.3. Descriptive statistics of sovereign yield spreads for 3-months, 36-months and 120-months maturity and the empirical yield spread level $L_t^{\Delta sy}$ and spread slope $S_t^{\Delta sy}$ for the time period from 1995:01 – 2014:12. Source: Bloomberg. See Section 5.4.2 for the calculation of the sovereign yield spreads and the spread level and slope.

C.2 Statistical Significance of Direction Accuracy

	Spread Level $L^{\Delta sy}$					Spread Slope $S^{\Delta sy}$				
	AUD	CAD	CHF	JPY	GBP	AUD	CAD	CHF	JPY	GBP
					h = 1					
IRD	0.00	0.00	**0.30**	**1.04**	**0.15**	-0.15	**0.75**	**0.15**	**1.19**	-0.90
PPP	**1.04**	**0.90**	**0.75**	**1.34**	**3.31***	**0.89**	**1.64**	**0.60**	**1.49**	2.26*
MON	**0.30**	-1.35	**0.60**	**1.34**	**1.64**	**0.15**	-0.60	**0.45**	**1.49**	0.60
TR	**1.49**	0.00	**0.45**	**0.90**	**0.75**	**1.34**	**0.75**	**0.30**	**1.04**	-0.30
RW	**1.19**	**0.15**	**0.45**	**1.49**	**1.49**	**1.04**	**0.89**	**0.30**	**1.64**	0.45
					h = 3					
IRD	-0.60	-1.50	**0.91**	**0.75**	0.00	**0.76**	**1.20**	-0.30	**0.75**	-0.60
PPP	**2.72***	**1.06**	**3.76***	**0.45**	**3.49***	**4.08***	**3.78***	**2.55***	**0.45**	**2.88***
MON	-0.30	-1.20	**2.70***	**4.59***	**2.25***	**1.05**	**1.50**	**1.50**	**4.59***	1.65
TR	**0.75**	-2.57"	**3.45***	**1.65**	**1.80**	**2.10***	**0.15**	**2.25***	**1.65**	1.20
RW	**1.05**	-0.75	**2.85***	**1.80**	**1.50**	**2.40***	**1.95**	**1.65**	**1.80**	0.90
					h = 6					
IRD	-0.61	-1.81	-0.15	-0.93	-0.46	**0.77**	**3.93***	-0.77	-1.56	-2.58"
PPP	**3.04***	**0.15**	**0.61**	**1.51**	**2.44***	**4.41***	**5.94***	0.00	**0.91**	**0.31**
MON	**1.97***	-3.68"	-0.31	**4.94***	**2.13***	**3.33***	**2.15***	-0.92	**4.32***	0.00
TR	**3.20***	-2.43"	**1.97***	**5.96***	**1.06**	**4.57***	**3.34***	**1.36**	**5.34***	-1.06
RW	**1.59**	-1.44	**1.89**	**2.19***	**0.68**	**2.95***	**4.31***	**1.29**	**1.59**	-1.44

Table C.4. Z-scores of a conventional test of the significance of proportions for the direction accuracy (DA) of the yield spread approaches against the benchmark models for the time period from 1995:01 – 2014:12 and h=1, h=3 and h=6 months-ahead forecasting horizons. The DA statistic reported in the corresponding Table 5.3 denote the proportion of forecasts that correctly predict the direction of the exchange rate movement over horizon h. Positive z-scores (**highlighted in bold**) indicate a higher direction accuracy of the yield spread models. Negative z-scores indicate a lower direction accuracy. Z-scores above/below 1.96/-1.96 (indicated with */") imply statistical significance of the results on a 5% or lower level. See section 5.4.1 for a detailed description of the models.

C.3 Forecasting Results excluding GFC period

	Spread Level $L^{\Delta sy}$					**Spread Slope** $S^{\Delta sy}$				
	AUD	CAD	CHF	JPY	GBP	AUD	CAD	CHF	JPY	GBP
					h = 1					
$L^{\Delta sy}/S^{\Delta sy}$	0.033	0.022	0.029	0.027	0.023	0.033	0.022	0.029	0.028	0.023
IRD	**0.971**	1.003	**0.972**	**0.984**	**0.994**	**0.988**	**0.987**	**0.981**	1.008	1.020
PPP	**0.988**	**0.990**	**0.984**	**0.982**	**0.967**	1.006	**0.974**	**0.993**	1.006	**0.993**
MON	1.003	1.018	**0.981**	**0.966**	**0.962**	1.022	1.002	**0.990**	**0.990**	**0.988**
TR	**0.966**	**0.980**	**0.991**	**1.000**	1.000	**0.984**	**0.964**	1.000	1.024	1.027
RW	1.017	1.017	1.011	1.011	1.013	1.035	**1.000**	1.020	1.035	1.040
					h = 3					
$L^{\Delta sy}/S^{\Delta sy}$	0.059	0.037	0.052	0.049	0.036	0.060	0.037	0.055	0.054	0.040
IRD	**0.972**	1.018	**0.966**	**0.932**	**0.981**	**0.986**	**0.994**	1.007	1.030	1.076
PPP	**0.970**	**0.950**	**0.994**	**0.917**	**0.927**	**0.985**	**0.927**	1.037	1.013	1.017
MON	1.019	1.012	**0.974**	**0.913**	**0.891**	1.034	**0.988**	1.016	1.009	0.977
TR	**0.983**	**0.976**	**0.948**	**0.971**	1.010	**0.997**	**0.953**	**0.989**	1.073	1.107
RW	1.087	1.018	1.064	1.033	1.040	1.103	**0.993**	1.110	1.141	1.141
					h = 6					
$L^{\Delta sy}/S^{\Delta sy}$	0.098	0.054	0.080	0.072	0.057	0.092	0.052	0.077	0.087	0.065
IRD	**0.994**	1.031	**0.926**	**0.893**	**0.972**	**0.929**	**0.982**	**0.895**	1.080	1.113
PPP	**0.949**	**0.909**	1.053	**0.845**	**0.925**	**0.886**	**0.865**	1.017	1.022	1.060
MON	1.076	**0.968**	1.037	**0.792**	**0.849**	1.005	**0.922**	1.002	**0.958**	**0.972**
TR	1.038	**0.925**	**0.908**	**0.855**	1.050	**0.970**	**0.881**	**0.877**	1.034	1.203
RW	1.330	1.029	1.179	1.031	1.112	1.242	**0.980**	1.138	1.247	1.274

Table C.5. Root mean squared errors (RMSEs) for the time period from 1995:01 – 2014:12 (**excluding the GFC period from 2007:8-2009:5**) and h=1, h=3 and h=6 months-ahead forecasting horizons. The first line reports the RMSE for the yield spread models $L^{\Delta sy}$ and $S^{\Delta sy}$. The RMSEs of all other models and the random walk are expressed as the ratio against the yield spread models. Hence, numbers smaller than one (**reported in bold**) indicate a smaller RMSE and accordingly superior forecasting performance of the yield spread model . Numbers larger than one indicate inferior forecasting performance in terms of the RMSE. See section 5.4.1 for a detailed description of the models.

	Spread Level $L^{\Delta sy}$					Spread Slope $S^{\Delta sy}$				
	AUD	CAD	CHF	JPY	GBP	AUD	CAD	CHF	JPY	GBP
					h = 1					
$L^{\Delta sy}/S^{\Delta sy}$	0.55	0.51	0.53	0.56	0.55	0.54	0.53	0.52	0.57	0.52
IRD	**0.54**	**0.50**	**0.53**	**0.52**	**0.54**	0.54	**0.50**	0.53	**0.52**	0.54
PPP	**0.50**	**0.47**	**0.48**	**0.51**	**0.42**	**0.50**	**0.47**	**0.48**	**0.51**	**0.42**
MON	**0.54**	0.56	**0.51**	**0.50**	**0.46**	0.54	0.56	**0.51**	**0.50**	**0.46**
TR	**0.49**	**0.49**	0.56	**0.54**	0.56	**0.49**	**0.49**	0.56	**0.54**	0.56
RW	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**
					h = 3					
$L^{\Delta sy}/S^{\Delta sy}$	0.54	0.49	0.60	0.58	0.56	0.59	0.58	0.58	0.58	0.54
IRD	0.56	0.54	**0.58**	**0.54**	0.56	**0.56**	**0.54**	0.58	**0.54**	0.56
PPP	**0.44**	**0.44**	**0.46**	**0.54**	**0.44**	**0.44**	**0.44**	**0.46**	**0.54**	**0.44**
MON	0.55	0.55	**0.48**	**0.38**	**0.45**	**0.55**	**0.55**	**0.48**	**0.38**	**0.45**
TR	**0.51**	0.56	**0.47**	**0.52**	0.57	**0.51**	**0.56**	**0.47**	**0.52**	0.57
RW	**0.50**	0.50	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**
					h = 6					
$L^{\Delta sy}/S^{\Delta sy}$	0.59	0.49	0.60	0.59	0.55	0.61	0.69	0.58	0.57	0.44
IRD	0.59	0.56	0.61	0.63	0.55	**0.59**	**0.56**	0.61	0.63	0.55
PPP	**0.43**	**0.44**	**0.54**	**0.50**	**0.38**	**0.43**	**0.44**	**0.54**	**0.50**	**0.38**
MON	**0.49**	0.60	**0.58**	**0.37**	**0.41**	**0.49**	**0.60**	0.58	**0.37**	**0.41**
TR	**0.52**	0.52	**0.47**	**0.48**	0.63	**0.52**	**0.52**	**0.47**	**0.48**	0.63
RW	**0.50**	0.50	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	**0.50**	0.50

Table C.6. Direction accuracy (DA) for the time period from 1995:01 – 2014:12 (**excluding the GFC period from 2007:8-2009:5**) and h=1, h=3 and h=6 months-ahead forecasting horizons. The DA-statistic reports the proportion of forecasts that correctly predict the direction of the exchange rate movement over horizon h. The higher the proportion the better the direction accuracy. The first line reports the direction accuracy of the yield spread models $L^{\Delta sy}$ and $S^{\Delta sy}$. Direction accuracy smaller than the yield spread models and accordingly superior forecasting performance is **indicated in bold**. A value above (below) 0.5 indicates a better (worse) forecasting accuracy than a random walk. See section 5.4.1 for a detailed description of the models.

	Spread Level $L^{\Delta sy}$					Spread Slope $S^{\Delta sy}$				
	AUD	CAD	CHF	JPY	GBP	AUD	CAD	CHF	JPY	GBP
					h = 1					
$L^{\Delta sy}/S^{\Delta sy}$	1.82	0.17	1.12	7.12	2.36	0.75	0.40	1.34	5.34	0.64
IRD	0.90	**0.59**	1.07	5.48	1.98	**0.90**	**0.59**	1.07	**5.48**	**1.98**
PPP	0.79	**2.39**	**2.36**	**8.32***	1.11	**0.79**	**2.39**	**2.36**	**8.32***	**1.11**
MON	0.55	**0.61**	0.81	6.22	2.07	0.55	**0.61**	0.81	**6.22**	**2.07**
TR	1.72	**0.29**	0.24	3.28	**2.70**	**1.72**	0.29	0.24	3.28	**2.70**

Table C.7. Berkowitz test statistics for the time period from 1995:01 – 2014:12 (**excluding the GFC period from 2007:8-2009:5**), all considered currencies against the USD and h=1 months-ahead forecasting horizons. The larger the test statistic, the more likely it is that the density forecasts are not correctly specified. The first line reports the Berkowitz test statistic for the yield spread models $L^{\Delta sy}$ and $S^{\Delta sy}$. The results for the benchmark models are reported below. Test statistics larger than for the yield spread models are **indicated in bold**. Rejection of the null of uniformity of the PITs at the 5% level is indicated with (*). See Section 5.4.1 for a detailed description of all models.

	Spread Level $L^{\Delta sy}$					Spread Slope $S^{\Delta sy}$				
	AUD	CAD	CHF	JPY	GBP	AUD	CAD	CHF	JPY	GBP
					h = 3					
$L^{\Delta sy}/S^{\Delta sy}$	709	1,937	1,051	2,113	759	651	2,672	3,001	3,115	2,493
IRD	**1,214**	**1,987**	**3,019**	1,137	**1,123**	**1,227**	1,892	**3,017**	1,120	1,159
PPP	**4,036**	**6,701**	**3,026**	**7,072**	**7,216**	**4,038**	**6,758**	2,975	**6,927**	**7,133**
MON	**759**	**3,068**	326	**3,889**	**6,470**	**793**	**3,139**	309	**3,651**	**6,403**
TR	**7,004**	**7,315**	**1,714**	203	714	**7,004**	**7,315**	1,714	203	714
					h = 6					
$L^{\Delta sy}/S^{\Delta sy}$	8,682	8,272	9,864	8,505	7,482	7,811	8,774	9,639	9,933	9,994
IRD	**9,361**	8,080	**9,992**	**8,598**	**9,436**	**9,392**	8,033	**9,993**	8,671	9,436
PPP	**9,989**	**9,974**	9,824	**10,000**	**9,999**	**9,990**	**9,982**	**9,816**	**9,997**	**10,000**
MON	**9,877**	**9,583**	7,616	**9,825**	**9,998**	**9,863**	**9,571**	7,555	9,841	**9,998**
TR	**10,000**	**10,000**	**1,0000**	**9,781**	5,854	**10,000**	**10,000**	**1,0000**	**9,781**	**5,854**

Table C.8. Number of rejections at the 5% level of significance for the null hypothesis of uniformity of the PITs based on 10,000 created i.i.d. bootstrap resamples and application of the Berkowitz test (with p-values adjusted following Dowd (2007)). Tests are applied for the time period 1995:01 – 2014:12 (**excluding the GFC period from 2007:8-2009:5**) and h=3 and h=6 months-ahead forecasting horizons. The smaller the number of rejections, the more likely it is that the density forecasts of the model are correctly specified. The first line reports the number of rejections for the yield spread models $L^{\Delta sy}$ and $S^{\Delta sy}$. Results for the benchmark models with more rejections than for the yield spread models are highlighted in bold. See Section5.4.1 for a detailed description of the models.

	Spread Level $L^{\Delta sy}$					**Spread Slope** $S^{\Delta sy}$				
	AUD	CAD	CHF	JPY	GBP	AUD	CAD	CHF	JPY	GBP
					h = 1					
$L^{\Delta sy}/S^{\Delta sy}$	0.06	0.07	0.05	0.00	0.06	0.05	0.16	0.05	0.12	-0.05
IRD	**0.03**	**0.06**	**0.02**	0.02	**-0.03**	**0.03**	**0.06**	**0.02**	**0.02**	-0.03
PPP	**-0.03**	**0.03**	**-0.04**	0.05	**-0.08**	**-0.03**	**0.03**	**-0.04**	**0.05**	**-0.08**
MON	**0.03**	0.09	**-0.02**	**-0.05**	**-0.02**	**0.03**	**0.09**	**-0.02**	**-0.05**	-0.02
TR	**0.02**	**0.02**	0.12	0.11	0.08	**0.02**	**0.02**	0.12	**0.11**	0.08
RW	**-0.05**	**0.05**	0.12	0.01	**-0.03**	**-0.05**	**0.05**	0.12	**0.01**	-0.03
					h = 3					
$L^{\Delta sy}/S^{\Delta sy}$	0.15	0.12	0.10	0.02	0.03	0.15	0.24	0.02	0.11	-0.07
IRD	0.16	0.26	**0.08**	**0.02**	**0.02**	0.16	0.26	0.08	**0.02**	0.02
PPP	**-0.18**	**-0.05**	**-0.04**	0.03	**-0.14**	**-0.18**	**-0.05**	**-0.04**	**0.03**	**-0.14**
MON	**-0.01**	**0.09**	**-0.02**	**-0.16**	**-0.11**	**-0.01**	**0.09**	**-0.02**	**-0.16**	**-0.11**
TR	**0.09**	0.19	**-0.02**	0.03	0.12	**0.09**	**0.19**	**-0.02**	**0.03**	0.12
RW	**-0.08**	**0.02**	0.22	0.11	**-0.02**	**-0.08**	**0.02**	0.22	**0.11**	-0.02
					h = 6					
$L^{\Delta sy}/S^{\Delta sy}$	0.19	0.16	0.20	0.05	-0.10	0.17	0.38	0.14	0.00	-0.24
IRD	0.20	0.25	0.21	0.07	**-0.13**	0.20	**0.25**	0.21	0.07	-0.13
PPP	**-0.18**	**-0.04**	**0.07**	**0.00**	**-0.16**	**-0.18**	**-0.04**	**0.07**	**0.00**	-0.16
MON	**0.01**	**0.02**	**0.15**	**-0.20**	**-0.19**	**0.01**	**0.02**	0.15	**-0.20**	-0.19
TR	**-0.01**	**0.09**	**0.04**	**-0.06**	0.12	**-0.01**	**0.09**	**0.04**	**-0.06**	0.12
RW	**-0.19**	**0.09**	0.36	0.07	-0.04	**-0.19**	**0.09**	0.36	0.07	-0.04

Table C.9. Annualized risk adjusted returns of a monthly trading strategy based on model exchange rate predictions for the time period from 1995:01 – 2014:12 (**excluding the GFC period from 2007:8-2009:5**) and h=1, h=3 and h=6 months-ahead forecasting horizons. The higher the risk adjusted returns the better is the risk-return relation of the model's forecasts. Risk adjusted returns's smaller than for the yield spread models $L^{\Delta sy}$ and $S^{\Delta sy}$, indicate a superior forecasting performance and are reported in bold. See section 5.4.1 for a detailed description of the models and section 5.6.1 for a description of the trading strategy.

Bibliography

Abhyankar, A., Sarno, L., Valente, G., 2005. Exchange rates and fundamentals: evidence on the economic value of predictability. Journal of International Economics 66 (2), 325–348.

Afonso, A., Martins, M. M., 2012. Level, slope, curvature of the sovereign yield curve, and fiscal behaviour. Journal of Banking & Finance 36 (6), 1789–1807.

Ang, A., Chen, J., 2010. Yield Curve Predictors of Foreign Exchange Returns. AFA 2011 Denver Meetings Paper.

Ang, A., Piazzesi, M., Wei, M., 2006. What does the yield curve tell us about GDP growth? Journal of Econometrics 131 (1-2), 359–403.

Bacchetta, P., van Wincoop, E., 2013. On the unstable relationship between exchange rates and macroeconomic fundamentals. Journal of International Economics 91 (1), 18–26.

Balke, N. S., Ma, J., Wohar, M. E., 2013. The contribution of economic fundamentals to movements in exchange rates. Journal of International Economics 90 (1), 1–16.

Berkowitz, J., 2001. Testing Density Forecasts, with Application for Risk Management. Journal of Business & Economic Statistics 19 (4), 465–474.

Bjørnland, H. C., Hungnes, H., 2006. The importance of interest rates for forecasting the exchange rate. Journal of Forecasting 25 (3), 209–221.

Bui, A. T., Fisher, L. A., 2016. The relative term structure and the Australian–US exchange rate. Studies in Economics and Finance (Forthcoming).

Cenedese, G., Payne, R., Sarno, L., Valente, G., 2015. What Do Stock Markets Tell Us about Exchange Rates? Review of Finance , published online, July 31, 2015.

Chen, Y.-C., Gwati, R., 2014. FX Options and Excess Returns: A Multi Moment Term Structure Model of Exchange Rate Dynamics. Working Paper.

Chen, Y.-C., Tsang, K. P., 2013. What Does the Yield Curve Tell Us about Exchange Rate Predictability? Review of Economics and Statistics 95 (1), 185–205.

Cheung, Y.-W., Chinn, M. D., Pascual, A. G., 2005. Empirical exchange rate models of the nineties: Are any fit to survive? Journal of International Money and Finance 24 (7), 1150–1175.

Christiansen, C., 2011. Intertemporal risk-return trade-off in foreign exchange rates. Journal of International Financial Markets, Institutions and Money 21 (4), 535–549.

Clark, T. E., McCracken, M. W., 2013. Advances in Forecast Evaluation. In: Elliott, G., Granger, C. W. J., Timmermann, A. (Eds.), Handbook of economic forecasting. Vol. 2. Elsevier, Amsterdam and Boston, pp. 1107–1201.

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

Corte, P. D., Sarno, L., Tsiakas, I., 2009. An Economic Evaluation of Empirical Exchange Rate Models. Review of Financial Studies 22 (9), 3491–3530.

Dewachter, H., Lyrio, M., 2006. Macro Factors and the Term Structure of Interest Rates. Journal of Money, Credit and Banking 38 (1), 119–140.

Diebold, F. X., Hahn, J., Tay, A. S., 1999. Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High-

Frequency Returns on Foreign Exchange. Review of Economics and Statistics 81 (4), p 661–673.

Diebold, F. X., Li, C., 2006. Forecasting the term structure of government bond yields. Journal of Econometrics 130 (2), 337–364.

Diebold, F. X., Rudebusch, G. D., Aruoba, S. B., 2006. The macroeconomy and the yield curve: a dynamic latent factor approach. Journal of Econometrics 131 (1-2), 309–338.

Dowd, K., 2007. Validating multiple-period density-forecasting models. Journal of Forecasting 26 (4), 251–270.

Engel, C., 2013. Exchange Rates and Interest Parity. NBER Working Paper 19336.

Engel, C., Mark, N. C., West, K. D., 2007. Exchange Rate Models Are Not as Bad as You Think. NBER Working Paper 13318.

Engel, C., Nelson, M. C., Kenneth, W. D., 2012. Factor Model Forecasts of Exchange Rates. NBER Working Paper 18382.

Engel, C., West, K. D., 2005. Exchange Rates and Fundamentals. Journal of Political Economy 113 (3), 485–517.

Erdogan, O., Bennett, P., Ozyildirim, C., 2015. Recession Prediction Using Yield Curve and Stock Market Liquidity Deviation Measures. Review of Finance 19 (1), 407–422.

Estrella, A., Mishkin, F. S., 1998. Predicting U.S. Recessions: Financial Variables as Leading Indicators. Review of Economics and Statistics 80 (1), 45–61.

Evans, M. D. D., Lyons, R. K., 2007. Exchange Rate Fundamentals and Order Flow. NBER Working Paper 13151.

Favero, C. A., Niu, L., Sala, L., 2012. Term Structure Forecasting: No-Arbitrage Restrictions versus Large Information Set. Journal of Forecasting 31 (2), 124–156.

Ferraro, D., Rogoff, K., Rossi, B., 2015. Can oil prices forecast exchange rates? An empirical analysis of the relationship between commodity prices and exchange rates. Journal of International Money and Finance 54, 116–141.

Fratzscher, M., 2009. What explains global exchange rate movements during the financial crisis? Journal of International Money and Finance 28 (8), 1390–1407.

Frenkel, J. A. (Ed.), 1983. Exchange rates and international macroeconomics. University of Chicago Press, Chicago.

Giacomini, R., Rossi, B., 2010. Forecast comparisons in unstable environments. Journal of Applied Econometrics 25 (4), 595–620.

Guidolin, M., Tam, Y. M., 2013. A yield spread perspective on the great financial crisis: Break-point test evidence. International Review of Financial Analysis 26, 18–39.

Guo, H., Savickas, R., 2008. Forecasting foreign exchange rates using idiosyncratic volatility. Journal of Banking & Finance 32 (7), 1322–1332.

Hong, Y., Li, H., Zhao, F., 2007. Can the random walk model be beaten in out-of-sample density forecasts? Evidence from intraday foreign exchange rates. Journal of Econometrics 141 (2), 736–776.

Inoue, A., Rossi, B., 2008. Monitoring and Forecasting Currency Crises. Journal of Money, Credit and Banking 40 (2-3), 523–534.

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

Mark, N. C., 1995. Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability. American Economic Review 85 (1), 201–218.

Meese, R. A., Rogoff, K., 1983. The out-of-sample failure of empirical exchange rate models: sampling error or misspecification? In: Frenkel, J. A.

(Ed.), Exchange rates and international macroeconomics. University of Chicago Press, Chicago, pp. 67–112.

Moench, E., 2012. Term structure surprises: the predictive content of curvature, level, and slope. Journal of Applied Econometrics 27 (4), 574–602.

Molodtsova, T., Papell, D. H., 2009. Out-of-sample exchange rate predictability with Taylor rule fundamentals. Journal of International Economics 77 (2), 167–180.

Molodtsova, T., Papell, D. H., 2012. Taylor Rule Exchange Rate Forecasting During the Financial Crisis. NBER Working Paper 18330.

Moosa, I. A., 2013. Why is it so difficult to outperform the random walk in exchange rate forecasting? Applied Economics 45 (23), 3340–3346.

Moosa, I. A., Burns, K., 2014. The unbeatable random walk in exchange rate forecasting: Reality or myth? Journal of Macroeconomics 40, 69–81.

Rapach, D. E., Wohar, M. E., 2006. The out-of-sample forecasting performance of nonlinear models of real exchange rate behavior. International Journal of Forecasting 22 (2), 341–361.

Rime, D., Sarno, L., Sojli, E., 2010. Exchange rate forecasting, order flow and macroeconomic information. Journal of International Economics 80 (1), 72–88.

Rosenblatt, M., 1952. Remarks on a Multivariate Transformation. The Annals of Mathematical Statistics 23 (3), 470–472.

Rossi, B., 2013. Exchange Rate Predictability. Journal of Economic Literature 51 (4), 1063–1119.

Rossi, B., 2015. Density Forecasts in Economics, Forecasting and Policymaking. Working Paper.

Rudebusch, G. D., Wu, T., 2007. Accounting for a Shift in Term Structure Behavior with No-Arbitrage and Macro-Finance Models. Journal of Money, Credit and Banking 39 (2-3), 395–422.

Rudebusch, G. D., Wu, T., 2008. A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy. Economic Journal 118 (530), 906–926.

Sarno, L., 2005. Viewpoint: Towards a solution to the puzzles in exchange rate economics: where do we stand? Canadian Journal of Economics 38 (3), 673–708.

Sarno, L., Valente, G., Leon, H., 2006. Nonlinearity in Deviations from Uncovered Interest Parity: An Explanation of the Forward Bias Puzzle. Review of Finance 10 (3), 443–482.

Stock, J. H., Watson, M. W., 2003. Forecasting output and inflation: the role of asset prices. Journal of Economic Literature 41 (3), 788–829.

Wang, J., Wu, J., 2012. The Taylor Rule and Forecast Intervals for Exchange Rates. Journal of Money, Credit and Banking 44 (1), 103–144.

Wang, Y.-C., Wu, J.-L., 2015. Fundamentals and Exchange Rate Prediction Revisited. Journal of Money, Credit and Banking 47 (8), 1651–1671.

6 Thesis Summary and Conclusions

This PhD thesis has investigated the predictive content entailed in the term structure of interest rates and interest rate differentials. In doing so, it has applied in-sample as well as out-of-sample forecasting methods to investigate major aspects of academic forecasting and provide a multifaceted analysis of the predictive power of the term structure assessed by several evaluation metrics. A particular emphasis of this analysis has been put on the impact of the global financial crisis of 2007-2009 (GFC).
This section aims at summarizing the main results of the three research papers and pointing out the major contributions this thesis has made to the overall discipline.

6.1 Main Results

The research paper '*Forecasting the Term Structure of Interest Rates near the Zero Bound - a New Era?*' (Chapter 3) provides a pioneer study in documenting the challenge which the current low interest rate environment following the GFC poses to popular dynamic factor yield curve forecasting models.
In order to examine the forecasting accuracy during this unique time period, we focus on variations of the parametric dynamic Nelson-Siegel model and regressions on principal components. We use a sample of US zero yields from January 1995 to December 2013 and apply sub-sample analysis and dynamic forecasting evaluation measures.
Our results for the pre-crisis and crisis periods are in line with findings from previous yield curve forecasting studies. The selected factor models perform relatively well for short term maturities, but all models fail to consistently beat the random walk. RMSEs are generally smaller for longer term maturi-

ties and the forecasting performance worsens with longer forecasting horizons. However, the forecasting accuracy of the applied dynamic factor models worsens dramatically after the GFC with short rates close to the zero bound. Dynamic forecasting metrics and sub-sample analysis of the forecasted time series reveal that the investigated dynamic factor models not only fail to beat the random walk but are notably outperformed in relative terms, with the RMSEs being up to ten times higher.

We identify several potential reasons for this behavior: first, we suspect that the cross-sectional structure of the dynamic factor models, which also includes additional information of other maturities, worsens the forecasting accuracy as the short end of the yield curve becomes more segmented from the rest of the curve. Further, the models were also estimated during periods when interest rates were significantly higher than during the post GFC period, so that forecasts created by the applied models will not only overstate the dynamics of the interest rate term structure, but possibly also interest rate levels

An additional, important aspect of our findings is that these results are not detected by traditional global forecasting metrics, such as the RMSE, which average the results over the entire forecasting period. As the forecasting errors for short and medium yields in the unique period after the GFC are relatively small in absolute terms, investigating only the global forecasting performance may fail to detect important information about the relative forecasting performance of competing models through time and may lead to entirely different conclusions about the predictive accuracy of econometric models. A thorough sub-sample analysis as well as dynamic forecasting measures are therefore crucial to truly expose a model's predictive abilities.

Finally, we identify forecast combination strategies as a promising mitigating approach. Simply combining all factor models already notably mitigates the inferior performance relative to the random walk. Strategically combining forecasts from one Nelson-Siegel variation, one PCA model and an AR(1) model is able to further improve the results. We also find that the forecasting accuracy generally increases when the combined forecasts are weighed based on the recent forecasting performance.

The research work titled '*Factors of the Term Structure of Sovereign Yield Spreads*' (Chapter 4) provides a novel analysis of the term structure of sovereign yield spreads. Sovereign yield spreads are the difference between two sovereign bond yields of equal maturity. We focus on the yield spreads of advanced economies with highly liquid markets of bonds issued in their own currency, little to no credit or default risk and free floating exchange rates (Australia, Canada, Switzerland, Japan, UK and US) for the time period from January 1995 to December 2013.

Our first objective is to derive and examine the latent factors driving the term structure of sovereign yield spreads, which, to the best of our knowledge, has not been thoroughly studied in the literature yet. We thus apply principal component analysis (PCA) on each of the five sovereign spread data sets. Our analysis shows that the term structure of all sovereign spreads is driven by similar factors and the first three estimated factors are sufficient to explain more than 99% of the variation in the entire spread term structure. Interestingly, the identified factors show a very similar shape to those reported in studies analysing the term structure of interest rates, see, e.g., Litterman and Scheinkman (1991); Bikbov and Chernov (2010), and can be related to the empirical spread level, spread slope and spread curvature.

We then proceed with investigating the predictive power of these factors for exchange rates and find that the extracted yield spread factors can explain and predict bilateral exchange rate movements and excess returns three months to two years ahead. The spread level and spread slope factor seem to be the most dominant. The negative signs of the predictive regression coefficients on these factors indicate that an increase in the spread level or spread slope factor, i.e. when the foreign yield curve shifts down or becomes steeper relative to the US, predicts a depreciation and smaller excess returns of the foreign currency against the US dollar. This is consistent with economic intuition and findings in the previous literature (Chen and Tsang, 2013; Bui and Fisher, 2016). Considering that exchange rate determination is traditionally based on macroeconomic fundamentals such as differences in price levels, output and monetary policy (Engel and West, 2005; Molodtsova and

Papell, 2009), we conclude that the estimated factors naturally summarize information about the same fundamentals reflected in the term structure of sovereign yield spreads and thus seem to proxy fundamental aspects of exchange rate determination.
When we compare the explanatory power of the extracted spread factors to the traditional UIRP approach, we find that the three factors provide statistically significant, additional predictive power for most currencies and horizons. We therefore infer that using the information of the entire spread curve summarized in the spread factors adds valuable information in particular about expected future fundamentals. These findings confirm the view that the exchange rate can be understood as an asset price and thus relies more on future than on current fundamentals.

The research work titled '*Exchange Rates and Unobservable Fundamentals: A New Approach to Out-of-Sample Forecasting*' (Chapter 5) is based on two main conclusions of Chapter 4. First, latent factors derived from the term structure of sovereign spreads – the difference between two economies' yield curves – have predictive power for exchange rates. Second, these factors are closely associated with the empirical sovereign spread level and slope. Based on these insights, the third research paper investigates whether the empirical yield spread level and slope can forecast exchange rates out-of sample when applied as unobservable macroeconomic fundamentals. Using the empirical factors is preferable to the estimated, latent factors in out-of-sample forecasting as it is notably less computationally extensive and straightforward to implement.
Our approach is favorable to traditional exchange rate models, based on observable macroeconomic fundamentals, such as differences in output, price levels or monetary aggregates, for several reasons. First, our approach is market based, as the expectations about future economic fundamentals reflect the views of a large number of market participants in highly liquid sovereign bond markets. Yield data is also readily and easily available on a daily basis as opposed to monthly and quarterly macroeconomic data, which is often published with a considerable time lag and revised in hindsight. Sec-

ond, previous research has shown that simple univariate specifications often deliver accurate forecasts in out-of sample forecasting studies (Clark and McCracken, 2013). Using a small number of variables which reflect a broad range of unobservable macroeconomic information and business conditions thus allows for parsimonious modelling and is more flexible to changes in business conditions over time and less vulnerable to the omitted variables bias. Finally, our parsimonious model is straightforward to implement and therefore an appealing approach for investment practitioners.

To assess the forecasting capability of our approach we conduct an extensive out-of-sample forecasting exercise and also investigate the trading profitability against the random walk and several traditional fundamental exchange rate models. In particular we consider models based on differences between interest rates, price levels, monetary aggregates and Taylor rule fundamentals. We again focus the major currencies of advanced countries with free floating exchange rates and highly liquid bond markets and little to no credit risk (Australian Dollar, the Canadian Dollar, the Swiss Franc, the Japanese Yen and the British Pound) against the US Dollar.

We assess the forecasting accuracy based on several different forecasting evaluation methods. Previous research and the findings from Chapter 3 have shown that the sole focus on the traditional RMSE metric may not be entirely appropriate (Cheung et al., 2005; Moosa and Burns, 2014). In addition to the RMSE, we therefore also apply a measure of direction accuracy and assess the uncertainty around the forecasts based on their density. We further implement a period-by-period trading strategy to assess the profitability of the forecasts.

We find promising forecasting results for our yield spread approach compared to traditional fundamental models. It is decisively superior in terms of the RMSE and direction accuracy and provides less uncertainty of the forecasts measured by its forecasting density. While it fails to beat the random walk in terms of the RMSE – which should hardly be surprising given the previous findings in the literature – it is also superior to a random walk in forecasting the direction of exchange rate changes.

In terms of trading profitability, we find positive returns for our approach for the majority of currencies and forecasting horizons. Compared to the fundamental models we also find consistently higher risk-adjusted profits.

6.2 Overall Contributions to the Discipline

The contributions of this thesis to the discipline can be assigned to the following research areas:

(1) **Term Structure Literature**
This thesis provides new and innovative insights into the predictive power that is contained in the term structure of interest rates and interest rate differentials.
It provides the first study to thoroughly explore the dynamics of the term structure of yield spreads for advanced economies. While spreads between certain maturities are subject to the enormous body of UIRP-literature, the term structure of yield spreads has been widely neglected so far. We have identified and successfully labeled the latent driving forces of the spread term structure and concluded that these factors summarize information about current and future macroeconomic differentials between economies.
This thesis further confirms that the macroeconomic information contained in the term structure of sovereign yield spreads has predictive power for exchange rates in-sample and is superior to traditional fundamental exchange models out-of-sample. The predictive power is also superior to an UIRP forecasting approach based on the information of maturities up to a certain horizon.
However, this thesis also reveals the poor performance of popular yield curve forecasting models which exploit the cross-sectional information of the term structure, when different maturities become more segmented and volatile in times of market turbulence. In the period after the GFC, this has a particularly strong impact on the forecasting accuracy for short- and medium term

yields.

(2) **Academic Forecasting Literature**

This thesis further makes several contributions to the academic forecasting literature.

To start with, we present an innovative, parsimonious, market driven approach to exchange rate forecasting based on readily and easily available data. This makes it a promising proposition in particular for market practitioners.

Another major contribution of this thesis is to demonstrate how sensitive inferences drawn from out-of-sample forecasts of financial time series are to the choice of forecasting evaluation metrics. This thesis shows that the choice and application of evaluation methods in forecasting financial time series does have a crucial impact on the conclusions. As different forecasting metrics may led to contrasting conclusions, several different forecasting evaluation metrics are required to provide a multidimensional and comprehensive picture of a model's forecasting accuracy. Furthermore, in the presence of time-varying dynamics, just minimizing the loss function of the forecasting errors over the entire sample will result in a significant loss of information. A thorough sub-sample analysis as well as dynamic forecasting measures are thus crucial to truly expose a model's predictive abilities.

Finally, this thesis identifies forecast combination strategies as a potential approach to mitigate poor forecasting performances in terms of market turbulence. While the forecasting accuracy of individual models varies heavily over time, combined forecasts are less affected by structural instability than either of the individual models.

(3) **Macro-Finance Literature**

With regards to the Macro-Finance literature, this thesis provides new evidence of the link between the term structure of interest rates, macroeconomic fundamentals and exchange rates. It shows that including the fundamental information embodied in the latent factors of sovereign yield spreads is a promising approach to forecast exchange rates and thus argues that these factors proxy current and future values of marcoeconomic fundamentals tra-

ditionally applied in exchange rate determination.
As standard empirical approaches commonly reduce the sum of expected future fundamentals to equal current fundamentals, these results also suggest that the recurrent problems faced by traditional fundamental exchange rate models in empirical forecasting are most likely due to overly restrictive assumptions, which fail to properly account for the forward-looking property of exchange rates.
This thesis thus provides further evidence that financial variables are important input factors for exchange rate prediction. As they are naturally forward looking and susceptible to the same macroeconomic risk as exchange rates, models based on financial variables are an innovative, promising approach to forecast exchange rates.

(4) **GFC Literature**
Finally, this thesis provides new insights into the impact of the GFC on financial markets and the performance of forecasting models traditionally applied to these markets.
To begin with, it thoroughly analyses and describes the impact the GFC has had on sovereign bond and foreign exchange markets during and after the GFC. While the crises has led to an unprecedented, prolonged period of historically abnormally low interest rates, it has also triggered sharp and unexpected currency movements, with significantly higher exchange rate volatility during the crises period.
This thesis further systematically documents and explains the poor forecasting performance for medium and short term US yields associated with the popular class of dynamic factor yield curve models in the low interest rate environment following the GFC. It also shows that the GFC's impact on the forecasting accuracy of exchange rate models is rather limited for point forecasts but significantly increases the uncertainty of these forecasts.

6.3 Directions for Future Research

The findings and contributions of this thesis naturally suggest a number of directions for future work in these areas.

First of all, more research is required to fully understand the dynamics of the term structure of yield spreads and its relation to macroeconomic fundamentals and exchange rates. While this thesis provides a pioneer analysis of the dynamics and an intuitive interpretation of the latent factors driving the spread term structure, further analysis may be helpful to fully understand the behavior of sovereign spreads curves for other time periods and countries, in particular developing economies. Furthermore, it may be valuable to tie the spread factors to specific macroeconomic variables.

Second, as this thesis provides further evidence that naturally forward looking financial variables are important input factors to be considered in exchange rate forecasting, it may be worthwhile to combine our yield spread approach with traditional fundamental exchange rate models or other financial variables, such as stock returns. Different financial variables may reflect other aspects of the business cycle and exchange rate determination, which may further improve the forecasting accuracy.

Another major finding of this thesis is the importance of applying several different forecasting evaluation measures to fully assess a model's forecasting accuracy. While numerous different evaluation metrics have been developed in the literature (see Clark and McCracken (2013) for a recent overview), many empirical forecasting studies still solely rely on one traditional evaluation metric, mostly the RMSE (see for example Bjørnland and Hungnes (2006) or Molodtsova and Papell (2012)). Future empirical forecasting studies for financial and macroeconomic time series should thus consider several different appropriate forecasting evaluation measures an carefully assess the dynamic forecasting accuracy throughout time.

The results of this thesis also point towards the potential benefit of utilizing the predictive power of several models instead of simply relying on one individual model. Besides forecast combinations – which have been applied in this thesis – adaptive forecasting techniques or regime switching models are other promising approaches to improve the forecasting accuracy for financial variables, especially in times of crises.

Finally, this thesis has also given an indication of the GFC's impact on financial markets, in particular bond and foreign exchange markets and the models applied within. While there is a growing body of "crisis literature" (see e.g. Guidolin and Tam (2013) or Contessi et al. (2014)), further research is required to fully understand the impact of the GFC on financial markets and models. One major takeaway of this thesis for future research is that studies conducted with financial time series encompassing the crises period have to carefully consider the crisis impact on results and inferences.

Bibliography

Bikbov, R., Chernov, M., 2010. Yield Curve and Volatility: Lessons from Eurodollar Futures and Options. Journal of Financial Econometrics 9 (1), 66–105.

Bjørnland, H. C., Hungnes, H., 2006. The importance of interest rates for forecasting the exchange rate. Journal of Forecasting 25 (3), 209–221.

Bui, A. T., Fisher, L. A., 2016. The relative term structure and the Australian–US exchange rate. Studies in Economics and Finance (Forthcoming).

Chen, Y.-C., Tsang, K. P., 2013. What Does the Yield Curve Tell Us about Exchange Rate Predictability? Review of Economics and Statistics 95 (1), 185–205.

Cheung, Y.-W., Chinn, M. D., Pascual, A. G., 2005. Empirical exchange rate models of the nineties: Are any fit to survive? Journal of International Money and Finance 24 (7), 1150–1175.

Clark, T. E., McCracken, M. W., 2013. Advances in Forecast Evaluation. In: Elliott, G., Granger, C. W. J., Timmermann, A. (Eds.), Handbook of economic forecasting. Vol. 2. Elsevier, Amsterdam and Boston, pp. 1107–1201.

Contessi, S., Pace, P. d., Guidolin, M., 2014. How did the financial crisis alter the correlations of U.S. yield spreads? Journal of Empirical Finance 28, 362–385.

Engel, C., West, K. D., 2005. Exchange Rates and Fundamentals. Journal of Political Economy 113 (3), 485–517.

Guidolin, M., Tam, Y. M., 2013. A yield spread perspective on the great financial crisis: Break-point test evidence. International Review of Financial Analysis 26, 18–39.

Litterman, R., Scheinkman, J., 1991. Common Factors Affecting Bond Returns. Journal of Fixed Income 1 (1), 54–61.

Molodtsova, T., Papell, D. H., 2009. Out-of-sample exchange rate predictability with Taylor rule fundamentals. Journal of International Economics 77 (2), 167–180.

Molodtsova, T., Papell, D. H., 2012. Taylor Rule Exchange Rate Forecasting During the Financial Crisis. NBER Working Paper 18330.

Moosa, I. A., Burns, K., 2014. The unbeatable random walk in exchange rate forecasting: Reality or myth? Journal of Macroeconomics 40, 69–81.