Erasmus University Rotterdam, Erasmus School of Economics
Econometrics & Management Science
Quantitative Finance

# The Stochastic Factor-Augmented Nelson-Siegel Model

A comparison between the extended and unscented Kalman filter

Master's Thesis

Author:
Justin A.L. Djodikromo
368090

Supervisor:
Prof. dr. Dick J.C. van Dijk

Co-reader:
Dr. Rutger-Jan Lange

**Abstract**

Accurately modeling and forecasting the term structure of interest rates is relevant for both academics and practitioners in the industry. The dynamic Nelson-Siegel model is suitable for this. Although the amount of research on the in-sample fit of the Nelson-Siegel model and its extensions is substantial, the number of studies examining the out-of-sample performance of these models is relatively little, particularly for the nonlinear class of Nelson-Siegel models. For this reason, the focus of this thesis is twofold. First, I examine the predictive performance of various extensions in the Nelson-Siegel framework relative to the standard dynamic Nelson-Siegel model. Second, I study the differences between the extended Kalman filter and the unscented Kalman filter in the context of nonlinear Nelson-Siegel models. I find that the results are maturity- and subsample-dependent. The greatest gain in predictive accuracy is found for the stochastic factor augmented Nelson-Siegel model, for which the improvement over the standard model attains values of a 28% decrease in the RMSPE when the yields are relatively volatile. Furthermore, the findings indicate that the use of the unscented Kalman filter rather than the extended Kalman filter for fitting nonlinear Nelson-Siegel models is beneficial for both ends of the yield curve and could have a positive impact on the accuracy in the predictive framework in some cases.

*Keywords*: Dynamic Nelson-Siegel Model, Data-Rich Environment, GARCH, State-Space Model, Kalman Filter

Rotterdam, October 2016

# Contents

# 1 Introduction

The focus of this study is on modeling and forecasting the term structure of interest rates. The term structure of interest rates, also known as the yield curve, denotes the relation between interest rates (or bond yields) and different terms (or maturities). The term structure reflects expectations of market participants about future changes in interest rates and their assessment of monetary policy conditions.

Understanding the dynamic evolution of the yield curve is important in theory and practice. For example, it is essential for conducting monetary policy as the yield curve could be used to predict future inflation (Fama, 1990; Mishkin, 1990), recession periods (Estrella and Mishkin, 1997) and real activity of the economy (Estrella and Hardouvelis, 1991). In addition, it is relevant for asset managers, traders and risk managers in the fixed-income markets as they use it to price financial assets and their derivatives, manage financial risk, hedge and allocate portfolios. Improvements in yield curve modeling translate into improvements in these tasks, and is therefore a huge area of interest in both academic and financial institutions.

The literature has produced a wide variety of models to study the dynamics of the yield curve. Among them, dynamic factor models, which trace to Sargent and Sims (1977), Geweke (1977) as well as Watson and Engle (1983), are found to be appealing for three key reasons. First, the structure performs empirically well as it describes the yield curve data accurately. Second, factor models are parsimonious as they are able to convert seemingly intractable high-dimensional situations (in this case the many yields across maturities) into a few constructed variables or factors. This is often related to the production of good out-of-sample forecasts, see Diebold (2007) for additional discussion. Lastly, financial economic theory suggests factor structure. For example, the capital asset pricing model (CAPM) is a factor model that is able to describe the risk premiums of thousands of financial assets in the equity market using only a single factor. Yield curve factor models are a natural bond market parallel.

There are several methods on constructing bond yield factors and factor loadings. One approach restricts both factors and factor loadings. We refer to this model as the no-arbitrage dynamic latent factor model, which is the model of choice in finance. The papers of Vasicek (1977), Cox et al. (1985) as well as Hull and White (1990), among others, lay the foundation for a vast literature on arbitrage-free models. Those models specify the dynamics of the yield curve factors using a risk-neutral measure. The affine versions of these models, as proposed by Duffie and Kan (1996), are particularly popular. Even though these affine term structure models are theoretically well-founded, in the sense that they impose restrictions that ensure absence of arbitrage, their forecasting performance is poor relative to the random walk (Duffee, 2002). In addition, and crucially, the estimation of those models is known to be problematic, in large part because of the existence of numerous likelihood maxima that have essentially identical fit to the data but very different implications for economic behavior.

A second approach places a specific structure only on the loadings. A popular example of this approach is the so-called Nelson-Siegel model as introduced in Nelson and Siegel (1987). This model was originally designed to describe the cross-sectional aspects of the yield curve by imposing a parsimonious three-factor structure on the links between yields of different maturities. For this model, the structure of the factor loadings in the DNS depend on a single loading parameter $\lambda$. This ensures the interpretation of the factors as level, slope and curvature. Diebold and Li (2006) introduce a suitable dynamic reformulation of this model (the dynamic Nelson-Siegel model), in which the latent factors are modelled by a stationary autoregressive process with a constant unconditional mean. They show that this model is simple to estimate, is quite flexible, fits both the cross-section and time series of yields remarkably well and provides forecasts that outperform the random walk and various alternative forecasting approaches. However, despite its good empirical performance, the dynamic Nelson-Siegel (DNS) model fails on an important theoretical property. That is, the model is not designed to rule out opportunities for riskless arbitrage which are important in financial markets (e.g. Filipovic, 1999; Diebold et al., 2005). This motivated Christensen, Diebold and Rudebusch (2011) to introduce the arbitrage-free Nelson-Siegel yield curve model. This class of model maintains the simplicity and empirical behaviour of the dynamic Nelson-Siegel model, while simultaneously enforcing the theoretically desirable property of absence of riskless arbitrage. However, Diebold and Rudebusch (2013) state that even though it may feel uncomfortable to work with non-'arbitrage-free' models for some tasks (e.g. pricing), the imposition of no-arbitrage appears theoretically unlikely to help at forecasting. Hence, as I put more emphasis on forecasting in this thesis, I consider the standard DNS to be the main model in this paper.

Drawing upon the great performance of the DNS, many studies introduce and examine extensions of this model. I discuss the main extensions briefly. Several extensions originated from the idea of including additional latent factors. Among such models is the Svensson (1995) extension of the DNS, which is widely used in industry and central banks.[i] In the dynamic Nelson-Siegel-Svensson model, a second curvature variable with a longer-maturity hump is added to obtain a better fit at longer maturities. Similarly, De Pooter (2007) adds a second slope factor to the DNS and finds that his four factor model provides more accurate forecasts and a better in-sample fit than the Diebold and Li (2006) three factor model. Instead of including additional latent factors, Diebold et al. (2006) write the DNS model in a state-space representation form, which consists of a measurement and a transition equation, and add three observable macroeconomic variables (specifically, real activity, inflation, and a monetary policy instrument) to the model. They find strong evidence of bidirectional causality from the latent factors to the macroeconomic variables and vice-versa, with a stronger causal direction from the macroeconomic variables to the yield curve. To estimate the parameters in the state-space representation of their models, the Kalman filter is used. This estimation method relies on an important assumption, i.e. the state-space model should be linear. This is true for the standard DNS as the loading parameter $\lambda$ is kept constant over time, resulting in constant factor loadings, for each maturity. Koopman et al. (2010) relax this and allow $\lambda$ to vary over time. As a result, the factor loadings vary over time and the standard DNS model becomes nonlinear, such that the standard Kalman filter is not applicable. Koopman et al. (2010) analyze models of this kind using the extended Kalman filter and find that a time-varying loading parameter in the DNS leads to a significant improvement of the in-sample fit. Concerning the addition of time-varying elements in the DNS, Hautsch and Ou (2008), Hautsch and Yang (2010) as well as Koopman et al. (2010) incorporate stochastic volatility in either the transition or measurement shocks. More recently, van Dijk et al. (2013) allow the unconditional mean to be time-varying and introduce the dynamic Nelson-Siegel model with so-called shifting endpoints. They find that this extension can provide gains in the predictive accuracy. Even though the amount of papers on the effect of these extensions on the in-sample fit is substantial, the number of studies on the out-of-sample forecasting performance is relatively scarce. In this study, therefore, I consider various extensions of the DNS model and analyze their performance in an out-of-sample setting using the Fama-Bliss zero-coupon yields dataset at the monthly frequency covering the

[i]For example, the U.S. Federal Reserve Board (see Gurkaynak et al, 2007) and the European Central Bank (see Coronea et al. (2011), among others.

period January 1970 to December 2009. Before I go into further details of the study, I will first elaborate on the extensions that I consider and the corresponding subquestions that will help in answering my main question. Moreover, the extensions that I consider are the inclusion of macro-economic variables, stochastic volatility and a time-varying loading parameter.

First, I consider the inclusion of observable macroeconomic variables in the DNS model for the out-of-sample framework. As mentioned, Diebold et al. (2006) include macroeconomic variables in the DNS (analogous to the inclusion of macroeconomic factors in the affine arbitrage-free model as introduced by And and Piazzesi, 2003). Based purely on the in-sample results, Diebold et al. (2006) find that the level factor is highly correlated with inflation and the slope factor is highly correlated with real activity. Moreover, the curvature factor appears to be unrelated to any of the main macroeconomic variables. More studies that include macroeconomic variables and study their explanatory power for yield movements are Kim and Wright (2005), Ang et al. (2006a), Dai and Philippon (2006), Hordahl et al. (2006), DeWachter and Lyrio (2006), Rudebusch and Wu (2007), Wu (2008), and Bikbov and Chernov (2010), among others. The natural questions that arise are which macro factors to include in the model and how many factors to include. Arguably, many macro variables may influence the dynamics of the yield curve. De Pooter et al. (2010) use principal components analysis (PCA) to extract a small number of factors and find that the inclusion of these factors lead to an improvement in forecast accuracy, compared to the use of single macroeconomic variables. Exterkate et al. (2013) use a wide variety of variable selection and dimension reduction techniques to extract macroeconomic information from an extended version of the Stock and Watson (2002) dataset. As a contribution to the literature, I use a different dataset; obtained from the database of the Federal Reserve Bank of St. Louis, which is updated on monthly basis, covering the same period as the yields. Moreover, I apply PCA on the set of 135 macroeconomic variables, to construct the factors that I include.

Another extension of the DNS on which I focus in this paper is the inclusion of stochastic volatility. In the state-space formulation of the DNS one can allow for time-varying volatility through the measurement shocks or the transition shocks, and different possibilities have been considered by different authors. Hautsch and Ou (2008) and Hautsch and Yang (2010) incorporate time-varying volatility in the transition equation. They find strong evidence of stochastic volatility in the transition shocks, and they show that accounting for it improves the conditional calibration of interval and density forecasts. Alternatively, Koopman et al. (2010) allow for stochastic volatility in the measurement shocks using the single-factor multivariate GARCH model of Harvey et al. (1992), and find that this leads to a significant improvement of the model fit. It would be interesting to allow for stochastic volatility in both the measurement and transition equation, but, Diebold and Rudebusch (2013) state that identification issues arise in this case. The out-of-sample performance of the dynamic Nelson-Siegel model with stochastic volatility incorporated in the transition shocks has been examined extensively by Hautsh and his co-authors. In contrast, Koopman et al. (2010) do not study their model in an out-of-sample setting. For this reason, I follow their approach and examine its predictive performance.

Finally, I examine the extension in which the loading parameter is time-varying. As an alternative to the use of the extended Kalman filter as in Koopman et al. (2010), the unscented Kalman filter could be used. Julier and Uhlmann (1997) proposed this filtering technique and found substantial performance gains against the extended Kalman filter in the context of state-estimation for nonlinear control. A lot of research, comparing these two methods, has been conducted ever since. Worth mentioning, Wan and van der Merwe provide a wide variety of papers in which they compare the unscented Kalman filter and the extended Kalman filter. They find that the unscented Kalman filter consistently achieves a better level of accuracy than the extended Kalman filter in a number of application domains, including state-space estimation, dual estimation and parameter estimation. However, its effectiveness for improving the in-sample fit and accuracy of the forecasts in the context of Nelson-Siegel models with a time-varying loading parameter has, to my knowledge, been unexplored. In this paper, I will study and compare the results from both filtering techniques to answer the question of whether the unscented Kalman filter is superior to the extended Kalman filter in the framework of modeling and forecasting the term structure of interest rates.

Summarizing, the two main questions in this research are as follows: (I) What is the best filtering technique, in the context of nonlinear Nelson-Siegel models, for modeling and forecasting the term structure of interest rates? And (II) among the models that I consider, what is the best Nelson-Siegel model in the forecasting framework? I answer these questions by examining how the forecasting performance of the different models vary across the yield curve relative to the standard DNS (e.g. do some models perform well for short-term yields and much worse for long-maturities) and how the relative forecasting performance change over time, particularly for stable and volatile periods.

By evaluating the root mean squared prediction error (RMSPE) of the DNS extensions relative to the standard DNS for forecast horizons of one, six and twelve months ahead, I find that the results are maturity- and subsample-dependent. First, for the full out-of-sample period the standard DNS is in general adequate in forecasting the yields. At the semiannually and annually forecast horizon it is even unbeatable. However, when the subsamples are considered the results are more promising. Moreover, when the yields are stable (as in January 1995 until December 1999), the GARCH extension of the DNS performs remarkably better than the DNS for all forecast horizons. This result corresponds to the finding Hautsch and Yang (2010), who incorporate stochastic volatility via the transition equation instead. They argue that the standard DNS has a higher forecasting uncertainty for periods of low-volatility, which stems from the fact that the ignored stochastic volatility and parameter uncertainty in periods of high-volatility spread to periods when the yields are more stable. For periods of relatively highly volatile yields, the most flexible model provides the most accurate forecasts, i.e. the stochastic factor-augmented Nelson-Siegel model (the combination of the mentioned models). By using this model, the gain in predictive accuracy reaches up to 28%. To understand why this model is able to attain such values, I focus on forecast combinations as well. Using three weighting schemes (equal weights, time-varying weights based on the cross-sectional average of the RMSPE, and maturity-specific time-varying weights) I find that the combined forecasts lead to additional gains in predictive accuracy, particularly for stable periods. Moreover, on average, the maturity-specific weighting scheme provides more accurate forecasts than the other two schemes for the short-end of the yield curve. However, the scheme-based combined forecasts are dominated by the forecasts from the stochastic factor-augmented Nelson-Siegel model in periods where the yields are highly volatile. A possible explanation is that the macro-factors contain valuable information for the loading parameter in such periods (this is the only property that the combined forecasts do not exhibit). Furthermore, I find that the extended models outperform the standard DNS most often for the short-end of the yield curve, followed by the long-term yields and the yields of intermediate maturity. Regarding the differences between the filtering techniques, the findings indicate that the use of the unscented Kalman filter rather than the extended Kalman filter for fitting nonlinear Nelson-Siegel models is beneficial for both ends of the yield curve (due to the inability of the extended Kalman filter to accurately approximate the curvature factor) and could have a positive impact on the accuracy in the predictive framework in some cases.

In addition to the contributions mentioned in the paragraph above, I find that the interpretation of the factors change due to the incorporation of a time-varying loading parameter, which results in a better (or worse) fit of the yield. Moreover, I find the estimated time-varying loading parameters from both filtering methods to vary considerably over time, producing some doubt on studies where $\lambda_t$ is fixed. Furthermore, as a contribution to the strand of forecast combination literature, I find that the stochastic factor augmented Nelson-Siegel model (essentially a combined model) is able to consistently outperform combined forecasts based on various weighting schemes when the yields are relatively highly volatile.

In the remainder of this paper, the next section begins with a description of the dynamic Nelson-Siegel model (Diebold and Li, 2006), the factor augmented Nelson-Siegel model as introduced by Diebold et al. (2006), the GARCH extension of the DNS (Koopman et al., 2010) and the corresponding estimation procedures. Then, I introduce the DNS with time-varying loadings as in Koopman et al., (2010) for which I consider and discuss both the extended- and unscented Kalman filter. The third section describes the data that I use for this research, consisting of the Fama-Bliss zero-coupon dataset and the indicators for the macro-economy. The fourth section contains the empirical results, and the fifth section concludes.

# 2 Methodology

In this section I present the methods that are used. First, I provide a brief review of the dynamic Nelson-Siegel model and its form in the state-space framework. Next, I elaborate on the factor augmented Nelson-Siegel model and principal components analysis. I continue by introducing the GARCH extension of the dynamic Nelson-Siegel model as introduced by Koopman et al. (2010), which allows for stochastic volatility. The estimation procedure of these models will then be explained. In addition, I discuss the usefulness of allowing the loading parameter to be time-varying. This extension introduces nonlinearities in the model, and to cope with this I elaborate on the estimation procedures that incorporate the extended Kalman filter and the unscented Kalman filter.

## 2.1 Dynamic Nelson-Siegel Model and the State-Space Framework

The original Nelson-Siegel (1987) model gives a static representation of the yield curve. However, to understand the evolution of the bond market over time, a dynamic version is required. For that reason, Diebold and Li (2006) modified the model such that it allows the coefficients to vary over time. The dynamic Nelson-Siegel (DNS) model is then given as

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t}\Big(\frac{1-\exp(-\lambda_t\tau)}{\lambda_t\tau}\Big) + \beta_{3,t}\Big(\frac{1-\exp(-\lambda_t\tau)}{\lambda_t\tau} - \exp(-\lambda_t\tau)\Big), \quad t = 1,\ldots,T, \quad (1)$$

where $y_t(\tau)$ is the yield at time $t$ for a maturity of $\tau$ months. This model decomposes the yield curve into two parts i.e. the part consisting of factors and a part consisting of their coefficients. The three dynamic, latent factors ($\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$) determine the dynamics of $y_t$ for any maturity $\tau$, while the cross-section of the yields $y(\tau)$ is formed by the coefficients (factor loadings) for any time $t$. An inspection of the factor loadings provides more information on the interpretation of the latent factors. These factor loadings are a function of maturity and plotted in Figure 1. First, I evaluate the loading on $\beta_{1,t}$. It is fixed and equal to 1 for all maturities. Moreover, an increase in $\beta_{1,t}$ results in an equal increase of all yields as the loading on this factor is identical for all maturities. Additionally, when the time to maturity approaches infinity, it is the only factor that affects the yields. Therefore, this factor is regarded as the level factor or the long-term factor. Now, consider the loading on $\beta_{2,t}$, $(1-e^{-\lambda_t\tau})/\lambda_t\tau$.

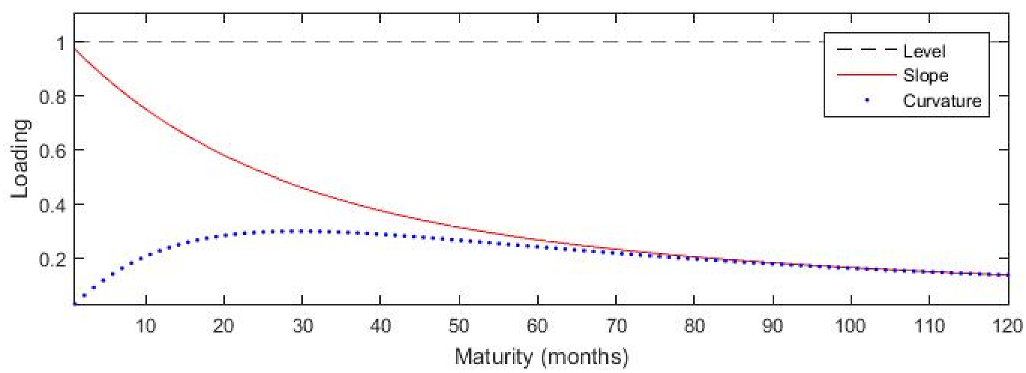



Figure 1: Factor loadings of the DNS model. This figure shows the factor loadings as a function of maturity for $\lambda_t = 0.0609$

For very low values of $\tau$, this function is near 1 and decays quickly to 0 as the maturity approaches infinity. In other words, the short-term yields load heavily on $\beta_{2,t}$, while the long-term yields load negligibly on this factor. Mostly affecting short-term yields, $\beta_{2,t}$ is often called the short-term factor. As mentioned by Diebold and Li (2006), an increase in the short-term factor, $\beta_{2,t}$, changes the slope due to the loading. Hence, they interpret the short-term factor as the slope factor. Finally, consider the loading on $\beta_{3,t}$, $(1-e^{-\lambda_t \tau})/\lambda_t \tau - e^{-\lambda_t \tau}$. For very short-term maturities this function is near 0, increases as $\tau$ increases, an then decays to zero again for very long-term maturities. Thus, an increase in $\beta_{3,t}$ will have minimal effect on both ends of the yield curve, but will increase medium-term yields as they load more heavily on it. Hence, this factor is often called the medium-term factor or the curvature factor.

The parameter $\lambda_t$ in the factor loadings determines the rate of exponential decay of the loading for the short-term factor $\beta_{2t}$ and the maturity where the loading for the medium-term factor $\beta_{3,t}$ reaches its maximum. Moreover, small values of $\lambda_t$ results in slow decay and can better fit the curve at long maturities, while large values of $\lambda_t$ have the reverse effect. Diebold and Li (2006), fix the loading parameter $\lambda_t$ for the standard DNS and set it equal to 0.0609. This value maximizes the loading on the medium-term factor at exactly 30 months (which is the average of the two- and three-year maturities as these are commonly regarded as the intermediate maturities). In this paper, I follow them and fix $\lambda_t$ at this value for the standard DNS model.

Following Diebold et al. (2006), I write the DNS model in a state-space representation form, which consists of a measurement and a transition equation. Consider a fixed set of maturities $(\tau_1, \tau_2, \ldots, \tau_m)$, by adding stochastic error terms $\epsilon_t(\tau)$, which I interpret as idiosyncratic or maturity-specific factors, to equation (1) I obtain the measurement equation,

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_m) \end{pmatrix} = \begin{pmatrix} 1 & \frac{1-\exp(-\lambda\tau_1)}{\lambda\tau_1} & \frac{1-\exp(-\lambda\tau_1)}{\lambda\tau_1} - \exp(-\lambda\tau_1) \\ 1 & \frac{1-\exp(-\lambda\tau_2)}{\lambda\tau_2} & \frac{1-\exp(-\lambda\tau_2)}{\lambda\tau_2} - \exp(-\lambda\tau_2) \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-\exp(-\lambda\tau_m)}{\lambda\tau_m} & \frac{1-\exp(-\lambda\tau_m)}{\lambda\tau_m} - \exp(-\lambda_t\tau_m) \end{pmatrix} \begin{pmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{pmatrix} + \begin{pmatrix} \epsilon_t(\tau_1) \\ \epsilon_t(\tau_2) \\ \vdots \\ \epsilon_t(\tau_m) \end{pmatrix},$$

which can be rewritten as

$$\boldsymbol{y}_t = \boldsymbol{\Lambda}(\lambda)\boldsymbol{\beta}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim NID(0, \boldsymbol{\Sigma}_\epsilon), \quad t = 1, \ldots T, \tag{2}$$

where $\boldsymbol{\Lambda}(\lambda)$ depends on $\lambda$ only, for given $\tau$. Doing so, we are able to partition the yield $y_t(\tau)$ in two parts; a part driven by the common factors $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$, and a part driven by its idiosyncratic factor $\epsilon_t(\tau)$. In addition to the measurement equation, I present the transition equation, which specifies the common factor dynamics. In accordance with Diebold and Li (2006), I assume the common factors to follow a autoregressive process of first order with mean vector $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)'$. The transition equation is then given as

$$\begin{pmatrix} \beta_{1,t+1} \\ \beta_{2,t+1} \\ \beta_{3,t+1} \end{pmatrix} = \begin{pmatrix} 1-\phi_{11} & 0 & 0 \\ 0 & 1-\phi_{22} & 0 \\ 0 & 0 & 1-\phi_{33} \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} + \begin{pmatrix} \phi_{11} & 0 & 0 \\ 0 & \phi_{22} & 0 \\ 0 & 0 & \phi_{33} \end{pmatrix} \begin{pmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{pmatrix} + \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \end{pmatrix},$$

which can be rewritten in matrix/vector notation as

$$\boldsymbol{\beta}_{t+1} = (\boldsymbol{I}_3 - \boldsymbol{\Phi})\boldsymbol{\mu} + \boldsymbol{\Phi}\boldsymbol{\beta}_t + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim NID(0, \boldsymbol{\Sigma}_\eta) \tag{3}$$

where $\mathbf{\Phi}$ governs the factor dynamics and $\boldsymbol{I}_3$ denotes the identity matrix of size $3 \times 3$. I will not include further lags of $\boldsymbol{\beta}_t$ in the transition equation as Diebold and Li (2006) find that it is sufficient to use only one lag to describe the evolution of the loading vector. As in Christensen, Diebold and Rudebusch (2011), I assume that the measurement disturbance $\boldsymbol{\epsilon}_t$ and transition disturbance $\boldsymbol{\eta}_t$ are zero-mean white noise and uncorrelated to each other. Furthermore, I assume that both the covariance matrix $\mathbf{\Sigma}_\epsilon$ as well as the covariance matrix $\mathbf{\Sigma}_\eta$ are diagonal. Diebold et al. (2006) explore the use of a more richly parametrized correlated structure of $\mathbf{\Phi}$, where the state variables $\boldsymbol{\beta}_t$ may interact dynamically and/or their shocks may be correlated. However, Christensen, Diebold and Rudebusch (2011) found that the latter structure of $\mathbf{\Phi}$ is dominated by its parsimonious version, as specified in (3), in an out-of-sample setting. As I put more emphasis on forecasting, I stick to only having the parsimonious version. I will use the DNS model (2) and (3) as the standard and use its empirical results as the benchmark for its extensions.

## 2.2 Factor Augmented Nelson-Siegel Model

The first extension of the DNS, in this paper, is the inclusion of macro-finance data in the model. In addition to the yield data, assume that a large number of $k$ macroeconomic variables is available at the monthly frequency, covering the same period as the yield data. Denote these variables by $\boldsymbol{X}_t = (x_{1,t}, \ldots, x_{k,t})'$ for each period $t$. As the estimation uncertainty increases with the number of variables in the model, a parsimonious model is preferred. Therefore, I summarize the large amount of information by a limited number of $p$ factors, $\boldsymbol{f}_t = (f_{1,t}, \ldots, f_{p,t})'$, with $p << k$. Moreover, its elements are normalized to have mean zero.

I follow Exterkate et al. (2013) in their procedure of including macro-economic information in the Nelson-Siegel models. This means that the measurement equation (2) remains unchanged, but that I have to adjust the transition equation (3) by adding the macro-factors $\boldsymbol{f}_t$ to the vector $\boldsymbol{\beta}_t$, such that the macro-factors affect the individual yields only via the Nelson-Siegel factors. For convenience, the adjusted state space representation is given by

$$\boldsymbol{y}_t = \mathbf{\Lambda}(\lambda)\boldsymbol{\beta}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim NID(0, \mathbf{\Sigma}_\epsilon), \quad t = 1, \ldots T, \tag{4}$$

$$\begin{pmatrix} \boldsymbol{\beta}_{t+1} \\ \boldsymbol{f}_{t+1} \end{pmatrix} = (\boldsymbol{I}_{3+p} - \mathbf{\Phi}^{FA}) \begin{pmatrix} \boldsymbol{\mu} \\ \mathbf{0}_p \end{pmatrix} + \mathbf{\Phi}^{FA} \begin{pmatrix} \boldsymbol{\beta}_t \\ \boldsymbol{f}_t \end{pmatrix} + \boldsymbol{\eta}_t^{FA}, \quad \boldsymbol{\eta}_t^{FA} \sim NID(0, \mathbf{\Sigma}_\eta^{FA}), \tag{5}$$

where $\boldsymbol{I}_{3+p}$ is the identity matrix of size $(3+p) \times (3+p)$, $\mathbf{0}_p$ is a $p \times 1$ vector consisting of zeros, the dimensions of $\mathbf{\Phi}$, $\boldsymbol{\eta}$ and $\mathbf{\Sigma}_\eta$ are increased as appropriate and they will be denoted by the superscript 'FA'. To be more specific, $\mathbf{\Phi}^{FA}$ is a $(3+p) \times (3+p)$ matrix, $\boldsymbol{\eta}^{FA}$ is a $(3+p) \times 1$ vector and $\mathbf{\Sigma}_\eta^{FA}$ is a $(3+p) \times (3+p)$ matrix. Moreover, $\mathbf{\Sigma}_\eta^{FA}$ remains diagonal and

$$\mathbf{\Phi}^{FA} = \left(\begin{array}{c|c} \text{Diagonal} & \text{Unrestricted} \\ \hline 0_{p\times 3} & \text{Diagonal} \end{array}\right),$$

where the block-structure corresponds to the partitioning of the state vector into $\boldsymbol{\beta}_t$ and $\boldsymbol{f}_t$. The structure of the coefficient matrix implies that the macro-factors affect the individual yields only via the Nelson-Siegel factors, and that there is no feedback from the yields to the macro-factors. Moreover, I restrict the bottom-right matrix to be diagonal to keep the model as parsimonious as possible. I will refer to this model as the factor-augmented DNS (FADNS) model.

### 2.2.1 Principal Component Analysis

Similar to Moench (2008) and DePooter et al. (2007), among others, I extract factors $\boldsymbol{f}_t$ from the dataset by applying principal component analysis (PCA), see Stock and Watson (2002). PCA reduces the dimension of the dataset by constructing factors that capture the common variance. The construction of the factors (or principal components), is executed by maximizing the explained variance in the dataset for each principal component. This maximization is done under the constraint that each principal component has to be orthogonal to preceding principal components.

In order to exclude the sensitivity of PCA to the scaling of variables, I normalize the time series of observations on each macro variable separately to have mean zero and unit variance over the estimation window. Then, I conduct PCA on the normalized macro-finance dataset. This is done by performing the eigenvalue decomposition on the dataset, $\boldsymbol{P}$. That is, the following problem is solved:

$$\boldsymbol{P}\boldsymbol{v} = \lambda\boldsymbol{v}, \quad \text{s.t.} \quad \boldsymbol{v}'\boldsymbol{v} = 1. \tag{6}$$

This provides a diagonal matrix $\boldsymbol{D}$ with the eigenvalues, $\hat{\lambda}_1 \leq \ldots \leq \hat{\lambda}_n$, of the correlation matrix $P$ on its diagonal. Moreover, $n$ denotes the number of variables used to construct the correlation matrix $\boldsymbol{P}$. In addition, we obtain $\boldsymbol{V} = (\hat{\boldsymbol{v}}_1, \ldots, \hat{\boldsymbol{v}}_n)$, where its columns are the eigenvectors corresponding to the eigenvalues. Subsequently, the factors $\boldsymbol{f}$ are constructed as linear combinations of the variables using the loadings in each eigenvector, that is

$$f_i = \hat{\boldsymbol{v}}_i'\boldsymbol{X}, \tag{7}$$

for $i = 1, \ldots, n$ and each period $t$. Moreover, I use the corresponding eigenvalues to estimate the proportion of variance explained by the factor, that is $\hat{\lambda}_i/n$.

## 2.3 Dynamic Nelson-Siegel Model and Stochastic Volatility

Another extension of the DNS is to incorporate stochastic volatility. To account for stochastic volatility, I follow the approach of Koopman et al. (2010). They allow for stochastic volatility in the measurement shocks $\boldsymbol{\epsilon}_t$ using the single-factor multivariate GARCH model of Harvey et al. (1992). More specifically, they decompose $\boldsymbol{\epsilon}_t$ as

$$\boldsymbol{\epsilon}_t = \boldsymbol{\Gamma}\epsilon_t^* + \boldsymbol{\epsilon}_t^+, \quad \epsilon_t^*|\boldsymbol{\Psi}_{t-1} \sim NID(0, h_t), \quad \boldsymbol{\epsilon}_t^+ \sim NID(0, \boldsymbol{\Sigma}_\epsilon^+), \quad t = 1, \ldots T, \tag{8}$$

where, $\boldsymbol{\Gamma}$ is a $m \times 1$ loading vector that determines the sensitivity of the yields to the common shock. Koopman et al. (2010) find that yields with short maturities are more heavily loaded on the common shock component than yields with longer maturities. In addition, $\boldsymbol{\epsilon}_t^+$ is a $m \times 1$ vector as well, representing the disturbance vector, with $\boldsymbol{\Sigma}_t^+$ as its corresponding diagonal covariance-matrix. The scalar $\epsilon_t^*$ represents the common disturbance term. The disturbance components are independent. The variance of $\epsilon_t^*$ conditional on all information up to time $t-1$ ($\boldsymbol{\Psi}_{t-1}$), denoted by $h_t$, is specified as the GARCH process introduced by Bollserslev (1986). This is given by

$$h_{t+1} = \gamma_0 + \gamma_1\epsilon_t^{*2} + \gamma_2 h_t, \quad t = 1, \ldots T, \tag{9}$$

where the unknown coefficients are subject to a few constraints. More specifically, $\{\gamma_0, \gamma_1, \gamma_2\} > 0$ and $\gamma_1 + \gamma_2 < 1$ to guarantee that $h_{t+1}$ is positive and stationary. The variance of the common component at time $t = 1$ is set equal to the unconditional variance, that is, $h_1 = \frac{\gamma_0}{1-\gamma_1-\gamma_2}$. By incorporating the GARCH specification in the DNS, the variance matrix of $\boldsymbol{\epsilon}_t$ is stochastic through $h_t$ and given by

$$\boldsymbol{\Sigma}_\epsilon(h_t) = h_t\boldsymbol{\Gamma}\boldsymbol{\Gamma}' + \boldsymbol{\Sigma}_\epsilon^*, \quad t = 1, \ldots T, \tag{10}$$

Koopman et al. (2010) suggest several possibilities to overcome identification issues. First, the loading vector $\mathbf{\Gamma}$ could be normalized such that $\mathbf{\Gamma}'\mathbf{\Gamma} = 1$. Another possibility is to assume that $\gamma_0$ is known and fixed at a certain value. Koopman et al. (2010) opt for the latter approach, and I choose to follow this. However, in practice the chosen approach to prevent identification problems does not matter as the outcomes of all methods are equal up to a scaling factor. By including the GARCH decomposition in the DNS model, the state-space equations slightly change. More specifically, the measurement equation and state equation are given by

$$\boldsymbol{y}_t = \mathbf{\Lambda}(\lambda)\boldsymbol{\beta}_t + \mathbf{\Gamma}\epsilon_t^* + \boldsymbol{\epsilon}_t^+, \quad \boldsymbol{\epsilon}_t^+ \sim NID(0, \mathbf{\Sigma}_\epsilon^+), \quad t = 1, \dots T, \tag{11}$$

$$\begin{pmatrix} \boldsymbol{\beta}_{t+1} \\ \epsilon_{t+1}^* \end{pmatrix} = \begin{pmatrix} (\boldsymbol{I}_3 - \mathbf{\Phi})\boldsymbol{\mu} \\ 0 \end{pmatrix} + \begin{pmatrix} \mathbf{\Phi} & \mathbf{0}_3 \\ \mathbf{0}_3' & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_t \\ \epsilon_t^* \end{pmatrix} + \begin{pmatrix} \boldsymbol{\eta}_t \\ \epsilon_{t+1}^* \end{pmatrix}, \quad \begin{pmatrix} \boldsymbol{\eta}_t \\ \epsilon_{t+1}^* \end{pmatrix} \sim NID\left(0, \begin{pmatrix} \mathbf{\Sigma}_\eta & \mathbf{0}_3 \\ \mathbf{0}_3' & h_{t+1} \end{pmatrix}\right), \tag{12}$$

where $\mathbf{0}_3$ denotes a $3 \times 1$ vector consisting of zeros. I refer to this model as DNS-GARCH.

### 2.4 Time-Varying Loading Parameter $\lambda$

In most studies, the value of $\lambda_t$ is fixed at a certain value. In this thesis, $\lambda_t$ is set equal to 0.0609 (for DNS, FADNS and DNS-GARCH) following Diebold and Li (2006). Alternatively, a constant $\lambda_t$ can be estimated along with the other model parameters as in Diebold et al. (2006) and De Pooter et al. (2007), among others. More specifically, Diebold et al. (2006) find that the estimated value is equal to 0.077. Following this, Yu and Zivot (2010) fix the value of $\lambda$ at 0.077. They argue that different values of $\lambda_t$ affect the factor loadings $\Lambda(\lambda)$ only by a small amount. By allowing for a time-varying $\lambda_t$, we move from a linear to a nonlinear state-space environment. Koopman et al. (2010) propose ways of allowing for time-varying $\lambda_t$. They find that they could be useful in situations that involve not only time-varying curvature, but also in situations where the location at which the curvature attains its maximum is time-dependent. A recent example of this is the changing part of the short end of the U.S. yield curve, which is at the zero bound. Such time-varying yield curve kinks may be captured by imposing time-varying $\lambda_t$. For this reason, I will also consider the loading parameter $\lambda_t$ to be time-varying.

Following Koopman et al. (2010), $\lambda_t$ is treated as a latent factor that is included in the original state vector. The inclusion then results in a change of the state-space equations. More specifically, the measurement equation and state equation are given by

$$\boldsymbol{y}_t = \mathbf{\Lambda}(\lambda_t)\boldsymbol{\beta}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim NID(0, \mathbf{\Sigma}_\epsilon), \quad t = 1, \dots T, \tag{13}$$

$$\boldsymbol{\alpha}_{t+1} = (\boldsymbol{I}_4 - \mathbf{\Phi}^{TVL})\boldsymbol{\mu}^{TVL} + \mathbf{\Phi}^{TVL}\boldsymbol{\alpha}_t + \boldsymbol{\eta}_t^{TVL}, \quad \boldsymbol{\eta}_t^{TVL} \sim NID(0, \mathbf{\Sigma}_\eta^{TVL}) \tag{14}$$

where $\boldsymbol{\alpha}_t = (\beta_{1,t}, \beta_{2,t}, \beta_{3,t}, \lambda_t)'$, and the dimensions of $\mathbf{\Phi}$, $\boldsymbol{\mu}$, $\boldsymbol{\eta}$ and $\mathbf{\Sigma}_\eta$ are increased as appropriate denoted by the superscript 'TVL'. To be more specific, $\mathbf{\Phi}^{TVL}$ is a $4 \times 4$ diagonal matrix, both $\boldsymbol{\mu}^{TVL}$ and $\boldsymbol{\eta}^{TVL}$ are $4 \times 1$ vectors and $\mathbf{\Sigma}_\eta^{TVL}$ is a $4 \times 4$ diagonal matrix. I will refer to this model as the DNS model with time-varying loadings (DNS-TVL).

Recall the interpretation of the loading parameter $\lambda_t$, it determines the rate of exponential decay of the loading for the short-term factor and the maturity where the loading for the medium-term factor reaches its maximum. Therefore, only positive values for $\lambda_t$ are expected. To guarantee this, I truncate the lower bound values of $\lambda$ at $5 \times 10^{-3}$. This value maximizes the loading on the medium-term factor at 360 months, which is the longest maturity of issued U.S. Treasury bonds. Furthermore, the upper bound values are truncated at 1.8 (the value that ensures the maximum of the loading on the medium-term factor to be at 1 month, the shortest maturity of issued U.S. Treasury bonds).

## 2.5 The Stochastic Factor-Augmented Nelson-Siegel Model

The last model that I consider in my thesis is the combination of all extensions into one model. I will refer to this model as the stochastic factor-augmented Nelson-Siegel model (SFADNS). The state-space form is then given by

$$\boldsymbol{y}_t = \boldsymbol{\Lambda}(\lambda_t)\boldsymbol{\beta}_t + \boldsymbol{\Gamma}\epsilon_t^* + \boldsymbol{\epsilon}_t^+, \quad \boldsymbol{\epsilon}_t^+ \sim NID(0, \boldsymbol{\Sigma}_\epsilon^+), \quad t = 1, \dots T, \tag{15}$$

$$\begin{pmatrix} \boldsymbol{\alpha}_{t+1} \\ \boldsymbol{f}_{t+1} \\ \epsilon_{t+1}^* \end{pmatrix} = \begin{pmatrix} (\boldsymbol{I}_{4+p} - \boldsymbol{\Phi}^{SFA}) & \boldsymbol{0}_{4+p} \\ \boldsymbol{0}'_{4+p} & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\mu}^{SFA} \\ 0 \end{pmatrix} + \begin{pmatrix} \boldsymbol{\Phi}^{SFA} & \boldsymbol{0}_{4+p} \\ \boldsymbol{0}'_{4+p} & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha}_t \\ \boldsymbol{f}_t \\ \epsilon_t^* \end{pmatrix} + \begin{pmatrix} \boldsymbol{\eta}_{t+1}^{SFA} \\ \epsilon_{t+1}^* \end{pmatrix}, \tag{16}$$

where

$$\begin{pmatrix} \boldsymbol{\eta}_{t+1}^{SFA} \\ \epsilon_{t+1}^* \end{pmatrix} \sim NID\left(0, \begin{pmatrix} \boldsymbol{\Sigma}_\eta^{SFA} & \boldsymbol{0}_{4+p} \\ \boldsymbol{0}'_{4+p} & h_{t+1} \end{pmatrix}\right).$$

Furthermore, the dimensions of $\boldsymbol{\Phi}$, $\boldsymbol{\eta}$ and $\boldsymbol{\Sigma}_\eta$ are increased as appropriate and they will be denoted by the superscript 'SFA'. To be more specific, $\boldsymbol{\Phi}^{SFA}$ is a $(4+p) \times (4+p)$ matrix, $\boldsymbol{\eta}^{SFA}$ is a $(4+p) \times 1$ vector and $\boldsymbol{\Sigma}_\eta^{SFA}$ is a $(4+p) \times (4+p)$ diagonal matrix. Moreover, the structure of

$$\boldsymbol{\Phi}^{SFA} = \left(\begin{array}{c|c} \text{Diagonal} & \text{Unrestricted} \\ \hline 0_{p\times 4} & \text{Diagonal} \end{array}\right), \quad \boldsymbol{\mu}^{SFA} = \left(\begin{array}{c} \boldsymbol{\mu}^{FA} \\ \hline 0_p \end{array}\right),$$

where the block-structure corresponds to the partitioning of the state vector into $\boldsymbol{\alpha}_t$ and $\boldsymbol{f}_t$. Almost similar to the interpretation of $\boldsymbol{\Phi}^{FA}$, the coefficient matrix $\boldsymbol{\Phi}^{SFA}$ implies that there is no feedback from the Nelson-Siegel factors and the time-varying loading parameter to the macro-factors.

## 2.6 Estimation Procedure for Linear Gaussian State-Space Models

Several procedures are studied for estimating and forecasting the Nelson-Siegel models, ranging from a two-step procedure to the class of one-step estimation procedures where all estimation is done simultaneously by exploiting the state-space structure of the dynamic Nelson-Siegel model. In this thesis, I will focus on the latter approach. This approach combines the Kalman Filter (KF) with Maximum Likelihood (ML) estimation. Before I will go into the details of the procedure, I will first introduce some new general notation. Consider the following general linear state-space model

$$\boldsymbol{y}_t = \boldsymbol{H}\boldsymbol{x}_t + \boldsymbol{v}_t, \quad \boldsymbol{v}_t \sim NID(0, \boldsymbol{\Sigma}_v), \quad t = 1, \dots, T, \tag{17}$$

$$\boldsymbol{x}_{t+1} = \boldsymbol{C} + \boldsymbol{K}\boldsymbol{x}_t + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim NID(0, \boldsymbol{\Sigma}_\omega). \tag{18}$$

The general procedure to find the values of the latent factors and the unknown parameters is recursive. The process is started by making an initial guess for the unknown parameters ($\boldsymbol{\theta}^{(0)}$), and run the KF. The optimal forecast (estimate) of the state vector $\boldsymbol{x}_t$ given the information known at time $t-1$ ($t$) are provided by the prediction and the update step of the KF. The results are then used as input in the log likelihood function such that ML estimation can be conducted to obtain new estimates of the unknown parameters ($\boldsymbol{\theta}^{(1)}$). These steps are then iterated until the parameter values $\boldsymbol{\theta}_{ML}$ are found for which the log likelihood function is maximized. The details of the procedure are discussed below.

**Initialisation**
The procedure starts with the initialisation of parameters. I use the two-step approach of Diebold and Li (2006) to obtain the parameter estimates for $\boldsymbol{C}$, $\boldsymbol{K}$, $\boldsymbol{\Sigma}_v$ and $\boldsymbol{\Sigma}_\omega$ and use these as the initial values for the Kalman filter. Furthermore, I have to initialise the state vector and its covariance matrix. More specifically, under the assumption of a stationary process, the initial value of the state vector is set equal to the unconditional mean, $\boldsymbol{x}_{0|0} = E[\boldsymbol{x}_t] = \boldsymbol{\mu}$, and the initial value of the covariance matrix, $\boldsymbol{\Sigma}_{x,0|0}$, is set equal to $\boldsymbol{V}$, where $\boldsymbol{V}$ solves $\boldsymbol{V} = \boldsymbol{K}\boldsymbol{V}\boldsymbol{K}' + \boldsymbol{\Sigma}_\omega$[ii].

**Prediction step**
The first step of the KF is the prediction step. Consider period $t-1$ and let $\boldsymbol{x}_{t-1|t-1}$ be the minimum mean square estimator (MMSE) of $\boldsymbol{x}_{t-1}$ at this period. Furthermore, let the covariance matrix of the estimation error, $\boldsymbol{x}_{t-1} - \boldsymbol{x}_{t-1|t-1}$, be denoted as $\boldsymbol{\Sigma}_{x,t-1|t-1}$. To obtain the optimal forecast of this state at time $t$ given the information known at time $t-1$, denoted by $\boldsymbol{x}_{t|t-1}$, one should calculate the expectation of the transition equation (18) conditional on $\boldsymbol{\Psi}_{t-1}$, i.e. the history of $\boldsymbol{y}$ and $\boldsymbol{x}$ up to and including the observations at $t-1$. As both the parameters in $C$ and $K$ as well as an estimate of the state vector at time $t-1$ are known, the conditional mean is equal to the transition equation itself as shown in (19). Next, the optimal forecast for the corresponding covariance matrix using the information known at time $t-1$ should be made. This is done by calculating the variance of the state vector conditional on $\boldsymbol{\Psi}_{t-1}$ using basic statistics as in (20). More formally, the prediction step consists of the following

$$\boldsymbol{x}_{t|t-1} = \mathrm{E}[\boldsymbol{x}_t|\boldsymbol{\Psi}_{t-1}] = \boldsymbol{C} + \boldsymbol{K}\boldsymbol{x}_{t-1|t-1}, \tag{19}$$

$$\boldsymbol{\Sigma}_{x,t|t-1} = \mathrm{Var}[\boldsymbol{x}_t|\boldsymbol{\Psi}_{t-1}] = \boldsymbol{K}\boldsymbol{\Sigma}_{x,t-1|t-1}\boldsymbol{K}' + \boldsymbol{\Sigma}_\omega. \tag{20}$$

**Update step**
Now consider the time-$t$ update step. In this step, the forecasts obtained from the prediction step are updated using information from the prediction error as this may hold information that is not yet contained in the forecasts themselves. The prediction error $\tilde{\boldsymbol{v}}_t$ is calculated by subtracting the forecasted yield $\boldsymbol{y}_{t|t-1}$ from the observed yield $\boldsymbol{y}_t$. Again, the optimal forecast of the yield is equal to the expectation of the yield conditional on $\boldsymbol{\Psi}_{t-1}$. Subsequently, the corresponding covariance matrix is calculated using basic statistics. These steps are shown in (21), respectively (22). The update of the predicted state vector, $\boldsymbol{x}_{t|t}$, and its covariance matrix, $\boldsymbol{\Sigma}_{x,t|t}$, are shown in (23), respectively (24). This is based on a property of a joint normal distribution. More specifically, to update the predicted state vector and its covariance matrix one should calculate the expectation, respectively the variance, of the state $\boldsymbol{x}_t$ conditional on $\boldsymbol{\Psi}_{t-1}$ *and* $\boldsymbol{v}_t$. More formally, the update step consists of the following

$$\tilde{\boldsymbol{v}}_t = \boldsymbol{y}_t - \mathrm{E}[\boldsymbol{y}_t|\boldsymbol{\Psi}_{t-1}] = \boldsymbol{y}_t - \boldsymbol{y}_{t|t-1} = \boldsymbol{y}_t - \boldsymbol{H}\boldsymbol{x}_{t|t-1}, \tag{21}$$

$$\boldsymbol{\Sigma}_{\tilde{v},t} = \mathrm{Var}[\boldsymbol{v}_t|\boldsymbol{\Psi}_{t-1}] = \boldsymbol{H}\boldsymbol{\Sigma}_{x,t|t-1}\boldsymbol{H}' + \boldsymbol{\Sigma}_v, \tag{22}$$

$$\boldsymbol{x}_{t|t} = \mathrm{E}[\boldsymbol{x}_t|\boldsymbol{v}_t, \boldsymbol{\Psi}_{t-1}] = \boldsymbol{x}_{t|t-1} + \boldsymbol{\Sigma}_{x,t|t-1}\boldsymbol{H}'\boldsymbol{\Sigma}_{\tilde{v},t}^{-1}\tilde{\boldsymbol{v}}_t, \tag{23}$$

$$\boldsymbol{\Sigma}_{x,t|t} = \mathrm{Var}[\boldsymbol{x}_t|\boldsymbol{v}_t, \boldsymbol{\Psi}_{t-1}] = \boldsymbol{\Sigma}_{x,t|t-1} - \boldsymbol{\Sigma}_{x,t|t-1}\boldsymbol{H}'\boldsymbol{\Sigma}_{\tilde{v},t}^{-1}\boldsymbol{H}\boldsymbol{\Sigma}_{x,t|t-1}. \tag{24}$$

[ii] See Appendix A for an explanation on how to solve for $\boldsymbol{V}$.

**Maximum Likelihood Estimation**
The information provided by the KF is then used for the estimation of the unknown parameters $\boldsymbol{\theta}$. To obtain maximum likelihood estimates of these parameters, numerical maximization of the log likelihood function is used. As the error terms are assumed to be Gaussian distributed, the distribution of $\boldsymbol{y}_t$ conditional on the information up to time $t-1$ is Gaussian as well. The log likelihood function for the observations are then obtained from the KF via the prediction error decomposition, see Harvey and Peters (1990). Hence the log likelihood is given by

$$\ell(\theta) = -\frac{mT}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}\log|\boldsymbol{\Sigma}_{\tilde{v},t}| - \frac{1}{2}\sum_{t=1}^{T}\tilde{\boldsymbol{v}}_t'\boldsymbol{\Sigma}_{\tilde{v},t}^{-1}\tilde{\boldsymbol{v}}_t, \tag{25}$$

which is maximized with respect to the unknown parameter $\boldsymbol{\theta}$. Some additional steps are required for DNS-GARCH. For this case $\boldsymbol{\Sigma}_\omega$ contains $h_{t+1}$ as modelled in (9), which is not computable at time $t$ as it depends on the latent variable $\epsilon_t^*$. Therefore, I replace $h_{t+1}$ by its estimate based on the observations $y_1, \ldots, y_t$, that is

$$\hat{h}_{t+1|t} = \gamma_0 + \gamma_1 e_t^2 + \gamma_2 \hat{h}_{t|t-1}, \quad t = 1, \ldots T, \tag{26}$$

where $e_t$ is an estimate of $\epsilon_t^*$ based on $y_1, \ldots, y_t$ and obtained from the update step of the KF. More specifically, it is the last element of $\boldsymbol{x}_{t|t}$ when DNS-GARCH is estimated. Past values of $\hat{h}_{t|t-1}$ can be stored outside the model and the variance $h_{t+1}$ in matrix $\boldsymbol{\Sigma}_\omega$ is replaced by $\hat{h}_{t+1|t}$ for the prediction step of the KF. As a result, the state estimates are sub-optimal. A more detailed discussion of this approach is provided by Harvey et al. (1992). The procedure for linear and non-linear state space models is similar, however, the KF can only be used for linear models (i.e. for the DNS, FADNS and DNS-GARCH).

## 2.7 Estimation Procedure for Nonlinear Gaussian State-Space Models

The standard KF is only applicable if the state-space is linear, which is not the case when the loading parameter $\lambda_t$ is stochastic. For this reason, I consider nonlinear Gaussian filtering techniques for the DNS-TVL and the SFADNS. More specifically, I discuss the extended Kalman filter (EKF) and the unscented Kalman filter (UKF). First, I introduce some general notation before I go into the details of the two nonlinear filtering techniques. As $\lambda_t$ is now treated as a latent factor the general nonlinear state-space model is given by

$$\boldsymbol{y}_t = \boldsymbol{H}(\boldsymbol{x}_t) + \boldsymbol{v}_t, \quad \boldsymbol{v}_t \sim NID(0, \boldsymbol{\Sigma}_v), \quad t = 1, \ldots, T, \tag{27}$$

$$\boldsymbol{x}_{t+1} = \boldsymbol{C} + \boldsymbol{K}\boldsymbol{x}_t + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim NID(0, \boldsymbol{\Sigma}_\omega). \tag{28}$$

The general idea to find the values of the latent factors and the unknown parameters remains the same as discussed in section 2.6. However, as the system is nonlinear, sophisticated methods are needed to calculate the exact statistics of the nonlinear transformation of a random Gaussian variable. Furthermore, the initialisation values and the likelihood function remain unchanged as well. Therefore, I will only elaborate on the steps of both filtering methods.

### 2.7.1 The Extended Kalman Filter

The EKF relaxes the requirement of a linear state-space model, but, the model need only be differentiable and Gaussian distributed. The EKF approximates the nonlinear system using first order Taylor expansions to locally linearize the model around the most current estimate of the state. Explicitly, the linearized model at $\boldsymbol{x}_t = \boldsymbol{x}_{t|t}$ is as follows

$$\boldsymbol{y}_t = \boldsymbol{H}_t(\boldsymbol{x}_{t|t}) + \boldsymbol{J}_t \cdot (\boldsymbol{x}_t - \boldsymbol{x}_{t|t}) + \boldsymbol{v}_t, \quad t = 1, \ldots, T$$

where $\boldsymbol{J}_t$ denotes the Jacobian of the nonlinear function $\boldsymbol{H}(\boldsymbol{x}_t)$ in the original measurement equation. More specifically, the Jacobian is given by

$$\boldsymbol{J}_t = \frac{\partial H_t(\boldsymbol{x}_t)}{\partial \boldsymbol{x}_t'}\Big|_{\boldsymbol{x}_t = \boldsymbol{x}_{t|t}}.$$

**Prediction step**

The first step of the EKF is the prediction step. Consider period $t-1$ and let $\boldsymbol{x}_{t-1|t-1}$ be the minimum mean square estimator (MMSE) of $\boldsymbol{x}_{t-1}$ at this period. Furthermore, let the covariance matrix of the estimation error, $\boldsymbol{x}_{t-1} - \boldsymbol{x}_{t-1|t-1}$, be denoted as $\boldsymbol{\Sigma}_{x,t-1|t-1}$. As the approximation is locally linear in the estimated state vector $\boldsymbol{x}_{t-1|t-1}$ and due to the fact that the transition equation remains unchanged, the prediction step of the EKF is similar to the prediction step of the KF. Therefore, I refer to subsection 2.6 for more details on this step.

**Update step**

Now consider the time-$t$ update step of the EKF. Just as for the standard KF, the prediction error $\tilde{\boldsymbol{v}}_t$ is calculated by subtracting the forecasted yield $\boldsymbol{y}_{t|t-1}$ from the observed yield $\boldsymbol{y}_t$. Again, the optimal forecast of the yield is equal to the expectation of the yield conditional on $\boldsymbol{\Psi}_{t-1}$. Subsequently, the corresponding covariance matrix is calculated using basic statistics. These steps are shown in (29), respectively (30). The update of the predicted state vector, $\boldsymbol{x}_{t|t}$, and its covariance matrix, $\boldsymbol{\Sigma}_{x,t|t}$, are shown in (31), respectively (32). Even though the idea is similar, there are some small modifications. Moreover, as the model is nonlinear in the measurement equation, the Jacobian should be used to calculate the variance of the prediction error. More formally, the update step consists of

$$\tilde{\boldsymbol{v}}_t = \boldsymbol{y}_t - \mathrm{E}[\boldsymbol{y}_t|\boldsymbol{\Psi}_{t-1}] = \boldsymbol{y}_t - \boldsymbol{H}_t(\boldsymbol{x}_{t|t-1}) - \boldsymbol{J}_t \cdot (\boldsymbol{x}_{t|t-1} - \boldsymbol{x}_{t|t-1}) = \boldsymbol{y}_t - \boldsymbol{H}_t(\boldsymbol{x}_{t|t-1}), \tag{29}$$

$$\boldsymbol{\Sigma}_{\tilde{v},t} = \mathrm{Var}[\boldsymbol{v}_t|\boldsymbol{\Psi}_{t-1}] = \boldsymbol{J}_t \boldsymbol{\Sigma}_{x,t|t-1} \boldsymbol{J}_t' + \boldsymbol{\Sigma}_v, \tag{30}$$

$$\boldsymbol{x}_{t|t} = \mathrm{E}[\boldsymbol{x}_t|\boldsymbol{v}_t, \boldsymbol{\Psi}_{t-1}] = \boldsymbol{x}_{t|t-1} + \boldsymbol{\Sigma}_{x,t|t-1} \boldsymbol{J}_t' \boldsymbol{\Sigma}_{\tilde{v},t}^{-1} \tilde{\boldsymbol{v}}_t, \tag{31}$$

$$\boldsymbol{\Sigma}_{x,t|t} = \mathrm{Var}[\boldsymbol{x}_t|\boldsymbol{v}_t, \boldsymbol{\Psi}_{t-1}] = \boldsymbol{\Sigma}_{x,t|t-1} - \boldsymbol{\Sigma}_{x,t|t-1} \boldsymbol{J}_t' \boldsymbol{\Sigma}_{\tilde{v},t}^{-1} \boldsymbol{J}_t \boldsymbol{\Sigma}_{x,t|t-1}. \tag{32}$$

The calculated $\tilde{\boldsymbol{v}}_t$ and $\boldsymbol{\Sigma}_{\tilde{v},t}$ are then plugged in the log likelihood function as given in (25) in order to continue the estimation procedure for the latent factors and unknown parameters. The EKF is an easy method to implement and very effective to estimate the state when the nonlinearities in the model are not too complex. However, when the assumption of local linearity is violated, the filter may be highly unstable. Furthermore, when the initial estimate of the state is wrong, the filter may quickly diverge

### 2.7.2 The Unscented Kalman Filter

The UKF addresses the approximation issues of the EKF, and many authors (Julier and Uhlmann (2004), among others) show that it is an improvement over the EKF, while the calculation complexity is kept equal. Unlike the EKF, this filtering technique does not approximate

$\boldsymbol{H}(\boldsymbol{x}_t)$ by local linearization but instead uses the exact nonlinear measurement equation. This approach follows the same idea of the other filtering techniques, i.e. the optimal predictions of the state and the corresponding covariance matrix are calculated using the conditional expectation and variance. However, in contrast to the KF and the EKF, a statistical method is needed to evaluate the expectation of the nonlinear function of a random Gaussian variable. This should be done in order to calculate the prediction error and its covariance matrix. More specifically, for the prediction error the optimal prediction of the yield is needed, which is the expectation of the yield conditional on $\boldsymbol{\Psi}_{t-1}$. Or, equivalently, $\mathrm{E}[\boldsymbol{y}_t|\boldsymbol{\Psi}_{t-1}] = \mathrm{E}[H(\boldsymbol{x}_t)+\boldsymbol{v}_t|\boldsymbol{\Psi}_{t-1}] = \mathrm{E}[H(\boldsymbol{x}_t)|\boldsymbol{\Psi}_{t-1}]$. One suitable method to evaluate the expectation of the nonlinear function is the unscented transformation (UT) as introduced by Julier and Uhlmann (1997). They argue that it is easier to approximate a Gaussian distribution rather than an arbitrary nonlinear function (which is done in the EKF).

**The Unscented Transformation**

Before I go into the details of the UT, I first elaborate on a simple example analogous to this method. Suppose you would like to find the mean and variance of the transformed variable $y = f(x)$ of a random variable $x$. Furthermore, assume the mean and variance of this random variable to be known. In the most basic simulation procedure, a large number of samples are drawn from the distribution of the random variable. Subsequently, the mean of the transformed variable is approximated by the average of the transformed sample points $\mathrm{E}[y] \approx 1/N \sum_{i=1}^{N} f(x_i) = \bar{y}$ and the variance is approximated in a similar manner $\mathrm{Var}[y] = \mathrm{E}[y - \mathrm{E}[y]]^2 \approx 1/N \sum_{i=1}^{N}[f(x_i - \bar{y})]^2$. For the UT, the expectation and the variance of the random variable are approximated by a weighted average. Furthermore, instead of a drawing a large sample from the distribution of $x$, a minimal set of points around the mean are chosen. These points are referred to as so-called sigma points. The next question is then; how to choose the rights set of sigma points and weights.

**Computing the sigma points and the weights**

Consider the $n_x$-dimensional state vector $\boldsymbol{x}$ and assume the mean is equal to $\boldsymbol{\mu}$ and the covariance is equal $\boldsymbol{\Sigma}_x$. The sigma points should be chosen such that the sample mean and sample covariance are equal to the assumed mean and covariance matrix of the state vector. Equivalently, the points $\boldsymbol{x}$ and the weights $\boldsymbol{w}$ should be chosen such that the following hold:

$$\sum_i w_i = 1, \quad \boldsymbol{\mu} = \sum_i w_i \boldsymbol{x}_i, \quad \boldsymbol{\Sigma}_x = \sum_i w_i \Big(\boldsymbol{x}_i - \boldsymbol{\mu}\Big)\Big(\boldsymbol{x}_i - \boldsymbol{\mu}\Big)'.$$

There is no unique solution for these requirements. But, Julier and Uhlmann (1997) derive and find that the following definitions provide the most accurate results. They suggest choosing a set of $2n_x + 1$ sigma points of the form

$$\boldsymbol{x}_0^s = \boldsymbol{\mu}, \quad \boldsymbol{x}_i^s = \boldsymbol{\mu} + \Big(\sqrt{(n_x+\xi)\boldsymbol{\Sigma}_x}\Big)_i, \quad \boldsymbol{x}_{i+n_x}^s = \boldsymbol{\mu} - \Big(\sqrt{(n_x+\xi)\boldsymbol{\Sigma}_x}\Big)_{i+n_x}, \tag{33}$$

for $i = 1, \ldots, n_x$. The corresponding weights to compute the first and second moment are

$$W_0^{(m)} = \frac{\xi}{n_x+\xi}, \quad W_i^{(m)} = \frac{1}{2(n_x+\xi)}, \quad i = 1, \ldots, 2n_x, \tag{34}$$

$$W_0^{(c)} = \frac{\xi}{n_x+\xi} + (1-\rho^2+\zeta), \quad W_i^{(c)} = W_i^m \quad i = 1, \ldots, 2n_x, \tag{35}$$

where $\xi = \rho^2(n_x+\kappa) - n_x$ and where $\Big(\sqrt{(n_x+\xi)\boldsymbol{\Sigma}_x}\Big)_i$ is the $i$th column of the matrix square root of $(n_x+\xi)\boldsymbol{\Sigma}_x$, which is obtained using the Cholesky decomposition of the matrix. Furthermore, the scaling parameter $\rho > 0$ is intended to minimize higher order effects and can be made arbitrary small, I set this equal to $10^{-3}$. The restriction $\kappa > 0$ guarantees the positivity of the covariance matrix, and I set this equal to 1. The parameter $\zeta \geq 0$ can capture higher order moments of the state distribution and it is equal to 2 for the Gaussian distribution. Furthermore, the superscript $m$ and $c$ at the weights denote for which steps it should be used in the algorithm, either for the means of for the covariance matrices. The steps for this filtering technique for the case of zero mean noise are below.

**Prediction step**

Let $\boldsymbol{\chi}_{t-1|t-1} = [\boldsymbol{x}^s_{0,t-1|t-1} \quad \boldsymbol{x}^s_{i,t-1|t-1} \quad \boldsymbol{x}^s_{i+n_x,t-1|t-1}]$ for $i = 1, \ldots, n_x$ be the set of sigma points calculated at time $t-1$ given all information known up until time $t-1$. In the UKF, the transformations (the mapping of the measurement and transition equation) are applied to the sigma points, equations (36) and (39), rather than to the state vector and observed yields. The optimal forecasts of the state vector and the yields, the expectation conditional on $\boldsymbol{\Psi}_{t-1}$, are defined as the weighted averages of these transformations, as seen in equation (37) and (40). The details are as follows

$$\boldsymbol{\chi}_{i,t|t-1} = \boldsymbol{C} + \boldsymbol{K}\boldsymbol{\chi}_{i,t-1|t-1}, \quad i = 1, \ldots 2n_x, \tag{36}$$

$$\boldsymbol{x}_{t|t-1} = \sum_{i=0}^{2n_x} W_i^{(m)} \boldsymbol{\chi}_{i,t|t-1}, \tag{37}$$

$$\boldsymbol{\Sigma}_{x,t|t-1} = \sum_{i=0}^{2n_x} W_i^{(c)} \Big[\boldsymbol{\chi}_{i,t|t-1} - \boldsymbol{x}_{t|t-1}\Big] \Big[\boldsymbol{\chi}_{i,t|t-1} - \boldsymbol{x}_{t|t-1}\Big]' + \boldsymbol{\Sigma}_w, \tag{38}$$

$$\boldsymbol{\Upsilon}_{i,t|t-1} = \boldsymbol{H}_t\Big(\boldsymbol{\chi}_{i,t|t-1}\Big), \quad i = 1, \ldots 2n_x, \tag{39}$$

$$\boldsymbol{y}_{t|t-1} = \sum_{i=0}^{2n_x} W_i^{(m)} \boldsymbol{\Upsilon}_{i,t|t-1}. \tag{40}$$

**Update step**

Now consider the time-$t$ update step of the UKF, in which the prediction error $\tilde{\boldsymbol{v}}_t$ and its covariance matrix $\boldsymbol{\Sigma}_{\tilde{v},t}$ are calculated first. In combination with $\boldsymbol{\Sigma}_{x\tilde{v},t}$, the conditional covariance matrix of the state vector and the prediction error, they provide the information used to improve the predicted state $\boldsymbol{x}_{t|t-1}$, to $\boldsymbol{x}_{t|t}$, and its covariance matrix/mean square error matrix $\boldsymbol{\Sigma}_{x,t|t-1}$, to $\boldsymbol{\Sigma}_{x,t|t}$. Moreover, for these calculations the same procedure as for the KF and EKF is used, i.e. the expectation and the variance of the state $\boldsymbol{x}_t$ conditional on $\boldsymbol{\Psi}_{t-1}$ and $\boldsymbol{v}_t$ is calculated. More formally, the update step consists of the following

$$\tilde{\boldsymbol{v}}_t = \boldsymbol{y}_t - \boldsymbol{y}_{t|t-1}, \tag{41}$$

$$\boldsymbol{\Sigma}_{\tilde{v},t} = \sum_{i=0}^{2n_x} W_i^{(c)} \Big[\boldsymbol{\Upsilon}_{i,t|t-1} - \boldsymbol{y}_{t|t-1}\Big] \Big[\boldsymbol{\Upsilon}_{i,t|t-1} - \boldsymbol{y}_{t|t-1}\Big]' + \boldsymbol{\Sigma}_v, \tag{42}$$

$$\boldsymbol{\Sigma}_{x\tilde{v},t} = \sum_{i=0}^{2n_x} W_i^{(c)} \Big[\boldsymbol{\chi}_{i,t|t-1} - \boldsymbol{x}_{t|t-1}\Big] \Big[\boldsymbol{\Upsilon}_{i,t|t-1} - \boldsymbol{y}_{t|t-1}\Big]', \tag{43}$$

$$\boldsymbol{x}_{t|t} = \boldsymbol{x}_{t|t-1} + \boldsymbol{\Sigma}_{x\tilde{v},t}\boldsymbol{\Sigma}_{\tilde{v},t}^{-1}\tilde{\boldsymbol{v}}_t, \tag{44}$$

$$\boldsymbol{\Sigma}_{x,t|t} = \boldsymbol{\Sigma}_{x,t|t-1} - \boldsymbol{\Sigma}_{x\tilde{v},t}\boldsymbol{\Sigma}_{\tilde{v},t}^{-1'}\boldsymbol{\Sigma}_{x\tilde{v},t}'. \tag{45}$$

Parameter estimation using the UKF can be executed in a similar fashion as for the KF and EKF. That is, plug in $\tilde{\boldsymbol{v}}_t$ and $\boldsymbol{\Sigma}_{\tilde{v},t}$ in the log likelihood function as given in (25).

# 3 Data

In this section, I introduce the data that I use for my research. First, the yields for the different maturities and their stylized facts will be examined. Afterwards, I discuss which macroeconomic variables have been used from which I extract the macro factors for the models.

## 3.1 U.S. Treasury Yields

For my research, I consider the unsmoothed Fama-Bliss zero-coupon yields dataset at the monthly frequency, obtained from the CRSP unsmoothed Fama and Bliss (1987) forward rates. The dataset covers the period January 1970 to December 2009, for maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108 and 120 months. This dataset is the same as the one analyzed by van Dijk et al. (2014). For more details on constructing the yields, I refer to Diebold and Li (2006).

As yield curves evolve dynamically, they do not only have a cross-sectional (i.e. the variation across different maturities), but also a temporal (i.e. the dynamics over time), dimension. In Figure 2, I show the resulting three-dimensional surface for the U.S., with yields shown as a function of maturity, over time. The figure shows an important yield curve fact: the yield curve can take on a variety of shapes, such as humped, decreasing almost flat and so on. The Nelson and Siegel (1987) model can accurately approximate all these shapes. The long term trend is downwards, with short term interest rates near zero. In Table 1, I present descriptive statistics for yields at various maturities. Several well-known an key yield curve facts emerge. First, the yield curve is on average upward sloping and concave. Second, yields are (highly) persistent, as shown not only by the sizable 1-month autocorrelations but also by the large 12- and 30-month autocorrelations. Third, the persistence of yields increase with maturity. Fourth, short-term yields are more volatile than long-term yields.

The upper triangular part of the cross-correlation matrix of the yields are reported in Table 2. It shows that the cross-correlations between the maturities are high. This finding implies that PCA is applicable to explain a large part of the fluctuation in the yields with only a few number of factors. By applying PCA on the dataset I find that the first three factors explain nearly all of the variation in bond yields, which corresponds to the finding of Litterman and Scheinkman (1991). They find that three factors can explain most of the variation, particularly since 1978. Furthermore, they interpret these factors as the level, slope and curvature factor. Comparing the descriptive statistics of the empirical proxies for these factors[iii] in Table 1 with those of the three PCA factors in Table 3, I find that they exhibit similar characteristics. Moreover, the first PCA factor (level) is the most variable but the most predictable, due to its high persistence. Followed by the second PCA factor (slope) and third PCA factor (curvature). As seen in Figure 3, the PCA factors are effectively the data-based level, slope and curvature factor. This finding is important for two reasons. First, it substantiates the use of the DNS model to estimate and forecast the yields. Second, it argues for the inclusion of macro-economic information in the model as Diebold and Rudebusch (2013) state that this finding implies that the factors are likely to have specific macroeconomic fundamentals. For example, inflation is related to the level of the yield curve, and a clear business cycle rhythm is displayed by the slope factor.

[iii] I define the proxy for level as the longest maturity yield (120 months), for slope it is the difference between the 120-month yield and the 3-month yield, and for curvature it is two times the 24-month yield minus the 3-month and 120-month yields.

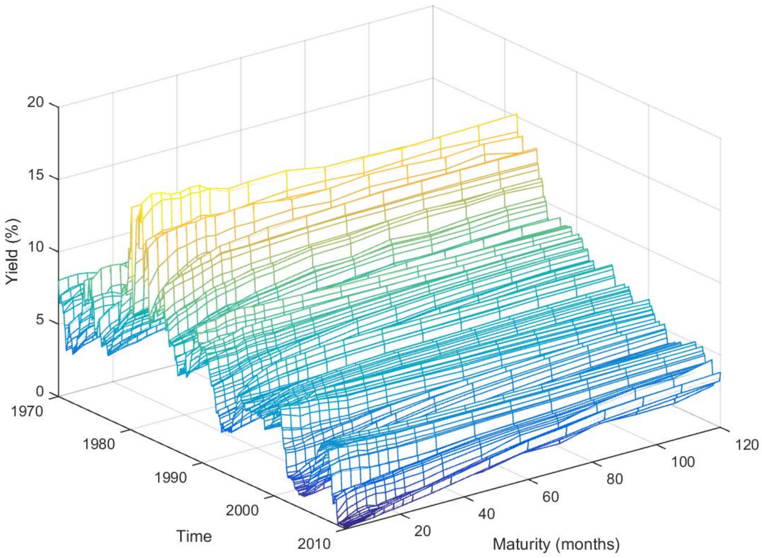



Figure 2: U.S. Treasury yield curves, 1970:01-2009:12. The sample consist of the unsmoothed Fama-Bliss zero-coupon monthly yield data, at maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108 and 120 months.

| Maturity (Months) | Mean | Std. dev. | Minimum | Maximum | $\hat{\rho}(1)$ | $\hat{\rho}(12)$ | $\hat{\rho}(30)$ |
|---|---|---|---|---|---|---|---|
| 3 | 5.766 | 3.071 | 0.041 | 16.019 | 0.979 | 0.749 | 0.411 |
| 6 | 5.969 | 3.098 | 0.150 | 16.481 | 0.980 | 0.763 | 0.442 |
| 9 | 6.083 | 3.089 | 0.193 | 16.394 | 0.981 | 0.771 | 0.467 |
| 12 | 6.166 | 3.053 | 0.245 | 16.101 | 0.981 | 0.777 | 0.483 |
| 15 | 6.253 | 3.029 | 0.377 | 16.055 | 0.982 | 0.785 | 0.504 |
| 18 | 6.324 | 3.009 | 0.438 | 16.219 | 0.983 | 0.792 | 0.522 |
| 21 | 6.387 | 2.990 | 0.532 | 16.173 | 0.983 | 0.797 | 0.537 |
| 24 | 6.418 | 2.943 | 0.532 | 15.814 | 0.983 | 0.799 | 0.550 |
| 30 | 6.512 | 2.878 | 0.819 | 15.429 | 0.983 | 0.808 | 0.570 |
| 36 | 6.600 | 2.832 | 0.978 | 15.538 | 0.984 | 0.814 | 0.586 |
| 48 | 6.756 | 2.755 | 1.019 | 15.599 | 0.984 | 0.822 | 0.614 |
| 60 | 6.852 | 2.671 | 1.556 | 15.129 | 0.985 | 0.832 | 0.636 |
| 72 | 6.964 | 2.638 | 1.525 | 15.108 | 0.987 | 0.842 | 0.653 |
| 84 | 7.026 | 2.573 | 2.179 | 15.024 | 0.987 | 0.841 | 0.666 |
| 96 | 7.069 | 2.536 | 2.105 | 15.052 | 0.988 | 0.850 | 0.673 |
| 108 | 7.095 | 2.519 | 2.152 | 15.114 | 0.988 | 0.853 | 0.677 |
| 120 (level) | 7.067 | 2.465 | 2.679 | 15.194 | 0.988 | 0.843 | 0.674 |
| Slope | 1.301 | 1.362 | -3.191 | 3.954 | 0.934 | 0.418 | -0.123 |
| Curvature | 0.003 | 0.863 | -2.174 | 2.905 | 0.877 | 0.441 | 0.130 |

Table 1: Descriptive statistics for monthly U.S. Treasury yields at different maturities over the period 1970:01-2009:12. The yields are constructed using the unsmoothed Fama-Bliss method. For each maturity I present the mean, standard deviation, minimum, maximum and the $j$th-order autocorrelation coefficients $\hat{\rho}(j)$ for $j = 1, 12$ and 30. In addition, I provide statistics for empirical proxies for the level, slope and curvature. I define the proxy for level as the longest maturity yield (120 months), for slope it is the difference between the 120-month yield and the 3-month yield, and for curvature it is two times the 24-month yield minus the 3-month and 120-month yields.

| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 1.000 | 0.997 | 0.994 | 0.990 | 0.984 | 0.980 | 0.976 | 0.972 | 0.964 | 0.957 | 0.944 | 0.934 | 0.924 | 0.917 | 0.909 | 0.904 | 0.902 |
| 6 | | 1.000 | 0.999 | 0.996 | 0.992 | 0.988 | 0.985 | 0.982 | 0.975 | 0.968 | 0.956 | 0.946 | 0.936 | 0.930 | 0.922 | 0.917 | 0.914 |
| 9 | | | 1.000 | 0.999 | 0.996 | 0.993 | 0.991 | 0.988 | 0.982 | 0.976 | 0.965 | 0.956 | 0.946 | 0.940 | 0.932 | 0.928 | 0.924 |
| 12 | | | | 1.000 | 0.999 | 0.997 | 0.995 | 0.993 | 0.988 | 0.983 | 0.973 | 0.965 | 0.955 | 0.950 | 0.942 | 0.938 | 0.934 |
| 15 | | | | | 1.000 | 0.999 | 0.998 | 0.997 | 0.993 | 0.990 | 0.981 | 0.974 | 0.965 | 0.960 | 0.953 | 0.949 | 0.945 |
| 18 | | | | | | 1.000 | 0.999 | 0.999 | 0.996 | 0.993 | 0.986 | 0.980 | 0.972 | 0.967 | 0.961 | 0.956 | 0.952 |
| 21 | | | | | | | 1.000 | 1.000 | 0.998 | 0.996 | 0.990 | 0.984 | 0.977 | 0.973 | 0.966 | 0.962 | 0.958 |
| 24 | | | | | | | | 1.000 | 0.999 | 0.997 | 0.992 | 0.987 | 0.981 | 0.976 | 0.971 | 0.966 | 0.962 |
| 30 | | | | | | | | | 1.000 | 0.999 | 0.996 | 0.992 | 0.987 | 0.983 | 0.978 | 0.974 | 0.969 |
| 36 | | | | | | | | | | 1.000 | 0.998 | 0.995 | 0.992 | 0.988 | 0.984 | 0.980 | 0.975 |
| 48 | | | | | | | | | | | 1.000 | 0.999 | 0.997 | 0.994 | 0.991 | 0.988 | 0.984 |
| 60 | | | | | | | | | | | | 1.000 | 0.999 | 0.997 | 0.995 | 0.992 | 0.988 |
| 72 | | | | | | | | | | | | | 1.000 | 0.999 | 0.997 | 0.995 | 0.991 |
| 84 | | | | | | | | | | | | | | 1.000 | 0.999 | 0.997 | 0.994 |
| 96 | | | | | | | | | | | | | | | 1.000 | 0.999 | 0.997 |
| 108 | | | | | | | | | | | | | | | | 1.000 | 0.999 |
| 120 | | | | | | | | | | | | | | | | | 1.000 |

Table 2: The upper triangular correlation-matrix of the yields from different maturities over the period 1970:01-2009:12.



| PCA factor | Mean | Std. dev. | Min. | Max. | $\hat{\rho}(1)$ | $\hat{\rho}(12)$ | $\hat{\rho}(30)$ |
|---|---|---|---|---|---|---|---|
| 1 | 0.000 | 4.074 | -8.009 | 13.039 | 0.985 | 0.816 | 0.583 |
| 2 | 0.000 | 0.598 | -1.337 | 1.648 | 0.948 | 0.491 | -0.018 |
| 3 | 0.000 | 0.171 | -0.493 | 0.695 | 0.863 | 0.366 | -0.062 |

Table 3: Descriptive statistics for the first three principal component analysis (PCA) factors of the unsmoothed Fama-Bliss zero-coupon yield dataset at maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 72, 84, 96, 108 and 120 months, for 1970:01-2009:12.

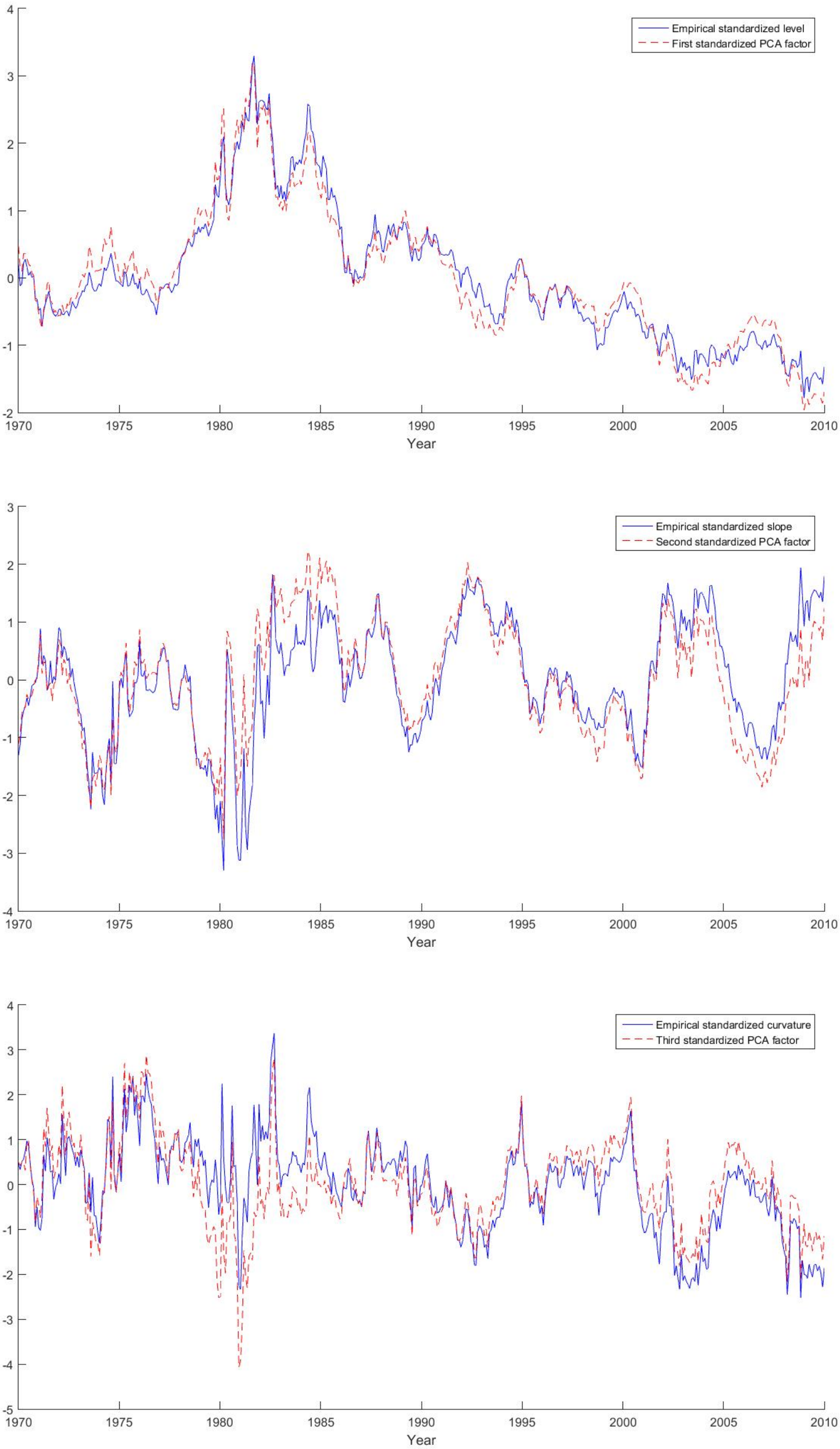



Figure 3: Empirical standardized level, slope and curvature factor against the first three standardized principal component analysis (PCA) factors for the period 1970:01-2009:12.

## 3.2 Indicators for the Macro-Economy

The macro factors that I include in the models are extracted from a set of 135 monthly variables covering the same period as the yield curve period. Moreover, I normalize the time series of observations on each variable separately to have mean zero and unit variance over the estimation window in order to rule out scale effects. These variables are obtained from the database of the Federal Reserve Bank of St. Louis (FRED).[iv] Moreover, these variables are categorized into eight groups by the FRED in an economically meaningful way. The groups are as follows: output and income; labor market; housing; consumption, orders and inventories; money and credit; prices; and stock market.

Attention should be paid to the timing of the macroseries relative to the yield series to prevent a potential look-ahead bias. Not coping with this, may lead to inaccurate results. Therefore, to make sure there is no information used that has not been released yet at the time when a forecast is being made, I consider two methods. First, I lag all macroseries by one month. Except for S&P variables, exchange rates and the federal funds rate, which are all monthly averages. Second, the included macro factors $\boldsymbol{f}_t$ in the models are actually the forecasted macro factors obtained using an AR(1) model, as shown in (5) and (16). Due to the fact that the macro factors are extracted from the full dataset, there might still be some look-ahead bias. However, this greatly facilitate the computational time.

The first macro factor extracted from the dataset, explains 56% of the variation in the panel for the full sample period. Furthermore, the second factor explains nearly 13%, and the third factor explains an additional 9%. Moreover, the first ten factors together explain a remarkable 95% of the variation. Following de Pooter et al. (2010), I regress the individual standardized series in the macro database of the FRED on each of the first four factors. Figure 11 in Appendix C shows the individual $R^2$'s of these regressions. This allows me to economically label the factors such that they can be interpreted more as representing meaningful economic variables instead of simply as results from applying a statistical procedure. The $R^2$ of the series in the real output and employment categories (groups 1 and 2), as well as groups 4 through 8, are quite high when they are regressed on the first factor. This factor can, therefore, be labelled as the business cycle or real activity factor. The second factor is mostly related to the unemployment part of the labor market group (group 2) and the housing category and could thus be labelled as the jobless claim factor. The third factor, although the correlations are much lower than for the first and second factor, closely resembles the federal funds related series in the money and credit category (group 6), which allows for the label of federal funds factor. Figure 4 presents graphically the interpretations through time-series plots of the first three factors together with real personal income, initial claims and the effective federal funds rate, respectively.

[iv] https://research.stlouisfed.org/econ/mccracken/fred-databases/

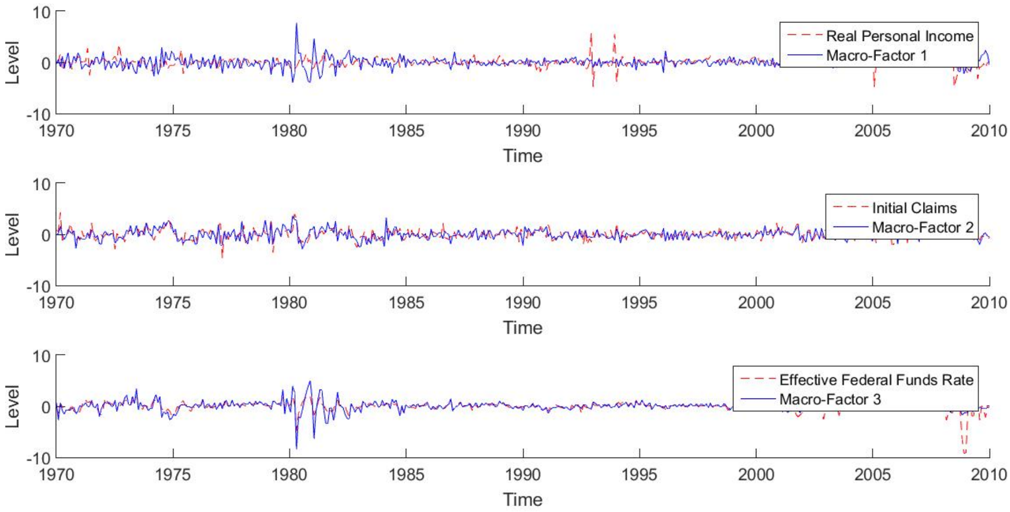



Figure 4: PCA factors versus individual macro series. The time-series plots of the first three factors (solid blue) together with the main individual macro series within the category to which the factor is most related (dashed red). The first factor is plotted with real personal income, the second factor with initial claims and the third factor with the effective federal funds rate.

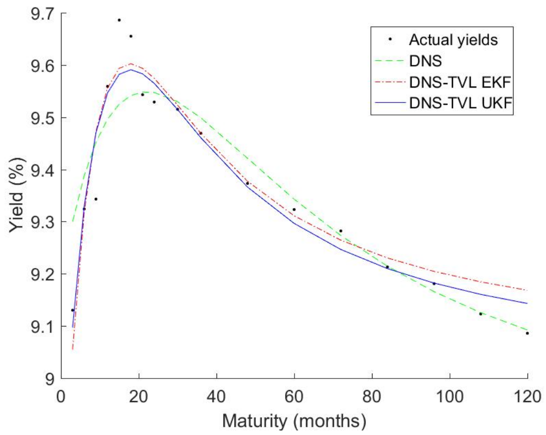
Various methods can be used to select the number of factors to be included in the model. De Pooter et al. (2010) evaluate how much variance each factor explains in the cross section of the macro series to choose the amount of factors. Similar to Diebold et al. (2006), they choose to include three factors in their models as the factors explains roughly 60% of the variance. Moreover, they find similar forecasting results when additional factors were included. Exterkate et al. (2013) use a dynamic scheme on choosing the number of factors. They conclude that the forecast accuracy for some models improve when the selection of factors is based on past predictive performance. I choose to follow the approach of Ludvigson and Ng (2009). That is, I use information criteria to evaluate whether a factor should be included in the model.

# 4 Empirical Results

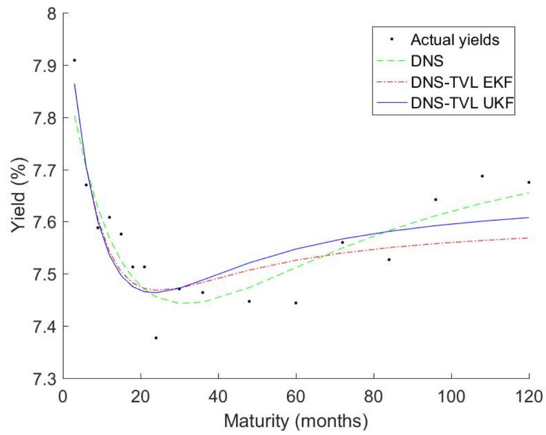
This section elaborates on the empirical results obtained for the DNS and its extensions as described in the methodology in twofold. First, I discuss the differences between the in-sample fit of the DNS, DNS-TVL-EKF and DNS-TVL-UKF. Furthermore, I put more emphasis on the dynamics of the loading parameter as obtained from the EKF and the UKF. In the second part of the results section, I elaborate on the out-of-sample forecasting performance of all models. Additionally, I consider various subsamples in which I evaluate this.

## 4.1 In-Sample Analysis

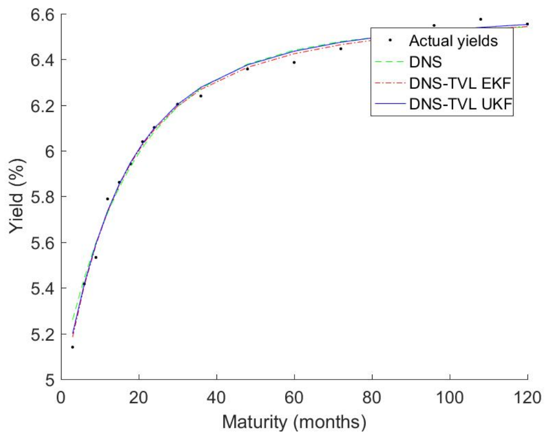
As the in-sample fit of the DNS and its extensions have been studied thoroughly by various authors, I choose to focus only on the models for which there is little information. That is, I study the effect of having a time-varying loading parameter in the DNS model on the in-sample fit. Moreover, I focus on the differences between the EKF and the UKF in particular. Figure 5 shows the actual yields and the fitted yield curves obtained from the DNS, DNS-TVL-EKF and DNS-TVL-UKF for the same dates that Diebold and Li (2006) use in their analysis. Overall,



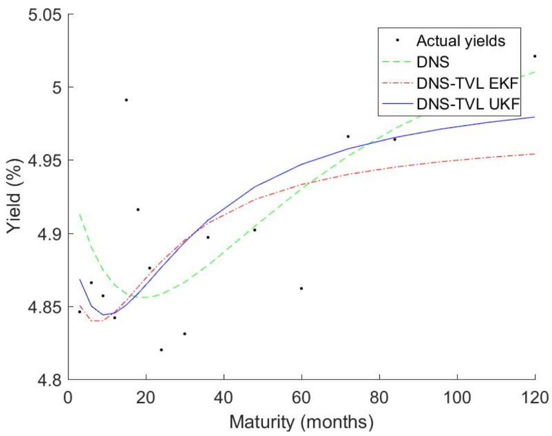
(a) 3/31/1989



(b) 31/7/1989



(c) 30/5/1997



(d) 31/8/1998

Figure 5: Actual yields against the fitted yield curves obtained from the DNS, DNS-TVL-EKF and DNS-TVL-UKF. This figure shows the differences in the fit of the yield curves from the considered models for selected dates.

the three models provide a good fit of the yields and are capable of replicating a variety of shapes. However, on the dates the yields are dispersed, the differences between the in-sample fit of the models become clear. In contrast to the DNS, the DNS-TVL-EKF and DNS-TVL-UKF are more accurate in fitting the short end of the yield curve. This finding confirms the statement made by Diebold and Rudebusch (2013) that allowing for a time-varying loading parameter may provide a better fit of the yield curve at short-term maturities. However, it seems to come at a cost i.e. the fit of the estimate yield curves from the DNS-TVL models are slightly less accurate than from the basic DNS model at longer maturities.

These findings may only hold for these selected dates, and therefore, to make general statements regarding the fit of the models, I provide the sample means and standard deviations of filtered errors in Table 4. The filtered errors are defined as the difference between the actual yields and the filtered estimate of the curve obtained from the different Kalman filters. By evaluating the filtered means of each model individually, I conclude that the 3-month rate is difficult to fit for all three models as the highest filtered mean can be found at this maturity. Moreover, the highest standard deviations can be found at this maturity as well, which corresponds to the finding of Koopman et al. (2010), among others. These findings could stem from the fact that the yield at this maturity is quite volatile, making it harder for the filters to estimate the Nelson-Siegel factors and, as a result, the yield itself. Comparing the filtered errors of the three models per maturity, I find that allowing for a time-varying loading parameter in the DNS model generally leads to an improvement of the in-sample fit. Especially for the short end of the yield curve, which corresponds with the findings for the selected dates. Furthermore, the table shows a decrease of the filtered error mean at the 120-month rate. It would be interesting to find out what the results are for maturities longer than ten years. However, this is not in the scope of this study. Interestingly, the errors of the models with a time-varying loading parameter are slightly more volatile than those of the standard DNS for all maturities except

| | DNS | | DNS-TVL-EKF | | DNS-TVL-UKF | |
|---|---|---|---|---|---|---|
| Maturity | Mean | Std. dev. | Mean | Std. dev. | Mean | Std. dev. |
| 3 | -0.140 | 0.271 | -0.068 | 0.161 | -0.076 | 0.143 |
| 6 | -0.027 | 0.098 | 0.005 | 0.068 | 0.001 | 0.055 |
| 9 | 0.002 | 0.062 | 0.016 | 0.099 | 0.011 | 0.093 |
| 12 | 0.006 | 0.067 | 0.011 | 0.104 | 0.005 | 0.096 |
| 15 | 0.020 | 0.079 | 0.022 | 0.094 | 0.014 | 0.085 |
| 18 | 0.023 | 0.069 | 0.023 | 0.081 | 0.014 | 0.070 |
| 21 | 0.023 | 0.063 | 0.024 | 0.077 | 0.014 | 0.066 |
| 24 | -0.004 | 0.050 | -0.003 | 0.076 | -0.014 | 0.068 |
| 30 | -0.015 | 0.044 | -0.012 | 0.076 | -0.024 | 0.074 |
| 36 | -0.016 | 0.050 | -0.012 | 0.077 | -0.025 | 0.081 |
| 48 | -0.003 | 0.070 | 0.003 | 0.083 | -0.010 | 0.095 |
| 60 | -0.015 | 0.069 | -0.006 | 0.082 | -0.019 | 0.098 |
| 72 | 0.016 | 0.086 | 0.026 | 0.090 | 0.014 | 0.103 |
| 84 | 0.016 | 0.074 | 0.027 | 0.094 | 0.015 | 0.097 |
| 96 | 0.011 | 0.044 | 0.023 | 0.082 | 0.012 | 0.072 |
| 108 | -0.003 | 0.111 | 0.012 | 0.106 | 0.001 | 0.099 |
| 120 | -0.063 | 0.176 | -0.045 | 0.150 | -0.057 | 0.145 |

Table 4: Filtered errors of the DNS, DNS-TVL-EKF and DNS-TVL-UKF. The table reports the mean and standard deviation of the filtered errors for each maturity. The filtered errors are defined as the difference between the actual yields and the filtered estimate of the curve obtained from the Kalman filter, respectively, the extended- and the unscented Kalman filter.

for the really short end of the yield curve (3- and 6-month rate) and the long end of the yield curve (108- and 120-month rate). A possible explanation could be that the overall fit of both the DNS-TVL-EKF and the DNS-TVL-UKF are more accurate than that of the DNS, but at some dates the estimates of these two deviate much from the actual yield. Digging deeper into the difference between the results obtained from the two nonlinear Gaussian filtering methods, I find that the EKF filters the bonds with intermediate maturity (two to five years) more accurately than the UKF does as the means of the filtered errors are lower here. However, for the remaining maturities the estimates from the UKF are more accurate than that of the EKF. The same structure is reflected in the standard deviations; for the intermediate maturities the filtered errors of the EKF are less volatile, while the filtered errors of the UKF are more stable for the other maturities. An analysis of the estimated latent factors could possibly provide an explanation for this. In general, the differences between the estimated factors are marginal, with an exception for the curvature factor. For this factor, I evaluate the estimated error defined as the data-based curvature factor minus the estimated curvature factor. Figure 6 shows that the UKF generally fits the curvature factor more accurately than the EKF as the errors are closer to zero for the UKF. Moreover, the negative errors indicate that the EKF tends to overestimate the data-based curvature factor, even more than the UKF does for some periods (e.g. 1976-1977). This means that the EKF either underestimates the 24-months yield or overestimates the yields at the ends of the yield curve (3-months yield and 120-months yield) by definition. As Table 4 shows that the mean of the filtered error of the 24-months yield is approximately zero, I therefore conclude that the EKF tends to overestimate the ends of the yield curve. This may stem from the fact that the linear approximation in the EKF, to calculate the statistics of this factor, is inferior to the unscented transformation in the UKF.

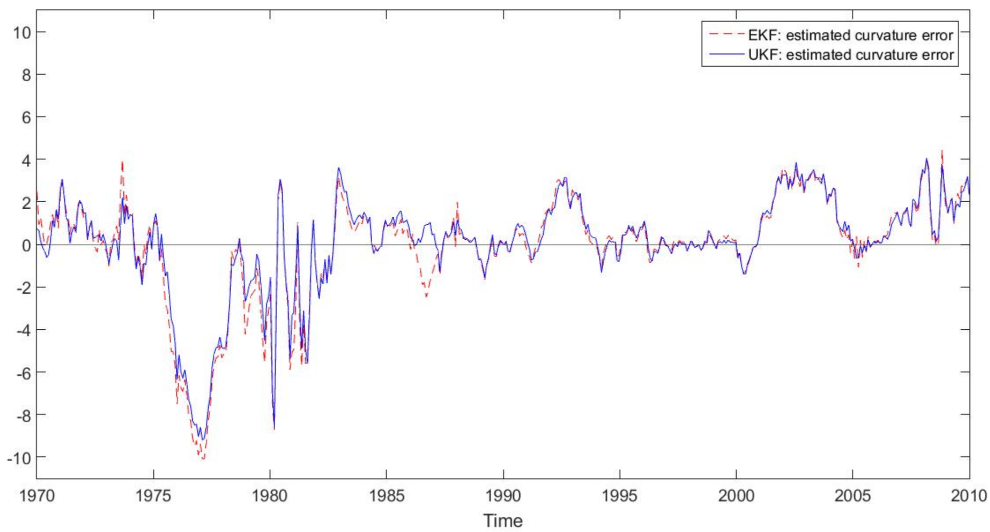



Figure 6: The (absolute) estimated curvature errors. The figure shows the (absolute) estimated curvature errors, which is defined as the (absolute value of the) data-based curvature factor minus the estimated curvature factor. The estimated curvature factors stems from the EKF and the UKF.

In Figure 7, I compare the filtered latent factors obtained from both the DNS-TVL-EKF and DNS-TVL-UKF with their data-based proxies on the left hand side. The figure shows that the estimated factors from both models describe the data-based proxy equally well. Furthermore, the filtered estimates of the loading parameter are presented. These estimates are particularly high in the early 1980s with a sample average equal to approximately 0.09. Moreover, the peaks and bottoms of the estimated loading parameter as obtained from the DNS-TVL-EKF are respectively higher and lower than from the DNS-TVL-UKF. On the right hand side, the difference between the fit of the models are highlighted. The most notable deviation is found in the late 1970s, during this period the differences of the latent factors show a similar shape. More specifically, for the slope and curvature factor a half head-shoulder-pattern is shown, with the head starting a little after 1975. For the level factor this shape is mirrored in the zero line. Interestingly is the fact that the difference of the estimated loading parameter is near zero for the same period. However, as there are no other periods where the same phenomenon can be seen, I am not able to make any statements regarding the correlation of these findings. For the other periods, the difference between the estimated factors of both models is marginal.

I also evaluate the performance of the models by considering the log likelihood, the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC) and the likelihood ratio (LR) test for model improvement. These values are presented in Table 5. First, I compare the DNS model with the two DNS-TVL models. As the standard model is nested in the models with a time-varying loading parameter, I can use the standard likelihood-ratio (LR) test. Under the null hypothesis of standard DNS, $LR = 2[\ell(\theta_{DNS-TVL}) - \ell(\theta_{DNS})] \sim \chi^2(3)$. The LR test statistics are presented in the right column of the table. Moreover, the associated $p$-values are both less than 0.0001, so I formally reject the restrictions imposed by the DNS model. In other words, allowing for a time-varying loading parameter results in a significant improvement of the in-sample fit. This is confirmed by the AIC, the BIC and LR values and corresponds to the finding of Koopman et al. (2010). Second, consider the fit of the filtered estimates as obtained from the EKF versus that of the UKF. The models are non-nested but contain equal numbers of parameters. Therefore, I compare their log likelihoods directly, with the clear result that the DNS-TVL-UKF is dominated by the DNS-TVL-EKF. Overall, the estimates of the yields as obtained by the EKF are more accurate than that of the UKF. Although, there are some differences when the fit is evaluated per maturity. That is, the EKF is able to fit the intermediate part of the yield curve better than the UKF (with the highest decrease in the mean of the filtered error of nearly 80% for the 24-months rate), while the latter provides remarkably more accurate estimates at both ends of the yield curve. Moreover, the biggest drop in the mean of the filtered error is approximately 96% for the 108-months rate.

| Model | Log likelihood | AIC | BIC | LR-test statistic |
|---|---|---|---|---|
| DNS | 5613.0 | -11174.0 | -11065.5 | - |
| DNS-TVL-EKF | 5811.3 | -11564.6 | -11443.6 | 396.6** |
| DNS-TVL-UKF | 5688.8 | -11319.6 | -11198.6 | 151.6** |

Table 5: Log likelihood, AIC, BIC and LR test statistics. The table reports the log likelihood, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and the likelihood ratio (LR) test statistics of comparing the standard DNS model to the models with a time-varying loading parameter (DNS-TVL-EKF and DNS-TVL-UKF). An asterisk (*) denotes significance at the 5% level or less and two asterisks (**) denote significance at the 1% or less.

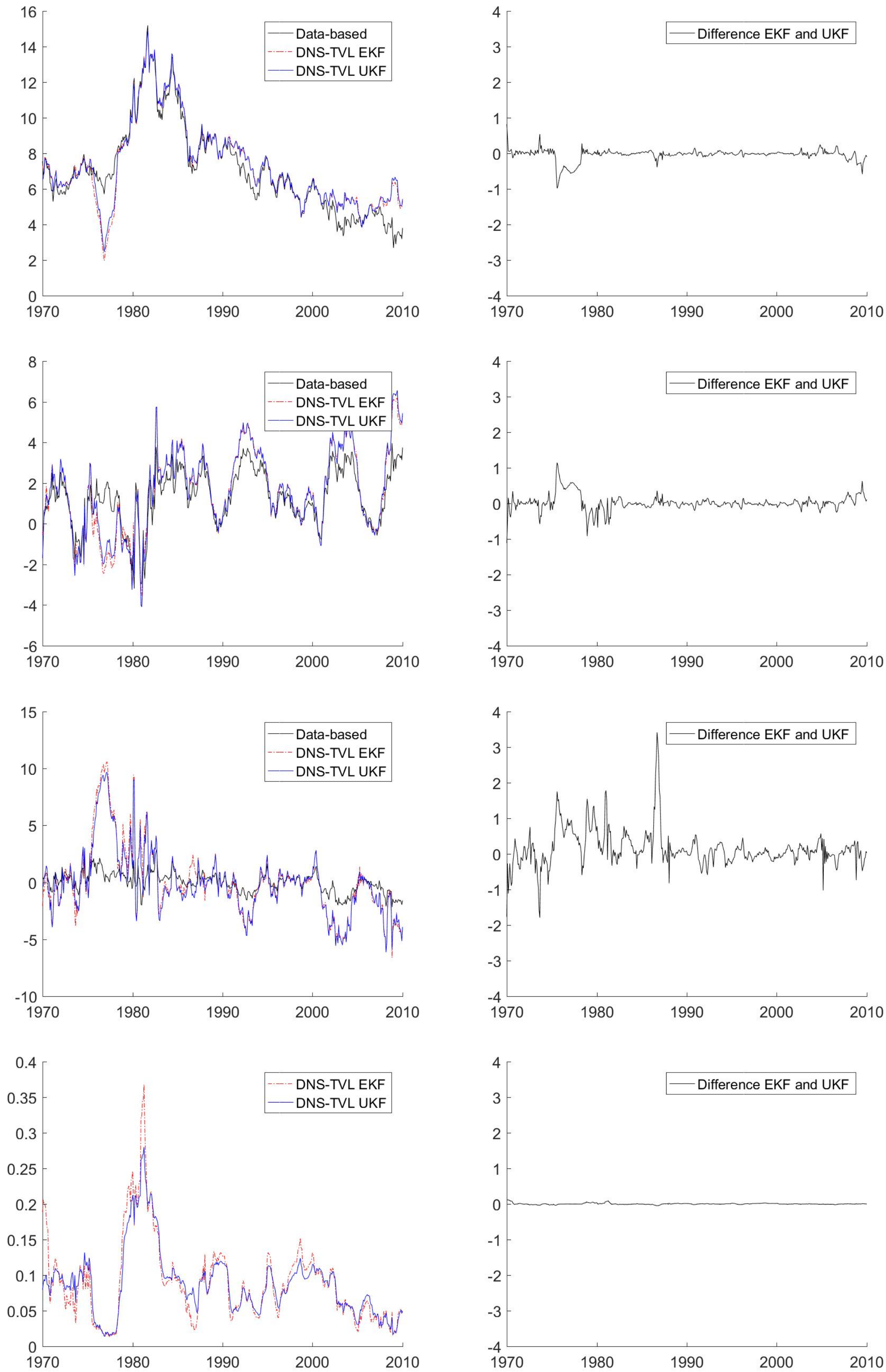



Figure 7: Latent factors. The figure shows the level, slope, curvature and the time-varying loading parameter as obtained from the DNS-TVL-EKF (red) and DNS-TVL-UKF (blue) on the left side. Moreover, the proxies of the first three factors are shown as well. On the right hand side, the difference between the estimated latent factors of the DNS-TVL-EKF and DNS-TVL-UKF are highlighted.

## 4.2 In-Sample Analysis: Factor Loading

Most of the time, the estimated level, slope and curvature factors corresponding to the estimated $\lambda_t$ as obtained from the DNS-EKF-TVL and DNS-UKF-TVL are able to accurately fit the data-based factors. However, for some values of the estimated time-varying loading parameter, a huge misfit is observed. The main problem is the fact that the factors are no longer uniquely identified due to multicolinearity when the factor loading takes on extreme values, see De Pooter (2007). A deeper analysis of the factor loadings, particularly the slope and curvature loading, provides an insight of the effect from the time-varying loading parameter on the interpretation of these factors. For this analysis, I consider the extreme values of $\lambda_t$. Let $\lambda_t^{EKF}$ and $\lambda_t^{UKF}$ denote the estimated time-varying loading parameter as obtained by the DNS-TVL-EKF and DNS-TVL-UKF, respectively. Figure 8 shows the slope and curvature loadings for the Diebold-Li (2006) fixed value of $\lambda_t$ and different estimated loading parameters. I look at the problem from a theoretical point-of-view and study the obtained estimates. Consider the factor loadings at their limits

$$\lim_{\lambda_t \downarrow 0}\left[\frac{1-\exp(-\lambda_t\tau)}{\lambda_t\tau}\right] = 1; \qquad \lim_{\lambda_t \downarrow 0}\left[\frac{1-\exp(-\lambda_t\tau)}{\lambda_t\tau} - \exp(-\lambda_t\tau)\right] = 0, \tag{46}$$

$$\lim_{\lambda_t \to \infty}\left[\frac{1-\exp(-\lambda_t\tau)}{\lambda_t\tau}\right] = 0; \qquad \lim_{\lambda_t \to \infty}\left[\frac{1-\exp(-\lambda_t\tau)}{\lambda_t\tau} - \exp(-\lambda_t\tau)\right] = 0, \tag{47}$$

where the loading on the left-hand side corresponds to the slope factor, while the loading on the right-hands side corresponds to the curvature factor. For very small values of $\lambda_t$, (46) implies that the curvature factor is non-identified. Furthermore, the level and slope factor can not be differentiated from each other as the loadings are equal. Put differently, they are jointly identified. This is the case in the period 1976 - 1978, in which the estimated $\lambda_t$'s attain their minimums. The minimum value of $\lambda_t^{EKF}$ is equal to 0.013 at June 1977, and the minimum value for $\lambda_t^{UKF}$ is equal to 0.014 at November 1976. These values are lower than the Diebold-Li (2006) value of 0.0609, which implies that the factor loadings of the slope and curvature factor decay to zero at a slower pace. Moreover, this results in the factor loading of the curvature factor peaking near the 132-months maturity. Thus, as shown in Figure 8, the factors take on very different roles in the fit of the model due to the shape of the corresponding loadings (indicated as 'Slope 2' and 'Curvature 2'). That is, in contrast to the Diebold-Li (2006) interpretation, for this period, the slope factor acts as both the short- and medium-term factor, while the curvature factor acts as a long-term factor. As a result, the level factor, which is still considered as a

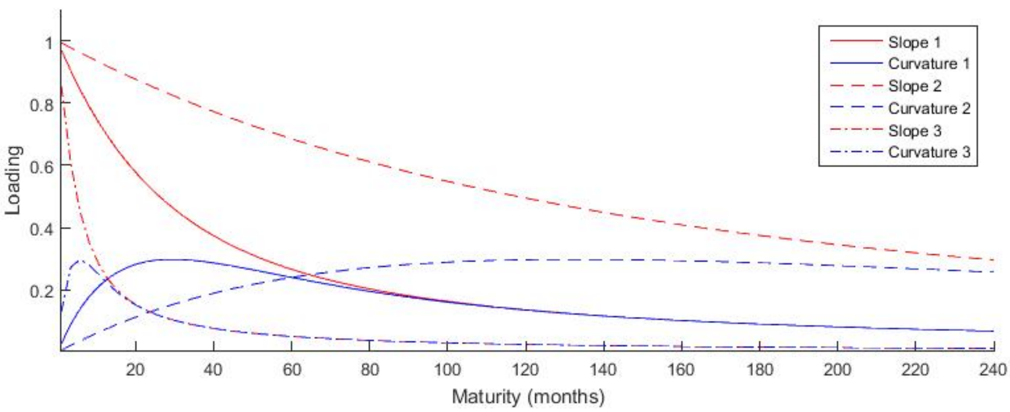



Figure 8: Slope and curvature loadings for different values of lambda. This figure shows the slope and curvature loadings as a function of maturity for $\lambda_1 = 0.0609$, $\lambda_2 = 0.0135$ and $\lambda_3 = 0.3238$.

long-term factor, is allowed to fit other areas of the yield curve (Christensen et al., 2009). The fact that the curvature factor is non-identified, and the slope and level factors are jointly identified due to the change of the factor loadings for this period clearly explains the significant misfit of the estimated level, slope and curvature factors on their proxies as shown in Figure 7. More specifically, this could be a possible explanation for the peaks with offsetting values in the estimated slope and level factor (in this period, I find a value of 2 for the estimated level factor, while at the same time a value of -2 is found for the estimated slope factor). This corresponds to the findings of Gimeno and Nave (2006), and explains it.

For very large values of $\lambda_t$, (47) implies that both the curvature and slope factors are non-identified. The estimated $\lambda_t$'s attain their maximums in the period 1980 - 1982. More specifically, the maximum of both $\lambda_t^{UKF}$ and $\lambda_t^{EKF}$ is found at April 1981, with corresponding values equal to 0.2799 and 0.3676, respectively. These values are nearly five to six times the value as used in Diebold and Li (2006), indicating that the slope and curvature factor decay to zero at a more rapid pace. Moreover, the peak of factor loading of the curvature factor is found near the 6-month maturity. Due to the identification problem in this period, the slope and curvature factor can not be differentiated from each other. This stems from the fact that the loadings for both factors suggest a short-term factor interpretation as the short-term yields load heavily on these factors, as shown in Figure 8. More specifically, the loadings are indicated as 'Slope 3' and 'Curvature 3'. This explains why the estimated slope factor from both filtering methods is able to accurately fit the data-based slope factor, but, the estimated curvature factor is not able to do so for his data-based counterpart.

### 4.3 Out-of-Sample Analysis

I construct one-, six- and twelve-months ahead forecasts for yields at all maturities using a recursive procedure. First, I fit the models from January 1970 to December 1984 and use the parameter estimates to construct the forecasts for the different horizons; then, one month of data is added, I re-estimate the models, and construct another set of forecasts. The largest estimation sample for the one-month ahead forecasts ends in November 2009 (300 forecasts), for the six-month horizon it ends in June 2009 (295 forecasts) and at a horizon of twelve months it ends in December 2008 (289 forecasts). The optimal forecast of the $\tau$-maturity yield made at time $t$ for time $t+h$ is equal to the conditional expectation

$$\hat{\boldsymbol{y}}_{t+h|t} = E[\boldsymbol{y}_{t+h}|\boldsymbol{\Psi}_t] = \boldsymbol{H}(E[\boldsymbol{x}_{t+h}|\boldsymbol{\Psi}_t]), \tag{48}$$

where recursive iteration (and i.i.d. innovations) imply that the conditional $h$-month ahead forecast of the state vector is given by

$$E[\boldsymbol{x}_{t+h}|\boldsymbol{\Psi}_t] = \sum_{i=0}^{h-1} \boldsymbol{K}^i \boldsymbol{C} + \boldsymbol{K}^h \boldsymbol{x}_t, \tag{49}$$

where the matrices of parameters are defined as in the methodology section. In examining the forecast performance, I am interested in two broad comparisons. First, to what extent do the forecasts improve after implementing the extensions in the DNS model, and second, what is the effect of a different filtering technique on the accuracy of the forecasts. I bring these questions into sharper focus by using the root mean squared prediction error (RMSPE) defined as

$$RMSPE(h,\tau) = \sqrt{\frac{1}{T}\sum_t [\hat{y}_{t+h|t}(\tau) - y_{t+h}(\tau)]^2}, \tag{50}$$

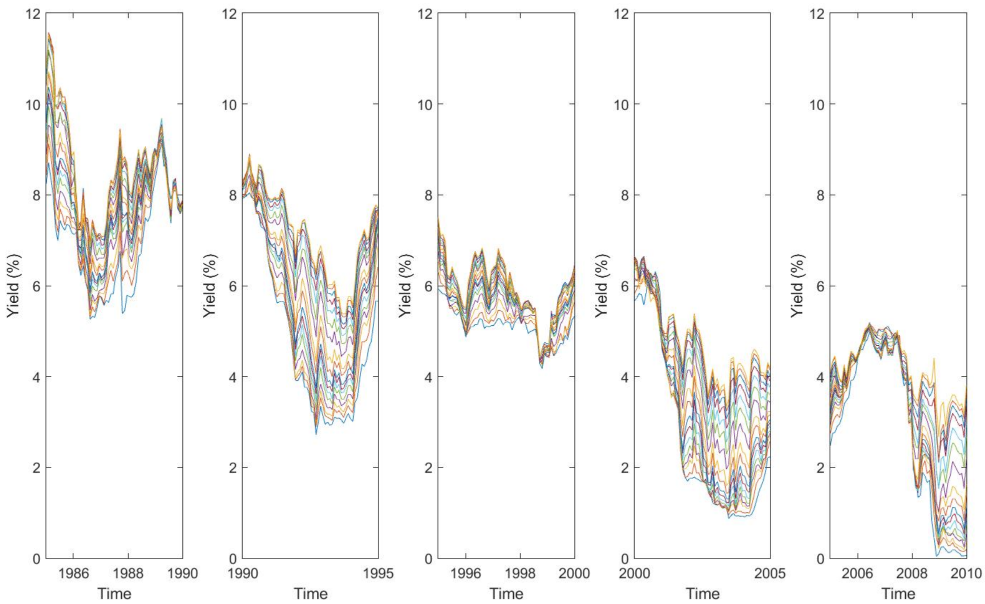



Figure 9: Subsamples for the out-of-sample analysis. This figure shows five subsamples of the yield on which the out-of-sample analysis is conducted. The subsamples cover the periods 1985:01-1989:12, 1990:01-1994:12, 1995:01-1999:12, 2000:01-2004:12 and 2005:01-2009:1, respectively.

to compare the models. More specifically, I use the DNS model as the benchmark and report the ratios of the RMSPE of the different models and the benchmark. By doing this, it is more convenient to evaluate whether the extensions provide more accurate forecasts than the DNS. In addition to the assessment over the full out-of-sample period, I evaluate the forecasting results for five subsamples, similar to Exterkate et al. (2013), each covering a period of five years. More specifically, the first subsample starts at January 1985 until December 1989, which contains the interest rate increase from 1986 to 1989 by the Federal Reserve as a response to the increasing inflation in the early 1980s. This restrictive monetary policy weakened the economic growth, and in combination with a loss of consumer and business it was one of the determinants of the following small recession in the early 1990s. The short recession started in July 1990 until March 1991, and is contained in the second subsample (January 1990 until December 1994). Even though the economic growth had returned in 1991, the unemployment rate kept rising through June 1992. As a response to this, the Federal Reserve lowered the interest rates to reduce this. In fact, the U.S. entered into its longest period (10 years) of economic expansion after the early 1990s recession, in which employment growth attained values of approximately 22,5%. This period is part of the third subsample, which starts at January 1995 and lasts through December 1999. Moreover, this period closely corresponds to the forecasting period as considered by Diebold and Li (2006). The long economic expansion was fueled by the Dot Com bubble, among others, and ended in the early 2000s due to the collapse of this bubble. In contrary to the long period of economic growth, the recession was brief and lasted from March 2001 until November 2001. In addition to this short recession, the fourth subsample (January 2000 until December 2004) contains a rare type of yield curve at the end of 2001; the inverted yield curve. This type of yield curve is often seen as a predictor of lower interest rates in the future as bonds with long maturities are being demanded, sending the yields down. Finally, the fifth subsample last from January 2005 until December 2009 containing the 2008-2009 crisis period. Moreover, as seen in Figure 9, the yields exhibit a lower volatility in the third subsample as compared to the other subsamples. I find similar results for the periods in which the yields are relatively highly volatile. Therefore, for the sake of brevity, I report only the results of the most recent volatile period, i.e. the fifth subsample.

To asses whether the differences in the RMSPE are significant, the statistic of Diebold and Mariano (1995) can be used. However, statistical issues may arise when the to be compared pair of prediction models are nested (West, 2006). This is the case in the thesis as the DNS model is nested in its extensions. Giacomini and White (2006) show that the Diebold-Mariano test statistic is valid if and only if a rolling scheme with a finite observation window is used. However, as I use a recursive scheme, I should consider other methods. Clark and McCracken (2005) introduce an approach to cope with this, which is based on the bootstrap bias adjustment of Kilian (1998). Similar to van Dijk et al. (2013), I use this bootstrap method to assess the statistical significance of the differences in the RMSPE. The steps are as follows:

(1) I estimate the factors $\boldsymbol{\beta}_t$ using the Kalman filter and then fit an AR(1) model to these, using the bootstrap bias adjustment of Kilian (1998).[v] Furthermore, I store the residuals.
(2) Now, I create bootstrap samples of the factors by resampling the residuals with replacements, again using Killian's bootstrap bias adjustment. In addition, I resample from the maturity-specific (or idiosyncratic) errors in the yields (as in equation (2)).
(3) For each bootstrap replication, I use the bootstrapped data to recursively estimate the restricted (DNS) and unrestricted forecasting models (extensions of the DNS). Subsequently, I compute the ratio of the recursive out-of-sample RMSPE using the several extensions of the DNS relative to that from the DNS.

The ratios are effectively the Diebold-Mariano test statistics (I, however, use the RMSPE instead of MSPE). I perform these steps for 500 bootstrap samples, giving the bootstrap approximation to the null distribution of the Diebold-Mariano test statistic. Furthermore, the test is regarded as a one-sided test. Similar to van Dijk et al. (2013), I expect the null distribution of the test statistic to be centered a bit above unity, as the predictive performance of the other models should be worse under the null hypothesis that the standard DNS is correctly specified.

Before I continue with the evaluation of the forecasts as obtained from the different models, I first elaborate on the number of macro-factors that I include in the FADNS, SFADNS-EKF, and SFADNS-UKF. After using the full sample to fit the FADNS for different numbers of macro-factors, I compare the obtained information criteria to choose the appropriate amount of factors to be included. Employing the same approach as Ludvigson and Ng (2009), I find that I should include the first two factors in the mentioned models. Moreover, these two factors explain roughly 69% of the variation in the panel for the full sample period, which is more or less of the same magnitude of percentage of variation explained as in De Pooter et al. (2010). Adding more factors worsens the in-sample fit, which may be the result of overfitting. However, keep in mind that a good in-sample fit does not inevitably lead to a good out-of-sample performance.

Table 6 reports the ratios of the RMSPEs of the extensions of the DNS relative to the DNS for out-of-sample forecasts at a horizon of one, six and twelve months for the full period. Similarly, the ratios for the stable period (January 1995 until December 1999) and the period of relatively high volatility (January 2005 until December 2009) are reported in Table 7 and 8, respectively. I start with the evaluation of the one-month ahead forecasts. On average, the extensions of the

[v] For more details and the exact procedure of the bootstrap bias adjustment, I refer to his paper.

DNS provide slightly more accurate forecasts than the DNS for the short-end of the yield curve when the full out-of-sample period is considered. In particular, the RMSPEs of the forecasts of the three-months yield from all models are lower than that of the benchmark. The gain in predictive accuracy ranges from 0.7% of the FADNS with two macro-factors to nearly 14% when the SFADNS is used to construct forecasts. Moreover, only for this maturity the differences are significant. In addition, for the really short-term yields, it holds that the accuracy of the forecasts increases when more extensions are employed. Put differently, the RMSPE ratios of the SFADNS for both filtering techniques are lower than that of the single extension models (FADNS, DNS-GARCH and DNS-TVL) for these maturities.[vi] However, for the yields with an intermediate maturity the results are the opposite i.e. the RMSPE ratios of the SFADNS are higher. Furthermore, the forecasts of these intermediate-term yields from the extensions of the DNS are less accurate than the standard DNS, but the degree of increase in the ratio is little with a loss in predictive accuracy of 3.5% for the SFADNS for the three-years yield being the maximum. For the long-end of the yield curve, the difference between the out-of-sample performance of the standard DNS and its extensions is marginal.

Now, consider the subsample covering the period January 1995 until December 1999, the stable period. For the short-end of the yield curve, the extended models are able to provide more accurate forecasts than the DNS in 75% of the cases. Even though the gain in predictive accuracy of a single model over the DNS is not that high as for the full out-of-sample period, the improvements are consistent and on average equal to 4% with nearly 10% being the maximal drop in the RMSPE when the DNS-TVL-UKF is used to construct forecasts for this period and these maturities. Similar to the findings for the full out-of-sample period, on average there is only a negligible difference between the forecast accuracy of the DNS and its extensions for the intermediate- and long-term yields. For the fifth subsample, the relatively highly volatile period, the results are in favor of the extensions of the DNS. The RMSPE ratios are either lower than one, indicating an outperformance over the DNS, or nearly equal to one with only a few outliers. Moreover, the highest gain in predictive accuracy is found at the short-end of the yield curve. In particular, the forecasts for the three-month yields of the considered extensions dominate the forecasts of the standard DNS, with a drop in the RMSPE ranging from 2% for the FADNS with two macro factors to almost 28% for the SFADNS. Only the forecasts for the long-end of the yield curve from some models are significantly more accurate than those of the DNS for both periods, as shown in the tables.

Overall, at a forecast horizon of one-month, the extensions of the DNS provide more accurate forecasts than the DNS for the short-end of the yield curve, particularly for the period January 2005 until December 2009. Moreover, for this period and maturities, nearly 81% of the cases the forecasts from the DNS are dominated. For the intermediate- and long-term yields, the forecasts provided by the DNS and its extensions do not deviate too much from each other as the ratios are near 1. There are some exceptions, with a gain of predictive accuracy of nearly 5%, in case the DNS-TVL-EKF is used to construct the forecast of the 6-year yield in the relatively highly volatile period.

The accuracy of the six-month ahead forecasts from the extensions of the DNS are relatively worse than that from the DNS for all maturities when the full out-of-sample period is considered. Moreover, I find a loss in predictive accuracy of nearly 10% when SFADNS is used to construct forecasts for the long-end of the yield curve. Contrariwise, the subsample results are more positive. Considering the stable period, in 70% of the cases the extensions of the DNS are

[vi] In subsection 4.4, I analyze why this holds.

able to provide more accurate six-month ahead forecasts of the short-term yield than the DNS can. The drop of the RMSPE ranges from 0.7% for the SFADNS to roughly 6% for the DNS-TVL-UKF. For yields with a longer maturity, only the FADNS with two macro-factors and the DNS-GARCH perform more or less the same as the standard DNS at providing accurate six-month ahead forecasts for this subsample. For the period January 2005 until December 2009, I find that the percentage of cases for which the forecasts of the short-end of the yield curve from the DNS are dominated by the forecasts from its extensions is roughly the same (67%) to that for the first subsample period. However, for this period the gain in predictive accuracy ranges from 0.9% for the FADNS with two macro-factors to 8% for the SFADNS. Moreover, only the SFADNS is able to outperform the DNS for yields with an intermediate-maturity. For the long-term yields, the DNS is unbeatable.

Summarizing, when the six-month forecast horizon is considered, the forecasts of the DNS are more accurate than that of its extensions for the full out-of-sample period. However, for the subsamples, the results are more promising. More specifically, the extensions of the DNS outperform the DNS in terms of RMSPE ratios for the short-term yields, with only a few models providing more accurate forecasts for yields with a longer maturity. In particular, this statement holds for the FADNS with two macro-factors and the DNS-GARCH for the period where the yields are relatively stable, as well as for the SFADNS when the yields are highly volatile. However, none of these forecast are significantly more accurate than the forecasts from the standard DNS.

For a forecast horizon of twelve months, I find that none of the models are able to provide forecasts more accurate than those of DNS when the full out-of-sample period, and the volatile periodsare considered. Remarkably, for the period in which the yield is quite stable, the DNS-GARCH outperforms the DNS for ten out the seventeen maturities; attaining a gain in predictive accuracy of 5% for the short-term yields. For the remaining maturities, the maximum increase in the RMSPE relatively to the DNS is equal to approximately 4% for the yield at a maturity of 10 years. However, the gain in predictive accuracy is not significant.

In general, irrespective of the forecast horizon, I find that it is quite hard to construct more accurate forecasts than those obtained from the standard DNS, with 51% being the highest percentage of obtaining a RMSPE ratio less than 1 at a forecast horizon of 1 month for the period January 2005 until December 2009. Moreover, only the forecasts from some models for the ends of the yield curve are more accurate than those from the DNS. In addition, the results are highly sensitive to the period that is considered. For periods when the yield is stable, the DNS-GARCH outperforms the DNS at forecasting remarkably well for all forecasting horizons. This result corresponds to the finding Hautsch and Yang (2010), who incorporate stochastic volatility via the transition equation instead. As a reason, they argue that the standard DNS has a higher forecasting uncertainty for periods of low-volatility, which stems from the fact that the ignored stochastic volatility and parameter uncertainty in periods of high-volatility spread to periods when the yields are more stable. When the yields are relatively highly volatile, the SFADNS performs extraordinary well as its forecasting accuracy dominates that of the DNS for approximately eleven out of the seventeen maturities when a forecast horizon of one- and six-months is considered. For the same volatile period at the twelve-months forecasting horizon, however, there is no gain in predictive accuracy when an extension is considered. In addition, I find no consistency regarding the increase (or decrease) of the RMSPE ratio when more flexibility is allowed in the model and observe that this is also highly sensitive to the chosen subsample for the evaluation. To be more precise, only for the high volatility period the average RMSPE

Panel A: 1-month ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| FADNS (2) | 0.993 | 1.004 | 1.013 | 1.009 | 1.001 | 1.004 | 1.009 | 1.018 | 1.016 | 1.019 | 1.015 | 1.019 | 1.011 | 1.011 | 1.007 | 1.006 | 1.015 |
| DNS-GARCH | **0.924** | 0.995 | 1.020 | 1.017 | 1.006 | 1.004 | 1.004 | 1.009 | 1.010 | 1.013 | 1.012 | 1.014 | 1.009 | 1.006 | 1.001 | 0.999 | 1.004 |
| DNS-TVL-EKF | **0.894** | 0.991 | 1.021 | 1.020 | 1.007 | 1.009 | 1.013 | 1.021 | 1.019 | 1.020 | 1.006 | 1.011 | 0.998 | 1.008 | 1.004 | 1.004 | 1.002 |
| DNS-TVL-UKF | **0.889** | 0.997 | 1.030 | 1.027 | 1.009 | 1.012 | 1.017 | 1.027 | 1.027 | 1.030 | 1.015 | 1.023 | 1.009 | 1.012 | 1.004 | 1.005 | 1.003 |
| SFADNS-EKF | **0.861** | 0.981 | 1.029 | 1.025 | 1.000 | 1.007 | 1.017 | 1.033 | 1.031 | 1.035 | 1.020 | 1.034 | 1.016 | 1.027 | 1.018 | 1.014 | 1.018 |
| SFADNS-UKF | **0.865** | 0.990 | 1.038 | 1.035 | 1.008 | 1.015 | 1.026 | 1.044 | 1.040 | 1.045 | 1.022 | 1.034 | 1.011 | 1.018 | 1.006 | 1.005 | 1.010 |

Panel B: 6-months ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| FADNS (2) | 1.011 | 1.017 | 1.022 | 1.020 | 1.016 | 1.019 | 1.023 | 1.029 | 1.030 | 1.033 | 1.033 | 1.037 | 1.034 | 1.035 | 1.031 | 1.031 | 1.040 |
| DNS-GARCH | 1.032 | 1.033 | 1.033 | 1.031 | 1.027 | 1.028 | 1.029 | 1.032 | 1.033 | 1.035 | 1.036 | 1.040 | 1.038 | 1.039 | 1.036 | 1.035 | 1.041 |
| DNS-TVL-EKF | 1.046 | 1.049 | 1.052 | 1.053 | 1.050 | 1.053 | 1.057 | 1.061 | 1.064 | 1.067 | 1.067 | 1.070 | 1.068 | 1.070 | 1.067 | 1.069 | 1.074 |
| DNS-TVL-UKF | 1.054 | 1.056 | 1.059 | 1.060 | 1.056 | 1.060 | 1.064 | 1.069 | 1.073 | 1.077 | 1.077 | 1.082 | 1.080 | 1.081 | 1.078 | 1.079 | 1.084 |
| SFADNS-EKF | 1.028 | 1.050 | 1.067 | 1.072 | 1.068 | 1.077 | 1.087 | 1.098 | 1.101 | 1.106 | 1.101 | 1.106 | 1.097 | 1.098 | 1.089 | 1.087 | 1.097 |
| SFADNS-UKF | 1.023 | 1.045 | 1.060 | 1.065 | 1.061 | 1.069 | 1.077 | 1.087 | 1.089 | 1.093 | 1.089 | 1.093 | 1.085 | 1.087 | 1.080 | 1.080 | 1.089 |

Panel C: 12-months ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| FADNS (2) | 1.050 | 1.050 | 1.051 | 1.047 | 1.043 | 1.044 | 1.046 | 1.050 | 1.051 | 1.054 | 1.056 | 1.061 | 1.060 | 1.061 | 1.058 | 1.059 | 1.065 |
| DNS-GARCH | 1.044 | 1.043 | 1.044 | 1.043 | 1.041 | 1.043 | 1.045 | 1.049 | 1.052 | 1.056 | 1.060 | 1.066 | 1.066 | 1.068 | 1.066 | 1.066 | 1.071 |
| DNS-TVL-EKF | 1.056 | 1.059 | 1.064 | 1.067 | 1.067 | 1.072 | 1.078 | 1.085 | 1.090 | 1.096 | 1.101 | 1.106 | 1.106 | 1.105 | 1.102 | 1.101 | 1.104 |
| DNS-TVL-UKF | 1.063 | 1.066 | 1.072 | 1.074 | 1.074 | 1.080 | 1.086 | 1.093 | 1.099 | 1.106 | 1.112 | 1.118 | 1.118 | 1.117 | 1.114 | 1.113 | 1.115 |
| SFADNS-EKF | 1.072 | 1.087 | 1.101 | 1.107 | 1.108 | 1.117 | 1.127 | 1.138 | 1.145 | 1.153 | 1.157 | 1.162 | 1.156 | 1.154 | 1.146 | 1.143 | 1.147 |
| SFADNS-UKF | 1.067 | 1.080 | 1.091 | 1.095 | 1.095 | 1.102 | 1.110 | 1.119 | 1.124 | 1.130 | 1.132 | 1.136 | 1.131 | 1.130 | 1.123 | 1.122 | 1.126 |

Table 6: Root mean squared prediction errors (RMSPE) ratios. The table shows the relative RMSPE of the considered extensions of the DNS against the standard DNS at a forecast horizon of one, six and twelve months. The results per forecast horizon are in Panel A, B and C, respectively, with corresponding periods 1985:01-2009:12, 1985:06-2009:12, 1985:12-2009:12. The shades of the cells indicate the rank of the model per maturity. The darker the shade, the better the model. Values in bold denote significance at the 5% level or less.

Panel A: 1-month ahead forecasts

| Model | Maturity 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| FADNS (2) | 1.027 | 1.024 | 1.023 | 0.995 | 0.973 | 0.981 | 0.988 | 1.001 | 0.999 | 1.002 | 1.000 | 1.012 | 1.004 | 0.999 | 0.992 | 0.994 | 1.007 |
| DNS-GARCH | **0.918** | 0.959 | 0.985 | 0.973 | 0.968 | 0.976 | 0.984 | 0.996 | 1.003 | 1.009 | 1.008 | 1.019 | 1.012 | 1.004 | 0.996 | 0.995 | 1.004 |
| DNS-TVL-EKF | **0.935** | 0.959 | 1.004 | 0.985 | 0.968 | 0.983 | 0.998 | 1.017 | 1.019 | 1.021 | 1.012 | 1.017 | 1.011 | 1.006 | 1.009 | 1.012 | 1.021 |
| DNS-TVL-UKF | **0.908** | **0.950** | 1.005 | 0.975 | 0.955 | 0.974 | 0.991 | 1.015 | 1.018 | 1.022 | 1.011 | 1.025 | 1.014 | 1.002 | 0.995 | 0.997 | 1.016 |
| SFADNS-EKF | **0.927** | 0.981 | 1.061 | 1.006 | 0.952 | 0.982 | 1.007 | 1.042 | 1.039 | 1.042 | 1.024 | 1.044 | 1.026 | 1.010 | 0.998 | 1.002 | 1.029 |
| SFADNS-UKF | **0.927** | 0.969 | 1.032 | 0.981 | **0.940** | 0.967 | 0.990 | 1.024 | 1.025 | 1.030 | 1.015 | 1.035 | 1.018 | 1.003 | 0.992 | 0.996 | 1.020 |

Panel B: 6-months ahead forecasts

| Model | Maturity 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| FADNS (2) | 1.038 | 1.022 | 1.016 | 0.990 | 0.969 | 0.973 | 0.978 | 0.988 | 0.987 | 0.991 | 0.992 | 1.005 | 1.002 | 1.000 | 0.997 | 1.002 | 1.014 |
| DNS-GARCH | 0.955 | 0.962 | 0.970 | 0.967 | 0.964 | 0.972 | 0.979 | 0.989 | 0.994 | 1.000 | 1.004 | 1.014 | 1.013 | 1.010 | 1.007 | 1.010 | 1.019 |
| DNS-TVL-EKF | **0.943** | 0.961 | 0.983 | 0.986 | 0.982 | 0.996 | 1.007 | 1.022 | 1.029 | 1.035 | 1.036 | 1.046 | 1.044 | 1.041 | 1.040 | 1.045 | 1.056 |
| DNS-TVL-UKF | **0.939** | 0.959 | 0.984 | 0.984 | 0.979 | 0.995 | 1.007 | 1.024 | 1.031 | 1.037 | 1.039 | 1.050 | 1.047 | 1.042 | 1.040 | 1.045 | 1.059 |
| SFADNS-EKF | 0.993 | 1.012 | 1.047 | 1.024 | 0.997 | 1.019 | 1.036 | 1.059 | 1.063 | 1.067 | 1.060 | 1.074 | 1.064 | 1.055 | 1.049 | 1.054 | 1.073 |
| SFADNS-UKF | 1.005 | 1.002 | 1.020 | 0.998 | 0.975 | 0.992 | 1.007 | 1.028 | 1.033 | 1.040 | 1.038 | 1.054 | 1.047 | 1.042 | 1.038 | 1.045 | 1.063 |

Panel C: 12-months ahead forecasts

| Model | Maturity 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| FADNS (2) | 1.097 | 1.071 | 1.056 | 1.021 | 0.994 | 0.994 | 0.996 | 1.004 | 1.003 | 1.007 | 1.010 | 1.025 | 1.024 | 1.023 | 1.022 | 1.029 | 1.045 |
| DNS-GARCH | **0.949** | 0.955 | 0.961 | 0.957 | **0.950** | 0.959 | 0.969 | 0.981 | 0.990 | 0.999 | 1.010 | 1.027 | 1.028 | 1.027 | 1.026 | 1.031 | 1.044 |
| DNS-TVL-EKF | 0.957 | 0.987 | 1.012 | 1.017 | 1.014 | 1.031 | 1.045 | 1.063 | 1.074 | 1.083 | 1.087 | 1.097 | 1.094 | 1.089 | 1.087 | 1.091 | 1.101 |
| DNS-TVL-UKF | 0.954 | 0.984 | 1.010 | 1.014 | 1.009 | 1.027 | 1.043 | 1.062 | 1.074 | 1.084 | 1.089 | 1.102 | 1.098 | 1.092 | 1.089 | 1.094 | 1.106 |
| SFADNS-EKF | 1.087 | 1.105 | 1.130 | 1.112 | 1.089 | 1.109 | 1.125 | 1.148 | 1.156 | 1.161 | 1.153 | 1.161 | 1.149 | 1.135 | 1.128 | 1.133 | 1.148 |
| SFADNS-UKF | 1.106 | 1.103 | 1.111 | 1.085 | 1.061 | 1.073 | 1.084 | 1.103 | 1.108 | 1.114 | 1.111 | 1.121 | 1.113 | 1.103 | 1.099 | 1.104 | 1.120 |

Table 7: Root mean squared prediction errors (RMSPE) ratios for the period 1995:01-1999:12. The table shows the relative RMSPE of the considered extensions of the DNS against the standard DNS at a forecast horizon of one, six and twelve months. The results per forecast horizon are in Panel A, B and C, respectively. The shades of the cells indicate the rank of the model per maturity. The darker the shade, the better the model. Values in bold denote significance at the 5% level or less.

Panel A: 1-month ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| FADNS (2) | 0.980 | 0.981 | 0.993 | 1.001 | 1.006 | 1.010 | 1.013 | 1.017 | 1.019 | 1.020 | 1.016 | 1.017 | 1.014 | 1.011 | 1.006 | 1.004 | 1.003 |
| DNS-GARCH | **0.870** | **0.901** | **0.944** | 0.967 | 0.978 | 0.987 | 0.996 | 1.004 | 1.018 | 1.025 | 1.029 | 1.028 | 1.023 | 1.010 | 1.001 | 0.994 | 0.988 |
| DNS-TVL-EKF | **0.778** | **0.848** | **0.933** | 0.980 | 1.007 | 1.018 | 1.019 | 1.016 | 1.004 | 0.986 | 0.951 | 0.955 | **0.950** | 0.988 | 0.973 | 0.990 | 0.955 |
| DNS-TVL-UKF | **0.759** | **0.835** | **0.926** | 0.975 | 1.000 | 1.015 | 1.018 | 1.018 | 1.015 | 1.001 | 0.974 | 0.982 | 0.988 | 1.016 | 1.005 | 1.032 | 0.983 |
| SFADNS-EKF | **0.721** | **0.793** | **0.899** | 0.957 | 0.967 | 0.986 | 0.997 | 1.001 | 0.995 | 0.996 | 0.972 | 1.013 | 1.007 | 1.080 | 1.057 | 1.057 | 1.024 |
| SFADNS-UKF | **0.721** | **0.796** | **0.895** | 0.950 | 0.968 | 0.983 | 0.992 | 0.994 | 0.989 | 0.980 | 0.950 | 0.966 | 0.966 | 1.007 | 0.992 | 1.012 | 0.976 |

Panel B: 6-months ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| FADNS (2) | 0.991 | 0.993 | 0.997 | 1.000 | 1.003 | 1.005 | 1.007 | 1.009 | 1.010 | 1.012 | 1.012 | 1.017 | 1.019 | 1.023 | 1.019 | 1.017 | 1.018 |
| DNS-GARCH | 0.980 | 0.986 | 0.991 | 0.995 | 1.000 | 1.003 | 1.007 | 1.010 | 1.015 | 1.019 | 1.023 | 1.027 | 1.028 | 1.028 | 1.024 | 1.022 | 1.016 |
| DNS-TVL-EKF | 0.980 | 0.995 | 1.005 | 1.012 | 1.017 | 1.019 | 1.021 | 1.022 | 1.023 | 1.026 | 1.029 | 1.046 | 1.065 | 1.093 | 1.111 | 1.139 | 1.173 |
| DNS-TVL-UKF | 0.983 | 0.998 | 1.008 | 1.016 | 1.022 | 1.025 | 1.029 | 1.031 | 1.036 | 1.041 | 1.051 | 1.075 | 1.100 | 1.135 | 1.157 | 1.189 | 1.225 |
| SFADNS-EKF | **0.921** | **0.941** | 0.956 | 0.966 | 0.972 | 0.976 | 0.981 | 0.985 | 0.987 | 0.992 | 0.996 | 1.021 | 1.040 | 1.073 | 1.083 | 1.101 | 1.128 |
| SFADNS-UKF | **0.916** | **0.937** | 0.951 | 0.959 | 0.963 | 0.966 | 0.969 | 0.971 | 0.973 | 0.976 | 0.981 | 1.005 | 1.027 | 1.059 | 1.075 | 1.098 | 1.129 |

Panel C: 12-months ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| FADNS (2) | 1.046 | 1.043 | 1.041 | 1.040 | 1.038 | 1.037 | 1.037 | 1.037 | 1.035 | 1.036 | 1.036 | 1.042 | 1.045 | 1.050 | 1.049 | 1.048 | 1.047 |
| DNS-GARCH | 1.003 | 1.003 | 1.005 | 1.008 | 1.011 | 1.014 | 1.017 | 1.021 | 1.026 | 1.031 | 1.039 | 1.048 | 1.053 | 1.060 | 1.060 | 1.061 | 1.060 |
| DNS-TVL-EKF | 1.019 | 1.023 | 1.028 | 1.033 | 1.037 | 1.042 | 1.047 | 1.051 | 1.059 | 1.069 | 1.089 | 1.116 | 1.141 | 1.176 | 1.203 | 1.231 | 1.260 |
| DNS-TVL-UKF | 1.025 | 1.029 | 1.034 | 1.039 | 1.045 | 1.050 | 1.056 | 1.062 | 1.072 | 1.084 | 1.109 | 1.144 | 1.173 | 1.217 | 1.249 | 1.282 | 1.316 |
| SFADNS-EKF | 1.023 | 1.029 | 1.034 | 1.037 | 1.040 | 1.043 | 1.047 | 1.051 | 1.055 | 1.061 | 1.076 | 1.101 | 1.123 | 1.154 | 1.173 | 1.192 | 1.211 |
| SFADNS-UKF | 1.022 | 1.027 | 1.030 | 1.031 | 1.031 | 1.031 | 1.033 | 1.034 | 1.035 | 1.039 | 1.051 | 1.074 | 1.097 | 1.128 | 1.150 | 1.175 | 1.196 |

Table 8: Root mean squared prediction errors (RMSPE) ratios for the period 2005:01-2009:12. The table shows the relative RMSPE of the considered extensions of the DNS against the standard DNS at a forecast horizon of one, six and twelve months. The results per forecast horizon are in Panel A, B and C, respectively. The shades of the cells indicate the rank of the model per maturity. The darker the shade, the better the model. Values in bold denote significance at the 5% level or less.

ratio of the SFADNS decreases relatively to the FADNS with two macro factors, the DNS-GARCH and the DNS-TVL, which is contrary to the findings for other subsamples. Moreover, for these subsamples the RMSPE ratios of the SFADNS are either in between the minimum and maximum of the RMSPE ratios of the latter three models or even higher.

Focusing on the difference between the effect of the filtering techniques on the forecasts, I find that the DNS-TVL-EKF provides more accurate forecasts than the DNS-TVL-UKF when the full out-of-sample period is considered, though the difference is marginal (on average 0.5%). However, for periods of highly volatile yields the percentage of the difference between these two filtering methods increases in favor of the DNS-TVL-EKF, ranging from 1% to 2%, depending on the forecast horizon and irrespective of the maturity. Contrary, the forecasts of the yield when the DNS-TVL-UKF is used are on average more accurate than those from the DNS-TVL-EKF when the yields are relatively stable, in particular for both ends of the yield curve. Remarkably, for the the SFADNS, I find the reversal of the findings. That is, the RMSPE ratios for the SFADNS-UKF are on average lower than those for the SFADNS-EKF, with improvements reaching up to 4%. This finding is independent on the forecast horizon and subsample that is considered. A notable exception is for the full-out-sample period at a forecast horizon of one month where the forecasts from the SFADNS-EKF are more accurate.

### 4.4 Out-of-Sample Analysis: On Forecast Combinations

It has often been found that combining forecasts lead to additional gains in predictive accuracy over the individual forecasting models. Introducing the SFADNS as a mixture of the other models (FADNS, DNS-GARCH and DNS-TVL) is one way of combining the characteristics with the aim of improving the forecasts. Another method is to consider the predicted yields from these models and combine them using specific weights. Moreover, the forecasted yields are constructed as

$$\hat{\boldsymbol{y}}^{c}_{t+h|t} = w^{1}_{t+h|t}\hat{\boldsymbol{y}}^{FADNS}_{t+h|t} + w^{2}_{t+h|t}\hat{\boldsymbol{y}}^{DNS-GARCH}_{t+h|t} + w^{3}_{t+h|t}\hat{\boldsymbol{y}}^{DNS-TVL}_{t+h|t} \tag{51}$$

where the superscript of the forecasted yields denote the models in the combination, and $w^i$ denote the weight that is used. In this thesis, I consider three weighting schemes. For the first scheme I use equal weights to combine the forecasts as several authors, Timmermann (2006) among others, find that simple combination schemes perform very well compared to more sophisticated combination schemes. I refer to the forecast combinations in which I use the equal weights scheme as EW-EKF and EW-UKF, depending on whether the DNS-TVL-EKF or DNS-TVL-UKF is used for the combination. For the second and third schemes the weights are varying over time and both dependent on the RMSPE. I use a rolling window of 2 years to compute this. Now, let the evaluation period be 2 years prior to time $t$ up to time $t$, the forecast from model $i$ at time $t+h$ then gets the weight

$$w^{i}_{t+h|t} = \frac{1/\text{RMSPE}^{i}_{t}}{\sum_{i=1}^{3} 1/\text{RMSPE}^{i}_{t}}, \tag{52}$$

to construct the combined forecasts for time $t+h$. Moreover, $\text{RMSPE}^{i}_{t}$ denotes the RMSPE of model $i$ over the period 2 years prior to time $t$ up until time $t$. Details on the different schemes; in the second scheme I use the average RMSPE across the maturities in the rolling window per model. While in the third scheme, I look at the RMSPE in the rolling window per maturity. This means that the weights not only differ over time, but also differ across the maturities. The reasoning behind the use of the latter scheme stems from the fact that I find a different best model per maturity when forecasting the term structure of interest rates. I refer to the combined forecasts using scheme 2 as RMSPE-AVG-EKF and RMSPE-AVG-UKF, and RMSPE-MAT-EKF and RMSPE-MAT-UKF when I use scheme 3.

Table 9 reports the ratios of the RMSPEs of both SFADNS and that of the combined forecasts using the various schemes relative to the DNS for out-of-sample forecasts at a horizon of one, six and twelve months for the full period. Just as for the DNS and its considered extensions, I examine the out-of-sample forecasting performance in the five subsamples. Again, I find similar results for the periods in which the yields are relatively highly volatile. Hence, I report only the results for the earlier mentioned subsamples for the sake of brevity. More specifically, the ratios for the stable period (January 1995 until December 1999) and the period of relatively high volatility (January 2005 until December 2009) are reported in Table 10 and 11, respectively.

I start with the evaluation of the one-month ahead forecasts. On average, the combined forecasts using the various weighting schemes are slightly more accurate than the standard DNS at both ends of the yield curve for the full out-of-sample period. In particular, the gain in predictive accuracy of the combined forecasts at the short-end of the yield curve ranges from 1,5% of the equally weighted forecasts to nearly 11% when the maturity-specific weighting scheme is used to combine the forecasts. Moreover, the combined forecasts are significantly more accurate than those from the standard DNS model at a significance level of five percent. Furthermore, for intermediate yields the difference between the out-of-sample performance of the standard DNS and the combined forecasts is marginal. For the stable period, January 1995 until December 1999, the forecasts obtained after conducting the various weighting schemes are on average more accurate than those from the DNS. Only at a maturity of three months I find the gain in predictive accuracy to be significant. When the volatile period is considered, January 2005 until December 2009, I find a greater amount of combined forecasts to be significantly more accurate, particularly for the short-end of the yield curve.

The accuracy of the six- and twelve-month ahead forecasts provided by the three weighting schemes are relatively worse than those from the DNS for all maturities when the full out-of-sample period is considered. For the subsamples, the results are more promising. Consider first the relatively highly volatile period. Even though the DNS dominates the combined forecasts at a forecast horizon of twelve months, gains in predictive accuracy of approximately 2% are attained when six-month ahead forecasts are made for the short-end of the yield curve. The results become better when the stable period is examined. Moreover, in nearly 42% of the cases the six-month ahead forecasts from the DNS are dominated by the combined forecasts. This holds in particular for the short- to intermediate-term yields. For the same yields, I find a gain in predictive accuracy when the twelve-month ahead forecasts are combined over the forecasts from the benchmark. However, for these forecasts horizons, none of the combination schemes are able to provide significantly more accurate forecasts than the standard DNS.

Overall, at a forecast horizon of one-month, the considered combination schemes are able to provide either forecasts with equal predictive accuracy relative to the standard DNS or forecasts that are more accurate. For a forecast horizon of six- and twelve-months, the results are in general not that promising, with an exception for the stable period. Moreover, for this subsample, I find the use of combinations schemes to be appealing for forecasting the short-end of the yield curve. Comparing the different combination schemes with each other, I find that the maturity-based combination scheme is able to provide, on average, more accurate forecasts than the other two schemes for the short-end of the yield curve, at a forecast horizon of one-month

| Panel A: 1-month ahead forecasts | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model | Maturity | | | | | | | | | | | | | | | | |
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | **0.917** | 0.985 | 1.013 | 1.010 | 1.001 | 1.003 | 1.006 | 1.012 | 1.012 | 1.013 | 1.007 | 1.010 | 1.004 | 1.003 | 0.997 | 0.996 | 0.995 |
| EW-UKF | **0.914** | 0.986 | 1.015 | 1.011 | 1.001 | 1.003 | 1.006 | 1.013 | 1.014 | 1.016 | 1.011 | 1.015 | 1.008 | 1.004 | 0.996 | 0.994 | 0.993 |
| RMSPE-AVG-EKF | **0.916** | 0.985 | 1.013 | 1.010 | 1.001 | 1.003 | 1.006 | 1.012 | 1.012 | 1.013 | 1.007 | 1.010 | 1.004 | 1.003 | 0.997 | 0.996 | 0.995 |
| RMSPE-AVG-UKF | **0.914** | 0.986 | 1.014 | 1.011 | 1.001 | 1.002 | 1.006 | 1.013 | 1.014 | 1.016 | 1.010 | 1.015 | 1.008 | 1.004 | 0.996 | 0.994 | 0.993 |
| RMSPE-MAT-EKF | **0.913** | 0.984 | 1.013 | 1.010 | 1.001 | 1.002 | 1.006 | 1.012 | 1.012 | 1.013 | 1.007 | 1.010 | 1.003 | 1.003 | 0.997 | 0.996 | 0.995 |
| RMSPE-MAT-UKF | **0.909** | 0.984 | 1.014 | 1.011 | 1.001 | 1.002 | 1.006 | 1.013 | 1.014 | 1.016 | 1.010 | 1.015 | 1.008 | 1.004 | 0.996 | 0.994 | 0.994 |
| SFADNS-EKF | **0.846** | 0.976 | 1.029 | 1.024 | 0.998 | 1.007 | 1.018 | 1.034 | 1.032 | 1.035 | 1.019 | 1.034 | 1.018 | 1.027 | 1.014 | 1.013 | 1.016 |
| SFADNS-UKF | **0.848** | 0.978 | 1.028 | 1.022 | 0.998 | 1.006 | 1.016 | 1.032 | 1.031 | 1.034 | 1.015 | 1.026 | 1.007 | 1.011 | 0.998 | 1.001 | 1.003 |

| Panel B: 6-months ahead forecasts | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model | Maturity | | | | | | | | | | | | | | | | |
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | 1.018 | 1.022 | 1.024 | 1.023 | 1.020 | 1.022 | 1.024 | 1.028 | 1.030 | 1.032 | 1.032 | 1.035 | 1.033 | 1.031 | 1.027 | 1.026 | 1.032 |
| EW-UKF | 1.020 | 1.024 | 1.026 | 1.025 | 1.022 | 1.024 | 1.026 | 1.031 | 1.032 | 1.035 | 1.035 | 1.039 | 1.037 | 1.035 | 1.031 | 1.030 | 1.036 |
| RMSPE-AVG-EKF | 1.018 | 1.022 | 1.024 | 1.023 | 1.020 | 1.022 | 1.024 | 1.028 | 1.029 | 1.032 | 1.031 | 1.035 | 1.032 | 1.031 | 1.027 | 1.026 | 1.032 |
| RMSPE-AVG-UKF | 1.020 | 1.024 | 1.026 | 1.025 | 1.021 | 1.023 | 1.026 | 1.030 | 1.032 | 1.035 | 1.035 | 1.039 | 1.037 | 1.035 | 1.030 | 1.030 | 1.035 |
| RMSPE-MAT-EKF | 1.018 | 1.022 | 1.024 | 1.023 | 1.020 | 1.022 | 1.024 | 1.028 | 1.029 | 1.032 | 1.031 | 1.035 | 1.032 | 1.031 | 1.027 | 1.026 | 1.031 |
| RMSPE-MAT-UKF | 1.020 | 1.024 | 1.026 | 1.025 | 1.021 | 1.023 | 1.026 | 1.030 | 1.032 | 1.035 | 1.035 | 1.039 | 1.036 | 1.035 | 1.030 | 1.029 | 1.035 |
| SFADNS-EKF | 1.018 | 1.037 | 1.051 | 1.055 | 1.050 | 1.057 | 1.066 | 1.077 | 1.079 | 1.083 | 1.078 | 1.083 | 1.074 | 1.073 | 1.063 | 1.062 | 1.074 |
| SFADNS-UKF | 1.013 | 1.029 | 1.039 | 1.040 | 1.034 | 1.039 | 1.046 | 1.055 | 1.056 | 1.060 | 1.056 | 1.062 | 1.056 | 1.057 | 1.051 | 1.052 | 1.064 |

| Panel C: 12-months ahead forecasts | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model | Maturity | | | | | | | | | | | | | | | | |
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | 1.038 | 1.039 | 1.041 | 1.041 | 1.039 | 1.041 | 1.044 | 1.049 | 1.052 | 1.056 | 1.060 | 1.066 | 1.066 | 1.066 | 1.062 | 1.063 | 1.069 |
| EW-UKF | 1.040 | 1.041 | 1.044 | 1.043 | 1.041 | 1.044 | 1.047 | 1.052 | 1.055 | 1.060 | 1.064 | 1.071 | 1.071 | 1.071 | 1.067 | 1.068 | 1.074 |
| RMSPE-AVG-EKF | 1.038 | 1.039 | 1.041 | 1.041 | 1.039 | 1.041 | 1.044 | 1.049 | 1.052 | 1.056 | 1.060 | 1.066 | 1.065 | 1.065 | 1.062 | 1.062 | 1.068 |
| RMSPE-AVG-UKF | 1.040 | 1.041 | 1.043 | 1.043 | 1.041 | 1.043 | 1.047 | 1.052 | 1.055 | 1.060 | 1.064 | 1.070 | 1.070 | 1.070 | 1.067 | 1.067 | 1.074 |
| RMSPE-MAT-EKF | 1.038 | 1.039 | 1.041 | 1.040 | 1.039 | 1.041 | 1.044 | 1.049 | 1.052 | 1.056 | 1.060 | 1.066 | 1.065 | 1.065 | 1.062 | 1.062 | 1.068 |
| RMSPE-MAT-UKF | 1.040 | 1.041 | 1.043 | 1.043 | 1.041 | 1.043 | 1.047 | 1.052 | 1.055 | 1.059 | 1.064 | 1.070 | 1.070 | 1.070 | 1.066 | 1.067 | 1.073 |
| SFADNS-EKF | 1.066 | 1.077 | 1.089 | 1.092 | 1.091 | 1.099 | 1.108 | 1.119 | 1.124 | 1.132 | 1.135 | 1.142 | 1.137 | 1.136 | 1.128 | 1.127 | 1.135 |
| SFADNS-UKF | 1.061 | 1.068 | 1.075 | 1.075 | 1.072 | 1.077 | 1.083 | 1.091 | 1.095 | 1.100 | 1.102 | 1.109 | 1.106 | 1.106 | 1.101 | 1.102 | 1.110 |

Table 9: Root mean squared prediction errors (RMSPE) ratios. The table shows the relative RMSPE of the combined forecasts using various weighting schemes against the standard DNS at a forecast horizon of one, six and twelve months. The results per forecast horizon are in Panel A, B and C, respectively, with corresponding periods 1987:01-2009:12, 1987:06-2009:12, 1987:12-2009:12. The shades of the cells indicate the rank of the model per maturity. The darker the shade, the better the model. Values in bold denote significance at the 5% level or less.

Panel A: 1-month ahead forecasts

| Model | Maturity 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | **0.938** | 0.974 | 1.001 | 0.982 | 0.968 | 0.978 | 0.989 | 1.004 | 1.006 | 1.010 | 1.006 | 1.016 | 1.009 | 1.002 | 0.998 | 0.999 | 1.008 |
| EW-UKF | **0.931** | 0.971 | 1.001 | 0.978 | 0.963 | 0.975 | 0.986 | 1.002 | 1.006 | 1.010 | 1.006 | 1.018 | 1.010 | 1.001 | 0.994 | 0.995 | 1.007 |
| RMSPE-AVG-EKF | **0.938** | 0.974 | 1.001 | 0.982 | 0.968 | 0.978 | 0.989 | 1.004 | 1.006 | 1.010 | 1.006 | 1.016 | 1.009 | 1.002 | 0.998 | 0.999 | 1.008 |
| RMSPE-AVG-UKF | **0.931** | 0.971 | 1.001 | 0.978 | 0.963 | 0.975 | 0.986 | 1.003 | 1.006 | 1.010 | 1.006 | 1.018 | 1.010 | 1.001 | 0.994 | 0.995 | 1.008 |
| RMSPE-MAT-EKF | **0.936** | 0.973 | 1.001 | 0.982 | 0.968 | 0.978 | 0.989 | 1.004 | 1.006 | 1.010 | 1.006 | 1.016 | 1.009 | 1.002 | 0.998 | 0.999 | 1.009 |
| RMSPE-MAT-UKF | **0.929** | 0.970 | 1.001 | 0.978 | 0.963 | 0.975 | 0.986 | 1.003 | 1.006 | 1.010 | 1.006 | 1.018 | 1.010 | 1.001 | 0.994 | 0.995 | 1.008 |
| SFADNS-EKF | **0.927** | 0.981 | 1.061 | 1.006 | 0.952 | 0.982 | 1.007 | 1.042 | 1.039 | 1.042 | 1.024 | 1.044 | 1.026 | 1.010 | 0.998 | 1.002 | 1.029 |
| SFADNS-UKF | **0.927** | 0.969 | 1.032 | 0.981 | **0.940** | 0.967 | 0.990 | 1.024 | 1.025 | 1.030 | 1.015 | 1.035 | 1.018 | 1.003 | 0.992 | 0.996 | 1.020 |

Panel B: 6-months ahead forecasts

| Model | Maturity 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | 0.974 | 0.979 | 0.988 | 0.979 | 0.969 | 0.978 | 0.986 | 0.998 | 1.002 | 1.007 | 1.010 | 1.021 | 1.019 | 1.016 | 1.014 | 1.018 | 1.029 |
| EW-UKF | 0.973 | 0.979 | 0.987 | 0.978 | 0.968 | 0.978 | 0.986 | 0.998 | 1.002 | 1.008 | 1.011 | 1.022 | 1.020 | 1.017 | 1.014 | 1.018 | 1.030 |
| RMSPE-AVG-EKF | 0.974 | 0.979 | 0.987 | 0.979 | 0.969 | 0.978 | 0.986 | 0.998 | 1.002 | 1.007 | 1.010 | 1.021 | 1.019 | 1.016 | 1.014 | 1.018 | 1.029 |
| RMSPE-AVG-UKF | 0.973 | 0.979 | 0.987 | 0.978 | 0.968 | 0.978 | 0.986 | 0.998 | 1.002 | 1.008 | 1.010 | 1.022 | 1.020 | 1.017 | 1.014 | 1.018 | 1.030 |
| RMSPE-MAT-EKF | 0.972 | 0.979 | 0.987 | 0.979 | 0.969 | 0.978 | 0.986 | 0.998 | 1.002 | 1.007 | 1.010 | 1.021 | 1.019 | 1.016 | 1.014 | 1.018 | 1.029 |
| RMSPE-MAT-UKF | 0.972 | 0.978 | 0.987 | 0.978 | 0.968 | 0.978 | 0.986 | 0.998 | 1.002 | 1.007 | 1.010 | 1.022 | 1.020 | 1.017 | 1.014 | 1.018 | 1.030 |
| SFADNS-EKF | 0.993 | 1.012 | 1.047 | 1.024 | 0.997 | 1.019 | 1.036 | 1.059 | 1.063 | 1.067 | 1.060 | 1.074 | 1.064 | 1.055 | 1.049 | 1.054 | 1.073 |
| SFADNS-UKF | 1.005 | 1.002 | 1.020 | 0.998 | 0.975 | 0.992 | 1.007 | 1.028 | 1.033 | 1.040 | 1.038 | 1.054 | 1.047 | 1.042 | 1.038 | 1.045 | 1.063 |

Panel C: 12-months ahead forecasts

| Model | Maturity 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | 0.997 | 1.001 | 1.007 | 0.996 | 0.983 | 0.992 | 1.001 | 1.014 | 1.020 | 1.028 | 1.034 | 1.049 | 1.048 | 1.045 | 1.044 | 1.050 | 1.062 |
| EW-UKF | 0.997 | 1.000 | 1.006 | 0.994 | 0.981 | 0.991 | 1.000 | 1.013 | 1.020 | 1.028 | 1.035 | 1.050 | 1.049 | 1.046 | 1.045 | 1.051 | 1.064 |
| RMSPE-AVG-EKF | 0.998 | 1.001 | 1.007 | 0.996 | 0.983 | 0.992 | 1.001 | 1.014 | 1.020 | 1.028 | 1.034 | 1.049 | 1.048 | 1.045 | 1.044 | 1.049 | 1.062 |
| RMSPE-AVG-UKF | 0.997 | 1.000 | 1.006 | 0.994 | 0.981 | 0.991 | 1.000 | 1.013 | 1.020 | 1.028 | 1.034 | 1.050 | 1.049 | 1.046 | 1.045 | 1.051 | 1.064 |
| RMSPE-MAT-EKF | 0.997 | 1.001 | 1.007 | 0.996 | 0.983 | 0.992 | 1.001 | 1.013 | 1.020 | 1.027 | 1.033 | 1.048 | 1.047 | 1.045 | 1.044 | 1.049 | 1.062 |
| RMSPE-MAT-UKF | 0.996 | 1.000 | 1.006 | 0.995 | 0.981 | 0.990 | 0.999 | 1.013 | 1.019 | 1.027 | 1.034 | 1.049 | 1.049 | 1.046 | 1.045 | 1.051 | 1.064 |
| SFADNS-EKF | 1.087 | 1.105 | 1.130 | 1.112 | 1.089 | 1.109 | 1.125 | 1.148 | 1.156 | 1.161 | 1.153 | 1.161 | 1.149 | 1.135 | 1.128 | 1.133 | 1.148 |
| SFADNS-UKF | 1.106 | 1.103 | 1.111 | 1.085 | 1.061 | 1.073 | 1.084 | 1.103 | 1.108 | 1.114 | 1.111 | 1.121 | 1.113 | 1.103 | 1.099 | 1.104 | 1.120 |

Table 10: Root mean squared prediction errors (RMSPE) ratios for the period 1995:01-1999:12. The table shows the relative RMSPE of the combined forecasts using various weighting schemes against the standard DNS at a forecast horizon of one, six and twelve months. The results per forecast horizon are in Panel A, B and C, respectively. The shades of the cells indicate the rank of the model per maturity. The darker the shade, the better the model. Values in bold denote significance at the 5% level or less.

Panel A: 1-month ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | **0.872** | **0.905** | 0.953 | 0.981 | 0.995 | 1.004 | 1.009 | 1.012 | 1.013 | 1.008 | 0.995 | 0.994 | 0.991 | 0.996 | 0.986 | 0.983 | 0.960 |
| EW-UKF | **0.865** | **0.900** | **0.950** | 0.978 | 0.992 | 1.002 | 1.008 | 1.012 | 1.016 | 1.014 | 1.004 | 1.005 | 1.004 | 1.004 | 0.991 | 0.988 | 0.956 |
| RMSPE-AVG-EKF | **0.871** | **0.904** | 0.953 | 0.981 | 0.995 | 1.004 | 1.009 | 1.012 | 1.013 | 1.008 | 0.995 | 0.993 | 0.991 | 0.996 | 0.986 | 0.983 | 0.959 |
| RMSPE-AVG-UKF | **0.865** | **0.899** | **0.950** | 0.978 | 0.992 | 1.002 | 1.008 | 1.012 | 1.016 | 1.014 | 1.004 | 1.005 | 1.004 | 1.004 | 0.991 | 0.988 | 0.955 |
| RMSPE-MAT-EKF | **0.865** | **0.902** | 0.953 | 0.981 | 0.995 | 1.004 | 1.008 | 1.012 | 1.013 | 1.008 | 0.995 | 0.993 | 0.991 | 0.996 | 0.986 | 0.984 | 0.961 |
| RMSPE-MAT-UKF | **0.857** | **0.896** | **0.949** | 0.978 | 0.992 | 1.002 | 1.008 | 1.012 | 1.016 | 1.014 | 1.004 | 1.004 | 1.004 | 1.004 | 0.992 | 0.988 | 0.958 |
| SFADNS-EKF | **0.721** | **0.793** | **0.899** | 0.957 | 0.967 | 0.986 | 0.997 | 1.001 | 0.995 | 0.996 | 0.972 | 1.013 | 1.007 | 1.080 | 1.057 | 1.057 | 1.024 |
| SFADNS-UKF | **0.721** | **0.796** | **0.895** | 0.950 | 0.968 | 0.983 | 0.992 | 0.994 | 0.989 | 0.980 | 0.950 | 0.966 | 0.966 | 1.007 | 0.992 | 1.012 | 0.976 |

Panel B: 6-months ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | 0.982 | 0.989 | 0.996 | 1.001 | 1.005 | 1.008 | 1.010 | 1.012 | 1.015 | 1.018 | 1.020 | 1.028 | 1.035 | 1.042 | 1.043 | 1.045 | 1.045 |
| EW-UKF | 0.983 | 0.991 | 0.997 | 1.002 | 1.007 | 1.010 | 1.013 | 1.015 | 1.019 | 1.023 | 1.028 | 1.037 | 1.046 | 1.055 | 1.056 | 1.060 | 1.060 |
| RMSPE-AVG-EKF | 0.982 | 0.989 | 0.996 | 1.001 | 1.005 | 1.008 | 1.010 | 1.012 | 1.015 | 1.018 | 1.020 | 1.028 | 1.034 | 1.042 | 1.042 | 1.045 | 1.045 |
| RMSPE-AVG-UKF | 0.983 | 0.990 | 0.997 | 1.002 | 1.006 | 1.010 | 1.013 | 1.015 | 1.019 | 1.023 | 1.027 | 1.037 | 1.045 | 1.054 | 1.055 | 1.059 | 1.059 |
| RMSPE-MAT-EKF | 0.982 | 0.990 | 0.996 | 1.001 | 1.005 | 1.008 | 1.010 | 1.012 | 1.015 | 1.018 | 1.020 | 1.028 | 1.034 | 1.041 | 1.042 | 1.043 | 1.043 |
| RMSPE-MAT-UKF | 0.983 | 0.991 | 0.997 | 1.002 | 1.007 | 1.010 | 1.013 | 1.015 | 1.019 | 1.023 | 1.027 | 1.037 | 1.045 | 1.053 | 1.053 | 1.056 | 1.055 |
| SFADNS-EKF | **0.921** | **0.941** | 0.956 | 0.966 | 0.972 | 0.976 | 0.981 | 0.985 | 0.987 | 0.992 | 0.996 | 1.021 | 1.040 | 1.073 | 1.083 | 1.101 | 1.128 |
| SFADNS-UKF | **0.916** | **0.937** | 0.951 | 0.959 | 0.963 | 0.966 | 0.969 | 0.971 | 0.973 | 0.976 | 0.981 | 1.005 | 1.027 | 1.059 | 1.075 | 1.098 | 1.129 |

Panel C: 12-months ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | 1.020 | 1.020 | 1.022 | 1.025 | 1.027 | 1.029 | 1.032 | 1.034 | 1.038 | 1.043 | 1.053 | 1.066 | 1.076 | 1.090 | 1.097 | 1.104 | 1.108 |
| EW-UKF | 1.022 | 1.022 | 1.024 | 1.027 | 1.029 | 1.032 | 1.035 | 1.038 | 1.043 | 1.048 | 1.059 | 1.075 | 1.087 | 1.103 | 1.111 | 1.120 | 1.125 |
| RMSPE-AVG-EKF | 1.020 | 1.021 | 1.023 | 1.025 | 1.027 | 1.029 | 1.032 | 1.034 | 1.038 | 1.043 | 1.052 | 1.065 | 1.076 | 1.089 | 1.096 | 1.102 | 1.106 |
| RMSPE-AVG-UKF | 1.022 | 1.023 | 1.025 | 1.027 | 1.029 | 1.032 | 1.035 | 1.038 | 1.042 | 1.048 | 1.059 | 1.074 | 1.085 | 1.101 | 1.109 | 1.117 | 1.122 |
| RMSPE-MAT-EKF | 1.021 | 1.021 | 1.023 | 1.025 | 1.027 | 1.029 | 1.032 | 1.034 | 1.038 | 1.043 | 1.052 | 1.065 | 1.075 | 1.088 | 1.094 | 1.099 | 1.101 |
| RMSPE-MAT-UKF | 1.023 | 1.023 | 1.025 | 1.027 | 1.029 | 1.032 | 1.035 | 1.038 | 1.042 | 1.048 | 1.058 | 1.073 | 1.084 | 1.099 | 1.106 | 1.112 | 1.115 |
| SFADNS-EKF | 1.023 | 1.029 | 1.034 | 1.037 | 1.040 | 1.043 | 1.047 | 1.051 | 1.055 | 1.061 | 1.076 | 1.101 | 1.123 | 1.154 | 1.173 | 1.192 | 1.211 |
| SFADNS-UKF | 1.022 | 1.027 | 1.030 | 1.031 | 1.031 | 1.031 | 1.033 | 1.034 | 1.035 | 1.039 | 1.051 | 1.074 | 1.097 | 1.128 | 1.150 | 1.175 | 1.196 |

Table 11: Root mean squared prediction errors (RMSPE) ratios for the period 2005:01-2009:12. The table shows the relative RMSPE of the combined forecasts using various weighting schemes against the standard DNS at a forecast horizon of one, six and twelve months. The results per forecast horizon are in Panel A, B and C, respectively. The shades of the cells indicate the rank of the model per maturity. The darker the shade, the better the model. Values in bold denote significance at the 5% level or less.

and irrespective of the subsample that is considered. For six-month ahead forecasts, this holds only when the stable period is considered. The contrary is found when the relatively highly volatile period is examined. For the other cases, the difference between the weighting schemes is marginal. At the same time, the various schemes can be compared to the SFADNS. In general, I find that the combined forecasts dominate the forecasts from the SFADNS in terms of RMSPE ratios, irrespective of the forecast horizon. A notable exception is the period January 2005 until December 2009 at a monthly and semiannually forecast horizon. For this case, the SFADNS provides remarkably more accurate forecasts then the combined forecasts. Stemming from the fact that the macro-factors contain valuable information for the loading parameter, which may result in a better fit. Moreover, this is the only property not included in the combined forecasts.

The results from the various weighting schemes can be used to further emphasize the difference between the EKF and the UKF in an out-of-sample setting. Based purely on the results reported in Tables 9, 10 and 11, I find the difference in the RMSPE ratios to be marginal. Therefore, I analyze the details of the weighting schemes. More specifically, I compare the relative number of times that a model gets the largest weight, during the forecasting period and per maturity, which is showed in Figure 10. As the only difference is the filtering technique that is used for the DNS with time-varying loading, a comparison between the techniques can be made based on the subfigures.

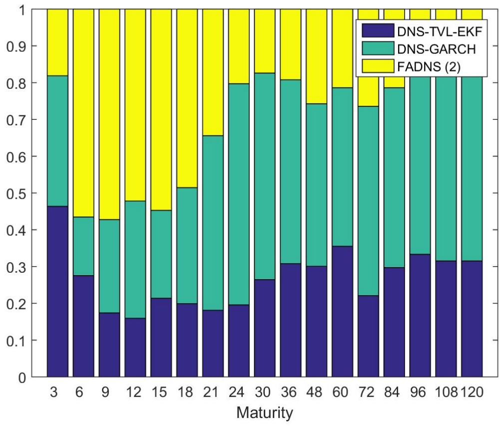
At the monthly forecast horizon, I observe that the DNS-TVL-UKF is often chosen to be the best model (in terms of RMSPE) for both ends of the yield curve. Interestingly, this observations corresponds to the finding for the in-sample framework. This makes sense as the forecast horizon is short. For the semi- and annually forecast horizon, the DNS-TVL-EKF appears to obtain rank 1 more often.

## 4.5 Out-of-Sample Analysis: Robustness Check

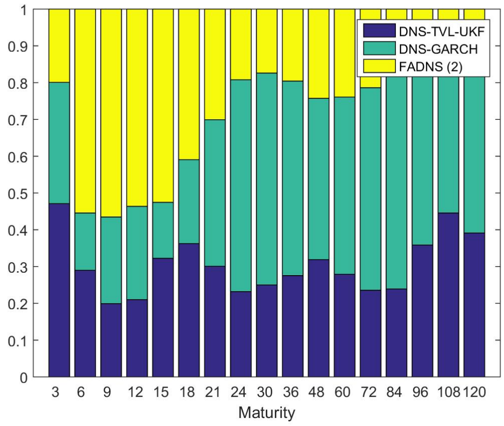
In this subsection, I describe the check for robustness of the main empirical results. More specifically, I focus on the choice of the amount of factors to be included in the FADNS and another weighting scheme based on the RMSPE. To preserve space I report the RMSPE ration in Appendix D and E.

### 4.5.1 Number of Included Macro-Factors

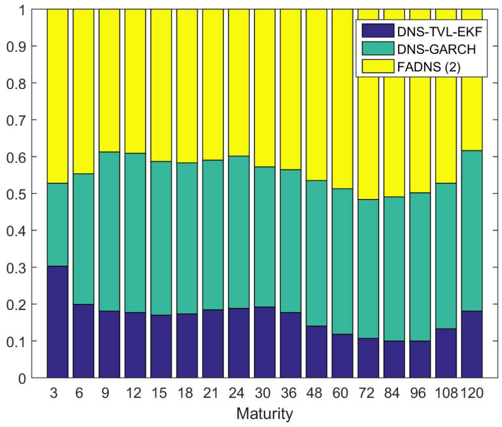
I consider the forecasts of the yields for the different forecasting horizons and subsamples for various choices of $p$. Tables 16, 17 and 18 report the RMSPE ratios for this robustness check. The evaluation of the forecasts made in the full out-of-sample period show that including four macro-factors in the FADNS results in a RMSPE ratio which is on average the lowest among those of the considered number of macro-factors. Moreover, for the short-end of the yield curve the FADNS with four macro-factors provides more accurate forecasts than the DNS at a forecast horizon of one- and six-months. Similar results are found for the January 2005 until December 2009, in which the yields are highly volatile. When the yields are relatively stable, the difference between the results of using two macro-factors in the FADNS and using either one or three macro-factors in the model is marginal. Although, the FADNS with one macro-factor slightly outperforms the others. Moreover, the latter model is in general able to beat the DNS at forecasting the yield for various maturities, irrespective of the forecast horizon that is considered. The gain in predictive accuracy rises to a maximum of 6%.



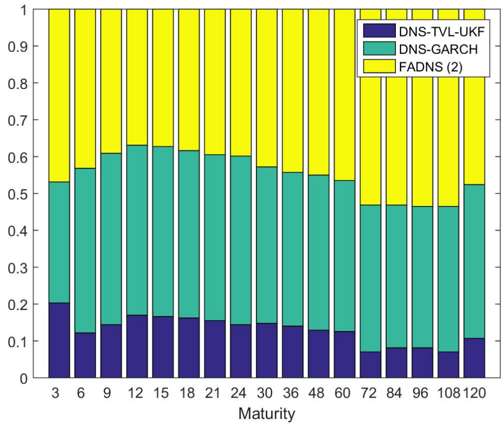

(a) EKF, h = 1



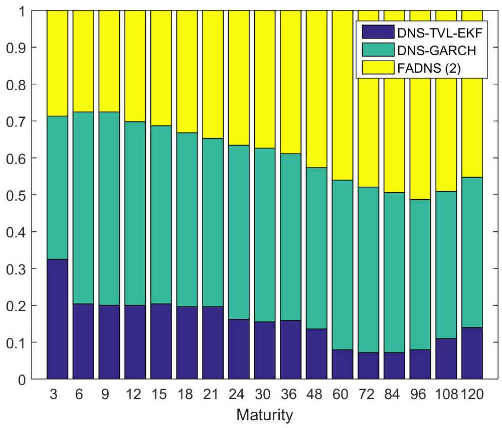

(b) UKF, h = 1



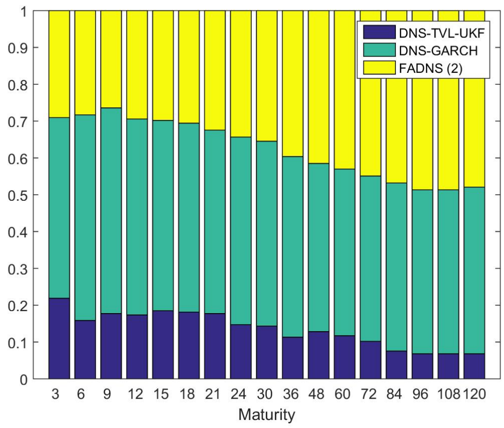

(c) EKF, h = 6




(d) UKF, h = 6




(e) EKF, h = 12




(f) UKF, h = 12

Figure 10: Relative number of times that the model gets a rank equal to 1 per maturity. This figure shows how often the DNS-TVL-EKF (or UKF), the DNS-GARCH or the FADNS with two macro factors is considered to be the relatively best forecasting model in terms of RMSPE (using a window of two years prior to the calculation of it), during the predictive period and per maturity. The best model then gets a rank 1 in the maturity-specific weighting scheme.

### 4.5.2 Rank-Based Combination Scheme

As the RMSPEs of the FADNS, the DNS-GARCH and the DNS-TVL are nearly equal to each other, the RMSPE-based weighting scheme is almost the same as the equal weighted combination scheme. To really differentiate the weights in the combination, one could opt for giving larger weights to the best model. Aiolfi and Timmermann (2006) do this by introducing a ranks in the combination scheme based on the RMSPE. Similar to them, I use a rolling window of 2 years to rank the models based on this statistical evaluation criterion. The best model gets a rank of 1, the second best a rank of 2, and the relative worst model gets a rank of 3. Now, let the evaluation period be 2 years prior to time $t$ up to time $t$, the forecast from model $i$ at time $t+h$ then gets the weight

$$w^i_{t+h|t} = \frac{1/\text{Rank}^i_t}{\sum_{i=1}^{3} 1/\text{Rank}^i_t}, \tag{53}$$

to construct the combined forecasts for time $t+h$. Moreover, $\text{Rank}^i_t$ denotes the rank of model $i$ according to the RMSPE over the period 2 years prior to time $t$ up until time $t$. Details on the different schemes; in the second scheme I look at the average RMSPE across the maturities in the rolling window per model, and rank the models according this average. While in the third scheme, I look at the RMSPE in the rolling window per maturity, and rank the models according to this. This means that the weights not only differ over time, but also differ across the maturities. The results are shown in Tables 19, 20 and 21. By comparing the rank-based combination scheme against the RMSPE-based combination scheme, I find that the rank-based combination scheme provides slightly more accurate forecasts, but the differences are marginal.

## 5 Conclusion

This study examines the out-of-sample forecast performance of various extensions in the Nelson-Siegel framework relative to the dynamic Nelson-Siegel (DNS) model. More specifically, I consider the factor augmented dynamic Nelson-Siegel (FADNS) model by Diebold et al. (2006), the GARCH extension of the DNS (DNS-GARCH) and the DNS with time-varying loadings (DNS-TVL), which are both introduced by Koopman et al. (2010). Additionally, I introduce the stochastic factor augmented DNS (SFADNS), which combines the characteristics of the mentioned DNS extensions.

By evaluating the RMSPE ratios of these models, I find that it is generally not easy to improve upon forecasts made by the standard DNS for the full out-of-sample period (January 1985 until December 2009), although gains in the predictive accuracy up to 14% can still be attained in several cases. The results are more promising when different subsamples are considered. When there is little volatility in yields, as in the period January 1995 until December 1999, the drop in the RMSPE reaches up to 10%. On the other hand, when volatility is relatively high, as in the period January 2005 until December 2009, gains of nearly 28% are attainable. Hence, the gains are highly sensitive to the period that is considered. Moreover, for stable periods and short- to intermediate-term yields, the DNS-GARCH is preferred over the DNS for all forecast horizons. This result corresponds to the finding Hautsch and Yang (2010), who incorporate stochastic volatility via the transition equation instead. As a reason, they argue that the standard DNS has a higher forecasting uncertainty for periods of low-volatility, which stems from the fact that the ignored stochastic volatility and parameter uncertainty in periods of high-volatility spread to periods when the yields are more stable.On the other hand, in case of relatively highly volatile yields, the SFADNS should be chosen as its out-of-sample forecast performance dominates that of the DNS at a forecast horizon of one and six months. When accurate forecasts for the long-end of the yield curve should be made the DNS is in general the most reliable model.

To understand why the SFADNS is able to attain such values, I focus on forecast combinations as well. Using three weighting schemes (equal weights, time-varying weights based on the cross-sectional average of the RMSPE, and maturity-specific time-varying weights) I find that the combining forecasts leads to additional gains in predictive accuracy, particularly for stable periods. Moreover, on average, the maturity-specific weighting scheme provides more accurate forecasts than the other two schemes for the short-end of the yield curve. However, the combined forecasts are dominated by the forecasts from the SFADNS in periods where the yields are relatively highly volatile. A possible explanation is that the macro-factors contain valuable information for the loading parameter in such periods (this is the only property that the combined forecasts do not exhibit).

The results show that the relative forecasting performance of the different models vary across the yield curve. Particularly, the forecasts from the extensions for the short-end of the yield curve perform remarkably well, dominating the forecasts from the DNS in up to 81% of the cases. For yields with an intermediate maturity, the standard DNS is sufficient for forecasting. In fact, there are only a few cases for which the extensions of the DNS provide more accurate forecasts of the intermediate-term yields.

There has been, to my knowledge, no comparison between two estimation techniques for nonlinear Gaussian state-space models, the extended Kalman filter (EKF) and the unscented Kalman filter (UKF), in the context of nonlinear Nelson-Siegel models. Therefore, my main contribution to the large strand of research on the Nelson-Siegel models is the examination of the differences between the filtering techniques. By evaluating the in-sample fit, I find that the estimates of the EKF are dominated by those of the UKF for both ends of the yield curve; observing a decrease in the mean of the filtered error of approximately 96%. On the other hand, for intermediate-term yields the EKF is preferred. A possible explanation for this is the inability of the extended Kalman filter to accurately approximate the curvature factor. In the out-of-sample framework, the results are contrasting. Moreover, I find that the forecasts from the EKF are generally more accurate than those from the UKF when the DNS-TVL is used. However, when I allow for more flexibility in the model, by considering the SFADNS, I find the converse with gains reaching up to 4% when the UKF is used. The contrasting results may stem from the fact that the EKF is not able to provide more accurate estimates and forecasts than those from the UKF when the number of latent factors and unknown parameters increase. A more in-depth analysis regarding the filtering techniques has been conducted in the forecast combination part. By using the ranking system of the maturity-specific weighting scheme, I find that the DNS-TVL-UKF is chosen more often to be the best model in terms of RMSPE for monthly forecasting both ends of the yield curve. For other cases, the DNS-TVL-UKF is dominated by the DNS-TVL-EKF.

To conclude, I find that there is no best Nelson-Siegel model in the forecasting framework. The results imply that the practitioner should take into account the circumstances of the current yield (stable or relatively volatile), and the maturity for which the yield is forecasted. Furthermore, the findings indicate that the use of the UKF over the EKF for fitting nonlinear Nelson-Siegel models is beneficial for both ends of the yield curve and could have a positive impact on the accuracy in the predictive framework in some cases.

# 6 Further Research

The findings presented in this thesis offer several directions for further research. I discuss six suggestions I find particularly interesting. First, PCA is commonly used to extract factors from a set of variables. However, it fails to take the prediction objective into account when constructing factors as argued by Heij et al. (2007). As an alternative, they propose principal covariate regression (PCovR) as a solution to this. Furthermore, other factor extraction techniques, e.g. partial least squares, can be considered as well. This is similar to the study of Exterkate et al. (2013), but conducted on a monthly updated dataset, which is easily accessible. Second, other specifications of the GARCH model can be considered. For example the threshold GARCH could be used, which accounts for the asymmetry in the financial markets. Also worth mentioning, is the exponential GARCH, for which no restrictions are needed to ensure that the volatility is positive and stationary. Third, the idea of allowing the loading parameter $\lambda_t$ to be time-varying could also be applied to the dynamic Nelson-Siegel-Svensson (DNSS) model (Svensson, 1995). This model extends the DNS with a second curvature factor such that the Svensson extension often fits better than the DNS at long maturities. The DNSS could then be adjusted by allowing the loading parameter to vary over time, while keeping it fixed for one of the curvature factors (to avoid identification issues). This should lead to an improved fit over the DNS-TVL as the DNSS is able to fit more maturities in the cross-section. Fourth, another interesting paths for further research is the inclusion of stochastic volatility and a time-varying loading parameter in the DNS with shifting endpoints, as introduced by van Dijk et al. (2013). Fifth, allowing the loading parameter to vary over time is one way of incorporating nonlinearities in the DNS model. Another method is to include nonlinear regime-switching yield factor dynamics in the model, in the tradition of Hamilton (1989). For these type of models, the difference between the EKF and the UKF can be examined. Finally, in this thesis I did not consider the relaxation of the assumption of Gaussian distributed shocks. For example, Mesters and Koopman (2014) show that the accuracy for the factor estimates greatly improves when the errors are assumed to be Student's $t$ distributed. Additionally, when the model is nonlinear and non-Gaussian different estimation techniques, such as the interesting particle filter, should be used.

## References


Ang, A. and Piazzesi, M. (2003). A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables. *Journal of Monetary Economics*, 50(4), pp. 74587

Ang, A., G. Bekaert and M. Wei (2007). Do macro variables, asset markets or surveys forecast inflation better?. *Journal of Monetary Economics*, 54, 1163-1212

Bikbov, R. and M. Chernov (2010). No-arbitrage macroeconomic determinants of the yield curve. *Journal of Econometrics*, 159, 166-182

Christensen, Jens H. E., Francis X. Diebold, and Glenn D. Rudebusch (2009). An Arbitrage Free Generalized Nelson-Siegel Term Structure Model. *Econometrics Journal*, Vol. 12, No. 3, C33-C64

Christensen, Jens H.E. & Diebold, Francis X. & Rudebusch, Glenn D. (2011). The affine arbitrage-free class of Nelson-Siegel term structure models. *Journal of Econometrics*, Elsevier, vol. 164(1), pages 4-20

Clark, T.E., McCracken, M.W., (2005). Evaluating direct multistep forecasts. *Econometric Reviews* 24, 369404

Cox, C.J, J.E Ingersoll, Jr. and S.A Ross (1985). A thepry of the term structure of interest rates. *Econometrica*, Volume 53, Issue 2, 385-408

Dai, Q. and T. Philippon (2006). Fiscal Policy and the Term Structure of Interest Rates. *National Bureau of Economic Research, Inc.*

Dewachter, H. and M. Lyrio (2006). Macro factors and the term structure of interest rates. *Journal of Money, Credit and Banking*, 38, 119-140

Diebold, F. X. and Li, C. (2006). Forecasting the Term Structure of Government Bond Yields. *Journal of Econometrics*, Volume 130, 337364

Diebold, F.X., & Rudebusch, G. (2013). Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach. *Princeton University Press*

Diebold, F.X. (2007). *Elements of Forecasting*, Thomson South-Western Publishing, 4th ed.

Diebold, F.X., Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics* 130, 337364

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13, 253263.

Diebold, F.X., M. Piazzesi and G.D. Rudebusch (2006). Modeling Bond Yields in Finance and Macroeconomics. *American Economic Review*, 95, 415-420

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

Dijk, D.J.C. van, Koopman, S.J., Wel, M. van der & Wright, J (2014). Forecasting interest rates with shifting endpoints. *Journal of Applied Econometrics*, 29, 693-712

Duffee, G.R (2002). Term Premia and Interest Rate Forecasts in Affine Models. *The journal of finance*, Volume 57, Issue 1, pages 405-443

Duffie, D and R. Kan (1996). A yield-factor model of interest rates. *Mathematical Finance*, Volume 6, Issue 4, pages 379-406

Engle, R.F. and Watson, M.W. (1983). Alternative algorithms for the estimation of dynamic factor, mimic and varying coefficient regression models. *Journal of Econometrics* 23(3):385-400

Estrella, A., and F. S. Mishkin (1997). The Predictive Power of the Term Structure of Interest Rates in Europe and The United States: Implications for The European Central Bank. *European Economic Review* 41:1375-1401

Estrella, A., and G. Hardouvelis (1991). The Term Structure as a Predictor of Real Economic Activity. *Journal of Finance* 46:555-576

Exterkate, P., Dijk, D.J.C. van, 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*, vol. 32(3), pages 193-214

Fama, E.F. (1990). Stock Returns, Expected Returns and Real Activity. *The Journal of Finance* Vol. 45, No.4 pp. 1089-1108

Fama, E.F. and R.R. Bliss (1987). The information in long-maturity forward rates. *American Economic Review* 77, pages 680-692

Filipovic, D. (1999). A Note on the NelsonSiegel Family. *Mathematical Finance*, Volume 9, Issue 4, pages 349-359

Geweke, J. (1977). The Dynamic Factor Analysis of Economic Time Series. *Latent Variables in Socio-Economic Models*[vii].

Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica* 74, 15451578.

Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57, 357-384

Harvey, A.C., E. Ruiz and E. Sentana (1992). Unobserved Component Time Series Models with ARCH Disturbances. *Journal of Econometrics*, 52, 129-158

Hautsch, N. & F. Yang (2012). Bayesian inference in a Stochastic Volatility NelsonSiegel model.

[vii]ed. by D.J. Aigner and A.S. Goldberger, Amsterdam: North-Holland

*Computational Statistics & Data Analysis, Elsevier*, vol. 56(11), pages 3774-3792

Hautsch, N. & Y. Ou, (2008). Yield Curve Factors, Term Structure Volatility, and Bond Risk Premia. *SFB 649 Discussion Papers SFB649DP2008-053, Sonderforschungsbereich 649, Humboldt University, Berlin, Germany*

Heij C., Groenen P.J.F., van Dijk D. (2007). Forecast comparison of principal component regression and principal covariate regression. *Computational Statistics and Data Analysis* 51: 36123625

Hordahl, P., O. Tristani and D. Vestin (2006). A joint Econometric model of macroeconomic and term-structure dynamics. *Journal of Econometrics*, 131, 405-444

Hull, J. and A. White (1990). Pricing interest-rate derivative securities. *The Review of Financial Studies*, Vol 3, No.4 pp. 573592

James H. Stock and Mark W. Watson (2002). Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the American Statistical Association*, Vol. 97, No. 460, pp. 1167-1179

Julier, S.J. (2000). The Scaled Unscented Transformation. *Automatica, February 2000*

Julier, S.J. and J.K. Uhlmann (1997). A New Extension of the Kalman Filter to Nonlinear Systems. *SPIE Proceedings* Vol. 3068

Julier, S.J. and J.K. Uhlmann (2004). Unscented Filtering and Nonlinear Estimation. *Proceedings of the IEEE* 92(3): 401-422

Kilian, L. (1998). Small-sample confidence intervals for impulse response functions. *Review of Economics and Statistics* 80, 218230.

Kim, D.H. and J.H. Wright (2005). An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates. *Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series*:200533.

Koopman, S.J., Mallee M., van der Wel, M. (2010). Analyzing the Term Structure of Interest Rates Using the Dynamic Nelson-Siegel Model With Time-Varying Parameters. *Journal of Business and Economic Statistics*, 28(3):329-343

Litterman, R., and J. A. Scheinkman. (1991). Common Factors Affecting Bond Returns. *Journal of Fixed Income* 1.

Mesters, G. and Koopman, S. J. (2014). Generalized Dynamic Panel Data Models with Random Effects for Cross-Section and Time. *Journal of Econometrics* 180, 127140.

Mishkin, F. S. (1990). What Does the Term Structure Tell Us about Future Inflation?. *Journal of Monetary Economics* 25, 77-95

Moench, E. (2008). Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach. *Journal of Econometrics*, Volume 146, Issue 1, Pages 2643

Nelson, C., and Siegel, A.F. (1987). Parsimonious Modeling of Yield Curves. *Journal of Business*, Volume 60, 473489

Nelson, C.R. and A.F. Siegel (1987). Parsimonious Modeling of Yield Curves. *The Journal of Business*, Volume 60, Issue 4, pages 473-489

Pooter, M.D. de, Ravazzolo, F. & Dijk, D.J.C. van (2010). Term structure forecasting using macro factors and forecast combination. (Preprints). *Washington, D.C.: Federal Reserve Board of Governors International Finance Discussion Papers* No. 2010-993.

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

Sargent, T.J., and C.A. Sims (1977). Business Cycle Modeling Without Pretending to Have Too Much A-Priori Economic Theory. *New Methods in Business Cycle Research,*[viii].

Stock, J.H., M.W. Watson (2002). Forecasting Using Principal Components From a Large Number of Predictors. *Journal of the American Statistical Association*, Vol. 97, No. 460, pages 1167-1179

Stock, J.H., M.W. Watson (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* 20: 147162

Svensson, L.E.O. (1995). Estimating Forward Interest Rates with the Extended Nelson-Siegel Method. *Sveriges Riksbank Quarterly Review*, 3, 13-26

Sydney C. Ludvigson & Serena Ng, (2009). Macro Factors in Bond Risk Premia. *Review of Financial Studies, Oxford University Press for Society for Financial Studies*, vol. 22(12), pages 5027-5067

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics* vol. 5, issue 2, pages 177-188

Wan, E. A. and van der Merwe, R. (2001) The Unscented Kalman Filter, in Kalman Filtering and Neural Networks. *John Wiley & Sons, Inc., New York, USA*

West, K. D. (2006). Forecast evaluation, in G. Elliott, C. W. Granger and A. Timmermann, eds, Handbook of Economic Forecasting, Elsevier, Amsterdam.

Wu, S. (2008). Monetary Policy And Long-Term Interest Rates. *Western Economic Association International*, vol. 26(3), pages 398-408

Yu, W. and E. Zivot (2010). Forecasting the Term Structures of Treasury and Corporate Yields: Dynamic Nelson-Siegel Models Evaluation. *International Journal of Forecasting*, Volume 27, Issue 2, Pages 579591

[viii] ed. by C. Sims et al., Minneapolis: Federal Reserve Bank of Minneapolis

# A The Unconditional Covariance Matrix of the State Vector

In order to solve $\boldsymbol{V} = \boldsymbol{KVK}' + \boldsymbol{\Sigma}_w$ for $\boldsymbol{V}$ I follow the approach of Christensen and van der Wel (2010). First, I rewrite the equation such that I obtain the unconditional covariance matrix at one side of the equal sign. This results in $\boldsymbol{V} - \boldsymbol{KVK}' = \boldsymbol{\Sigma}_w$. Subsequently, I apply the vectorization operator to both side of the equation, giving

$$vec(\boldsymbol{V}) - vec(\boldsymbol{KVK}') = vec(\boldsymbol{\Sigma}_w),$$

which can be rewritten as

$$\begin{aligned} &\boldsymbol{I}_{dim^2} vec(\boldsymbol{V}) - (\boldsymbol{K} \otimes \boldsymbol{K}) vec(\boldsymbol{V}) = \\ &\quad \Big[\boldsymbol{I}_{dim^2} - (\boldsymbol{K} \otimes \boldsymbol{K})\Big] vec(\boldsymbol{V}) = vec(\boldsymbol{\Sigma}_w), \end{aligned}$$

where $dim^2$ is the squared dimension of the state vector. Now, the above can be solved for $\boldsymbol{V}$, which result in the following

$$vec(\boldsymbol{V}) = \Big[\boldsymbol{I}_{dim^2} - (\boldsymbol{K} \otimes \boldsymbol{K})\Big]^{-1} vec(\boldsymbol{\Sigma}_w).$$

# B The Macroeconomic Dataset

| Panel A: | Output and income | |
|---|---|---|
| No. | Transformation | Variable |
| 1 | 5 | Real Personal Income |
| 2 | 5 | Real personal income ex transfer receipts |
| 3 | 5 | IP Index |
| 4 | 5 | IP: Final Products and Nonindustrial Supplies |
| 5 | 5 | IP: Final Products (Market Group) |
| 6 | 5 | IP: Consumer Goods |
| 7 | 5 | IP: Durable Consumer Goods |
| 8 | 5 | IP: Nondurable Consumer Goods |
| 9 | 5 | IP: Business Equipment |
| 10 | 5 | IP: Materials |
| 11 | 5 | IP: Durable Materials |
| 12 | 5 | IP: Nondurable Materials |
| 13 | 5 | IP: Manufacturing (SIC) |
| 14 | 5 | IP: Residential Utilities |
| 15 | 5 | IP: Fuels |
| 16 | 1 | ISM Manufacturing: Production Index |
| 17 | 2 | Capacity Utilization: Manufacturing |

Table 12: The macroeconomic dataset. The table lists the individual macro series of the dataset that are used to contruct macro factors. The series are categorized in 8 groups: (A) output and income. The transformations applied to original series are coded by the FRED as: (1) no transformation (levels are used); (2) $\Delta x_t$ (the difference); (3) $\Delta^2 x_t$ (the squared differences); (4) $\log(x_t)$ (logarithm of series); (5) $\Delta \log(x_t)$ (difference of the logarithm of series); (6) $\Delta^2 \log(x_t)$ (squared difference of the logarithm of series); (7) $\Delta(x_t/x_{t-1} - 1)$ (difference of the growth). The description of the remaining groups are down below.

| Panel B: | Labor market | |
|---|---|---|
| No. | Transformation | Variable |
| 18 | 2 | Help-Wanted Index for United States |
| 19 | 2 | Ratio of Help Wanted/No. Unemployed |
| 20 | 5 | Civilian Labor Force |
| 21 | 5 | Civilian Employment |
| 22 | 2 | Civilian Unemployment Rate |
| 23 | 2 | Average Duration of Unemployment (Weeks) |
| 24 | 5 | Civilians Unemployed - Less Than 5 Weeks |
| 25 | 5 | Civilians Unemployed for 5-14 Weeks |
| 26 | 5 | Civilians Unemployed - 15 Weeks & Over |
| 27 | 5 | Civilians Unemployed for 15-26 Weeks |
| 28 | 5 | Civilians Unemployed for 27 Weeks and Over |
| 29 | 5 | Initial Claims |
| 30 | 5 | All Employees: Total nonfarm |
| 31 | 5 | All Employees: Goods-Producing Industries |
| 32 | 5 | All Employees: Mining and Logging: Mining |
| 33 | 5 | All Employees: Construction |
| 34 | 5 | All Employees: Manufacturing |
| 35 | 5 | All Employees: Durable goods |
| 36 | 5 | All Employees: Nondurable goods |
| 37 | 5 | All Employees: Service-Providing Industries |
| 38 | 5 | All Employees: Trade, Transportation & Utilities |
| 39 | 5 | All Employees: Wholesale Trade |
| 40 | 5 | All Employees: Retail Trade |
| 41 | 5 | All Employees: Financial Activities |
| 42 | 5 | All Employees: Government |
| 43 | 1 | Avg Weekly Hours : Goods-Producing |
| 44 | 2 | Avg Weekly Overtime Hours : Manufacturing |
| 45 | 1 | Avg Weekly Hours : Manufacturing |
| 46 | 1 | ISM Manufacturing: Employment Index |
| 47 | 6 | Avg Hourly Earnings : Goods-Producing |
| 48 | 6 | Avg Hourly Earnings : Construction |
| 49 | 6 | Avg Hourly Earnings : Manufacturing |

| Panel C: | Housing | |
|---|---|---|
| No. | Transformation | Variable |
| 50 | 4 | Housing Starts: Total New Privately Owned |
| 51 | 4 | Housing Starts, Northeast |
| 52 | 4 | Housing Starts, Midwest |
| 53 | 4 | Housing Starts, South |
| 54 | 4 | Housing Starts, West |
| 55 | 4 | New Private Housing Permits (SAAR) |
| 56 | 4 | New Private Housing Permits, Northeast (SAAR) |
| 57 | 4 | New Private Housing Permits, Midwest (SAAR) |
| 58 | 4 | New Private Housing Permits, South (SAAR) |
| 59 | 4 | New Private Housing Permits, West (SAAR) |

Table 13: The macroeconomic dataset (continued).

| Panel D: | Consumption, orders, and inventories | |
|---|---|---|
| No. | Transformation | Variable |
| 60 | 5 | Real personal consumption expenditures |
| 61 | 5 | Real Manu. and Trade Industries Sales |
| 62 | 5 | Retail and Food Services Sales |
| 63 | 1 | ISM : PMI Composite Index |
| 64 | 1 | ISM : New Orders Index |
| 65 | 1 | ISM : Supplier Deliveries Index |
| 66 | 1 | ISM : Inventories Index |
| 67 | 5 | New Orders for Consumer Goods |
| 68 | 5 | New Orders for Durable Goods |
| 69 | 5 | New Orders for Nondefense Capital Goods |
| 70 | 5 | Unlled Orders for Durable Goods |
| 71 | 5 | Total Business Inventories |
| 72 | 2 | Total Business: Inventories to Sales Ratio |
| 73 | 2 | Consumer Sentiment Index |

| Panel E: | Money and credit | |
|---|---|---|
| No. | Transformation | Variable |
| 74 | 6 | M1 Money Stock |
| 75 | 6 | M2 Money Stock |
| 76 | 5 | Real M2 Money Stock |
| 77 | 6 | St. Louis Adjusted Monetary Base |
| 78 | 6 | Total Reserves of Depository Institutions |
| 79 | 7 | Reserves Of Depository Institutions |
| 80 | 6 | Commercial and Industrial Loans |
| 81 | 6 | Real Estate Loans at All Commercial Banks |
| 82 | 6 | Total Nonrevolving Credit |
| 83 | 2 | Nonrevolving consumer credit to Personal Income |
| 84 | 6 | MZM Money Stock |
| 85 | 6 | Consumer Motor Vehicle Loans Outstanding |
| 86 | 6 | Total Consumer Loans and Leases Outstanding |
| 87 | 6 | Securities in Bank Credit at All Commercial Banks |

Table 14: The macroeconomic dataset (continued).

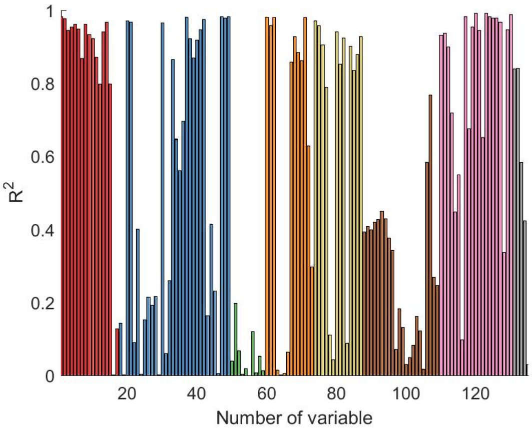
| Panel F: | Interest and exchange rates | |
|---|---|---|
| No. | Transformation | Variable |
| 88 | 2 | Effective Federal Funds Rate |
| 89 | 2 | 3-Month AA Financial Commercial Paper Rate |
| 90 | 2 | 3-Month Treasury Bill: |
| 91 | 2 | 6-Month Treasury Bill: |
| 92 | 2 | 1-Year Treasury Rate |
| 93 | 2 | 5-Year Treasury Rate |
| 94 | 2 | 10-Year Treasury Rate |
| 95 | 2 | Moody's Seasoned Aaa Corporate Bond Yield |
| 96 | 2 | Moody's Seasoned Baa Corporate Bond Yield |
| 97 | 1 | 3-Month Commercial Paper Minus FEDFUNDS |
| 98 | 1 | 3-Month Treasury C Minus FEDFUNDS |
| 99 | 1 | 6-Month Treasury C Minus FEDFUNDS |
| 100 | 1 | 1-Year Treasury C Minus FEDFUNDS |
| 101 | 1 | 5-Year Treasury C Minus FEDFUNDS |
| 102 | 1 | 10-Year Treasury C Minus FEDFUNDS |
| 103 | 1 | Moody's Aaa Corporate Bond Minus FEDFUNDS |
| 104 | 1 | Moody's Baa Corporate Bond Minus FEDFUNDS |
| 105 | 5 | Trade Weighted U.S. Dollar Index: Major Currencies |
| 106 | 5 | Switzerland / U.S. Foreign Exchange Rate |
| 107 | 5 | Japan / U.S. Foreign Exchange Rate |
| 108 | 5 | U.S. / U.K. Foreign Exchange Rate |
| 109 | 5 | Canada / U.S. Foreign Exchange Rate |

| Panel G: | Prices | |
|---|---|---|
| No. | Transformation | Variable |
| 110 | 6 | PPI: Finished Goods |
| 111 | 6 | PPI: Finished Consumer Goods |
| 112 | 6 | PPI: Intermediate Materials |
| 113 | 6 | PPI: Crude Materials |
| 114 | 6 | Crude Oil, spliced WTI and Cushing |
| 115 | 6 | PPI: Metals and metal products: |
| 116 | 1 | ISM Manufacturing: Prices Index |
| 117 | 6 | CPI : All Items |
| 118 | 6 | CPI : Apparel |
| 119 | 6 | CPI : Transportation |
| 120 | 6 | CPI : Medical Care |
| 121 | 6 | CPI : Commodities |
| 122 | 6 | CPI : Durables |
| 123 | 6 | CPI : Services |
| 124 | 6 | CPI : All Items Less Food |
| 125 | 6 | CPI : All items less shelter |
| 126 | 6 | CPI : All items less medical care |
| 127 | 6 | Personal Cons. Expend.: Chain Index |
| 128 | 6 | Personal Cons. Exp: Durable goods |
| 129 | 6 | Personal Cons. Exp: Nondurable goods |
| 130 | 6 | Personal Cons. Exp: Services |

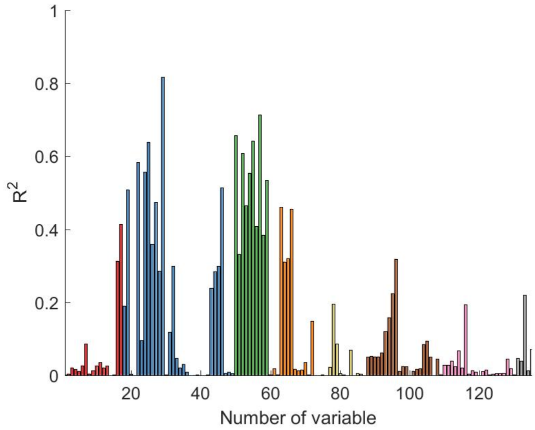
| Panel H: | Stock market | |
|---|---|---|
| No. | Transformation | Variable |
| 131 | 5 | S&P's Common Stock Price Index: Composite |
| 132 | 5 | S&P's Common Stock Price Index: Industrials |
| 133 | 2 | S&P's Composite Common Stock: Dividend Yield |
| 134 | 5 | S&P's Composite Common Stock: Price-Earnings Ratio |
| 135 | 1 | VXO |

Table 15: The macroeconomic dataset (continued).

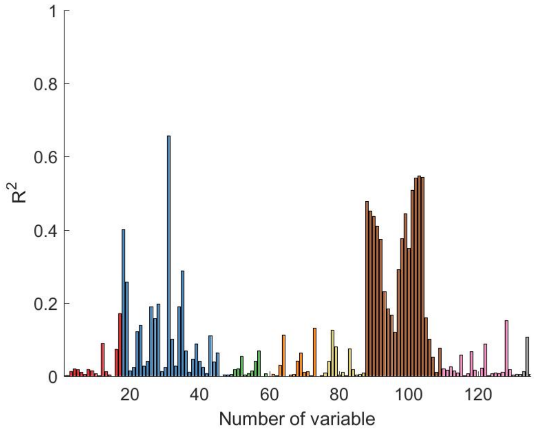
## C $R^2$ in Regressions of Individual Macro Series on PCA Factors




(a) First factor obtained via PCA




(b) Second factor obtained via PCA




(c) Third factor obtained via PCA

Figure 11: $R^2$ of the regression of individual standardized macro series on PCA factors. The figure show the $R^2$ when regressing the individual standardized series in the macro database of the FRED on each of the first three PCA factors. The macro dataset consists of 135 series (transformed to rule out scale effects) over the period 1970:01-2009:12. Panels (a), (b), and (c) show the results for the first, second, and third PCA factor, respectively. In each panel the macro series are grouped and given different colors according to the 8 categories as indicated on the FRED website. The group categories are (red) output and income; (blue) labor market; (green) housing; (orange) consumption, (yellow) orders and inventories; (brown) money and credit; (pink) prices; and (grey) stock market.

## D Robustness Check: Number Included Macro-Factors

Panel A: 1-month ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| FADNS (1) | 1.004 | 1.021 | 1.030 | 1.026 | 1.015 | 1.017 | 1.021 | 1.029 | 1.028 | 1.030 | 1.023 | 1.027 | 1.018 | 1.016 | 1.010 | 1.009 | 1.018 |
| FADNS (2) | 0.993 | 1.004 | 1.013 | 1.009 | 1.001 | 1.004 | 1.009 | 1.018 | 1.016 | 1.019 | 1.015 | 1.019 | 1.011 | 1.011 | 1.007 | 1.006 | 1.015 |
| FADNS (3) | 0.992 | 1.001 | 1.011 | 1.006 | 0.998 | 1.002 | 1.008 | 1.016 | 1.014 | 1.017 | 1.013 | 1.017 | 1.010 | 1.010 | 1.007 | 1.006 | 1.015 |
| FADNS (4) | 0.992 | 0.996 | 1.005 | 0.997 | 0.988 | 0.994 | 1.002 | 1.014 | 1.012 | 1.015 | 1.013 | 1.018 | 1.011 | 1.013 | 1.011 | 1.011 | 1.019 |

Panel B: 6-months ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| FADNS (1) | 1.038 | 1.044 | 1.049 | 1.049 | 1.045 | 1.048 | 1.052 | 1.057 | 1.059 | 1.062 | 1.060 | 1.062 | 1.057 | 1.056 | 1.050 | 1.049 | 1.057 |
| FADNS (2) | 1.011 | 1.017 | 1.022 | 1.020 | 1.016 | 1.019 | 1.023 | 1.029 | 1.030 | 1.033 | 1.033 | 1.037 | 1.034 | 1.035 | 1.031 | 1.031 | 1.040 |
| FADNS (3) | 1.005 | 1.010 | 1.015 | 1.012 | 1.008 | 1.012 | 1.015 | 1.021 | 1.022 | 1.025 | 1.026 | 1.031 | 1.029 | 1.032 | 1.030 | 1.030 | 1.039 |
| FADNS (4) | 0.998 | 0.996 | 0.998 | 0.992 | 0.987 | 0.992 | 0.997 | 1.005 | 1.007 | 1.012 | 1.018 | 1.028 | 1.029 | 1.035 | 1.035 | 1.037 | 1.046 |

Panel C: 12-months ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| FADNS (1) | 1.052 | 1.055 | 1.059 | 1.059 | 1.057 | 1.061 | 1.065 | 1.071 | 1.074 | 1.080 | 1.083 | 1.089 | 1.086 | 1.087 | 1.083 | 1.082 | 1.089 |
| FADNS (2) | 1.050 | 1.050 | 1.051 | 1.047 | 1.043 | 1.044 | 1.046 | 1.050 | 1.051 | 1.054 | 1.056 | 1.061 | 1.060 | 1.061 | 1.058 | 1.059 | 1.065 |
| FADNS (3) | 1.051 | 1.048 | 1.047 | 1.042 | 1.037 | 1.038 | 1.040 | 1.043 | 1.043 | 1.046 | 1.048 | 1.054 | 1.054 | 1.057 | 1.055 | 1.057 | 1.065 |
| FADNS (4) | 1.063 | 1.050 | 1.042 | 1.032 | 1.023 | 1.023 | 1.024 | 1.028 | 1.028 | 1.032 | 1.039 | 1.050 | 1.055 | 1.062 | 1.065 | 1.069 | 1.078 |

Table 16: Root mean squared prediction errors (RMSPE) ratios. The table shows the relative RMSPE of the considered extensions of the DNS against the standard DNS at a forecast horizon of one, six and twelve months. The results per forecast horizon are in Panel A, B and C, respectively, with corresponding periods 1985:01-2009:12, 1985:06-2009:12, 1985:12-2009:12. The shades of the cells indicate the rank of the model per maturity. The darker the shade, the better the model. Values in bold denote significance at the 5% level or less.

Panel A: 1-month ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| FADNS (1) | 1.014 | 1.018 | 1.018 | 0.991 | 0.972 | 0.979 | 0.986 | 0.998 | 0.997 | 1.001 | 0.999 | 1.011 | 1.005 | 0.999 | 0.991 | 0.993 | 1.006 |
| FADNS (2) | 1.027 | 1.024 | 1.023 | 0.995 | 0.973 | 0.981 | 0.988 | 1.001 | 0.999 | 1.002 | 1.000 | 1.012 | 1.004 | 0.999 | 0.992 | 0.994 | 1.007 |
| FADNS (3) | 1.040 | 1.032 | 1.027 | 0.996 | 0.973 | 0.981 | 0.988 | 1.000 | 0.998 | 1.000 | 0.998 | 1.008 | 1.002 | 0.998 | 0.992 | 0.994 | 1.006 |
| FADNS (4) | 1.097 | 1.084 | 1.066 | 1.019 | 0.980 | 0.990 | 0.998 | 1.013 | 1.005 | 1.006 | 1.001 | 1.009 | 1.002 | 0.998 | 0.994 | 0.996 | 1.001 |

Panel B: 6-months ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| FADNS (1) | 0.985 | 0.985 | 0.986 | 0.970 | 0.954 | 0.962 | 0.969 | 0.980 | 0.982 | 0.987 | 0.991 | 1.005 | 1.003 | 1.001 | 0.998 | 1.003 | 1.015 |
| FADNS (2) | 1.038 | 1.022 | 1.016 | 0.990 | 0.969 | 0.973 | 0.978 | 0.988 | 0.987 | 0.991 | 0.992 | 1.005 | 1.002 | 1.000 | 0.997 | 1.002 | 1.014 |
| FADNS (3) | 1.068 | 1.042 | 1.030 | 0.998 | 0.974 | 0.978 | 0.982 | 0.991 | 0.989 | 0.991 | 0.992 | 1.004 | 1.001 | 1.000 | 0.998 | 1.003 | 1.014 |
| FADNS (4) | 1.197 | 1.136 | 1.102 | 1.044 | 1.002 | 1.001 | 1.002 | 1.009 | 1.001 | 1.000 | 0.996 | 1.005 | 1.001 | 1.000 | 0.999 | 1.002 | 1.010 |

Panel C: 12-months ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| FADNS (1) | 0.967 | 0.968 | 0.971 | 0.955 | 0.940 | 0.950 | 0.960 | 0.975 | 0.982 | 0.993 | 1.004 | 1.025 | 1.026 | 1.025 | 1.025 | 1.032 | 1.048 |
| FADNS (2) | 1.097 | 1.071 | 1.056 | 1.021 | 0.994 | 0.994 | 0.996 | 1.004 | 1.003 | 1.007 | 1.010 | 1.025 | 1.024 | 1.023 | 1.022 | 1.029 | 1.045 |
| FADNS (3) | 1.133 | 1.100 | 1.081 | 1.041 | 1.010 | 1.008 | 1.009 | 1.015 | 1.012 | 1.014 | 1.014 | 1.027 | 1.026 | 1.025 | 1.025 | 1.032 | 1.045 |
| FADNS (4) | 1.263 | 1.203 | 1.163 | 1.096 | 1.048 | 1.039 | 1.034 | 1.037 | 1.027 | 1.024 | 1.019 | 1.029 | 1.027 | 1.026 | 1.027 | 1.033 | 1.044 |

Table 17: Root mean squared prediction errors (RMSPE) ratios for the period 1995:01-1999:12. The table shows the relative RMSPE of the considered extensions of the DNS against the standard DNS at a forecast horizon of one, six and twelve months. The results per forecast horizon are in Panel A, B and C, respectively. The shades of the cells indicate the rank of the model per maturity. The darker the shade, the better the model. Values in bold denote significance at the 5% level or less.

Panel A: 1-month ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| FADNS (1) | 1.023 | 1.052 | 1.057 | 1.060 | 1.055 | 1.055 | 1.053 | 1.054 | 1.049 | 1.045 | 1.039 | 1.037 | 1.032 | 1.025 | 1.021 | 1.021 | 1.028 |
| FADNS (2) | 1.001 | 1.015 | 1.036 | 1.021 | 1.017 | 1.022 | 1.028 | 1.033 | 1.029 | 1.026 | 1.023 | 1.023 | 1.019 | 1.016 | 1.014 | 1.014 | 1.021 |
| FADNS (3) | 1.008 | 1.012 | 1.035 | 1.015 | 1.010 | 1.017 | 1.024 | 1.031 | 1.027 | 1.023 | 1.020 | 1.021 | 1.016 | 1.014 | 1.012 | 1.013 | 1.020 |
| FADNS (4) | 1.033 | 1.004 | 1.031 | 0.992 | 0.986 | 0.999 | 1.013 | 1.022 | 1.020 | 1.018 | 1.018 | 1.020 | 1.017 | 1.015 | 1.015 | 1.016 | 1.023 |

Panel B: 6-months ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| FADNS (1) | 1.056 | 1.066 | 1.078 | 1.085 | 1.087 | 1.093 | 1.097 | 1.100 | 1.106 | 1.109 | 1.110 | 1.108 | 1.109 | 1.109 | 1.107 | 1.108 | 1.113 |
| FADNS (2) | 1.023 | 1.029 | 1.039 | 1.035 | 1.034 | 1.040 | 1.044 | 1.049 | 1.051 | 1.054 | 1.057 | 1.058 | 1.059 | 1.062 | 1.062 | 1.064 | 1.073 |
| FADNS (3) | 1.024 | 1.026 | 1.034 | 1.026 | 1.024 | 1.029 | 1.033 | 1.038 | 1.039 | 1.040 | 1.044 | 1.046 | 1.047 | 1.050 | 1.052 | 1.055 | 1.065 |
| FADNS (4) | 1.031 | 1.018 | 1.019 | 0.999 | 0.993 | 0.998 | 1.003 | 1.008 | 1.009 | 1.011 | 1.020 | 1.028 | 1.034 | 1.042 | 1.047 | 1.054 | 1.064 |

Panel C: 12-months ahead forecasts

| Model | Maturity | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| FADNS (1) | 1.066 | 1.071 | 1.079 | 1.083 | 1.084 | 1.089 | 1.094 | 1.099 | 1.105 | 1.110 | 1.119 | 1.123 | 1.126 | 1.130 | 1.133 | 1.137 | 1.145 |
| FADNS (2) | 1.052 | 1.052 | 1.055 | 1.051 | 1.049 | 1.051 | 1.053 | 1.056 | 1.058 | 1.060 | 1.066 | 1.072 | 1.075 | 1.080 | 1.083 | 1.088 | 1.098 |
| FADNS (3) | 1.056 | 1.052 | 1.052 | 1.045 | 1.041 | 1.041 | 1.042 | 1.045 | 1.045 | 1.047 | 1.053 | 1.059 | 1.063 | 1.069 | 1.073 | 1.079 | 1.091 |
| FADNS (4) | 1.078 | 1.058 | 1.046 | 1.027 | 1.017 | 1.013 | 1.011 | 1.012 | 1.009 | 1.010 | 1.020 | 1.034 | 1.044 | 1.057 | 1.067 | 1.078 | 1.092 |

Table 18: Root mean squared prediction errors (RMSPE) ratios for the period 2005:01-2009:12. The table shows the relative RMSPE of the considered extensions of the DNS against the standard DNS at a forecast horizon of one, six and twelve months. The results per forecast horizon are in Panel A, B and C, respectively. The shades of the cells indicate the rank of the model per maturity. The darker the shade, the better the model. Values in bold denote significance at the 5% level or less.

# E Robustness Check: Rank-Based Combination Scheme

| Panel A: 1-month ahead forecasts | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model | Maturity | | | | | | | | | | | | | | | | |
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | **0.917** | 0.985 | 1.013 | 1.010 | 1.001 | 1.003 | 1.006 | 1.012 | 1.012 | 1.013 | 1.007 | 1.010 | 1.004 | 1.003 | 0.997 | 0.996 | 0.995 |
| EW-UKF | **0.914** | 0.986 | 1.015 | 1.011 | 1.001 | 1.003 | 1.006 | 1.013 | 1.014 | 1.016 | 1.011 | 1.015 | 1.008 | 1.004 | 0.996 | 0.994 | 0.993 |
| RMSPE-AVG-EKF-RANK | **0.913** | 0.983 | 1.011 | 1.009 | 1.000 | 1.002 | 1.005 | 1.012 | 1.012 | 1.013 | 1.006 | 1.009 | 1.001 | 1.002 | 0.995 | 0.995 | 0.993 |
| RMSPE-AVG-UKF-RANK | **0.914** | 0.984 | 1.012 | 1.008 | 0.999 | 1.001 | 1.005 | 1.012 | 1.014 | 1.016 | 1.011 | 1.015 | 1.009 | 1.004 | 0.997 | 0.996 | 0.992 |
| RMSPE-MAT-EKF-RANK | **0.898** | 0.974 | 1.007 | 1.007 | 0.998 | 1.001 | 1.004 | 1.011 | 1.010 | 1.010 | 1.003 | 1.006 | 1.000 | 1.003 | 0.996 | 0.995 | 0.992 |
| RMSPE-MAT-UKF-RANK | **0.892** | 0.974 | 1.008 | 1.007 | 0.998 | 1.001 | 1.004 | 1.012 | 1.013 | 1.014 | 1.007 | 1.012 | 1.009 | 1.005 | 0.999 | 0.996 | 0.997 |
| SFADNS-EKF | **0.846** | 0.976 | 1.029 | 1.024 | 0.998 | 1.007 | 1.018 | 1.034 | 1.032 | 1.035 | 1.019 | 1.034 | 1.018 | 1.027 | 1.014 | 1.013 | 1.016 |
| SFADNS-UKF | **0.848** | 0.978 | 1.028 | 1.022 | 0.998 | 1.006 | 1.016 | 1.032 | 1.031 | 1.034 | 1.015 | 1.026 | 1.007 | 1.011 | 0.998 | 1.001 | 1.003 |

| Panel B: 6-months ahead forecasts | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model | Maturity | | | | | | | | | | | | | | | | |
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | 1.018 | 1.022 | 1.024 | 1.023 | 1.020 | 1.022 | 1.024 | 1.028 | 1.030 | 1.032 | 1.032 | 1.035 | 1.033 | 1.031 | 1.027 | 1.026 | 1.032 |
| EW-UKF | 1.020 | 1.024 | 1.026 | 1.025 | 1.022 | 1.024 | 1.026 | 1.031 | 1.032 | 1.035 | 1.035 | 1.039 | 1.037 | 1.035 | 1.031 | 1.030 | 1.036 |
| RMSPE-AVG-EKF-RANK | 1.018 | 1.021 | 1.023 | 1.021 | 1.018 | 1.019 | 1.022 | 1.026 | 1.027 | 1.029 | 1.029 | 1.032 | 1.030 | 1.028 | 1.024 | 1.023 | 1.029 |
| RMSPE-AVG-UKF-RANK | 1.019 | 1.022 | 1.023 | 1.021 | 1.018 | 1.020 | 1.022 | 1.026 | 1.028 | 1.031 | 1.031 | 1.035 | 1.032 | 1.031 | 1.026 | 1.026 | 1.032 |
| RMSPE-MAT-EKF-RANK | 1.015 | 1.021 | 1.022 | 1.021 | 1.017 | 1.019 | 1.021 | 1.025 | 1.027 | 1.029 | 1.029 | 1.032 | 1.029 | 1.028 | 1.024 | 1.023 | 1.027 |
| RMSPE-MAT-UKF-RANK | 1.017 | 1.021 | 1.023 | 1.022 | 1.018 | 1.020 | 1.023 | 1.027 | 1.028 | 1.030 | 1.030 | 1.034 | 1.032 | 1.031 | 1.026 | 1.025 | 1.031 |
| SFADNS-EKF | 1.018 | 1.037 | 1.051 | 1.055 | 1.050 | 1.057 | 1.066 | 1.077 | 1.079 | 1.083 | 1.078 | 1.083 | 1.074 | 1.073 | 1.063 | 1.062 | 1.074 |
| SFADNS-UKF | 1.013 | 1.029 | 1.039 | 1.040 | 1.034 | 1.039 | 1.046 | 1.055 | 1.056 | 1.060 | 1.056 | 1.062 | 1.056 | 1.057 | 1.051 | 1.052 | 1.064 |

| Panel C: 12-months ahead forecasts | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model | Maturity | | | | | | | | | | | | | | | | |
| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | 1.038 | 1.039 | 1.041 | 1.041 | 1.039 | 1.041 | 1.044 | 1.049 | 1.052 | 1.056 | 1.060 | 1.066 | 1.066 | 1.066 | 1.062 | 1.063 | 1.069 |
| EW-UKF | 1.040 | 1.041 | 1.044 | 1.043 | 1.041 | 1.044 | 1.047 | 1.052 | 1.055 | 1.060 | 1.064 | 1.071 | 1.071 | 1.071 | 1.067 | 1.068 | 1.074 |
| RMSPE-AVG-EKF-RANK | 1.038 | 1.038 | 1.040 | 1.039 | 1.036 | 1.038 | 1.041 | 1.045 | 1.048 | 1.052 | 1.055 | 1.061 | 1.061 | 1.061 | 1.058 | 1.059 | 1.065 |
| RMSPE-AVG-UKF-RANK | 1.039 | 1.039 | 1.041 | 1.040 | 1.037 | 1.039 | 1.042 | 1.047 | 1.050 | 1.054 | 1.058 | 1.064 | 1.064 | 1.065 | 1.061 | 1.062 | 1.069 |
| RMSPE-MAT-EKF-RANK | 1.036 | 1.036 | 1.038 | 1.038 | 1.036 | 1.038 | 1.041 | 1.045 | 1.048 | 1.052 | 1.055 | 1.061 | 1.060 | 1.060 | 1.057 | 1.058 | 1.064 |
| RMSPE-MAT-UKF-RANK | 1.037 | 1.037 | 1.039 | 1.038 | 1.036 | 1.039 | 1.042 | 1.047 | 1.049 | 1.054 | 1.057 | 1.064 | 1.063 | 1.064 | 1.061 | 1.061 | 1.068 |
| SFADNS-EKF | 1.066 | 1.077 | 1.089 | 1.092 | 1.091 | 1.099 | 1.108 | 1.119 | 1.124 | 1.132 | 1.135 | 1.142 | 1.137 | 1.136 | 1.128 | 1.127 | 1.135 |
| SFADNS-UKF | 1.061 | 1.068 | 1.075 | 1.075 | 1.072 | 1.077 | 1.083 | 1.091 | 1.095 | 1.100 | 1.102 | 1.109 | 1.106 | 1.106 | 1.101 | 1.102 | 1.110 |

Table 19: Root mean squared prediction errors (RMSPE) ratios. The table shows the relative RMSPE of the combined forecasts using various weighting schemes against the standard DNS at a forecast horizon of one, six and twelve months. The results per forecast horizon are in Panel A, B and C, respectively, with corresponding periods 1987:01-2009:12, 1987:06-2009:12, 1987:12-2009:12. The shades of the cells indicate the rank of the model per maturity. The darker the shade, the better the model. Values in bold denote significance at the 5% level or less.

Panel A: 1-month ahead forecasts

| Model | Maturity 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | **0.938** | 0.974 | 1.001 | 0.982 | 0.968 | 0.978 | 0.989 | 1.004 | 1.006 | 1.010 | 1.006 | 1.016 | 1.009 | 1.002 | 0.998 | 0.999 | 1.008 |
| EW-UKF | **0.931** | 0.971 | 1.001 | 0.978 | 0.963 | 0.975 | 0.986 | 1.002 | 1.006 | 1.010 | 1.006 | 1.018 | 1.010 | 1.001 | 0.994 | 0.995 | 1.007 |
| RMSPE-AVG-EKF-RANK | **0.942** | 0.970 | 0.994 | 0.974 | 0.964 | 0.974 | 0.984 | 1.000 | 1.005 | 1.010 | 1.007 | 1.017 | 1.010 | 1.002 | 0.997 | 0.997 | 1.007 |
| RMSPE-AVG-UKF-RANK | **0.928** | 0.967 | 0.999 | 0.976 | 0.965 | 0.977 | 0.988 | 1.005 | 1.009 | 1.013 | 1.009 | 1.021 | 1.012 | 1.003 | 0.995 | 0.996 | 1.009 |
| RMSPE-MAT-EKF-RANK | **0.927** | 0.976 | 1.001 | 0.979 | 0.964 | 0.975 | 0.985 | 1.001 | 1.005 | 1.010 | 1.006 | 1.017 | 1.011 | 1.006 | 0.998 | 0.999 | 1.010 |
| RMSPE-MAT-UKF-RANK | **0.918** | 0.969 | 1.003 | 0.979 | 0.965 | 0.978 | 0.988 | 1.005 | 1.008 | 1.013 | 1.008 | 1.020 | 1.011 | 1.004 | 0.996 | 0.998 | 1.011 |
| SFADNS-EKF | **0.927** | 0.981 | 1.061 | 1.006 | 0.952 | 0.982 | 1.007 | 1.042 | 1.039 | 1.042 | 1.024 | 1.044 | 1.026 | 1.010 | 0.998 | 1.002 | 1.029 |
| SFADNS-UKF | **0.927** | 0.969 | 1.032 | 0.981 | **0.940** | 0.967 | 0.990 | 1.024 | 1.025 | 1.030 | 1.015 | 1.035 | 1.018 | 1.003 | 0.992 | 0.996 | 1.020 |

Panel B: 6-months ahead forecasts

| Model | Maturity 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | 0.974 | 0.979 | 0.988 | 0.979 | 0.969 | 0.978 | 0.986 | 0.998 | 1.002 | 1.007 | 1.010 | 1.021 | 1.019 | 1.016 | 1.014 | 1.018 | 1.029 |
| EW-UKF | 0.973 | 0.979 | 0.987 | 0.978 | 0.968 | 0.978 | 0.986 | 0.998 | 1.002 | 1.008 | 1.011 | 1.022 | 1.020 | 1.017 | 1.014 | 1.018 | 1.030 |
| RMSPE-AVG-EKF-RANK | 0.969 | 0.975 | 0.984 | 0.977 | 0.969 | 0.978 | 0.986 | 0.997 | 1.001 | 1.006 | 1.009 | 1.020 | 1.018 | 1.015 | 1.013 | 1.016 | 1.027 |
| RMSPE-AVG-UKF-RANK | 0.970 | 0.975 | 0.984 | 0.977 | 0.968 | 0.977 | 0.986 | 0.997 | 1.002 | 1.007 | 1.010 | 1.021 | 1.019 | 1.016 | 1.013 | 1.016 | 1.027 |
| RMSPE-MAT-EKF-RANK | 0.964 | 0.970 | 0.981 | 0.977 | 0.970 | 0.978 | 0.985 | 0.994 | 0.998 | 1.003 | 1.005 | 1.017 | 1.015 | 1.012 | 1.010 | 1.013 | 1.025 |
| RMSPE-MAT-UKF-RANK | 0.961 | 0.969 | 0.980 | 0.974 | 0.968 | 0.977 | 0.986 | 0.997 | 0.998 | 1.003 | 1.006 | 1.017 | 1.015 | 1.012 | 1.010 | 1.014 | 1.028 |
| SFADNS-EKF | 0.993 | 1.012 | 1.047 | 1.024 | 0.997 | 1.019 | 1.036 | 1.059 | 1.063 | 1.067 | 1.060 | 1.074 | 1.064 | 1.055 | 1.049 | 1.054 | 1.073 |
| SFADNS-UKF | 1.005 | 1.002 | 1.020 | 0.998 | 0.975 | 0.992 | 1.007 | 1.028 | 1.033 | 1.040 | 1.038 | 1.054 | 1.047 | 1.042 | 1.038 | 1.045 | 1.063 |

Panel C: 12-months ahead forecasts

| Model | Maturity 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | 0.997 | 1.001 | 1.007 | 0.996 | 0.983 | 0.992 | 1.001 | 1.014 | 1.020 | 1.028 | 1.034 | 1.049 | 1.048 | 1.045 | 1.044 | 1.050 | 1.062 |
| EW-UKF | 0.997 | 1.000 | 1.006 | 0.994 | 0.981 | 0.991 | 1.000 | 1.013 | 1.020 | 1.028 | 1.035 | 1.050 | 1.049 | 1.046 | 1.045 | 1.051 | 1.064 |
| RMSPE-AVG-EKF-RANK | 1.008 | 1.008 | 1.011 | 0.999 | 0.986 | 0.993 | 1.001 | 1.013 | 1.018 | 1.025 | 1.031 | 1.046 | 1.045 | 1.043 | 1.042 | 1.047 | 1.059 |
| RMSPE-AVG-UKF-RANK | 1.008 | 1.008 | 1.010 | 0.998 | 0.984 | 0.992 | 1.000 | 1.012 | 1.018 | 1.025 | 1.031 | 1.046 | 1.046 | 1.043 | 1.042 | 1.047 | 1.060 |
| RMSPE-MAT-EKF-RANK | 1.000 | 1.003 | 1.009 | 0.996 | 0.982 | 0.992 | 0.999 | 1.013 | 1.015 | 1.022 | 1.026 | 1.042 | 1.040 | 1.038 | 1.039 | 1.046 | 1.061 |
| RMSPE-MAT-UKF-RANK | 1.000 | 1.003 | 1.008 | 0.995 | 0.980 | 0.991 | 0.999 | 1.013 | 1.017 | 1.024 | 1.026 | 1.042 | 1.040 | 1.038 | 1.041 | 1.048 | 1.063 |
| SFADNS-EKF | 1.087 | 1.105 | 1.130 | 1.112 | 1.089 | 1.109 | 1.125 | 1.148 | 1.156 | 1.161 | 1.153 | 1.161 | 1.149 | 1.135 | 1.128 | 1.133 | 1.148 |
| SFADNS-UKF | 1.106 | 1.103 | 1.111 | 1.085 | 1.061 | 1.073 | 1.084 | 1.103 | 1.108 | 1.114 | 1.111 | 1.121 | 1.113 | 1.103 | 1.099 | 1.104 | 1.120 |

Table 20: Root mean squared prediction errors (RMSPE) ratios for the period 1995:01-1999:12. The table shows the relative RMSPE of the combined forecasts using various weighting schemes against the standard DNS at a forecast horizon of one, six and twelve months. The results per forecast horizon are in Panel A, B and C, respectively. The shades of the cells indicate the rank of the model per maturity. The darker the shade, the better the model. Values in bold denote significance at the 5% level or less.

Panel A: 1-month ahead forecasts

| Model | Maturity 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | **0.872** | **0.905** | 0.953 | 0.981 | 0.995 | 1.004 | 1.009 | 1.012 | 1.013 | 1.008 | 0.995 | 0.994 | 0.991 | 0.996 | 0.986 | 0.983 | 0.960 |
| EW-UKF | **0.865** | **0.900** | **0.950** | 0.978 | 0.992 | 1.002 | 1.008 | 1.012 | 1.016 | 1.014 | 1.004 | 1.005 | 1.004 | 1.004 | 0.991 | 0.988 | 0.956 |
| RMSPE-AVG-EKF-RANK | **0.855** | **0.885** | **0.939** | 0.971 | 0.992 | 1.003 | 1.008 | 1.011 | 1.011 | 1.004 | 0.985 | 0.984 | 0.980 | 0.990 | 0.977 | 0.979 | 0.951 |
| RMSPE-AVG-UKF-RANK | **0.858** | **0.887** | **0.939** | 0.969 | 0.989 | 1.001 | 1.007 | 1.012 | 1.018 | 1.016 | 1.008 | 1.008 | 1.008 | 1.001 | 0.989 | 0.988 | **0.938** |
| RMSPE-MAT-EKF-RANK | **0.836** | **0.881** | **0.943** | 0.974 | 0.990 | 0.999 | 1.007 | 1.010 | 1.009 | 1.000 | 0.979 | 0.978 | 0.977 | 1.000 | 0.988 | 0.989 | 0.962 |
| RMSPE-MAT-UKF-RANK | **0.826** | **0.874** | **0.941** | 0.973 | 0.987 | 1.000 | 1.007 | 1.011 | 1.018 | 1.013 | 0.994 | 0.999 | 1.009 | 1.009 | 1.007 | 0.994 | 0.984 |
| SFADNS-EKF | **0.721** | **0.793** | **0.899** | 0.957 | 0.967 | 0.986 | 0.997 | 1.001 | 0.995 | 0.996 | 0.972 | 1.013 | 1.007 | 1.080 | 1.057 | 1.057 | 1.024 |
| SFADNS-UKF | **0.721** | **0.796** | **0.895** | 0.950 | 0.968 | 0.983 | 0.992 | 0.994 | 0.989 | 0.980 | 0.950 | 0.966 | 0.966 | 1.007 | 0.992 | 1.012 | 0.976 |

Panel B: 6-months ahead forecasts

| Model | Maturity 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | 0.982 | 0.989 | 0.996 | 1.001 | 1.005 | 1.008 | 1.010 | 1.012 | 1.015 | 1.018 | 1.020 | 1.028 | 1.035 | 1.042 | 1.043 | 1.045 | 1.045 |
| EW-UKF | 0.983 | 0.991 | 0.997 | 1.002 | 1.007 | 1.010 | 1.013 | 1.015 | 1.019 | 1.023 | 1.028 | 1.037 | 1.046 | 1.055 | 1.056 | 1.060 | 1.060 |
| RMSPE-AVG-EKF-RANK | 0.981 | 0.988 | 0.994 | 0.999 | 1.003 | 1.006 | 1.009 | 1.011 | 1.014 | 1.017 | 1.020 | 1.026 | 1.030 | 1.035 | 1.033 | 1.034 | 1.031 |
| RMSPE-AVG-UKF-RANK | 0.981 | 0.989 | 0.995 | 0.999 | 1.003 | 1.007 | 1.010 | 1.012 | 1.016 | 1.019 | 1.023 | 1.031 | 1.036 | 1.041 | 1.040 | 1.040 | 1.037 |
| RMSPE-MAT-EKF-RANK | 0.987 | 0.995 | 0.998 | 1.002 | 1.004 | 1.006 | 1.008 | 1.009 | 1.013 | 1.016 | 1.018 | 1.025 | 1.028 | 1.034 | 1.031 | 1.031 | 1.030 |
| RMSPE-MAT-UKF-RANK | 0.989 | 0.996 | 0.999 | 1.002 | 1.004 | 1.007 | 1.009 | 1.011 | 1.015 | 1.018 | 1.021 | 1.030 | 1.034 | 1.041 | 1.039 | 1.039 | 1.038 |
| SFADNS-EKF | **0.921** | **0.941** | 0.956 | 0.966 | 0.972 | 0.976 | 0.981 | 0.985 | 0.987 | 0.992 | 0.996 | 1.021 | 1.040 | 1.073 | 1.083 | 1.101 | 1.128 |
| SFADNS-UKF | **0.916** | **0.937** | 0.951 | 0.959 | 0.963 | 0.966 | 0.969 | 0.971 | 0.973 | 0.976 | 0.981 | 1.005 | 1.027 | 1.059 | 1.075 | 1.098 | 1.129 |

Panel C: 12-months ahead forecasts

| Model | Maturity 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 30 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DNS | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| EW-EKF | 1.020 | 1.020 | 1.022 | 1.025 | 1.027 | 1.029 | 1.032 | 1.034 | 1.038 | 1.043 | 1.053 | 1.066 | 1.076 | 1.090 | 1.097 | 1.104 | 1.108 |
| EW-UKF | 1.022 | 1.022 | 1.024 | 1.027 | 1.029 | 1.032 | 1.035 | 1.038 | 1.043 | 1.048 | 1.059 | 1.075 | 1.087 | 1.103 | 1.111 | 1.120 | 1.125 |
| RMSPE-AVG-EKF-RANK | 1.019 | 1.020 | 1.022 | 1.024 | 1.026 | 1.028 | 1.030 | 1.032 | 1.035 | 1.039 | 1.047 | 1.058 | 1.066 | 1.076 | 1.080 | 1.084 | 1.085 |
| RMSPE-AVG-UKF-RANK | 1.021 | 1.022 | 1.023 | 1.025 | 1.027 | 1.029 | 1.031 | 1.034 | 1.037 | 1.041 | 1.050 | 1.061 | 1.070 | 1.081 | 1.086 | 1.090 | 1.092 |
| RMSPE-MAT-EKF-RANK | 1.020 | 1.020 | 1.022 | 1.023 | 1.025 | 1.027 | 1.029 | 1.031 | 1.034 | 1.037 | 1.044 | 1.055 | 1.062 | 1.072 | 1.075 | 1.078 | 1.079 |
| RMSPE-MAT-UKF-RANK | 1.021 | 1.022 | 1.023 | 1.024 | 1.026 | 1.028 | 1.030 | 1.032 | 1.036 | 1.040 | 1.048 | 1.060 | 1.068 | 1.079 | 1.082 | 1.086 | 1.088 |
| SFADNS-EKF | 1.023 | 1.029 | 1.034 | 1.037 | 1.040 | 1.043 | 1.047 | 1.051 | 1.055 | 1.061 | 1.076 | 1.101 | 1.123 | 1.154 | 1.173 | 1.192 | 1.211 |
| SFADNS-UKF | 1.022 | 1.027 | 1.030 | 1.031 | 1.031 | 1.031 | 1.033 | 1.034 | 1.035 | 1.039 | 1.051 | 1.074 | 1.097 | 1.128 | 1.150 | 1.175 | 1.196 |

Table 21: Root mean squared prediction errors (RMSPE) ratios for the period 2005:01-2009:12. The table shows the relative RMSPE of the combined forecasts using various weighting schemes against the standard DNS at a forecast horizon of one, six and twelve months. The results per forecast horizon are in Panel A, B and C, respectively. The shades of the cells indicate the rank of the model per maturity. The darker the shade, the better the model. Values in bold denote significance at the 5% level or less.