Insper | Instituto de Ensino e Pesquisa

Inspirar para Transformar

Forecasting the term structure of the Euro Market using Principal Component Analysis

Alexander Dauwe

Marcelo L. Moura

Insper Working Paper

WPE: 233/2011

Insper [Instituto
de Ensino
e Pesquisa

Inspirar para Transformar

Forecasting the term structure of the Euro Market using Principal Component Analysis

Alexander Dauwe [†]

Marcelo L. Moura [‡]

January 27, 2011

Abstract

We forecast the monthly Euro Interest Rate Swap Curve with an autoregressive principal component model. We compare its predictability accuracy against the Diebold and Li's dynamic Nelson Siegel, the auto-regressive direct regression of the yield levels and the random walk model. After a robust set of specifications and regression windows, we conclude that our proposed model achieve forecasts that significantly outperform the competitor models, mainly for short run horizons.

Keywords: Term structure forecasting; Principal component model; Nelson-Siegel model; AR(1), VAR(1); Model selection; Out-of-sample forecasting evaluations

JEL: E43, E47; G17

[†]Insper - Institute of Education and Research; International Commercial Banking Trainee ING GROUP; Hadewijchlaan, 22 - Kortrijk, BELGIUM, 8500; tel.: +32-56-223-178; alexanderdauwe@hotmail.com

[‡]Author for correspondence: Insper - Institute of Education and Research, Associate Professor; Rua Quata, 300 - Sao Paulo/SP, BRAZIL, 04546-042; tel.: 55-11-4504-2435; marcelom@insper.edu.br.

1   Introduction

Understanding the movements and being able to make accurate forecasts of the term structure of interest rates is crucial amongst bond portfolio management, monetary policy and debt policy. Mainly because accurate forecasts allow more effective forms of hedging and because the understanding of the interest rate movements increases the control on the shape of the yield curve. The latter is important for central banks when setting the short rate or for governments when determining the maturity of the debt that will be issued. The importance of yield curve forecasting, together with the fact that the Euro Swap market has become one of the largest and most liquid markets in the world, makes forecasting the Euro Swap curve very exciting.

The main contribution of this work is to evaluate the performance of a Principal Component forecast model when forecasting the Euro Interest Rate Swap curve for different time horizons. Specifically, it will be compared with the results of tough competitor models such as the Nelson Siegel, the Yield regression model, and the naive Random Walk hypothesis. In addition, first difference adaptations of the original models will be postulated and evaluated.

Due to the very high persistence of the yields, the naive 'no change' model, also known as the Random Walk, is very successful at forecasting the term structure. However, evidence exists that the current term structure contains information about future term structures. As an example, Duffee (2002) states that long-maturity bond yields tend to fall over time when the slope of the yield curve is steeper than usual. Next to this internal evidence, the success of, for example, Taylor rules, has demonstrated a strong connection between the yield curve and the observable macro variables. Although the evidence is there and many improvements have been made during the past decade, none of the researchers have succeeded yet in finding a single model that consistently outperforms the Random Walk for all forecast horizons and for all maturities.

In what follows next, we will summarize a selection of the term structure forecast papers that are important references to this work. We have structured them according to their core discussion points, in chronological order.

In 2002 Duffee (Duffee, 2002) reports that the popular 'completely' affine/linear term structure models do not perform well at forecasting. He discovers that this poor result is caused by the models' fundamental assumption that the market price of the risk is a fixed multiple of the variance of the risk. By relaxing this assumption, Duffee creates a new model that seems to outperform the Random Walk. In this new 'essentially' affine model, the market price of risk is no longer a fixed multiple of the variance of risk, but a linear combination of the state vector. Duffee names this model 'essentially' affine, because only the variance of the market price of risk loses its linearity in relation with the state vector.

In 2003 Ang and Piazzesi (2003) derive a no-arbitrage affine/linear term structure model, in which the state vector contains both observable macro factors and unobservable latent factors. Ang and Piazzesi are able to forecast the term structure by assuming that the dynamics of the state space vector are driven by a Gaussian VAR process. This paper is very original because it demonstrates a way to include macro variables directly into the term structure model. Important results of the paper are that the no-arbitrage restriction improves the forecast results and that macro factors explain up to 85% of the movements in the short and middle parts of the yield curve. In a one month ahead forecast exercise, their Macro model seems to beat the Random Walk.

Moench (2008) follows a similar procedure to Ang and Piazzesi. He also uses the com-

bination of no-arbitrate affine/linear model and a VAR of the state vector to forecast the interest rate term structure. Differences lie within the state vector. Moench uses the short term interest rate and the first four principal components of a large panel of macroeconomic time series. He justifies the use of factors by proving that the Fed's monetary policy is better simulated with a Taylor rule based on macro factors, instead of macro variables. In relation to forecasting the term structure, Moench reports that relative to the Random Walk, his model reduces RMSE up to 50% at the short end of the term structure and still up to 20% at the long end.

In 2006 Diebold and Li (2006) break from the traditional affine forecast models and opt for a dynamic variation of the Nelson-Siegel exponential model. To avoid over-fitting, they estimate the dynamics of the Nelson-Siegel factors with an autoregressive model. Their results are promising and they justify the break from the traditional affine models by promoting the simplicity and usability of their model. In particular, at long forecast horizons, the Diebold and Li model appears to be more accurate than benchmark models. In a follow-up paper Diebold et al. (2006) extend the original Diebold and Li model by including real activity, inflation and the monetary policy instrument. This time, they use a VAR to model the dynamics. With this set up, the authors are able to study the dynamic connection between macro variables and the term structure. They find a stronger effect form macro variables on future movements in the yield curve and than for the reverse.

Although Diebold and Li (2006) develop a one-step state space estimation approach, they did not perform any out-of-sample forecast with this technique. This was only explored by Yu and Zivot (2011) who estimate the term structure for Treasury and corporate bonds for nine different ratings. Their findings indicate the one-step estimation using the Kalman Filter as the model for for high-yield bonds at short-term horizons forecast. Possibly, this result is due to the instability of parameter estimation in speculative ratings. Another interesting incursion in more volatile markets can be found in Vicente and Tabak (2008) who compares affine term structure models with the Diebold and Li model for the Brazilian Economy. Their results suggests that the parsimonious dynamic Nelson Siegel model has a superior performance for rates of up to three months and at 12 month-ahead forecasts.

In 2008, Christensen et al. (2009) wrote a paper in which they develop an arbitrage free version of the Nelson Siegel model. Next to getting better results than for the original Diebold an Li set-up, the algorithm of this new no-arbitrage model, is much faster than other affine no-arbitrage approaches.

Although many researchers report outperforming the Random Walk, none of them prove the consistency of their model. In 2007 de Pooter et al. (2007) set up a showdown between the Random walk, the original Nelson Siegel model (with and without macro factors) and the no-arbitrage models (with and without macro factors). They use U.S. Treasury zero-coupon bond data. Interestingly, the authors are not able to select the best model. Mainly because the forecast accuracy of each model is inconsistent over time. Their answer to making a better model, lies in determining logical combinations of all of the models. More specifically, they combine forecast models with a weighting scheme that is based on relative historical performance. Results are now consistent and highly accurate, especially for longer maturities. In relation with the macro factors, the authors notice a positive effect on out-of-sample forecasting.

Another attempt to use model selection is made by Blaskowitz and Herwatz (2009). The importance of their paper to the current lies in the fact that the authors also implement the combination of a principal component yield curve model and an autoregressive model

(AR). In addition, they too study the Euribor Swap Term Structure (Daily rates). Their adaptive technique consists of first creating a pool of models, by changing the time window, the number of principal components and the lags in the AR, next evaluating past performance of all of the models in the pool and finally selecting the best model to make the future forecast. They conclude that the adaptive approach offers additional forecast accuracy in terms of directional accuracy and big hit ability over the Random Walk and the Diebold and Li approach. However, the Root Mean Squared forecast Errors (RMSE) are not compared.

Next to disagreement in model selection, authors also do not yet agree on how to extract the macro factors. In 2008 Exterkate (2008) addresses this area. He evaluates the effect of different macro extraction techniques on the forecast performance of a Nelson Siegel model combined with a factor augmented VAR. Exterkate studies the effect of grouping macro variables before factor extraction and the effect of a technique called thresholding. The latter selects the macro variables with the highest forecasting potential. Exterkate reports a positive effect from both techniques. Next to this small victory, Exterkate has to report that including macro factors did not improve his forecast results compared to the original Diebold and Li setup. Additionally, he is not able to reproduce results that were previously achieved.

In 2008, Duffee (2008) investigates whether imposing no-arbitrage helps when using the term structure to forecast future bond yields. He does this by testing a no-arbitrage and an unrestricted three factor discrete-time Gaussian model, in practice and with a Monte Carlo simulation. In addition, Duffee also investigates the Diebold and Li model that imposes specific analytical functions onto the factors. After testing, Duffee concludes that both in practice and in simulation, imposing no-arbitrage does not improve forecasting performance. Imposing different restrictions, like Diebold and Li, does have a negative result on the forecasts. Duffee explains that the irrelevance of the no-arbitrage restrictions comes directly from the fact that in any n-factor affine model, yields are linear functions of a constant and n other yields. Deviations from this linear equation are so small that its parameters can be estimated with minimal uncertainty even without imposing no arbitrage cross-equation restrictions.

Important lessons to be learned from the literature review: First, Duffee (2008) proves that, for a three factor discrete-time Gaussian model, the imposition of no-arbitrage onto the factors does not improve the forecasts. Second, de Pooter et al. (2007) shows that including the factors of macro variables into the factor transition equation has a positive result on the forecasts for all of his models.

This work will exploit the first of these results. We will use an unrestricted factor model, based on the principal components of the interest rate data, together with an autoregressive transition equation. In the next section, we will specify the autoregressive principal component model of the term structure in detail.

Another lesson is that, surprisingly, while there is an extensive well known literature of forecasting term structure U.S. bonds, the literature is relatively scarce for Euro markets. Therefore our paper also contributes to fill this gap.

The paper is structured as follows. The next section contains the methodology that has been used to make our yield forecasting exercise. Section three is devoted to the data that has been used. In the fourth section we will present and discuss the results. We conclude in section six.

2 Methodology

Tests have shown that naive (vector) autoregressive models of the yields do not make use of the internal structure of the data when forecasting and that they produce bad forecast accuracy compared to the Random Walk (de Pooter et al., 2007). On the other hand, some models impose restrictions that do give better forecast results. Most of these successful models, impose restrictions on the parameters of a general linear yield curve model:

$$
\begin{aligned}
Y_t &= A + BX_t + \epsilon_t & (1) \\
X_t &= \Gamma_1 X_{t-1} + \ldots + \Gamma_l X_{t-l} + \eta_t \qquad \eta_t \sim \aleph(0,1) & (2)
\end{aligned}
$$

The first equation models the yield curve, the second its dynamics. In this equation, $Y_t = [y_{1t}, \ldots, y_{mt}]'$ is the yield vector at time t, m denotes the number of maturities, $X_t = [x_{1t}, \ldots, x_{nFt}]'$ is the state vector and l denotes the number of lags that are included in the transition equation. Concerning the errors, in this work, ϵ_t will be called the approximation error, while η_t will be called the regression error. This general model assumes that each yield is a linear combination of nF factors from the state vector X_t. All yield curve models that we will discuss in this section are created by imposing a different constraint on A, B, Γ and the number of factors nF in X_t.

The main focus of this section is to explain how to specify equations (1) and (2) in case of the principal component model. Next, we will discuss the implementation of the model and finally, we will present a summary of all the models that are used in this work.

2.1 Principal Component Model of the Term Structure

Principal component analysis (PCA) is defined as a linear transformation of a number of correlated variables into a smaller number of uncorrelated variables called principal components. Basically, making a principal component analysis comes down to computing the eigenvalues/eigenvectors of the covariance/correlation matrix of the variables.

The idea behind principal component analysis is to determine the linear combination of variables that has the highest variance. This linear combination of variables forms a new variable that is called 'component' or 'factor' ($x_{it} = \alpha_i' Y_t$) and the coefficients of the linear combination are called loadings (α_i).

The whole principal component algorithm can be summarized by making an eigenvector analysis on the data covariance matrix:

$$
\begin{aligned}
\Sigma &= var(Y) = B_{PC}' \Lambda B_{PC} \\
X_t &= \begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_m \end{pmatrix}' Y_t = B_{PC}' Y_t & (3)
\end{aligned}
$$

With B_{PC} being the eigenvector matrix of Σ and $X_t = [x_{1t}, x_{2t}, \ldots, x_{mt}]'$ being the factor matrix. The yields are now presented as a linear combination of a state vector, remember equation (1) with $A = 0$ and $B = B_{PC}$.

Remember that the principal component factors are independent and are extracted from the data in a natural way. If the no-arbitrage condition lies within the data, it does not need to be imposed when forecasting (Duffee, 2008). Also note that X_t still has the same dimension as the yields Y_t. In the next section we will discuss why and how it is possible to reduce the dimension of X_t.

Inverting equation (3) [1] and then splitting up X_t is particularly helpful to understand how and why principal components can be used to reduce the dimensionality of the data. Let's also look at a regression equation of the factors onto the yields:

$$Y_t = B_{1PC}x_{1t} + B_{2PC}x_{2t} + \ldots + B_{mPC}x_{mt} \tag{4}$$

$$Y_t = \Gamma_0 + \Gamma_1 x_{1t} + \Gamma_2 x_{2t} + \ldots + \Gamma_{mt} x_{mt} + \epsilon_t \tag{5}$$

Due to equation (4), regression(5) will explain all of the variance in Y_t. In addition all coefficients will be estimated as $\Gamma_0 = 0$ and $\Gamma_i = B_{iPC}$, for all i different than 0. The potential to reduce the number of variables in this equation lies in the fact that the principal components are independent (orthogonal). This means that omitting a variable does not cause bias on the other coefficients and that each factor contributes a specific independent part to R^2.

If a factor is truly important depends on the required model accuracy. We implement three automatic factor selection methods: the Kaiser criterium, next the Scree plot and finally the mean square error, for more details see (Field, 2009). When implementing these methods, we get values ranging from 2 to 3. For calculations we choose 3 factors because this is common in literature ($nF = 3$). Following factor analysis standard procedures, we estimate factors with zero mean and unit variance.

It is also possible to apply linear transformations onto the loadings to improve their economical interpretation. Imagine that the loadings span a subspace in the data space, then any set of vectors that can also span that subspace are equivalent to the loadings, without loss of accuracy. In general the idea is to look for vectors that have more economical meaning than the current loadings. For this work we have tested the varimax rotation on the three principal factors of the term structure. The results are not shown because the rotation of the loadings did not enhance their economical interpretability.

2.2   Out-of-sample forecasting

Imagine having a time series $X = [x_1, \ldots, x_n]$ on which you want to apply a forecast model to make out-of-sample forecasts. The first step to achieve 'testable' out-of-sample forecasts is to divide X into an in-sample part $X_{in} = [x_1, \ldots, x_m]$ and an out-of-sample part $X_{out} = [x_{m+1}, \ldots, x_n]$.

The transition equation is fully specified when the forecast step h and the number of lags l are known. For an autoregressive equation, we get the following in-sample regression:

$$x_{t+h} = \beta_1 + \beta_2 x_t + \ldots + \beta_{l+2} x_{t-l} + \eta_t \qquad \eta_t \sim \aleph(0,1) \qquad \text{for t} = l+1 : m-h$$

For this regression to be in-sample, both the left hand side and the right hand side of this equation need to be in-sample, hence the domain of t. The next step is to use the coefficients of this equation to make an out-of-sample forecast. There are many ways to do this, but in this work we only use in-sample observations of X_{in} to make the out-of-sample forecast with the following equation:

$$\tilde{x}_{t+h} = \beta_1 + \beta_2 x_t + \ldots + \beta_{l+2} x_{t-l} \qquad \text{for t} = m-h+1 : m$$

Note that the number of out-of-sample forecasts (domain of t) depends on the considered time step. We retain the forecast based on the latest in-sample observations x_{m+h}.

---

[1]Eigenvectors are orthogonal so $B'_{PC}B_{PC} = I_m$

Consecutive out-of-sample forecast are made by moving the history. In this work, we estimate the model with a 'rolling' history and with an 'increasing' history. A 'rolling' history has a fixed length and moves through time e.g. $history_1 = (x_1, \ldots, x_m)$, $history_2 = (x_2, \ldots, x_{m+1})$,... An 'increasing' history has an increasing length in time e.g. $history_1 = (x_1, \ldots, x_m)$, $history_2 = (x_1, \ldots, x_{m+1})$,... For each history, the principal components are calculated, regressed and used for one out-of-sample forecast as described. Note that this process is quite time consuming.

One of our main tasks will be to model the h-step out of sample of the transition equation. Most of the authors use a (vector) auto regressive model to forecast the factors. In our forecasting exercise, we will achieve the h-step out-of-sample forecasts evaluatimg four variations:

Direct Regression (Scheme 1)

$$x_{it+h} \quad = \quad \beta_1 + \beta_2 x_{it} + \ldots + \beta_{l+2} x_{it-l} + \eta_t \qquad \eta_t \sim \aleph(0,1) \qquad \text{for i} = 1 : nF$$

With l the number of lags and nF the number of factors that are included in the model.

Direct Differenced Regression (Scheme 2)

$$\delta_h x_{it+h} \quad = \quad x_{it+h} - x_{it}$$
$$\delta_h x_{it+h} \quad = \quad \beta_1 + \beta_2 \delta_h x_{it} + \ldots + \beta_{l+2} \delta_h x_{it-l} + \eta_t \qquad \eta_t \sim \aleph(0,1) \qquad \text{for i} = 1 : nF$$

With l the number of lags and nF the number of factors that are included in the model.

Iterative Differenced Regression (Scheme 3)

$$\Delta x_{it} \quad = \quad x_{it} - x_{it-1}$$
$$\delta_h x_{it+h} \quad = \quad x_{it+h} - x_{it} = \sum_{j=it+1:it+h} \Delta x_j$$
$$\Delta x_{it+h} \quad = \quad \beta_1 + \beta_2 \Delta x_{it} + \ldots + \beta_{l+2} \Delta x_{it-l} + \eta_t \qquad \eta_t \sim \aleph(0,1) \qquad \text{for i} = 1 : nF$$

With l the number of lags and nF the number of factors that are included in the model.

Indirect Differenced Regression (Scheme 4)

$$\delta_h x_{it+h} \quad = \quad x_{it+x} - x_{it}$$
$$\Delta x_{it} \quad = \quad x_{it} - x_{it-1}$$
$$\delta_h x_{it+h} \quad = \quad \beta_1 + \beta_2 \Delta x_{it} + \ldots + \beta_{l+2} \Delta x_{it-l} + \eta_t \qquad \eta_t \sim \aleph(0,1) \qquad \text{for i} = 1 : nF$$

With l the number of lags and nF the number of factors that are included in the model.

Scheme 3 and 4 have been implemented because in theory, AR models were developed to forecast stationary series. As will be seen in the results section, the yields and some of the factors are only stationary after taking the first difference.

It is now time to quantify the error of the forecast exercise. We discuss the performance of a specific model and then the relative performance between nested models.

Performance We use the Root Mean Square forecast Error (RMSE) to determine the accuracy of the forecast model.

$$RMSE_{out-of-sample} \quad = \quad \sqrt{\frac{\sum_{t=m-h+1:n-h} \left(Y_{t+h} - \tilde{Y}_{t+h} \right)^2}{lhistory}}$$

Note that in literature the RMSE is the most used evaluation technique. Other evaluation technique, such as the directional accuracy and the big hit ability (Blaskowitz and Herwatz, 2009) are not tested in this work.

Relative Performance When more than one model is used, it is quite natural to want to rank them according to their forecast accuracy. For nested models West and Clark (2007) have developed a specific statistic based on the squared errors:

$$f_{t+h} = (Y_{t+h} - \tilde{Y}_{1t+h})^2 - \left((Y_{t+h} - \tilde{Y}_{2t+h})^2 - (\tilde{Y}_{1t+h} - \tilde{Y}_{2t+h})^2\right)$$

with $\tilde{Y}_{1t+h}$ being the estimation of the nested model and $\tilde{Y}_{2t+h}$ being the estimation of the more general model. In order to know if the general model is significantly better than the nested model, just regress f_{t+h} on a constant for the out-of-sample domain and check if a is significantly larger than zero. In this work, we use the P-value of the t-statistics to evaluate the relative performance of all models with the Random Walk. Models with a P-value lower than 10% are assumed to be significantly better than the Random Walk.

2.3   Summary of the models

This section describes the Principal Component Forecast model and the comparison models: the Diebold and Li model, the Yield Regression model and the the Random Walk model. Next follows a brief description these four methods

2.3.1   Principal Component Forecast Model

In this work we use five adaptations of the principal component forecast model called: PC AR, PC REG, DPC AR, DPC VAR and DPC REG. Note that due to the fact that transition scheme 3 yielded the best results, only this scheme will be presented for the differenced models.

PC AR

$$
\begin{aligned}
Y_{t+h} &= B_{t,PC}X_{t+h} + \bar{Y}_t + \epsilon_t \\
x_{it+h} &= \beta_1 + \beta_2 x_{it} + ... + \beta_{l+2}x_{it-l} + \eta_t \qquad \eta_t \sim \aleph(0,1)
\end{aligned}
$$

With B_{PC} being a specific set of eigenvectors of the covariance matrix of the data, $\bar{Y}_t$ being the mean of the in-sample time series of Y_{insamp}, h being the forecast step and l being the number of lags. According to equation (1), the principal component model imposes restrictions $A = \bar{Y}_t$ and $B = B_{PC}$. These restrictions come naturally from the eigenvalue/eigenvector decomposition of the covariance in the data.

PC REG

$$y_{jt+h} = \beta_1 + \sum_{i=1:nF} (\beta_{i2}x_{it} + ... + \beta_{il+2}x_{it-l}) + \eta_t \qquad \eta_t \sim \aleph(0,1) \qquad \text{for j} = 1:m$$

With m being the number of maturities of the yields.

DPC (V)AR

$$Y_{t+h} = Y_t + (Y_{t+h} - Y_t) = Y_t + \sum_{j=t+1:t+h} (Y_j - Y_{j-1}) = Y_t + \sum_{j=t+1:t+h} \Delta Y_j$$

$$\Delta Y_{t+h} = B_{t,PC} X_{t+h} + \overline{\Delta Y}_t + \epsilon_t$$

$$x_{it+h} = \beta_1 + \beta_2 x_{it} + ... + \beta_{l+2} x_{it-l} + \left(\sum_{j=1:nF} (\alpha_{j1} x_{jt} + ... + \alpha_{jk} x_{jt-k}) \right) + \eta_t$$

$$\eta_t \sim \aleph(0,1) \qquad \text{for i} = 1 : nF \qquad \text{and j} \neq i$$

DPC REG

$$Y_{t+h} = Y_t + (Y_{t+h} - Y_t) = Y_t + \sum_{j=t+1:t+h} (Y_j - Y_{j-1}) = Y_t + \sum_{j=t+1:t+h} \Delta Y_j$$

$$\Delta Y_{t+h} = \beta_1 + \sum_{i=1:nF} (\beta_{i2} x_{it} + ... + \beta_{il+2} x_{it-l}) + \eta_t \qquad \eta_t \sim \aleph(0,1)$$

2.3.2 Diebold and Li Model

The 2006 Diebold and Li (2006) forecast model is a three factor dynamic Nelson Siegel model that imposes exponential restrictions on parameter B of equation (2).

NS AR The famous model of Diebold and Li.

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} C_{1t} + \begin{pmatrix} \frac{1-e^{-\tau_1 \lambda_t}}{\tau_1 \lambda_t} \\ \frac{1-e^{-\tau_2 \lambda_t}}{\tau_2 \lambda_t} \\ \vdots \\ \frac{1-e^{-\tau_N \lambda_t}}{\tau_N \lambda_t} \end{pmatrix} C_{2t} + \begin{pmatrix} \frac{1-e^{-\tau_1 \lambda_t}}{\tau_1 \lambda_t} - e^{-\tau_1 \lambda_t} \\ \frac{1-e^{-\tau_2 \lambda_t}}{\tau_2 \lambda_t} - e^{-\tau_2 \lambda_t} \\ \vdots \\ \frac{1-e^{-\tau_N \lambda_t}}{\tau_N \lambda_t} - e^{-\tau_N \lambda_t} \end{pmatrix} C_{3t} + \begin{pmatrix} \epsilon_t(\tau_1) \\ \epsilon_t(\tau_2) \\ \vdots \\ \epsilon_t(\tau_N) \end{pmatrix}$$

$$Y_t = B_{NS} X_t + \epsilon_t$$

$$x_{it+h} = \beta_1 + \beta_2 x_{it} + ... + \beta_{l+2} x_{it-l} + \eta_t \qquad \eta_t \sim \aleph(0,1) \qquad \text{for i} = 1 : 3$$

Some important facts about the Diebold an Li model: First, λ_t is a fourth parameter that determines the speed of decay of the elements in the coefficient B. Although this parameter may vary with time, Diebold and Li give it a fixed value based on the data. In this work λ_t is fixed at the value that gives the highest R^2 for the OLS regression of the Neslon Siegel equations. The R^2 of the OLS have been calculated for $\lambda_t = (0.02 : 0.001 : 0.07)$.Second, all loadings are time independent. This enhances the speeds of out-of-sample forecasting dramatically. Third, the original transition equation is auto-regressive even though the Nelson Siegel factors are not independent and even though some have a unit root. Note that the transition equation has changed into a VAR. Due to the fact that the purpose of this work is not to test transition equations, we will continue to use an autoregressive transition equation.

2.3.3 The Yield Regression

In order to test the real contribution of the PCA/NS structure in the data, we have also included an autoregressive model of the yields and of the 1st difference of the yields.

Y REG   Regression of the yield levels.

$$y_{it+h} \quad = \quad \beta_1 + \beta_2 y_{it} + \dots + \beta_{l+2} y_{it-l} + \eta_t \qquad \eta_t \sim \aleph(0,1) \qquad \text{for i} = 1:m$$

With y being a yield and m the number of maturities.

DY REG   Regression of the 1st difference of the yield levels.

$$Y_{t+h} \quad = \quad Y_t + (Y_{t+h} - Y_t) = Y_t + \sum_{j=t+1:t+h} (Y_j - Y_{j-1}) = Y_t + \sum_{j=t+1:t+h} \Delta Y_j$$

$$\Delta Y_{t+h} \quad = \quad \beta_1 + \beta_2 y_{it} + \dots + \beta_{l+2} y_{it-l} + \eta_t \qquad \eta_t \sim \aleph(0,1) \qquad \text{for i} = 1:m$$

2.3.4   The Random Walk

This model assumes that the best way to forecast future yields is to look at their current value:

$$Y_{t+h} \quad = \quad Y_t + \eta_t \qquad \eta_t \sim \aleph(0,I)$$

With h being the forecast step. Due to the fact that interest rates have very high autocorrelations, the Random Walk is a good benchmark model. Remember that de Pooter et al. (2007) report that separately none of the models they tested consistently outperforms the Random Walk. In the results section, it will become clear that for the Euro Interest Rates Swaps, η_t is not a $\aleph(0,I)$ process.

3   Data

Forecast literature has mainly dealt with the U.S. treasury market because of the U.S. market dominance and because of its relatively large historical database. Nevertheless, in 2003 Remolona and Wooldridge (2003) point out that with a total outstanding amount of 26.3 trillion euro at the end of June 2002, the Euro Interest Rate Swap market has become one of the largest and most liquid financial markets in the world. The total outstanding amount of US-dollar interest rate swap instruments was slightly smaller, with 26.2 trillion euro. In 2009, the Euro Interest Rate Swap market stays the most liquid with an outstanding amount of 115 trillion dollar (Packer and Mesny, 2009). As a consequence the swap yield curve is becoming a benchmark yield curve in Euro financial markets against which even government bonds are now often referenced.

Euro Interest Rate Swap data   The Euro Interest Rate Swap data consists of monthly last price Euribor continuously compounded rates for maturities of 1, 3, 6 and 12 months and of monthly last price continuously compounded Euro Interest Rate Swap rates for maturities of 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30 years. Euro Interest Rate Swap can be seen as a long term extension of the Euribor rates. Moreover, they represent the fixed rate that settles a variable rate when engaging into an Interest Rate Swap. All rates have been retrieved with Bloomberg for the period of January 1999 until January 2009. Figure 1 illustrates the evolution of all Euro Interest Rates Swaps over the full sample period.

When looking at this figure, it is important to notice that the level of the short rate (lowest
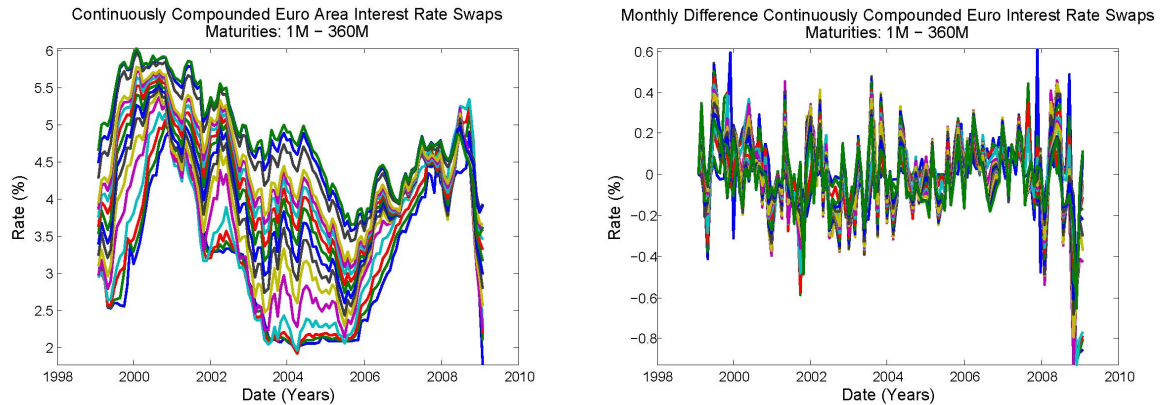
Figure 1: Left: EVOLUTION OF THE EURO SWAP RATES. Right: EVOLUTION OF THE MONTLY DIFFERENCE OF EURO SWAP RATES

line on the left figure) seems to follow an almost seasonal path due to the internet bubble of 2000-2001 and the financial crisis 2008-2009. In addition, yields seem to have a negative trend, especially the ones with a higher maturity. Both of these effects disappear when looking at the monthly differences. In the next section of this work the yield curve and its behavior in relation with forecasting will be discussed in more detail.

4   Results

This section is structured as follows: first, the Euro Interest Rate Swaps will be examined from a forecast perspective. Specifically stationarity and AR order will be examined. Next, the components will be examined from this perspective and finally, the out-of-sample forecast models will be compared against each other for a forecast horizon of 1 and 12 months.

4.1   The Euro Interest Rate Swaps

Due to the fact that autoregressive models have been developed for stationary series, the two main goals of this subsection are: first to find out which of the series are stationary and next to use the Auto Correlation Function (ACF) and the Partial Auto Correlation Function (PACF) to determine the order of the autoregressive model. Let's start with stationarity.

Augmented Dickey Fuller tests are a common tool to evaluate the stationarity of a series. Before applying these tests, the user must specify the ADF model on either having an intercept, a trend, or both. For the case of the yield levels, we assume that yield series do not have a trend, but that they have an intercept. We make this assumption based on two facts. First, comparing the mean with the standard deviations of the levels, leads to the conclusion that the mean is different from zero. Secondly, comparing the mean and standard deviation of the 1st difference, leads to the conclusion that there is no trend. When testing the 1st differences of the levels for stationarity, we assume the 1st differences neither to have a trend, nor to have an intercept. The evolution of the yield levels and their 1st differences, presented in figure 1, confirm these assumptions. The results of the Augmented Dickey Fuller tests can be found in table 1 and lead to the conclusion that all yields are I(1), namely that their 1st

difference is stationary.

Level	$\hat{\mu}$ (%)	$\hat{\sigma}$ (%)	$\hat{\rho}(1)$	PACF	ADF	$\Delta Level$	$\hat{\mu}$ (%)	$\hat{\sigma}$(%)	ACF	PACF	ADF
Euribor 1m	3,197	0,922	0,962	1	-1,741	Δ 1m	-0,011	0,220	1,2	1	**-3,876**
Euribor 3m	3,289	0,973	0,972	1,-2	-1,820	Δ 3m	-0,008	0,204	1,2,3	1	**-4,668**
Euribor 6m	3,340	0,980	0,972	1,-2	-1,804	Δ 6m	-0,006	0,204	1,2,3	1	**-4,805**
Euribor 12m	3,443	0,985	0,969	1,-2	-2,209	Δ 12m	-0,005	0,219	1,2,3	1	**-5,478**
Eurswap 24m	3,604	0,891	0,953	1,-2,5	-1,736	Δ 24m	-0,006	0,237	1,3	1,-4	**-7,512**
Eurswap 36m	3,767	0,828	0,950	1,-2	-1,812	Δ 36m	-0,004	0,232	1	1	**-7,829**
Eurswap 48m	3,913	0,781	0,947	1,-2	-1,836	Δ 48m	-0,004	0,224	1	1	**-8,050**
Eurswap 60m	4,036	0,747	0,948	1,-2	-1,837	Δ 60m	-0,003	0,213	1	1	**-8,117**
Eurswap 72m	4,149	0,724	0,951	1,-2	-1,785	Δ 72m	-0,003	0,202	1	1,3	**-8,195**
Eurswap 84m	4,251	0,710	0,954	1,-2	-1,731	Δ 84m	-0,003	0,191	1	1,3	**-8,176**
Eurswap 96m	4,340	0,698	0,957	1,-2	-1,695	Δ 96m	-0,003	0,182	1	1,3	**-8,162**
Eurswap 108m	4,416	0,686	0,958	1,-2	-1,623	Δ 108m	-0,002	0,176	1	1,3	**-8,346**
Eurswap 120m	4,480	0,676	0,960	1,-2	-1,592	Δ 120m	-0,002	0,170	1	1,3	**-8,353**
Eurswap 180m	4,706	0,661	0,964	1,-2	-1,118	Δ 180m	-0,003	0,158	1	1,3	**-8,716**
Eurswap 240m	4,812	0,657	0,965	1,-2	-0,929	Δ 240m	-0,005	0,150	1	1	**-8,829**
Eurswap 360m	4,844	0,663	0,959	1,-2	-0,592	Δ 360m	-0,009	0,152	1	1	**-8,875**
5% ADF	-2,886					5% ADF	-1,944				

Table 1: EURO SWAP STATISTICS, period 29/01/1999-28/01/2009. The numbers in the ACF and PACF column denote the lags of the significant spikes in the correlogram, their signs denotes the sign of the spikes. The Augmented Dickey Fuller tests (ADF) were made, based on the assumption that yields have no trend and that their 1st difference has no trend and no intercept. The series are assumed to be stationary if the t-statistic is smaller than the 5% t-statistic (bold ADF column).

To determine the order of the autoregressive model, the Auto Correlation function (ACF) and the Partial Auto Correlation function (PACF) provide a lot of information. Technically, the ACF provides information about the correlation between the series and its lagged series, while the PACF provides information about the incremental autocorrelation added by an additional lag. Therefore, it is the highest significant lag in the PACF that provides the order of the Auto Regressive process (AR). Column ACF and PACF in table 1 indicate the lags of the significant spikes and their sign in the Autocorrelation and the Partial Autocorrelation function of the levels of the yields and their 1st difference. Note that in case of the levels, the PACF has a spike at 1 and -2. The negative sign for the spike at lag 2 implies that the level of the yields can be explained by the 1st difference of the yields:

$$Y_t = aY_{t-1} - bY_{t-2} + \epsilon_t = (a-b)Y_{t-1} + b(Y_{t-1} - Y_{t-2}) + \eta_t \qquad (6)$$

In case of the 1st difference model, $(a-b)$ in equation (6) is restricted to be 1. Figure 2 compares the RMSE of a Yield Regression model in level and 1st difference. This figure confirms the hypothesis that the level model mimics the 1st difference model, even though the less restricted level model performs slightly better.

Now that the yields have been investigated, let's investigate their factors.

4.2 Components

Another way of looking at a yield series, is to look at them as a sum of other series, components. In this work components are determined with the Principal Component and the Nelson Siegel method. In what follows next, we will first define the criterium that has been used to determine the number of components, than discuss stationarity and AR order of the components and finally, discuss the cross correlation between the components.
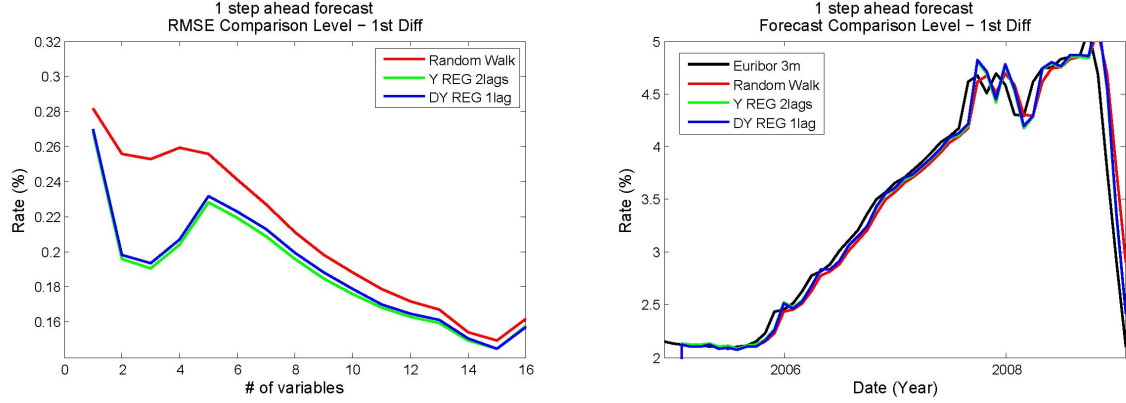
Figure 2: Left: RMSE EURO SWAP CURVE FORECAST. Right: EURIBOR3M forecast. Forecast horizon 1 month, in-sample history 29/01/1999 - 29/12/2004, increasing history.

4.2.1   Number/Importance of components

Even though various criteria exist to automatically determine the number of components that are required to simulate the yield curve, we have used a different criterium when forecasting. The criterium is based on comparing the forecast accuracy of the Random Walk model with
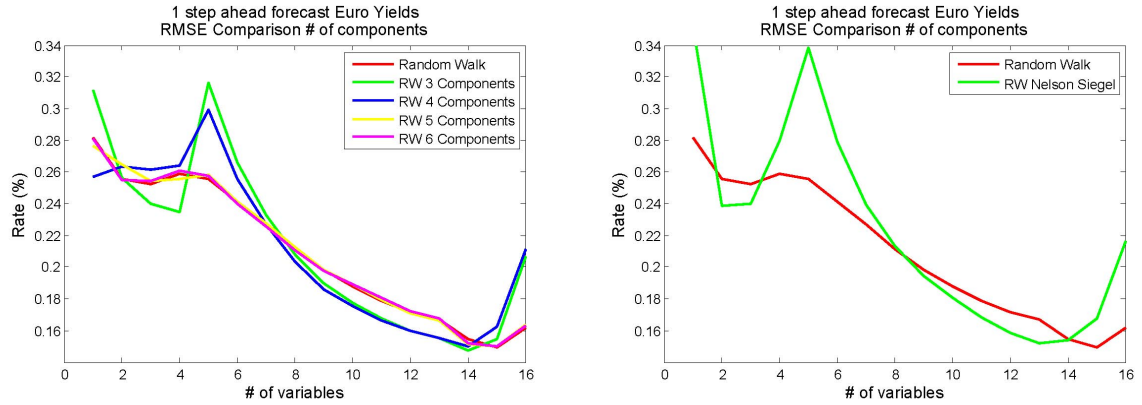
Figure 3: RMSE EURO SWAP CURVE FORECAST. Left: Comparison of Random Walk and Random Walk of components. Right: Comparison of Random Walk and Random Walk of Nelson Siegel components. Forecast horizon 1 month, in-sample history 29/01/1999 - 29/12/2004, increasing history.

the forecast accuracy of a Principal Component model and a Nelson Siegel model that assumes that the factors follow a Random Walk:

$$Y_{t+h} = B_{t,PC}X_{t+h} + \bar{Y}_t + \epsilon_t$$
$$x_{it+h} = x_{it} + \eta_t \qquad \eta_t \sim \aleph(0,1) \qquad \text{for i} = 1 : nF$$

Figure 3 illustrates the effect of the number of components on the RMSE when making an out-of-sample forecast. Apparently, for the case of the Euro Interest Rate Swap curve, the

Random Walk Neslon Siegel model does not lead to a good approximation of the Random Walk of the yields. Moreover, the Random Walk Principal Component model only makes a good approximation of the Random Walk, when considering 6 components. The reason that 6 components have to be used, is because the data set is made up out of Euribor rates and Euro Rate Swaps.

Even though 6 components, may seem like a lot a components, not all components have a similar importance. In order to identify the importance of these components, an idea about the loadings might be helpful. Table 2 presents the loadings of the 1st six components of the Euro Interest Rate Swaps.

LEVEL	f_1	f_2	f_3	f_4	f_5	f_6	df_1	df_2	df_3	df_4	df_5	df_6
1 month	-0,82	-0,36	0,20	0,06	0,05	0,02	-0,11	-0,17	0,05	0,06	0,00	0,00
3 months	-0,87	-0,41	0,13	-0,01	0,00	-0,03	-0,13	-0,15	0,02	-0,03	0,01	-0,03
6 months	-0,89	-0,41	0,05	-0,03	-0,03	0,00	-0,15	-0,13	0,01	-0,03	0,00	0,00
12 months	-0,91	-0,38	-0,05	-0,05	-0,04	0,03	-0,19	-0,10	-0,01	-0,03	-0,01	0,03
24 months	-0,86	-0,15	-0,15	0,06	-0,03	0,00	-0,23	-0,01	-0,06	0,01	-0,02	0,00
36 months	-0,81	-0,03	-0,14	0,03	0,00	0,00	-0,22	0,01	-0,05	0,01	-0,01	0,00
48 months	-0,77	0,06	-0,12	0,01	0,01	-0,01	-0,22	0,03	-0,04	0,01	0,00	-0,01
60 months	-0,73	0,12	-0,09	0,00	0,02	-0,01	-0,21	0,04	-0,02	0,01	0,01	-0,01
72 months	-0,70	0,17	-0,06	-0,01	0,02	0,00	-0,20	0,04	-0,01	0,00	0,01	-0,01
84 months	-0,67	0,22	-0,03	-0,01	0,02	0,00	-0,18	0,05	0,00	0,00	0,01	0,00
96 months	-0,65	0,25	-0,01	-0,02	0,02	0,00	-0,17	0,05	0,02	0,00	0,01	0,00
108 months	-0,63	0,27	0,01	-0,03	0,02	0,00	-0,17	0,05	0,02	0,00	0,01	0,01
120 months	-0,61	0,29	0,03	-0,03	0,01	0,00	-0,16	0,06	0,03	0,00	0,01	0,01
180 months	-0,56	0,35	0,08	-0,03	-0,01	0,00	-0,14	0,06	0,05	-0,01	0,00	0,00
240 months	-0,52	0,38	0,11	0,00	-0,03	0,00	-0,13	0,06	0,06	-0,01	-0,01	0,00
360 months	-0,49	0,42	0,13	0,06	-0,06	0,00	-0,12	0,06	0,06	0,00	-0,03	-0,01

Table 2: LOADINGS EURO SWAP COMPONENTS. Loadings of the principal components of the yields and of the 1st difference of the yields.

Due to the fact that all component series have unit variance, the importance of a component for a specific yield is directly related to the size of its loading. As an example, the 60 months Euro Interest Rate Swap level, is mainly influenced by the 1st component, while the 360 months Euro Interest Rate Swap level is influenced by both the 1st and the 2nd component. On average, it is the eigenvalues that reveals the importance of their related components. For the components of the yield levels, for example, the eigenvalues reveal that the 1st component explains 84.2%, the 2nd 13.7% and the 3rd 1.6% of the variance in the yield levels. So, basically, it is the accuracy with which the 1st and the 2nd component can be forecasted that determines the accuracy of the yield forecast. Component 3, 4, 5 and 6 will improve the accuracy of the forecasts, by reducing the approximation error (ϵ_t), as can be seen in figure 3 and by having cross correlation with the 1st and the 2nd component. This will be discussed later.

In the next section, when discussing stationarity and AR order, we will only discuss 3 components for simplicity. Later, in the forecast exercise, we will include all 6 factors.

4.2.2 Stationarity and AR order

Similar to the yields, an Augmented Dickey Fuller test can be used to check the stationarity of the components and the ACF and PACF functions can be used to determine the order of the Auto Regressive models. Table 3 presents the results of such analysis. Please note that when calculating the ADF statistics, levels are assumed to have an intercept, while 1st differences are assumed not to have an intercept.

Level	$\hat{\mu}$ (%)	$\hat{\sigma}$ (%)	$\hat{\rho}(1)$	PACF	ADF	$\Delta Level$	$\hat{\mu}$ (%)	$\hat{\sigma}$(%)	ACF	PACF	ADF
f_1 (PC)	0,000	1,000	0,962	1,-2	-1,593	Δf_1 (PC)	0,007	0,233	1,3	1	**-6,849**
f_2 (PC)	0,000	1,000	0,962	1,-2	-1,511	Δf_2 (PC)	0,002	0,273	1	1	**-8,152**
f_3 (PC)	0,000	1,000	0,863	1	**-2,995**	Δf_3 (PC)	0,015	0,510	-4	-4	**-10,343**
						f_1 (ΔPC)	0,000	1,000	1	1	**-7,260**
						f_2 (ΔPC)	0,000	1,000	1	1	**-6,570**
						f_3 (ΔPC)	0,000	1,000	-4	-4	**-5,868**
β_1 (NS)	5,085	0,703	0,972	1	-0,802	$\Delta\beta_1$ (NS)	-0,006	0,703	3	3	**-10,451**
β_2 (NS)	-1,850	1,074	0,972	1,-2,-4	-2,518	$\Delta\beta_2$ (NS)	-0,004	1,074	1,3	1,3	**-6,791**
β_3 (NS)	-1,201	1,286	0,870	1	**-2,987**	$\Delta\beta_3$ (NS)	0,016	1,286	-4	-4	**-10,161**
λ (NS)	0,041					λ (NS)	0,041				
5% ADF	-2,886					5% ADF	-1,944				

Table 3: EURO SWAP COMPONENT STATISTICS over the period 29/01/1999-28/01/2009. Note: The numbers in the ACF and PACF column denote the lags of the relevant spikes in the correlogram, their signs denotes the sign of the spikes. The Augmented Dickey Fuller tests (ADF) were made, based on the assumption that levels have no trend and that their 1st difference has no trend and no intercept. The series are assumed to be stationary if the t-statistic is smaller than the 5% t-statistic (bold ADF column).

Note that the Nelson Siegel method of the 1st difference of the yields is equal to the 1st difference of the Nelson Siegel components of the level of the yields. This is not the case for the principal components, hence the extra rows for the principal components in table 3.

Dickey Fuller results in table 3 lead to the conclusion that f_1 and f_2, β_1 and β_2 are not stationary, while f_3, β_3 are stationary. As discussed before, forecast improvements will mainly be due to a more accurate forecast of component 1 and 2. Similar to the yields, those two factors have an AR2 behavior in level and an AR1 behavior in 1st difference as can be seen in table 4.

RMSE (%)	RW	AR1	AR2	AR3	AR4	DAR1	DAR2	DAR3	DAR4
f_1 (PC)	0,254	0,254	0,216	0,217	0,210	0,219	0,220	0,215	**0,213**
f_2 (PC)	0,276	0,289	0,276	0,276	0,286	**0,267**	0,267	0,277	0,276
f_3 (PC)	0,543	0,527	0,530	0,535	0,546	0,552	0,561	0,569	**0,518**
β_1 (NS)	0,152	0,154	0,154	0,155	0,153	0,153	0,153	**0,151**	0,151
β_2 (NS)	0,278	0,283	0,256	0,258	0,258	**0,255**	0,258	0,262	0,262
β_3 (NS)	0,640	0,620	0,623	0,624	0,627	0,650	0,656	0,663	**0,608**

Table 4: RMSE EURO SWAP COMPONENT FORECAST. Forecast horizon 1 month, in-sample history 29/01/1999 - 29/12/2004, increasing history.

In contrast with the yields, component 1 and 2 are forecasted with more accuracy when using the 1st difference of the components and 1 lag, compared to using the levels and 2 lags. For the 3rd component, differencing is not necessary and does not yield a better result.

4.2.3 Cross Correlation

Remembering what Duffee (2002) states: 'the long-maturity bond yields tend to fall over time when the slope of the yield curve is steeper than usual', it is natural to check wether our components have an impact on each other. Table 5 presents the cross correlations of the 1st difference of the Euro components.

It seems, that both the component 1 and 2 have cross correlations with each other and with the other components. Moreover, component 4 and 5 seem to influence component 1. Including them in the forecast model should yield better forecast results.

i	0	1	2	3	4	5	6	7	8	9	10	11	12
$df_1, df_2(-i)$	0,00	0,10	0,05	-0,11	-0,12	-0,08	-0,06	0,07	-0,08	-0,03	0,04	-0,05	-0,02
$df_1, df_3(-i)$	0,00	0,02	0,19	0,12	-0,07	-0,06	-0,10	-0,12	-0,06	-0,04	-0,02	0,04	0,00
$df_1, df_4(-i)$	0,00	-0,13	0,08	0,02	0,00	-0,07	-0,04	0,00	0,05	0,02	-0,03	-0,06	0,00
$df_1, df_5(-i)$	0,00	**0,26**	**0,33**	0,01	0,10	0,11	-0,04	-0,09	-0,02	-0,14	-0,03	0,08	-0,06
$df_1, df_6(-i)$	0,00	0,16	-0,10	0,06	0,11	-0,08	-0,04	0,13	0,00	0,04	0,08	0,12	0,08
$df_2, df_1(-i)$	0,00	**0,24**	**0,36**	**0,29**	**0,21**	0,14	0,06	-0,10	0,01	0,10	0,10	0,15	0,13
$df_2, df_3(-i)$	0,00	-0,02	-0,13	**0,22**	0,15	-0,01	0,03	-0,14	-0,13	0,07	-0,05	0,13	0,06
$df_2, df_4(-i)$	0,00	**0,26**	0,03	-0,09	-0,06	-0,08	-0,01	0,18	0,04	-0,13	-0,10	-0,03	0,00
$df_2, df_5(-i)$	0,00	0,08	0,18	**0,28**	0,22	0,08	0,15	0,00	-0,16	-0,09	-0,15	-0,10	0,03
$df_2, df_6(-i)$	0,00	0,01	-0,02	0,00	0,12	0,01	0,12	-0,03	0,04	0,08	0,09	0,04	0,00

Table 5: Cross correlation of the components of the 1st difference of the EURO SWAP CURVE. Significant cross correlation are bold.

4.3   Forecasting the Euro Interest Rate Swaps

This subsection consists of three parts. First, the additional parameters to fully specify a forecast model will be discussed. Next, a 1-step ahead forecast exercise will be achieved and finally, a 12-step ahead forecast will be evaluated and discussed.
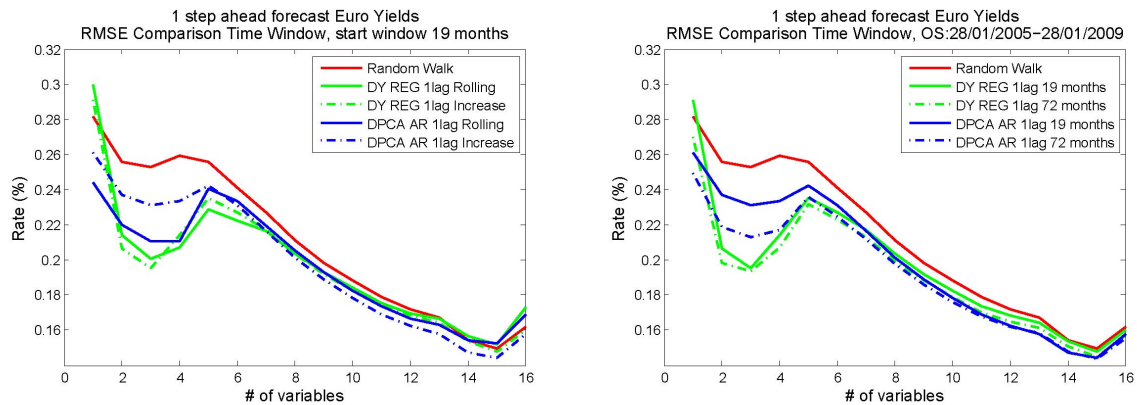
Figure 4: RMSE EURO SWAP CURVE FORECAST. Left: Comparison of Rolling and Increasing Time window. Right: Comparison of short and long history, increasing time window used. Forecast horizon 1 month, 6 components, out-of-sample forecast: 28/01/2005-28/01/2009

4.3.1   Parameters of the forecast models

When forecasting, the importance of the size of the in-sample history is twofold: firstly, the autoregressive models use it to regress their coefficients, secondly, the principal component models use it to determine the principal components.

Figure 4 illustrates the effect of the size of the in-sample history on the accuracy of the forecast. Forecast results tend to be more accurate for large in-sample histories that increase through time. This is probably due to the fact that for short in-sample histories the Auto Regressive coefficients and Principal Component loadings change a lot when the history moves through time. In what follows, forecasts will be based on a large in-sample history that increases when moving through time.

4.4 Forecasting the Euro Interest Rate Swap curve

In this subsection the forecast accuracy of all of the different forecast models will be evaluated for a 1-step ahead and a h-step ahead forecast.

4.4.1 1-step ahead forecasts

Significant Auto Correlations and Cross Correlations above suggest that a model that includes those effects would outperform the Random Walk. Table 6 illustrates the RMSE and the Clark and West p-values of the forecast exercise. As suspected, the multivariate 1st difference models

Root Mean Square out of Sample Error (%)								
Maturity	RW	Y REG	DYREG	PC AR	PC REG	DPC AR	DPC REG	DPC VAR	NS AR
1 month	0,281	0,283	0,270	0,276	**0,208**	0,249	0,239	0,239	0,349
3 months	0,256	0,257	0,199	0,259	**0,188**	0,219	0,197	0,198	0,228
6 months	0,253	0,254	0,193	0,254	**0,177**	0,213	0,196	0,196	0,225
12 months	0,259	0,260	0,207	0,259	**0,187**	0,217	0,204	0,203	0,274
24 months	0,256	0,256	0,232	0,270	0,240	0,236	**0,228**	0,229	0,342
36 months	0,241	0,241	0,223	0,247	0,228	0,225	**0,213**	**0,212**	0,290
48 months	0,227	0,227	0,213	0,230	0,214	0,212	**0,196**	**0,196**	0,251
60 months	0,212	0,211	0,199	0,213	0,200	0,198	**0,179**	0,180	0,223
72 months	0,199	0,198	0,188	0,199	0,188	0,186	**0,165**	**0,165**	0,202
84 months	0,188	0,188	0,179	0,189	0,182	0,176	**0,156**	**0,155**	0,187
96 months	0,179	0,179	0,170	0,179	0,175	0,168	0,146	**0,145**	0,175
108 months	0,172	0,172	0,165	0,171	0,171	0,162	0,141	**0,140**	0,165
120 months	0,167	0,168	0,161	0,166	0,167	0,158	**0,136**	**0,136**	0,159
180 months	0,155	0,156	0,151	0,151	0,162	0,148	**0,129**	**0,129**	0,158
240 months	0,150	0,151	0,145	0,151	0,159	0,144	0,128	**0,127**	0,171
360 months	0,162	0,164	0,157	0,168	0,161	0,155	**0,138**	0,140	0,222

Clark and West P-value								
Maturity	RW	Y REG	DYREG	PC AR	PC REG	DPC AR	DPC REG	DPC VAR	NS AR
1 month	NaN	0,948	0,254	**0,089**	**0,041**	**0,069**	**0,066**	**0,066**	0,986
3 months	NaN	0,880	**0,058**	0,911	**0,049**	**0,042**	**0,048**	**0,047**	**0,073**
6 months	NaN	0,839	**0,046**	0,719	**0,043**	**0,038**	**0,054**	**0,054**	0,195
12 months	NaN	0,842	**0,042**	0,497	**0,058**	**0,036**	**0,059**	**0,058**	0,624
24 months	NaN	0,643	**0,026**	0,955	0,202	**0,062**	**0,045**	**0,054**	0,999
36 months	NaN	0,505	**0,049**	0,934	0,291	**0,063**	**0,071**	**0,063**	0,999
48 months	NaN	0,461	**0,074**	0,901	0,288	**0,063**	**0,051**	**0,052**	0,995
60 months	NaN	0,454	**0,088**	0,718	0,302	**0,065**	**0,050**	**0,049**	0,922
72 months	NaN	0,478	0,103	0,507	0,311	**0,075**	**0,048**	**0,046**	0,663
84 months	NaN	0,538	0,113	0,590	0,375	**0,084**	**0,050**	**0,046**	0,455
96 months	NaN	0,575	0,114	0,539	0,415	**0,094**	**0,049**	**0,047**	0,358
108 months	NaN	0,618	0,123	0,317	0,480	0,115	**0,050**	**0,050**	0,271
120 months	NaN	0,655	0,136	0,293	0,483	0,127	**0,053**	**0,054**	0,227
180 months	NaN	0,780	0,185	0,219	0,659	0,182	**0,057**	**0,055**	0,636
240 months	NaN	0,896	0,109	0,650	0,692	0,194	**0,084**	**0,083**	0,990
360 months	NaN	0,958	0,152	0,884	0,483	0,192	0,123	0,135	0,962

Table 6: RMSE AND C&W P-VALUE EURO SWAP CURVE. Forecast horizon 1 month, 6 components, out-of-sample forecast: 28/01/2005-28/01/2009, first history: 29/01/1999-29/12/2004, increasing history, $\lambda = 0.041$ for NS model, 1 lag. Bold West and Clark (2007) p-values denote a significant forecast improvement relative to the Random Walk at 10% level.

provide more accurate forecasts of the yields. The Clark and West p-values indicates that the DPC VAR and the DPC REG model outperform the Random Walk significantly for all maturities. Moreover, on average, the DPC REG and DPC VAR model have a 16,7% lower RMSE than the Random Walk. Note that their results are almost equal, which implies that the PCA within the yields is very strong. In addition, the use of the PCA structure provides more accurate forecast results than the simple Auto Regressive models of the yields (Y REG and DY REG).

4.4.2 h-step ahead forecast

The story for h-step ahead forecasts is a little different, because several schemes exist to make an h-step out-of-sample forecast. In this work, we have implemented the following schemes:

direct regression (scheme 1), Direct Differenced Regression (scheme 2), Iterative Differenced Regression (scheme 3) and Indirect Differenced Regression (scheme 4).

The schemes are discussed in detail in section 2. Concerning the 1st and 2nd scheme, we deal with the regression of non-stationary series. For the 3rd and 4th scheme, the series are stationary. Table 7 presents the RMSE for a 12-step ahead forecast with a Yield Regression model.

Maturity	RW (/)	Y REG (S # 1)	DYREG (S # 2)	**DYREG** **(S # 3)**	DYREG (S # 4)
1 month	0,961	0,904	1,256	1,152	1,188
3 months	1,009	1,009	1,324	1,136	1,191
6 months	0,996	1,018	1,344	1,157	1,224
12 months	0,984	1,018	1,383	1,211	1,283
24 months	0,909	**0,825**	1,316	1,176	1,236
36 months	0,815	**0,696**	1,229	1,094	1,146
48 months	0,749	**0,607**	1,163	1,036	1,080
60 months	0,690	**0,542**	1,094	0,975	1,012
72 months	0,639	**0,497**	1,035	0,922	0,956
84 months	0,601	0,470	0,990	0,883	0,912
96 months	0,566	0,453	0,948	0,845	0,872
108 months	0,539	0,441	0,914	0,815	0,839
120 months	0,519	0,435	0,886	0,790	0,813
180 months	0,472	0,441	0,809	0,726	0,743
240 months	0,469	0,470	0,766	0,697	0,709
360 months	0,507	0,528	0,736	0,690	0,694
Average	0,714	0,647	1,075	0,957	0,994

Table 7: RMSE EURO SWAP CURVE. Forecast horizon 12 month, 6 components, out-of-sample forecast: 29/12/2005-28/01/2009, first history: 29/01/1999-29/12/2004, increasing history, 1 lag.

According to this table, level models give more accurate forecast results than the 1st difference models. The third scheme seems to be the most accurate for the methods using differenced yields. A reason for the different performance could lay in the fact that only the levels have Auto Correlation at higher lags (table 1).

Finally, table 8 presents the results of a 12-step out-of-sample forecast for all of the models that have been developed in this work. According with the results of table 7, level models have used the 1st scheme and differenced models have used the 3rd scheme. Level Yield Regression models and Principal Componenents auto-regressive model seem to perform the best. Unfortunately, the more accurate forecast results might be due to the sample history that has been studied. Yield levels are not stationary, so it is not certain that future forecasts of the yields will be as accurate as the forecasts in the sample period (January 1999 - January 2009). For this, we should also consider the possibility of using the differenced principal components models that have the best forecasting performance at short maturities.

Root Mean Square out of Sample Error (%)

Maturity	RW (/)	Y REG (S # 1)	DYREG (S # 3)	PC AR (S # 1)	PC REG (S # 1)	DPC AR (S # 3)	DPC REG (S # 3)	DPC VAR (S # 3)	NS AR (S # 1)
1 month	0,961	0,904	1,152	1,055	1,697	0,910	0,973	**0,828**	2,127
3 months	1,009	1,009	1,136	1,234	1,826	0,960	1,051	**0,901**	2,238
6 months	0,996	1,018	1,157	1,326	1,920	0,963	1,030	**0,951**	2,312
12 months	0,984	1,018	1,211	1,406	1,984	**0,970**	1,015	1,012	2,368
24 months	0,909	**0,825**	1,176	1,178	1,812	0,939	0,935	1,030	2,248
36 months	0,815	**0,696**	1,094	0,946	1,619	0,868	0,836	0,969	2,046
48 months	0,749	**0,607**	1,036	0,745	1,450	0,829	0,770	0,924	1,846
60 months	0,690	**0,542**	0,975	0,594	1,312	0,782	0,708	0,872	1,664
72 months	0,639	**0,497**	0,922	0,478	1,198	0,740	0,655	0,820	1,507
84 months	0,601	0,470	0,883	**0,401**	1,107	0,710	0,613	0,782	1,375
96 months	0,566	0,453	0,845	**0,362**	1,030	0,680	0,577	0,744	1,267
108 months	0,539	0,441	0,815	**0,347**	0,961	0,657	0,549	0,714	1,176
120 months	0,519	0,435	0,790	**0,350**	0,912	0,637	0,528	0,689	1,104
180 months	0,472	0,441	0,726	**0,433**	0,772	0,589	0,481	0,618	0,893
240 months	0,469	0,470	0,697	0,521	0,740	0,569	0,476	0,585	0,808
360 months	0,507	0,528	0,690	0,631	0,766	0,581	0,513	0,588	0,765

Clark and West P-value

Maturity	RW (/)	Y REG (S # 1)	DYREG (S # 3)	PC AR (S # 1)	PC REG (S # 1)	DPC AR (S # 3)	DPC REG (S # 3)	DPC VAR (S # 3)	NS AR (S # 1)
1 month	NaN	0,156	0,965	0,948	0,989	**0,010**	0,996	**0,028**	1,000
3 months	NaN	0,499	0,990	1,000	0,999	**0,006**	0,994	**0,030**	1,000
6 months	NaN	0,643	0,999	1,000	1,000	**0,036**	0,986	**0,193**	1,000
12 months	NaN	0,698	1,000	1,000	1,000	0,250	0,973	0,716	1,000
24 months	NaN	**0,077**	1,000	1,000	1,000	0,860	0,920	0,997	1,000
36 months	NaN	**0,019**	1,000	0,997	1,000	0,962	0,901	1,000	1,000
48 months	NaN	**0,006**	1,000	0,463	1,000	0,992	0,909	1,000	1,000
60 months	NaN	**0,003**	1,000	**0,020**	1,000	0,997	0,911	1,000	1,000
72 months	NaN	**0,003**	1,000	**0,001**	1,000	0,999	0,922	1,000	1,000
84 months	NaN	**0,005**	1,000	**0,001**	1,000	0,999	0,921	1,000	1,000
96 months	NaN	**0,009**	1,000	**0,001**	1,000	1,000	0,936	1,000	1,000
108 months	NaN	**0,016**	1,000	**0,003**	1,000	1,000	0,955	1,000	1,000
120 months	NaN	**0,027**	1,000	**0,008**	1,000	1,000	0,968	1,000	1,000
180 months	NaN	0,172	1,000	0,295	1,000	1,000	0,997	1,000	1,000
240 months	NaN	0,517	1,000	0,761	1,000	1,000	0,998	1,000	1,000
360 months	NaN	0,838	1,000	0,965	1,000	0,995	0,986	0,994	1,000

Table 8: RMSE EURO SWAP CURVE. Forecast horizon 12 month, 6 components, out-of-sample forecast: 29/12/2005-28/01/2009, first history: 29/01/1999-29/12/2004, increasing history, $\lambda = 0.041$ for NS model, 1 lag. Bold West and Clark (2007) p-values denote a significant forecast improvement relative to the Random Walk at 10% level.

5 Conclusion

This work contributes to the literature by in two ways: first, the set-up of the Principal Component forecast model has been discussed in detail. Second, the out-of-sample RMSE of the Principal Component forecast model has been compared against the RMSE of the Nelson-Siegel forecast model, the Random Walk and the Yield Regression forecast models.

The study of the set-up of the Principal Component forecast model has lead to the following conclusions:First, the number of the components can be determined with the use of a Principal Component forecast model that assumes the components to follow a Random Walk. With this technique, 6 components are required to model the Euro Interest Rate Swap curve.

Second,The loadings and the eigenvalues help to determine the importance of each of the components on the forecast accuracy. The two first components have the most impact when forecasting the Euro Interest Rate Swap curve.

Third, ACF and PACF diagrams are helpful to determine the order of the Auto Regressive transition equation. Fourth, including components with Cross Correlation improves the forecast results. For the Euro Interest Rate Swap curve, components 3, 4 and 5 influence the 1st and the 2nd component. Finally, longer in-sample histories that increase through time improve forecast results.

Comparing the different models for a 1-step and a 12-step out-of-sample forecast, leads to the following conclusions:First, for the 1-step out-of-sample forecasts, DPCREG and DPC-VAR models are the most accurate. These models are based on the stationary 1st difference of the yields and make use of their Principal Component structure. Specifically, these methods

use ACF, PACF and Cross Correlation between the components to achieve superior forecast results.

Second, For the 12-step out-of-sample forecasts, the level Yield Regression forecast model and the Principal Component regression leads to the best results. These results are similar to the ones found by Yu and Zivot (2011) where a yields level model outperforms dynamic Nelson Siegel models and other competitors for corporate bonds with A and A- ratings. However, those optimistic results might be due to the sample history that has been studied. Yield levels are not stationary, so it is not certain if future forecasts of the yields will be as accurate as the forecasts in the sample period. Therefore, the second best forecasting performance of the principal components differenced models can be a more robust choice.

A probable extension of this work could be the inclusion of Macro factors into the transition models. More specifically, for the Euro Interest Rate Swap curve, the cross correlation of the Macro factors with the 1st component (Level) and the 2nd component (Slope) could lead to an improvement of the short term forecast results. This could be achieved, by including the Macro variables into the Principal Component structure, or by including them as external factors in the transition equations. Another possible extension would be the investigation of the stability of the level Yield Regression and the level Principal Component models for the long term forecast.

References

ANG, Andrew; PIAZZESI, Monika, A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. Journal of Monetary Economics, 2003. vol. 50, no. 4, pp. 745–787.

BLASKOWITZ, Oliver; HERWATZ, Helmut, Adaptive Forecasting of the EURIBOR Swap Term Structure. Journal of Forecasting, 2009. vol. 28, no. 7, pp. 575–594.

CHRISTENSEN, Jens H. E.; DIEBOLD, Francis X.; RUDEBUSCH, Glenn D., An arbitrage-free generalized Nelson-Siegel term structure model. Econometrics Journal, 2009. vol. 12, no. 3, pp. C33–C64.

DIEBOLD, Francis X.; LI, Canlin, Forecasting the term structure of government bond yields. Journal of Econometrics, 2006. vol. 130, no. 2, pp. 337–364.

DIEBOLD, Francis X.; RUDEBUSCH, Glenn D.; ARUOBA, S. Boragan, The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach. Journal of Econometrics, 2006. vol. 131, no. 2, pp. 309–338.

DUFFEE, Gregory R., Term Premia and Interest Rate Forecasts in Affine Models. Journal of Finance, 2002. vol. 57, no. 1, pp. 405–443.

DUFFEE, Gregory R., Forecasting with the term structure: The role of no-arbitrage. Working paper, Johns Hopkins University, 2008, URL `http://www.econ.jhu.edu/people/Duffee/duffeeNoArbJuly2009.pdf`.

EXTERKATE, Peter, Forecasting the yield curve in a data-rich environment, Working paper. Working paper, Cemmap, 2008, URL `http://www.cemmap.ac.uk/forms/ufm08/exterkate.pdf`.

FIELD, Andy, Discovering Statistics Using SPSS (Introducing Statistical Methods series). Sage Publications Ltd, third edition ed., 2009.

MOENCH, Emanuel, Forecasting the yield curve in a data-rich environment - a no-arbitrage factor-augmented VAR approach. Journal of Econometrics, 2008. vol. 146, no. 1, pp. 26–43.

PACKER, Frank; MESNY, Philippe, International banking and financial market developments. Bis quarterly review, Bank for international settlements, 2009.

DE POOTER, Michiel D.; RAVAZZOLO, Francesco; VAN DIJK, Dick, Predicting the Term Structure of Interest Rates: Incorporating Parameter Uncertainty, Model Uncertainty and Macroeconomic Information. Tinbergen institute discussion papers, Tinbergen Institute, 2007.

REMOLONA, Eli M.; WOOLDRIDGE, Philip D., The Euro Interest Rate Swap market. 2003.

VICENTE, Jose; TABAK, Benjamim M., Forecasting bond yields in the Brazilian fixed income market. International Journal of Forecasting, 2008. vol. 24, no. 3, pp. 490–497.

WEST, Kenneth D.; CLARK, Todd, Approximately Normal Tests for Equal Predictive Accuracy in Nested Models. Journal of Econometrics, 2007. vol. 138, no. 1, pp. 291–311.

Yu, Wei-Choun; Zivot, Eric, Forecasting the term structures of Treasury and corporate yields using dynamic Nelson-Siegel models. International Journal of Forecasting, 2011. vol. 27, no. 2, pp. 579–591.