ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Koopman, Siem Jan; Mallee, Max I.P.; van der Wel, Michel

Working Paper

Analyzing the Term Structure of Interest Rates using the Dynamic Nelson-Siegel Model with Time-Varying Parameters

Tinbergen Institute Discussion Paper, No. 07-095/4

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Koopman, Siem Jan; Mallee, Max I.P.; van der Wel, Michel (2007) : Analyzing the Term Structure of Interest Rates using the Dynamic Nelson-Siegel Model with Time-Varying Parameters, Tinbergen Institute Discussion Paper, No. 07-095/4, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
https://hdl.handle.net/10419/86192

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

TI 2007-095/4
Tinbergen Institute Discussion Paper

Analyzing the Term Structure of Interest Rates using the Dynamic Nelson-Siegel Model with Time-Varying Parameters

*Siem Jan Koopman**
Max I.P. Mallee
*Michel van der Wel**

VU University Amsterdam.

** Tinbergen Institute.*

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

Analyzing the Term Structure of Interest Rates using the Dynamic Nelson-Siegel Model with Time-Varying Parameters*

Siem Jan Koopman[(a,c)] *Max I.P. Mallee*[(a)] *Michel van der Wel*[(b,c)]

[(a)] Department of Econometrics, VU University Amsterdam

[(b)] Department of Finance, VU University Amsterdam

[(c)] Tinbergen Institute

December 2007

Some keywords:

Yield Curve; Time-varying Volatility; Spline Functions; Kalman Filter; Missing Values.

JEL classification:

C32, C51, E43.

*We are grateful to Francis X. Diebold, Dick van Dijk and Michiel de Pooter for their comments on an earlier version of this paper. We further would like to thank for the comments of participants in seminar presentations at the Erasmus University Rotterdam, the University of Pennsylvania and the NAKE 2007 day in Utrecht. All errors are our own.

Analyzing the Term Structure of Interest Rates using the Dynamic Nelson-Siegel Model with Time-Varying Parameters

Abstract

In this paper we introduce time-varying parameters in the dynamic Nelson-Siegel yield curve model for the simultaneous analysis and forecasting of interest rates of different maturities, known as the term structure. The Nelson-Siegel model has been recently reformulated as a dynamic factor model where the latent factors are interpreted as the level, slope and curvature of the term structure. The factors are modelled by a vector autoregressive process. We propose to extend this framework in two directions. First, the factor loadings are made time-varying through a simple single step function and we show that the model fit increases significantly as a result. The step function can be replaced by a spline function to allow for more smoothness and flexibility. Second, we investigate empirically whether the volatility in interest rates across different time periods is constant. For this purpose, we introduce a common volatility component that is specified as a spline function of time and scaled appropriately for each series. Based on a data-set that is analysed by others, we present empirical evidence where time-varying loadings and volatilities in the dynamic Nelson-Siegel framework lead to significant increases in model fit. Improvements in the forecasting of the term structure are also reported. Finally, we provide an illustration where the model is applied to an unbalanced dataset. It shows that missing data entries can be estimated accurately.

1 Introduction

Fitting and predicting time-series of a cross-section of yields has proven to be a challenging task. As with many topics in empirical economic analysis there is the trade-off between the goodness of fit that is obtained by employing statistical models without a reference to economic theory, and the lack of fit by economic models that do provide a basis for the underlying economic theory.

For many decades work on the term structure of interest rates has mainly been theoretical in nature. In the early years work focused on the class of affine term structure models.

The classical models are Vasicek (1977) and Cox, Ingersoll, and Ross (1985). Duffie and Kan (1996) generalized the literature and Dai and Singleton (2000) characterized the set of admissable and identifiable models. Later a class of models was introduced that focused on fitting the term structure at a given point in time to ensure no arbitrage opportunities exist (Hull and White (1990) and Heath, Jarrow, and Morton (1992)). It has been shown that the forecasts obtained using the first class of models do not outperform the random walk, see for example Duffee (2002). The second class of models focuses on the cross-section dimension of yields but not on the time series dimension. Time series models aim to describe the dynamical properties and are therefore more suited for forecasting. This may partly explain the renewed interest in statistical time series models for yields.

The papers of Diebold and Li (2006, DL) and Diebold, Rudebusch, and Aruoba (2006, DRA) have shifted attention back to the Nelson and Siegel (1987, NS) model. DL and DRA consider a statistical three factor model to describe the yield curve over time. The three factors represent the level, slope and curvature of the yield curve and thus carry some economical interpretation. More importantly, they show that the model-based forecasts outperform many other models including standard time series models such as vector autoregressive models and dynamic error-correction models. In DRA, the Nelson-Siegel framework is extended to include non-latent factors such as inflation. Further they frame the Nelson-Siegel model into a state space model where the three factors are treated as unobserved processes and modelled by vector autoregressive processes. A wide range of statistical methods associated with the state space model can be exploited for maximum likelihood estimation and signal extraction, see Durbin and Koopman (2001). We will follow this approach in which the state space representation of the Nelson and Siegel (1987) model plays a central role.

Parameter estimation in Diebold and Li (2006) and Diebold, Rudebusch, and Aruoba (2006) relies on two simplifying assumptions. First, to allow the time-varying factors to be estimated in a linear setting the factor loadings are kept constant over time for each maturity. In the original Nelson and Siegel (1987) model the factor loadings depend on a parameter that is also time-varying, which in DL is restricted to be constant to keep the factor loadings constant. Second, volatility is kept constant over the full sample period.

We contribute to the literature by introducing time-varying factor loadings and time-varying volatility in the dynamic Nelson-Siegel model. First, a time-varying loading para-

meter is introduced and specified as a step-function through time. This specification leads to an improvement of model fit. Then the factor loadings are allowed to change gradually over time via a flexible cubic spline function. Improvements in model fit are achieved here as well. Second, we introduce a time-varying volatility specification in the model. Empirically it is found that during high volatility periods, the yields for all maturities are highly volatile although some maturities are more volatile than others. An cubic spline function of time is considered for the overall or common volatility. This spline is multiplied by a scaling factor and a constant level of volatility is added for each maturity. Similar to the introduction of time-varying factor loadings, time-varying volatility significantly increases model fit.

These findings may shed some light on more recent developments in the term structure literature. The dynamic Nelson-Siegel model does not rely on theoretical concepts such as the absence of arbitrage, see also the discussion in Ang and Piazzesi (2003). Recently, Christensen, Diebold, and Rudebusch (2007) have modified the Nelson-Siegel framework to impose the abritrage-free condition. As a result, a new class of affine dynamic term structure models is defined. An important condition for the risk-free rate to exist in this framework is that loadings are constant over time. This condition may be validated by allowing the factor loadings to be time-varying as we do in this paper. Also in the work of Diebold, Li, and Yue (2007) on the global yield curve, constant factors are an important condition.

A third contribution to the literature is that we show how easily the Nelson and Siegel (1987) model in state space form treats missing observations. This is a general property of state space models, but has not yet been explored in this context. Besides the standard unsmoothed Fama-Bliss monthly yields dataset for the period 1972-2000 (as used by DRA and others), we also estimate the time-varying model for U.S. Treasury yields over the period January 1972 up to June 2007 obtained from the Federal Reserve Economic Database (FRED). The latter dataset is interesting as it has more recent data, but can not easily be used in the OLS framework due to its many missing values. We show that in the state space framework unbalanced datasets can be treated in a straightforward manner. In particular, by combining the two datasets we show how well the smoothed values for the missing data approximates the true value. Using the state space framework allows us to include the longest maturity bond (maturing in 30 years), which was not issued during the period February 2002 until January 2006.

There are a number of papers that extend the work of DL and DRA for the NS model. Bianchi, Mumtaz, and Surico (2006) allow for time-varying variance for the factors. It is implied that the factor loadings for the term-structure are also appropriate weights for the variation in the term-structure. This is a strong assumption that needs to be validated. To avoid this restricted framework, we introduce for each yield in the observation equation, a different factor loading for the common volatility factor that is modelled as a flexible function of time. Yu and Zivot (2007) extend the Nelson-Siegel framework by including corporate bonds. De Pooter (2007) examines the dynamic NS model that is extended by additional factors. It is shown that such extentions can improve both the in-sample fit and the post-sample forecasting performance. De Pooter, Ravazzolo, and Van Dijk (2007) study the forecasting ability of the Nelson-Siegel model by focusing on the predictive gains that can be obtained when macroeconomic variables are included and forecasts of different model specifications are combined. Without adopting the Nelson-Siegel framework, Bowsher and Meeks (2006) introduce a 5-factor model where splines are used to model the yield curve and where the knots for these splines act as factors. While their approach allows for a more flexible yield curve some economic intuition of the factors is lost. Moreover, also in this framework, volatility is kept fixed over time.

The rest of the paper is organised as follows. Section 2 describes the baseline dynamic Nelson-Siegel latent factor model and Section 3 discusses our new extensions. In Section 4 we present, discuss and compare estimation and forecasting results for different model specifications. Section 5 presents an illustration for an unbalanced dataset. Section 6 concludes.

2 The Nelson-Siegel Latent Factor Yield Curve Model

In this section we introduce the latent factor model that Nelson and Siegel (1987) develop for the yield curve. First we focus on the model that is slightly adjusted in terms of factorization by Diebold and Li (2006). Second we discuss the state space approach as proposed by Diebold, Rudebusch, and Aruoba (2006).

2.1 The Nelson-Siegel model

Interest rates are denoted by $y_t(\tau)$ at time t and maturity τ. For a given time t, the yield curve $\theta_t(\tau)$ is some smooth function representing the interest rates (yields) as a function of maturity τ. A parsimonious functional description of the yield curve is proposed by Nelson and Siegel (1987). The Nelson-Siegel formulation of the yield is modified by Diebold and Li (2006, henceforth DL) to lower the coherence between the components of the yield curve.

The DL formulation is given by

$$\theta_t(\tau) = \theta(\tau; \lambda, \beta_t) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right), \tag{1}$$

where $\beta_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})'$, for given time t, maturity τ and fixed coefficient λ that determines the exponential decay of the second and third component in (1).

The shape and form of the yield curve is determined by the three components and their associated weights in β_t. The first component takes the value 1 (constant) and can therefore be interpreted as the overall level that influences equally the short and long term interest rates. The second component converges to one as $\tau \downarrow 0$ and converges to zero as $\tau \to \infty$ for a given t. Hence this component mostly influences short-term interest rates. The third component converges to zero as $\tau \downarrow 0$ and as $\tau \to \infty$ but is concave in τ, for a given t. This component is therefore associated with medium-term interest rates.

Since the first component is the only one that equals one as $\tau \to \infty$, its corresponding β_{1t} coefficient is usually linked with the long-term interest rate. By defining the slope of the yield curve as $\theta_t(\infty) - \theta_t(0)$, it is easy to verify that the slope converges to $-\beta_{2t}$ for a given t. Finally, the shape of the yield can be defined by $[\theta_t(\tau^*) - \theta_t(0)] - [\theta_t(\infty) - \theta_t(\tau^*)]$ for a medium maturation τ^*, say, two years, and for a given t. It can be shown that this shape approximately equals β_{3t}.

In case we observe a series of interest rates $y_t(\tau_i)$ for a set of N different maturities $\tau_1 < \ldots < \tau_N$ available at a given time t, we can estimate the yield curve by the simple

regression model

$$\begin{aligned} y_t(\tau_i) &= \theta_t(\tau_i) + \varepsilon_{it} \\ &= \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) + \varepsilon_{it}, \end{aligned} \tag{2}$$

for $i = 1, \ldots, N$. The disturbances $\varepsilon_{1t}, \ldots, \varepsilon_{Nt}$ are assumed to be independent with mean zero and constant variance σ_t^2 for a given time t. The least squares method provides estimates for the β_{jt} coefficients $j = 1, 2, 3$. These cross-section estimates can be obtained as long as sufficient interest rates for different maturities are available at time t.

The series of regression estimates for β_t, for all time periods $t = 1, \ldots, T$, appear to be strongly correlated over time. In other words, the coefficients are forecastable and hence the Nelson-Siegel framework can be used for forecasting in this way. This has been recognized by DL who implemented the following two-step procedure: first, estimate the β_t by cross-section least squares for each t; second, treat these estimates as three time series and apply time series methods for forecasting β_t and hence the yield curve $\theta(\tau; \lambda, \beta_t)$.

DL compare forecasts obtained using this method with other methods, such as forecasts based on the random walk model, the univariate autoregressive models and the trivariate vector autoregressive models. These different methods produce rather similar results. Nevertheless, the two-step forecasting approach does better than forecasting the different interest rates series directly, especially for the longer maturities.

2.2 The dynamics of the latent factors

Diebold, Rudebusch, and Aruoba (2006, henceforth DRA) go a step further by recognizing that the Nelson-Siegel framework can be represented as a state space model when treating β_t as a latent vector. For this purpose, the regression equation (2) is rewritten by

$$y_t = \Gamma(\lambda)\beta_t + \varepsilon_t, \qquad y_t = [y_t(\tau_1), \ldots, y_t(\tau_N)]' \tag{3}$$

with $N \times 3$ factor loading matrix $\Gamma(\lambda)$ where its (i, j) element is given by

$$\Gamma_{ij}(\lambda) = \begin{cases} 1, & j = 1, \\ \left(1 - e^{-\lambda \cdot \tau_i}\right) \ / \ \lambda \cdot \tau_i, & j = 2, \\ \left(1 - e^{-\lambda \cdot \tau_i} - \lambda \cdot \tau_i e^{-\lambda \cdot \tau_i}\right) \ / \ \lambda \cdot \tau_i, & j = 3. \end{cases}$$

The observation disturbance vector is given by

$$\varepsilon_t \sim NID(0, \Sigma_\varepsilon), \qquad \varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})', \qquad t = 1, \ldots, n.$$

The time series process for the 3×1 vector β_t can be modelled by the vector autoregressive process

$$\beta_{t+1} = (I - \Phi)\mu + \Phi\beta_t + \eta_t, \qquad \eta_t \sim NID(0, \Sigma_\eta), \tag{4}$$

for $t = 1, \ldots, n$, with mean vector μ and initial condition $\beta_1 \sim N(\mu, \Sigma_\beta)$ where coefficient matrix Φ and variance matrix Σ_β are chosen such that $\Sigma_\beta - \Phi\Sigma_\beta\Phi' = \Sigma_\eta$ and stationarity of the vector autoregressive process is ensured. An effective reparameterisation for this is developed by Ansley and Kohn (1986). This approach differs from that of DRA but we have found that our estimates of the parameter coefficients are similar to those reported by DRA.

2.3 Estimation and forecasting using state space methods

The dynamic Nelson-Siegel model is represented by the equations (3) and (4) with an appropriate initial condition for β_1. These equations can be regarded as a special case of the linear Gaussian state space model as discussed, for example, in Durbin and Koopman (2001). The unobserved factors of β_t are placed in the state vector and can be estimated by the Kalman filter and the associating smoothing algorithm. The model parameters, including the 3×3 autoregressive coefficient matrix Φ, the 3×3 variance matrix Σ_η and the $N \times N$ diagonal variance matrix Σ_ε, are treated as fixed and are unknown. Estimation of these parameters is based on the method of maximum likelihood. The likelihood function is routinely evaluated by the Kalman filter for a given value of the parameter vector. A quasi-Newton optimization method is then employed to maximize the likelihood function with respect to the parameter vector. The score function is evaluated numerically for this purpose. This approach is

implemented in the matrix programming language `Ox` of Doornik (2001) with the use of the `SsfPack` state space functions developed by Koopman, Shephard, and Doornik (1999).

The state space framework allows for many other variations of the dynamic Nelson-Siegel model. In particular, different dynamic processes for vector β_t can be considered. Also, the variance matrix of the factor disturbances Σ_η can be taken as a full or as a diagonal matrix. Diebold, Rudebusch, and Aruoba (2006) assume that the measurement disturbance variance matrix Σ_ε is diagonal so that the yield equations of different maturities are uncorrelated, given the factors f_t. This assumption is often used since it reduces the number of coefficients that need to be estimated. It therefore leads to computational tractability of the estimation process when a potentially large number of yields is available.

Finally, the Kalman filter also forms the basis for forecasting. The Nelson-Siegel model where β_t is modeled by the vector autoregressive process (4) has been developed primarily for the forecasting of yield curves. Yu and Zivot (2007) conclude that forecasting results have not improved considerably compared to the two-step approach of Diebold and Li (2006). This applies to forecasting performances for short-term and long-term maturities. De Pooter (2007) confirms these findings but he shows that forecasts can be improved by considering more than three factors.

3 Time-Varying Factor Loadings and Volatility

In the previous section we have discussed the dynamic Nelson-Siegel model for the yield curve as proposed by Diebold, Rudebusch, and Aruoba (2006). Next we propose two extensions of their time series modelling approach of interest rates: (i) introduction of time-varying factor loadings and (ii) introduction of time-varying volatility. Motivations and further details of these extensions are given below.

3.1 Time-varying factor loadings

In the dynamic Nelson-Siegel model, the parameter λ determines the shape of the yield curve. In the earlier studies, the default is to pre-fix a value for λ and not necessarily estimate it. For example, Diebold and Li (2006) fix λ at 0.0609, Diebold, Rudebusch, and Aruoba (2006) estimate λ to be 0.077. Yu and Zivot (2007) adopt these estimates in their empirical study

concerning corporate bonds. They argue that the loadings $\Gamma_{ij}(\lambda)$ are not very sensitive to different values of λ as can be illustrated graphically. Hence there is no need to estimate λ and they fix λ so that it maximizes the loading on the curvature component at some medium term (that is, 30 months for $\lambda = 0.0609$ and 23.3 months for $\lambda = 0.077$).

We conclude that the value for λ is usually fixed in empirical studies although different values for λ are considered in different studies and sometimes it is estimated. We emphasize that the estimation of λ is straightforward in a state space framework as it can be included in the parameter vector. Its estimation by maximum likelihood methods is described in the previous section. Keeping λ fixed over the full sample period may be too restrictive as the data usually spans over a long time period. The characteristic of the yield curve may have changed over a longer time period and λ determines the shape of the yield curve directly. The factors themselves also influence the yield curve and they are time-varying and modelled by a vector autoregressive process. However, the maturity at which the curvature factor is maximized and the speed of decay of the slope parameter depend only on λ and are fixed as a result. The importance of λ and its constancy over time is also discussed in Christensen, Diebold, and Rudebusch (2007, Proposition 1) where an arbitrage-free version of the Nelson-Siegel framework is proposed. Given the importance of λ, we study its role in more detail by considering possible changes of λ over time.

The time-varying λ is formalized as follows. We let λ depend on some function of time t, that is

$$\lambda_t = f(t; \lambda^*), \qquad t = 1, \ldots, n,$$

where $f()$ is a function depending on time-index t and $k \times 1$ vector of coefficients λ^*. We consider two possible specifications for $f()$ below. The first specification is a step function as represented by

$$\lambda_t = \ell_t \lambda^*, \qquad t = 1, \ldots, n,$$

where ℓ_t is a particular row of the $k \times k$ identity matrix for $t = 1, \ldots, n$. The sequence of ℓ_t is known and is typically chosen such that λ_t represents a step-function of k steps or, in other words, k different constants for k different periods.

An example of a smooth time function for $f()$ is the cubic spline function. The regression representation of a cubic spline function is particular convenient for our purposes, see Poirier

(1976). It allows the linear specification of $f()$ as given by

$$\lambda_t = w_t' \lambda^*, \qquad t = 1, \ldots, n,$$

where w_t is the $k \times 1$ vector of interpolating weights which are determined by certain smoothness conditions at the knot points. These knot points can be regarded as the break-points in the function. In both formulations of λ_t, the break-points need to be determined a-priori. In our empirical work reported below, we will consider both of these specifications.

3.2 Time-varying volatility

Another key aspect in the analysis of the term structure is the recognition that interest rates are the result of trading. Therefore, the volatility in the series may change over time. In most empirical work on the yield curve, monthly time series of interest rates are analyzed under the assumption that the volatility in the time series is constant over time. A few exceptions are Engle, Ng, and Rothschild (1990) and Bianchi, Mumtaz, and Surico (2006). However, investigating time-varying volatility in the context of the dynamic Nelson-Siegel model is a novelty.

In our empirical work below we do find changes in the volatility of interest rate series for different maturities but they appear to evolve slowly over time. Furthermore, the changes between high and low volatilities in the yields for the different maturities appear to occur at the same time which leads to common patterns of time-varying volatility. We therefore adapt the model and allow the variances of the yields for different maturities to be driven by two components. The first component v_t is the time-varying log-volatility component which is common to all maturities. It is multiplied by a constant scaling parameter for each maturity α_i^2. The second component σ_i^2 is a constant and measures the overall individual volatility of the yield for maturity i with $i = 1, \ldots, N$.

The variance matrix of the observation disturbances is given by Σ_ε. Below, we will treat it as a time-varying matrix, that is $\Sigma_{\varepsilon,t}$ and it will be specified as a diagonal matrix with elements h_{it} for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. Each element h_{it} represents the variance of the series of yields for the ith maturity at time t. The diagonal elements are thus modelled

by

$$h_{it} = \sigma_i^2 + \alpha_i^2 \exp(v_t), \tag{5}$$

where the common time-varying log-volatility component v_t is specified by a flexible smooth function of time. A typical example is the smoothing spline function but other functional forms can also be considered.

4 Data and Empirical Findings

For our first and main empirical analysis of yield curves we use the dataset as used by Diebold, Rudebusch, and Aruoba (2006). We will first give a short description of the data, with some summary statistics. Then we will outline our empirical results.

4.1 Standard Fama-Bliss dataset

The dataset we use is the unsmoothed Fama-Bliss zero-coupon yields dataset, obtained from the CRSP unsmoothed Fama and Bliss (1987) forward rates. We study U.S. Treasury yields with maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108 and 120 months over the period from January 1972 to December 2000. This dataset is the same as used by Diebold, Rudebusch, and Aruoba (2006)[1], Diebold and Li (2006) provide more details on how it is obtained.

[insert Table 1]

Table 1 provides summary statistics for our dataset. For each maturity we show the mean, standard deviation, minimum, maximum and some autocorrelation coefficients. In addition we show these statistics for proxies of the level, slope and curvature coefficients. These proxies are chosen in accordance with the construction of the factors from the Nelson and Siegel (1987) model, see the discussion in Section 2.1, and have previously been used, among others, by Diebold and Li (2006).

From the table we see that the average yield curve is upward sloping. Volatility decreases by maturity, with the exception of the 6 month being more volatile than the 3 month. Impor-

[1]We thank Francis X. Diebold for making the dataset available on his website: `www.ssc.upenn.edu/ fdiebold/`.

tant for econometric analyses, yields for all maturities are very persistent. The persistence is most notable for long term bonds. However, with a first-order autocorrelation of 0.970, the 3 month bill is still highly persistent. The level, slope and curvature proxies are persistent but to a lesser extent. The curvature and slope proxies are least persistent given the twelfth-order autocorrelation coefficients of 0.259 and 0.410, respectively.

[insert Figure 1]

Figure 1 shows the cross-section of yields over time. This is a graphical representation of our data-set. In addition to the findings of Table 1 we see a few interesting characteristics. The first noticable fact is that yields vary significantly over time from which various common dynamics across all yields can be deduced. Especially in the years 1978-1987 interest rates are remarkable high and volatile. Secondly, the shape of the yield curve is not constant over time. Though on average it is upward sloping, there are periods when it is downward sloping or humped.

4.2 Empirical Results

In Sections 4.2.2 and 4.2.3 we analyze how our extensions affect the performance of the Nelson-Siegel latent factor model. Section 4.2.4 discusses the results from the Nelson-Siegel with both of our model extensions jointly included. In Section 4.2.5 we discuss the forecasting performances. Before we do this we will first provide results for the baseline Nelson-Siegel latent factor model in Section 4.2.1.

4.2.1 The baseline dynamic Nelson-Siegel model

[insert Table 2]

Table 2 presents the estimates of the vector autoregression (VAR) model for the latent factors. The high persistence from the proxies for the level, slope and curvature that we report in Table 1 are confirmed by the high diagonal elements of the VAR coefficient matrix. The estimates in this table are almost identical to those in Diebold, Rudebusch, and Aruoba (2006, Table 1, p.316). The slight difference stems from our use of the Ansley and Kohn (1986) method to ensure stationarity, see Section 2.2.

The factor loadings parameter λ is estimated as 0.0778, with a standard error of 0.00209. The high significance of this estimate confirms that interest rates can be informative about λ and that small changes in the loadings can have a significant effect on the likelihood value.

[insert Table 3]

Table 3 reports sample means and standard deviations of measurement and prediction errors. Panel A of this table focuses on the prediction errors, while Panel B looks at the measurement errors. The prediction errors are obtained from the Kalman filter. The measurement errors are defined as the actual yields minus the yields that are obtained using the estimated parameters and smoothed estimates of level, slope and curvature. We find that in particular the 3 month rate is difficult to fit: it has the highest mean prediction and measurement error. The standard deviations reported in Table 3 indicate that the relatively long bonds are predicted most accurately. The shortest and longest bonds are however difficult to fit. The medium term notes of around 24 month are modelled adequately and produce the best fit.

4.2.2 Time-varying factor loadings

To obtain some indication whether the λ parameter varies over time in our dataset, we estimate the baseline model for four equally sized subperiods. The four estimates of λ are 0.0397, 0.126, 0.0602 and 0.0695. The corresponding standard errors are sufficiently small for concluding that the estimates are distinct from each other (except for the last two subsamples). This finding confirms that datasets of interest rates can be informative about the λ parameter and that the assumption of constant factor loadings over time does not necessarily hold.

[insert Table 4]

Table 4 presents the estimated factor loadings parameter for different specifications of a time-varying λ_t together with standard errors. The likelihood ratio (LR) tests are also reported. They enable model comparisons. The estimates of the factor loadings parameter λ parameterized as a step function, see Section 3.1, confirm the earlier finding based on the subsample estimates of λ. All other parameters in the model are assumed to be constant

over the full sample period and therefore the estimates differ from the subsample estimates of λ. However, again we find low standard errors and distinctly different estimates for λ over time. The LR-test shows significant improvement in model fit over the baseline model with constant λ, even at the 1% level.

[insert Figure 2]

Finally, we estimate the factor loadings parameter λ in its most flexible form. This is the spline function as defined in Section 3.1. For the knots we choose to divide the interval in four equally spaced intervals. Together with knots at the begin- and end-point of the sample, we place knots at April 1979, July 1986 and October 1993.[2] Figure 2 presents the estimates of the time-varying λ in a graph. From this figure it is clear that even within each subperiod, the factor loadings parameter λ is not constant over time. In Table 4 we show the mean λ for each of the periods. On average, the mean λ for these periods varies less than the estimates obtained from models with subsamples and with the step function. However, there still is considerable evidence of a varying λ parameter over time. The LR test shows that the model with a spline for the factor loadings parameter improves the fit most significantly compared to both the constant λ model and the model with a step function.

To gain further insights in the impact of a varying λ parameter, we present in Figure 3 the slope and curvature factor loadings for the different maturities, that is $\Gamma_{i2}(\lambda_t)$ and $\Gamma_{i3}(\lambda_t)$ for $i = 1, \ldots, N$ in equation (3) for a particular time point t. The factor loadings are presented for the minimum value of λ_t and the maximum value of λ_t in panel (A). Since the maximum of λ_t occurs at the end of the sample where the uncertainty with respect to the estimate of λ_t is high, we also present the loadings for a local maximum of λ_t in panel (B). It is apparent from these graphs that the time-varying λ_t can lead to significantly different factor loading patterns for slope and curvature. We conclude from this finding that our modifications to the dynamic Nelson-Siegel model are not necessarily marginal.

[2] We have tried various different number of knots, see Figure 1 in the Appendix (available from the authors on request). Adding more knots will improve the model looking at both the likelihood and Akaike Information Criterion (AIC). However, to keep the number of parameters tractable while allowing the loadings parameter to vary over time we choose a number of knots which provides a shape that represented a wide range of values for the number of knots. The main change in shape of the spline when using a large number of knots is in the end of the sample period. Adding more knots shows the high value of the spline there is mainly caused by the inverted shape of the yield curve around 1999.

Table 3 compares the prediction and measurement error of this model with the baseline constant λ model. For 10 (9) out of the 17 maturities the mean prediction (measurement) error is lower. This is particularly true for short maturities. The standard deviation of the prediction (measurement) error is lower for 13 (9) out of the 17 maturities.

4.2.3 Time-varying volatility

The second modification we propose in Section 3 is to allow for a time-varying volatility. In a similar way as we obtain time-varying factor loadings, we introduce a spline function for the time-varying common volatility pattern in the observation disturbances, see Section 3.2 for details. The knots are set to equal those of Section 4.2.2.[3]

[insert Figure 4]

Panel (A) of Figure 4 presents the spline for the time-varying common volatility component. This spline shows how the average overall or common volatility varies over time. When mulitplied by a loading for each individual maturity, it provides an estimate of the volatility pattern for each maturity up to a constant. Panel (B) in Figure 4 gives the resulting volatilities for a selection of maturities. Interestingly, we find that volatility is especially high in the period between 1980 and 1987. Thereafter it is almost completely constant for all maturities.

Table 3 compares the measurement and prediction error of this model with the baseline constant volatility model and the time-varying λ model. For 9 (10) out of the 17 maturities the mean prediction (measurement) error is lower. It is interesting to note that the improvements of a time-varying λ are mainly obtained for the short maturities while those for time-varying volatility are most pronounced for long bonds. The standard deviation of the prediction (measurement) error is lower than the baseline case for only 2 (8) out of the 17 maturities.

[insert Table 5]

In Table 5 we report the performance of the various models. We report the loglikelihood and the Akaike Information Criterion (AIC) value, together with an LR-test for model

[3]See Footnote 2 for more details. The shape of the spline for volatility depends less on the number of knots chosen than was the case for the factor loadings parameter, as can be seen from Figure 2 in an Appendix of this paper (available from the authors on request).

improvement. We find a highly significant improvement of the model over the baseline model without time-varying volatility. The increase in likelihood value and decrease in AIC is higher than is the case for the baseline model versus the time-varying factor loadings model. This indicates that most gains in describing the yield curve in this dataset are obtained by introducing time-varying volatility.

4.2.4 Time-varying factor loadings and volatility

In this section we report our results of analyzing the dataset using the dynamic Nelson-Siegel model with both factor loadings and volatility time-varying. As the parameters of the vector autoregression are similar to those in Table 2 and the splines similar to Figures 2 and 4 we do not repeat these here[4].

In Table 3 the measurement and prediction error of this model are given. In comparison with the baseline Nelson-Siegel model, we observe that the prediction (measurement) error is lower for 11 (11) out of the 17 maturities. Finally, we find that the standard deviations of the prediction errors are improved for 13 of the 17 maturities while only 6 maturities show such improvements for the measurement errors.

The loglikelihood and Akaike Information Criterion (AIC) values reported in Table 5 show strong significant improvements compared to the baseline model. Also when we benchmark them against the model with only time-varying factor loadings or only time-varying volatility we obtain significant improvements. We therefore conclude that, though adding time-varying volatility is most significant, both model extensions significantly contribute to improving the Nelson-Siegel latent factor model.

[insert Figure 5]

In Figure 5 the latent factors obtained from the model are compared with their data-based proxies. Each of the factors agrees with their data-based proxy: the level factor is close to the 120 month yield, the slope is close to spread of 3 month over 120 month yields and the curvature is close to the 24 month yield minus the 3 and 120 month yield. Finally, in Figure 6 we discuss the four selected fitted yield curves as earlier reported in Diebold and Li (2006, Figure 5). We report the yield curve obtained from the DL OLS model, DRA SSF

[4]These additional results are reported in an Appendix of this paper (available on request).

model and our extended SSF model with both the factor loadings and volatility time-varying. Especially in the August 1998 it is clear that our model extensions allow for more flexibility and improve the fit of the model.

[insert Figure 6]

4.2.5 Some forecasting evidence

Although our modifications are not primarily motivated by improving the forecast performance of the dynamic Nelson-Siegel model, it is however interesting to empirically investigate the quality of our forecasts for this dataset. For this purpose we have created similar tables as in DL. To limit the number of tables, we only present root mean squared error (RMSE) measures of our forecasts in Table 6, see Footnote 4. The forecasting exercise is carried out for two post-samples: (A) from 1994 onwards and (B) from 1989 onwards. We present our forecasting results for these two samples since (A) is used in DL and DRA and (B) illustrates the variations in the forecasting results that can be obtained. Therefore, these forecasting results should be interpreted with some reservation.

[insert Table 6]

The forecasting results reported in Panel (A) of Table 6 confirm that the two time-varying extensions (time-varying λ and time-varying volatility) do not lead to clear forecast improvements compared to the baseline model DRA. However, some improvements can be observed. These results also confirm that the forecast performances of the univariate random walk model and the DL method are overall superior to DRA. Similar findings have been reported by Yu and Zivot (2007). To balance such conclusions, we also investigated the forecasting performances for a longer post-sample starting in 1989, see Panel (B) of Table 6. Here the results look more promising with respect to DRA and our modifications. In some cases the forecast performances are highly improved as a result of our extensions of the model.

[insert Figure 7]

Another interesting aspect in forecasting is the precision at which a forecast is produced. In the dynamic Nelson-Siegel model-based framework, the standard error can be produced

for each forecast. The standard error is a precision measure as it determines the confidence interval around the forecast. Different model specifications produce different forecast standard errors. In Figure 7 the forecasts for four maturities, for two different model specifications and for the months in 1984 are presented together with confidence intervals. In forecasting 1 year and 3 year maturities, the extended model produces more precise forecasts compared to the baseline model while for 3 month and 10 year the precision is more or less equal. Graphs as in Figure 7 can be presented for other years, for other maturities and for other model specifications. It is therefore not straightforward to present an overall picture on improvements in forecast precision. However, Figure 7 does illustrate that our modifications can produce more precise forecasts.

5 An illustration for an unbalanced dataset

An attractive feature of models in state space form is that they can allow for missing values. For OLS estimation of the Nelson-Siegel model (as put forward by Diebold and Li (2006)) data must be available for all periods to avoid ad-hoc measures. However, this is not the case for models in state space form.

The smoothing algorithm associated with the Kalman filter[5] produces the smoothed estimates of the latent factors for all periods and based on the available observations in the dataset. The estimation procedure itself does not change depending on data availability. Moreover, the smoothed estimates of the factors do also generate smoothed estimates of the interest rates for all maturities. It implies that when data is missing for a certain maturity, we are still able to obtain estimates for this maturity which are then based on data from this maturity at other time periods and data from other maturities. It is therefore expected that we still can obtain accurate forecasts for a maturity although the data is not complete for this maturity. This is a strong property of Kalman filter methods.

To illustrate this we use a publicly available dataset of fixed maturity U.S. Treasury yields. The dataset is obtained from the Federal Reserve Economic Data (FRED) online database, maintained by the Federal Reserve Bank of St. Louis. We look at the fixed

[5]See Section 2.3 for our discussion of state space models, and Durbin and Koopman (2001) for a discussion on state space models generally.

maturity interest rates, over the period January 1972 up to June 2007, with maturities of 1, 3, 6, 12, 24, 36, 60, 84, 120, 240 and 360 months. With the convenience of being freely available and covering a long time horizon, there comes the inconvenience that many missing entries are present in the dataset. For example, the dataset for the 3 month bill starts only in January 1982, for the 24 month this is June 1976 and the 360 month starts February 1977 with missings in the period March 2002 until January 2006 (the period when it was not issued).

[insert Figure 8]

We have estimated the Nelson-Siegel latent factor model with both time-varying factor loadings and time-varying volatility for this dataset. Figure 8 shows the time series of the maturities with missings that are mentioned above. We see that using the smoothed latent factors, based on data available not only in that period but also during other periods, an estimate of the missing yield can be obtained.

To obtain an indication of the reliability of this estimate, we compare it to its value obtained from the dataset of Section 4.1. Unfortunately, this dataset does not contain the 30 year bond and therefore we can only make such comparisons for the 3 month and 24 month yields. We find that in both cases the smoothed yield obtained provides a reliable estimate of the missing data.

6 Conclusion

In this paper we propose two extensions for the dynamic Nelson-Siegel model for the yield curve of Diebold, Rudebusch, and Aruoba (2006). We first consider the loadings parameter of the latent factors which is usually assumed fixed at some constant known value or estimated as constant through time. Our empirical results show that this parameter can be estimated with a small standard error. This implies that the data can be highly informative about the shape of the factor loadings. We then investigate whether this parameter is time-varying. For this purpose we consider a simple step function and a spline function. We show that these extensions can lead to significant improvements in model fit. Next we turn our attention to the volatility for each of the maturities. We propose to include a common time-varying

volatility component that is specified as a spline function. This common volatility component is multiplied by a loading parameter and a constant variance is added for each maturity. This extension also provides a significant improvement in model fit. We also show that forecast results are also affected by these modifications of the baseline model. Finally we illustrate that missing values can be easily treated in this modelling framework. For example, we consider a dataset where four years of the 30 year bond has not been issued. However, with the use of both data from other periods of the same maturity and data from other maturities we obtain accurate estimates of the missing values. Given the general framework of the dynamic Nelson-Siegel model, other modifications can be considered in future work.

References

Ang, A. and M. Piazzesi (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *J. Monetary Economics 50*, 745–87.

Ansley, C. F. and R. Kohn (1986). A note on reparameterizing a vector autoregressive moving average model to enforce stationarity. *J. Statistical Computation and Simulation 24*, 99–106.

Bianchi, F., H. Mumtaz, and P. Surico (2006). The UK Great Stability: a View from the Term Structure of Interest Rates. Working Paper.

Bowsher, C. and R. Meeks (2006). High Dimensional Yield Curves: Models and Forecasting. Working Paper.

Christensen, J., F. Diebold, and S. Rudebusch (2007). The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models. Manuscript University of Pennsylvania.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica 53*, 385–407.

Dai, Q. and K. J. Singleton (2000). Specification Analysis of Affine Term Structure Models. *J. Finance 55*(5), p1943–1978.

De Pooter, M. (2007). Examining the Nelson-Siegel Class of Term Structure Models. Working Paper.

De Pooter, M., F. Ravazzolo, and D. Van Dijk (2007). Predicting the term structure of interest rates. TI Discussion Paper.

Diebold, F. and C. Li (2006). Forecasting the Term Structure of Government Bond Yields. *J. Econometrics 130*, p337–364.

Diebold, F., C. Li, and V. Yue (2007). Global Yield Curve Dynamics and Interactions: A Generalized Nelson-Siegel Approach. Manuscript University of Pennsylvania.

Diebold, F., S. Rudebusch, and S. Aruoba (2006). The Macroeconomy and the Yield Curve. *J. Econometrics 131*, p309–338.

Doornik, J. A. (2001). *Object-Oriented Matrix Programming using Ox 3.0* (4th ed.). London: Timberlake Consultants Ltd. See `http://www.doornik.com`.

Duffee, G. (2002). Term Premia and Interest Rate Forecasts in Affine Models. *J. Finance 57*, p405–443.

Duffie, D. and R. Kan (1996). A Yield-factor Model of Interest Rates. *Journal of Mathematical Finance 6*(2), p379–406.

Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods.* Oxford: Oxford University Press.

Engle, R. F., V. K. Ng, and M. Rothschild (1990). Asset pricing with a factor ARCH covariance structure: empirical estimates for treasury bills. *J. Econometrics 45*, 213–238.

Fama, E. F. and R. R. Bliss (1987). The Information in Long-Maturity Forward Rates. *American Economic Review 77*, p680–692.

Heath, D., R. Jarrow, and A. Morton (1992). Bond Pricing and the Term Structure of Interest Rates: a New Methodology for Contingent Claims Valuation. *Econometrica 60*, p77–105.

Hull, J. and A. White (1990). Pricing interest rate derivative securities. *Rev. Financial Studies 3*, 573–92.

Koopman, S. J., N. Shephard, and J. A. Doornik (1999). Statistical algorithms for models in state space form using SsfPack 2.2. *Econometrics Journal 2*, 113–66.

http://www.ssfpack.com/.

Nelson, C. and A. Siegel (1987). Parsimonious Modelling of Yield Curves. *Journal of Business 60-4*, p473–489.

Poirier, D. J. (1976). *The Econometrics of Structural Change: with Special Emphasis on Spline Functions*. Amsterdam: North-Holland.

Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *J. Financial Economics 5*, p177–188.

Yu, W. and E. Zivot (2007). Forecasting the Term Structures of Treasury and Corporate Yields: Dynamic Nelson-Siegel Models Evaluation. Working Paper Uiversity of Washington.

Table 1: Summary Statistics

The table reports summary statistics for U.S. Treasury yields over the period 1972-2000. We examine monthly data, constructed using the unsmoothed Fama-Bliss method. Maturity is measured in months. For each maturity we show mean, standard deviation (*Std.dev.*), minimum, maximum and three autocorrelation coefficients, 1 month ($\hat{\rho}(1)$), 1 year ($\hat{\rho}(12)$) and 30 months ($\hat{\rho}(30)$).

Summary Statistics for each Maturity							
Maturity	Mean	Std.dev.	Minimum	Maximum	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(30)$
3	6.851	2.695	2.732	16.020	0.970	0.700	0.319
6	7.079	2.702	2.891	16.481	0.972	0.719	0.355
9	7.201	2.679	2.984	16.394	0.972	0.726	0.378
12	7.302	2.602	3.107	15.822	0.971	0.729	0.394
15	7.408	2.548	3.288	16.043	0.973	0.737	0.415
18	7.481	2.532	3.482	16.229	0.974	0.743	0.431
21	7.544	2.520	3.638	16.177	0.975	0.747	0.442
24	7.558	2.474	3.777	15.650	0.975	0.745	0.450
30	7.647	2.397	4.043	15.397	0.975	0.755	0.470
36	7.724	2.375	4.204	15.765	0.977	0.761	0.480
48	7.861	2.316	4.308	15.821	0.977	0.765	0.499
60	7.933	2.282	4.347	15.005	0.980	0.779	0.514
72	8.047	2.259	4.384	14.979	0.980	0.786	0.524
84	8.079	2.215	4.352	14.975	0.980	0.768	0.526
96	8.142	2.201	4.433	14.936	0.982	0.793	0.535
108	8.176	2.209	4.429	15.018	0.982	0.794	0.540
120(level)	8.143	2.164	4.443	14.925	0.982	0.771	0.532
slope	1.292	1.461	-3.505	4.060	0.929	0.410	-0.099
curvature	0.121	0.720	-1.837	3.169	0.788	0.259	0.076

Table 2: Baseline Model - Estimates of VAR Model for Latent Factors

The table reports the estimates of the vector autoregressive (VAR) model for the latent factors. The results shown correspond to the latent factors of the baseline Nelson-Siegel latent factor model. Panel A shows the estimates for the constant vector μ and autoregressive coefficient matrix Φ, Panel B shows the variance matrix Σ_η.

Panel A: Baseline Model - Constant and Autoregressive Coefficients of VAR				
	Level$_{t-1}$ ($\beta_{1,t-1}$)	Slope$_{t-1}$ ($\beta_{2,t-1}$)	Curvature$_{t-1}$ ($\beta_{3,t-1}$)	Constant (μ)
Level$_t$ ($\beta_{1,t}$)	0.997^{**}	0.0271^{**}	-0.0216^{*}	8.03^{**}
	0.00811	0.00889	0.0105	1.27
Slope$_t$ ($\beta_{2,t}$)	-0.0236	0.942^{**}	0.0392	-1.46^{**}
	0.0167	0.0176	0.0212	0.527
Curvature$_t$ ($\beta_{3,t}$)	0.0255	0.0241	0.847^{**}	-0.425
	0.023	0.0257	0.0312	0.537

An asterisk (*) denotes significance at the 5% level or less and two asterisks (**) denote significance at the 1% level or less. The standard errors are reported below the estimates.

Panel B: Baseline Model - Variance Matrix of VAR			
	Level$_t$ ($\beta_{1,t}$)	Slope$_t$ ($\beta_{2,t}$)	Curvature$_t$ ($\beta_{3,t}$)
Level$_t$ ($\beta_{1,t}$)	0.0949^{**}	-0.014	0.0439^{*}
	0.00841	0.0113	0.0186
Slope$_t$ ($\beta_{2,t}$)		0.384^{**}	0.00927
		0.0306	0.0344
Curvature$_t$ ($\beta_{3,t}$)			0.801^{**}
			0.0812

An asterisk (*) denotes significance at the 5% level or less and two asterisks (**) denote significance at the 1% level or less. The standard errors are reported below the estimates.

Table 3: Prediction and Measurement Errors

The table reports the prediction and measurement errors from the four Nelson-Siegel latent factor models we estimate. The *Baseline Model* corresponds to the baseline Nelson-Siegel latent factor model with constant factor loadings and volatility. The *Time-Varying Factor Loading* model corresponds to the model with a spline for λ. The *Time-Varying Volatility* model corresponds to the model with a spline for the volatility. The *Both Time-Varying* model corresponds to the model with a spline for both the factor loadings parameter and the volatility. For each maturity we show mean and standard deviation (*Std.dev.*). We summarize these per model with three statistics: the mean, median and number of maturities for which the absolute value is lower than that of the baseline model (*#Lower*). In Panel A we report prediction errors, in Panel B measurement errors.

Panel A: Prediction Errors (in basis points)								
	Baseline Model		Time-Varying Factor Loading		Time-Varying Volatility		Both Time-Varying	
Maturity	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.
3	-11.91	64.84	-9.32	63.63	-13.80	66.71	-11.38	64.17
6	-0.69	60.52	0.05	60.07	-1.86	61.18	-0.55	60.16
9	1.09	59.05	0.73	58.80	0.53	59.17	1.09	58.69
12	1.82	59.25	0.84	58.94	1.74	59.41	1.81	59.00
15	4.15	56.39	2.87	56.05	4.44	56.57	4.22	56.22
18	3.98	54.17	2.61	53.74	4.56	54.30	4.16	53.95
21	3.55	52.43	2.21	51.94	4.33	52.52	3.85	52.13
24	-1.16	52.28	-2.40	51.63	-0.25	52.33	-0.75	51.85
30	-2.57	48.73	-3.51	47.86	-1.56	48.76	-1.99	48.12
36	-3.29	46.83	-3.93	45.92	-2.35	46.87	-2.65	46.14
48	-2.01	44.29	-2.21	43.38	-1.47	44.43	-1.48	43.58
60	-3.60	40.85	-3.61	40.10	-3.60	41.09	-3.41	40.36
72	1.53	39.20	1.55	38.81	0.99	39.53	1.33	39.04
84	0.21	39.40	0.17	39.52	-0.83	39.49	-0.41	39.43
96	2.98	37.79	2.86	38.12	1.53	37.92	2.00	38.03
108	3.66	36.72	3.44	37.40	1.86	36.35	2.36	37.01
120	-1.95	37.91	-2.27	38.64	-4.05	37.40	-3.53	38.12
Mean	-0.25	48.86	-0.58	48.50	-0.58	49.06	-0.31	48.59
Median	0.21	48.73	0.17	47.86	-0.25	48.76	-0.41	48.12
#Lower			10	13	9	2	11	13

Table 3 (continued)

Panel B: Measurement Errors after Smoothing (in basis points)								
	Baseline Model		Time-Varying Factor Loading		Time-Varying Volatility		Both Time-Varying	
Maturity	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.
3	-12.66	22.42	-9.89	19.91	-14.80	28.43	-12.38	22.90
6	-1.37	5.08	-0.31	2.80	-2.63	10.50	-1.37	7.18
9	0.50	8.11	0.55	8.83	-0.04	7.97	0.44	8.25
12	1.31	9.86	0.82	10.21	1.33	9.58	1.31	9.96
15	3.71	8.72	3.00	8.49	4.18	8.91	3.85	8.69
18	3.63	7.28	2.86	6.71	4.42	7.82	3.92	7.35
21	3.27	6.51	2.57	6.08	4.28	7.31	3.71	6.72
24	-1.38	6.40	-1.95	6.54	-0.21	6.89	-0.80	6.64
30	-2.67	6.06	-2.92	6.42	-1.39	5.74	-1.89	6.08
36	-3.28	6.57	-3.24	6.84	-2.08	5.98	-2.43	6.42
48	-1.82	9.71	-1.41	9.66	-1.07	9.64	-1.12	9.69
60	-3.29	8.04	-2.77	7.48	-3.14	8.14	-2.95	7.78
72	1.94	9.14	2.41	9.04	1.50	10.03	1.83	9.77
84	0.68	10.38	1.03	10.38	-0.29	10.52	0.13	10.52
96	3.50	9.05	3.72	9.88	2.09	8.73	2.56	9.44
108	4.23	13.64	4.30	13.83	2.44	12.59	2.93	13.33
120	-1.35	16.44	-1.41	16.38	-3.46	14.94	-2.95	15.68
Mean	-0.30	9.61	-0.16	9.38	-0.52	10.22	-0.31	9.79
Median	0.50	8.72	0.55	8.83	-0.21	8.91	0.13	8.69
#Lower			9	9	10	8	11	6

Table 4: Estimates of Time-Varying Factor Loadings Parameter

The table reports various estimates of the time-varying factor loadings parameter λ. The *Baseline* model corresponds to the baseline Nelson-Siegel latent factor model (with constant factor loadings) as estimated for the full sample. The *Step function* model corresponds to the Nelson-Siegel latent factor model with a step function for λ. The *Spline* model corresponds to the Nelson-Siegel latent factor model with a spline for the λ.

Estimates of Time-Varying Factor Loadings Parameter			
	(a) Baseline	(b) Step Function	(c) Splines$^{\#}$
Full Sample	0.0778** 0.00209		
01/72 - 03/79		0.0932** 0.00533	0.0938
04/79 - 06/86		0.116** 0.00393	0.116
07/86 - 09/93		0.0638** 0.00240	0.0698
10/93 - 12/00		0.0717** 0.00419	0.0861
Loglikelihood	3185.4	3259.0	3289.0
AIC	-6304.8	-6446.1	-6503.9
LR-test (a) vs. (b)		147.2** 0.000	
LR-test (a) vs. (c)		207.2** 0.000	
LR-test (b) vs. (c)		60.0** 0.000	

An asterisk (*) denotes significance at the 5% level or less and two asterisks (**) denote significance at the 1% level or less. The standard errors are reported below the estimates, for the tests this is the probability H_0 is accepted. The $^{\#}$ indicates that the λ's from the model with a spline are the average over the period, see Figure 2.

Table 5: Loglikelihood and AIC of Model Extensions

The table reports the loglikelihood and Akaike Information Criterion (AIC) for the various model extensions proposed. The *Baseline* model corresponds to the baseline Nelson-Siegel latent factor model (with constant factor loadings) as estimated for the full sample. The *TV Factor Loadings* model corresponds to the baseline Nelson-Siegel latent factor model with a spline for the factor loadings parameter. The *TV Volatility* model corresponds to the Nelson-Siegel latent factor model with a common time-varying volatility component. The *TV Loadings & Volatility* model corresponds to the Nelson-Siegel latent factor model with both factor loading and volatility time-varying.

Performance of Model Extensions			
	Loglikelihood	AIC	LR-test vs. Baseline
Baseline	3185.4	-6304.8	
TV Factor Loadings	3289.0	-6503.9	207.2** 0.000
TV Volatility	4144.2	-8180.5	1917.6** 0.000
TV Loadings & Volatility	4187.3	-8258.5	2003.8** 0.000

An asterisk (*) denotes significance at the 5% level or less and two asterisks (**) denote significance at the 1% level or less. The probability H_0 is accepted is reported below the test-statistic.

Table 6: Forecast Performance of Various Model Extensions

The table reports the out-of-sample forecasting results from various specifications of the Nelson-Siegel model, with some benchmarks. The *RW* corresponds to forecasts obtained using the Random Walk model. *DL* corresponds to forecasts obtained using the Diebold and Li (2006) forecasting method with the factors obtained by OLS and prediced using univariate AR(1) models. *DRA* correspond to the forecasts obtained using the state space model as introduced by Diebold, Rudebusch, and Aruoba (2006). *TV* λ corresponds to our model with the time-varying factor loadings parameter λ, *TV Vol* to our model extension with time-varying volatility and *Both TV* to forecasts from a model with both extensions jointly. We compute the root mean squared error (*RMSE*) of the forecasts and summarize the forecasting performance by presenting the RMSE for 1-, 6- and 12-months-ahead forecasts.

Panel A: from January 1994						
Maturity	RW	DL	DRA	TV λ	TV Vol	Both TV
RMSE 1-month-ahead forecasting results						
3 months	0.179	0.160	0.180	0.179	0.307	0.279
1 year	0.240	0.227	0.236	0.238	0.232	0.241
3 years	0.277	0.263	0.271	0.289	0.270	0.273
5 years	0.275	0.279	0.280	0.295	0.278	0.284
10 years	0.253	0.249	0.268	0.292	0.266	0.264
RMSE 6-months-ahead forecasting results						
3 months	0.603	0.464	0.525	0.527	0.539	0.588
1 year	0.775	0.612	0.696	0.682	0.674	0.687
3 years	0.874	0.712	0.765	0.801	0.760	0.758
5 years	0.856	0.750	0.773	0.817	0.775	0.776
10 years	0.754	0.698	0.728	0.767	0.726	0.723
RMSE 12-months-ahead forecasting results						
3 months	1.013	0.812	0.834	0.860	0.831	0.897
1 year	1.190	0.907	0.953	0.969	0.925	0.937
3 years	1.230	1.026	0.999	1.057	1.000	0.996
5 years	1.184	1.087	1.013	1.070	1.024	1.029
10 years	1.045	1.056	0.978	1.020	0.994	0.994

Table 6 (continued)

Panel B: from January 1989						
Maturity	RW	DL	DRA	TV λ	TV Vol	Both TV
RMSE 1-month-ahead forecasting results						
3 months	0.207	0.245	0.221	0.234	0.258	0.232
1 year	0.266	0.342	0.661	0.350	0.307	0.292
3 years	0.299	0.355	0.782	0.430	0.329	0.311
5 years	0.289	0.347	0.611	0.421	0.325	0.313
10 years	0.257	0.283	0.340	0.369	0.279	0.279
RMSE 6-months-ahead forecasting results						
3 months	0.735	1.008	0.752	0.944	0.846	0.751
1 year	0.827	1.136	0.835	1.219	0.972	0.854
3 years	0.843	1.067	0.856	1.362	0.920	0.815
5 years	0.806	0.986	0.825	1.311	0.856	0.783
10 years	0.686	0.790	0.702	1.093	0.705	0.681
RMSE 12-months-ahead forecasting results						
3 months	1.315	1.653	1.334	1.773	1.383	1.260
1 year	1.330	1.717	1.336	1.970	1.440	1.255
3 years	1.203	1.554	1.210	1.961	1.273	1.070
5 years	1.113	1.425	1.127	1.809	1.156	0.990
10 years	0.961	1.187	0.967	1.470	0.963	0.875

Figure 1: Yield Curves from January 1972 up to December 2000

In this figure the U.S. Treasury yields over the period 1972-2000 are shown. We examine monthly data, constructed using the unsmoothed Fama-Bliss method. The maturities we show are 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108 and 120 months.

Figure 2: Time-Varying Factor Loadings Parameter

In this figure we present the time-varying factor loadings parameter. It is estimated with a spline, with knots at both the beginning and end of the sample, and at April 1979, July 1986 and October 1993.

Figure 3: Minimum and Maximum Factor Loadings

In this figure we present the slope and curvature factor loadings using the minimum and maximum value of the time-varying factor loadings parameter λ. In Panel (A) we do this for the maximum factor loadings parameter over the whole sample, in Panel (B) we exclude the last part of the sample for this.

(A) Minimum and Maximum over whole sample

(B) Minimum and Maximum, excluding sample after 1997

Figure 4: Time-Varying Volatility

In this figure we present the time-varying volatility. The volatility is estimated with a spline, with knots both at the beginning and end of the sample, and at April 1979, July 1986 and October 1993. The spline is loaded onto each maturity by a scalar and added to a constant volatility level per maturity. Panel (A) shows the estimated spline, Panel (B) depicts for a few maturities (3 months, 12 months, 36 months and 120 months) the volatility that is obtained using the spline and the loadings of each maturity.

(A) Spline for the Volatility

(B) Estimated Volatility for Some Maturities

Figure 5: Time-Varying Model - Level, Slope and Curvature

This figure reports the level, slope and curvature as obtained from the Nelson-Siegel latent factor model with both time-varying factor loadings and volatility. Panel (A) shows these together in one plot. Panels (B), (C) and (D) report them with their proxies from the data. For the level this is the 120 month treasury yield, for slope this is the spread of 3 month over 120 month yields and for curvature this is twice the 24 month yield minus the 3 and 120 month yield.

(A) Level, Slope and Curvature

(B) Level, with Proxy from Data

(C) Slope, with Proxy from Data

(D) Curvature, with Proxy from Data

Figure 6: Fitted Yield Curves for Four Months

This figure presents the fitted yield curve obtained from the Nelson-Siegel latent factor model with both time-varying factor loadings and volatility. The dots represent the actual yield curve, the solid line the fitted yield curve obtained from the Nelson-Siegel latent factor with both our extensions, the dashed line the model as put in state space form by Diebold, Rudebusch, and Aruoba (2006) and the dotted line the OLS model as in Diebold and Li (2006). We show these for four different months: March 1989, July 1989, May 1997 and August 1998.

Figure 7: Confidence Intervals of Forecasts

In this figure we present the forecasts with confidence intervals of the yield curve made at December 1983 for the full year 1984. The *DRA* model corresponds to the Nelson-Siegel model in state space form as introduced by Diebold, Rudebusch and Aruoba (2006). The *Both TV* model is this model with both time-varying factor loadings and time-varying volatility. We show the forecasts up to 12 months ahead with their 95% confidence intervals.

Figure 8: An Illustration with Missing Values

This figure illustrates how the Nelson-Siegel latent factor model deals with missing values. The data is the FRED fixed maturity U.S. Treasury Yields dataset from January 1972 up to June 2007 (note that all other tables and figures in this paper are based on the unsmoothed Fama-Bliss data). We show the yield from the data, the smoothed yield using the Nelson-Siegel latent factor model with time-varying factor loadings and volatility and, if available, the yield from the unsmoothed Fama-Bliss dataset.