The Dynamic Nelson-Siegel Model with Time-Varying Loadings and Volatility

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

September 2007 - Preliminary First Draft, Do Not Quote

Abstract

In this paper we explore time-varying parameter extensions of the dynamic Nelson-Siegel yield curve model for forecasting multiple sets of interest rates with different maturities. The Nelson-Siegel model is recently reformulated as a dynamic factor model where the latent factors level, slope and curvature are modelled simultaneously 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 a significantly increase of the model fit. 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 the time-series of a cross-section of yields has proven to be a challenging task. As with many topics in empricial 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. This model does not rely on theoretical frameworks such as the concept of no-arbitrage. DL and DRA consider a statistical three factor model to describe the yield curve over time. It can be argued that the three factors represent the level, slope and curvature of the yield curve and thus carry some economical interpretation. Moreover, 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 the 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 by OLS 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 introducing time-varying factor loadings and time-varying volatility. First, we look at the estimate of the parameter that defines the factor loadings in various subsamples and find strong evidence that it is not constant over time. Moreover, when we introduce a simple step function for this parameter we already find a highly significant improvement of model fit. Therefore, to allow the factor loadings to change gradually over time we estimate a cubic spline for their parameter.

Second, graphs of the data provide some evidence that the volatility in interest rate series is not constant over time. Interestingly, in high volatility periods the yields at all maturities are very volatile. However, some maturities are more volatile than others. To incorporate both of these observations in the model we also introduce a flexible cubic spline function for the volatility. This spline scales all volatilities over time, and is interacted with a constant level of volatility for each maturity. Similar to the introduction of time-varying factor loadings, time-varying volatility also significantly increases the fit of the model.

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 illustrated in this context. Besides the standard unsmoothed Fama-Bliss monthly yields dataset for the period 1972-2000 (as also used by DRA), 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 SSF framework thus allows 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, while we model the volatility for each of the yields individually in the measurement equation. Moreover, they only look at UK data and their approach is numerically intensive as it is estimated in a Bayesian framework. Yu and Zivot (2007) extend the NS model to include corporate bonds. More statistical in nature, Bowsher and Meeks (2006) introduce a 5-factor model that uses splines to model the yield curve and has the knots of these splines as factors. While this allows for a more flexible yield curve some economic intuition of the factors is lost. Moreover, volatility is kept fixed over time.

The rest of the paper is build up as follows. Section 2 describes the baseline Nelson-Siegel latent factor model, Section 3 the extensions we propose. In Section 4 we describe our dataset and provide estimates of the various models. Section 5 provides an illustration with missing values, 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. We will focus on the model as slightly adjusted in terms of factorization by Diebold and Li (2006). First we will introduce this model, after which we will go into the model in state space form 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

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

where $\beta_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})'$, for given time t, maturity τ and fixed coefficient λ that determines the exponentially 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 the 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

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

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

Plotting the three series of regression estimates for β_t shows that these series are 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 modeled 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 variance matrix Σ_β is chosen such that $\Sigma_\beta - \Phi\Sigma_\beta\Phi' = \Sigma_\eta$.[1] Equations (3) and (4) with the initial condition for β_1 constitute a special case of the general state space model. The 3×1 mean vector μ, the 3×3 autoregressive coefficient matrix Φ, the 3×3 variance matrix Σ_η and the $N \times N$ variance matrix Σ_ε are unknown and need to be estimated.

To estimate the model we maximize the likelihood function from the above set of equations. For models in state space this can be done using the output from the Kalman Filter, our initializations are the same as DRA. Our estimations are done in Ox (see Doornik (2001)), we use the functions of the SsfPack (see Koopman, Shephard, and Doornik (1999)). For more on state space models see a.o. Durbin and Koopman (2001).

The general state space framework allows for other, more general, dynamic processes for β_t. Diebold, Rudebusch, and Aruoba (2006) assume that Σ_ε is diagonal, which implies that the disturbances for yields of different maturities are uncorrelated for a given time t. This assumption is often used but is also convenient for computational tractability during estimation given the potentially large number of yields available. The variance matrix Σ_η is assumed non-diagonal.

The Nelson-Siegel model with β_t modeled as the vector autoregressive process (4) has been developed for the forecasting of yield curves, see Diebold, Rudebusch, and Aruoba (2006). They conclude that forecasting results have not improved considerably compared to the two-step approach of Diebold and Li (2006). This is confirmed by Yu and Zivot (2007) although they distinguish different forecasting performances for short-term and long-term maturities.

3 Time-Varying Factor Loadings and Volatility

In the above section we have introduced the latent factor model for the yield curve as put forward by Nelson and Siegel (1987) and proposed to be estimated in state space form by Diebold, Rudebusch, and Aruoba (2006). Now we are ready to state the two extensions to this model that we propose. The first is the introduction of time-varying factor loadings,

[1] This enforces stationarity of the vector autoregressive process, and follows from Ansley and Kohn (1986). By doing so we slightly deviate from DRA. However, our results do not depend on this.

the second is introducing time-varying volatility.

3.1 Extension 1: Time-Varying Factor Loadings

In the latent factor yield curve model the parameter λ decides the shape of the factor loadings. 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 and following this Yu and Zivot (2007) fix it at 0.077. 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 $\lambda = 0.077$).

As is clear from the previous paragraph, the value at which λ is typically fixed is not the same across studies, and differs from values estimated from the data. With the model in state space form it is straightforward to estimate λ in addition to the other parameters. However, fixing λ over the full sample period restricts the factor loadings to be constant over time. Though this does not mean the shape of the yield curve can not change (as it also depends on the factors themselves) it does fix the maturity at which the curvature factor is maximized and the speed of decay of the slope parameter. Therefore the λ is informative about the shape of the yield curve. Given these arguments, we study the role of λ further and we also consider possible changes in λ over time.

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

$$\lambda_t = f(t; \lambda^*),$$

where $f()$ is a function with time-index t as argument and λ^* is a $k \times 1$ vector of coefficients. Two examples of specifications for $f()$ are (i) a step function as represented by $\lambda_t = \ell_t \lambda^*$ where ℓ_t is a particular row of the $k \times k$ identity matrix and (ii) a cubic spline function as in Poirier (1976) that can be represented by $\lambda_t = w_t' \lambda^*$ where w_t is the $k \times 1$ vector of interpolating weights (determined by certain smoothing conditions). In our empirical results we will try both of these forms and test whether there is a significant improvement in model fit.

3.2 Extension 2: Time-Varying Volatility

Another aspect of the analysis of the term structure is the recognition that the data is the result of trading, where the clustering of large shocks over time is often found. Therefore it is assumed that the volatility in the time series is not constant over time, see the seminal works of Engle (1982) and Bollerslev (1986). We also observe volatility changes in interest rate series of different maturities but they are more longer-term and are changing slowly. We therefore will account for this by implementing a model, in which the volatilities of the yields at different maturities are each driven by two components. The first component v_t is the time-varying component. This component is equal for all the maturities, because the yields at all maturities show high volatilities in the same periods. It is interacted with an individual constant scaling parameter per maturity α_i^2. The second component σ_i^2 is a constant, which measures the constant overall individual volatility of yield with maturity i, for $i = 1, \ldots, N$.

Each of the diagonal elements of Σ_ε, that is, $h_{i,i}$ represents the variance of the series of yields at one of the N maturities. The diagonal elements are modeled by

$$h_{i,i} = \sigma_i^2 + v_t \alpha_i^2, \tag{5}$$

where the values of v_t are determined by a smoothing spline function.

4 Data and Empirical Findings

For our main analyses we use the exact same dataset as used in 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.

This dataset is the exact same as used by Diebold, Rudebusch, and Aruoba (2006)[2], 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 for the level, slope and curvature coefficients. These proxies are chosen conform the construction of the factors from the Nelson and Siegel (1987) model, see the discussion in Section 2.1, and have previously been used a.o. 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. Important for econometric analyses, yields for all maturities are very persistent. The persistence is most notable for long term bonds, but with a first-order autocorrelation of 0.970 the 3 month bill is still highly persistent. Also the level, slope and curvature proxies are persistent. The curvature and slope are least persistent, with twelfth-order autocorrelation coefficients of 0.259 and 0.410 respectively.

[insert Figure 1]

Figure 1 shows the cross-section of yields we study over time. In addition to the findings of Table 1 we see a few interesting characteristics. The first thing to note is that the yields vary significantly over time, with a large common component across all yields. 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 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

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

with both of our model extensions jointly included. 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 Nelson-Siegel Latent Factor Yield Curve Model

[insert Table 2]

Table 2 shows 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 shows the measurement error. Panel A of this table focuses on the prediction errors, while Panel B looks at the measurement errors. The measurement errors are defined as the actual yields minus the yields that are obtained using the estimated parameters and smoothed 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. Looking at the standard deviation we see that for prediction the long bonds are best predicted. The standard errors of the measurement errors however indicate that especially the shortest and longest bonds are difficult to be fit. Overall, the medium term notes of around 24 months are best fit by the model.

4.2.2 Time-Varying Factor Loadings

We are now ready to look at the first extension we propose: time-varying factor loadings. Before we discuss the results of estimating the model of Section 3.1 we will take an intermediate step, and estimate the baseline model for three subperiods.

[insert Table 4]

Table 4 shows the estimated factor loadings parameter λ for various models, together with LR-tests to compare them. If we divide the full sample period in four equally spaces subsamples we see the full sample λ estimate of 0.778 breaks down into 0.0397, 0.126, 0.0602 and 0.0695. With their small standard errors they again confirm our earlier observation that interest rates can be informative about λ. Moreover, these estimates confirm how the assumption of constant factor loadings is contradicted by the data.

The next step is to estimate the factor loadings parameter λ with a step function, as defined in Section 3.1. As all other parameters are assumed to be constant over the sample period the estimates differ from the estimates obtained for each subsample. However, again we find low standard errors and high variation for λ over time. The LR-test shows significant improvement in model fit over the baseline model with constant λ, even at a 1% level.

[insert Figure 2]

Finally, we estimate the factor loadings parameter λ in its most flexible form, with the spline function as defined in Section 3.1. For the knots we choose to divide the interval in four equally spaced intervals. Besides knots at the begin- and end-point of the sample we put them at April 1979, July 1986 and October 1993.[3] Figure 2 shows the estimated λ over time. From this figure it is clear that even within each subperiod λ is not constant over time. In Table 4 we show the mean λ for each of the periods. Though on average the mean λ for these periods varies less than the models with subsamples and the step function there still is considerable variation for λ. The LR test shows that the model with a spline for the factor loadings parameter improves the fit significantly compared to both the constant λ model and the model with step function.

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)

[3] We have tried various different number of knots, see Figure **??** 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.

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 extension we propose in Section 3 is to make volatility time-varying. Similar to our approach to make the factor loadings time-varying we introduce a spline to this end, see Section 3.2 for details. The knots are set to equal those of Section 4.2.2.[4]

[insert Figure 3]

Figure 3A shows the spline that represents the time-varying common volatility component. On itself its values represent how the average overall volatility varies over time. When combined with the factor loadings for each individual maturity, it will provide an estimate of volatility for each maturity and each period. Figure 3B shows the resulting volatility for some maturities. Interestingly we find that volatility is especially high in the period 1980 until 1987, but thereafter 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. Interestingly, where the improvement with a time-varying λ was mainly in short maturities for the model with time-varying volatility this is mostly 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 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 was the case when comparing the baseline model with the time-varying factor loadings model. This indicates that most gain in describing the yield curve can be gained by introducing time-varying volatility.

[4]See Footnote 3 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 **??** in the Appendix (available from the authors on request).

4.2.4 Both Factor Loadings and Volatility Time-Varying

We are now ready to estimate the Nelson-Siegel latent factor 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 3 we do not repeat these here.

In Table 3 the measurement and prediction error of this model are given. Compared to the baseline Nelson-Siegel model we see the prediction (measurement) error is lower for 11 (11) out of the 17 maturities. Looking at the standard deviation we see the improvement is for 13 out of the 17 maturities for the prediction errors, but for only 6 for the measurement errors.

Looking at the loglikelihood and Akaike Information Criterion (AIC), reported in Table 5, we see a highly significant improvement compared to the baseline model. Also when benchmarked against the model with only time-varying factor loadings or time-varying volatility we see the improvement is significant. We can 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 4]

In Figure 4 we compare the latent factors obtained from the model 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.

[insert Figure 5]

Finally, in Figure 5 we re-visit 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 model and our extended SSF model with both the factor loadings and volatility time-varying. Especially in the August 1998 it is clear our model extensions allow for more flexibility and improve the model's fit.

5 An Illustration with Missing Values

An attractive feature of models in state space form is that they 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.

With the Kalman Smoother[5] the smoothed latent factors can be obtained for all periods based on all available data. The estimation procedure does not change depending on data availability. Moreover, with these smoothed factors a smoothed value for all maturities can be obtained. Thus, for periods when data is missing for a certain maturity data from other maturities can be used to obtain an estimate.

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 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 big advantage of being freely available and covering a long time horizon comes the disadvantage that there are many missing values. 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 6]

We have estimated the Nelson-Siegel latent factor model with both time-varying factor loadings and time-varying volatility for this dataset. Figure 6 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 get a feeling for how reliable this estimate is we compare it to the value we have from our dataset of Section 4.1. As that dataset does not contain the 30 year bond we can only do this for the 3 month and 24 month yields. We see that in both cases the smoothed yield

[5] See Section 2.1 for our discussion of state space models, and Durbin and Koopman (2001) for more on state space models in general.

obtained provides a reliable estimate of the missing data.

6 Conclusion

We propose two extensions to the Nelson and Siegel (1987, NS) latent factor yield curve model as suggested to be estimated in State Space Form (SSF) by Diebold, Rudebusch, and Aruoba (2006). We first look at the factor loadings parameter of the latent factors, which in the literature is either fixed at some value or estimated constant over time. Our empirical results show that this parameter is estimated with low standard error, consistent with the data being highly informative about the parameter that decides the shape of the factor loadings. We propose this parameter to be time-varying in two ways. We first introduce a simple step function for the parameter that governs the factor loadings. Secondly, we introduce a spline function to model the factor loading parameter. We show that these extensions provide a highly significant improvement in model fit.

Next we turn our attention to the volatility for each of the maturities. Like many financial data, there are signs of volatility clustering and commonality for the yields. We propose to take this into account by introducing a common volatility component, modeled by a spline function. This common volatility component is then loaded onto each of the maturities by a constant individual loading parameter. Also this extension provides a significant improvement in model fit.

In addition we illustrate how easily the NS model in SSF deals with missing values, such as the four years in which the 30 year bond was not issued. For a dataset with many missing values we show how data both from other periods of the same maturity and data from other maturities is used to obtain an estimate of the missing value. Comparing the obtained estimate to the value in our main dataset we see the estimate is very accurate.

References

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.

Bollerslev, T. (1986). Generalised autoregressive conditional heteroskedasticity. *J. Econometrics 51*, 307–327.

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

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.

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

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. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom inflation. *Econometrica 50*, 987–1007.

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.

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_ε.

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

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 *Subsamples* model corresponds to the baseline Nelson-Siegel latent factor model as estimated for each of the four subsamples individually. 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) Subsamples	(c) Step Function	(d) Splines#
Full Sample	0.0778** 0.00209			
01/72 - 03/79		0.0397** 0.00257	0.0932** 0.00533	0.0938
04/79 - 06/86		0.126** 0.00603	0.116** 0.00393	0.116
07/86 - 09/93		0.0602** 0.00202	0.0638** 0.00240	0.0698
10/93 - 12/00		0.0695** 0.00238	0.0717** 0.00419	0.0861
Loglikelihood	3185.4	4524.2	3259.0	3289.0
AIC	-6304.8	-8784.4	-6446.1	-6503.9
LR-test (a) vs. (c)		147.2** 0.000		
LR-test (a) vs. (d)		207.2** 0.000		
LR-test (c) vs. (d)		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 λ 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.

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.

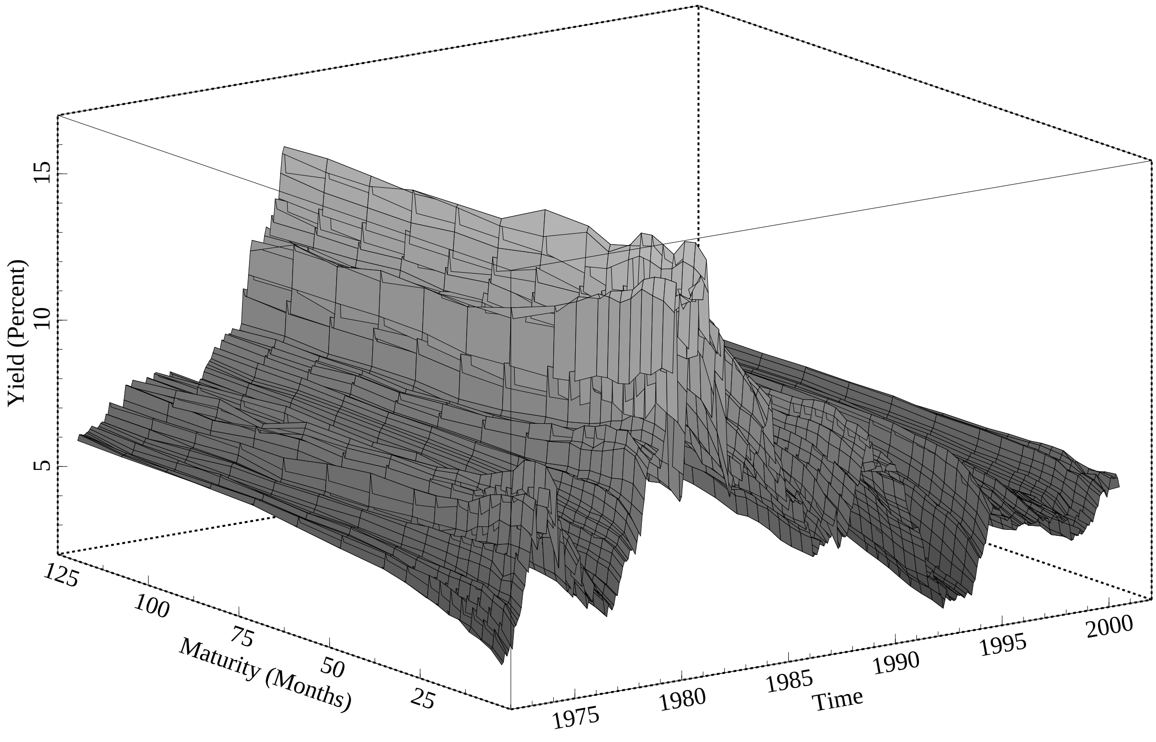


Figure 2: Time-Varying Factor Loadings Parameter

In this figure we show 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.

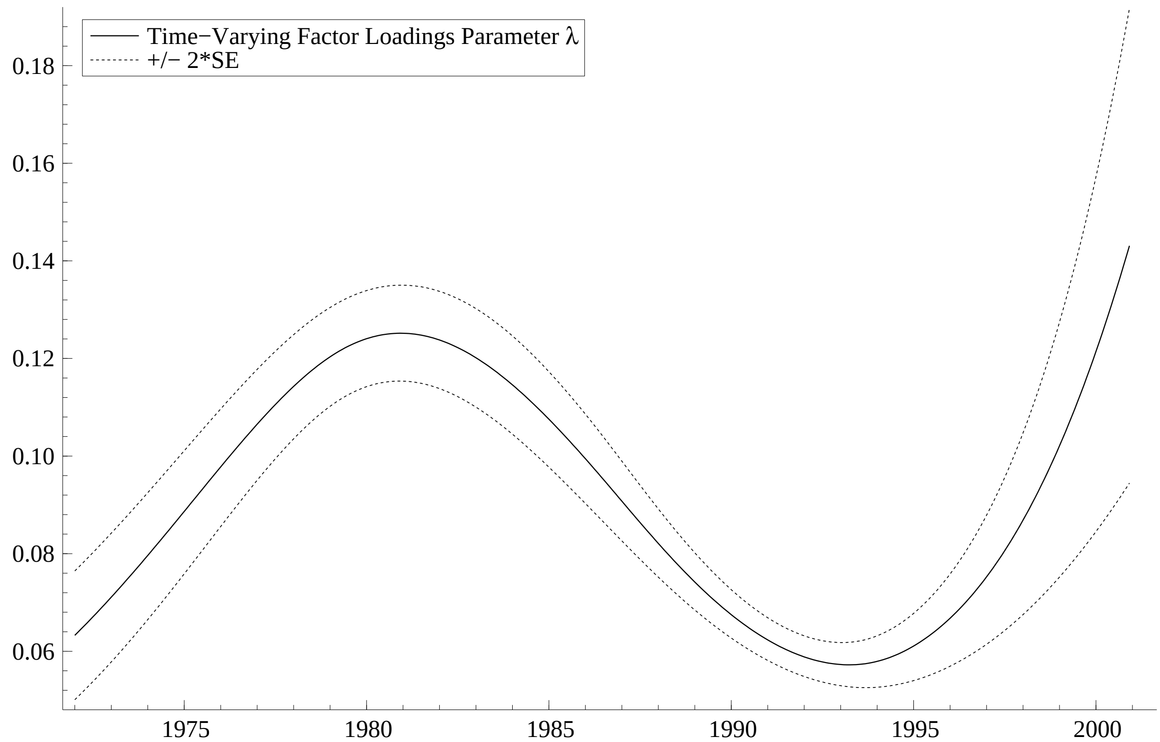
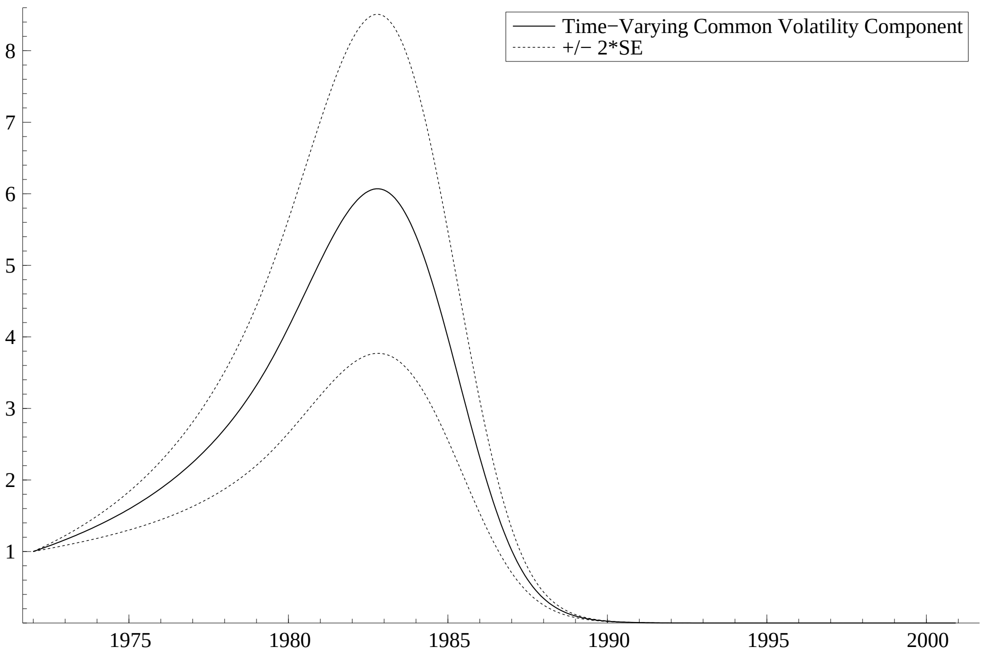


Figure 3: Time-Varying Volatility

In this figure we show 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

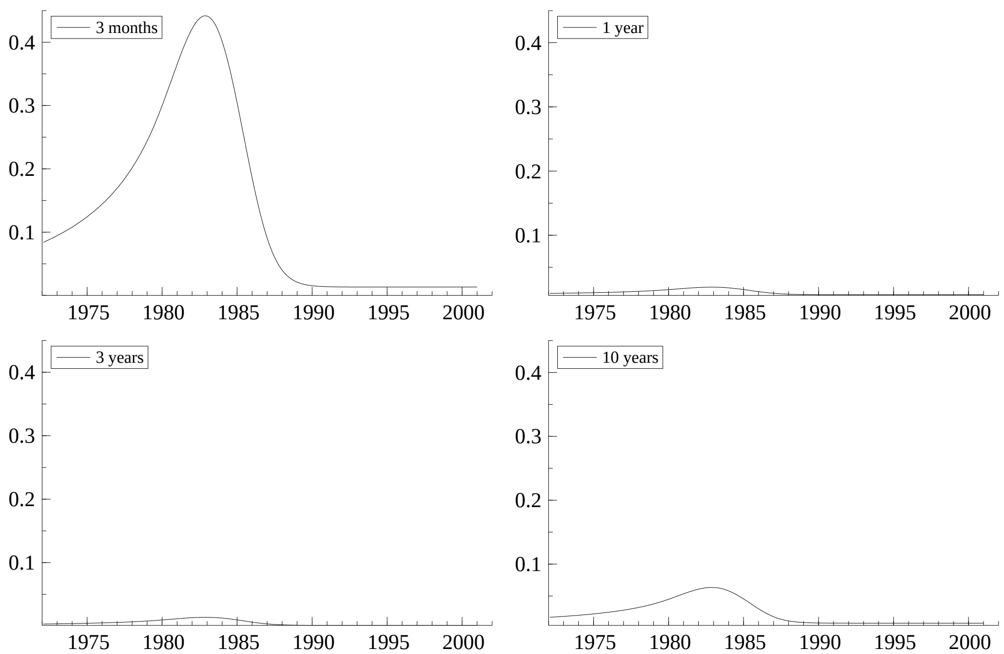
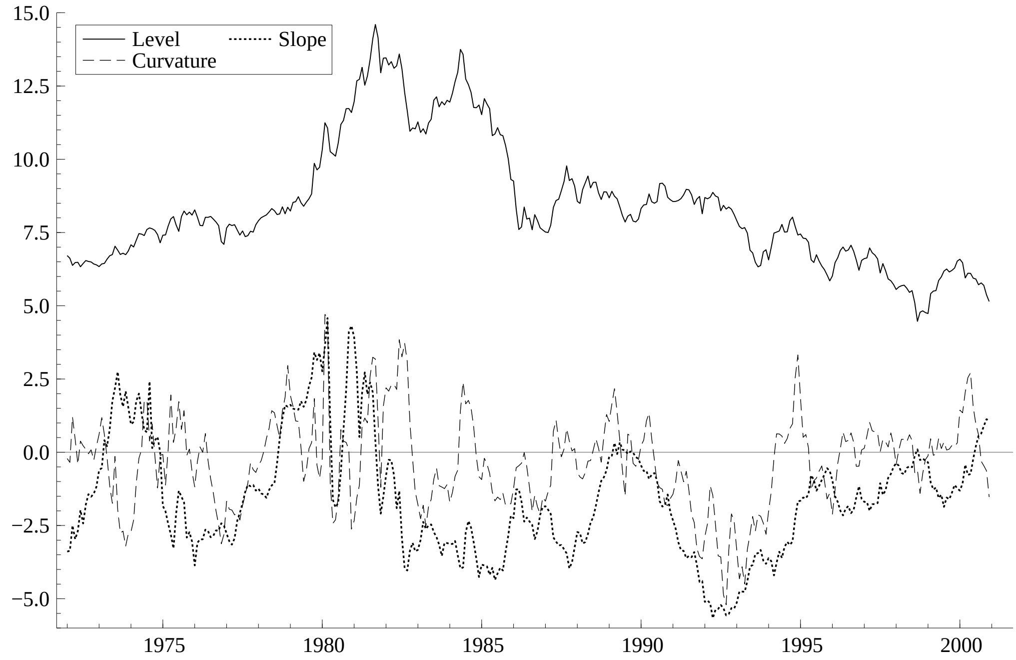


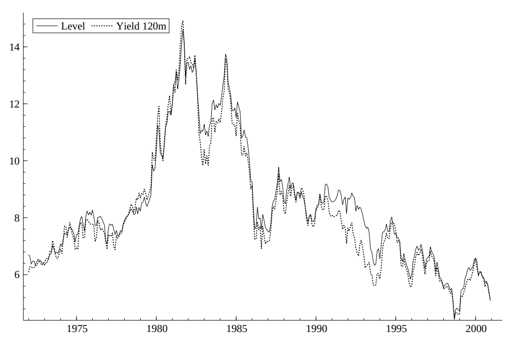
Figure 4: Time-Varying Model - Level, Slope and Curvature

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.

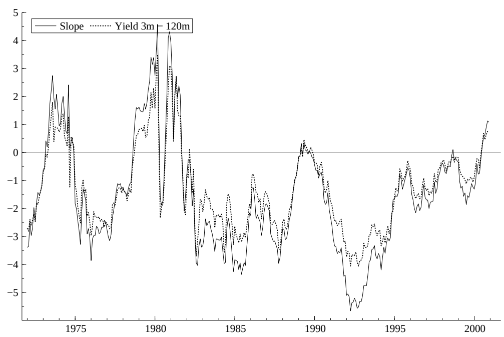
(A) Level, Slope and Curvature



(B) Level, with Proxy from Data



(C) Slope, with Proxy from Data



(D) Curvature, with Proxy from Data

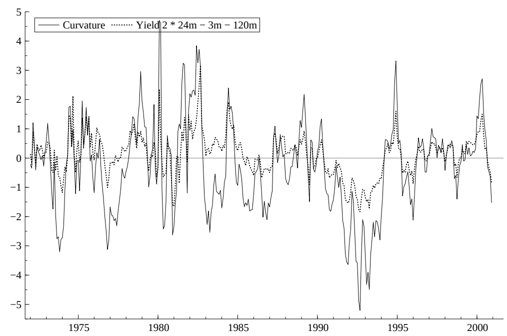


Figure 5: Fitted Yield Curves for Four Months

This figure shows 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.

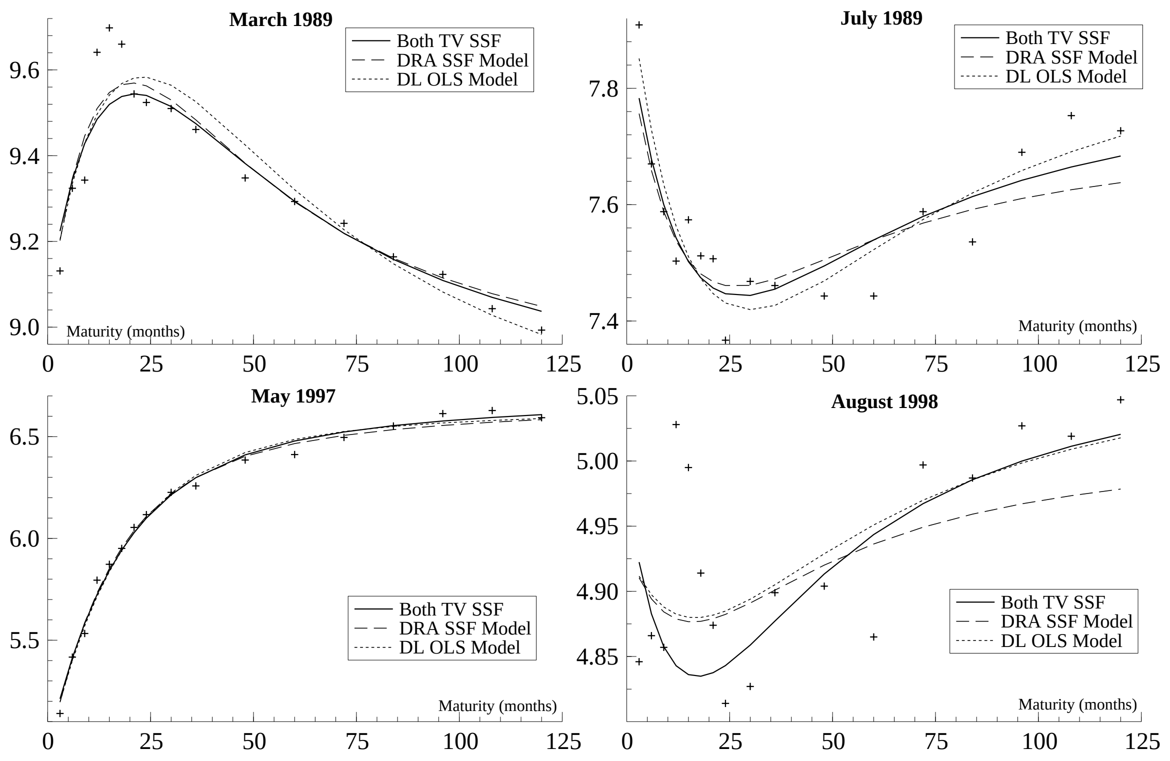


Figure 6: An Illustration with Missing Values

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.