UNIVERSITEIT
AMSTERDAM

VU Research Portal

Forecasting Interest Rates with Shifting Endpoints

van Dijk, D.; Koopman, S.J.; van der Wel, M.; Wright, J.H.

2012

document version
Early version, also known as pre-print

Link to publication in VU Research Portal

Forecasting Interest Rates with Shifting Endpoints

Dick van Dijk[1,2]
Siem–Jan Koopman[3,2]
Michel van der Wel[1,2,5]
Jonathan Wright[4]

[1] Econometric Institute, Erasmus University Rotterdam;
[2] Tinbergen Institute;
[3] VU University Amsterdam;
[4] Johns Hopkins University;
[5] CREATES, Aarhus.

Forecasting Interest Rates with Shifting Endpoints

Dick van Dijk$^{(a,b)}$, Siem Jan Koopman$^{(b,c)}$,

Michel van der Wel$^{(a,b,d)}$ & Jonathan H. Wright$^{(e)*}$

$^{(a)}$ Econometric Institute, Erasmus School of Economics, NL 3000 DR Rotterdam

$^{(b)}$ Tinbergen Institute, Amsterdam & Rotterdam

$^{(c)}$ Department of Econometrics, VU University, NL 1081 HV Amsterdam

$^{(d)}$ CREATES, Aarhus

$^{(e)}$ Department of Economics, Johns Hopkins University, Baltimore MD 21218

July 2012

Abstract

Many economic studies on inflation forecasting have found favorable results when inflation is modeled as a stationary process around a slowly time-varying trend. In contrast, the existing studies on interest rate forecasting either treat yields as being stationary, without any shifting endpoints, or treat yields as a random walk process. In this study we consider the problem of forecasting the term structure of interest rates with the assumption that the yield curve is driven by factors that are stationary around a time-varying trend. We compare alternative ways of modeling the time-varying trend. We find that allowing for shifting endpoints in yield curve factors can provide gains in the out-of-sample predictive accuracy, relative to stationary and random walk benchmarks. The results are both economically and statistically significant.

---

*Email: djvandijk@ese.eur.nl; s.j.koopman@vu.nl; vanderwel@ese.eur.nl; wrightj@jhu.edu. Michel van der Wel is grateful to Netherlands Organisation for Scientific Research (NWO) for a Veni grant; and acknowledges support from CREATES, funded by the Danish National Research Foundation.

Corresponding author: Dick van Dijk, Econometric Institute, Erasmus University Rotterdam, Burg. Oudlaan 50, NL-3062 PA Rotterdam, The Netherlands. Email: djvandijk@ese.eur.nl; Telephone: +31-10-4081263; Fax: +31-10-4089162

1   Introduction

Forecasting government bond yields is a topic of great practical importance, both to investors and to monetary policymakers, who wish to decompose yields into expectations of future interest rates and risk premia. However, yield forecasting is a challenging task. Duffee (2002) has shown that widely-used affine term structure models actually produce less precise out-of-sample forecasts compared to forecasts of a simple random walk model with a "no change" prediction.

It is well known that the cross-section of yields can be accurately approximated by three factors, representing the level, slope and curvature of the yield curve, see Litterman and Scheinkman (1991). Accordingly, a parsimonious factor structure is a key feature of many term structure models, including no-arbitrage affine models (see, e.g., Duffie and Kan (1996) and Dai and Singleton (2000)), principal-components based methods (Duffee, 2011), spline-based models (Bowsher and Meeks, 2008) and the Nelson-Siegel model (Nelson and Siegel, 1987). The last approach has recently gained popularity, especially in the context of forecasting. This is mostly due to Diebold and Li (2006), who extend the Nelson-Siegel model to a forecasting device by combining the factor representation of the yields with autoregressive specifications for the dynamics of the three factors.

In the so-called *dynamic Nelson-Siegel* approach of Diebold and Li (2006), the autoregressions are assumed to be stationary processes with constant unconditional mean. Given the historical behavior of interest rates this assumption may be inappropriate, in particular for the level factor. Indeed, Duffee (2011) obtains superior out-of-sample forecasts when a random walk is used for the first principal components of yields (which closely corresponds to the Nelson-Siegel level factor).

In this paper we also investigate whether allowing for nonstationarity in the Nelson-Siegel factor dynamics is useful for forecasting interest rates. But, instead of imposing a random walk structure as in Duffee (2011), we consider autoregressive specifications with a time-varying unconditional mean or a *shifting endpoint*. This

idea is adopted from the strand of the economic forecasting literature that argues that certain macroeconomic variables, notably inflation, should be modeled as the sum of a transitory component and a slowly-varying trend or permanent component; see, among others, Kozicki and Tinsley (2001, 2005), Cogley and Sargent (2005), Stock and Watson (2007, 2010), Clark (2011), Orphanides and Wei (2012), Wright (2012) and Faust and Wright (2012). This decomposition may, in particular, reflect time-variation in the central bank's implicit inflation target, which was high around 1980 but then declined steadily over the next two decades. Among others, Stock and Watson (2010) advocate forecasting inflation in "gap" form, as deviations from the trend. In a similar fashion, Cieslak and Povala (2010) explore the decomposition of yields into a long-term "expected inflation" component and a transitory "risk-premium" component for the forecasting of annual bond returns. The two components partly depend on macroeconomic and financial variables.

We argue that the close relation of the term structure of interest rates with macroeconomic variables such as inflation makes it plausible to consider the possibility that the factors driving the term structure of nominal interest rates, especially the level factor, have some permanent component as well. Accordingly, we adapt the dynamic Nelson-Siegel framework by allowing some or all of the factors to exhibit time-varying means or shifting endpoints. A key issue in this approach is of course the specification of the shifting endpoints. We consider three possibilities, where shifts in the permanent component of the yield factors are captured using either (i) time series methods (exponential smoothing) or, (ii) long-range survey forecasts of either interest rates or macroeconomic variables such as inflation and output growth, or (iii) exponentially smoothed realizations of these macro variables.

We focus on the impact of allowing for a permanent component in the Nelson-Siegel factors on out-of-sample forecasting for the monthly U.S. Treasury yield curve over the period 1994-2009. We compare the accuracy of forecasts obtained from the dynamic shifting endpoint Nelson-Siegel model to forecasts from the 'standard'

dynamic Nelson-Siegel approach and to random walk forecasts. In a similar way that allowing for shifting endpoints gives large gains in the accuracy of forecasting inflation, we find that it also gives substantial improvements in the out-of-sample accuracy of interest rate forecasts. The improvements are both economically and statistically significant. The forecast improvements are largest for long-maturity interest rates and for long-horizon forecasts. For example, in forecasting ten-year yields at the one-year-ahead horizon, the best "shifting endpoints" forecast reduces the out-of-sample root mean square prediction error by 28 percent and 23 percent relative to the Diebold-Li and random walk forecasts, respectively. At the two-year horizon, the gains are even larger. The best forecasts are those obtained from shifting endpoint methods that relate the level factor in the Nelson-Siegel model to the trend component of inflation. This trend can be taken from Blue Chip survey forecasts of inflation or from exponential smoothing of realized inflation.

Our analysis relates to three strands of literature. First, a few papers have applied the shifting endpoints approach to the term structure of interest rates, in particular Kozicki and Tinsley (2001) and Orphanides and Wei (2012). The focus in these papers is on describing the yield curve *in-sample*, and they are set in a no-arbitrage affine framework. Second, a growing literature considers the relation between macroeconomic variables and the term structure of interest rates. Duffee (2012) provides a recent overview, notable contributions include Ang and Piazzesi (2003), Diebold et al. (2006) and Joslin et al. (2011). In the context of yield curve forecasting with Nelson-Siegel type models, De Pooter et al. (2010) and Exterkate et al. (2012) document that macro factors contain valuable predictive information. Third, several recent papers use survey data in analysis of the term structure. Survey forecasts are used in term structure models by, among others, Piazzesi and Schneider (2011), Chun (2011), Ehling et al. (2012) and Kim and Orphanides (2012). Chun (2012) compares the individual survey forecasts themselves with models for the yield curve. We contribute to all three strands of literature with our explicit focus

on forecasting the yield curve using survey data. In our analysis we start from the Nelson-Siegel setting (as opposed to arbitrage-free settings) as this has been documented to provide favorable forecasting performance. Further, we consider modeling shifting endpoints using exponential smoothing of interest rates, long-horizon surveys of interest rates, long-horizon surveys of inflation and real GDP growth, and exponential smoothing of inflation and growth.

The plan for the remainder of this paper is as follows. In Section 2, we briefly summarize the dynamic Nelson-Siegel method as advocated by Diebold and Li (2006). In Section 3 we introduce our extension of this framework by allowing for shifting endpoints of the yield factors. This includes a description of the different approaches we adopt to capture time-variation in the unconditional means. In Section 4 we present the empirical findings for our dataset of U.S. Treasury yields. We conclude in Section 5.

2   The dynamic Nelson-Siegel model

Our point of departure for modeling and forecasting the yield curve is the dynamic Nelson-Siegel model, as popularized by Diebold and Li (2006). In Section 2.1 we discuss the model specification, and in Section 2.2 we apply the model to a panel of U.S. government bond yields.

2.1   The dynamic Nelson-Siegel set-up

Let $y_t(\tau)$ denote the continuously compounded yield to maturity on a zero coupon bond with maturity of τ periods at time t. Following Nelson and Siegel (1987) and Diebold and Li (2006), we consider the following three factor model for the yield curve

$$y_t(\tau) = \beta_{1t} + \beta_{2t} \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} \right) + \beta_{3t} \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau} \right) + \varepsilon_t(\tau). \qquad (1)$$

As argued by Diebold and Li (2006), β_{1t}, β_{2t} and β_{3t} may be interpreted as latent dynamic factors. The loading on β_{1t} does not depend on the maturity τ, such that it represents a level factor.[1] The loading on β_{2t} starts at 1 for $\tau = 0$ and declines monotonically to zero as the maturity increases; thus, it may be considered as slope factor.[2] Finally, the loading on β_{3t} is equal to zero at maturities zero and infinity and positive in between; hence, it may be viewed as a curvature factor. The parameter λ_t determines both how fast the loading on β_{2t} decays to zero and the maturity at which the loading on β_{3t} achieves its maximum value. In our empirical application, we follow Diebold and Li (2006) and assume that λ_t is constant and set it equal to 0.0609.[3] Finally, the disturbances $\varepsilon_t(\tau)$ in (1) represent measurement error, and as such are assumed to have mean zero, a variance σ_t^2 and to be independent over time and across maturities.

Diebold and Li (2006) interpret the Nelson-Siegel yield curve in (1) as the measurement equation that details the relation between the observed yields and factors to be estimated. The model is completed by specifications of the dynamics of the level, slope and curvature factors. Specifically, for this purpose Diebold and Li (2006) suggest separate univariate first-order autoregressive processes, given by

$$\beta_{j,t+1} = \mu_j + \phi_j \left(\beta_{jt} - \mu_j \right) + \eta_{j,t+1}, \tag{2}$$

for $j = 1, 2, 3$, where the disturbances $\eta_{j,t+1}$ have zero mean and variance σ_j^2 and are assumed to be mutually and serially independent at all time periods.

The dynamic Nelson-Siegel model offers a straightforward approach to obtain forecasts of future yields. First, we generate forecasts of the factors $\beta_{j,t+h}$ by iterating

---

[1]Alternatively, β_{1t} may be considered as a long-term factor as the loadings on β_{2t} and β_{3t} converge to zero as τ increases.

[2]This also follows by observing that at maturity zero the yield is given by $\beta_{1t} + \beta_{2t}$ (plus the noise term ε_t), such that β_{2t} measures the difference between the instantaneous and the long-term yields.

[3]This implies that the loading on β_{3t} achieves its maximum value at a maturity τ of 30 months. Alternatively one could estimate it along with the parameters, or treat it as time-varying as in Koopman et al. (2010). To focus on our time-varying specifications for the trend we treat it as constant.

equation (2) for the desired forecast horizon of h periods.[4] Second, we obtain multi-step forecasts of the interest rates via equation (1). This procedure turns out to produce accurate yield forecasts in various studies (see, e.g., Diebold and Li (2006), De Pooter (2007) and Exterkate et al. (2012)) and has quickly become an important benchmark in the yield-curve forecasting literature.

2.2 The dynamic Nelson-Siegel model and the U.S. government bond yield curve

We apply the dynamic Nelson-Siegel model to a panel of unsmoothed Fama and Bliss (1987) U.S. government bond yields at the monthly frequency for the period January 1970 through December 2009. The data set is constructed by applying the Fama-Bliss algorithm to end-of-month CRSP data on prices of individual Treasury bonds. Table 1 present some descriptive statistics for the resulting yields for maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108 and 120 months. These reveal the usual stylized facts of the yield curve, in particular that (i) on average the yield curve is upward sloping and concave; (ii) yields are persistent, with long-term yields showing slightly larger autocorrelations than short-term yields; and (iii) the volatility of yields declines with maturity. Also included in Table 1 are statistics for empirical proxies for the level, slope and curvature of the yield curve,[5] showing that the persistence of the level is higher than the persistence of the slope, which in turn is more persistent than the curvature.

---

[4]Thus, we use so-called iterated forecasts. An alternative would be to consider 'direct' forecasts, by estimating separate autoregressive models for each forecast horizon

$$\beta_{j,t+h} = \mu_{jh} + \phi_{jh} \left(\beta_{jt} - \mu_j h \right) + \eta_{j,t+h},$$

see Diebold and Li (2006). We consider iterated forecasts as time-varying specifications for the trend are introduced more naturally into this setting. For direct forecasts one would estimate a different specification for the trend for each forecasting horizon; this is somewhat more complicated but nevertheless feasible. To enable fair comparison of the models in this paper, we consider iterated forecasts for all models.

[5]The proxy for level is the longest maturity yield (120 months), for slope it is the longest (120 months) minus the shortest (3 months) maturities, and for curvature it is two times the 24 month yield minus the sum of the 3 month and 120 month yields.

Table 1: Descriptive Statistics

We report descriptive statistics for U.S. Treasury yields over the period 1970-2009, based on monthly data, constructed using the unsmoothed Fama-Bliss method. The maturity is measured in months. For each maturity we present mean, standard deviation (St.d), minimum, maximum and the j-th order autocorrelation coefficients $\hat{\rho}(j)$ for $j = 1$, 12 and 30. We also include statistics for empirical proxies for the level, slope and curvature of the yield curve. The proxy for level is the longest maturity yield (120 months), for slope it is the 120 month yield minus the 3 month yield, and for curvature it is two times the 24 month yield minus the 3 month and 120 month yields.

Maturity	Mean	St.d	Min	Max	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(30)$
3	5.766	3.068	0.041	16.019	0.979	0.749	0.411
6	5.969	3.095	0.150	16.481	0.980	0.763	0.442
9	6.083	3.086	0.193	16.394	0.981	0.771	0.467
12	6.166	3.050	0.245	16.101	0.981	0.777	0.483
15	6.253	3.026	0.377	16.055	0.982	0.785	0.504
18	6.324	3.006	0.438	16.219	0.983	0.792	0.522
21	6.387	2.987	0.532	16.173	0.983	0.797	0.537
24	6.418	2.940	0.532	15.814	0.983	0.799	0.550
30	6.512	2.875	0.819	15.429	0.983	0.808	0.570
36	6.600	2.829	0.978	15.538	0.984	0.814	0.586
48	6.756	2.752	1.019	15.599	0.984	0.822	0.614
60	6.852	2.668	1.556	15.129	0.985	0.832	0.636
72	6.964	2.636	1.525	15.108	0.987	0.842	0.653
84	7.026	2.570	2.179	15.024	0.987	0.841	0.666
96	7.069	2.534	2.105	15.052	0.988	0.850	0.673
108	7.095	2.517	2.152	15.114	0.988	0.853	0.677
120(level)	7.067	2.462	2.679	15.194	0.988	0.843	0.674
slope	1.301	1.361	-3.191	3.954	0.934	0.418	-0.123
curvature	0.003	0.862	-2.174	2.905	0.877	0.441	0.130

Figure 1 plots the time series of a subset of the yields in our panel. Yields of all maturities trended upwards over the 1970s and downwards since about 1980, in line with shifts in inflation and long-run inflation expectations. It is these low-frequency patterns that we are attempting to take into account with our shifting-endpoint specification.

Figure 1: U.S. Treasury Yields from January 1970 up to December 2009
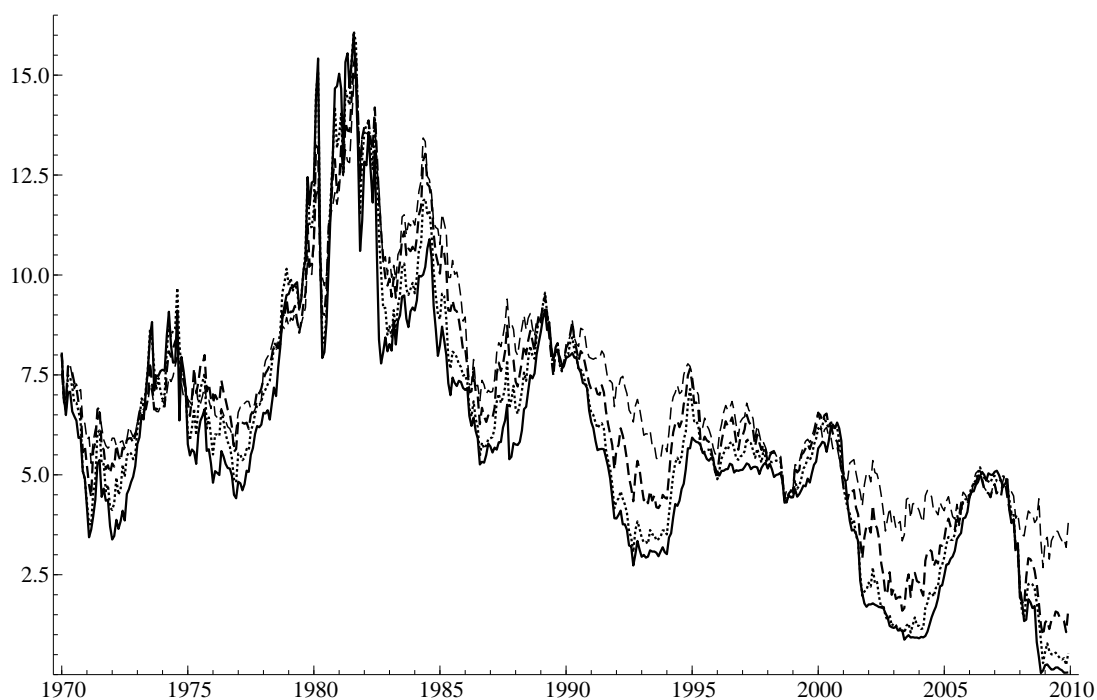
We present monthly U.S. Treasury yields over the period January 1970 - December 2009, constructed using the unsmoothed Fama-Bliss method. The maturities we show are 3 (solid), 12 (dotted), 60 (slash thick) and 120 (slash thin) months.

Given that we fix the parameter λ_t, we can obtain estimates of the factors β_{1t}, β_{2t} and β_{3t} by fitting the model (1) to the cross-section of yields for each month t

using ordinary least squares, that is,

$$
\begin{pmatrix} \hat{\beta}_{1t} \\ \hat{\beta}_{2t} \\ \hat{\beta}_{3t} \end{pmatrix} = \left(\sum_{i=1}^{N} x_i x_i' \right)^{-1} \sum_{i=1}^{N} x_i y_t(i), \qquad \text{where } x_i = \begin{pmatrix} 1 \\ \left[1 - \exp(-\lambda \tau_i) \right] / \lambda \tau_i \\ \left[1 - \exp(-\lambda \tau_i) \right] / \lambda \tau_i - \exp(-\lambda \tau_i) \end{pmatrix},
$$

for the N available yields with maturities τ_i, $i = 1, \ldots, N$ and $y_t(i)$ the yield of maturity τ_i at time t. Figure 2 presents the resulting factor estimates for our data set of monthly U.S. government bond yields, together with the empirical proxies for the level, slope and curvature. The level factor was fairly stable around 7 percent during the first half of the 1970s, then increased rapidly to reach a maximum of 15 percent in 1982, and has declined steadily ever since. The slope factor generally is positive, reflecting the fact that most of the time the yield curve is upward sloping. The slope typically declines and turns negative towards the onset of recession periods. This is most pronounced for the earlier recessions during our sample period, i.e. the crisis in 1974 due to the OPEC oil price shocks and the double-dip recession in 1980-81. Prior to the more recent recessions in 1990-1991, 2001 and 2007-9, the slope did turn negative but only just. Correspondingly, the mean of the slope factor changes from 0.94 for the period 1970-1989 to 1.66 for the period 1990-2009. The mean of the curvature factor is close to zero over the full sample period, although a slight downward trend may be noted from the bottom panel of Figure 2. Indeed, while the mean curvature was positive and equal to 0.44 during the period 1970-1989, it was negative and equal to -0.43 during the period 1990-2009.

Figure 2: Estimated Level, Slope and Curvature from Nelson-Siegel Model
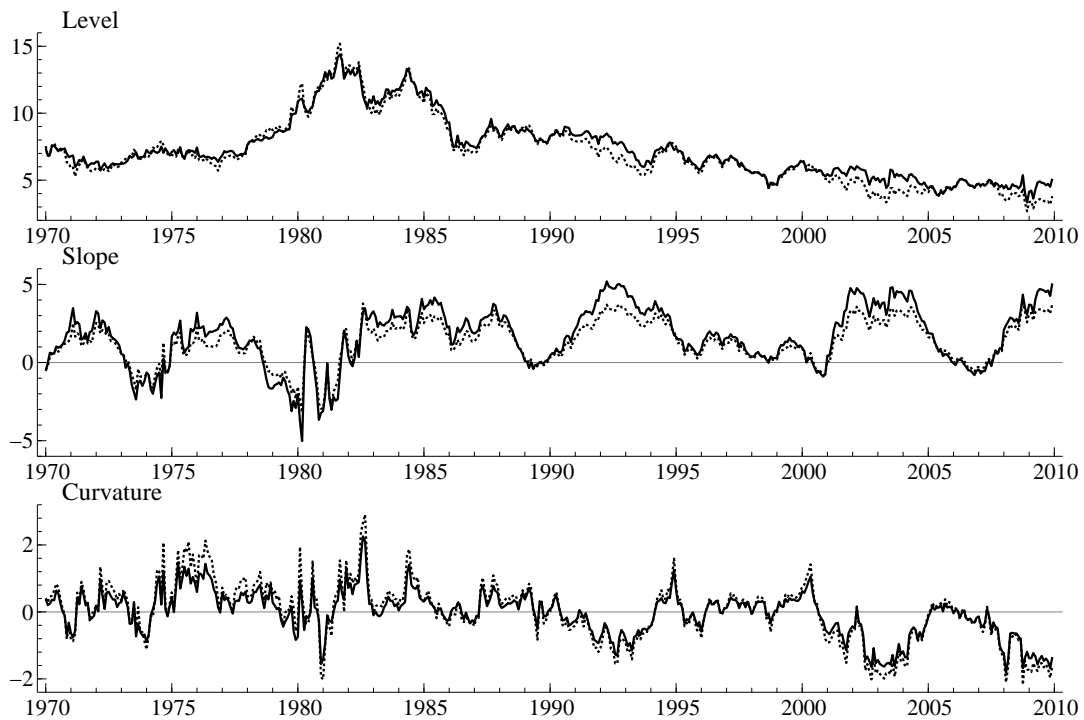
We present the estimated level, slope and curvature factors (solid) from the Nelson and Siegel (1987) model applied to monthly U.S. Treasury yields over the period January 1970 - December 2009, constructed using the unsmoothed Fama-Bliss method. In addition we include proxies for level, slope and curvature obtained directly from the data (dotted). The proxy for level is the longest maturity yield (120 months), for slope it is the longest (120 months) minus the shortest (3 months) maturities, and for curvature it is two times the 24 month yield minus the 3 month and 120 month yields. For ease of interpretation we present the negative slope from the Nelson-Siegel model.

10

3   Shifting endpoints in the dynamic Nelson-Siegel model

Under the (implicit) assumption that $|\phi_j| < 1$, the factor β_{jt} of the autoregressive specification (2) is a stationary, mean-reverting process with constant unconditional mean equal to μ_j. Given the historical behavior of U.S. Treasury yields, this may be regarded inappropriate. This is confirmed by the factor estimates discussed above, showing that the mean shows substantial variation over time in particular for the level factor, but to a lesser extent also for slope and curvature. To accommodate this feature in the dynamic Nelson-Siegel model, we assume that the factors have permanent as well as transitory components. Specifically, we modify the specification in (2) to allow for a time-varying unconditional mean, that is,

$$\beta_{j,t+1} = \mu_{j,t+1} + \phi_j \left(\beta_{jt} - \mu_{jt} \right) + \eta_{j,t+1}. \tag{3}$$

We label this specification with a time-varying unconditional mean as *shifting endpoint* model, consistent with Kozicki and Tinsley (2001), among others. A key element determining the success of this approach obviously is the specification of the time-varying unconditional mean $\mu_{j,t+1}$ in (3). We consider several possibilities, based on exponential smoothing methods, economic surveys, or functions of realized macroeconomic variables. In the remainder of this section we describe each of these in turn.

3.1   Shifting endpoints from exponential smoothing

A straightforward approach to allow for a permanent component in the factors β_{jt} is to consider the so-called local level model, where the unconditional mean $\mu_{j,t+1}$ is

generated by the exponential smoothing recursion

$$\mu_{j,t+1} = \alpha\beta_{jt} + (1-\alpha)\mu_{jt}, \qquad \text{for } t = 2, 3, \ldots, \tag{4}$$

with decay parameter $0 < \alpha < 1$ and starting with $\mu_{j1} = \beta_{j1}$. Note that by recursive substitution we can write

$$\mu_{j,t+1} = \alpha \sum_{l=0}^{t-2} (1-\alpha)^l \beta_{j,t-l} + (1-\alpha)^{t-1}\beta_{j1},$$

showing that the unconditional mean at time t is an exponentially weighted moving average of past factor values. Similarly, by substituting (4) into (3), we obtain

$$\beta_{j,t+1} = \omega_j \beta_{jt} + (1-\omega_j)\mu_{jt} + \eta_{j,t+1}, \qquad \text{with } \omega_j = \phi_j + \alpha, \tag{5}$$

such that the conditional expectation of the factor at time $t+1$ is a weighted average of the unconditional expectation and the realization of the factor at time t.

The equations (4) and (5) can be iterated forwards to obtain multi-step forecasts of β_{jt}, that is

$$\hat{\beta}_{j,t+h|t} = \omega_j \hat{\beta}_{j,t+h-1|t} + (1-\omega_j)\hat{\mu}_{j,t+h-1|t}, \qquad h = 1, 2, \ldots,$$

where

$$\hat{\mu}_{j,t+h-1|t} = \alpha\hat{\beta}_{j,t+h-2|t} + (1-\alpha)\hat{\mu}_{j,t+h-2|t},$$

with, for $h = 1$ and 2, $\hat{\mu}_{j,t+h-1|t} = \mu_{j,t+h-1} = \alpha\beta_{j,t+h-2} + (1-\alpha)\mu_{j,t+h-2}$ as given by (4).

We implement two variants of this approach. First, we apply the exponential smoothing only for the unconditional mean of the level factor β_{1t} and continue to use the AR(1) equation (2) for forecasting the slope and curvature factors β_{2t} and β_{3t}. Second, we use exponential smoothing to allow for shifting endpoints for all

three factors. We label these variants Exponential Smoothing for Level (ESL) and for Level, Slope and Curvature (ESLSC), respectively.

Obviously, the smoothness of the shifting endpoints and, eventually, the yield forecasts depends on the smoothing parameter α. When α is chosen closer to zero, the exponentially decaying weighting pattern $(1 - \alpha)^{l-1}$ converges more slowly to zero as l increases, resulting in a higher level of smoothness of the forecasts. In our empirical analysis, we set $\alpha = 0.1$ for monthly data, which is effectively close to 0.7 for yearly series.

3.2 Shifting endpoints from survey expectations

The shifting endpoint μ_{jt} in (3) represents the steady-state, or long-run mean, of the yield curve factors. From that perspective, we may consider linking them to long-term survey expectations of interest rates or related macroeconomic variables.

3.2.1 Blue Chip Financial Forecasts of Yields

Twice a year since 1984, Blue Chip Financial Forecasts conducts a survey of analysts' expectations of thirty-year yields from five to ten years hence. These long-range forecasts can be regarded as representing the time-varying steady-state predictions for long-term interest rates. In terms of the yield curve, the forecasts reflect the respondents' beliefs about the steady-state of the level factor, μ_{1t}. We may expect that these long-range survey forecasts would adapt quickly to structural breaks, making them particularly instrumental for measuring trends. Therefore we consider equation (3) for the level factor (i.e. $j = 1$) with μ_{1t} set equal to the long-range forecasts of the thirty-year yield from this Blue Chip survey.[6]

When we use this approach for forecasting the level factor, the unconditional mean is fixed by setting future values $\mu_{1,t+h}$ equal to the end-of-sample value of

---

[6]The long-range forecasts of the ten-year yield are also available in the Blue Chip survey but start later (in 1988). For this reason we take the thirty-year yield to represent forecasts on the long end of the yield curve.

the long-range survey forecast, treating the survey-implied μ_{1t} as a random walk. It seems natural to specify that the long-run expectation of the level factor is a random walk, because if it is truly a long-run expectation, then by the law of iterated expectations, revisions to it ought not to be forecastable. Meanwhile, the forecasts for the slope and curvature factors are generated from equation (2). We label this method Blue-Chip for Yields (BCY).

3.2.2 Blue Chip Economic Indicators: inflation and output growth

The yield curve is closely related to other macroeconomic variables. This carries over to the factors in the dynamic Nelson-Siegel model. Specifically, Diebold et al. (2006) argue that the level factor is correlated with inflation, while the slope factor seems to have some correlation with measures of real economic activity. Long-term forecasts of these variables are available from the Blue Chip Economic Indicators, which are obtained from surveys of analysts' expectations on a number of macroeconomic variables from five to ten years' since 1984. These surveys also are carried out twice a year. Let π_t and γ_t denote the most recent long-range forecast of inflation (GDP deflator) and of real GDP growth at time t. A natural approach to link the shifting endpoints to these long-term survey expectations is to specify that

$$\mu_{1t} = \theta_{0,1} + \theta_{1,1}\pi_t, \tag{6}$$

and

$$\mu_{2t} = \theta_{0,2} + \theta_{1,2}\gamma_t. \tag{7}$$

Estimates of the coefficients in equation (6) may be obtained from the regression

$$\beta_{1t} = \theta_{0,1} + \theta_{1,1}\pi_t + \xi_{1t}. \tag{8}$$

We can then again use equation (3) for forecasting the level factor, setting $\mu_{1t} = \theta_{0,1} + \theta_{1,1}\pi_t$ and projecting that μ_{1t} will remain constant at its end-of-sample value (as above, treating a long-run expectation as a random walk). In other words, we assume that the long-run survey inflation expectation π_t follows a random walk and is cointegrated with β_{1t} with cointegrating regression given by equation (8), and with a cointegrating error ξ_{1t} that is an AR(1).[7] Meanwhile, we obtain forecasts of the slope and curvature factors from AR(1) processes with constant mean, as in equation (2). In a variant of this, we continue to use inflation expectations to pin down μ_{1t}, but obtain estimates of the coefficients $\theta_{0,2}$ and $\theta_{1,2}$ in equation (7) by regressing β_{2t} onto a constant and γ_t, and set $\mu_{2t} = \theta_{0,2} + \theta_{1,2}\gamma_t$. In this approach, we still use equation (2) for the curvature factor. We label these two approaches Blue Chip Inflation for level (BCI) and Blue Chip Inflation for level and real GDP growth for slope (BCIG), respectively.

3.3 Shifting endpoints from realized measures

Instead of using long-run surveys to measure the trend components of macroeconomic variables, we could use exponential smoothing of realized inflation and growth data. Let π_t^{ES} and γ_t^{ES} denote real-time exponentially smoothed realized inflation and industrial production (IP) growth, respectively [8] in month t and consider the regressions

$$\beta_{1t} = \theta_{0,1} + \theta_{1,1}\pi_t^{ES} + \xi_{1t}. \tag{9}$$

and

$$\beta_{2t} = \theta_{0,2} + \theta_{1,2}\gamma_t^{ES} + \xi_{2t}. \tag{10}$$

We can again use equation (3) for forecasting the level factor, setting $\mu_{1t} =$

---

[7]A variant on this theme would include other survey forecasts on the right-hand-side of equation (8). We investigated this, but found that it worsens forecast performance. Perhaps this is not surprising given the close association between the level factor of interest rates and inflation.

[8]We use real-time CPI inflation data from the dataset of the Federal Reserve Bank of Philadelphia (month t vintage data) and from the dataset collected by Norman Swanson, Dick Van Dijk and Miles Callan. Real-time data IP data are taken from the Federal Reserve Bank of Philadelphia.

$\theta_{0,1} + \theta_{1,1}\pi_t^{ES}$ and projecting that μ_{1t} will remain constant at its end-of-sample value. We can forecast the slope and curvature factors from AR(1) processes with constant mean in equation (2). Alternatively, we could use (3) to forecast both the level and slope factors, setting $\mu_{2t} = \theta_{0,2} + \theta_{1,2}\gamma_t^{ES}$, projecting that μ_{2t} will remain constant at its last value, and only using the constant-mean AR(1) specification for the curvature.[9] We label these two variants smooth RealiZed Inflation for level (RZI) and for Inflation for level and real GDP growth (RZIG), respectively.

4   Empirical results

We now turn to the results of our empirical analysis. In Section 4.1 we discuss several implementation details of the out-of-sample forecasting exercise. In Section 4.2 we discuss our main results. In Section 4.3 we examine the statistical significance of differences in forecast accuracy across models, while in Section 4.4 we provide additional insight on how the shifting endpoints forecasts differ from the usual Nelson-Siegel forecasts.

4.1   Implementation

We examine the predictive ability of the Nelson-Siegel model with different shifting endpoint specifications in a recursive out-of-sample forecasting experiment using the bond data from Section 2.2. Due to availability of the Blue Chip survey data we consider the period from January 1985 through December 2009. We forecast using an expanding window, using in all cases data back to January 1985 for estimating the model parameters. Our first forecast is made for interest rates in January 1994 and then recursive out-of-sample forecasts are made for all subsequent months up

---

[9]We also investigated using the exponentially smoothed first principal component of a small set of real-time realized inflation measures as a proxy for π_t^{ES} and the smoothed first principal component of a small set of real-time realized activity measures as a proxy for γ_t^{ES}. However, this worked slightly less well than the two methods utilizing realized measures that are currently considered in the paper.

to and including December 2009. The forecasts are made for interest rates of all maturities at horizons $h = 6, 12, 24$ months ahead.

The different forecasting methods are listed in Table 2, which also provides the labels for the methods that we use in the discussions below. Besides the models from Section 3 we consider three models that can be viewed as benchmarks. First, we obtain forecasts with the Dynamic Nelson-Siegel (or Diebold-Li) method as described in Section 2.1. Second, we use two variants of random walk specifications for the yield curve. In a first variant we adopt a random walk process for the level, slope and curvature factors in the Nelson-Siegel model (which amounts to equation (2) with $\mu_j = 0$ and $\phi_j = 1$ for all j). In a second variant we simply use a random walk for each of the yields individually. We label these two variants RW and RWY, respectively.

Table 2: Econometric methods and their labels

We list the forecasting methods considered in this paper and give its acronym.

Label	Description
DL	Dynamic Nelson-Siegel method or Diebold-Li method
RW	Random Walk process for the level, slope and curvature factors
RWY	Random Walk process for interest rates of each maturity
ESL	Exponential smoothing for the level factor only
ESLSC	Exponential smoothing for the level, slope and curvature factors
BCY	Blue Chip Survey for thirty-year interest rate
BCI	Blue Chip Economic Indicator of inflation for level factor
BCIG	BCI plus its indicator of GDP growth for slope factor
RZI	Smooth realized inflation for level factor
RZIG	RZI plus smooth realized GDP growth for slope factor

4.2   Results

We present forecasting results for a representative selection of yields, at maturities $n = 3, 12, 36, 60, 120$ months.[10] The root mean square prediction errors (RMSPE)

---

[10]Detailed results for other maturities are available upon request.

of the different forecasts are presented in Table 3. Several interesting conclusions emerge from these numbers. First, over the sample for which we are assessing the forecasts, the random walk predictions (RW or RWY) generally do better than the DL forecasts. For example, at the 12-month forecast horizon, either of the two random walk forecasts gives a reduction in out-of-sample RMSPE of 2 to 7 percent, relative to DL. This corroborates the findings of Duffee (2011) that imposing non-stationarity for the level factor (and in our case also the slope and curvature factors) enhances the predictive ability of the yield curve. Also note that the RMSPEs of the two random walk specifications differ only marginally. Hence, once nonstationarity of the yield curve is imposed, the factor structure of the Nelson-Siegel framework does not provide added value.

Second, incorporating shifting endpoints via exponential smoothing does not improve forecast accuracy at short and medium horizons. Both the ESL and ESLSC methods generally do not perform better than the RW(Y) specifications for 6 and 12 months ahead forecasts. However, at the long 24 month horizon improvements in RMSPE are found, up to 10 percent, especially for short maturities. Interestingly, including shifting endpoints for the slope and curvature factors leads to slightly worse forecasts, especially for $h = 24$ months ahead.

Third, the Blue Chip survey expectations of the thirty-year yield are not informative in this context. The RMSPE of the BCY approach exceeds that of the RW(Y) methods, often by quite a margin. In fact, overall BCY seems to render the least accurate forecasts among the methods considered here.

Fourth, the best forecasts in our comparison are those obtained from the BCI(G) and RZI(G) shifting endpoint methods. These are all based on relating the level factor in the Nelson-Siegel model to the trend component of inflation (and relating the slope factor to the trend component of GDP growth). All four specifications give more accurate forecasts than the DL and random walk models for all maturities at the longer horizons of 12 and 24 months, and for the long-maturity interest rates

Table 3: Root Mean Square Prediction Errors of Alternative Interest Rate Forecasts

We report the out-of-sample root mean square prediction error of alternative forecasts of interest rates for the period from January 1994 to December 2009. Units are in percentage points. The forecast evaluation exercise is recursive, using in all cases data back to January 1985 for model estimation. The forecast method labels are given in Table 2.

Maturity (in months)	3	12	36	60	120	3	12	36	60	120	3	12	36	60	120
Forecast Horizon	$h = 6$ months					$h = 12$ months					$h = 24$ months				
DL	0.96	0.98	0.93	0.85	0.68	1.60	1.54	1.35	1.18	0.94	2.27	2.16	1.82	1.55	1.22
RW	0.90	0.91	0.88	0.81	0.67	1.56	1.51	1.29	1.12	0.88	2.43	2.30	1.79	1.40	0.93
RWY	0.89	0.92	0.88	0.81	0.66	1.55	1.52	1.29	1.11	0.88	2.42	2.32	1.78	1.39	0.93
ESL	0.92	0.96	0.90	0.82	0.67	1.54	1.52	1.31	1.13	0.90	2.20	2.07	1.63	1.31	0.93
ESLSC	0.92	0.95	0.89	0.82	0.67	1.57	1.53	1.30	1.13	0.90	2.31	2.15	1.67	1.33	0.94
BCY	1.11	1.10	1.04	0.95	0.77	1.69	1.61	1.41	1.23	0.97	2.19	2.07	1.75	1.47	1.11
BCI	0.99	0.98	0.90	0.81	0.62	1.54	1.46	1.22	1.02	0.73	2.07	1.94	1.54	1.23	0.83
BCIG	0.98	0.97	0.90	0.80	0.61	1.52	1.44	1.20	1.00	0.73	1.97	1.86	1.49	1.19	0.81
RZI	0.97	0.97	0.87	0.78	0.59	1.55	1.47	1.19	0.98	0.68	2.07	1.93	1.49	1.15	0.72
RZIG	0.97	0.97	0.87	0.77	0.59	1.55	1.47	1.19	0.97	0.68	2.06	1.92	1.46	1.13	0.70

19

at the short 6-month horizon. For example, in forecasting ten-year yields at the one-year horizon, the BCI(G) forecasts reduce RMSPE by a substantial 20 percent relative to the random walk forecasts, while the RZI(G) forecasts show even larger improvements. Comparing the forecasts based on Blue Chip economic indicators and based on realized measures across all maturities and horizons, it seems that the latter dominate, with differences in RMSPE of up to about 10 percent. Also, for both the BC and RZ approaches it seems that allowing for a shifting endpoint in the level factor is the crucial aspect of the success of these methods, in the sense that allowing for a shifting endpoint in the slope factor does not lead to further improvements in forecast accuracy (with the possible exception of the short end of the curve at long horizons, where the BCIG approach outperforms the BCI method by some margin).

In sum, it appears to be beneficial to forecast interest rates by modeling the level factor in "gap" form around a linear function of the trend component of inflation. This trend can be taken from surveys (BCI(G)) or from exponential smoothing of realized inflation (RZI(G)). It may seem surprising that it is better to use the trend component of *inflation* rather than the trend component of the *level of interest rates* directly (ESL(SC) and BCY). In this regard, it is worth noting that several authors have found a range of survey inflation forecasts to have excellent predictive performance, see, for example, Croushore (2010) and Ang et al. (2007). The evidence on the quality of survey forecasts of interest rates is more mixed, see Bacchetta et al. (2009).

4.3   Statistical Significance of Forecast Improvements

The statistic of Diebold and Mariano (1995) can be used to compare the predictive accuracy of the different forecast methods that we are assessing. The statistic is simply the ratio of the out-of-sample MSPEs of one of the forecasts relative to that of another. Assessing the statistical significance of these Diebold-Mariano statistics can

however involve tricky statistical issues because several pairs of prediction models that are being compared are nested, see West (2006). We accordingly establish the statistical significance of the Diebold-Mariano statistic via two bootstraps. The first bootstrap tests the null hypothesis that the data generating process is an AR(1) for each of the factors, i.e. that the DL forecast is correctly specified. The steps in our bootstrap procedure are as follows:

(i) We estimate the factors $\{\beta_{jt}\}$ using cross-sectional OLS (as explained in Section 2.2) and then fit AR(1)s to these, using the bootstrap bias-adjustment of Kilian (1998).

(ii) Resampling from the residuals with replacement, we create bootstrap datasets of the factors, again using Kilian's bias-adjustment. We also resample from the idiosyncratic errors in yields (the $\varepsilon_t(\tau)$ disturbances of equation (1)) and resample from the long-run survey predictions, assuming that they are random walks. We draw the same shock for the factors, the $\varepsilon_t(\tau)$'s and the first differences of the survey predictions, and so preserve any cross-sectional dependence among these variables.

(iii) On this bootstrap sample, we then compute the ratio of the recursive out-of-sample MSPE using the ESL(SC), BCY, BCI(G) and RZI(G) methods relative to the MSPE of the DL method.

(iv) This is repeated over 500 bootstrap replications, giving the bootstrap approximation to the null distribution of the Diebold-Mariano test statistic.

Table 4 presents the p-values from the Diebold-Mariano statistic comparing the DL forecast with all shifting endpoint specifications using this bootstrap approximation to the null distribution. This is a one-sided test. Note that we would expect the null distribution of the Diebold-Mariano statistic to be centered a bit above unity, because under the null that the DL model is correctly specified, these

21

other more complicated models should tend to have *worse* out-of-sample forecasting performance.

As can be seen in Table 4, at conventional significance levels the improvements in forecast accuracy offered by the BCI(G) and RZI(G) methods over the standard DL approach are statistically significant for maturities beyond one year across all forecasting horizons. The improvement in forecasting performance in using ESL(SC) rather than DL is significant only for predicting the ten-year yield at the two-year horizon.

The second bootstrap tests the null hypothesis that the data generating process for each factor is a random walk. The bootstrap imposes that the three factors and all long-run survey projections follow random walks. The recursive out-of-sample Diebold-Mariano statistic is then computed in 500 bootstrap replications. Table 5 shows the p-values from the Diebold-Mariano statistic comparing the RW forecast with DL and all shifting endpoint specifications, using this bootstrap approximation to the null distribution. Again this is a one-sided test. At the two-year horizon, the predictions obtained with survey forecasts or realized measures of inflation as shifting endpoint of the level factor give significant improvements relative to the RW prediction across all maturities. At forecast horizons of 6 and 12 months, the differences in RMSPE become insignificant for short maturities, but remain significant for yields at longer maturities.

Overall we conclude that the gains in interest rate forecast accuracy from capturing shifting endpoints via centering the Nelson-Siegel level factor around survey projections or realized measures of inflation is both economically and statistically significant.

4.4 Bias

As an illustration of the performance of the different interest rate forecasts, Figure 3 shows the time series of two-year-ahead forecasts of the ten-year yield from DL, ESL

Table 4: Bootstrap p-values testing null of equal accuracy of Diebold-Li and selected forecasts

We report the bootstrap p-values from Diebold-Mariano tests comparing the hypothesis of equal forecast accuracy of the Diebold-Li forecast with selected alternatives. The bootstrap imposes the null hypothesis that the Dynamic Nelson-Siegel model is correctly specified, and is described further in Section 4.3. The forecast method labels are given in Table 2.

Maturity (in months)	3	12	36	60	120	3	12	36	60	120	3	12	36	60	120
Forecast Horizon	h = 6 months					h = 12 months					h = 24 months				
ESL	0.17	0.29	0.30	0.31	0.48	0.25	0.34	0.34	0.33	0.38	0.31	0.32	0.23	0.13	0.05
ESLSC	0.40	0.37	0.23	0.26	0.46	0.66	0.52	0.31	0.29	0.35	0.72	0.54	0.24	0.14	0.04
BCY	0.85	0.85	0.89	0.92	0.93	0.49	0.45	0.55	0.53	0.47	0.25	0.24	0.25	0.24	0.20
BCI	0.46	0.26	0.08	0.04	0.01	0.21	0.11	0.03	0.01	0.00	0.14	0.13	0.05	0.02	0.00
BCIG	0.25	0.20	0.08	0.03	0.00	0.13	0.09	0.03	0.01	0.00	0.12	0.08	0.04	0.01	0.00
RZI	0.36	0.18	0.04	0.01	0.00	0.20	0.13	0.03	0.00	0.00	0.17	0.12	0.03	0.01	0.00
RZIG	0.31	0.24	0.05	0.01	0.00	0.25	0.19	0.04	0.00	0.00	0.21	0.18	0.04	0.01	0.00

**Table 5: Bootstrap *p*-values testing null of equal accuracy of Random Walk and selected forecasts**

We report the bootstrap *p*-values from Diebold-Mariano tests comparing the hypothesis of equal forecast accuracy of the Random Walk forecast with selected alternatives. The bootstrap imposes the null hypothesis that the Random Walk model is correctly specified, and is described further in Section 4.3. The forecast method labels are given in Table 2.

Maturity (in months)	3	12	36	60	120	3	12	36	60	120	3	12	36	60	120
Forecast Horizon	$h = 6$ months					$h = 12$ months					$h = 24$ months				
DL	0.54	0.56	0.62	0.58	0.37	0.25	0.28	0.44	0.53	0.53	0.12	0.13	0.31	0.57	0.92
ESL	0.32	0.55	0.53	0.46	0.41	0.17	0.25	0.36	0.43	0.62	0.09	0.07	0.09	0.14	0.40
ESLSC	0.78	0.93	0.76	0.54	0.49	0.41	0.63	0.61	0.69	0.87	0.03	0.00	0.01	0.01	0.56
BCY	0.91	0.92	0.96	0.96	0.96	0.37	0.39	0.58	0.64	0.67	0.09	0.09	0.19	0.38	0.69
BCI	0.31	0.26	0.20	0.08	0.01	0.05	0.04	0.04	0.03	0.00	0.01	0.01	0.03	0.03	0.03
BCIG	0.23	0.24	0.17	0.08	0.01	0.05	0.04	0.02	0.01	0.00	0.01	0.00	0.01	0.02	0.06
RZI	0.40	0.31	0.15	0.06	0.00	0.12	0.11	0.07	0.02	0.00	0.03	0.03	0.03	0.03	0.01
RZIG	0.35	0.32	0.12	0.05	0.00	0.14	0.11	0.05	0.02	0.00	0.05	0.03	0.03	0.03	0.01

and BCI. The actual value is also plotted. The OLS estimates of the autoregressions fitted to the Nelson-Siegel factors are stationary and hence by construction give forecasts of yields that are tending toward the sample mean of interest rates over the estimation period (since 1985). Given the downward trend of the ten-year yield over this period, of course these predictions turn out to be consistently too high. The ESL and BCI forecasts proposed in this paper do not make this same mistake, because they allow for shifting endpoints: these forecasts are consistently lower, and exactly this feature makes them more accurate. The forecasting methods that we propose are intended to be simple yet robust ways of controlling for intercept shifts in inflation, inflation-expectations and the term structure of interest rates, of the sort considered by Clements and Hendry (1998, 1999) and Kozicki and Tinsley (2001, 2005).

The point is made more generally in Table 6, which shows the bias of the different forecasts over the complete out-of-sample period from January 1994 to December 2009. The DL forecast show a consistent and large upward bias at all maturities. The bias is around 65 basis points at the one-year forecast horizon, and over a full percentage point at the two-year horizon. All other forecasting methods that allow for some form of nonstationarity in the level factor exhibit upward bias too, but it is much less pronounced. The RZIG predictions have the smallest bias, but the bias of the BCI, BCIG and RZI forecasts are also modest. Reducing this bias is clearly part of the reason for the good performance of these forecasting methods. Note however that the differences in bias do not completely explain the differences in forecast accuracy, as we also observe that the RW(Y) and ESL(SC) methods lead to a substantial bias reduction compared to the standard DL approach. In fact, the bias of these forecasts is lower than the bias of the BCI(G) specifications. Clearly, comparing Tables 3 and 6, the advantage that the BCI(G) and RZI(G) forecasts all have relative to the random walk and exponential smoothing projections in terms of RMSPE is that they are generally less variable.

Figure 3: Selected Two-Year-Ahead forecasts of Ten-Year Yields
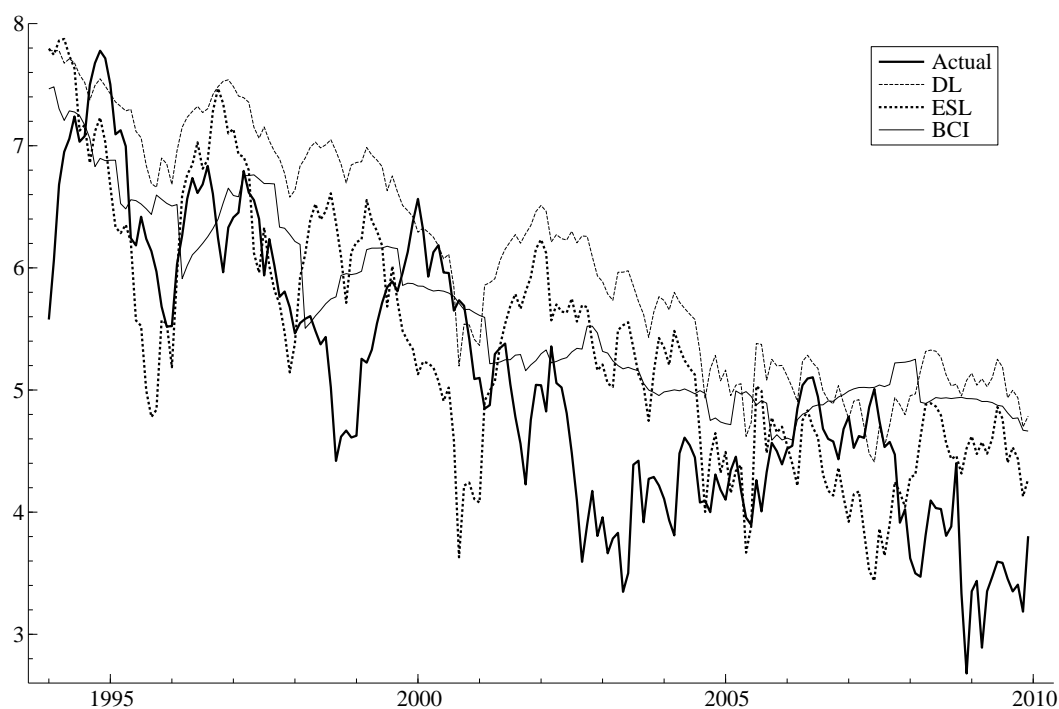
We present the forecasts from the Diebold-Li, exponential smoothing for level only (ESL) and Blue Chip Economic Indicators survey-centered (BCI) forecasts. The forecasts are for ten-year yields, at the two-year forecast horizon. Actual values are also shown. Yields are in percentage points. All forecasts are shown as of the date for which the forecast is being made. For example, a two-year-ahead forecast made in January 1996 is marked as applying to January 1998.

Table 6: Bias of Alternative Interest Rate Forecasts

We present the out-of-sample bias of alternative forecasts of interest rates for the period from January 1994 to December 2009 (forecast value less actual subsequent realization). Units are in percentage points. The forecast evaluation exercise is recursive, using in all cases data back to January 1985 for model estimation. The forecast method labels are given in Table 2.

Maturity (in months)	3	12	36	60	120	3	12	36	60	120	3	12	36	60	120
Forecast Horizon	$h = 6$ months					$h = 12$ months					$h = 24$ months				
DL	0.38	0.32	0.38	0.37	0.28	0.65	0.61	0.69	0.67	0.56	1.00	1.02	1.14	1.13	1.01
RW	0.15	0.07	0.12	0.12	0.04	0.23	0.16	0.21	0.21	0.14	0.35	0.31	0.41	0.42	0.35
RWY	0.09	0.10	0.09	0.08	0.07	0.18	0.18	0.19	0.18	0.15	0.32	0.32	0.39	0.39	0.35
ESL	0.15	0.09	0.15	0.14	0.05	0.21	0.18	0.26	0.24	0.13	0.30	0.32	0.45	0.43	0.32
ESLSC	0.13	0.08	0.15	0.14	0.05	0.18	0.16	0.25	0.23	0.12	0.24	0.29	0.44	0.43	0.31
BCY	0.45	0.39	0.46	0.45	0.35	0.65	0.62	0.70	0.68	0.57	0.79	0.81	0.94	0.92	0.81
BCI	0.27	0.21	0.27	0.26	0.17	0.40	0.37	0.44	0.42	0.31	0.52	0.54	0.66	0.65	0.53
BCIG	0.22	0.17	0.25	0.25	0.16	0.30	0.29	0.40	0.39	0.29	0.36	0.41	0.59	0.60	0.51
RZI	0.11	0.05	0.11	0.10	0.01	0.19	0.16	0.23	0.21	0.10	0.25	0.27	0.39	0.38	0.26
RZIG	0.05	0.01	0.09	0.09	0.00	0.09	0.08	0.19	0.18	0.08	0.09	0.15	0.32	0.33	0.24

5   Conclusions

The literature on modeling and forecasting inflation has found it to be important to model inflation as being stationary about a slowly-varying trend, reflecting the preferences and credibility of the central bank. That trend can be proxied by time series methods or by surveys. For example, exponential smoothing and unobserved components models have been used by Stock and Watson (2007). Forecasts of inflation that are constructed in "gap" form, as deviations from the trend, do better out of sample than forecasts that model inflation as simply being a stationary process, especially at longer horizons; see also the discussion in Stock and Watson (2010).

In contrast, the literature on forecast interest rates generally treats yields as being stationary, without any shifting endpoints, or alternatively treats interest rates as following random walks. That seems a disconnect, since long-term nominal interest rates are heavily influenced by inflation expectations and risk premia. We have considered interest rate forecasting that takes account of shifting endpoints. We find that the best approach to forecasting interest rates is to treat the level factor of interest rates as having a trend that is in turn a linear function of the trend in inflation. Especially for forecasting at long horizons and forecasting long-maturity interest rates, this method gives substantial gains in out-of-sample predictive performance, relative to the forecasts obtained by the dynamic Nelson-Siegel approach of Diebold and Li (2006) or random walk predictions, that are both economically and statistically significant.

References

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

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

Bacchetta, P., Mertens, E. and van Wincoop, E. (2009), 'Predictability in financial markets: What do survey expectations tell us?', *Journal of International Money and Finance* **28**, 406–426.

Bowsher, C. G. and Meeks, R. (2008), 'The dynamics of economic functions: modeling and forecasting the yield curve', *Journal of the American Statistical Assocation* **103**, 1419–37.

Chun, A. L. (2011), 'Expectations, bond yields, and monetary policy', *Review of Financial Studies* **24**, 208–247.

Chun, A. L. (2012), Forecasting interest rates and inflation: Blue Chip clairvoyants or econometrics? Working paper.

Cieslak, A. and Povala, P. (2010), Understanding bond risk premia. SSRN: http://ssrn.com/abstract=1709636.

Clark, T. E. (2011), 'Real-time density forecasts from VARs with stochastic volatility', *Journal of Business and Economic Statistics* **29**, forthcoming.

Clements, M. P. and Hendry, D. F. (1998), *Forecasting Economic Time Series*, Cambridge University Press, Cambridge.

Clements, M. P. and Hendry, D. F. (1999), *Forecasting Nonstationary Economic Time Series*, MIT Press, Cambridge, Massachusetts.

Cogley, T. and Sargent, T. J. (2005), 'Drifts and volatilities: Monetary policies and outcomes in the post World War II U.S.', *Review of Economic Dynamics* **8**, 262–302.

Croushore, D. (2010), 'An evaluation of inflation forecasts from surveys using real-time data', *BE Journal of Macroeconomics* **10**.

Dai, Q. and Singleton, K. J. (2000), 'Specification analysis of affine term structure models', *Journal of Finance* **55**, 1943–78.

De Pooter, M. (2007), Examining the Nelson-Siegel class of term structure models. Tinbergen Institute Discussion Paper.

De Pooter, M., Ravazzolo, F. and van Dijk, D. (2010), Term structure forecasting using macro factors and forecast combination. Federal Reserve Board of Governors International Finance Discussion Papers No. 2010-993.

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

Diebold, F. X. and Mariano, R. S. (1995), 'Comparing predictive accuracy', *Journal of Business and Economic Statistics* **13**, 253–263.

Diebold, F. X., Rudebusch, G. D. and Aruoba, S. B. (2006), 'The macroeconomy and the yield curve', *Journal of Econometrics* **131**, 309–338.

Duffee, G. R. (2002), 'Term premia and interest rate forecasts in affine models', *Journal of Finance* **57**, 405–443.

Duffee, G. R. (2011), Forecasting with the term structure: The role of no-arbitrage restrictions. Working paper.

Duffee, G. R. (2012), Bond pricing and the macroeconomy. Working paper, prepared for the 'Handbook of the Economics of Finance'.

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

Ehling, P., Gallmeyer, M., Heyerdahl-Larsen, C. and Illeditsch, P. (2012), Beliefs about Inflation and the Term Structure of Interest Rates. Working paper.

Exterkate, P., Van Dijk, D., Heij, C. and Groenen, P. J. F. (2012), 'Forecasting the yield curve in a data-rich environment using the factor-augmented Nelson-Siegel model', *Journal of Forecasting* p. forthcoming.

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

Faust, J. and Wright, J. H. (2012), Forecating Inflation. Working paper, prepared for the 'Handbook of the Economic Forecasting'.

Joslin, S., Le, A. and Singleton, K. J. (2011), Why Gaussian macro-finance term structure models are (nearly) unconstrained factor-VARs. Working paper.

Kilian, L. (1998), 'Small-sample confidence intervals for impulse response functions', *Review of Economics and Statistics* **80**, 218–230.

Kim, D. H. and Orphanides, A. (2012), 'Term Structure Estimation with Survey Data on Interest Rate Forecasts', *Journal of Financial and Quantitative Analysis* **47**, 241–272.

Koopman, S. J., Mallee, M. and Van der Wel, M. (2010), 'Analyzing the term structure of interest rates using the dynamic Nelson-Siegel model with time-varying parameters', *Journal of Business and Economic Statistics* **28**, 329–343.

Kozicki, S. and Tinsley, P. A. (2001), 'Shifting endpoints in the term structure of interest rates', *Journal of Monetary Economics* **47**, 613–652.

Kozicki, S. and Tinsley, P. A. (2005), 'What do you expect? Imperfect policy credibility and tests of the expectations hypothesis', *Journal of Monetary Economics* **52**, 421–447.

Litterman, R. and Scheinkman, J. A. (1991), 'Common factors affecting bond returns', *Journal of Fixed Income* **1**, 54–61.

Nelson, C. R. and Siegel, A. F. (1987), 'Parsimonious modeling of yield curves', *Journal of Business* **60**, 473–489.

Orphanides, A. and Wei, M. (2012), 'Evolving macroeconomic perceptions and the term structure of interest rates', *Journal of Economic Dynamics and Control* **36**, 239–254.

Piazzesi, M. and Schneider, M. (2011), Trend and Cycle in Bond Premia. Working paper.

Stock, J. H. and Watson, M. W. (2007), 'Why has U.S. inflation become harder to forecast?', *Journal of Money, Credit and Banking* **39**, 3–33.

Stock, J. H. and Watson, M. W. (2010), Modeling inflation after the crisis. Macroeconomic challenges: the decade ahead, Federal Reserve Bank of Kansas City Economic Policy Symposium.

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

Wright, J. H. (2012), 'Evaluating real-time VAR forecasts with an informative democratic prior', *Journal of Applied Econometrics* p. forthcoming.