Forecasting the term structures of Treasury and corporate yields using dynamic Nelson-Siegel models

Wei-Choun Yu\textsuperscript{a,\ast}, Eric Zivot\textsuperscript{b}

\textsuperscript{a}Winona State University, MN, United States
\textsuperscript{b}University of Washington, Seattle, WA, United States

Abstract

We extend Diebold and Li’s dynamic Nelson-Siegel three-factor model to a broader empirical prospective by including the evaluation of the state space approach and by using nine different ratings for corporate bonds. We find that the dynamic Nelson-Siegel factor AR(1) model outperforms other competitors on the out-of-sample forecast accuracy, especially on the investment-grade bonds for the short-term forecast horizon and on the high-yield bonds for the long-term forecast horizon. The dynamic Nelson-Siegel factor state space model, however, becomes appealing on the high-yield bonds in the short-term forecast horizon, where the factor dynamics are more likely time-varying and parameter instability is more probable in the model specification.

\footnote{Corresponding author. Tel.: +\textcolor{red}{1} 507 457 2982; fax: +\textcolor{red}{1} 507 457 5697. E-mail address: wyu@winona.edu (W.-C. Yu).}

Keywords: Term structures; Treasury yields; Corporate yields; Nelson-Siegel model; Factor model; AR(1); VAR(1); Out-of-sample forecasting evaluations

1. Introduction

Forecasting the term structure of interest rates plays a crucial role in portfolio management, household finance decisions, business investment planning, and policy formulation. Improved methods for yield curve modeling and forecasting have recently been gaining more and more attention. It is well known that the no-arbitrage models (Ho & Lee, 1986; Hull & White, 1990) focus only on term-structure cross-sectional fitting. Although they are the most common, the affine equilibrium models (Cox, Ingersoll, & Ross, 1985; Vasicek, 1977; Dai & Singleton, 2000) only look at instantaneous short rates, and thus result in poor yield curve forecasts (Duffee, 2002). To provide an improved forecasting model for the yield curve, Diebold and Li (2006, hereafter DL) extended the parsimonious three-factor (exponential components) yield curve model of Nelson and Siegel (1987) to the dynamic form. Compared to the no-arbitrage and affine equilibrium models, DL’s dynamic Nelson-Siegel factor autoregressive of order 1 (AR(1)) model produced superior out-of-sample forecasts of US Treasury yields, especially at the 1-year-ahead horizon.

DL used a simple two-step approach, in which they first estimate three factors, and then model and forecast them. Diebold, Rudebusch, and Aruoba (2006, hereafter DRA) proposed a one-step approach that uses a state space model to perform factor estimation, modeling, and forecasting simultaneously. They argued that the one-step state space approach provides a unified framework, and should improve the out-of-sample forecasts. However, they did not perform any out-of-sample forecast evaluations in their study. In this paper, we bridge the empirical gap and evaluate the out-of-sample forecasting performances of one-step and two-step dynamic Nelson-Siegel models, as well as those of some common competing models. We find, surprisingly, that the state space approach does not improve the out-of-sample forecasts for Treasury yields: the simple two-step AR(1) method of DL is still the most accurate forecasting model.

Many financial institutions and academies have devoted substantial resources to the measurement and management of credit (default) risk over the past decade. However, less attention has been paid to the modeling and time series forecasting of credit yield curves. To partially fill this gap, we use the dynamic Nelson-Siegel three-factor model, with both the one-step and two-step approaches, to model and forecast AAA, AA, A+, A, A-, BB+, BB-, B, and B- rated corporate bond yields. The different nature of interest rate data relative to various credit rating bonds provides a broader universe in which to evaluate the models’ predictability. We find that Treasury yields and investment-grade (AAA, AA, and A+) corporate yields have similar results: the two-step AR(1) method produces superior and robust out-of-sample forecasts. We find, however, that the one-step state space approach beats the AR(1) approach for the high-yield (speculative) bonds (A-, BB+, BB-, B, and B-), where the model parameters are more likely to be time-varying.

It is desirable to know the long-term forecast performances of yield curve models for long-term investments and portfolio management, especially for longer maturity bonds. A successful model for short-term forecasting, however, does not necessarily produce good long-term forecasts. DL only provided short-term out-of-sample forecasts (1-, 6-, and 12-month-ahead forecasts). In addition to short-term forecasts, we also evaluate long-term (36- and 60-month-ahead) forecasts of Treasury and corporate yields. We find that the performances of the models’ short-run and long-run predictions are different, and that the long-term forecasts of high yield bonds from the two-step AR(1) model are very robust.

In summary, we make four contributions to the literature. First, we evaluate and compare the forecasting performances of the one-step and two-step dynamic Nelson-Siegel three-factor models. Second, we model and forecast Treasury and corporate bond yields. Third, we evaluate both short-run and long-run forecasts. Fourth, we provide evidence that the level factors include the additional credit risk, which is asymmetrically proportional to non-investment grade bonds with a higher leverage ratio.

The rest of the paper is organized as follows. Section 2 presents the model and the data; Section 3 reports the forecasting evaluation; and Section 4 concludes.

2. Model and data

2.1. Nelson-Siegel model

Nelson and Siegel (1987) introduced a parsimonious and influential three-factor model for zero coupon bond yields:

$$y_t(\tau_i) = \beta_{1t} + \beta_{2t} \left( \frac{1 - e^{-\lambda \tau_i}}{\lambda \tau_i} \right) + \beta_{3t} \left( \frac{1 - e^{-\lambda \tau_i}}{\lambda \tau_i} - e^{-\lambda \tau_i} \right),$$

(1)

where $y_t$ denotes the yield at time $t$, and $\tau_i$ is the maturity of bond $i$, which usually ranges from 3 months to 30 years. The parameter $\lambda$ determines the rate of exponential decay. The three factors are $\beta_{1t}$, $\beta_{2t}$, and $\beta_{3t}$. The factor loading on $\beta_{1t}$ is 1, and loads equally at all maturities. A change in $\beta_{1t}$ changes all yields uniformly, and it is therefore called the level factor. As the value of $\tau_i$ increases, $\beta_{1t}$ plays a more important role in forming yields relative to the smaller factor loadings on $\beta_{2t}$ and $\beta_{3t}$. In the limit, $y_t(\infty) = \beta_{1t}$, and thus $\beta_{1t}$ is also called the long-term factor. The factor loading on $\beta_{2t}$ is $(1 - e^{-\lambda \tau_i})/\lambda \tau_i$, which is a

---

1 The original Nelson-Siegel model is slightly different from Eq. (1), which is a modified equation from DL. DL explain the reasons for this revision.
function which decays quickly and monotonically to zero as \(\tau\) increases. It loads short rates more heavily than long rates, and as a consequence, it changes the slope of the yield curve. Hence, \(\beta_{2t}\) is a short-term factor, which is also called the slope factor.\(^2\) The factor loading on \(\beta_{3t}\) is \((1 - e^{-\lambda\tau_t})/\lambda\tau_t - e^{-\lambda\tau_t}\), which is a function starting at zero (not short-term) and decaying to zero (not long-term), with a humped shape in the middle. It loads medium rates more heavily. Accordingly, \(\beta_{3t}\) is the medium-term factor. It is also called the curvature factor, because an increase in \(\beta_{3t}\) will increase the yield curve curvature.

The Nelson-Siegel three factor model can capture various different kinds of yield curves, including upward sloping, downward sloping, humped and inverted humped. It is well known that the exponential component model has competitive in-sample predictions, in spite of its parsimonious structure.

2.2. VAR(1) and AR(1) factor models—two-step approach

Despite its success in the cross-sectional interpolation of yields across maturities, the Nelson-Siegel model is not well suited to out-of-sample time series forecasts, because the factors \(\beta_{1t}, \beta_{2t}\) and \(\beta_{3t}\) do not stay constant over time. This motivates DL to extend the Nelson-Siegel model to a dynamic form in which the factors are time varying. They proposed two benchmark models for the evolution of the factors: a multivariate vector autoregressive VAR(1) model for \(\beta_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})^\top\), and univariate AR(1) models for \(\beta_{it}\) \((i = 1, 2, 3)\). DL followed a two-step method in which they first estimate the three factors \(\hat{\beta}_{1t}, \hat{\beta}_{2t},\) and \(\hat{\beta}_{3t}\) by least squares period by period, assuming a fixed value for \(\lambda\), and then use the estimated factors \(\hat{\beta}_{1t}, \hat{\beta}_{2t},\) and \(\hat{\beta}_{3t}\) to forecast future values of the factors. For example, the forecasts from the unconstrained VAR(1) model are computed using

\[
\hat{\beta}_{t+h|t} = \hat{c} + \hat{\Gamma}\hat{\beta}_t,
\]

where \(\hat{c}\) and \(\hat{\Gamma}\) are estimated by regressing \(\hat{\beta}_t\) on an intercept and \(\hat{\beta}_{t-k}\). Similarly, the forecasted factors from the univariate AR(1) models are computed using

\[
\hat{\beta}_{j,t+h|t} = \hat{c}_j + \hat{\gamma}_j\hat{\beta}_{jt}, \quad j = 1, 2, 3.
\]

Given the projected factors computed in Eqs. (2) and (3), the predicted term structure of yields is formed using

\[
y_{t+h|t}(\tau_t) = \hat{\beta}_{1,t+h|t} + \hat{\beta}_{2,t+h|t}\left(\frac{1-e^{-\lambda\tau_t}}{\lambda\tau_t}\right) + \hat{\beta}_{3,t+h|t}\left(1 - e^{-\lambda\tau_t} - e^{-\lambda\tau_t}\right). \tag{4}
\]

We follow this two-step method for forecast horizons of \(h = 1, 6, 12, 36,\) and 60 months. Following DRA, we fix \(\lambda = 0.077\) to simplify the computations.\(^3\) We note that a prudent choice of \(\lambda\) may be able to enhance the model’s forecasting ability. For example, Yu and Salyards (2009) found that there is a substantial difference in the optimal value of \(\lambda\) between investment-grade and speculative-grade bonds. Koopman, Mallee, and der Wel (in press) allowed \(\lambda\) to be time-varying by treating it as a fourth factor which is modeled jointly with the other factors. Because the main focus of this study is on evaluating the one-step state space approach relative to the two-step approach, we fix \(\lambda\), as was done by DL and DRA.

2.3. State space model—one-step approach

DRA argue that the two-step procedure used by DL suffers from the fact that the parameter estimation and signal extraction uncertainty associated with the first step are not considered in the second step. They propose a linear Gaussian state space approach which uses a one-step Kalman filter, a recursive procedure for computing the optimal estimator of the state vector at time \(t\) given the information available at time \(t\), to perform parameter estimation and signal extraction simultaneously in the dynamic Nelson-Siegel model. They suggest that the one-step approach is better than the two-step approach because the simultaneous estimation of all parameters produces correct inference via standard theory.

The measurement (observation) equation of the state space form of the dynamic Nelson-Siegel model

\(^{2}\) Some authors define the yield curve slope as \(y_t(\infty) - y_t(0)\), which (from Eq. (1)) equals \(\beta_{1t} - (\beta_{1t} + \beta_{3t}) = -\beta_{2t}\).

\(^{3}\) Using \(\lambda = 0.077\) implies that the loading on the curvature factor is maximized at a maturity of 23.3 months. If \(\lambda = 0.061\), the value used by DL, the curvature factor loading reaches its maximum at a maturity of 30 months.
predictions of the state vector, we can compute the out-of-sample forecasts from these equations by extending the data set with a set of missing values. For example, a sequence of \( h \) missing values at the end of the sample will produce a set of \( h \)-step-ahead forecasts.  

2.4. Data

We use the end-of-month zero-coupon bond rates of US Treasury bonds, and Standard and Poor’s rating AAA, AA, A+, A, A−, BB+, BB−, BB, and B− corporate bonds, taken from Bloomberg over the period December 1994 to April 2006. The maturities of bonds include 1–10, 15, 20, and 30 years. For a further description of this dataset, see Yu and Salyards (2009).

2.5. Dynamic Nelson-Siegel models—estimation results

For the two-step approach we first use least squares to estimate the level, slope, and curvature factors period by period using Eq. (1). These factors are used as data in the VAR(1) and AR(1) models in the second step. In the one-step approach, we estimate the state space model (Eqs. (6) and (8)) by maximum likelihood. Figs. 1 to 4, respectively, show the time series of estimated factors from the two approaches, as well as their common empirical proxies for the Treasury, A+ corporate, BB+ corporate, and B− corporate yields. The empirical counterpart of the level factor is the average of the short-term (1 year), medium-term (2 year), and long-term (10 year) yields: \( \gamma_1 + \gamma_2 + \gamma_{10}/3 \); the empirical counterpart of the slope factor is \( \gamma_1 - \gamma_{10} \); and the empirical counterpart of the curvature factor is \( 2 \times \gamma_2 - \gamma_1 - \gamma_{10} \).

Dewachter and Lyrio (2006) provide a macroeconomic interpretation of the latent factors of the Treasury yields: the level factor expresses the long-term inflation expectation; the slope factor represents business cycle conditions; and the curvature factor represents a clear independent monetary policy factor. For the slope factor series that we obtained from the AR(1)
and state space approaches, the Treasury and corporate bonds have similar dynamics, capturing real-time business cycle conditions. For the level factor series, Treasury and A+ corporate also have similar dynamics, which reconciles the real-time long-term inflation expectation. However, the level factor departs from the inflation expectation dynamics under the non-investment corporate bonds (BB+, and B−). Because of the surge of level factors between 1998 and 2003 for BB+ and B− since the 1998 Russian default, we argue that the level factors include the additional credit...
risk, which is asymmetrically proportional to non-investment grade bonds with a higher leverage ratio. The context is consistent with the findings of Yu and Salyards (2009), who argue that three factors are sufficient to explain the variation in all corporate yields in different ratings. Regarding the curvature factor series, Treasury, A+, and BB+ have similar dynamics, responding to independent monetary policy. The different pattern of the curvature factor of B− suggests that highly risky bonds are not sensitive to monetary policy.
3. Forecast evaluation

In this section, we assess the out-of-sample predictive accuracies of various models. We consider forecasts from models that fit to yields-only, as well as one model that utilizes additional macroeconomic information.

3.1. Yields-only models

In addition to the one-step and two-step dynamic Nelson-Siegel models, we also consider a subset of competing forecasting models used by DL. The yields-only models we consider are: Model 1: VAR(1) NS factor; Model 2: AR(1) NS factor; Model 3: state space NS factor; Model 4: random walk; and Model 5: VAR(1) on yield levels. The random walk model uses the current value as the forecast at all horizons. The VAR(1) on yield levels presumes that there are neither V AR(1) nor co-integrated series in the data. Instead of analyzing the in-sample interactions and variance decompositions of the yields-macro model, and analyze the impulse response functions to forecast future factors. We do not incorporate macro variables into Model 3 (state space NS factor), because we assume that the transition matrix $T$ is diagonal, and thus there will be no information gain by including macro variables. We also do not incorporate macro variables into Model 5 (VAR(1) on yield levels), for two reasons. First, we believe that macro variables will not improve the prediction accuracy of an unrestricted VAR(1) on yield levels, as was concluded by Ang and Piazzesi (2003). Second, Model 5 already has 13 variables for different maturities; therefore, adding three more variables will further reduce the degrees of freedom of an already high-dimensional VAR.

3.2. The yields-macro model

In addition to the yields-only models, we also consider Model 6 — a yield curve model with macroeconomic variables. It is well known that macro variables are related to the dynamics of yields curves, and their inclusion in yields-only models should improve the forecasts. For example, Ludvigson and Ng (2009) find that real and inflation macro variables have predictive power for future government bond yields. Ang and Piazzesi (2003) find that the 1-month-ahead out-of-sample VAR forecast performance of Treasury yields is improved when no-arbitrage restrictions are imposed and macro variables are incorporated. Cochrane and Piazzesi (2005) study the time-varying risk premia of government bonds and find an additional return-forecasting factor. DRA also incorporate three macroeconomic variables (manufacturing capacity utilization, the federal funds rate, and annual price inflation) into their state space model, and analyze the impulse response functions and variance decompositions of the yields-macro model. Instead of analyzing the in-sample interactions between yields and these macroeconomic variables, as DRA did, we focus on the out-of-sample forecast gain from including the macroeconomic variables.

Our Model 6 is a two-step model that integrates three macro variables into Model 1 (VAR(1) NS factor): the annual personal consumption expenditure inflation rate ($INFL$), the log of manufacturing capacity utilization ($CU$), and the effective federal funds rate ($FFR$). The first step is the same as in Model 1. The second step of Model 6 uses

$$\begin{bmatrix}
\beta_{1t+h|t} \\
\beta_{2t+h|t} \\
\beta_{3t+h|t} \\
CU_{t+h|t} \\
FFR_{t+h|t} \\
INFL_{t+h|t}
\end{bmatrix} = \begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\alpha_3 \\
\alpha_4 \\
\alpha_5 \\
\alpha_6
\end{bmatrix}$$

(10)

$$+ \begin{bmatrix}
\gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} & \gamma_{15} & \gamma_{16} \\
\gamma_{21} & \gamma_{22} & \gamma_{23} & \gamma_{24} & \gamma_{25} & \gamma_{26} \\
\gamma_{31} & \gamma_{32} & \gamma_{33} & \gamma_{34} & \gamma_{35} & \gamma_{36} \\
\gamma_{41} & \gamma_{42} & \gamma_{43} & \gamma_{44} & \gamma_{45} & \gamma_{46} \\
\gamma_{51} & \gamma_{52} & \gamma_{53} & \gamma_{54} & \gamma_{55} & \gamma_{56} \\
\gamma_{61} & \gamma_{62} & \gamma_{63} & \gamma_{64} & \gamma_{65} & \gamma_{66}
\end{bmatrix} \begin{bmatrix}
\beta_{1t} \\
\beta_{2t} \\
\beta_{3t} \\
CU_t \\
FFR_t \\
INFL_t
\end{bmatrix}$$

Macro data were obtained online from the Federal Reserve Economic Data (FRED) database.
to 2004:4 to forecast 2004:5, data from 1995:1 to 2004:5 to forecast 2004:6, and so on. For longer forecast horizons, we have to decrease the in-sample size in order to keep 24 out-of-sample predictions. For instance, to compute the 6-month-ahead forecast, we use data from 1994:12 to 2003:11, and our first prediction will be for 2004:5.

To evaluate the out-of-sample forecasting performances, we use the root mean squared forecast errors (RMSE),

$$\left[ \sum (y_{t+h} - y_{t+h|t})^2 / 24 \right]^{1/2},$$

where \(y_{t+h}\) is the realized yield and \(y_{t+h|t}^{\text{Model}}\) is the prediction made by the models. In Table 1, we show the evaluation results for 1-, 3-, 5-, 10-, and 30-year maturities for Treasury bonds. The bold numbers in the bottom row of each model represent their average RMSEs across the 1-, 3-, 5-, 10-, and 30-year maturities. To simplify the evaluation output for the nine corporate bonds in Tables 2 and 3, we only present the average RMSEs of the 1-, 3-, 5-, 10-, and 30-year maturities. For the last two columns in each table, we also present the average short-term RMSEs: (1-month-ahead + 6-month-ahead + 12-month-ahead RMSEs)/3; and the average long-term RMSEs: (36-month-ahead + 60-month-ahead RMSEs)/2.

3.4. Forecast results

First, consider the results for Treasury yields in Table 1. Model 2 (NS factor AR(1) model) has the lowest average short-term RMSE, 0.485, and the second lowest average long-term RMSE, 1.954. Model 4 (random walk) has the lowest long-term RMSE, 1.663. Table 2 illustrates similar results for AAA and A+ corporate yields. Model 2 produces the best short-term forecasts while Model 4 produces the best long-term forecasts. For AA corporate yields, Model 2 is the best for both short- and long-term forecasts. For the remaining corporate yields, the results are mixed, especially for short-term forecasts (1-, 6-, and 12-month-ahead forecasts). For the short-term forecast horizon, Model 5 (VAR(1) on yields level) is best for the A (RMSE: 0.569) and A- (RMSE: 0.513) corporate bonds. Table 3 shows that Model 4 is the best for the BB- (RMSE: 0.638) and B yields (RMSE: 0.560). Model 3 (NS factor state space approach) is superior for BB+ (RMSE: 0.556) and B- (RMSE: 0.739).

For the long-term forecast horizon (36 and 60 months ahead), Model 2 is the best for the BB+ (RMSE: 2.408), BB- (RMSE: 2.42), B (RMSE: 3.023), and B- (RMSE: 3.375) yields. 6

3.5. Short-term forecasts for the one-step and two-step approaches

Our results show that a model which has a good in-sample fit may not necessarily produce the best out-of-sample forecasts. In addition, when the credit risks (spreads) increase, the NS factor AR(1) short-term forecast performance deteriorates and the state space model performance improves. The average short-term RMSE increases from 0.485 for Treasury yields to 1.056 for B- yields. Although Model 3 is always suboptimal for forecasting all types of investment-grade bonds, its short-term forecasts are very stable for non-investment grade bonds. Its lowest RMSEs are for BB+ (0.556) and B- (0.739) yields, while it is the second lowest for the A- (0.562), BB- (0.562), and B (0.602) yields. In general, Model 3 provides the most reliable out-of-sample forecasts for high-risk bonds.

A possible explanation for why the state space model predicts better on high-risk yields while the AR(1) model predicts better on investment-grade yields is as follows. The state space model, which was designed for time-varying parameters, is most robust in the environment of parameter instability, nonlinearity, and non-normality for which speculative bonds are most likely to exist. As Metz and Cantor (2006) point out, although the distribution of bond ratings is fairly stable over time, the relationship between the ratings and their corresponding credit factors (such as interest coverage) is not stable over time, particularly for speculative bonds. In addition, Cantor and Packer (1996) argue that rating announcements have immediate effects on market pricings for speculative bonds, leading to more unstable dynamics. In the context of the term structure theory, Dai and Singleton (2003) summarize that one would use the diffusion process for modeling.

---

6 Finally, it is worth noting that Model 1 (NS factor VAR(1)) is inferior to Model 2 (NS factor AR(1)) for all kinds of bonds because the three factors (level, slope, curvature) constructed by exponential components are supposed to be uncorrelated. Consequently, this result supports our assumption regarding Model 3, in which the transition matrix \(T\) is diagonal.
Table 1
Treasury yields out-of-sample forecast evaluation: Root Mean Squared Errors (RMSEs).

<table>
<thead>
<tr>
<th>Maturities</th>
<th>1-month-ahead</th>
<th>6-month-ahead</th>
<th>12-month-ahead</th>
<th>36-month-ahead</th>
<th>60-month-ahead</th>
<th>Average 1, 6, 12 m</th>
<th>Average 36, 60 m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>1 year</td>
<td>0.230</td>
<td>0.904</td>
<td>1.903</td>
<td>11.977</td>
<td>2.526</td>
<td>1.012</td>
</tr>
<tr>
<td>NS factors</td>
<td>3 year</td>
<td>0.352</td>
<td>0.814</td>
<td>1.442</td>
<td>6.852</td>
<td>1.946</td>
<td>0.869</td>
</tr>
<tr>
<td>VAR(1)</td>
<td>5 year</td>
<td>0.278</td>
<td>0.381</td>
<td>0.678</td>
<td>4.086</td>
<td>1.835</td>
<td>0.446</td>
</tr>
<tr>
<td></td>
<td>10 year</td>
<td>0.248</td>
<td>0.331</td>
<td>0.408</td>
<td>2.226</td>
<td>1.417</td>
<td>0.329</td>
</tr>
<tr>
<td></td>
<td>30 year</td>
<td>0.318</td>
<td>0.649</td>
<td>0.931</td>
<td>2.653</td>
<td>1.413</td>
<td>0.633</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td><strong>0.285</strong></td>
<td><strong>0.616</strong></td>
<td><strong>1.072</strong></td>
<td><strong>5.559</strong></td>
<td><strong>1.827</strong></td>
<td><strong>0.658</strong></td>
</tr>
<tr>
<td>Model 2</td>
<td>1 year</td>
<td>0.217</td>
<td>0.519</td>
<td>0.786</td>
<td>4.062</td>
<td>2.524</td>
<td>0.507</td>
</tr>
<tr>
<td>NS factors</td>
<td>3 year</td>
<td>0.292</td>
<td>0.400</td>
<td>0.448</td>
<td>2.207</td>
<td>2.010</td>
<td>0.380</td>
</tr>
<tr>
<td>AR(1)</td>
<td>5 year</td>
<td>0.291</td>
<td>0.393</td>
<td>0.437</td>
<td>1.554</td>
<td>1.908</td>
<td>0.374</td>
</tr>
<tr>
<td></td>
<td>10 year</td>
<td>0.257</td>
<td>0.464</td>
<td>0.630</td>
<td>1.085</td>
<td>1.480</td>
<td>0.450</td>
</tr>
<tr>
<td></td>
<td>30 year</td>
<td>0.335</td>
<td>0.732</td>
<td>1.078</td>
<td>1.256</td>
<td>1.457</td>
<td>0.715</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td><strong>0.278</strong></td>
<td><strong>0.502</strong></td>
<td><strong>0.676</strong></td>
<td><strong>2.033</strong></td>
<td><strong>1.876</strong></td>
<td><strong>0.485</strong></td>
</tr>
<tr>
<td>Model 3</td>
<td>1 year</td>
<td>0.695</td>
<td>0.481</td>
<td>1.108</td>
<td>6.622</td>
<td>1.792</td>
<td>0.761</td>
</tr>
<tr>
<td>NS factors</td>
<td>3 year</td>
<td>0.417</td>
<td>0.975</td>
<td>1.726</td>
<td>6.956</td>
<td>1.223</td>
<td>1.039</td>
</tr>
<tr>
<td>State space</td>
<td>5 year</td>
<td>0.296</td>
<td>0.625</td>
<td>1.180</td>
<td>4.992</td>
<td>0.953</td>
<td>0.700</td>
</tr>
<tr>
<td></td>
<td>10 year</td>
<td>0.257</td>
<td>0.432</td>
<td>0.752</td>
<td>2.644</td>
<td>0.714</td>
<td>0.480</td>
</tr>
<tr>
<td></td>
<td>30 year</td>
<td>0.432</td>
<td>0.375</td>
<td>0.650</td>
<td>0.749</td>
<td>0.650</td>
<td>0.486</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td><strong>0.420</strong></td>
<td><strong>0.578</strong></td>
<td><strong>1.083</strong></td>
<td><strong>4.392</strong></td>
<td><strong>1.066</strong></td>
<td><strong>0.693</strong></td>
</tr>
<tr>
<td>Model 4</td>
<td>1 year</td>
<td>0.263</td>
<td>0.887</td>
<td>1.544</td>
<td>2.170</td>
<td>2.761</td>
<td>0.898</td>
</tr>
<tr>
<td>Random walk</td>
<td>3 year</td>
<td>0.281</td>
<td>0.633</td>
<td>1.087</td>
<td>1.561</td>
<td>2.294</td>
<td>0.667</td>
</tr>
<tr>
<td></td>
<td>5 year</td>
<td>0.282</td>
<td>0.484</td>
<td>0.708</td>
<td>1.128</td>
<td>2.182</td>
<td>0.492</td>
</tr>
<tr>
<td></td>
<td>10 year</td>
<td>0.248</td>
<td>0.356</td>
<td>0.506</td>
<td>0.760</td>
<td>1.662</td>
<td>0.370</td>
</tr>
<tr>
<td></td>
<td>30 year</td>
<td>0.207</td>
<td>0.477</td>
<td>0.755</td>
<td>0.900</td>
<td>1.211</td>
<td>0.480</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td><strong>0.256</strong></td>
<td><strong>0.568</strong></td>
<td><strong>0.920</strong></td>
<td><strong>1.303</strong></td>
<td><strong>2.022</strong></td>
<td><strong>0.581</strong></td>
</tr>
<tr>
<td>Model 5</td>
<td>1 year</td>
<td>0.275</td>
<td>0.949</td>
<td>1.880</td>
<td>38.387</td>
<td>2.880</td>
<td>1.035</td>
</tr>
<tr>
<td>Yields level</td>
<td>3 year</td>
<td>0.331</td>
<td>0.709</td>
<td>1.247</td>
<td>29.743</td>
<td>2.442</td>
<td>0.762</td>
</tr>
<tr>
<td>VAR(1)</td>
<td>5 year</td>
<td>0.327</td>
<td>0.499</td>
<td>0.793</td>
<td>26.793</td>
<td>2.556</td>
<td>0.540</td>
</tr>
<tr>
<td></td>
<td>10 year</td>
<td>0.304</td>
<td>0.425</td>
<td>0.483</td>
<td>16.127</td>
<td>1.637</td>
<td>0.404</td>
</tr>
<tr>
<td></td>
<td>30 year</td>
<td>0.244</td>
<td>0.648</td>
<td>0.905</td>
<td>2.869</td>
<td>1.207</td>
<td>0.599</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td><strong>0.297</strong></td>
<td><strong>0.646</strong></td>
<td><strong>1.062</strong></td>
<td><strong>22.784</strong></td>
<td><strong>2.145</strong></td>
<td><strong>0.668</strong></td>
</tr>
<tr>
<td>Model 6</td>
<td>1 year</td>
<td>0.216</td>
<td>0.541</td>
<td>0.632</td>
<td>3.735</td>
<td>3.048</td>
<td>0.463</td>
</tr>
<tr>
<td>NS factors +</td>
<td>3 year</td>
<td>0.289</td>
<td>0.545</td>
<td>0.685</td>
<td>3.084</td>
<td>2.479</td>
<td>0.506</td>
</tr>
<tr>
<td>Macvo variables</td>
<td>5 year</td>
<td>0.313</td>
<td>0.628</td>
<td>0.888</td>
<td>2.683</td>
<td>2.198</td>
<td>0.610</td>
</tr>
<tr>
<td>VAR(1)</td>
<td>10 year</td>
<td>0.273</td>
<td>0.645</td>
<td>0.970</td>
<td>1.980</td>
<td>1.743</td>
<td>0.630</td>
</tr>
<tr>
<td></td>
<td>30 year</td>
<td>0.330</td>
<td>0.804</td>
<td>1.289</td>
<td>1.813</td>
<td>1.790</td>
<td>0.808</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td><strong>0.284</strong></td>
<td><strong>0.632</strong></td>
<td><strong>0.893</strong></td>
<td><strong>2.659</strong></td>
<td><strong>2.251</strong></td>
<td><strong>0.603</strong></td>
</tr>
</tbody>
</table>

1. The table reports the out-of-sample forecast evaluation based on the root mean squared error (RMSE).
3. RMSE = \[ \frac{\sum (y_{t+h} - \hat{y}_{t+h|t})^2}{24} \]^{1/2}, where \( y_{t+h} \) is the realized data and \( \hat{y}_{t+h|t} \) is the predicted data.
4. The table evaluates the 1-, 6-, 12-, 36-, and 60-month-ahead forecasts.
5. The average RMSE for 1, 6, and 12 m is the mean of the (1 m + 6 m + 12 m) RMSEs, while the average RMSE for 36, 60 m is the mean of the (36 m + 60 m) RMSEs.

default-free bond yields, but the jump-diffusion process for modeling defaultable bond yields. For example, Fig. 4 shows that the level factor of the bond yield exhibits a steep structural shift (jump) from 18 to 10 in early 2003. It would therefore follow to conclude that the state space model is more capable of a jump-diffusion data generating process (DGP), while the AR(1) is superior in estimating a diffusion DGP.

Moreover, Yu and Salyards (2009) use the principal component method and find a dichotomy of factor
### Table 2
Investment-grade corporate yields out-of-sample forecast evaluation: Root Mean Squared Errors (RMSEs).

<table>
<thead>
<tr>
<th>Bond rating</th>
<th>Model</th>
<th>1-month-ahead</th>
<th>6-month-ahead</th>
<th>12-month-ahead</th>
<th>36-month-ahead</th>
<th>60-month-ahead</th>
<th>Average 1, 6, 12 m</th>
<th>Average 36, 60 m</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>Model 1</td>
<td>0.567</td>
<td>0.879</td>
<td>1.233</td>
<td>4.209</td>
<td>2.105</td>
<td>0.893</td>
<td>3.157</td>
</tr>
<tr>
<td></td>
<td>Model 2</td>
<td>0.271</td>
<td>0.568</td>
<td>0.761</td>
<td>1.878</td>
<td>1.960</td>
<td><strong>0.533</strong></td>
<td>1.919</td>
</tr>
<tr>
<td></td>
<td>Model 3</td>
<td>0.230</td>
<td>0.627</td>
<td>1.117</td>
<td>3.688</td>
<td>3.151</td>
<td>0.658</td>
<td>2.520</td>
</tr>
<tr>
<td></td>
<td>Model 4</td>
<td>0.232</td>
<td>0.576</td>
<td>0.948</td>
<td>1.420</td>
<td>2.392</td>
<td>0.585</td>
<td><strong>1.906</strong></td>
</tr>
<tr>
<td></td>
<td>Model 5</td>
<td>0.247</td>
<td>0.525</td>
<td>0.960</td>
<td>28.323</td>
<td>1.935</td>
<td>0.577</td>
<td>15.129</td>
</tr>
<tr>
<td></td>
<td>Model 6</td>
<td>0.294</td>
<td>0.672</td>
<td>0.951</td>
<td>2.744</td>
<td>2.910</td>
<td>0.639</td>
<td>2.827</td>
</tr>
<tr>
<td></td>
<td>Model 7</td>
<td>0.245</td>
<td>0.563</td>
<td>1.027</td>
<td>4.703</td>
<td>1.892</td>
<td>0.612</td>
<td>3.298</td>
</tr>
<tr>
<td></td>
<td>Model 8</td>
<td>0.257</td>
<td>0.553</td>
<td>0.762</td>
<td>1.629</td>
<td>1.924</td>
<td><strong>0.524</strong></td>
<td><strong>1.776</strong></td>
</tr>
<tr>
<td>AA</td>
<td>Model 1</td>
<td>0.245</td>
<td>0.563</td>
<td>1.027</td>
<td>4.703</td>
<td>1.892</td>
<td>0.612</td>
<td>3.298</td>
</tr>
<tr>
<td></td>
<td>Model 2</td>
<td>0.257</td>
<td>0.553</td>
<td>0.762</td>
<td>1.629</td>
<td>1.924</td>
<td><strong>0.524</strong></td>
<td><strong>1.776</strong></td>
</tr>
<tr>
<td></td>
<td>Model 3</td>
<td>0.252</td>
<td>0.649</td>
<td>1.091</td>
<td>3.244</td>
<td>1.482</td>
<td>0.664</td>
<td>2.363</td>
</tr>
<tr>
<td></td>
<td>Model 4</td>
<td>0.230</td>
<td>0.584</td>
<td>0.935</td>
<td>1.432</td>
<td>2.388</td>
<td>0.583</td>
<td>1.910</td>
</tr>
<tr>
<td></td>
<td>Model 5</td>
<td>0.220</td>
<td>0.494</td>
<td>0.959</td>
<td>5.746</td>
<td>1.905</td>
<td>0.558</td>
<td>3.825</td>
</tr>
<tr>
<td></td>
<td>Model 6</td>
<td>0.288</td>
<td>0.675</td>
<td>0.922</td>
<td>2.918</td>
<td>3.985</td>
<td>0.628</td>
<td>3.452</td>
</tr>
<tr>
<td></td>
<td>Model 7</td>
<td>0.575</td>
<td>0.899</td>
<td>1.219</td>
<td>2.084</td>
<td>2.165</td>
<td>0.898</td>
<td>2.124</td>
</tr>
<tr>
<td></td>
<td>Model 8</td>
<td>0.253</td>
<td>0.577</td>
<td>0.808</td>
<td>2.946</td>
<td>1.981</td>
<td><strong>0.546</strong></td>
<td>2.464</td>
</tr>
<tr>
<td>A+</td>
<td>Model 1</td>
<td>0.253</td>
<td>0.665</td>
<td>1.152</td>
<td>4.815</td>
<td>1.530</td>
<td>0.690</td>
<td>3.172</td>
</tr>
<tr>
<td></td>
<td>Model 2</td>
<td>0.253</td>
<td>0.665</td>
<td>1.152</td>
<td>4.815</td>
<td>1.530</td>
<td>0.690</td>
<td>3.172</td>
</tr>
<tr>
<td></td>
<td>Model 3</td>
<td>0.218</td>
<td>0.590</td>
<td>0.933</td>
<td>1.442</td>
<td>2.483</td>
<td>0.581</td>
<td><strong>1.962</strong></td>
</tr>
<tr>
<td></td>
<td>Model 4</td>
<td>0.227</td>
<td>0.498</td>
<td>0.982</td>
<td>11.587</td>
<td>2.251</td>
<td>0.569</td>
<td>6.919</td>
</tr>
<tr>
<td></td>
<td>Model 5</td>
<td>0.227</td>
<td>0.498</td>
<td>0.982</td>
<td>11.587</td>
<td>2.251</td>
<td>0.569</td>
<td>6.919</td>
</tr>
<tr>
<td></td>
<td>Model 6</td>
<td>0.284</td>
<td>0.670</td>
<td>0.910</td>
<td>2.964</td>
<td>3.435</td>
<td>0.621</td>
<td>3.199</td>
</tr>
<tr>
<td></td>
<td>Model 7</td>
<td>0.246</td>
<td>0.585</td>
<td>1.014</td>
<td>3.125</td>
<td>2.019</td>
<td>0.615</td>
<td>2.572</td>
</tr>
<tr>
<td></td>
<td>Model 8</td>
<td>0.256</td>
<td>0.615</td>
<td>0.888</td>
<td>2.275</td>
<td>2.042</td>
<td>0.586</td>
<td>2.158</td>
</tr>
<tr>
<td>A</td>
<td>Model 1</td>
<td>0.249</td>
<td>0.664</td>
<td>1.121</td>
<td>3.375</td>
<td>1.566</td>
<td>0.678</td>
<td>2.470</td>
</tr>
<tr>
<td></td>
<td>Model 2</td>
<td>0.219</td>
<td>0.585</td>
<td>0.921</td>
<td>1.442</td>
<td>2.559</td>
<td>0.575</td>
<td>2.001</td>
</tr>
<tr>
<td></td>
<td>Model 3</td>
<td>0.242</td>
<td>0.544</td>
<td>0.920</td>
<td>1.644</td>
<td>2.175</td>
<td><strong>0.569</strong></td>
<td><strong>1.909</strong></td>
</tr>
<tr>
<td></td>
<td>Model 4</td>
<td>0.284</td>
<td>0.668</td>
<td>0.895</td>
<td>2.960</td>
<td>3.106</td>
<td>0.616</td>
<td>3.033</td>
</tr>
<tr>
<td></td>
<td>Model 5</td>
<td>0.564</td>
<td>0.891</td>
<td>1.220</td>
<td>5.991</td>
<td>2.262</td>
<td>0.892</td>
<td>4.127</td>
</tr>
<tr>
<td></td>
<td>Model 6</td>
<td>0.253</td>
<td>0.646</td>
<td>0.927</td>
<td>4.040</td>
<td>2.049</td>
<td>0.609</td>
<td>3.044</td>
</tr>
<tr>
<td>A−</td>
<td>Model 1</td>
<td>0.213</td>
<td>0.574</td>
<td>0.899</td>
<td>1.420</td>
<td>2.597</td>
<td>0.562</td>
<td><strong>2.008</strong></td>
</tr>
<tr>
<td></td>
<td>Model 2</td>
<td>0.251</td>
<td>0.650</td>
<td>1.063</td>
<td>3.783</td>
<td>1.834</td>
<td>0.655</td>
<td>2.808</td>
</tr>
<tr>
<td></td>
<td>Model 3</td>
<td>0.227</td>
<td>0.469</td>
<td>0.843</td>
<td>4.300</td>
<td>20.295</td>
<td><strong>0.513</strong></td>
<td>12.297</td>
</tr>
<tr>
<td></td>
<td>Model 4</td>
<td>0.261</td>
<td>0.624</td>
<td>0.816</td>
<td>2.741</td>
<td>2.956</td>
<td>0.567</td>
<td>2.848</td>
</tr>
</tbody>
</table>

1. See the notes to Table 1. Model 1: NS factors VAR(1); Model 2: NS factors AR(1); Model 3: NS factors state space; Model 4: Random walk; Model 5: Yields level VAR(1); Model 6: NS factors + macroeconomic variables VAR(1).

2. Unlike Table 1, to simplify the output, the RMSE in each cell represents the average of the 1, 3, 5, 10, and 30 year maturities of each investment-grade corporate bond.

Loadings between investment-grade and speculative-grade yields. This finding supports multiple optimal solutions for $\lambda$ across different bond ratings. If the model was purposely constrained (e.g., for simplicity we fix $\lambda$ in this study), then the state space approach will mitigate the loss of inferior model selections.

In general, we find that the risky corporate bonds are not much harder to predict than Treasury bonds if we choose an appropriate model. Model 2 is preferred to Model 3 for investment-grade bonds. However, Model 3 is robust for non-investment-grade bonds at short-term forecast horizons.

### 3.6. Long-term forecasts from the one-step and two-step approaches

It is surprising to see that Model 2 is superior for the long-term forecasting of BB+, BB−, B, and
Table 3
Non-investment-grade corporate yields out-of-sample forecast evaluation: Root Mean Squared Errors (RMSEs).

<table>
<thead>
<tr>
<th></th>
<th>RMSE 1-month-ahead</th>
<th>RMSE 6-month-ahead</th>
<th>RMSE 12-month-ahead</th>
<th>RMSE 36-month-ahead</th>
<th>RMSE 60-month-ahead</th>
<th>Average 1, 6, 12 m</th>
<th>Average 36, 60 m</th>
</tr>
</thead>
<tbody>
<tr>
<td>BB+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>0.287</td>
<td>0.572</td>
<td>0.890</td>
<td>2.320</td>
<td>27.319</td>
<td>0.583</td>
<td>14.820</td>
</tr>
<tr>
<td>Model 2</td>
<td>0.320</td>
<td>0.876</td>
<td>1.465</td>
<td>2.432</td>
<td>2.385</td>
<td>0.887</td>
<td>2.408</td>
</tr>
<tr>
<td>Model 3</td>
<td>0.286</td>
<td>0.567</td>
<td>0.816</td>
<td>2.397</td>
<td>3.094</td>
<td>0.556</td>
<td>2.746</td>
</tr>
<tr>
<td>Model 4</td>
<td>0.249</td>
<td>0.589</td>
<td>0.910</td>
<td>2.136</td>
<td>3.298</td>
<td>0.583</td>
<td>2.717</td>
</tr>
<tr>
<td>Model 5</td>
<td>0.262</td>
<td>0.600</td>
<td>0.922</td>
<td>2.966</td>
<td>13899</td>
<td>0.594</td>
<td>6950.9</td>
</tr>
<tr>
<td>Model 6</td>
<td>0.311</td>
<td>0.727</td>
<td>0.820</td>
<td>5.795</td>
<td>24.045</td>
<td>0.619</td>
<td>14.920</td>
</tr>
</tbody>
</table>

| BB−       |                    |                    |                     |                    |                     |                   |                  |
| Model 1   | 0.741              | 0.968              | 1.370               | 2.516              | 208.08              | 1.026             | 105.30           |
| Model 2   | 0.390              | 0.994              | 1.657               | 2.539              | 2.301               | 1.014             | 3.054            |
| Model 3   | 0.347              | 0.630              | 0.978               | 2.639              | 3.469               | 0.652             | 2.420            |
| Model 4   | 0.366              | 0.644              | 0.905               | 2.677              | 2.757               | 0.638             | 2.717            |
| Model 5   | 0.337              | 0.657              | 1.037               | 3.053              | 72.486              | 0.677             | 37.770           |
| Model 6   | 0.361              | 0.668              | 0.929               | 6.198              | 32.881              | 0.652             | 19.539           |

| B         |                    |                    |                     |                    |                     |                   |                  |
| Model 1   | 0.586              | 1.052              | 1.500               | 2.741              | 12.701              | 1.046             | 7.721            |
| Model 2   | 0.338              | 0.955              | 1.818               | 3.092              | 2.955               | 1.037             | 3.023            |
| Model 3   | 0.349              | 0.634              | 0.823               | 3.233              | 3.760               | 0.602             | 3.497            |
| Model 4   | 0.300              | 0.586              | 0.793               | 3.637              | 4.006               | 0.560             | 3.822            |
| Model 5   | 0.336              | 0.826              | 1.503               | 170.71             | 5002.4              | 0.888             | 2586.5           |
| Model 6   | 0.340              | 0.666              | 0.881               | 5.759              | 17.279              | 0.629             | 11.519           |

| B−        |                    |                    |                     |                    |                     |                   |                  |
| Model 1   | 0.546              | 1.202              | 2.081               | 3.221              | 5.119               | 1.276             | 4.170            |
| Model 2   | 0.402              | 0.880              | 1.885               | 3.694              | 3.056               | 1.056             | 3.375            |
| Model 3   | 0.383              | 0.696              | 1.137               | 4.291              | 4.565               | 0.739             | 4.428            |
| Model 4   | 0.512              | 0.860              | 1.346               | 4.657              | 5.317               | 0.906             | 4.987            |
| Model 5   | 0.445              | 1.114              | 2.462               | 5.198              | 3.451               | 1.340             | 4.325            |
| Model 6   | 0.418              | 0.729              | 1.234               | 4.014              | 15.207              | 0.794             | 9.611            |

1. See the notes to Tables 1 and 2.

B− high-yield bonds. Although Model 2 has the lowest RMSE for AA yields among investment-grade yields, it is the second best model for Treasury, AAA, and A+ yields for long-term forecasts. In contrast to the reliable forecasting ability of Model 3 at the short-term horizon, Model 2 provides more reliable long-term forecasts. A possible reason for this dichotomy is that the long-term forecasts of yields weigh more in the steady state equilibrium level than short-term forecasts. The simple AR(1) model assumes that time series are stationary, and therefore long-term forecasts revert to long-run levels. On the other hand, the state space approach is based on the unified structural analysis and recursive one-step-ahead in-sample forecast updates from the Kalman filter. It is supposed to forecast well out-of-sample in the short-run, especially for complex data-generating structures, such as high-yield corporate bonds. The signal extracted from updating data via the state space approach is less informative for the long-term horizons.

3.7. Forecasts from the yields-macro model

We find that introducing macroeconomic variables into the yields level VAR improves some yield curve forecasts. That is, Model 6 outperforms Model 1 on the Treasury yields and many of the corporate yields. For short-term forecasts, Model 6’s RMSEs are lower than Model 1’s on Treasury (0.603 vs. 0.658), AAA (0.639 vs. 0.893), A+ (0.621 vs. 0.898), A− (0.567 vs. 0.892), BB− (0.652 vs. 1.026), B (0.629 vs. 1.046), and B− (0.794 vs. 1.276) yields. Model 6’s RMSEs are similar to Model 1’s on AA (0.628 vs. 0.612) and A (0.616 vs. 0.615). Finally, Model 6’s RMSE is higher than Model 1’s on BB+ only (0.619 vs. 0.583). For long-term forecasts, the benefits of using additional macroeconomic variables are not as clear cut. Model 6’s RMSEs are lower than Model 1’s on Treasury.
(2.455 vs. 3.693), AAA (2.827 vs. 3.157), A− (2.848 vs. 4.127), and BB− (19.539 vs. 105.3) yields, but higher on A+ (3.199 vs. 2.124), A (3.033 vs. 2.572), B (11.519 vs. 7.721), and B− (9.611 vs. 4.170) yields. This result is consistent with the findings of Dewachter and Lyrio (2006), who provide a macroeconomic interpretation for the latent factors of yield curves. They argue that the level factor represents the long-run inflation expectation of agents. In other words, in the context of long-term out-of-sample forecasting, macroeconomic variables have been incorporated into the latent factors well, especially the level factor. Consequently, there will be little information gain and forecast improvement from including macroeconomic variables.

3.8. Forecast improvements

Some of the long-term forecast errors, especially for 36 months ahead, are extraordinarily large. For example, in Table 1, Model 1’s 1- and 3-year RMSEs at the 36-month-ahead horizon are 11.9777 and 6.852, respectively. There are three possible reasons for these extreme forecasts: (1) the model parameters exhibit structural breaks during the out-of-sample periods; (2) there are some outliers in the yield data; or (3) the sample size is small.

Stock and Watson (1999) point out that “crazy” forecasts would be noticed and adjusted by human judgment. To mimic a forecast adjustment procedure, we use an interpolation method to replace extreme forecasts (defined as values over some absolute number for 36-month-ahead forecasts). Using such a method, we can reduce the 36-month-ahead RMSEs of Model 1 substantially. For the 1-, 3-, 5-, 10-, and 30-year bonds, the RMSEs drop from 11.977, 6.852, 4.086, 2.226, and 2.653 to 2.148, 1.691, 1.685, 1.470, and 1.655, respectively.

The interpolation method is ad hoc and requires human intervention. As an alternative, we consider a purely statistical approach based on robust regression methods to detect and adjust for potential outliers and level shifts. Using such a method, the average 36-month-ahead RMSE of Model 1 for AAA bonds, which was 4.209 using non-robust methods, declines to 2.143 using the robust method. Although the robust method reduced the RMSEs of these forecasts, it increased the RMSEs of the shorter-term forecasts. For example, the 1-month-ahead RMSE rises from 0.567 to 1.068. This increase occurs because the robust method modifies some data information as the outliers and level shifts are adjusted, which is useful for short-term forecasts. The ad hoc interpolation and robust methods can be applied to Models 1, 2, 3, 5, and 6, and were found to beat the random walk model, which produces the best long-term forecasts of Treasury (RMSE: 1.663), AAA (RMSE: 1.906), and A+ (RMSE: 1.962) yields. Even with the use of these methods, the conclusions we reach regarding the models’ competitiveness remain the same.

4. Conclusions

We provide a comprehensive short- and long-term forecasting evaluation of the two-step and one-step approaches of Diebold and Li (2006) and Diebold et al. (2006) using Treasury yields and nine different ratings of corporate bonds. We find that the simple two-step dynamic Nelson-Siegel factor AR(1) model survives a long list of robustness checks on the out-of-sample forecast accuracy, especially for investment-grade bonds at short-term horizons and for high yield bonds at long-term horizons. The one-step dynamic Nelson-Siegel factor state space model has the best performance for high yield bonds at short-term horizons, in which credit risks have more influence on the yields dynamics, especially in the level factor, and model parameter instability is more likely to exist. In short, the state space model is more capable of estimating a jump-diffusion DGP of defaultable bond yields, while the AR(1) is superior for estimating a diffusion DGP of default-free bond yields. We also show that forecasts from the Nelson-Siegel factor VAR(1) model can be improved by incorporating macroeconomic variables. Nevertheless, this yield-macro model remains suboptimal compared with the parsimonious NS factors AR(1) model for investment-grade yields in the short run.

Our findings suggest that for real-time forecasts, the multivariate state space model computed by the

---

7 For details, see Zivot and Wang (2006, Chapter 17).

8 If the random walk model dominates, this implies that the other forecasting methods are useless.
Kalman filter is not necessarily better than the simple univariate AR(1) model computed by least squares. The preferred model depends on the nature of the data and the forecast horizons. In the context of the Kalman filter, which updates the mean and covariance matrix of the conditional distribution of the state vector as new observations become available, the state space model can predict more accurately if the data’s implied state or hyperparameter is more likely to be time-varying and unstable.

Acknowledgements

We would like to thank Frank Russell Investment Company for providing the data. We would also like to thank the editor, two anonymous referees, Tom Goodwin, Steve Murray, Patrick Rowland and Michael Wenz for their helpful comments and discussions. We are grateful to Christina Flaherty for providing assistance.

References


Wei-Choun Yu is an Assistant Professor in the Department of Economics at Winona State University. He also serves on the board of directors at the Minnesota Economic Association. He received his Ph.D. in economics at the University of Washington in 2006. Yu’s research interests are in time series econometrics, finance and macroeconomics, and in particular linear, nonlinear, and volatility modeling and forecasting. He has publications in various journals in the fields of econometrics and finance, including the *Journal of International Money and Finance, Journal of Forecasting, Applied Economics Letters, Global Economic Review and Economic Affairs.*

Eric Zivot is a Professor and Gary Waterman Distinguished Scholar in the Department of Economics, an Adjunct Professor of Statistics, and an Adjunct Professor of Finance at the University of Washington. He regularly teaches courses on econometric theory, financial econometrics and time series econometrics, and is a recipient of the Henry T. Buechel Award for Outstanding Teaching. He is an associate editor of *Studies in Nonlinear Dynamics and Econometrics*. He has published papers in the leading econometrics journals, including *Econometrica, Econometric Theory, Journal of Business and Economic Statistics, Journal of Econometrics, and Review of Economics and Statistics.*