A regime-switching Nelson–Siegel term structure model of the macroeconomy

Xiaoneng Zhu a,⇑, Shahidur Rahman b

a School of Finance, Shanghai University of Finance and Economics, China
b Kazakhstan University, Kazakhstan

ARTICLE INFO

Article history:
Received 22 May 2014
Accepted 23 December 2014
Available online 11 February 2015

JEL classification:
E43
E44
G12

Keywords:
Economic activity
Macro factors
MCMC
Nelson–Siegel model
The term structure

ABSTRACT

This paper presents a regime-switching Nelson–Siegel term structure model with macro factors and introduces a Markov chain Monte Carlo procedure to estimate the model. We find that regime shifts are important for understanding the interaction between the yield curve and economic activity. We also find that two regimes are closely related to the business cycle and monetary policy. Finally, we find that the proposed regime-switching model with macro factors is competitive in the out-of-sample forecasting of bond yields.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Understanding the joint dynamics of the yield curve and economic activity is important for bond pricing, monetary policy, and fiscal policy. To date, standard term structure models with macro factors successfully describe the joint dynamics of the yield curve and macro factors. 1 On the one hand, macro factors sharpen our understanding on the economic driving force of yield curve movements (e.g., Ang and Piazzesi, 2003; Joslin et al., 2014; Kozicki and Tinsley, 2001; Cooper and Priestly, 2009; Carriero et al., 2006; Wright, 2011) and improve the out-of-sample forecasting performance of term structure models (e.g., Ang and Piazzesi, 2003; Guidolin and Timmermann, 2009; Ludvigson and Ng, 2009; Hordahl et al., 2006). On the other hand, since interest rate movements contain important information about economic activity, term structure models with macro factors illuminate various macroeconomic issues (e.g., Bekaert et al., 2010; Rudebusch and Wu, 2003) and produce superior out-of-sample GDP growth forecasts (e.g., Ang et al., 2006; Estrella and Hardouvelis, 1991).

This paper proposes a regime-switching Nelson–Siegel term structure model with macro factors (RSDNS-X). We build the RSDNS-X model upon the dynamic Nelson–Siegel (DNS) model proposed by Nelson and Siegel (1987) and reinterpreted by Diebold and Li (2006, DL hereafter) and Diebold et al. (2006, DRA hereafter). We also build the RSDNS-X model upon the

⇑ Corresponding author at: School of Finance, Shanghai University of Finance and Economics, Guoding Road 777, Shanghai 200433, China. Tel.: +86 21 65908403; fax: +86 21 65109325.

E-mail address: xiaonengz@gmail.com (X. Zhu).

1 Duffee (2013) reviews the academic literature that links nominal and real term structures with macroeconomic fundamentals.
regime-switching Nelson–Siegel model without macro factors (RSDNS) proposed by Xiang and Zhu (2013) and other regime-switching term structure models.²

We contribute to the literature in three ways. First, the RSDNS-X model may sharpen our understanding on the joint dynamics of macro factors and the yield curve. Though a large literature has investigated the joint dynamics of the term structure and economic activity, the regime-switching relationship between interest rates and macroeconomic fundamentals is not explicitly accounted for. Indeed, economic theories (e.g., Buraschi and Jiltsov, 2007; Wachter, 2006) generally suggest that the yield curve should vary with macroeconomic fundamentals. Naturally, shifts between distinct economic states or monetary policy regimes might lead to the regime-switching interaction between the yield curve and macro factors.³ This article thus fills the gap and provides a characterization of the regime-dependent interaction between the term structure and the macroeconomy. Second, the RSDNS-X can provide new insights on the economic driving force of regime shifts by directly incorporating macro factors in a regime-switching term structure model. Third, the RSDNS-X model can shed light on the role of regime shifts and macro factors in interest rate forecasts.

The RSDNS-X model provides greater flexibility for understanding interactions between the yield curve and the macroeconomy. However, statistical inference of RSDNS-X is complicated by two unobserved components: latent regimes and latent yield factors. Toward this end, we develop a Bayesian Markov chain Monte Carlo approach to estimate the RSDNS-X model. The Bayesian approach is a natural way of taking parameter uncertainty into consideration. In particular, we use the Kalman filter to extract yield factors. Moreover, we employ the Hamilton (1994) filter to construct regime probabilities. The MCMC method allows us to estimate the RSDNS-X model efficiently and to extract latent yield factors and unobserved regimes simultaneously.

There is substantial prior uncertainty concerning the specification of model that is most appropriate for describing the dynamics of bond yields. To take model uncertainty into consideration, we do not assume two regimes in the RSDNS-X model. Alternatively, the number of regimes could be one, two, or three a priori. We choose the optimal number of regimes using the Bayes factor. The empirical analysis supports a two-regime RSDNS-X model. Our empirical analysis based on the two-regime RSDNS-X model suggests that regime shifts are important for understanding interactions between the yield curve and macro factors. To anticipate our results, we find that the portion of the forecasting variance attributable to macro factors is higher in the turbulent regime than in the tranquil regime. For example, 18% of the 3-month yield forecast variance is attributable to macro factors at a 60-month forecast horizon in the low-volatility regime. In contrast, macro factors explain 37% of the forecast variance in the high-volatility regime. The impulse response analysis also suggests the regime-dependent interaction between the yield curve and macro factors. Moreover, the statistical tests confirm bidirectional interactions between yield and macro factors in both regimes. Finally, we find that two regimes are intimately related to the business cycle and monetary policy.

Forecasting interest rates plays a crucial role in bond portfolio management, interest rate derivative pricing, and risk management. To shed light on the role of regime shifts and macro factors in interest rate forecasts, we therefore use the two regime RSDNS-X model to forecast bond yields. The empirical analysis shows that the two-regime RSDNS-X model usually produces more accurate out-of-sample forecasts of yields than the single-regime DNS model with macro factors and other competing models such as the random walk or the first-order autoregressive model. These results highlight the importance of incorporating regime shifts into the DNS model.

The rest of this paper is organized as follows. Section 2 presents the RSDNS-X model. Section 3 discusses data issues. Section 4 introduces the Markov chain Monte Carlo method and presents empirical results. Section 5 investigates the forecasting performance of the RSDNS-X model. Section 6 concludes.

2. The RSDNS-X model

In the framework of Nelson and Siegel (1987), the yield curve can be formulated as follows:

\[ i_{t} = L_t + S_t \left( \frac{1 - e^{-\lambda t}}{\lambda} \right) + C_t \left( \frac{1 - e^{-\lambda t}}{\lambda} - e^{-\lambda t} \right) + \nu_t, \]

where \( i_{t} \) denotes the bond yield with maturity \( t \) at time \( t \). The three factors \( L_t, S_t, \) and \( C_t \) are denoted as level, slope, and curvature, respectively. The parameter \( \lambda \) determines the rate of exponential decay and the maturity at which the curvature loading achieves its maximum. Empirically, the level factor is corresponding to the long-term interest rate, so the level factor is a long-term factor. By construction, the slope factor has a maximal impact on short maturities and minimal effect on the longer maturity yields, so the slope factor is a short-term factor. In addition, the curvature is a medium-term factor since the factor loading of the curvature achieves its maximum at medium maturity.

The original Nelson–Siegel model is static and built for fitting the cross section of yields, but not for out-of-sample forecasts. DL extends the Nelson–Siegel model to a dynamic form by assuming yield factors following an AR(1) process.


³ For example, conventional wisdom views that the Federal Reserve is accommodative for growth in recessions and proactive in controlling inflation in booms. Hence, the Taylor (1993) rule may suggest a regime-dependent interaction between macro factors and the yield curve due to shifts in monetary policy regimes.
\[ F_t = \alpha + \beta F_{t-1} + \epsilon_t, \text{ for } F_t = L_t, S_t \text{ or } C_t. \]

A two step estimation method can be used to estimate the dynamic Nelson–Siegel model and produce forecasts. The first step estimates three yield factors by using some fixed value of \( \lambda \). The second step is to model and forecast yield factors and interest rates.

The DNS model is simple and easy to estimate. However, for modelling the entire yield curve consistently and simultaneously, we need a state-space representation for the DNS model. A one-step method to estimate \( \lambda \) also calls for a state-space representation for the DNS model. DRA present a state-space representation for the DNS model. For the entire yield curve, the observation equation is

\[
\begin{bmatrix}
L_{t(t_1)} \\
S_{t(t_2)} \\
C_{t(t_3)} \\
\vdots \\
L_{t(s_n)} \\
S_{t(s_{n+1})} \\
C_{t(s_{n+2})}
\end{bmatrix} = \begin{bmatrix}
1 & -e^{x^2 t_1} & e^{x^2 t_1} & \vdots & \vdots & \vdots & e^{x^2 t_1} \\
1 & -e^{x^2 t_2} & e^{x^2 t_2} & \vdots & \vdots & \vdots & e^{x^2 t_2} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
1 & -e^{x^2 t_{n+1}} & e^{x^2 t_{n+1}} & \vdots & \vdots & \vdots & e^{x^2 t_{n+1}}
\end{bmatrix} \begin{bmatrix}
L_t \\
S_t \\
C_t \\
\vdots \\
L_{t(s_n)} \\
S_{t(s_{n+1})} \\
C_{t(s_{n+2})}
\end{bmatrix} + \begin{bmatrix}
\epsilon_{t(t_1)} \\
\epsilon_{t(t_2)} \\
\epsilon_{t(t_3)} \\
\vdots \\
\epsilon_{t(s_n)}
\end{bmatrix},
\]

where the \( N \times 1 \) error term vector \( \epsilon_t = (\epsilon_{t(t_1)}, \ldots, \epsilon_{t(s_n)}) \) follows a multivariate normal distribution with mean 0 and the variance–covariance matrix \( \Omega \), namely, \( \epsilon_t \sim N(0, \Omega) \). The state equation for the DNS model is a vector autoregression (VAR)

\[
\begin{bmatrix}
L_t \\
S_t \\
C_t
\end{bmatrix} = \begin{bmatrix}
\mu_t \\
\mu_5 \\
\mu_C
\end{bmatrix} + \begin{bmatrix}
\phi_{11} & \phi_{12} & \phi_{13} \\
\phi_{21} & \phi_{22} & \phi_{23} \\
\phi_{31} & \phi_{32} & \phi_{33}
\end{bmatrix} \begin{bmatrix}
L_{t-1} \\
S_{t-1} \\
C_{t-1}
\end{bmatrix} + \begin{bmatrix}
\eta_{t1} \\
\eta_{t2} \\
\eta_{t3}
\end{bmatrix},
\]

where \( \mu = (\mu_t, \mu_5, \mu_C) \) is the intercept vector of the state equation. The error term vector \( \eta = (\eta_{t1}, \eta_{t2}, \eta_{t3}) \) follows a multivariate normal distribution.

The two-regime RSDNS-X model is an extension of the DNS model along two directions. The current term structure literature is actively progressing to study the joint behavior of the yield curve and macroeconomic factors. The standard approach for examining the joint behavior is to extend existing term structure models by including observed macro factors. Following this strand of literature, the first extension is to include macro factors in the DNS model. To be consistent with DRA, three macro factors are respectively capacity utilization (CU), the effective federal funds rate (FFR), and inflation (INFL).

Our selection of macro factors follows common practice. To date, a flood of research has linked the yield curve to the macroeconomy. Macro factors that have received the most attention in prior studies include: a real factor for capturing real economic activity; an inflation factor for describing the price level; and a monetary policy factor for capturing the effect of monetary policy. In our setting, CU is a measure of real economic activity; INFL is a measure of price level, FFR is a natural measure of monetary policy. Though there are many different ways for constructing real and inflation factors (e.g., Ang and Piazzesi, 2003; Joslin et al., 2014; Ludvigson and Ng, 2009), recent empirical research generally finds that these factors are highly correlated. Why we use this particular set of macro factors is because DRA use this set of macro variables. Since the RSDNS-X model is an extension of the Nelson–Siegel model with macro factor proposed by DRA, using this set of macro factors makes our results directly comparable with those presented by DRA. The second extension is to allow the dynamic behavior of interest rates and macro variables to be regime-dependent. The regime-dependence implies richer dynamic movements of yields and macro factors, and therefore offers greater econometric flexibility for the term structure to simultaneously account for the cross section and time series properties of interest rates, macro factors, and interactions among these factors. Specifically, the RSDNS-X model allows for regime-dependent means and heteroscedasticity in the dynamics of factors.

If we stack the state variables in a \( 6 \times 1 \) vector \( X_t = (L_t, S_t, C_t, CU_t, FFR_t, INFL_t) \) and stack the yields in \( y_t = (L_{t(t_1)}, \ldots, L_{t(s_n)}) \). The observation equation of the RSDNS-X model can be succinctly written as

\[
y_t = AX_t + \epsilon_t, \]

where \( A \) is an \( N \times 6 \) coefficient matrix with the right most three columns contain only zeros and the left most three columns given in Eq. (3). Thus, yields only load on the yield factors, as in the DNS model. This specification is consistent with DRA and consistent with the view that three factors are needed to distill the information in the yield curve. The joint dynamics of yield and macro factors are assumed to follow a Markov-switching VAR process in order to identify possibly turbulent and tranquil periods in the bond market. Hence, the state equation in the two-regime RSDNS-X model is

\[
X_t = \mu_t + \Phi X_{t-1} + \eta_t; \quad \epsilon_t = H \text{ or } L,
\]

---

4 CU (INFL) is highly correlated with the real (inflation) factor constructed by Ang and Piazzesi (2003), with a correlation coefficient of 0.88 (0.92). Furthermore, CU (INFL) is also highly correlated with the real and inflation factors used in Joslin et al. (2014).

5 If we employ the real and inflation factors used in Ang and Piazzesi (2003) or Joslin et al. (2014), we obtain similar results.

6 Note that term structure models with unspanned macro risks (e.g., Joslin et al., 2014) also recognize the interaction between yield factors and macro factors in the physical measure.
where \( \xi_t = H, L \) indicates a high or low volatility regime prevailing at time \( t \) and \( \eta_t = (\eta_{t,1}, \eta_{t,2}, \eta_{t,3})' \) allows regime-dependent heteroscedasticity,

\[
\begin{align*}
\eta_H &\sim N(0, \Sigma_H), \\
\eta_L &\sim N(0, \Sigma_L).
\end{align*}
\] (7)

In addition, the intercept vector \( \mu_t = (\mu_{t,1}, \mu_{t,2}, \mu_{t,3}, \mu_{t,4}, \mu_{t,5}, \mu_{t,6}, \mu_{t,7})' \) is also regime-dependent. The regime-dependent intercepts lead to regime-dependent means of interest rates. For optimality of the Kalman filter, we assume the disturbances \( \eta_t \) and \( \varepsilon_t \) are uncorrelated with each other, and initial state variables \( X_0 \) is orthogonal to the realization of \( \eta_t \) and \( \varepsilon_t \).

\[
\begin{align*}
E(\varepsilon_t \eta_t) &= 0 \text{ for } t = 1, 2, \ldots, T; \\
E(\varepsilon_t X_0) &= 0 \text{ for } t = 1, 2, \ldots, T.
\end{align*}
\] (8)

A discrete Markov chain governs switches between two regimes with the transition matrix given by

\[
p = \begin{bmatrix}
    p & 1 - q \\
    1 - p & q
\end{bmatrix}.
\]

Now the standard Hamilton (1994) algorithm can be used to extract the probabilities of staying in each regime.

The coefficient matrix \( A \) plays three roles in the RSDNS-X model. First, \( A \) provides identification restrictions. Three yield factors and two regimes are unobserved components in system (5) and (6). As usual, there are some identification conditions that must be imposed to estimate a model with latent factors. \( A \) provides such identification restrictions. Second, \( A \) also gives three latent yield factors a nice interpretation, respectively, level, slope and curvature. This interpretation of three factors in the DNS model seems to be stable over sample selection and set of yields chosen. In contrast, an unrestricted vector autoregression does not provide us such a clear interpretation. In addition, Admissibility constitutes a third role of \( A \). As discussed in DRA, the Nelson–Siegel form avoids a negative forward rate at all horizons.

The RSDNS-X model achieves parsimony by a diagonal \( \Omega \) assumption. Since three underlying latent factors explain a large fraction of yield variations (e.g., Diebold and Li, 2006; Diebold et al., 2006), the diagonal \( \Omega \) is expected to be a good approximation. Thus, the efficiency loss from diagonal restrictions should be insignificant. Indeed, this is an usual strategy in affine term structure models with no-arbitrage conditions. For example, Ang and Piazzesi (2003) assume some yields are measured with errors. Computational tractability is a second reason for the diagonal covariance matrix assumption.

### 3. Yields and macro factors

The yield curve consists of the monthly observations of 1, 3, 6, 12, 24, 36, 60, 84, 120 months zero-coupon yields on Treasury securities. The sample covers a period from January 1983 to September 2014. The data source is the economic database, Federal Reserve Bank of St. Louis. Fig. 1 plots bond yields and Table 1 presents some summary statistics of bond yields. The bond yields are characterized by some standard stylized facts: the average yield curve is increasing and concave; the yield curve shows a variety of shapes through time, for example, upward sloping, downward sloping, humped shape, and so on; the bond yields are highly autocorrelated, with increasing autocorrelation at longer maturity; the bond yields show positive skewness, it is particularly so at the long end of the yield curve; The kurtosis of the bond yields increases with maturity.

---

**Fig. 1.** Bond yields. The sample covers the period from January 1983 to September 2014. The yields plotted in this graph include, from the lowest to the highest line (with occasional cross-overs), 1-, 3-, 6-, 12-, 24-, 36-, 60-, 84-, 120-month treasury zero-coupon bond yields.
Three proxies for economic activity and monetary policy are capacity utilization, the effective Fed funds rate, and inflation. The CU, FFR and consumer price index (CPI) data are obtained from the economic database, Federal Reserve Bank of St. Louis. The year-over-year inflation rate is defined by taking the yearly percentage change in the CPI index, 
\[ \text{INFL}_t = 100 \left( \frac{\ln \text{CPI}_t - \ln \text{CPI}_{t-12}}{\ln \text{CPI}_t - \ln \text{CPI}_{t-12}} \right) \].

Inflation is included in the RSDNS-X model because it is a key variable in shaping the nominal yield curve through the level of inflation and inflation risk premia. CU is a measure of the deviation of economic activity from its natural level. FFR is an indicator of monetary policy.

4. Empirical analysis

This section conducts the empirical analysis. Section 4.1 introduces the Markov chain Monte Carlo method for estimating the RSDNS-X model. Section 4.2 uses the Bayes factor to determine the optimal number of regimes. Section 4.3 reports the estimation results. Section 4.4 tests the interaction restrictions between the yield curve and macro factors. Sections 4.5 and 4.6 conduct the impulse response analysis and variance decomposition, respectively. Section 4.7 relates two regimes to the business cycle and monetary policy.

4.1. Model estimates

The state-space system of the RSDNS-X model is estimated by the Markov chain Monte Carlo method (MCMC), which is a combination of a Gibbs sampling step and a Random Walk Metropolis step. The MCMC method is efficient and numerically trustworthy. In particular, the MCMC approach incorporates parameter uncertainty into the estimation by integrating over posterior distributions. Appendix A provides the details of the MCMC method.

To evaluate the reliability of the MCMC estimation method, we turn to some diagnostics. The basic idea of most convergence statistics is to compare moments of sampled parameters. A visual check on the plot of sampled parameters can provide information about convergence. For an MCMC implementation that converges, parameter drawings should not deviate from some mean for a long period, although this is subjective in the sense that there is no clear measure of deviation and duration. The second statistic is the plot of CUSUM path. The third criterion for judging convergence is the relative numerical efficiency (RNE).

4.2. Model selection

There is substantial prior uncertainty concerning the specification of model that is most appropriate for describing the dynamics of interest rates. To take model uncertainty into consideration, we use the Bayes factor, a widely employed method in the Bayesian statistics literature, to select the optimal number of regimes. Kass and Raftery (1995) provide a detailed discussion of the Bayes factor. An advantage of the use of the Bayes factor is that it automatically, and quite naturally, includes a penalty for including too much model structure. It thus guards against in-sample overfitting. Consider two models, \( M_1 \) and \( M_2 \), given the observed data \( Y \) and prior odds \( p(M_1) \), the Bayesian method compares the models using the posterior odds
\[ \frac{p(M_1 | Y_t)}{p(M_2 | Y_t)} = \frac{p(Y_t | M_1) p(M_1)}{p(Y_t | M_2) p(M_2)} \] (9)

If we set prior odds \( p(M_1) = p(M_2) = 0.5 \), we compare the models using the Bayes factor. Hence, the Bayes factor is a summary of the evidence provided by the data in favor of one model as opposed to another.

Let \( M_1 \) denote the two-regime RSDNS-X model and \( M_2 \) denote the one-regime (three-regime) RSDNS-X model, the Bayes factor is 19.2 (7.3). A value of \( \frac{p(M_1 | Y_t)}{p(M_2 | Y_t)} > 1 \) means that the two-regime RSDNS-X model is more strongly supported by the data.

Table 1
Summary statistics of bond yields.

<table>
<thead>
<tr>
<th>Maturity (months)</th>
<th>Mean</th>
<th>Std.dev.</th>
<th>Skew.</th>
<th>Kuro.</th>
<th>( \hat{\rho}(1) )</th>
<th>( \hat{\rho}(12) )</th>
<th>( \hat{\rho}(30) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.94</td>
<td>2.73</td>
<td>0.08</td>
<td>2.10</td>
<td>0.99</td>
<td>0.79</td>
<td>0.39</td>
</tr>
<tr>
<td>3</td>
<td>4.15</td>
<td>2.87</td>
<td>0.08</td>
<td>2.08</td>
<td>0.99</td>
<td>0.80</td>
<td>0.42</td>
</tr>
<tr>
<td>6</td>
<td>4.33</td>
<td>2.95</td>
<td>0.10</td>
<td>2.14</td>
<td>0.99</td>
<td>0.80</td>
<td>0.42</td>
</tr>
<tr>
<td>12</td>
<td>4.50</td>
<td>3.00</td>
<td>0.14</td>
<td>2.24</td>
<td>0.99</td>
<td>0.81</td>
<td>0.45</td>
</tr>
<tr>
<td>24</td>
<td>4.88</td>
<td>3.07</td>
<td>0.20</td>
<td>2.35</td>
<td>0.99</td>
<td>0.82</td>
<td>0.50</td>
</tr>
<tr>
<td>36</td>
<td>5.10</td>
<td>3.02</td>
<td>0.24</td>
<td>2.44</td>
<td>0.99</td>
<td>0.82</td>
<td>0.53</td>
</tr>
<tr>
<td>60</td>
<td>5.49</td>
<td>2.88</td>
<td>0.35</td>
<td>2.65</td>
<td>0.99</td>
<td>0.83</td>
<td>0.55</td>
</tr>
<tr>
<td>84</td>
<td>5.79</td>
<td>2.77</td>
<td>0.45</td>
<td>2.75</td>
<td>0.99</td>
<td>0.83</td>
<td>0.56</td>
</tr>
<tr>
<td>120</td>
<td>6.01</td>
<td>2.62</td>
<td>0.57</td>
<td>2.86</td>
<td>0.99</td>
<td>0.83</td>
<td>0.57</td>
</tr>
</tbody>
</table>

This table presents summary statistics for zero-coupon bond yields. Data are obtained from the Federal Reserve. The sample period is from 1983:01 to 2014:09. \( \rho_{ij} \) is autocorrelation coefficient with lag length \( i \).
under consideration than the competitors. Indeed, the literature has relied on the following interpretations of the Bayes factor in evaluating models: less than 1, negative support for \( M_1 \); 1–3.2, barely worth mentioning; 3.2–10, substantial; 10–30, strong; 30–100 very strong; and greater than 100, decisive. Our empirical analysis indicates that the two-regime RSDNS-X model is supported by the Bayes factor. Hence, we will focus on the two-regime RSDNS-X model in the following empirical analysis.

4.3. Estimation results

We use the MCMC method to estimate the two-regime RSDNS-X model. Three yield factors are drawn based on the multistep Gibbs sampling algorithm (Cater and Kohn, 1994) where the entire conditional posterior distributions are from other parameters and the Kalman filter. This method simplifies the MCMC simulation because we can draw yield factors jointly by a recursive method. Specifically, we use the Kalman filter to process yield factors forward, then we take random draws of the posterior distributions backward. The forward filtering and backward sampling (FFBS) method makes the simulation more efficient because this scheme draws serially correlated yield factors jointly. Using the FFBS scheme combined with the Hamilton (1994) filter, we can also generate the unobserved regimes prevailing at each time point \( t \). The smoothed regimes are usually parameters of interest, the FFBS scheme combined with the Kim (1994) filter produces drawings of smoothed regimes.

To facilitate the convergence of MCMC iterations, we initialized the MCMC procedure by a two-step estimation. The first step estimates the DNS model by fixing \( \lambda \) at 0.0609, as in DL. With estimated yield factors, the state equation can be estimated by the Gaussian maximum likelihood method. The estimated parameters from the two-step estimation are then catered to the MCMC scheme. This initialization makes the MCMC converged quickly. We also try other initials in the estimation. Results are similar, though the MCMC scheme converges slowly. In the MCMC experiment, all nonstationary drawings were dropped to ensure that the estimated system is stationary. Specifically, the MCMC estimation consists of 50,000 stationary iterations, the number of burn-in iteration is 20,000. The CUSUM and RNE indicate the convergence of the estimation.

The estimation results are reported in Table 2. Three latent factors are all persistent though the degree of persistence differs. The autoregressive coefficient is 0.97 for the most persistent level factor, while for the least persistent curvature factor it is 0.90. These results are consistent with those typically found in empirical term structure models. The value of \( \lambda \) is 0.053. This is consistent with our prior belief that it maximizes curvature in medium-term.\(^7\)

Fig. 2 plots the yield factors extracted from the RSDNS-X model, macro factors, and the empirical yield factors. It is evident that the slope factor is closely linked to the detrended CU with a correlation coefficient of 0.53. The level (curvature) factor is correlated with inflation (FFR) with a correlation coefficient of 0.51 (0.78). In addition, the estimated yield factors are closely related to the empirical factors. Specifically, the empirical level factor is defined as an average of short-, medium- and long-term yields \((h_{t(3)} + h_{t(24)} + h_{t(120)})/3\); the empirical slope factor is defined as the difference between the 3-month and 120-month interest rates \((h_{t(3)} - h_{t(120)})\); the empirical curvature factor is defined as \(2h_{t(24)} - h_{t(3)} - h_{t(120)}\). The simple correlation between the estimated and empirical level factors is 0.92; \( S_1 \) is closely linked to the empirical slope with a correlation of 0.98; \( C_1 \) is intimately correlated with the empirical curvature with a correlation coefficient 0.97.

Two regimes continue to be labeled as \( L \) and \( \bar{H} \) regimes according to the estimation results. It is clear that the residual variances of the yield factors in regime \( L \) are significantly smaller than those in regime \( \bar{H} \). Statistically, the Wald statistic reported in Table 2 reject the null hypothesis of equal variance in two regimes. The rejection of equal variance implies the existence of regime-dependent heteroscedasticity. The Wald statistics also indicate that neither the covariance matrix \( \Sigma_L \) nor the matrix \( \Sigma_H \) are diagonal. Fig. 3 plots the smoothed probabilities of being in regime \( H \). The regime classification results suggest that NBER economic recessions are usually observed during the high-volatility regime.

Table 3 reports the fits of the RSDNS-X model. For comparison, the fits of the RSDNS model are also presented in Table 3. For each model, we present the estimated means and standard deviations of the measurement equation residuals. On average, the fitting errors are usually only a few basis points. It seems that both models fits the yield curve well.

4.4. Testing interactions across the yield curve and macro factors

There are three interesting null hypothesis about interactions across yield and macro factors. The first hypothesis is totally no interaction between yield and macro factors. A less strong assumption is that the dynamics of yield factors do affect the dynamics of macro factors, but not vice versa. Opposed to the second assumption, the last hypothesis postulates that the unidirectional linkage is from macro factors to yield factors.

Following DRA, three hypotheses can be formalized by zero restrictions on the autoregressive matrix and the variance-covariance matrix of the state equation. Specifically, we partition the \((6 \times 6)\) matrix \( \Phi \) into four \((3 \times 3)\) blocks

\[
\Phi = \begin{bmatrix}
\Phi_1 & \Phi_2 \\
\Phi_3 & \Phi_4 
\end{bmatrix},
\]

\(^7\) The scaling factor \( \xi \) in the Random Walk Metropolis steps (see Appendix A) for generating \( \lambda \) is set to be 0.00028. The acceptance rate of the Random Walk Metropolis steps is 0.31.
Table 2
Parameter estimates of the RSDNS-X model and tests.

<table>
<thead>
<tr>
<th>Autoregressive coefficient matrix $\Phi$</th>
<th>$L_{t-1}$</th>
<th>$S_{t-1}$</th>
<th>$C_{t-1}$</th>
<th>$CU_{t-1}$</th>
<th>$FFR_{t-1}$</th>
<th>$INFL_{t-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_t$</td>
<td>0.97</td>
<td>0.04</td>
<td>0.02</td>
<td>-0.01</td>
<td>-0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>(18.94)</td>
<td>(0.65)</td>
<td>(1.24)</td>
<td>(-2.05)</td>
<td>(0.33)</td>
<td>(1.74)</td>
<td></td>
</tr>
<tr>
<td>$S_t$</td>
<td>-0.02</td>
<td>0.92</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>(0.38)</td>
<td>(16.74)</td>
<td>(4.53)</td>
<td>(2.33)</td>
<td>(0.32)</td>
<td>(-0.29)</td>
<td></td>
</tr>
<tr>
<td>$C_t$</td>
<td>0.72</td>
<td>0.68</td>
<td>0.90</td>
<td>0.03</td>
<td>-0.68</td>
<td>-0.01</td>
</tr>
<tr>
<td>(4.26)</td>
<td>(4.71)</td>
<td>(38.71)</td>
<td>(2.75)</td>
<td>(-4.57)</td>
<td>(-0.09)</td>
<td></td>
</tr>
<tr>
<td>$CU_t$</td>
<td>0.18</td>
<td>0.10</td>
<td>0.05</td>
<td>0.98</td>
<td>-0.16</td>
<td>-0.06</td>
</tr>
<tr>
<td>(2.35)</td>
<td>(1.24)</td>
<td>(2.26)</td>
<td>(95.68)</td>
<td>(-2.14)</td>
<td>(-1.94)</td>
<td></td>
</tr>
<tr>
<td>$FFR_t$</td>
<td>0.28</td>
<td>0.26</td>
<td>0.07</td>
<td>0.01</td>
<td>0.71</td>
<td>0.00</td>
</tr>
<tr>
<td>(3.29)</td>
<td>(7.83)</td>
<td>(11.28)</td>
<td>(0.53)</td>
<td>(18.92)</td>
<td>(0.32)</td>
<td></td>
</tr>
<tr>
<td>$INFL_t$</td>
<td>-0.05</td>
<td>-0.09</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.09</td>
<td>0.89</td>
</tr>
<tr>
<td>(-1.47)</td>
<td>(-1.66)</td>
<td>(-0.32)</td>
<td>(2.68)</td>
<td>(1.98)</td>
<td>(41.75)</td>
<td></td>
</tr>
</tbody>
</table>

Estimated covariance matrix $\Sigma_l$:

<table>
<thead>
<tr>
<th>Estimated covariance matrix $\Sigma_l$</th>
<th>$L_t$</th>
<th>$S_t$</th>
<th>$C_t$</th>
<th>$CU_t$</th>
<th>$FFR_t$</th>
<th>$INFL_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_t$</td>
<td>0.066</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$S_t$</td>
<td>-0.058</td>
<td>0.063</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$C_t$</td>
<td>-0.008</td>
<td>-0.014</td>
<td>0.238</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$CU_t$</td>
<td>0.024</td>
<td>0.004</td>
<td>-0.011</td>
<td>0.157</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$FFR_t$</td>
<td>0.002</td>
<td>0.001</td>
<td>-0.001</td>
<td>0.018</td>
<td>0.006</td>
<td>–</td>
</tr>
<tr>
<td>$INFL_t$</td>
<td>0.033</td>
<td>-0.019</td>
<td>-0.022</td>
<td>0.045</td>
<td>0.014</td>
<td>0.063</td>
</tr>
</tbody>
</table>

Estimated covariance matrix $\Sigma_H$:

<table>
<thead>
<tr>
<th>Estimated covariance matrix $\Sigma_H$</th>
<th>$L_t$</th>
<th>$S_t$</th>
<th>$C_t$</th>
<th>$CU_t$</th>
<th>$FFR_t$</th>
<th>$INFL_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_t$</td>
<td>0.081</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$S_t$</td>
<td>-0.033</td>
<td>0.106</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$C_t$</td>
<td>-0.014</td>
<td>-0.022</td>
<td>0.905</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$CU_t$</td>
<td>0.008</td>
<td>-0.002</td>
<td>0.086</td>
<td>0.312</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$FFR_t$</td>
<td>0.020</td>
<td>0.048</td>
<td>0.058</td>
<td>0.019</td>
<td>0.087</td>
<td>–</td>
</tr>
<tr>
<td>$INFL_t$</td>
<td>0.005</td>
<td>-0.004</td>
<td>0.013</td>
<td>0.011</td>
<td>0.007</td>
<td>0.218</td>
</tr>
</tbody>
</table>

Transition probabilities, Wald test of diagonality, and test for no regime-dependent heteroscedasticity:

<table>
<thead>
<tr>
<th>$p$</th>
<th>$q$</th>
<th>$\Sigma_l$</th>
<th>$\Sigma_H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.92</td>
<td>0.091</td>
<td>68.72</td>
<td>74.39</td>
</tr>
<tr>
<td>(15.83)</td>
<td>(12.71)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
</tbody>
</table>

This table reports estimates of parameters of the RSDNS-X model. The table also reports Wald test statistics for diagonality of $\Sigma_l$ and $\Sigma_H$ and no regime-dependent heteroscedasticity. Bold entries indicate significance at a 5% level. $t$-values are in parenthesis. For Wald tests, $p$-values are in parentheses.

Fig. 2. Yield factors, empirical yield factors, and macroeconomic factors. The upper panel shows the level factor implied by the RSDNS-X model, the empirical level factor, and inflation. The middle panel shows the slope factor implied by the RSDNS-X model, the empirical slope factor, and demeaned capacity utilization. The lower panel shows the curvature factor implied by the RSDNS-X model, the empirical curvature factor, and the Federal fund rate.
and similarly partitioning the covariance matrices \( \Sigma_H \) and \( \Sigma_L \)

\[
\Sigma_{i_t} = \begin{bmatrix} \Sigma_1 & \Sigma_2 \\ \Sigma_3 & \Sigma_4 \end{bmatrix}; \quad \xi_t = L \text{ or } H,
\]

(11)

where \( \Sigma_3 \) is the transpose of the \( \Sigma_2 \). Given the prevailing regime, \( \Phi_2 = \Phi_3 = \Sigma_2 = 0 \) is the equivalence of the first null hypothesis. The second hypothesis can be rewritten as \( \Phi_2 = 0 \). The restrictions for the third hypothesis are \( \Phi_3 = \Sigma_2 = 0 \). The Wald test is easily implemented for testing these hypotheses. Table 4 reports the Wald statistics for three hypotheses in regime \( L \) and \( H \). All hypotheses are overwhelmingly rejected. This implies bidirectional linkages across yield and macro factors in both regimes. The finding is consistent with a growing literature that relates the term structure of interest rates with economic activity.

4.5. Impulse responses

One toolkit for investigating the dynamic relationship between yield and macro factors is impulse responses (IR). For computing impulse responses, we take two strategies. The first strategy is to calculate general impulse responses (Pesaran and Shin, 1998). When one variable is shocked, other variables also vary as is implied by the covariance. The general impulse response function computes the mean by integrating out all other shocks. As such, general IRs are neutral with respect to orderings of factors. The second strategy is to calculate orthogonalized IRs. Producing impulse responses from this recursive contemporaneous structure requires an identification of the covariance since innovations are contemporaneously

Table 3
Measurement errors of yields.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>RSDNS Mean</th>
<th>Std. dev.</th>
<th>RSDNS-X Mean</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-month</td>
<td>-0.098</td>
<td>0.018</td>
<td>-0.095</td>
<td>0.022</td>
</tr>
<tr>
<td>3-month</td>
<td>0.032</td>
<td>0.008</td>
<td>0.029</td>
<td>0.007</td>
</tr>
<tr>
<td>6-month</td>
<td>0.082</td>
<td>0.009</td>
<td>0.075</td>
<td>0.008</td>
</tr>
<tr>
<td>12-month</td>
<td>0.023</td>
<td>0.007</td>
<td>0.016</td>
<td>0.006</td>
</tr>
<tr>
<td>24-month</td>
<td>0.015</td>
<td>0.002</td>
<td>0.016</td>
<td>0.002</td>
</tr>
<tr>
<td>36-month</td>
<td>-0.052</td>
<td>0.003</td>
<td>-0.061</td>
<td>0.003</td>
</tr>
<tr>
<td>60-month</td>
<td>-0.055</td>
<td>0.004</td>
<td>-0.053</td>
<td>0.004</td>
</tr>
<tr>
<td>84-month</td>
<td>0.015</td>
<td>0.002</td>
<td>0.014</td>
<td>0.002</td>
</tr>
<tr>
<td>120-month</td>
<td>0.037</td>
<td>0.005</td>
<td>0.024</td>
<td>0.006</td>
</tr>
</tbody>
</table>

The table reports the filtered errors from the RSDNS-X and RSDNS models. Errors are defined as the difference between the actual yield curve and its filtered estimates from the Kamlan filter. The RSDNS model is the regime switching DNS model without macro factors.
correlated and have overlapping information to some extent. We use two orderings in the Cholesky decomposition by respectively placing the yield curve factors prior to the macro factors and by placing the macro factors prior to the yield curve factors: \((L_t, S_t, C_t, C_{U_t}, FFR_t, \text{INFL}_t)\) and \((C_{U_t}, \text{INFL}_t, FFR_t, L_t, S_t, C_t)\).

Each IR is measured in terms of one percentage point shock to residuals. We consider two classifications of impulse responses. Since one focus of the RSDNS-X model is regime shifts, the first classification is to compare impulse responses in regime \(L\) and \(H\). According to another focus of the RSDNS-X model, that is, linkages between yield and macro factors, the second classification splits impulse responses into four groups as in DRA: macro-to-macro responses, macro-to-yield responses, yield-to-yield responses and yield-to-macro responses.

To save space, we report the general IRs.⁹ Figs. 4 and 5 plot the general impulse responses of the yield and macro factors on each other in regime \(L\) and \(H\). We find that the yield factors usually respond to the macro factors. It is evident that the magnitude of the impulse responses varies across regimes. Similarly, the yield factors also respond to shocks to the macro factors in both regimes.

To shed light on how short- and long-term yields react to yield and macro factors in two regimes, Fig. 6 plots the impulse responses of 3, 12, and 60 month yields to shocks on the yield and macro factors in the low volatility regimes. Fig. 7 plots the IRs of 3, 12, and 60 month yields to factor shocks in the high volatility regime. The impulse responses of bond yields to level shocks are flat, this is consistent with the effect of level in the yield curve. The IRs of bond yields to slope shocks decay quickly, in line with the interpretation of the slope. Similar to the hump shape of the curvature factor, the IRs of bond yields to curvature shocks show hump shapes. It is worth noting that the magnitude of IRs in two regimes are different, though the shape of IRs are similar. The figures shows that bond yields respond to CU shocks positively and significantly in both regimes. In addition, the responses to CU shocks are stronger in the high volatility regime than in the low volatility regime. Inflation shock usually has positive effect on bond yields in both regimes, though the effect is small. The initial responses of bond yields to FFR shocks are positive. However, the IRs of bond yields to FFR shocks become negative quickly.

### 4.6. Variance decompositions

Another toolkit for understanding interactions between yield and macro factors is variance decomposition. Variance decomposition can tell us the relative contributions of the macro and yield factors to forecast errors. We calculate variance decomposition using the averaging method of Pesaran and Shin (1998), which is consistent the calculation of impulse responses.⁹ Table 5 reports the variance decomposition results. The results suggest that the proportions of medium and long-term yield forecast variances attributable to the macro factors are significantly higher in the high-volatility regime than in the low-volatility regime. For example, 18% of the 3-month yield forecast variance is attributable to the macro factors at a 60-month forecast horizon in the low-volatility regime. In contrast, the macro factors explain 37% of the forecast variance in the high-volatility regime.

At the long end of the yield curve, proportions of forecast variances attributable to the macro factors is also higher in the turbulent regime than in the tranquil regime. For example, 16% of the 60-month yield forecast variance is attributable to the macro factors at a 1-month forecast horizon in the low-volatility regime. On the other hand, the proportion is 33% in the high-volatility regime. It is also worth noting that, in the high-volatility regime, the macro factors explain up to 39% of the 60-month yield forecast variance at a 60-month horizon. Overall, the variance decomposition analysis suggests that the macro factors are important for understanding the forecast variance of bond yields.

### 4.7. Understanding regimes

In the two-regime RSDNS-X model, the most straightforward interpretation for the two regimes is a low volatility regime and a high volatility regime. However, this classification does not offer an insight on their economic interpretation. In this section, we attempt to relate the regimes to the business cycle and to a time-varying monetary policy.

---

⁹ For the IRs computed using a Cholesky orthogonalization, we obtain qualitatively similar results. We consistently find that the interaction between the yield curve and the macro factors varies across the two regimes.

⁹ Changing the ordering of the yield and macro factors just has small effect on the results.
A visual check seems to suggest that the regime classification by and large coincides with the NBER business cycles. Having qualitatively related two regimes to business cycles, the next issue is to quantitatively investigate whether or not the probabilities of regimes are related to economic activity. In the spirit of Bansal and Zhou (2002) and Zhu (2015a), we esti-
Fig. 6. Impulse response functions. Impulse Responses (IR’s) for 3 month (top panel), 12 month (middle panel) and 60 month (bottom panel) yields in the low volatility regime.

Fig. 7. Impulse response functions. Impulse Responses (IR’s) for 3 month (top panel), 12 month (middle panel) and 60 month (bottom panel) yields in the high volatility regime.
mate a logit model to examine the relation between economic activity and regime classification. The binary variable is defined to be one when the average monthly filtered probability of being in a low volatility regime is smaller than one-half and to be zero when the average monthly filtered probability of being in a low volatility regime is greater than or equal to one-half. Because the GDP data are not available at a monthly frequency, the explanatory proxy variable for real economic activity is the growth rate of industrial production (IP), which is obtained from the economic database, Federal Reserve Bank of St. Louis. The logit model is as follows:

$$P_t(L) = \frac{\exp(\delta_0 + \delta_1 IP_t)}{1 + \exp(\delta_0 + \delta_1 IP_t)};$$

where $P_t(L)$ is the implied probability of being in a low volatility regime. The estimation shows that $\delta_1 = 0.12$ with a $t$-value 11.2. The pseudo-$R^2$ of the logit regression is 0.25. The result confirms the traditional wisdom that the probability of being in a low volatility regime is higher when an economy is in a boom. The logit regression thus provides supporting evidence on the business cycle interpretation of two regimes.

Two regimes are also likely to be related to monetary policy. A large body of narrative and empirical evidence (e.g., Ang et al., 2011; Li et al., 2011; Clarida et al., 2006; Zhu, 2014) suggests that a time-varying monetary policy has important implications for the term structure of interest rates. The transmission channels include: (1) a time-varying monetary policy largely determines short-term interest rates through the Taylor (1993) rule; (2) a time-varying monetary policy affects the expected inflation and inflation risk premium. Eventually, it affects long-term interest rates; (3) a time-varying monetary policy affects expected future short-term interest rates, so it finally exerts an influence on long-term interest rates. Since a time-varying monetary policy affects the term structure of interest rates, two regimes identified from the term structure of interest rates should contain information about time-varying monetary policy.

For investigating the relation between time-varying monetary policy and regimes, we assume that the dynamic behavior of the short-term interest rate (the 1-month interest rate) follows the Taylor rule where monetary authority sets the short rate as a function of inflation ($\pi_t$) and the output gap ($g_t$). We respectively estimate the Taylor rule in the high volatility regime and in the low volatility regime. Based on the observations in the high volatility regime, the Taylor rule is:

$$i_{t(1)} = 0.14 + 0.98\pi_t + 0.32g_t + \epsilon_t;$$

(0.85) (10.8) (3.95)

where $t$-values are reported in parentheses. The Taylor rule in the low volatility regime is:

$$i_{t(2)} = 0.51 + 1.53\pi_t + 0.06g_t + \epsilon_t;$$

(5.37) (16.7) (1.18)

Table 5: Variance decompositions.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>Level</th>
<th>Slope</th>
<th>Curvature</th>
<th>CU</th>
<th>FFR</th>
<th>INFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-month yield</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 L</td>
<td>0.14</td>
<td>0.76</td>
<td>0.05</td>
<td>0.00</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>24 L</td>
<td>0.25</td>
<td>0.67</td>
<td>0.03</td>
<td>0.00</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>60 L</td>
<td>0.35</td>
<td>0.27</td>
<td>0.16</td>
<td>0.05</td>
<td>0.07</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12-month yield</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 L</td>
<td>0.22</td>
<td>0.27</td>
<td>0.42</td>
<td>0.02</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>24 L</td>
<td>0.31</td>
<td>0.16</td>
<td>0.33</td>
<td>0.07</td>
<td>0.12</td>
<td>0.02</td>
</tr>
<tr>
<td>60 L</td>
<td>0.39</td>
<td>0.15</td>
<td>0.36</td>
<td>0.10</td>
<td>0.08</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60-month yield</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 L</td>
<td>0.27</td>
<td>0.18</td>
<td>0.39</td>
<td>0.10</td>
<td>0.05</td>
<td>0.01</td>
</tr>
<tr>
<td>24 L</td>
<td>0.36</td>
<td>0.12</td>
<td>0.38</td>
<td>0.06</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>60 L</td>
<td>0.58</td>
<td>0.08</td>
<td>0.23</td>
<td>0.06</td>
<td>0.04</td>
<td>0.01</td>
</tr>
</tbody>
</table>

The table reports the contributions of factor $i$ to the $h$-step-ahead forecast of 3-, 12-, 60-month yields in the high-volatility and low-volatility regime. The variance decomposition is based on the RSDNS-X model.

10 The output is measured by capacity utilization.
It is interesting to see that most coefficients are significantly different under the two regimes. The coefficient of inflation is 0.98 under the high volatility regime, while the coefficient is 1.53 in the low volatility regime. Therefore, the Federal Reserve is more (less) active in controlling inflation in the low (high) volatility regime. The coefficient of output gap is about 0.32 (0.06) under the high volatility regime (the low volatility regime). In particular, the coefficient of output gap is statistically insignificant in the low volatility regime, suggesting that the Fed is more accommodative for growth under the high volatility regime than the low volatility regime. Since the high volatility regime is related to economic recessions, the result is reasonable. The Taylor-rule estimates suggest that the Fed is proactive in controlling inflation in the low volatility regime but is more accommodative for growth in the high volatility regime. The Taylor-rule analysis thus provides us supporting evidence on the monetary policy interpretation of regimes.

### 5. Interest rate forecasts

A good term structure model should not only fit the yield curve well in-sample but also forecast well out-of-sample. In this section, we evaluate the out-of-sample forecast accuracy of the two-regime RSDNS-X model. The empirical benchmark for evaluating the out-of-sample predictive accuracy of the RSDNS model is a naive random walk (RW), an autoregression (AR), a vector autoregression (VAR), the DNS model, and the DNS model with macro factors (DNS-X) model. In the RW, any \( h \)-step-ahead forecast of yield \( i_{t+h|t} \) is simply equal to the most recently observed value \( i_{t|t} \). This no-change model is a good benchmark for judging the prediction power of other models. Because yields are nonstationary or near nonstationary, in practice, it is difficult to beat the RW in terms of out-of-sample forecasting accuracy.

The forecasting procedure for examining out-of-sample forecasts over the last 10 years of our sample is as follows. We estimate and forecast recursively, using data from 1983:01 to the time that the forecast is made, beginning in 2004:10 and extending through 2014:09. We then compare the 1- and 6-month-ahead forecasts from the RSDNS-X model to those of competitors, for all nine yields. To evaluate the out-of-sample forecasting performance, we calculate the popular error metric, per maturity and per forecast horizon. Specifically, the metric is root mean squared forecast errors (RMSE).

#### Table 6
 **Forecast comparisons.**

<table>
<thead>
<tr>
<th>Maturity</th>
<th>RW</th>
<th>AR(1)</th>
<th>DNS-X</th>
<th>DNS</th>
<th>VAR</th>
<th>RSDNS-X</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-month</td>
<td>0.2202</td>
<td>0.2233</td>
<td>0.2184</td>
<td>0.2099</td>
<td>0.2203</td>
<td>0.2277*</td>
</tr>
<tr>
<td>3-month</td>
<td>0.1924</td>
<td>0.1941</td>
<td>0.1865</td>
<td>0.1893</td>
<td>0.1894</td>
<td>0.1932*</td>
</tr>
<tr>
<td>6-month</td>
<td>0.1759</td>
<td>0.1775</td>
<td>0.1883</td>
<td>0.1858</td>
<td>0.1734</td>
<td>0.1712*</td>
</tr>
<tr>
<td>12-month</td>
<td>0.1749</td>
<td>0.1765</td>
<td>0.1819</td>
<td>0.1805</td>
<td>0.1768</td>
<td>0.1707*</td>
</tr>
<tr>
<td>24-month</td>
<td>0.1879</td>
<td>0.1890</td>
<td>0.1897</td>
<td>0.2048</td>
<td>0.1931</td>
<td>0.1879*</td>
</tr>
<tr>
<td>36-month</td>
<td>0.2033</td>
<td>0.2038</td>
<td>0.2248</td>
<td>0.2348</td>
<td>0.2257</td>
<td>0.2137*</td>
</tr>
<tr>
<td>60-month</td>
<td>0.2202</td>
<td>0.2196</td>
<td>0.2418</td>
<td>0.2463</td>
<td>0.2271</td>
<td>0.2318*</td>
</tr>
<tr>
<td>84-month</td>
<td>0.2235</td>
<td>0.2226</td>
<td>0.2257</td>
<td>0.2286</td>
<td>0.2311</td>
<td>0.2205*</td>
</tr>
<tr>
<td>120-month</td>
<td>0.2167</td>
<td>0.2155</td>
<td>0.2430</td>
<td>0.2420</td>
<td>0.2338</td>
<td>0.2123*</td>
</tr>
</tbody>
</table>

Panel A of Table 6 reports the 1-month-ahead and 6-month-ahead out-of-sample forecasting results. We forecast over the last 120 months of our sample and record the root mean square error (RMSE) of the forecast versus the actual values. Lower RMSE denotes better forecasts, with the best statistics highlighted in boldface. We first estimate models on the in-sample, and update the estimations at each observation in the out-of-sample.

\* Indicates that the Diebold and Mariano (1995) test statistic is significant at 10% level.

This table reports the 1-month-ahead and 6-month-ahead out-of-sample forecasting results. We forecast over the last 120 months of our sample and record the root mean square error (RMSE) of the forecast versus the actual values. Lower RMSE denotes better forecasts, with the best statistics highlighted in boldface. We first estimate models on the in-sample, and update the estimations at each observation in the out-of-sample.

It seems to suggest that the RSDNS model is competitive in the out-of-sample forecasting of bond yields.
6. Conclusions

This paper proposes a regime-switching term structure model with latent yield factors and macro factors. The model allows for the regime-dependent interaction between the yield curve and macro factors. The model’s convenient state-space representation facilitates estimation, the extraction of latent yield factors, and unobserved regimes. An MCMC procedure is used to estimate the model. The empirical analysis suggests that the response of the yield curve to macro factors changes as regime switches. The variance decomposition analysis confirms the regime-dependent interaction between yield and macro factors. We also find that the RSDNS-X model is very competitive in terms of out-of-sample forecasting.

Though the RSDNS-X model is empirically almost arbitrage-free, it does not explicitly impose no-arbitrage restrictions. Theoretically, this is a drawback of the RSDNS-X model. Indeed, imposing no-arbitrage restrictions allows us to explicitly investigate the importance of macro factors in understanding time-varying bond premia. Recently, Christensen et al. (2011) present a class of arbitrage-free dynamic Nelson–Siegel term structure models. In future work, we hope to explicitly impose no-arbitrage conditions on RSDNS-X and explore whether the imposition of restrictions is helpful for forecasting and for understanding dynamic interactions between the yield curve and the macroeconomy. In addition, examining the performance of RSDNS-X in international markets (e.g., Zhu, 2015b) constitute an interesting research topic.

Acknowledgment

We thank Ronald Gallant, Pok-sang Lam, Yangru Wu, Anindya Banerjee, Christian Rober, Luigi Spezia, Choy Keen Meng, Dennis Alba, Jie Zhu, Wang Donghua, Weixing Zhou, Yongping Ruan, Jianjun Li, and seminar participants at many universities for helpful comments. Xiaoneng Zhu acknowledges the financial support from the Natural Science Foundation of China (Grant No. 71473281).

Appendix A. The MCMC algorithm

We estimate the RSDNS-X model by the MCMC approach. The parameters that we need to estimate are

$$\Theta = \{\Phi, \mu_l, \mu_H, \Omega_l, \Omega_H, \Sigma, \lambda, p, q, \{X_{t}\}_{t=1}^{T}, \{z_{t}\}_{t=1}^{T}\},$$

where $[X_{t}]_{t=1}^{T}$ are the latent yield factors and $[z_{t}]_{t=1}^{T}$ are the unobserved regimes. We need to filter the latent yield factors and regimes based on the observed yields $Y$ and the other parameters $\Theta$. We now detail the procedure for drawing each of these parameters.

A.1. Generation of coefficient matrix $\Phi$

Assume the prior distribution of $vec(\Phi)$ is a normal distribution $N(a_0, \Omega_0)$, conditional on all state variables $X$ (a $T \times k$ matrix, $k$ is the number of factors), the observed yields $Y = (y_1, y_2, \ldots, y_T)$, and other parameters $\Psi_{-\Phi}$, the posterior distribution of $vec(\Phi)$ is also a normal distribution $N(a_1, \Omega_1)$, with

$$\Omega_1^{-1}Y, X, \Psi_{-\Phi} = \Omega_0^{-1} + U'U$$

$$a_1|Y, X, \Psi_{-\Phi} = \Omega_1[\Omega_0^{-1}a_0 + U'W].$$

For simplifying the expressions of $U$ and $W$, we define

$$V_i = (I_T \otimes \Sigma_i^{-1/2}),$$

where $i = H$. $L$ represents a high or low volatility regime. $I_T$ is an identity matrix with dimension $T$. $\Sigma_i$ is defined in Eq. (7). Furthermore, let $Z$ be

$$Z_i = z_{i} \otimes I_{k} \text{ for } i = L \text{ or } H,$$

where $z_{i}$ is a $k \times 1$ column vector of $1$s, and $z_{i}$ is a $T \times 1$ column vector from the Hamilton filter consisting of the probabilities in regime $i$. Moreover, let $t_{i}^{1}$ be a $k^2 \times 1$ column vector of $1$s, then we have

$$U = V_L[(X - \mu_L) \otimes I_{k}] \otimes (Z_{L} \otimes t_{i}^{1}) + V_H[(X - \mu_H) \otimes I_{k}] \otimes (Z_{H} \otimes t_{i}^{1})$$

and

$$W = V_L vec(X - \mu_L) \otimes Z_{L} + V_H vec(X - \mu_H) \otimes Z_{H}.$$
A.2. Generation of \( \mu_k \) and \( \mu_l \)

Assume that the prior distribution of \( \mu_k \) (\( \mu_l \)) is a normal distribution \( N(a_0, \Omega_0) \), then the posterior distribution of \( \mu_k \) (\( \mu_l \)) is also a multivariate normal distribution \( N(a_1, \Omega_1) \), with

\[
\Omega_1^{-1} = \Omega_0^{-1} + \Omega_k^{-1},
\]

\[
a_1 = \Omega_1 (\Omega_0^{-1} a_0 + \Omega_k^{-1} \mu_k),
\]

where \( \Omega_k \) is the covariance matrix of \( \mu_k \) (\( \mu_l \)) and \( \mu \) is the average of \( (X_t - \Phi X_{t-1}) \) in regime \( L \) (\( H \)).

A.3. Generation of regimes \( \xi \)

We use the multimove Gibbs sampling method to generate regimes. Based on Carter and Kohn (1994) and Kim and Nelson (1998) partition the joint distribution of regimes \( \xi = (\xi_1, \xi_2, \ldots, \xi_T) \) conditional on \( X \) and other generated parameters \( \Psi_{-\xi} \),

\[
g(\xi|X, \Psi_{-\xi}) = g(\xi_1|X, \Psi_{-\xi}) \prod_{t=1}^T g(\xi_t|\xi_{t+1}, X_t, \Psi_{-\xi}).
\]

The forward filtering and backward sampling (FFBS) approach therefore can be applied in two steps. The first step is to run Hamilton’s (1994) filter to get filtered probabilities \( g(\xi_t|X_t, \Psi_{-\xi}) \). The last iteration of the filter is exactly \( g(\xi_T|X_T, \Psi_{-\xi}) \), from which \( \xi_T \) is generated with a uniform distribution generator. The second step is to generate \( \xi_t \) conditional on \( \xi_{t+1} \) and \( X_t \). We can make use of the following result:

\[
g(\xi_t|\xi_{t+1}, X_t, \Psi_{-\xi}) \propto g(\xi_{t+1}|\xi_t) g(\xi_t|X_t, \Psi_{-\xi}),
\]

combined with the fact that \( g(\xi_{t+1}|\xi_t) \) is the transition probability and \( g(\xi_t|X_t, \Psi_{-\xi}) \) has been provided by the Hamilton filter, we have

\[
g(\xi_t = 1|X_t) = \frac{g(\xi_{t+1} = 1|\xi_t = 1) g(\xi_t = 1|X_t, \Psi_{-\xi})}{\sum_{j=0}^T g(\xi_{t+1} = j|\xi_t = 1) g(\xi_t = j|X_t, \Psi_{-\xi})}.
\]

Then we can generate all regimes recursively.

A.4. Generation of state variables \( X \)

For generating the state vector, we still employ the FFBS approach. Kim and Nelson (1998) employ Carter and Kohn’s multimove Gibbs sampling method and provide the partition of joint distribution. There are also two steps like in the generation of the regimes \( L \), but we run the Kalman filter instead of the Hamilton filter. Given the measurement equation, the state equation, and all observed yields and macro factors \( Y \), the \( X_t \) have a conditional normal posterior distribution:

\[
X_T|\xi_T, Y_T \sim N(X_T|\beta, P_{T|T}),
\]

where \( X_T|\beta \) is the conditional expectation of \( X_T \) from the last step of the Kalman filter. \( P_{T|T} \) is the covariance matrix of \( X_T|\beta \). Consequently, we have

\[
X_{t-1}|\{\xi_{t-1}, X_t, Y_{t-1}\} \sim N(X_{t-1}|\beta, P_{t|t-1}),
\]

with

\[
X_{t|t-1} = X_{t-1} + P_{t|t} (\Phi P_{t|t} \Phi' + \Sigma_t)^{-1} (X_{t-1} - \mu_t - \Phi X_{t-1}),
\]

and

\[
P_{t|t-1} = P_{t-1} - P_{t-1} (\Phi P_{t-1} \Phi' + \Sigma_t)^{-1} \Phi P_{t-1},
\]

where \( \Sigma_t \) is a weighted average of \( \Sigma_0 \) and \( \Sigma_l \), and \( \mu_t \) is a weighted average of \( \mu_k \) and \( \mu_l \). Specifically, we have

\[
\Sigma_t = \Pr(\xi_t = H) \Sigma_H + \Pr(\xi_t = L) \Sigma_L,
\]

\[
\mu_t = \Pr(\xi_t = H) \mu_H + \Pr(\xi_t = L) \mu_L.
\]

A.5. Generation of the rate of factor loading changes \( \lambda \)

To draw \( \lambda \), we use a Random Walk Metropolis step

\[
\lambda_{m+1} = \lambda_m + \zeta v,
\]

where \( v \sim N(0, 1) \) and \( \zeta \) is the scaling factor used to adjust the acceptance rate. The acceptance probability \( \pi \) for \( \lambda \) is given by
\[ x = \min \left\{ \frac{g(\lambda_{m+1}|Y, \Psi_{-j})q(\lambda_{m}|\lambda_{m+1})}{g(\lambda_{m}|Y, \Psi_{-j})q(\lambda_{m}|\lambda_{m-1})}, 1 \right\} = \min \left\{ \frac{g(\lambda_{m+1}|Y, \Psi_{-j})}{g(\lambda_{m}|Y, \Psi_{-j})}, 1 \right\}, \]

where \( q() \) is a symmetric proposal distribution in the Metropolis step. Furthermore, the posterior \( g(\lambda_{m}|Y, \Psi_{-m+1}) \) is given by

\[ g(\lambda_{m}|Y, \Psi_{-j}) \propto g(Y|\Psi)g(\lambda_{m}). \]

Thus, in case of the draw of \( \lambda \), the acceptance rate is the posterior ratio of the new and old draws of \( \lambda \).

### A.6. Generation of diagonal covariance matrix \( \Omega \)

Since \( \Omega \) is diagonal, it can be generated element-by-element. Assume \( \sigma_{i}^{2} \), the \( i \)-th element in the diagonal of \( \Omega \), has an inverted Gamma prior distribution, \( \sigma_{i}^{2} \sim IG(v_{i}/2, \delta_{i}/2) \), the posterior distribution of \( \sigma_{i}^{2} \) is still an inverted Gamma distribution, \( \sigma_{i}^{2} \sim IG(v_{i}/2, \delta_{i}/2) \), with

\[ v_{i} = v_{0} + T \]

and

\[ \delta_{i} = \delta_{0} + (y_{i} - x_{i} \Phi_{i})^{T}(y_{i} - x_{i} \Phi_{i}), \]

where \( y_{i}, x_{i} \), and \( \Phi_{i} \) are the appropriate columns of \( Y \), \( X \) and \( \Phi \).

### A.7. Generation of non-diagonal covariance matrix \( \Sigma_{\beta} \) and \( \Sigma_{L} \)

The covariance matrix is sampled from the inverted Wishart distribution. With an informative prior, the posterior distribution of covariance matrix follows

\[ \Sigma_{\beta|Y, \Psi_{-i}} = \frac{IW(T, i, \sum_{t=1}^{T} \eta_{t}\eta_{t}^{T})}{\sum_{t=1}^{T} \eta_{t}\eta_{t}^{T}}, \]

where \( i \) denotes two regimes.

### A.8. Generation of transition probabilities \( p \) and \( q \)

The conjugate prior distribution for \( p \) and \( q \) is a beta distribution.

\[ p \sim \text{beta}(u_{11}, u_{10}) \]

\[ q \sim \text{beta}(u_{00}, u_{01}). \]

As discussed in Kim and Nelson (1998, pp. 214–215), the posterior distributions are

\[ p \sim \text{beta}(u_{11} + n_{11}, u_{10} + n_{10}) \]

\[ q \sim \text{beta}(u_{00} + n_{00}, u_{01} + n_{01}) \]

with \( n_{ij} \) referring to the transitions from state \( i \) to \( j \), which can be calculated by counting the generated regimes \( \zeta \).

### References


---

The text spans a complex statistical framework involving the Metropolis step, gamma distributions, and beta distributions, aiming to model the dynamics of interest rates and macroeconomic indicators. It integrates various econometric models and theories, including the term structure of interest rates, macroeconomic models, and financial factor models. The references cite key works in the field of macroeconomics and finance, providing a comprehensive backdrop for the methodologies discussed.