Modeling and forecasting serially dependent yield curves

Hao Li[*]

[*]University of Amsterdam

ASSA 2021

January 03–05, 2021

Serially dependent yield curves

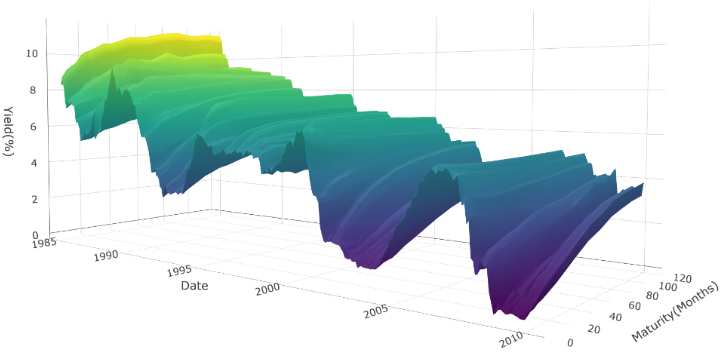

Figure: Monthly unsmoothed Fama-Bliss zero-coupon yields of U.S. Treasuries

Examples of two consecutive yield curves

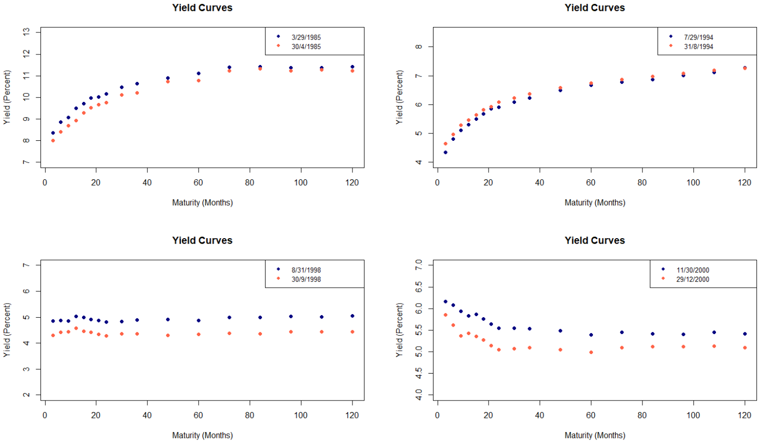

Figure: Monthly unsmoothed Fama-Bliss zero-coupon yields of U.S. Treasuries

Forecasting yield curves

Accurate forecasts of yield curves are of great use for investment decision, risk management, financial derivative pricing and inflation targeting.

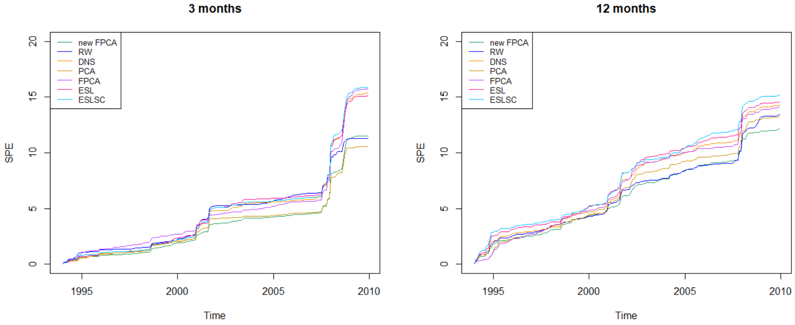

Figure: Cumulative sum of Squared Prediction Errors (SPE) of monthly yield data at maturities of 3, 12, 24, 36, 60, 120 months from January 1994 to December 2009.

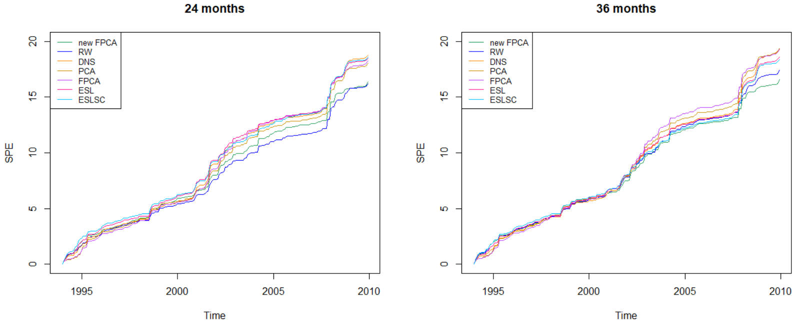

FIGURE: Cumulative sum of Squared Prediction Errors (SPE) of monthly yield data at maturities of 3, 12, 24, 36, 60, 120 months from January 1994 to December 2009.

Forecasting yield curves

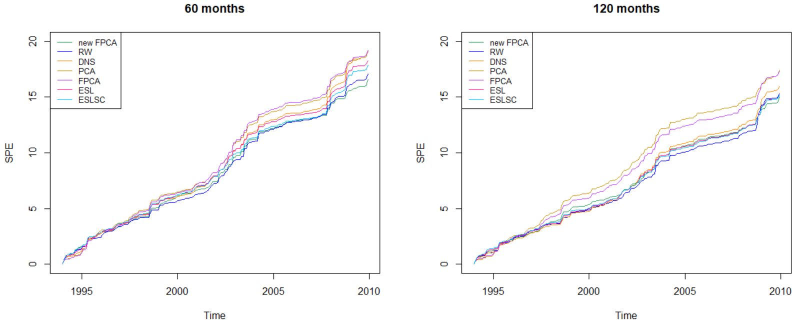

Figure: Cumulative sum of Squared Prediction Errors (SPE) of monthly yield data at maturities of 3, 12, 24, 36, 60, 120 months from January 1994 to December 2009.

Motivation

- Inadequate dimension reduction when the traditional FPCA is applied to serially dependent data (Hörmann et al., JRSS Series B, 2015)

Motivation

- Inadequate dimension reduction when the traditional FPCA is applied to serially dependent data (Hörmann et al., JRSS Series B, 2015)
- Debate on the number of factors underlying yield curve dynamics, see Andreasen et al.(JoE, 2019) and Crump and Gospodinov (2020)

Motivation

- Inadequate dimension reduction when the traditional FPCA is applied to serially dependent data (Hörmann et al., JRSS Series B, 2015)
- Debate on the number of factors underlying yield curve dynamics, see Andreasen et al.(JoE, 2019) and Crump and Gospodinov (2020)
- Yield curve residuals do not adhere to the familiar white noise assumption, see, e.g., Diebold and Li (JoE, 2006), van Dijk et al. (JoAE, 2014), and Andreasen et al. (JoE, 2019)

Preview the main findings

Literature

Preview the main findings

2 PCs here VS 16 PCs from the traditional FPCA

Preview the main findings

Selection of dimension and lag order

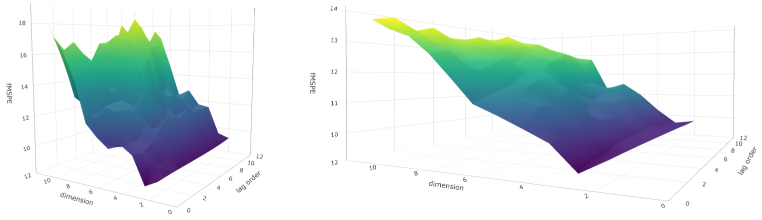

Figure: Three-dimensional surface plots of functional Mean Squared Prediction Errors (fMSPEs) depending on different values of dimension and lag order. The minimum value of fMSPEs is reached when $d = 3$ and $p = 1$ for both the validation set and the time series cross-validation data.

Preview the main findings

Favorable in-sample and out-of-sample properties

- Yield curve residuals from this new model's fit exhibit less autocorrelation and have zero mean
- The forecasts of this new model have superiority:
  (i) less non-zero mean in prediction errors,
  (ii) less autocorrelated prediction errors at different maturities, and
  (iii) smaller root mean squared prediction errors (RMSPE)
  over time and across term structure of interest rates at the
  1-month-ahead horizon.

Contribution

- First work considering factors that drive serial dependence across yield curves into the modeling, estimation and forecasting of the term structure of interest rates.

Contribution

- First work considering factors that drive serial dependence across yield curves into the modeling, estimation and forecasting of the term structure of interest rates.
- A data-driven method is proposed to determine the lag order and dimensionality of yield curves simultaneously.

Contribution

- First work considering factors that drive serial dependence across yield curves into the modeling, estimation and forecasting of the term structure of interest rates.
- A data-driven method is proposed to determine the lag order and dimensionality of yield curves simultaneously.
- This new model is the most preferred from in-sample and out-of-sample perspectives.

Preview the main findings

Literature

Literature

- Parametric approach:
  Chambers et al. (JFQA, 1984), Nelson and Siegel (JB, 1987), Svensson (NBER, 1994), Duffee (JF, 2002), Diebold and Li (JoE, 2006), Christensen et al. (JoE, 2011), Joslin et al. (RFS, 2011), Van Dijk et al. (JoAE, 2014), Joslin et al. (JF, 2014), Jungbacker et al. (JoAE, 2014), Almeida et al. (JoFE, 2018), Andreasen et al.(JoE, 2019)

To the best of my knowledge, no existing/recent work employs the same concept of dealing with serial dependence of yield curves

Literature

- Parametric approach:
  Chambers et al. (JFQA, 1984), Nelson and Siegel (JB, 1987), Svensson (NBER, 1994), Duffee (JF, 2002), Diebold and Li (JoE, 2006), Christensen et al. (JoE, 2011), Joslin et al. (RFS, 2011), Van Dijk et al. (JoAE, 2014), Joslin et al. (JF, 2014), Jungbacker et al. (JoAE, 2014), Almeida et al. (JoFE, 2018), Andreasen et al.(JoE, 2019)
- Nonparametric approach:
  Linton et al. (JoE, 2001), Hays et al. (AAS, 2012), Bardsley et al. (EJ, 2017), Caldeira and Torrent (JoF, 2017), Otto and Salish (2019), Sen and Klüppelbergg (2019), Crump and Gospodinov (2020), and Koo et al. (JoE, 2020)

To the best of my knowledge, no existing/recent work employs the same concept of dealing with serial dependence of yield curves

Literature

- Parametric approach:
  Chambers et al. (JFQA, 1984), Nelson and Siegel (JB, 1987), Svensson (NBER, 1994), Duffee (JF, 2002), Diebold and Li (JoE, 2006), Christensen et al. (JoE, 2011), Joslin et al. (RFS, 2011), Van Dijk et al. (JoAE, 2014), Joslin et al. (JF, 2014), Jungbacker et al. (JoAE, 2014), Almeida et al. (JoFE, 2018), Andreasen et al.(JoE, 2019)
- Nonparametric approach:
  Linton et al. (JoE, 2001), Hays et al. (AAS, 2012), Bardsley et al. (EJ, 2017), Caldeira and Torrent (JoF, 2017), Otto and Salish (2019), Sen and Klüppelbergg (2019), Crump and Gospodinov (2020), and Koo et al. (JoE, 2020)
- Semiparametric approach:
  Bowsher and Meeks (JASA, 2008) and Härdle and Majer (EJF, 2016)

To the best of my knowledge, no existing/recent work employs the same concept of dealing with serial dependence of yield curves

Outline

1. Methodology

Outline

1. Methodology
2. Empirical illustrations

Outline

1. Methodology
2. Empirical illustrations
3. Conclusions

1. Methodology

New dynamic functional factor model

The sequence of observed yield curves

$$Y_t(u) = X_t(u) + \epsilon_t(u), \qquad t = 1, \ldots, T, \qquad u \in [a, b],$$

where $\epsilon_t(\cdot)$ is assumed to be the noise term, in the sense that (i) $\mathbb{E}\{\epsilon_t(u)\} = 0$ for all t and all $u \in [a, b]$; (ii) $\mathrm{Cov}\{\epsilon_t(u), \epsilon_{t+k}(v)\} = 0$ for all $u, v \in [a, b]$ provided that $k \neq 0$, (iii) $\mathrm{Cov}\{X_t(u), \epsilon_{t+k}(v)\} = 0$ for all $u, v \in [a, b]$ and $k \neq 0$ and $Y_t(u)$ is square-integrable on the domain $[a, b]$ in a Hilbert space $\mathcal{L}_2$.

New dynamic functional factor model

DEPENDING ON THE SERIAL DEPENDENCE ACROSS YIELD CURVES can be defined as

$$Y_t(u) = \mu(u) + \sum_{j=1}^{d} \eta_{tj}\psi_j(u) + \epsilon_t(u), \qquad t = 1, \ldots, T, \qquad u \in [a, b],$$

where $\mu(u) \equiv \mathbb{E}\{Y_t(u)\}$, η_{tj} denotes the jth factor at time t, analogous to factors in the dynamic Nelson-Siegel (NS) model proposed by Diebold and Li (2006), and $\psi_j(u)$ denotes the jth loading for time to maturity u. Here, $\psi_1(u)$,..., $\psi_d(u)$ are eigenfunctions of the non-negative operator

$$K(u, v) = \sum_{k=1}^{p} \int_a^b c_k(u, z)\, c_k(v, z)\, \mathrm{d}z \qquad k = 1, \ldots, p,$$

where $c_k(u, z) \equiv \mathrm{Cov}\{X_t(u), X_{t+k}(v)\}$ is the kth lag autocovariance function of the yield curve process.

Covariance Kernel:

$$c_Y(u,v) = c_X(u,v) + c_\epsilon(u,v)$$

$\widehat{c}_Y(u,v)$ is not a consistent estimator of $c_X(u,v)$.

However,
the Autocovariance Kernel:

$$c_k(u,v), k \in \mathbb{Z}$$

of $X(\cdot)$ and of $Y(\cdot)$ are equal for $k \neq 0$.
$\widehat{c}_k(u,v)$ using observed yield data is a consistent estimator for the true yield curve process.

New dynamic functional factor model

Following Bathia et al. (2010, Annals of Statistics) to find a d-dimensional orthonormal system in a square-integrable function space, let

$$\widehat{K}(u,v) = \sum_{k=1}^{p} \int_a^b \widehat{c}_k(u,z)\,\widehat{c}_k(v,z)\,\mathrm{d}z, \qquad k = 1,\dots,p, \qquad u,v,z \in [a,b],$$

where

$$\widehat{c}_k(u,v) := \frac{1}{(T-p)} \sum_{t=1}^{T-p} (Y_t(u) - \widehat{\mu}(u))(Y_{t+k}(v) - \widehat{\mu}(v)),$$

be the k lags sample autocovariance operator. Then using the observed yields data, $\widehat{K}(u,v)$ can be written as

$$\widehat{K}(u,v) = \frac{1}{(T-p)^2} \sum_{t,s=1}^{T-p} \sum_{k=1}^{p} (Y_t(u) - \widehat{\mu}(u))(Y_s(v) - \widehat{\mu}(v))$$
$$\times \langle Y_{t+k}(u) - \widehat{\mu}(u), Y_{s+k}(v) - \widehat{\mu}(v)\rangle, \qquad k = 1,\dots,p, \qquad u,v,z \in [a,b]. \tag{1}$$

However, $\widehat{K}(u, v)$ may be an $\infty \times \infty$ matrix if the centered yield curve $Y_t(\cdot) - \widehat{\mu}(\cdot)$ is a functional curve evaluated at fine grid such as an $\infty \times 1$ vector on the domain [a,b].
Thus, applying the duality property here to make the eigenanalysis tractable in a $\mathcal{L}_2$. That is the $\infty \times \infty$ matrix of $\widehat{K}(u, v)$ has the same $\widehat{d}$ nonzero eigenvalues of the $(T-p) \times (T-p)$ matrix

$$\widehat{\mathbf{K}}(u, v) := \frac{1}{(T-p)^2} \sum_{k=1}^{p} \mathbf{Y}_k \mathbf{Y}_0,$$

where $\mathbf{Y}_k$ is the $(T-p) \times (T-p)$ matrix with the (t, s)th element $\langle Y_{t+k} - \widehat{\mu}, Y_{s+k} - \widehat{\mu} \rangle$ for $k = 0, \dots, p$.
Furthermore, let $\widehat{\boldsymbol{\gamma}}_j = (\widehat{\gamma}_{1,j}, \dots, \widehat{\gamma}_{T-p,j})$, $j = 1, \dots, \widehat{d}$, be the eigenvectors of $\widehat{\mathbf{K}}(u, v)$ corresponding to the $\widehat{d}$ largest nonzero eigenvalue $\widehat{\theta}_1, \dots, \widehat{\theta}_{\widehat{d}}$. Then, the eigenfunctons of $\widehat{K}(u, v)$ can be defined by

$$\widehat{\psi}_j := \sum_{t=1}^{T-p} \widehat{\gamma}_{t,j} \times (Y_t(\cdot) - \widehat{\mu}(\cdot)), \qquad j = 1, \dots, \widehat{d},$$

where $\widehat{\psi}_j$ satisfy the identity $\int \widehat{K}(u, v) \widehat{\psi}_j(v) = \widehat{\theta}_j \widehat{\psi}(v)$.

New dynamic functional factor model

Thus, the fitted yield curve can be defined by

$$\widehat{Y}_t(u) = \widehat{\mu}(u) + \sum_{j=1}^{\widehat{d}} \widehat{\eta}_{tj}\widehat{\psi}_j(u), \qquad t = 1, \ldots, T, \qquad j = 1, \ldots, \widehat{d}, \qquad u \in [a, b],$$

where the eigenfunction $\widehat{\psi}_j$ is the jth empirical functional principal component (FPC), the empirical FPC score $\widehat{\eta}_{tj} = \langle Y_t - \mu, \widehat{\psi}_j \rangle$ is analogous to the factor $\beta_{i,t}$, $i = 1, 2, 3$, in the dynamic NS model.

Link with the dynamic NS model:
Let $\mu(u) = 0$, $d = 3$, $\psi_1(u) = 1$, $\psi_2(u) = \frac{1-\mathrm{e}^{-\lambda u}}{\lambda u}$, $\psi_3(u) = \frac{1-\mathrm{e}^{-\lambda u}}{\lambda u} - \mathrm{e}^{-\lambda u}$, where the maturities $u = i$, i ranges from the maturities $\{3, \ldots, 120\}$ and $\lambda = 0.0609$.

CONSISTENCY OF ESTIMATORS

Following the regularity conditions in Bathia et al. (AoS, 2010) such as ψ-mixing condition, $\mathbb{E}\{\int_a^b Y_t(u)^2\, du\}^2 < \infty$, $\mathrm{Cov}\{X_t(u), \epsilon_{t+k}(v)\} = 0$ for all $u, v \in [a, b]$ and $k \in \mathbb{Z}^+$, and the non-zero eigenvalues θ_j, $j \in \mathbb{Z}^+$ of operator K are monotonically decreasing as j increases. Let conditions above hold true. Then as $T \to \infty$, the following theorem hold,

(a) $\sup_{u\in[a,b]} |\widehat{\mu}(u) - \mu(u)| = O_P(T^{-1/2})$,

(b) The Hilbert-Schmidt norm for the operator $\widehat{K} - K$, $\|\widehat{K} - K\|_{\mathcal{S}} = O_P(T^{-1/2})$,

(c) $|\widehat{\theta}_j - \theta_j| = O_P(T^{-1/2})$,

(d) $\left(\int_a^b \{\widehat{\psi}_j(u) - \psi_j(u)\}^2\, du\right)^{1/2} = O_P(T^{-1/2})$,

for all $j = 1, \ldots, d$ *and* $\theta_1 > \ldots > \theta_d > 0$.

(e)$|\widehat{\theta}_j - \theta_j| = O_P(T^{-1})$,

(f) $\left(\int_a^b \{\widehat{\psi}_j(u) - \sum_{j=d+1}^{\infty} \langle \psi_j, \widehat{\psi}_j \rangle \psi_j(u)\}^2\, \mathrm{d}u\right)^{1/2} = O_P(T^{-1/2})$,

for all $j > d$*,and* $j \in \mathbb{Z}^+$.

And the discrepancy between $\widehat{\mathcal{M}} = \mathrm{span}\{\widehat{\psi}_1, \ldots, \widehat{\psi}_d\}$ *and* $\mathcal{M} = \mathrm{span}\{\psi_1, \ldots, \psi_d\}$,

$D(\widehat{\mathcal{M}}, \mathcal{M}) = \sqrt{1 - \frac{1}{d}\sum_{j=k=1}^{d} (\langle \psi_k, \widehat{\psi}_j \rangle)^2}$

THE PROCEDURE AIMING AT PREDICTING THE THE SERIALLY DEPENDENT YIELD CURVES

- Transform the observed yield curves $Y_1, \ldots, Y_t$ into a d-dimensional vector time series of the empirical FPC scores $\widehat{\boldsymbol{\eta}}_t = (\widehat{\eta}_{t,1}, \ldots, \widehat{\eta}_{t,d})'$, where d and p are fixed.
- Use a d-dimensional VAR model without the constant term for a stationary process of empirical FPC scores to forecast the h-step-ahead out-of-sample FPC scores.
- Employ the h-step-ahead out-of-sample FPC scores forecast $\widehat{\boldsymbol{\eta}}_{t+h}$ to yield the h-step-ahead yield curve forecast $\widehat{Y}_{t+h}$ according to the Karhunen–Loève expansion, jointly with the d-dimensional empirical functional principal component(s) $\psi_j(u)$ and the empirical mean function of the observed yield curves $\widehat{\mu}(u)$ for $u \in [a, b]$ available at time t.

Decomposing functional squared prediction error

As the eigenfunctions ψ_j are orthonormal and FPC scores η_{tj} are uncorrelated by construction, according to Aue et al.(2015, JASA), the h-step-ahead functional squared prediction error (fSPE) can be decomposed as

$$\begin{aligned}
\mathbb{E}\{\|Y_{t+h}-\widehat{Y}_{t+h}\|^2\} &= \mathbb{E}\left\{\left\|\sum_{j=1}^{\infty}\eta_{t+h,j}\psi_j-\sum_{j=1}^{d}\widehat{\eta}_{t+h,j}\psi_j\right\|^2\right\} \\
&= \mathbb{E}\left\{\left\|\boldsymbol{\eta}_{t+h}-\widehat{\boldsymbol{\eta}}_{t+h}\right\|^2\right\}+\sum_{j=d+1}^{\infty}\theta_j \qquad (2) \\
&\approx \frac{t+pd}{t-pd}\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{e}})+\sum_{j=d+1}^{\infty}\widehat{\theta}_j,
\end{aligned}$$

where $\|\cdot\|$ represents the Euclidean norm and $\|\psi_j\|=1$ for orthonormal eigenfunctions, $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{e}}$ is obtained from a d-dimensional VAR model for a stationary process of empirical FPC scores, $\widehat{\boldsymbol{\eta}}_{t+h}=\sum_{k=1}^{q}\widehat{\boldsymbol{\Gamma}}_k\widehat{\boldsymbol{\eta}}_{T+h-k}+\widehat{\boldsymbol{e}}_{t+h}$, where $\boldsymbol{\Gamma}_k$ is a $d\times d$ matrix of coefficients.

Forecasting the serially dependent yield curves

Functional Mean Squared Prediction Error

Similar to studies such as Hyndman and Ullah (CSDA, 2007) and Aue et al. (JASA, 2015), I use the functional MSPE (fMSPE) in the setting of functional data analysis, and thus the fMSPE jointly determining the dimension d and lag order d can be defined as

$$\text{fMSPE}(p, d) = \frac{1}{P} \sum_{t=1}^{P} \int_a^b (Y_{t+h}(u) - \widehat{Y}_{t+h}(u))^2 \, du, \qquad u \in [a, b], \quad (2)$$

where P is the size of validation set. To choose the dimension and lag order of the yield curves, I use information set $I_t = \{Y_1, \ldots, Y_{L+m}\}$, where L is the size of training set, which is big enough to produce reliable empirical FPC scores and thus forecasts of yield curve, and $m = 0, \ldots, P-1$, to obtain the next P periods of h-step-ahead yield curve forecasts $\widehat{Y}_{L+m+h}$. The lag order p and the dimension d of the observed yield curves, therefore, can be simultaneously determined by the minimum of the fMSPEs produced by different combinations of p and d.

Finally, using the selected combination of d and p to make the h-step ahead out-of-sample forecasts of yield curves and evaluating this novel method's performance over $T - P - L$ periods, where $T - P - L$ is the size of test set.

2. Empirical illustrations

Data

- Monthly unsmoothed Fama-Bliss zero-coupon yields of U.S. Treasuries from January 1985 to December 2009 ($T = 300$) used in van Dijk et al (JoAE, 2014)
- 17 fixed maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, and 120 months
- The conventional cubic B-spline expansion is applied to create $Y_t(u)$.
- d and p are determined according to the first 108 observations of yield curves from January 1985 to December 1993
- The size of training set L is set to 60 to ensure that the selection of d and p based on the validation set is effective

Factor loadings

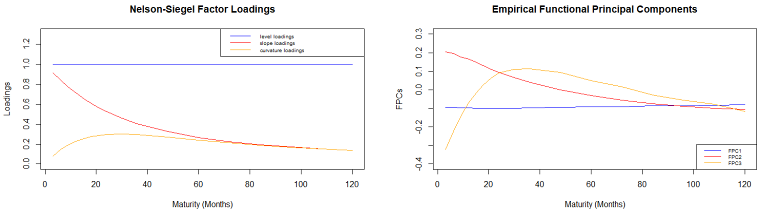

Figure: Nelson-Siegel factor loadings and the empirical functional principal components that account for serial dependence (lag order $p = 1$) across yield curves. The data set is from January 1985 to December 2009.

Empirical illustrations

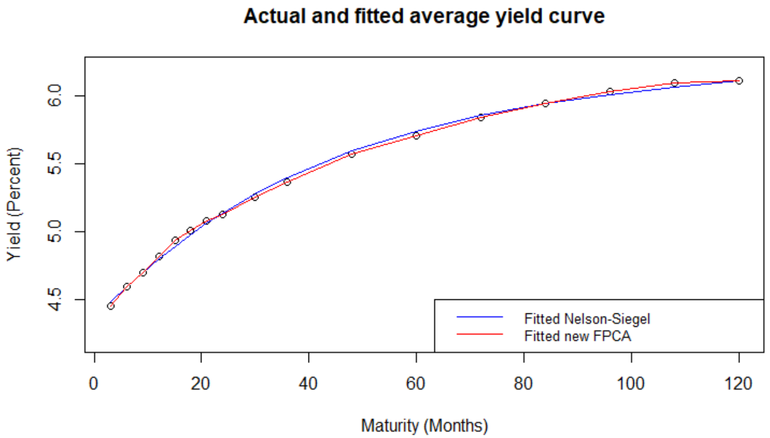

Figure: Actual and fitted yield curves. Blue line: DNS ($\lambda = 0.0609$). Red line: new functional factor model based on FPCA.

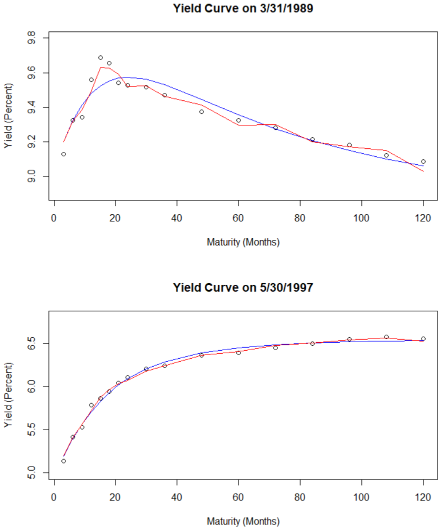

Figure: Actual and fitted yield curves. Blue line: DNS ($\lambda = 0.0609$). Red line: new FPCA.

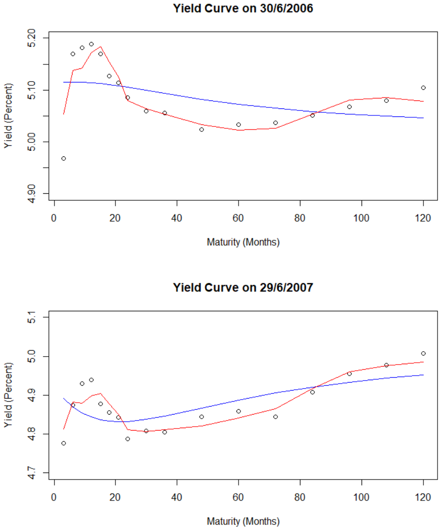

FIGURE: Actual and fitted yield curves. Blue line: DNS ($\lambda = 0.0609$). Red line: new FPCA.

Fitting yield curves

Table: Descriptive statistics for yield curve residuals from the new model of three dynamic functional factors, $p = 1$

Maturity (Month)	Mean	SD	Min	Max	MAE	RMSE	$\widehat{\rho}(1)$	$\widehat{\rho}(12)$
3	-0.000	0.152	-0.705	0.287	0.112	0.152	0.826	0.182
6	0.000	0.047	-0.187	0.173	0.035	0.047	0.517	0.224
9	-0.000	0.053	-0.188	0.205	0.040	0.053	0.665	0.171
12	-0.000	0.069	-0.167	0.318	0.051	0.069	0.555	0.173
15	0.000	0.061	-0.190	0.390	0.042	0.061	0.648	0.116
18	-0.000	0.048	-0.152	0.212	0.036	0.048	0.647	0.159
21	0.000	0.044	-0.143	0.179	0.032	0.044	0.654	0.133
24	-0.000	0.046	-0.189	0.201	0.033	0.046	0.670	0.146
30	0.000	0.029	-0.178	0.125	0.021	0.029	0.400	0.077
36	-0.000	0.028	-0.076	0.103	0.021	0.028	0.554	-0.073
48	0.000	0.043	-0.316	0.148	0.029	0.043	0.687	0.080
60	-0.000	0.035	-0.100	0.125	0.027	0.035	0.630	-0.044
72	0.000	0.045	-0.237	0.152	0.032	0.045	0.851	0.194
84	-0.000	0.038	-0.142	0.250	0.025	0.038	0.651	0.006
96	0.000	0.025	-0.107	0.145	0.017	0.025	0.608	0.100
108	0.000	0.031	-0.113	0.108	0.024	0.031	0.745	0.022
120	-0.000	0.071	-0.241	0.322	0.052	0.071	0.813	0.171

Fitting yield curves

Table: Descriptive statistics for yield curve residuals from Diebold and Li (2006)'s Nelson–Siegel model

Maturity (Month)	Mean	SD*	Min	Max	MAE*	RMSE*	$\widehat{\rho}(1)^*$	$\widehat{\rho}(12)^*$
3	-0.036	1.526	-0.507	0.174	1.446	1.435	0.995	0.532
6	-0.001	1.148	-0.138	0.218	1.180	1.148	1.455	0.749
9	0.001	0.751	-0.198	0.287	0.728	0.751	0.859	0.436
12	**0.016**	0.908	-0.171	0.363	0.860	0.889	0.877	0.596
15	**0.043**	1.017	-0.178	0.359	0.739	0.827	0.918	0.479
18	**0.033**	1.253	-0.111	0.147	0.846	0.951	1.014	0.556
21	**0.020**	1.583	-0.095	0.122	1.179	1.278	1.464	2.750
24	-**0.01**2	1.242	-0.162	0.097	1.176	1.182	1.218	1.411
30	-**0.018**	0.760	-0.200	0.119	0.673	0.684	0.732	0.918
36	-**0.031**	0.619	-0.184	0.134	0.482	0.514	0.774	-0.530
48	-**0.021**	0.642	-0.477	0.153	0.587	0.612	0.874	0.357
60	-**0.036**	0.736	-0.175	0.135	0.562	0.587	0.854	-0.463
72	-**0.014**	0.708	-0.366	0.216	0.719	0.691	0.945	0.521
84	**0.003**	0.969	-0.140	0.234	0.934	0.966	1.018	0.127
96	**0.023**	0.898	-0.097	0.168	0.562	0.689	0.972	0.741
108	**0.030**	0.814	-0.123	0.133	0.622	0.643	0.981	0.416
120	0.001	0.869	-0.275	0.380	0.898	0.869	0.931	0.491

Fitting yield curves

Table: Descriptive statistics for yield curve residuals from Otto and Salish (2019)'s dynamic functional factor model

Maturity (Month)	Mean	SD*	Min	Max	MAE*	RMSE*	$\widehat{\rho}(1)^*$	$\widehat{\rho}(12)^*$
3	**-0.051**	0.866	-0.830	0.287	0.828	0.832	0.989	0.878
6	**-0.037**	0.537	-0.368	0.305	0.479	0.495	0.835	1.066
9	**-0.033**	0.927	-0.185	0.162	0.754	0.801	0.860	0.608
12	**-0.038**	1.035	-0.246	0.277	0.833	0.896	0.837	0.666
15	**-0.026**	0.868	-0.286	0.350	0.772	0.812	0.976	0.577
18	**-0.021**	0.854	-0.233	0.173	0.799	0.801	0.929	0.573
21	**-0.014**	0.860	-0.179	0.148	0.837	0.830	0.975	0.494
24	**-0.014**	1.025	-0.185	0.136	0.983	0.976	0.903	0.519
30	**-0.016**	0.888	-0.218	0.127	0.820	0.796	1.031	0.905
36	**-0.012**	0.945	-0.110	0.097	0.851	0.876	0.956	0.952
48	**-0.009**	0.991	-0.325	0.148	0.961	0.973	0.997	1.433
60	**-0.005**	1.015	-0.103	0.121	1.013	1.003	1.013	0.836
72	-0.004	1.003	-0.241	0.147	1.023	0.999	1.002	1.093
84	-0.002	1.005	-0.136	0.235	0.981	1.003	1.000	0.538
96	-0.001	1.001	-0.100	0.126	0.936	1.000	0.977	0.660
108	-0.000	0.964	-0.132	0.095	0.968	0.964	0.988	0.574
120	-0.000	0.987	-0.239	0.301	0.951	0.987	0.998	1.040

Fitting yield curves

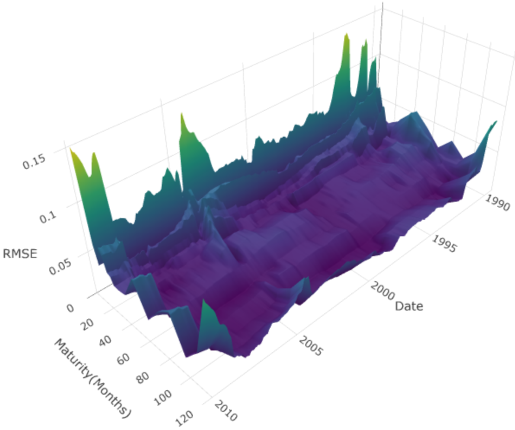

Figure: Five years rolling RMSEs, $p = 1$ and $d = 3$

Robustness checks

Table: Summary statistics for yield curve residuals from the new functional factor model ($d = 4$, $p = 1$)

Maturity (Month)	Panel A: relative to DNS				Panel B: relative to FPCA			
	MAE*	RMSE*	$\widehat{\rho}(1)^*$	$\widehat{\rho}(12)^*$	MAE*	RMSE*	$\widehat{\rho}(1)^*$	$\widehat{\rho}(12)^*$
3	0.585	0.556	0.699	0.530	0.335	0.322	0.695	0.874
6	0.789	0.774	0.664	0.630	0.320	0.334	0.381	0.897
9	0.635	0.646	0.772	0.323	0.658	0.688	0.773	0.450
12	0.562	0.566	0.388	0.570	0.545	0.570	0.370	0.637
15	0.547	0.562	0.646	0.626	0.572	0.552	0.687	0.753
18	0.575	0.609	0.659	0.631	0.543	0.513	0.604	0.651
21	0.766	0.829	0.789	-1.093	0.544	0.538	0.525	-0.196
24	0.774	0.782	0.662	-0.572	0.647	0.646	0.491	-0.210
30	0.631	0.655	0.615	0.258	0.769	0.763	0.866	0.254
36	0.482	0.513	0.774	-0.526	0.851	0.876	0.955	0.945
48	0.391	0.422	0.641	-0.080	0.640	0.670	0.732	-0.320
60	0.476	0.507	0.762	-1.139	0.858	0.864	0.904	2.055
72	0.464	0.423	0.754	0.066	0.659	0.611	0.799	0.139
84	0.929	0.967	1.020	0.044	0.976	1.004	1.002	0.185
96	0.551	0.664	0.972	0.439	0.918	0.964	0.978	0.391
108	0.451	0.543	0.917	-1.987	0.702	0.813	0.924	-2.740
120	0.547	0.526	0.690	0.256	0.580	0.597	0.740	0.543

Robustness checks

Table: Summary statistics for yield curve residuals from the new functional factor model ($d = 3$, $p = 2$)

Maturity (Month)	Panel A: relative to DNS				Panel B: relative to FPCA			
	MAE*	RMSE*	$\widehat{\rho}(1)^*$	$\widehat{\rho}(12)^*$	MAE*	RMSE*	$\widehat{\rho}(1)^*$	$\widehat{\rho}(12)^*$
3	1.430	1.428	0.995	0.540	0.819	0.828	0.989	0.890
6	1.168	1.138	1.444	0.745	0.474	0.491	0.828	1.062
9	0.727	0.753	0.862	0.454	0.753	0.803	0.863	0.634
12	0.859	0.890	0.877	0.593	0.833	0.898	0.838	0.664
15	0.735	0.828	0.920	0.462	0.768	0.812	0.978	0.557
18	0.842	0.950	1.015	0.552	0.795	0.800	0.930	0.569
21	1.173	1.275	1.461	2.710	0.832	0.828	0.972	0.486
24	1.170	1.177	1.215	1.350	0.978	0.972	0.901	0.496
30	0.671	0.682	0.725	0.880	0.817	0.794	1.021	0.868
36	0.483	0.514	0.774	-0.519	0.853	0.878	0.955	0.931
48	0.588	0.617	0.876	0.375	0.962	0.980	1.000	1.504
60	0.563	0.589	0.855	-0.460	1.015	1.005	1.015	0.830
72	0.719	0.692	0.945	0.522	1.022	0.999	1.002	1.096
84	0.934	0.966	1.018	0.110	0.981	1.004	1.000	0.466
96	0.565	0.691	0.970	0.760	0.941	1.003	0.976	0.677
108	0.626	0.646	0.983	0.484	0.974	0.968	0.990	0.667
120	0.901	0.873	0.932	0.496	0.954	0.992	0.999	1.050

Empirical illustrations

Robustness checks

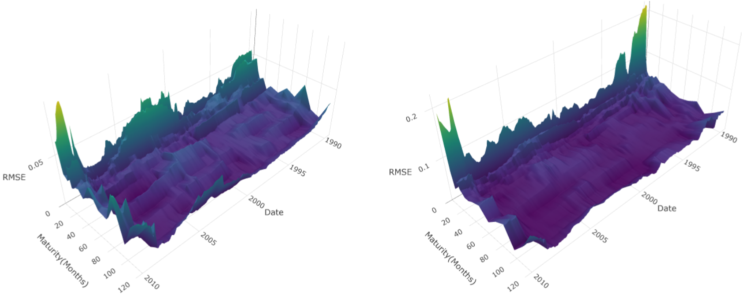

Figure: Five years rolling RMSEs. ($p = 1$, $d = 4$) and ($p = 2$ and $d = 3$).

Forecasting yield curves

Table: Out-of-sample 1-month-ahead forecasting results, new functional factor model ($p = 1$, $d = 3$), $\mathrm{RMSPE}_i = \sqrt{\frac{1}{P}\sum_{t \in P}(Y_{t+1,i} - \widehat{Y}_{t+1|t}(u_i))^2}$, $u_i = i$

Maturity (Month)	Mean	SD	RMSPE	$\widehat{\rho}(1)$	$\widehat{\rho}(12)$
3	**-0.052**	0.240	0.245	0.202	0.008
6	-0.004	0.217	0.217	0.107	-0.041
9	0.014	0.235	0.235	0.136	-0.016
12	0.016	0.252	0.252	0.079	-0.025
15	0.008	0.263	0.263	0.100	0.009
18	0.009	0.273	0.272	0.081	0.004
21	0.011	0.284	0.283	0.095	-0.002
24	0.013	0.292	0.292	0.101	-0.006
30	0.004	0.295	0.294	0.066	0.007
36	-0.005	0.294	0.294	0.055	0.004
48	-0.007	0.306	0.305	0.066	0.034
60	-0.011	0.295	0.294	0.060	-0.015
72	-0.007	0.295	0.294	0.053	-0.002
84	0.0001	0.281	0.280	0.001	-0.029
96	-0.016	0.281	0.280	-0.006	-0.028
108	-0.006	0.278	0.277	0.015	-0.006
120	-0.001	0.280	0.279	0.023	0.040

TABLE: Out-of-sample 1-month-ahead forecasting results relative to the new model

Maturity	Panel A: RW					Panel B: DNS				
	Mean	SD*	RMSPE*	$\widehat{\rho}(1)^*$	$\widehat{\rho}(12)^*$	Mean	SD*	RMSPE*	$\widehat{\rho}(1)^*$	$\widehat{\rho}(12)^*$
3	-0.016	1.012	0.991	1.286	8.983	**-0.115**	1.082	1.157	1.964	12.975
6	-0.016	1.093	1.095	3.325	0.414	**-0.056**	1.136	1.164	4.288	-0.431
9	-0.016	1.060	1.060	2.672	3.755	**-0.045**	1.103	1.118	3.542	0.232
12	-0.016	1.050	1.050	3.899	1.951	-0.032	1.078	1.084	5.478	-0.031
15	-0.016	1.033	1.034	3.118	0.185	-0.030	1.094	1.100	4.440	8.250
18	-0.016	1.021	1.022	3.557	-0.400	-0.040	1.068	1.078	4.933	14.106
21	-0.016	1.000	1.001	2.867	-1.861	**-0.049**	1.050	1.063	3.915	-11.337
24	-0.016	0.995	0.995	2.521	1.702	**-0.073**	1.043	1.071	3.568	-2.080
30	-0.014	1.014	1.015	3.520	1.486	**-0.085**	1.040	1.080	4.624	4.553
36	-0.015	1.024	1.025	3.840	1.378	**-0.091**	1.033	1.078	4.820	5.717
48	-0.014	1.018	1.019	2.303	0.987	**-0.091**	1.036	1.078	3.792	1.262
60	-0.013	1.014	1.014	2.382	-1.100	**-0.099**	1.016	1.070	3.378	0.217
72	-0.012	1.013	1.014	2.306	-6.771	**-0.093**	1.035	1.082	4.023	-16.386
84	-0.011	1.023	1.024	116.196	0.473	**-0.055**	1.024	1.043	165.069	0.496
96	-0.010	1.036	1.035	-14.466	0.559	-0.031	1.028	1.032	-19.280	0.432
108	-0.010	1.003	1.003	7.434	6.098	-0.025	1.021	1.025	9.696	0.484
120	-0.010	1.010	1.011	2.148	-0.545	-0.033	1.026	1.032	5.565	0.709

Table: Out-of-sample 1-month-ahead forecasting results relative to the new model

Maturity	Panel C: FPCA					Panel D: PCA				
	Mean	SD*	RMSPE*	$\widehat{\rho}(1)^*$	$\widehat{\rho}(12)^*$	Mean	SD*	RMSPE*	$\widehat{\rho}(1)^*$	$\widehat{\rho}(12)^*$
3	**-0.147**	1.029	1.170	1.062	7.499	**-0.074**	0.929	0.957	0.628	-0.372
6	**-0.096**	1.017	1.109	0.649	0.133	**-0.045**	1.005	1.026	1.205	-0.611
9	**-0.087**	1.012	1.075	0.553	-0.375	**-0.045**	1.022	1.038	1.016	-3.301
12	**-0.094**	1.013	1.078	0.422	0.916	**-0.057**	1.023	1.046	0.897	-0.487
15	**-0.097**	1.016	1.081	0.699	2.211	**-0.074**	1.022	1.059	0.787	4.183
18	**-0.097**	1.015	1.076	0.692	4.059	**-0.079**	1.018	1.058	0.731	7.602
21	**-0.092**	1.013	1.063	0.708	-3.821	**-0.082**	1.015	1.055	0.798	-6.563
24	**-0.091**	1.010	1.056	0.622	-0.588	**-0.084**	1.012	1.051	0.744	-1.902
30	**-0.105**	1.012	1.073	0.655	1.636	**-0.095**	1.021	1.070	0.991	2.664
36	**-0.111**	1.011	1.079	0.366	2.307	**-0.103**	1.022	1.081	0.948	3.833
48	**-0.110**	1.010	1.072	0.353	1.130	**-0.103**	1.025	1.079	1.018	1.320
60	**-0.111**	1.009	1.076	0.168	0.148	**-0.105**	1.013	1.073	0.541	-0.435
72	**-0.106**	1.022	1.084	0.476	-7.025	**-0.098**	1.021	1.074	0.790	-10.159
84	**-0.098**	1.024	1.082	-20.854	0.652	**-0.091**	1.021	1.071	-16.255	0.214
96	**-0.111**	1.025	1.097	4.535	0.895	**-0.104**	1.020	1.083	3.497	0.301
108	**-0.100**	1.023	1.084	-0.492	2.545	**-0.092**	1.023	1.075	0.478	-1.847
120	**-0.091**	1.024	1.074	-0.093	0.304	**-0.082**	1.037	1.078	1.522	1.546

Table: Out-of-sample 1-month-ahead forecasting results relative to the new model

Maturity	Panel A: ESL					Panel B: ESLSC				
	Mean	SD*	RMSPE*	$\widehat{\rho}(1)^*$	$\widehat{\rho}(12)^*$	Mean	SD*	RMSPE*	$\widehat{\rho}(1)^*$	$\widehat{\rho}(12)^*$
3	**-0.090**	1.111	1.147	2.144	9.533	**-0.082**	1.153	1.175	2.428	16.240
6	-0.031	1.163	1.172	4.603	0.216	-0.018	1.212	1.215	5.070	-1.208
9	-0.020	1.123	1.124	3.744	2.190	-0.004	1.166	1.164	4.000	-2.352
12	-0.008	1.095	1.094	5.915	0.905	0.010	1.118	1.117	6.073	-0.790
15	-0.005	1.106	1.106	4.775	5.810	0.015	1.117	1.118	4.757	8.248
18	-0.016	1.078	1.079	5.367	6.839	0.005	1.090	1.089	5.312	14.210
21	-0.024	1.059	1.062	4.295	0.844	-0.003	1.069	1.068	4.230	-11.615
24	**-0.048**	1.051	1.063	3.925	3.404	-0.026	1.064	1.067	3.883	-2.701
30	**-0.061**	1.044	1.064	5.197	0.403	-0.038	1.055	1.063	5.094	4.216
36	**-0.066**	1.034	1.058	5.522	-0.593	**-0.044**	1.042	1.053	5.343	4.748
48	**-0.067**	1.035	1.058	4.413	0.506	**-0.047**	1.038	1.049	4.202	0.947
60	**-0.074**	1.019	1.049	4.235	2.282	**-0.057**	1.020	1.037	3.985	1.356
72	**-0.069**	1.032	1.058	4.922	-1.444	**-0.053**	1.031	1.047	4.655	-4.808
84	-0.030	1.020	1.025	221.163	1.571	-0.016	1.018	1.020	204.262	1.481
96	-0.006	1.019	1.018	-26.996	1.388	0.006	1.017	1.015	-24.784	1.316
108	-0.001	1.009	1.009	12.808	4.547	0.010	1.008	1.009	12.052	4.118
120	-0.008	1.007	1.007	7.151	0.133	0.002	1.007	1.007	6.827	0.224

Forecasting yield curves

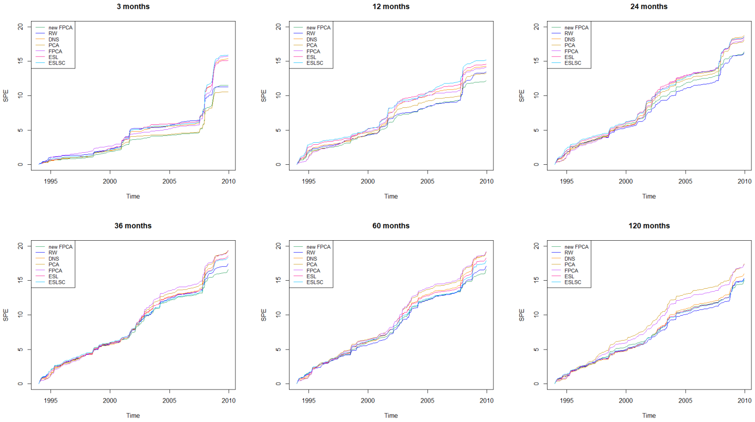

Figure: Cumulative sum of Squared Prediction Errors (SPE) of monthly yield data at maturities of 3, 12, 24, 36, 60, 120 months from January 1994 to December 2009.

3. Conclusions

Conclusions

- This method produces adequate dimension reduction for the serially dependent yield data.

Conclusions

- This method produces adequate dimension reduction for the serially dependent yield data.
- Yield curve residuals from this new model over time exhibit zero mean and less autocorrelation.

Conclusions

- This method produces adequate dimension reduction for the serially dependent yield data.
- Yield curve residuals from this new model over time exhibit zero mean and less autocorrelation.
- The forecasts of this new model have superiority:
  (i) less non-zero mean in prediction error,
  (ii) less autocorrelated prediction errors at different maturities, and
  (iii) smaller root mean squared prediction errors (RMSPE)
  over time and across term structure of interest rates at the 1-month-ahead horizon.

Thank you for your attention!