Bank i Kredyt 55(4), 2024, 459-478

BANK I KREDYT
BANK & CREDIT

Forecasting the yield curve for Poland with the PCA and machine learning

Tomasz Piotr Kostyra*

Submitted: 12 December 2023. Accepted: 27 May 2024.

Abstract

The article examines the application of the Principal Component Analysis (PCA) and machine learning method, the Long Short-Term Memory (LSTM), in the prediction of the yield curve for Poland. The PCA was applied to decompose the yield curve, forecast its components using the LSTM, and obtain the yield curve predictions upon recomposition. The results from the PCA-LSTM model were compared to predictions generated directly by the LSTM model, simple autoregression and random walk, which serves as a benchmark. Overall, LSTM predictions are the most accurate with PCA-LSTM being a close second, nonetheless PCA-LSTM is more accurate in short-term forecasting of interest rates at long maturities. Both methods outperform the benchmark, while autoregression usually underperforms. For these reasons, the PCA-LSTM as well as the LSTM can be useful in interest rate management or building trading strategies. The PCA-LSTM has the advantage that it can focus on particular components of the yield curve, such as variability of the yield curve's level or steepness.

Keywords: forecasting, LSTM, machine learning, PCA, yield curve

JEL: C53, G12, G17

* Warsaw School of Economics; e-mail: tomasz.p.kostyra@gmail.com.

1. Introduction

The Principal Component Analysis (PCA) is a dimensionality reduction technique that enables better pattern recognition. It can be applied to decompose yield curve movements into key factors, hence it provides a low-dimensional, parsimonious model. These factors can be further forecast with other techniques (i.e. autoregression, vector-autoregression). In a way, it is an alternative approach to a very popular Nelson and Siegel (1987) (NS) model and its dynamic extension proposed by Diebold and Li (2006) (DL). The NS model decomposes the yield curve into latent factors (level, slope, curvature), which in the DL model are forecast with autoregressive techniques. Both methods offer a decomposition of the yield curve into principal components, which are then separately forecast and aggregated back into yield curves.

However, contrary to the NS model's factors, the PCA factors are uncorrelated, which is particularly useful in interest rate risk management. Risk managers can focus on those components which are the source of variability and eliminate those which can be perceived as noise (Nath 2012).

According to Redfern and McLean (2014) another advantage of using the PCA in yield curve estimation is that it can be done analytically, and no numerical optimization is required. The drawback is the loss of accuracy; with a reduced number of the principal components, the chosen number of components may be insufficient.

In this paper, we examine the usefulness of the PCA framework in yield forecasting for the Polish market. First, we collect zero-coupon rates from Refinitiv for the period 2012–2021. Next, we produce four sets of forecasts. The first set contains predictions, under the assumption that interest rates will not change, so-called naive random walk, which serves as the benchmark. The second includes the predictions produced by simple autoregression. The third comprises predictions obtained by forecasting the yields with the machine learning neural network Long Short-Term Memory (LSTM) technique based directly on past yield values. The final set contains predictions which are obtained with the PCA method in a two-step procedure. In the first step, we decompose yields into principal components and verify their contributions to select those whose contributions are material. In the second, treating the selected principal components as time series, we predict their values using the LSTM method. Finally, we recompose the predicted PCA factors into yield forecasts. The backtesting results indicate that both LSTM-based methods are better than the benchmark, hence can be used in forecasting. Overall, the LSTM approach delivers more accurate forecasts, but the PCA-LSTM outperforms the LSTM in short-term forecast at long-term maturities. The PCA-LSTM approach has the advantage as it allows to focus on variability of the yield curve factors. This might be of particular use in the case of certain trading strategies or stress testing exercises. It is also less computationally expensive compared to the LSTM. The autoregression model generally underperforms the benchmark.

This study contributes to the existing literature on interest rates forecasting for Poland. In particular, it provides new insights due to combined use of the PCA and machine learning methods in this area.

The article is organized as follows. In the second section, we review the literature on the application of the PCA. The third section outlines the PCA and forecasting methodologies. The fourth section describes the data used in the research. The fifth section provides the results. The sixth section concludes.

2. Literature review

The Principal Component Analysis was first proposed by Pearson (1901). In the context of the yield curve, it was popularized by Litterman and Scheinkman (1991), who applied the principal component analysis and found that US bond returns are mainly (82%) determined by three factors, which correspond to level, steepness (slope), and curvature movements in the term structure. More recent studies which applied this methodology confirmed that the yield curve structure movements are captured by these three factors, for example 99.78% for the US, 98.74% for the UK (Chae, Choi 2022), 96.84% for Romania (Oprea 2022) or 98.99% for India (Nath 2012). In Poland, on the other hand, the explanation of 95% of the changes in the yield curve (measured in the years 2003–2011) required more than three principal components according to Lusztyn (2013).

Mallick and Mishra (2019) used the PCA to derive the PC1 (level) component which explains almost 88% of Indian yield curve variability and forecast it with the autoregressive integrated moving average – ARIMA (2,1,1). They concluded that the model delivers good forecasting results in- and out-of-sample and in the period of financial stress.

The PCA decomposition is also used to forecast interest rates. Pimentel, Risstad and Westgaard (2022) combined estimates of conditional principal component volatilities in a quantile regression (QREG) framework to infer distributional yield estimates. Their proposed PCA-QREG model offers predictions of high accuracy for most maturities while retaining simplicity in application and interpretability. They showed that the model is robust even though the sample period covered multiple major global economic events.

Blaskowitz, Herwartz and de Cadenas Santiago (2005) decomposed the FIBOR/EURIBOR swap term structure by applying the PCA to generate an interest rate forecast with the vector-autoregression model. They found that PCA techniques can be fruitfully exploited for term structure forecasting, as their approach outperformed the benchmark model.

Similarly, Dauwe and Moura (2011) forecast the monthly Euro Interest Rate Swap Curve with an autoregressive principal component model. They compared its predictive power against the Diebold and Li's dynamic Nelson-Siegel, the auto-regressive direct regression of the yield levels and the random walk model. They concluded that the proposed model significantly outperformed other models, in particular for short-run horizons.

We intend to apply a similar approach and combine the PCA method to decompose the yield curve into principal components and forecast them instead of original yields. However, contrary to the approaches outlined above, we will not use the autoregressive techniques to forecast the PCA factors, but machine learning, specifically the LSTM model. This approach of combining the PCA and LSTM techniques was successfully tested by Zhang (2022) as well as by Mi, Xu and Gao (2023) in stock time series. The study showed that the use of a reduced number of variables enabled by the PCA, improved forecasting accuracy of the LSTM model, when compared to direct forecasting of stock prices using multivariable LSTM.

The LSTM model proved to be an effective tool in interest rate forecasting. Suimon et al. (2020) constructed Japanese government bond trading strategies based on interest rate forecasts using machine learning techniques (LSTM, autoencoder) and vector-autoregression. The LSTM model delivered better results than other approaches.

3. Methodology

3.1. PCA decomposition

Principal Component Analysis. Alexander (2008) derives the principal components of yield curve time series in the following way. Suppose **X** be a T by M matrix of daily yield time series, such that:

$$\mathbf{X} = \begin{pmatrix} R_{1,\tau_1} & R_{1,\tau_2} & \cdots & R_{1,\tau_M} \\ R_{2,\tau_1} & R_{2,\tau_2} & \cdots & R_{2,\tau_M} \\ \vdots & \vdots & \ddots & \vdots \\ R_{T,\tau_1} & R_{T,\tau_2} & \cdots & R_{T,\tau_M} \end{pmatrix} \tag{1}$$

where $t = 1, 2, ..., T$ (days in a daily time series), $\tau_i = \tau_1, \tau_2, ..., \tau_M$ (the set of defined maturities in years), and **V** is its covariance matrix.

Further, let **A** be the M by M orthogonal matrix of eigenvectors of **V**. The T by M matrix **P**, where columns that correspond to principal components (as an exact linear combination) of **X**, is then given by the relation:

$$\mathbf{P} = \mathbf{X} \cdot \mathbf{A} \tag{2}$$

The aim of the PCA is to use only a few principal components to sufficiently explain the original yields. Therefore, we only take K-first columns of matrix P, let's denote it by $\mathbf{P}^*$ which has size T by K and similarly reduce the size of the matrix **A** to $\mathbf{A}^*$. We can set explainability at some level (i.e. 95%) and adjust K appropriately. Since **A** is orthogonal, $\mathbf{A}^{-1} = \mathbf{A}^T$, we obtain approximated yields:

$$\mathbf{X}_{approx} = \mathbf{P}^* \cdot \mathbf{A}^{*T} \tag{3}$$

In the yield curve modelling setting, the principal components represent how the interest rates that form the yield curve can deviate from their mean levels (Redfern, McLean 2014). This means that each component contributes to the variability of the yield curve. The first component is thought of as a parallel interest rate shift or level shifts. The second component represents tilts, i.e. steepening/ flattening of the curve. The third one, twists, i.e. changes in the curvature of the yield curve.

3.2. Forecasting

Random Walk (RW). Random walk was put forward by Pearson (1905) and popularized in finance by Malkiel (1973). The method serves as the benchmark, and in its naive interpretation supposes that 'nothing changes' over the forecasting horizon. The dynamics of a rate at maturity τ ($r_{t,\tau}$) are assumed to be:

$$r_{t,\tau} = r_{t-1,\tau} + \varepsilon_t \tag{4}$$

where $E(\varepsilon_t) = 0$.

Autoregression (AR). Autoregression was proposed by Yule (1927) and Walker (1931). This model assumes that the data generating process for the variable (in our case the interest rate at maturity τ) is a simple autoregression of order P:

$$r_{t,\tau} = \alpha + \sum_{p=1}^{P} \varrho_p r_{t-p,\tau} + \varepsilon_t \tag{5}$$

where $E(\varepsilon_t) = 0$. Given the estimates of parameters α and ϱ_p for $p = 1, 2, \ldots, P$ the forecast can be calculated recursively.

Long Short-Term Memory (LSTM). Long Short-Term Memory is a machine learning technique proposed by Hochreiter and Schmidhuber (1997). We follow the specification as presented by Sak, Senior and Beaufays (2014). The LSTM is a type of the recurrent neural network (RNN) technique. RNN has the ability to feed the network activations from a previous time step as inputs to the network to impact predictions at the current time step. These activations are stored in the internal states of the network which can hold long-term temporal information. The LSTM has units called memory blocks in the recurrent hidden layer. The memory blocks contain memory cells with self-connections storing the temporal state of the network in addition to special multiplicative units called gates to control the flow of information. Each memory block contains an input gate, output gate and a forget gate. The input gate controls the flow of input activations into the memory cell. The output gate controls the output flow of cell activations into the rest of the network. The forget gate scales the internal state of the cell before adding it as input to the cell through the self-recurrent connection of the cell, therefore adaptively forgetting information.

The LSTM model maps an input $x = (x_1, \ldots, x_{T-P})$ to an output $y = (y_{1+P}, \ldots, y_T)$ by computing the network unit activations. It uses the following equations iteratively from $t = 1$ to T:

$$i_t = \sigma\left(W_{ix} x_t + W_{im} m_{t-1} + W_{ic} c_{t-1} + b_i\right) \tag{6}$$

$$f_t = \sigma\left(W_{fx} x_t + W_{fm} m_{t-1} + W_{fc} c_{t-1} + b_f\right) \tag{7}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g\left(W_{cx} x_t + W_{cm} m_{t-1} + b_c\right) \tag{8}$$

$$o_t = \sigma\left(W_{ox} x_t + W_{om} m_{t-1} + W_{oc} c_t + b_o\right) \tag{9}$$

$$m_t = o_t \odot h\left(c_t\right), \ \ y_t = \phi\left(W_{ym} m_t + b_y\right) \tag{10}$$

where:

x_t – input vector to LSTM unit,
y_t – output vector of LSTM unit,
i_t – input/update gate's activation vector,
f_t – forget gate's activation vector,
o_t – output gate's activation vector,
b_i – input gate's bias vector,
c_t – cell input activation vector,
m_t – cell output activation vector,

$\odot$ – Hadamard product (element-wise product),
σ – sigmoid function,
g – cell input activation function (hyperbolic tangent),
h – cell output activation functions (hyperbolic tangent),
ϕ – network output activation (softmax),
W – weight matrices.

4. Data

We use the zero-coupon Polish yield curve sourced from Refinitiv from the period January 2012 – December 2021 in the forecasting exercise and we also include 2022 for backtesting purposes. Collected data are of weekly frequency for the following maturities 3M, 6M, 9M, 1Y, 2Y, 3Y, 4Y, 5Y, 7Y, 10Y (Figure 1). Table 1 provides statistics of weekly yield series for the period January 2012 – December 2021.

We selected this timeframe as the intention is to focus on more recent data, but to also include the full economic cycle. Since joining the European Union in 2004, Poland experienced almost continuous GDP growth. There were, however, years where GDP growth significantly slowed down, or even turned negative. The lowest (-2%) growth was recorded in 2020, which was a consequence of the COVID-19 pandemic. The second lowest (0.9%) in 2013, was caused by the stagnation in the euro area that was accompanied by a GDP growth rate decline in Germany, Poland's main trading partner. There were also three GDP growth peaks in this period, namely 2007 (7.1%) due to joining the EU, 2021 (6.9%) due to the fiscal and monetary measures implemented to counter the negative effects of COVID on the economy, and 2018 (5.9%) – a peak of continuous economic expansion driven by the Germany economy that started after 2012–2013 slowdown. The selection of the 2012–2021 period captures economic activity from peak/downturn to peak/downturn.

The forecasting exercise is limited to four maturities, i.e. 3-month, 1-year, 5-year and 10-year. For each day we transform the series of yields R_τ for ten fixed maturities τ into continuously compounded yields r_τ with the following formula:

$$r_\tau = \ln\left(1 + \frac{R_\tau}{100}\right) \cdot 100 \tag{11}$$

In order to establish the number of lags for forecasting purposes, we compute the Partial Auto--Correlation Function (PACF) for selected maturities. Based on the outcomes, we use the following number of lags: 1st and 2nd lag for 3-month, 1st and 5th lag for 1-year, 1st lag for both 5-year and 10-year maturity.

5. Results

5.1. PCA decomposition

Table 2 and Figure 2 present the results of the Principal Component Analysis. We can see that practically 100% of the yield curve dynamic is captured by the first three factors, which are also referred to as level (PC1), slope (PC2) and curvature (PC3). Already the first principal component captures 96% of variability of yield curve movements, the yellow line on Figure 2 denotes 95% level of explainability. We conclude that we are able to capture nearly 100% of yield curve variability by predicting the evolution of the first three PCA components.

It would be acceptable to limit the forecasting exercise to the first two components as their explanatory power meets two commonly used thresholds of explainability, i.e. 95% and 99%. However, looking at PCA components' time series plots in Figure 3, we note that PC3's volatility increases from 2019. This may contribute to greater forecast accuracy, therefore we have decided to include it in forecasting.

Table 3 provides descriptive statistics of the time series for PCA components. After verification of their stationarity, we conclude that these are sufficiently stable across time, hence no additional transformations are required.

Figure 3 shows the time series of the PCA components. We compute the PACF for them to establish the number of lags for forecasting purposes. Based on the outcomes, we will use the 1st lag for PC1, and 1st and 2nd lag both for PC2 and PC3.

5.2. Forecasting

In the forecasting exercise, we use the time series of the three PCA components obtained above. We produce forecasts for weekly horizons starting from week 1 up to week 12 for each week in the period January 2015 – December 2021. The estimation is done on data from the period January 2012 – December 2021. Models are estimated based on a 3-year (156-week) rolling window applied to input data. To illustrate, the first set of 12 forecasts produced in December 2014 for the period January 2015 – March 2015 is generated using the models estimated based on observations from January 2012 to December 2014 (156 weeks). To compute 12 weekly forecasts, first we produce a one-step ahead forecast (1st week forecast) and use this prediction to re-estimate the model in order to produce another one--step ahead forecast (2nd week forecast). We continue until the end of the forecasting horizon (12th week) is reached. Once the first set of 12 forecasts is completed, the estimation window rolls by one week and the procedure is repeated. This continues for each week for the period January 2015 – December 2021. When completed, we obtain 366 sets of 12 weekly forecasts.

The results were backtested against actual values from the period January 2015 – March 2022. Figure 4 shows the 12-week projections for AR, LSTM and PCA-LSTM models. Table 4 presents the root mean squared forecast errors (RMSFE). They are complemented with the outcomes of Diebold and Mariano (1995) test, which verifies the accuracy of these forecasts against the forecasts generated by the RW model. We compare the RMSFE results of each model to the benchmark (RW model) by scaling their error values with the error value of the benchmark.

We observe that the AR model typically underperforms the benchmark, and the results are statistically significant.

The LSTM consistently beats the RW. It generates predictions with lower RMSFE for all forecasting horizons and maturities. The results are statistically significant. We note that the RMSFE increases for longer forecast horizons, hence the advantage of the LSTM model over the benchmark diminishes, yet the differences remain statistically significant.

The PCA-LSTM model that uses a two-step forecasting procedure (PCA decomposition and prediction of the PCA factors with LSTM) overall performs better than the benchmark as indicated by the RMSFE. However, the RMSFE increases for longer forecast horizons, and the results become not statistically significant for the 12W horizon, unlike the results of the LSTM model.

We verified the PCA-LSTM and LSTM models in more detail. Table 5 presents the results of the direct comparison between the LSTM and PCA-LSTM models. Statistically significant differences occur for short horizons. The LSTM model generates more accurate results for 1W horizon, for 3-month and 1-year maturities, while the PCA-LSTM produces better results for 1W and 4W horizons, for 5-year maturity. For 10-year maturity the accuracy level is similar for both models. Overall, the RMSFE error is lower for the LSTM. Therefore, we conclude that in general, the LSTM model delivers slightly more accurate forecasts than the PCA-LSTM.

6. Discussion and conclusions

After more than a decade of near-zero interest rate policies, central banks started to adjust nominal interest rates in response to the external events and their expected impact on the economies such as COVID-19 pandemic or the war in Ukraine. In Poland, we could observe dramatic changes in nominal interest rates set by Narodowy Bank Polski (NBP) and inflation rates. Due to COVID-19, the Polish economy quickly transitioned from the long period of low, or even negative inflation, to double digit inflation rates by March 2022. In response, NBP began a series of nominal interest rates hikes in 2021, which continued in 2022, reaching levels last time observed 20 years ago. These events increased the volatility of Treasury yields in Poland, and as such propelled interest in yield forecasting. Consequently, we are re-examining the yield forecasting approaches, now incorporating machine learning.

The aim of this study is to investigate the application of the PCA and LSTM methods in yield time series forecasting for the Polish economy. In particular, we want to verify whether the combination of the PCA and LSTM methods improves forecasting accuracy compared to direct use of the LSTM. Interest rates for the Polish Treasury bonds are forecast at various maturities. Both methods perform better than the naive RW model, which assumes that interest rates will stay at prior levels, and simple autoregression. When comparing the PCA-LSTM with the LSTM, we note that the LSTM model performed slightly better. However, the application of the PCA-LSTM has other advantages. Since the first three PCA components represent the changes in the level, slope and curvature of the yield curve, their analysis provides additional insights into the variability of yield curve structure. For certain applications it can be sufficient to focus on forecasting of a specific component of the yield curve, for example, trading strategies can be constructed based on expectations of yield curve steepening/flattening. The PCA-LSTM model also has the operational advantage as LSTM simulations are computationally expensive. It requires only LSTM-simulations of three PCA factors, which is sufficient to construct forecasts for all ten maturities, while with the LSTM model, ten simulations are required (one for each maturity). Therefore, PCA-LSTM forecasts can be generated faster.

This analysis can be helpful for money managers in their prediction and analysis of interest rate evolution. It can also be employed in managing the interest rate risk in banks as well as in a regulatory Internal Capital Adequacy Assessment Process. The study can be extended by incorporating macro or financial variables to verify if their inclusion can improve the accuracy of forecasts.

References

Alexander C. (2008), *Market Risk Analysis. Practical Financial Econometrics,* John Wiley and Sons.

Blaskowitz O.J., Herwartz H., de Cadenas Santiago G. (2005), *Modeling the FIBOR/EURIBOR swap term structure: an empirical approach*, Technical Report, SFB Discussion Paper, 649.

Chae S.C., Choi S.Y. (2022), Analysis of the term structure of major currencies using principal component analysis and autoencoders, *Axioms*, 11(3), 135.

Dauwe A., Moura M.L., *Forecasting the term structure of the euro market using principal component analysis,* Insper Working Paper, 233/2011, Insper – Institute of Education and Research.

Diebold F.X., Li C. (2006), Forecasting the term structure of government bond yields, *Journal of Econometrics*, 130(2), 337–364.

Diebold F.X., Mariano R.S. (1995), Comparing predictive accuracy, *Journal of Business & Economic Statistics*, 13(3), 253–263.

Hochreiter S., Schmidhuber J. (1997), Long short-term memory, *Neural Computation*, 9(8), 1735–1780.

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

Lusztyn M. (2013), *Zastosowanie metody składowych głównych w analizie ryzyka stopy procentowej na przykładzie krzywej dochodowości w Polsce w latach 2003–2011*, Studia i Prace Kolegium Kolegium Zarządzania i Finansów SGH, 47.

Malkiel B.G. (1973), *A Random Walk down Wall Street*, W.W. Norton & Co.

Mallick A.K., Mishra A.K. (2019), Interest rates forecasting and stress testing in India: a PCA-ARIMA approach, *Palgrave Communications*, 5(1), 1–17.

Mi Y., Xu D., Gao T. (2023), Application of PCA-LSTM algorithm for financial market stock return prediction and optimization model, *International Journal for Simulation and Multidisciplinary Design Optimization*, 14, 8.

Nath G.C. (2012), *Estimating term structure changes using principal component analysis in Indian sovereign bond market,* SSRN 2075635.

Nelson C.R., Siegel A.F. (1987), Parsimonious modeling of yield curves, *The Journal of Business*, 60(4), 473–489.

Oprea A. (2022), The use of Principal Component Analysis (PCA) in building yield curve scenarios and identifying relative-value trading opportunities on the Romanian government bond market, *Journal of Risk and Financial Management*, 15(6), 247.

Pearson K. (1901), On lines and planes of closest fit to systems of points in space, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2, 559–579.

Pearson K. (1905), The problem of the random walk, *Nature*, 72(1865), 294–294.

Pimentel R., Risstad M., Westgaard S. (2022), Predicting interest rate distributions using PCA & quantile regression, *Digital Finance*, 4(4), 291–311.

Redfern D., McLean D. (2014), *Principal component analysis for yield curve modelling reproduction of out--of-sample yield curves*, White Paper, Moody's Analytics.

Sak H., Senior A. W., Beaufays F. (2014), *Long short-term memory recurrent neural network architectures for large scale acoustic modeling*, https://research.google/pubs/pub43905/.

Suimon Y., Sakaji H., Izumi K., Matsushima H. (2020), Autoencoder-based three-factor model for the yield curve of Japanese government bonds and a trading strategy, *Journal of Risk and Financial Management*, 13(4), 82.

Walker G.T. (1931), On periodicity in series of related terms, *Proceedings of the Royal Society of London*, Series A, Containing Papers of a Mathematical and Physical Character, 131(818), 518–532.

Yule G.U. (1927), VII. On a method of investigating periodicities disturbed series, with special reference to Wolfer's sunspot numbers, *Philosophical Transactions of the Royal Society of London,* Series A, Containing Papers of a Mathematical or Physical Character, 226 (636–646), 267–298.

Zhang Z. (2022), Research on stock price prediction based on PCA-LSTM model, *Academic Journal of Business & Management*, 4(3), 42–47.

Appendix

Table 1

Zero-coupon yields – descriptive statistics

Maturity	Mean	St. dev.	Min.	Max.	ACF	ADF
3M	1.55	1.21	-0.59	4.94	0.98	-1.98**
6M	1.60	1.18	-0.29	4.73	0.99	-2.12**
9M	1.66	1.16	-0.21	4.71	0.99	-2.11**
1Y	1.71	1.14	-0.14	4.69	0.99	-1.91*
2Y	1.91	1.11	-0.03	4.70	0.99	-1.51
3Y	2.12	1.08	0.05	4.90	0.99	-1.44
4Y	2.32	1.06	0.19	5.10	0.99	-1.41
5Y	2.50	1.05	0.37	5.28	0.99	-1.41
7Y	2.80	1.04	0.74	5.58	0.99	-1.48
10Y	3.10	1.04	1.18	5.87	0.95	-1.57

Notes:
ACF and ADF refer to the values of the autocorrelation coefficient and the Augmented Dickey Fuller.
Asterisks ***, **, and * denote the rejection of the null that series is non-stationary at the 1%, 5% and 10% significance level, respectively.
Weekly time series from the period January 2012 – December 2021.

Source: own calculations.

Table 2

Principal Component Analysis

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Variance explained	0.964	0.030	0.005	0.001	0.000	0.000	0.000	0.000	0.000	0.000
Cumulative variance explained	0.964	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Notes:
ACF and ADF refer to the values of the autocorrelation coefficient and the Augmented Dickey Fuller.
Asterisks ***, **, and * denote the rejection of the null that series is non-stationary at a 1%, 5% and 10% significance level, respectively.
Weekly time series from the period January 2012 – December 2021.

Source: own calculations.

Table 3

PCA factors – descriptive statistics

PC	Mean	St. dev.	Min.	Max.	ACF	ADF
PC1 (L)	1.063	3.435	-7.651	7.146	0.99	-2.41**
PC2 (S)	-0.004	0.608	-1.446	1.628	0.93	-3.02***
PC3 (C)	-0.050	0.290	-0.904	1.249	0.84	-2.44**

Notes:
ACF and ADF refer to the values of the autocorrelation coefficient and the Augmented Dickey Fuller.
Asterisks ***, **, and * denote the rejection of the null that series is non-stationary at a 1%, 5% and 10% significance level, respectively.
Weekly time series from the period January 2012 – December 2021.

Source: own calculations.

Table 4

RMSFE

Model	1W	4W	8W	12W
3-month maturity				
RW	0.13450	0.22488	0.31328	0.41340
AR	0.98715	1.06428**	1.10705**	1.12743**
LSTM	0.74189***	0.93315***	0.96703***	0.98400**
PCA-LSTM	0.88439***	0.93236***	0.98410	1.00715
1-year maturity				
RW	0.07611	0.19961	0.33864	0.48907
AR	0.95486	1.00656	1.06115	1.12507
LSTM	0.91100***	0.97943***	0.98872**	0.99042*
PCA-LSTM	1.06792***	0.98570	0.98108	0.98158
5-year maturity				
RW	0.10193	0.21770	0.34719	0.48473
AR	1.01104*	1.02939*	1.04041*	1.03395
LSTM	0.97989***	0.98816***	0.86191**	0.91657*
PCA-LSTM	0.93053***	0.96322***	0.97425**	0.98445*
10-year maturity				
RW	0.11235	0.22770	0.33248	0.44113
AR	1.01560**	1.05106***	1.07948**	1.09749*
LSTM	0.97968***	0.98338***	0.98617***	0.98965***
PCA-LSTM	0.95029**	0.99188	0.99604	0.99936

Notes:
RMSFE are computed for the following horizons: 1W, 4W, 8W, 12W across four maturities 3M, 1Y, 5Y, 10Y. RMSFE for forecasting models are scaled with the RMSFE of the benchmark (RW), hence the value of 1 must be taken for comparison with other models. To arrive at the actual value of the RMFSE of the forecasting model, multiply the RMSFE of the benchmark with the scalar of the forecasting model. The Diebold-Mariano test was conducted to verify forecast accuracy against the benchmark (RW), with null hypothesis that the forecast accuracy of models is the same as the benchmark's, against the alternative that the forecast accuracy of models is different than the benchmark's.
Asterisks ***, **, and * denote the rejection of the null at a 1%, 5% and 10% significance level, respectively.

Source: own calculations.

Table 5

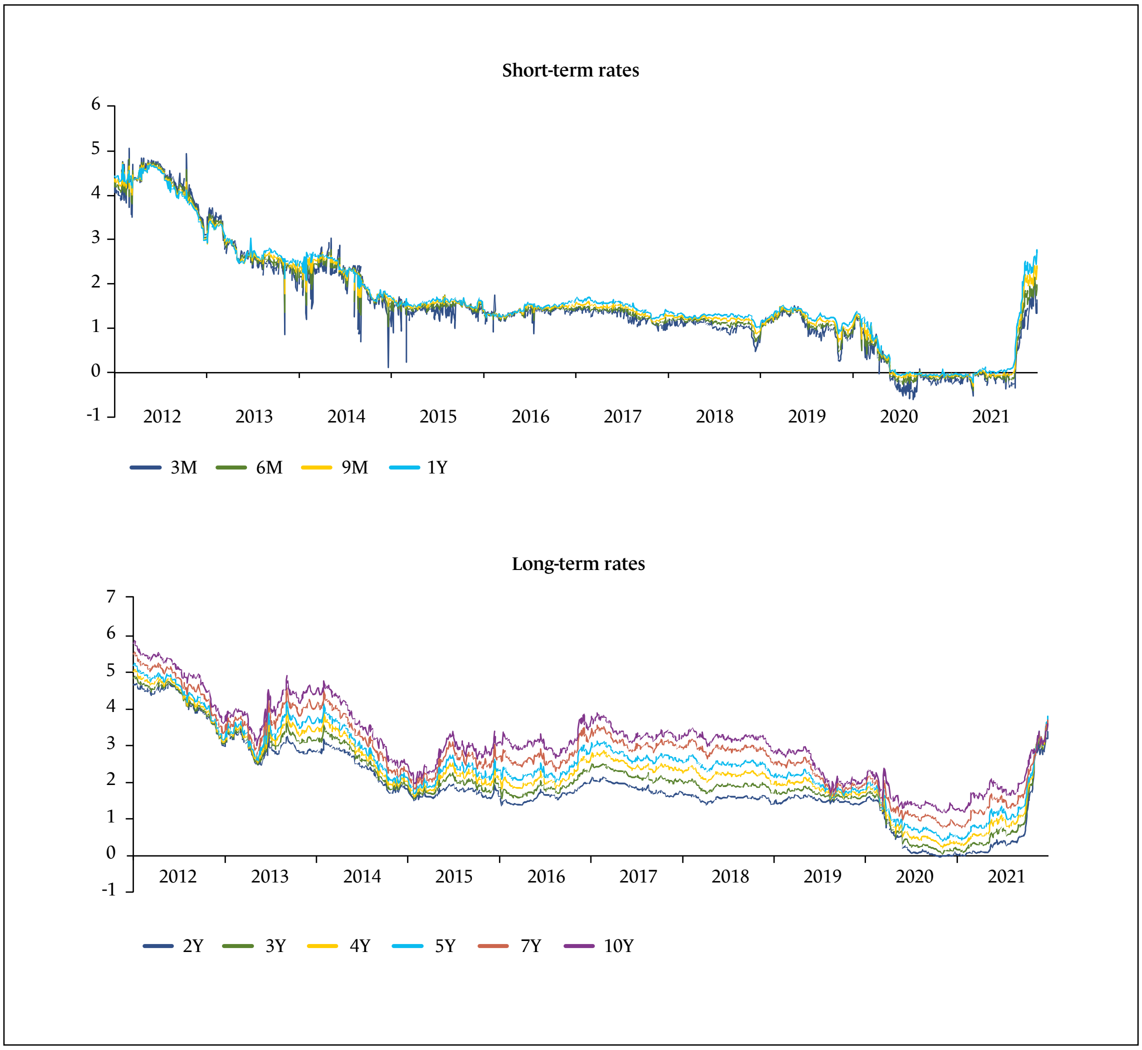
PCA-LSTM vs. LSTM comparison

Maturity	1W	4W	8W	12W
3 month	LSTM***	PCA-LSTM	LSTM	LSTM
1 year	LSTM***	LSTM	PCA-LSTM	PCA-LSTM
5 year	PCA-LSTM***	PCA-LSTM**	LSTM*	LSTM
10 year	PCA-LSTM	LSTM	LSTM	LSTM

Notes:
RMSFE are compared for PCA-LSTM and LSTM models. Model with lower RMSFE is shown with the result of the Diebold-Mariano test of forecast accuracy. The null hypothesis is that the forecast accuracy of both models is the same, against the alternative that the forecast accuracy is different.
Asterisks ***, **, and * denote the rejection of the null at 1%, 5% and 10% significance level, respectively.

Source: own calculations.

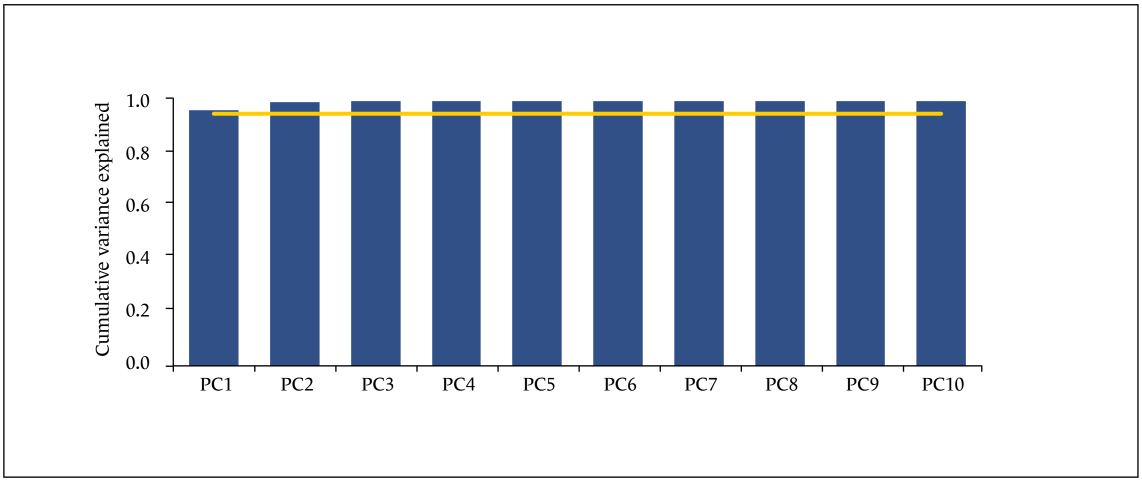
Figure 1
Zero-coupon rates



Note: daily time series from the period January 2012 – December 2021.

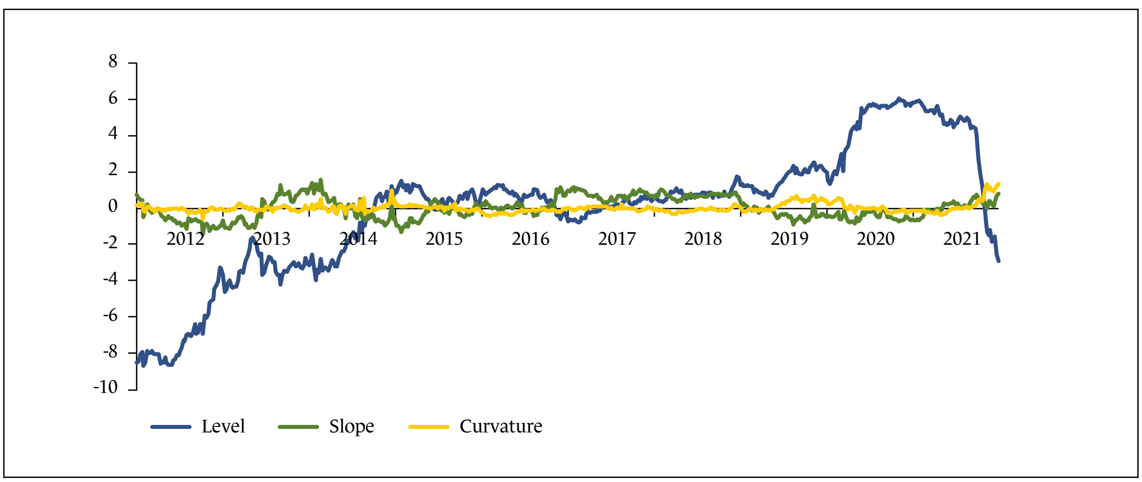
Source: own calculations based on Refinitiv's data.

Figure 2
Principal Component Analysis



Note: based on the data from the period January 2012 – December 2021.

Figure 3
Principal Component Time Series



Note: weekly time series from the period January 2012 – December 2021.

Source: own calculations.

Figure 4
Forecasts

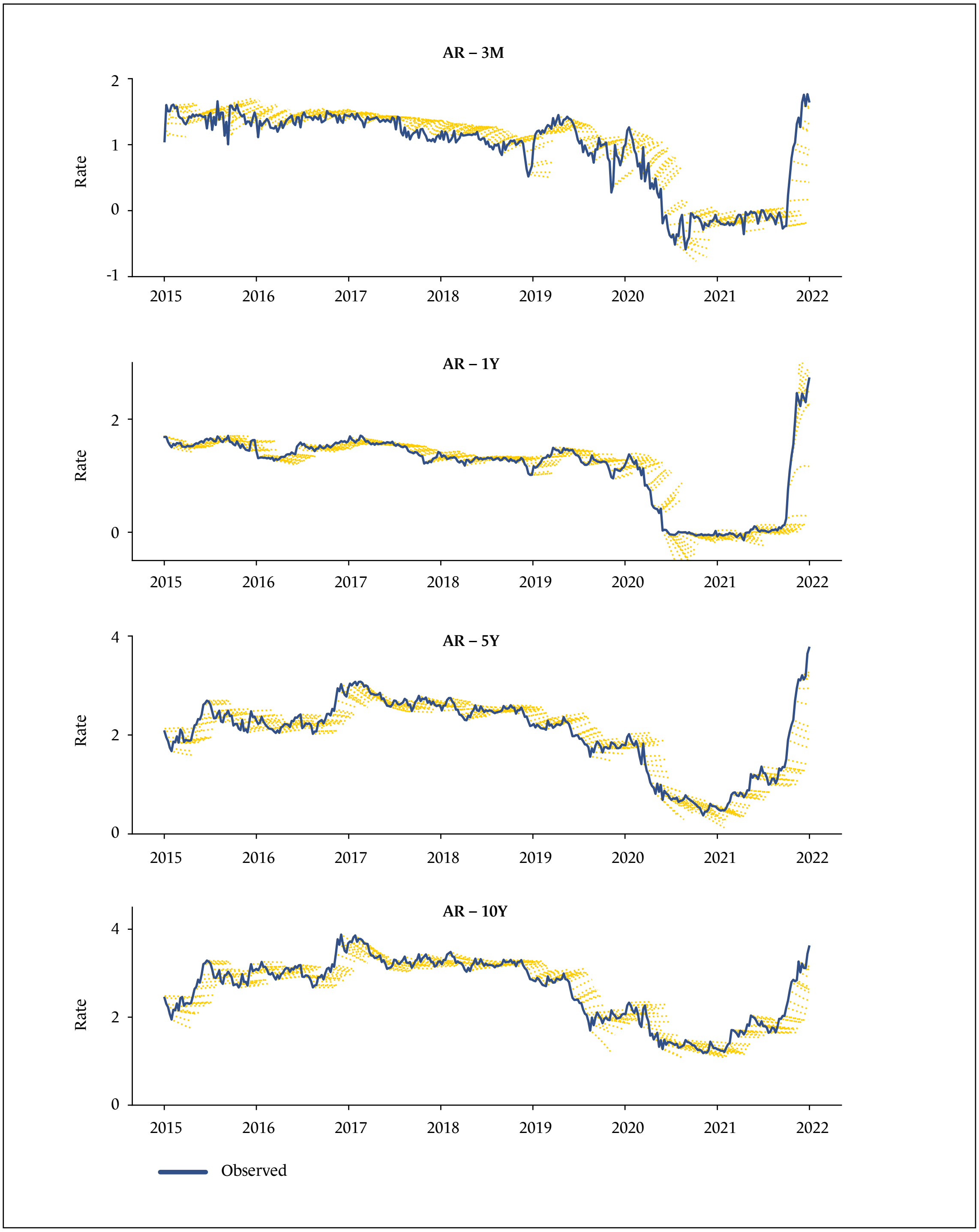


Figure 4, cont'd

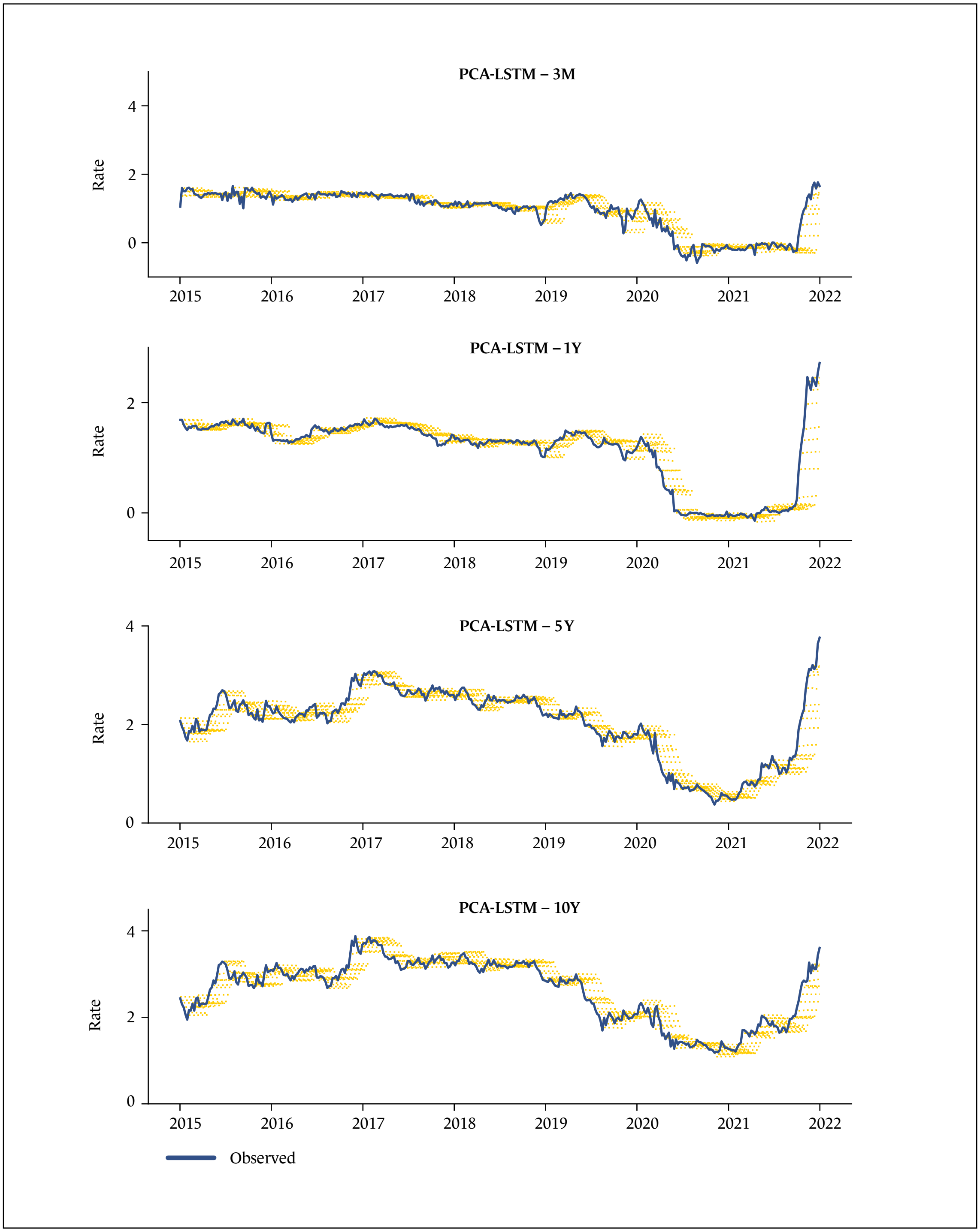
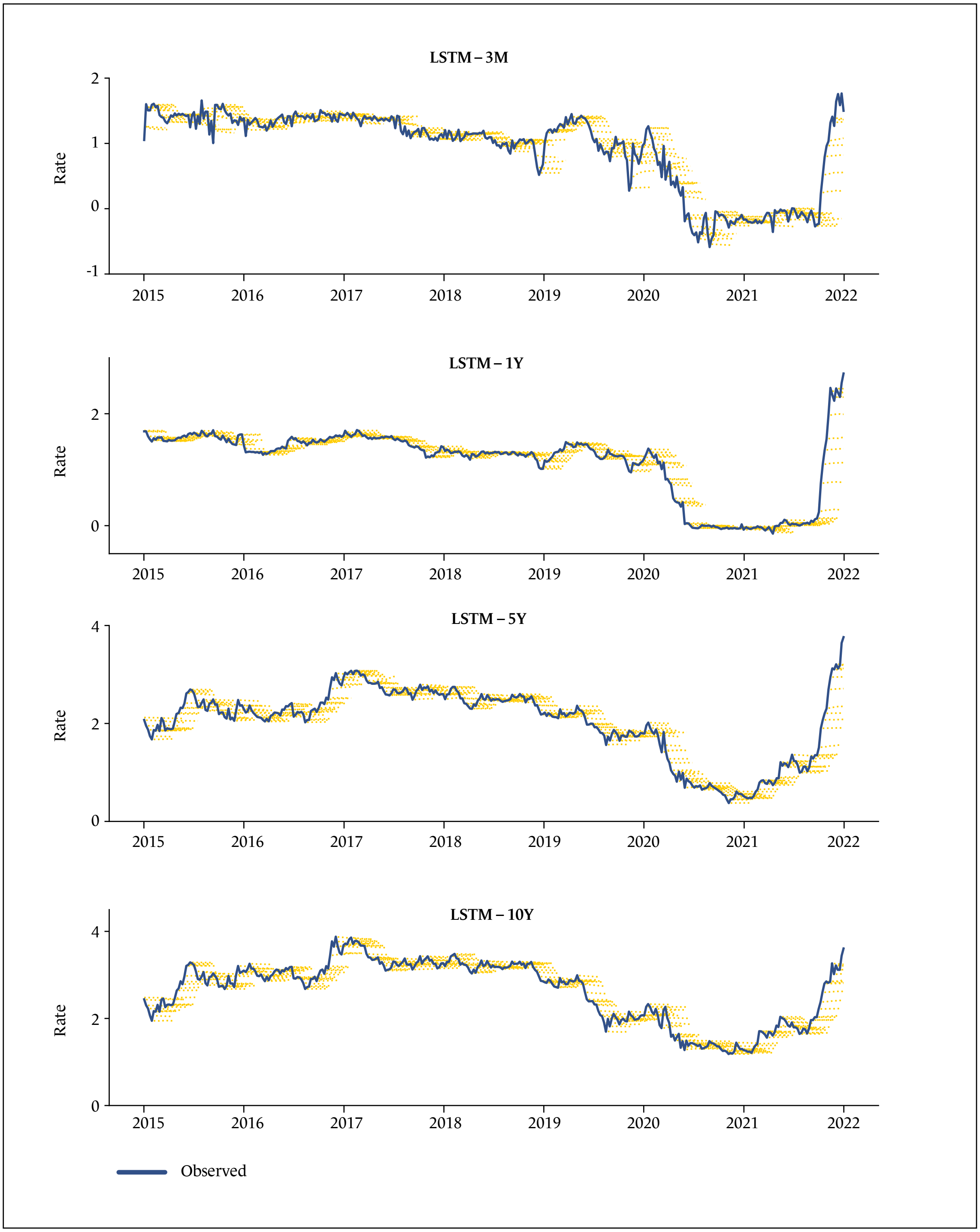


Figure 4, cont'd



Notes:
Projections based on the data from the period January 2012 – December 2021. Data of weekly frequency.

Source: own calculations.

Prognozowanie krzywej dochodowości dla Polski za pomocą PCA oraz uczenia maszynowego

Streszczenie

Metoda PCA służy do uzyskania czynników krzywej dochodowości (poziom, nachylenie, krzywizna) za pomocą algebraicznej dekompozycji krzywej. Tego rodzaju zastosowanie PCA zostało zapoczątkowane przez Littermana i Scheinkmana (1991), którzy dowiedli, że w USA zwroty z obligacji są w 82% zdeterminowane przez te trzy czynniki. W Polsce czynniki te wyjaśniają 95% zmienności krzywej dochodowości (Lusztyn 2013).

Metoda ta jest często łączona z metodami autoregresyjnymi w celu sporządzenia prognoz czynników, a następnie ich rekompozycji, by uzyskać prognozę krzywej dochodowości. Dauwe i Moura (2011) zastosowali takie podejście do prognozowania krzywej swapowej dla euro. Wykazali, że prognozy uzyskane za pomocą modelu łączącego PCA i metody autoregresyjne są lepsze od prognoz generowanych przez dynamiczny model Nelsona-Siegla, bezpośrednich modeli autoregresyjnych oraz modelu błądzenia losowego.

Zastosowanie PCA umożliwia też skupienie się na analizie i prognozach najbardziej znaczącego czynnika. Takie podejście zastosowali Mallick i Mishra (2019). Mianowicie po dokonaniu rozkładu krzywej dochodowości opracowali prognozy za pomocą modelu ARIMA tylko dla pierwszego czynnika, tj. poziomu stóp procentowych, gdyż odpowiada za 88% ich zmienności. Wykazali, że wyniki predykcji są na zadowalającym poziomie.

Technikę PCA zaczęto łączyć z metodami uczenia maszynowego, których popularność w prognozach finansowych stale rośnie. Zhang (2022) zastosował PCA do redukcji liczby zmiennych objaśniających ceny akcji, których prognozy wykonano z użyciem modelu LSTM. Prognozy generowane przez model ze zredukowaną liczbą zmiennych za pomocą metody PCA okazały się dokładniejsze od prognoz modelu wykorzystującego wszystkie zmienne.

W przypadku Polski zidentyfikowaliśmy lukę badawczą w zakresie stosowania metody PCA oraz metod uczenia maszynowego w predykcji stóp procentowych.

W opracowaniu odpowiadamy na pytanie, czy prognozowanie stóp procentowych w Polsce przez przewidywanie ich czynników składowych jest bardziej efektywne od bezpośredniego prognozowania stóp procentowych. Do prognozowania zastosowaliśmy model uczenia maszynowego LSTM, który jest oparty na sieciach neuronowych.

W artykule weryfikujemy tezę, czy model PCA-LSTM pozwala na efektywne prognozowanie stóp procentowych.

Za pomocą techniki PCA rozkładamy krzywą dochodowości na czynniki składowe, a następnie prognozujemy czynniki za pomocą modelu LSTM dla horyzontów do 12. tygodnia. Prognozy czynników zostają następnie złożone w prognozy stóp procentowych. Do celów porównawczych dokonujemy prognoz stóp procentowych dla określonych terminów zapadalności za pomocą trzech dodatkowych modeli (modelu błądzenia losowego, prostej autoregresji oraz modelu LSTM). W modelach porównawczych jako dane wsadowe wykorzystujemy tylko szeregi czasowe stóp procentowych. Wyniki prognoz porównaliśmy z faktycznymi wartościami stóp procentowych z danych okresów. Na tej podstawie za pomocą testu Diebolda-Mariano konkludujemy, który model generuje najlepsze prognozy.

Modele PCA-LSTM oraz LSTM pozwalają na efektywne prognozowanie stóp procentowych, jednak to model LSTM generuje nieco trafniejsze prognozy. Z kolei prognozy dokonane za pomocą modelu autoregresyjnego są mniej dokładne w porównaniu z benchmarkiem. Model PCA-LSTM umożliwia ponadto analizę czynników krzywej dochodowości (poziom, nachylenie, krzywizna) oraz szybsze uzyskanie prognoz. W związku z powyższym modele PCA – LSTM i LSTM – mogą zostać wykorzystane do budowy strategii tradingowych, zarządzania ryzykiem stóp procentowych oraz w regulacyjnych ćwiczeniach szacowania poziomu kapitału wewnętrznego (ICAAP).

Metoda PCA pozwala się skupić na głównych czynnikach mających wpływ na dane zjawisko, w tym przypadku krzywą dochodowości. Prowadzi to do odrzucenia pewnych czynników jako nieistotnych; tutaj jest to czynnik PC4. Takie podejście powoduje jednak utratę części informacji, co może negatywnie wpływać na trafność prognoz.

Słowa kluczowe: krzywa dochodowości, LSTM, PCA, prognozowanie, uczenie maszynowe