Forecasting the Yield Curve for Poland

Tomasz Piotr Kostyra [*♣] and **Michał Rubaszek** [†♣]

♣SGH Warsaw School of Economics

ABSTRACT: This paper evaluates the accuracy of forecasts for Polish interest rates of various maturities. We apply the traditional autoregressive Diebold-Li framework as well as its extension, in which the dynamics of latent factors are explained with machine learning techniques. Our findings are fourfold. Firstly, they show that all methods have failed to predict the declining trend of interest rates. Secondly, they suggest that the dynamic affine models have not been able to systematically outperform standard univariate time series models. Thirdly, they indicate that the relative performance of the analyzed models has depended on yield maturity and forecast horizon. Finally, they demonstrate that, in comparison to the traditional time series models, machine learning techniques have not systematically improved the accuracy of forecasts.

1   Introduction

Developing a model that would provide accurate forecasts for financial time series is the Holy Grail for financial markets participants. The reason behind this is straightforward: such a model would enable the development of profitable investment strategies. In the same vein, the development of financial markets, innovations in financial products and

---

*Email: tomasz.p.kostyra@gmail.com
†Email: michal.rubaszek@sgh.waw.pl

changes in the regulatory requirements continue to stimulate interest in successful forecasting techniques.

In this paper, we examine the effectiveness of some of the most popular forecasting methods and offer a new extension of the Diebold and Li (2006, DL) framework to explain the dynamics and forecasting money market rates in Poland. More precisely, we evaluate the accuracy of Polish interest rate curve forecasts generated from two sets of models. The first one contains univariate models which assume that the current levels of interest rates depend only on their lagged values. In this group we consider a random walk (RW), autoregression (AR) and a random forest (RF) i.e. supervised machine learning model based on decision trees (Breiman, 2001). The second set comprises dynamic variants of the affine DL model, where the three factors (level, curvature, and slope) are extracted from the interest rate time series and then approximated using the autoregression, vector-autoregression (VAR) and the random forest. We also consider interest rate forecasts which are calculated using the yield curve interpolated with the Nelson and Siegel (1987) (NS) method assuming no arbitrage condition.

In this respect our contribution is twofold. First of all, we focus on an emerging economy, hence our results might be of interest to academics, as well as market participants, who focus on countries under economic transformation. Secondly, we provide new evidence on the usefulness of machine learning techniques in modeling financial markets. To the best of our knowledge, we are the first to apply this method as the extension to Diebold-Li framework in the interest rate forecasting context for Polish market.

Our main findings are fourfold. Firstly, we show that all methods have failed to predict the declining trend of interest rates. Secondly, we indicate that the dynamic affine models have not been able to systematically outperform standard univariate time series models. Thirdly, we suggest that the performance of models depends on the maturity and the forecast horizon. Finally, we demonstrate that, in comparison to traditional time series models, applied machine learning techniques have not systematically improved the accuracy of forecasts.

This article is organized into five sections. Section 2 presents forecasting methodologies. Section 3 describes the data. Section 4 contains the results of the forecasting competition. The final section is dedicated to the discussion and conclusions.

2   Literature overview

Our work builds on a number of studies in which yields at different maturities are assumed to be a linear function of few latent factors, which are called affine models. Our main focus is on the interpretable affine models proposed by Nelson and Siegel (1987, NS) and Diebold and Li (2006), which have gained a lot of recognition in yield curve modeling.

The properties of the NS model were discussed by Geyer and Mader (1999) while describing the interest rate structure in Western Europe, USA and Japan. The authors

compared its performance to its extension proposed by Svensson (1994), which allows for more than one local extremum along the maturity profile of a yield curve. Overall, they showed that the parsimonious NS approach outperformed the Svensson's one in terms of parameters stability in time, as it was less sensitive to outliers. Hladíková and Radová (2012) pointed out good proprieties of the NS model in describing the yield curve for the Czech Republic, a country that underwent a similar transformation to Poland, and which was characterized by an illiquid market. In the same vein, Zoricic and Badurina (2013) indicated that the NS model could be successfully applied to describe the yield curve in Croatia. Finally, Marciniak (2006) conducted a comparative analysis to show that the Svensson extension of the NS model performed relatively well in comparison to the B-spline–Variables Roughness Penalty model in explaining the interest rate structure in Poland.

The dynamic version of the NS model, proposed by Diebold and Li (2006), has been widely applied in forecasting yields at different maturities. For instance, Gurkaynak and Wright (2012) used this framework to explore the relationship between the interest rates and macroeconomic variables in the forecasting context. In a similar vein, Rubaszek (2016) studied dynamic affine models with autoregressive, vector-autoregressive and Bayesian autoregressive processes, also exploring the predictive content of external, macroeconomic regressors. The author showed that the dynamic NS model performed better when factors were described by the univariate AR as opposed to VAR processes. Moreover, he demonstrated that allowing for the interaction of latent factors with macroeconomic variables did not improve the accuracy of forecasts. Yu and Zivot (2010), on the other hand, compared the two-step estimation procedure proposed by DL to the one-step Kalman filtering estimation method proposed by Diebold, Rudebusch, and Aruoba (2006). They concluded that two-step approach delivered better forecasts for default-free bonds, while one–step approach did so for defaultable bonds. Finally, Christensen and Rudebusch (2015) focused on the arbitrage free version of the DL model and showed that it was possible to deliver relatively accurate forecasts.

The studies presented above were conducted either for developed economies with a focus on American economy with its large, diverse, and liquid bond market, or much smaller, emerging economies of central Europe. Therefore their results might not be directly applicable to the Polish market. Studies concerning the Polish economy are relatively scarce. Hence, in our study, we would like to address this gap and explore the usefulness of the Diebold–Li framework for Poland, a country, which underwent a transformation from a command economy, with no capital markets in 1989, into to a free market economy and which, in 2018, was classified as the developed market in indices run by FTSE Russell. Moreover, we would like to examine the usefulness of machine learning techniques within the affine model setup. Reflecting on the fact that decision tree models are widely used by financial market participants (Jung et al., 2019) and considering the results of Martin, Póczos, and Hollifield (2018), who compared a random forest approach

with a support vector regression in yield forecasting, we have chosen a random forest technique to forecast the latent factors in our study.

3 Forecasting Methodologies

We consider the following competitors in the forecasting horse race.

3.1 Univariate models

Random Walk (RW). This method serves as the benchmark and assumes that 'nothing changes' over the forecast horizon, so that the dynamics of a rate at maturity m ($R_{m,t}$) are assumed to be:

$$R_{m,t} = R_{m,t-1} + \epsilon_t. \tag{1}$$

This model is commonly used as a benchmark in the forecasting literature. It can be expected to be especially tough for Poland since the reference interest rates set by the Polish central bank haven't changed since March 2015.

Autoregression (AR). This model assumes that the data generating process for the variable (in our case the interest rate at maturity m, $R_{m,t}$) is a simple autoregression of order P:

$$R_{m,t} = \alpha + \sum_{p=1}^{P} \rho_p R_{m,t-p} + \epsilon_t. \tag{2}$$

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

Random Forest (RF). Let x denote the set of our input variables and y the response variable. The procedure to forecast from RF is as follows (Breiman, 2001). We randomly draw a sample from the training data (y, x) so that each data point has an equal probability of getting selected. All the samples have the same size as the original training set. These samples are called Bootstrap samples and they are taken with replacement from the training data set. For each draw, which is indexed by $k = 1, 2, ..., K$, we grow a random forest of decision trees depending on a random vector θ_k, which provides a predictor $h(x, \theta_k)$. A prediction from the random forest is subsequently calculated as:

$$\hat{y} = \frac{1}{k} \sum_{k=1}^{K} h(\mathbf{x}, \theta_k) \tag{3}$$

We use this approach to forecast interest rates by building separately RF for each maturity m and horizon h, so that $y = R_{m,t+h}$. We use a model with own P lags so that $x = [R_{m,t}\ R_{m,t-1}\ \ldots\ R_{m,t-P+1}]$.

3.2 Forecasts based on expectations

Nelson – Siegel (NS).   In the expectation hypothesis method forecasts are based on the assumption that arbitrage is not possible on the market. In this case a forward rate contract beginning in period $t + h$ for interest rate at maturity m should stand at:

$$F_{t,h,m} = \frac{(m+h)R_{t,m+h} - hR_{t,h}}{m} \tag{4}$$

We then construct the forecast as:

$$R^f_{t+h,m} = F_{t,h,m} \tag{5}$$

In this method we need the values of interest rates at different maturities. We derive them by interpolating the yield curve with the Nelson and Siegel (1987) model of the form:

$$R_m = L + S\left(\frac{1 - e^{-m\lambda}}{m\lambda}\right) + C\left(\frac{1 - e^{-m\lambda}}{m\lambda} - e^{-m\lambda}\right), \tag{6}$$

The parameters L, S, C are estimated using observations for the spot rate at different maturities, whereas for λ, we fix its value at 0.0609, following Diebold and Li (2006).

3.3 Diebold–Li framework (DL)

We consider few dynamic affine yield curve models, which are based on the seminal work of Diebold and Li (2006). It is a two-step procedure. In the first step we estimate the Nelson-Siegel factors L, S and C for each moment in time:

$$R_{m,t} = L_t + S_t\left(\frac{1 - e^{-m\lambda}}{m\lambda}\right) + C_t\left(\frac{1 - e^{-m\lambda}}{m\lambda} - e^{-m\lambda}\right). \tag{7}$$

In the second step we forecast factors and use the results to forecast the entire yield curve:

$$R_{m,t+h} = L^f_{t+h} + S^f_{t+h}\left(\frac{1 - e^{-m\lambda}}{m\lambda}\right) + C^f_{t+h}\left(\frac{1 - e^{-m\lambda}}{m\lambda} - e^{-m\lambda}\right). \tag{8}$$

The forecasts for latent factors are formulated in three variants, which we describe below.

DL-AR.   Here we predict each factor $f \in \{L, C, S\}$ with an autoregressive model:

$$f_t = \alpha + \sum_{p=1}^{P} \rho_p f_{t-p} + \epsilon_t. \tag{9}$$

DL-RF.   Here we predict each factor with a random forest method described in the previous subsection.

DL-VAR. In the last method we allow for a dynamic interaction between the factors by assuming that the law of motion for the vector $Z_t = [L_t \, S_t \, C_t]'$ is well described by the vector autoregression (VAR) process:

$$Z_t = A + \sum_{p=1}^{P} B_p Z_{t-p} + v_t. \tag{10}$$

Table 1: Descriptive statistics

	Mean	St. Dev.	Min.	Max.	ACF	ADF
Swap interest rate curves						
1M	4.70	3.44	1.25	18.22	0.99	-0.01 D***
3M	4.65	3.38	1.43	18.21	0.99	-0.01 D***
6M	4.65	3.29	1.48	18.33	0.99	-0.01 D***
9M	4.63	3.24	1.46	18.42	0.99	-0.01 D***
1Y	4.63	3.21	1.43	18.46	0.99	-0.01 D***
2Y	4.67	2.91	1.46	17.26	0.99	-0.02 D,T***
3Y	4.71	2.73	1.51	16.10	0.99	-0.02 D,T***
4Y	4.76	2.60	1.53	15.19	0.99	-0.02 D,T***
5Y	4.80	2.48	1.52	14.24	0.99	-0.01 D***
7Y	4.83	2.27	1.52	13.19	0.99	-0.01 D***
10	4.84	2.01	1.53	12.00	0.99	-0.02 D,T***
DL factors						
L	4.98	1.71	1.5 0	11.23	0.99	-0.02 D,T*
C	-0.48	2.09	-5.48	12.11	0.95	-0.04 N***
S	-0.30	2.18	-2.61	10.91	0.99	-0.01 N***

Notes: ACF and ADF refer to the values of the autocorrelation coefficient for the first lag and the Augmented Dickey Fuller tests. Asterisks ***, **and *denote the rejection of the null that series is a non-stationary series at 1%, 5% and 10% significance level, respectively. D,T,N denote whether the series exhibits drift, trend or none of these.

4 Data

To assess the predictability of the above models, we have collected daily data on swap yields from Polish intrabank market. More precisely, we have gathered daily series of average rates for the following tenors: 1M, 3M, 6M, 9M, 1Y, 2Y, 3Y, 4Y, 5Y, 7Y, 10Y over the period from 2000:11 to 2019:10 from the Thomson Reuters Eikon database. They were later converted into weekly data. We have transformed the series (r_t) into

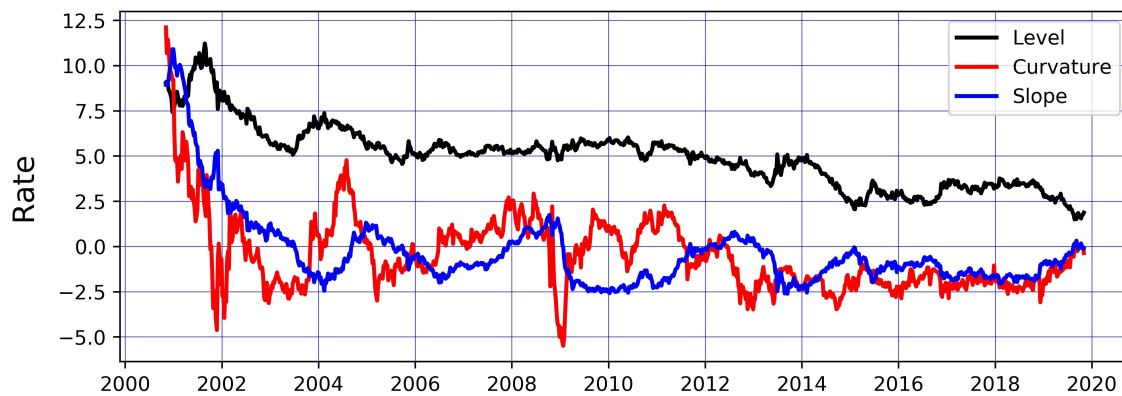continuously compounded yields with the following formula:

$$R_t = \log(1 + \frac{r_t}{100}) \times 100. \tag{11}$$

Figure 1: Swap interest rate curves on Polish intrabank market over the period 2000-2019
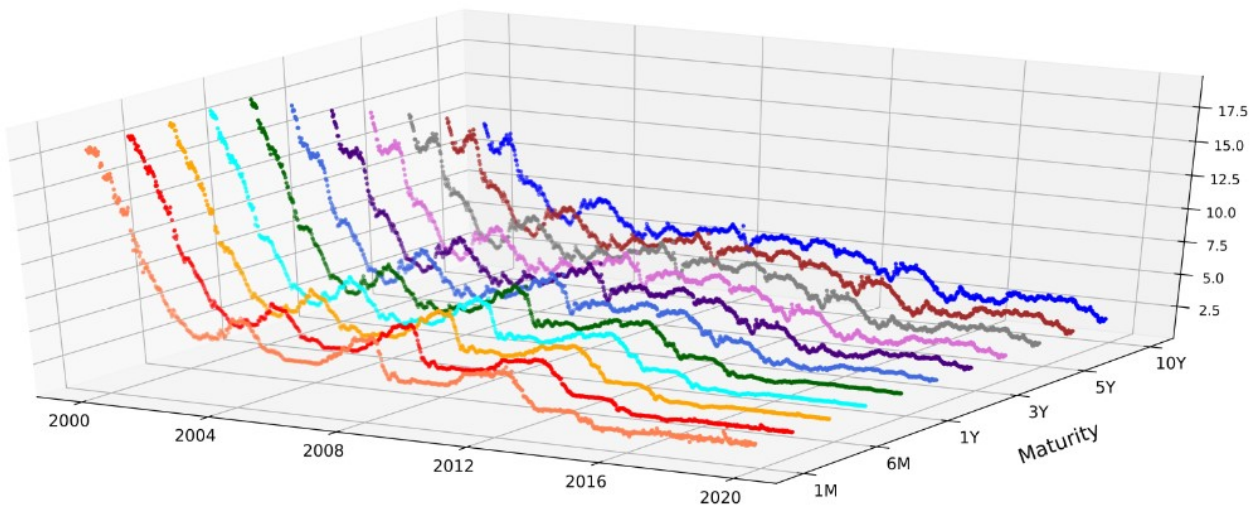
Source: Thomson Reuters Eikon.

Figure 2: Level, Curvature, Slope factors over the period 2000-2019

Source: Own calculations.

The values of the rates at different maturities are presented in Figure 1. It shows that interest rates at all maturities have steadily decreased from nearly 20%, observed nineteen years ago, to levels below 2% at the end of the sample. Table 1 shows descriptive statistics of the swap interest rate set.

In the next step, we calculated the series for the level, slope and curvature factors (L_t, S_t and C_t) for each t. We plotted them in Figure 2. It shows that the level of interest rate declined over the period, whereas slope and curvature tended to fluctuate without trend or drift. Table 1 provides descriptive statistics for these factors.

Table 2: Mean Forecast Errors (MFE)

	1	4	12	24	32	52
One-month maturity						
RW	0.00	-0.02	-0.05	-0.09	-0.12	-0.19
NS	0.01	-0.01	-0.07*	-0.17	-0.24	-0.43
AR	0.00	-0.01	-0.04	-0.07	-0.08	-0.12
RF	0.00	-0.01	-0.04	-0.09	-0.12	-0.19
DL-AR	0.00	-0.03	-0.10	-0.20	-0.25	-0.38
DL-VAR	0.02**	0.02**	0.02	0.01	-0.01	-0.09
DL-RF	0.02*	0.01	-0.03	-0.07	-0.09	-0.17
Two-year maturity						
RW	0.00	-0.02	-0.05	-0.09	-0.12	-0.22
NS	0.02***	-0.01	-0.09	-0.19	-0.26	-0.45*
AR	-0.01	-0.02	-0.07	-0.12	-0.15	-0.24
RF	-0.01	-0.02	-0.06	-0.10	-0.13	-0.22
DL-AR	0.01	-0.02	-0.11	-0.21	-0.27	-0.42
DL-VAR	0.01**	0.00	-0.05	-0.12	-0.17	-0.32
DL-RF	0.02**	0.01	-0.02	-0.06	-0.09	-0.19
Ten-year maturity						
RW	0.00	-0.02	-0.06	-0.10	-0.13	-0.22
NS	0.00	-0.02	-0.07	-0.14	-0.19	-0.31
AR	-0.01	-0.03	-0.08	-0.13	-0.16	-0.25
RF	-0.01	-0.02	-0.05	-0.10	-0.13	-0.21
DL-AR	-0.01	-0.04	-0.11	-0.19	-0.25	-0.38
DL-VAR	-0.01	-0.04	-0.11	-0.20	-0.26	-0.42
DL-RF	0.00	-0.01	-0.04	-0.08	-0.11	-0.20

Notes: Asterisks ***, **and *denote the rejection of the null that the MFE is equal to zero at 1%, 5% and 10% significance level, respectively. AR, RF, DL-AR, DL-VAR, DL-RF model were estimated with using two lagged values (p=2)

5   Results

We produced forecasts for weekly horizons starting from week 1 up to week 52. The evaluation was based on data covering the period from 2005:11 to 2019:10, henceforth called the evaluation sample. The models were estimated based on 5-year (260-week) rolling window. To illustrate, the first set of 52 forecasts produced in 2005:11, for the period between 2005:11 and 2006:10, was generated using the models estimated based on observations from 2000:11 to 2005:10. This procedure was repeated for each week from the period between 2005:11 and 2019:10. The forecasts of factors are presented in Figure 3 and forecasts of interest rates at different maturities in Figures 4, 5 and 6.

Table 3: Root Mean Squared Forecast Errors (RMSFE)

	1	4	12	24	32	52
One-month maturity						
RW	1.00	1.00	1.00	1.00	1.00	1.00
NS	0.65***	0.67***	0.67*	0.75	0.78	0.85*
AR	0.97**	1.00	1.00	1.00	1.00	0.99
RF	1.08***	1.05*	1.02	1.02*	1.01*	1.01***
DL-AR	1.00	1.00	1.08***	1.11**	1.10*	1.06
DL-VAR	0.96	0.77**	0.71	0.83	0.85	0.83
DL-RF	1.61***	1.15***	1.02	1.01	1.01	1.03***
Two-year maturity						
RW	1.00	1.00	1.00	1.00	1.00	1.00
NS	0.36***	0.86**	0.97	1.01	1.03	1.07
AR	1.01*	1.02	1.04	1.07	1.08	1.12
RF	1.25***	1.09***	1.01	1.00	1.00	1.00
DL-AR	1.08***	1.12***	1.21**	1.23	1.23	1.25
DL-VAR	1.07***	1.07*	1.12	1.17	1.16	1.18
DL-RF	1.56***	1.21***	1.06***	1.04**	1.04**	1.04*
Ten-year maturity						
RW	1.00	1.00	1.00	1.00	1.00	1.00
NS	0.40***	0.89***	0.97	1.00	1.02	1.05
AR	1.00	1.02	1.06	1.09	1.11	1.14
RF	1.21***	1.08***	1.02	1.01	1.01	1.01
DL-AR	1.06***	1.06*	1.10	1.14	1.15	1.17
DL-VAR	1.09***	1.09***	1.14**	1.19	1.22	1.33
DL-RF	1.31***	1.1***	1.02	1.01	1.01	1.01

Notes: Forecasts are scaled with benchmark value equal to 1. We use the Diebold-Mariano test to verify forecast accuracy, with the null hypothesis that a method has the same forecast accuracy level as the RW benchmark against the alternative that it has a different level of accuracy. Asterisks ***, **and *denote the 1%, 5% and 10% significance levels, respectively. AR, RF, DL-AR, DL-VAR, DL-RF models were estimated using two lagged values (p=2).

5.1   Mean Forecast Errors

We began the forecasting contest by calculating the mean forecast errors (MFE) for yields at three maturities: 1 month (short-term rate), 2 years (mid-term rate) and 10 years (long-term rate). The MFE values, complemented with the results of the unbiasedness test, with the null hypothesis stating that the MFE equals zero, are presented in Table 2. We observed that all models for the short and mid-term maturities, as well as for horizons longer than 1 week, tended to produce forecasts that are higher than observed values. For horizons equal to 1 week, NS as well as DL models tended to underpredict. However, statistically, with the exception of DL-VAR model for short horizon, the MFE are not significantly different than zero.

Table 4: RMSFE Ranking table

|	1	4	12	24	32	52
	Models with the lowest RMSFE					
1M	NS	NS	NS	NS	NS	DL-VAR
3M	NS	NS	NS	NS	NS	DL-VAR
6M	NS	NS	NS	NS	NS	DL-VAR
9M	NS	NS	NS	NS	NS	DL-VAR
1Y	NS	NS	NS	NS	NS	NS
2Y	NS	NS	NS	RF	RF	RF
3Y	NS	NS	NS	RW	RW	RW
4Y	NS	NS	NS	RW	RW	RW
5Y	NS	NS	NS	RW	RW	RW
7Y	NS	NS	NS	RW	RW	RW
10Y	NS	NS	NS	RW	RW	RW
	Models with the highest RMSFE					
1M	DL-RF	DL-RF	DL-AR	DL-AR	DL-AR	DL-AR
3M	DL-RF	DL-RF	DL-RF	DL-AR	DL-AR	DL-AR
6M	DL-RF	DL-RF	DL-RF	DL-RF	DL-RF	DL-RF
9M	DL-RF	DL-RF	DL-RF	DL-RF	DL-RF	DL-RF
1Y	DL-RF	DL-RF	DL-AR	DL-AR	DL-AR	DL-AR
2Y	DL-RF	DL-RF	DL-AR	DL-AR	DL-AR	DL-AR
3Y	DL-RF	DL-RF	DL-AR	DL-AR	DL-AR	DL-AR
4Y	DL-RF	DL-RF	DL-AR	DL-AR	DL-VAR	DL-VAR
5Y	DL-RF	DL-RF	DL-AR	DL-VAR	DL-VAR	DL-VAR
7Y	DL-RF	DL-RF	DL-AR	DL-VAR	DL-VAR	DL-VAR
10Y	DL-RF	DL-RF	DL-VAR	DL-VAR	DL-VAR	DL-VAR
Ranking	Averaged RMSFE across maturities					
1	NS	NS	NS	NS	NS	NS
2	RW	RW	RW	RW	RW	RW
3	AR	AR	AR	AR	AR	RF
4	DL-VAR	DL-VAR	DL-VAR	DL-VAR	DL-VAR	AR
5	DL-AR	DL-AR	DL-AR	RF	RF	DL-VAR
6	RF	RF	RF	DL-RF	DL-RF	DL-RF
7	DL-RF	DL-RF	DL-RF	DL-AR	DL-AR	DL-AR

Notes: AR, RF, DL-AR, DL-VAR, DL-RF models were estimated using two lagged values (p=2).

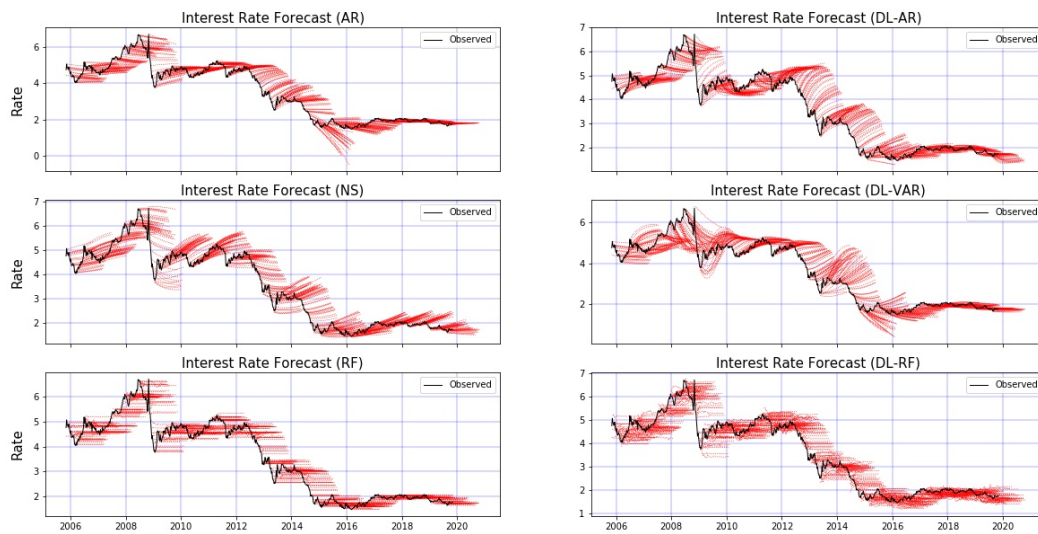Figure 3: Estimated factor forecasts for the period 2006-2019
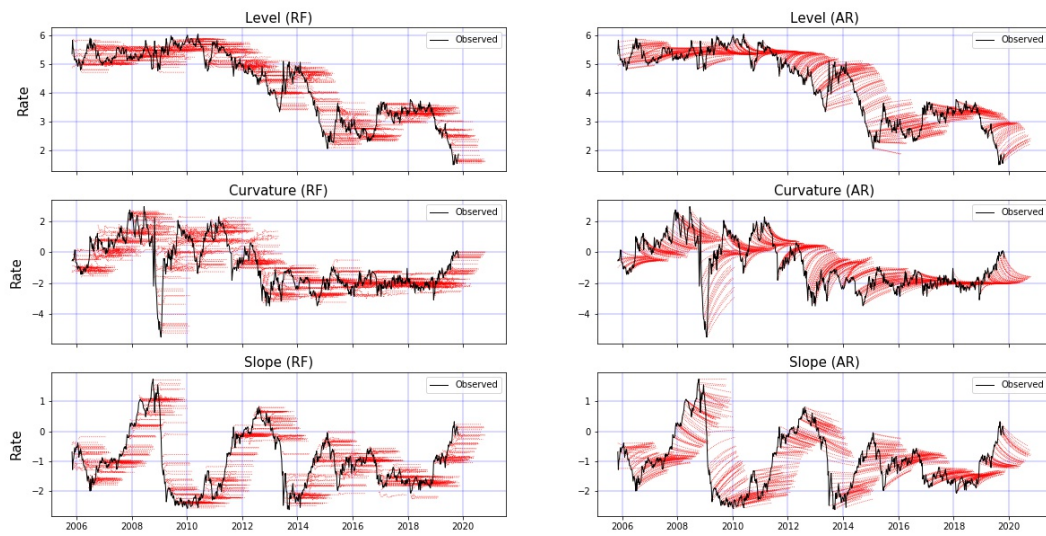
Source: Own calculations.

Figure 4: Swap interest rates forecasts over the period 2006-2019 for 1 month maturity
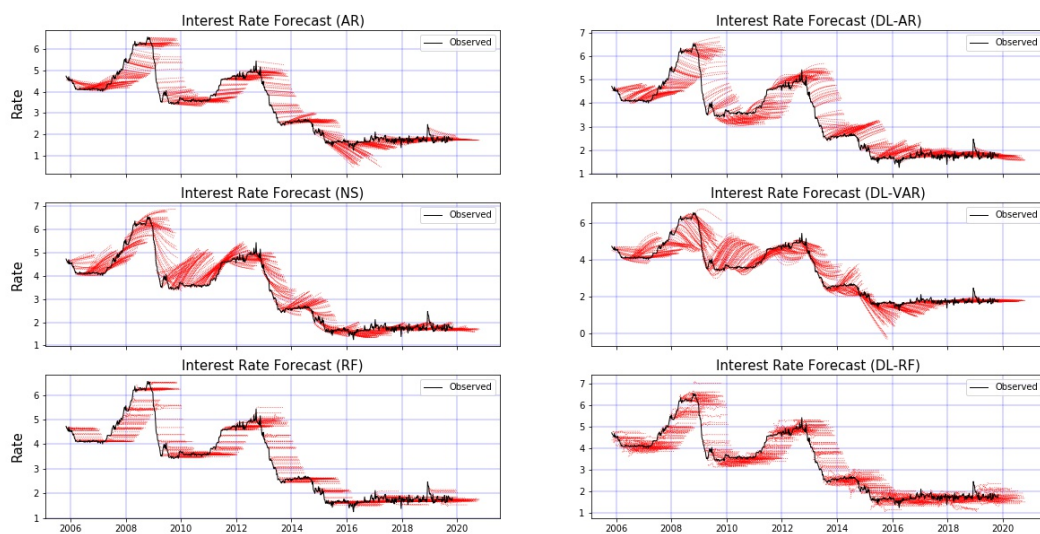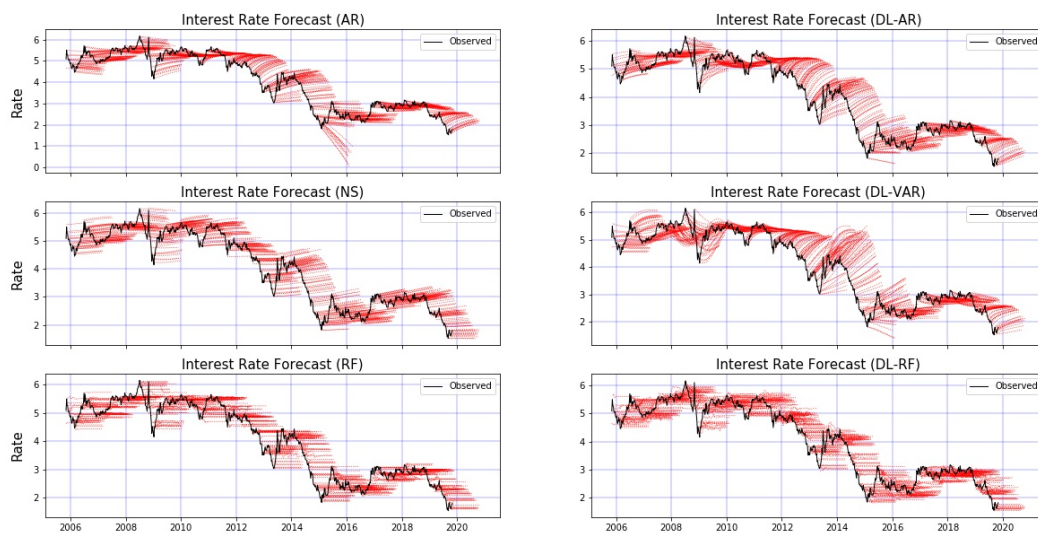

Source: Own calculations.

Figure 5: Swap interest rates forecasts over the period 2006-2019 for 2 year maturity

Source: Own calculations.

Figure 6: Swap interest rates forecasts over the period 2006-2019 for 10 year maturity
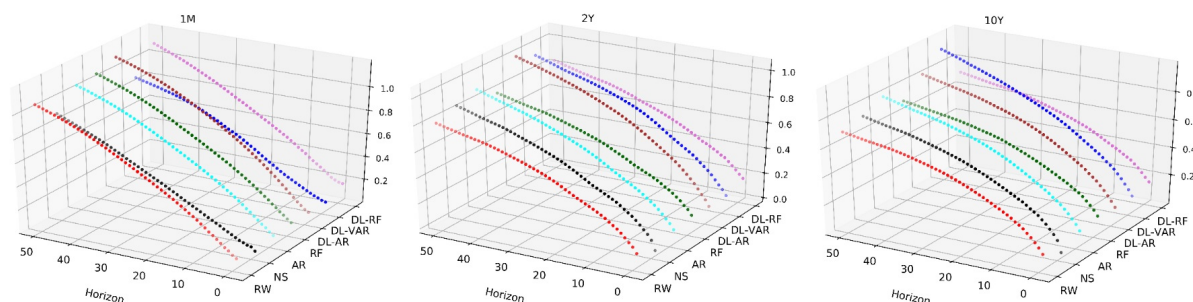

Source: Own calculations.

5.2 Root Mean Squared Forecast Errors

The root mean squared forecast (RMSFE) results are presented in Tables 3 and 4, and in Figure 7. We complemented them with accuracy of forecast analysis based on the Diebold

and Mariano (1995) tests. Our null hypothesis is that the benchmark's forecast accuracy is not significantly different from the other models, with the alternative that forecasts generated by benchmark and the other models have different levels of accuracy (two-tail test). Our findings are as follows: For maturities up to and including 1 year, NS model dominates other models. Furthermore, for 1-week horizon, it provides the results that are statistically more accurate than the benchmark according to the Diebold and Mariano (1995) test. The NS model is followed by AR and DL-VAR models, where AR tend to perform better for short horizons and DL-VAR for longer ones, even beating NS model for the 52-week horizon. However, in the case of AR and DL-VAR models, the accuracy is not significantly different from the baseline, DL-RF is the worst performing models. With respect to maturities longer than 1 year, the NS model still dominates, yet only for short term horizons, for which it delivers forecasts of higher statistical accuracy than the benchmark. For horizons longer than 12 weeks. it is the RF and RW that provide the most accurate predictions. The worst performing models for short horizons are DL-RF and DL-AR, and DL-VAR for longer horizons. Interestingly, for the machine learning models, both RF and DL-RF produce the forecasts across all the maturities that tend to be statistically worse than a benchmark for horizons shorter than 12 weeks. Their forecast accuracy increases for longer horizons where it becomes comparable with the benchmark. Finally, in general forecasts, generated by dynamic affine models, are statistically less accurate than the benchmark or those produced by the NS model.

Figure 7: Evolution of Root Mean Squared Forecast Error

Source: Own calculations.

All in all, our results suggest that, from the set of models presented, the NS performs best. It produces the best predictions for short term maturities practically across all horizons and for the maturities equal or longer than two years, it is the most accurate for the horizons up to 12 weeks and just slightly worse for longer horizons. It is also the only model whose forecasts are statistically more accurate than those from the benchmark (for short horizons) or are at the same level of accuracy. For mid- and long-term maturities, the benchmark and the RF provide best predictions for horizons equal and longer than 12 weeks. On average, it appears that DL models performs worse than univariate models. Interestingly, for longer horizons, the DL extension based on a random forest provides

more accurate forecasts compared to other DL extensions such as simple autoregression and vector-autoregression.

6   Discussion and Conclusions

This paper compares the accuracy of interest rates forecasts for different maturities based on univariate models, and on various specifications of the Diebold–Li approach. The analysis was performed for Polish intrabank swap curves. We have found that Nelson–Siegel expectation hypothesis outperformed other models in terms of the accuracy of the forecasts as measured by RMSE. The NS approach turned out to be statistically more accurate than the benchmark for short forecast horizons across all maturities, producing the results that were on par with other approaches for longer horizons. For maturities over 1 year and longer horizons, however, it was the benchmark that delivered the most accurate results. In terms of DL extensions, with the exception of DL-VAR for short maturities, they underperformed, compared to univariate models. In general, dynamic latent factors did not capture the declining trend of the level factor. The period under investigation exhibited a steady decline of the interest rates, which dropped from nearly 20% to 2% over the last two decades. This decline might be attributed to more favorable inflation developments, but it may also be linked to a steady decline of real interest rates (Summers, 2014). Finally, we have not found strong evidence regarding the superiority of the machine learning techniques based on the random forest over traditional time series models.

References

Breiman, Leo. 2001. "Random forests." *Machine learning* 45 (1):5–32.

Christensen, Jens H. E. and Glenn D. Rudebusch. 2015. "Estimating Shadow-Rate Term Structure Models with Near-Zero Yields." *Journal of Financial Econometrics* 13 (2):226–259.

Diebold, Francis X. and Canlin Li. 2006. "Forecasting the term structure of government bond yields." *Journal of Econometrics* 130 (2):337–364.

Diebold, Francis X., Glenn D. Rudebusch, and Boragan S. Aruoba. 2006. "The macroeconomy and the yield curve: a dynamic latent factor approach." *Journal of Econometrics* 131 (1-2):309–338.

Geyer, Alois and Richard Mader. 1999. "Estimation of the term structure of interest rates - A parametric approach." Working Papers 37, Oesterreichische Nationalbank.

Gurkaynak, Refet S. and Jonathan H. Wright. 2012. "Macroeconomics and the Term Structure." *Journal of Economic Literature* 50 (2):331–67.

Hladíková, Hana and Jarmila Radová. 2012. "Term structure modelling by using Nelson-Siegel model." *European Financial and Accounting Journal* 7 (2):36–55.

Jung, Carsten, Henrike Mueller, Simone Pedemonte, Simone Plances, and Oliver Thew. 2019. *Machine learning in UK financial services.* Bank of England and Financial Conduct Authority.

Marciniak, Marek. 2006. "Yield curve estimation at the National Bank of Poland." *Bank i Kredyt* 10:52–74.

Martin, Daniel, Barnabás Póczos, and Burton Hollifield. 2018. "Machine learning-aided modeling of fixed income instruments." .

Nelson, Charles R. and Andrew F. Siegel. 1987. "Parsimonious Modeling of Yield Curves." *The Journal of Business* 60 (4):473–89.

Rubaszek, Michał. 2016. "Forecasting the yield curve with macroeconomic variables." *Econometric Research in Finance* 1 (1):1–21.

Summers, Lawrence H. 2014. "U.S. Economic Prospects: Secular Stagnation, Hysteresis, and the Zero Lower Bound." *Business Economics* 49 (2):65–73.

Svensson, Lars E.O. 1994. "Estimating and Interpreting Forward Interest Rates: Sweden 1992 - 1994." NBER Working Papers 4871, National Bureau of Economic Research, Inc.

Yu, William and Eric Zivot. 2010. "Forecasting the term structures of treasury and corporate yields: Dynamic nelson-siegel models evaluation." *International Journal of Forecasting, Forthcoming* .

Zoricic, Davor and Marko Badurina. 2013. "Nelson-Siegel Yield Curve Model Estimation And The Yield Curve Trading In The Croatian Financial Market." *UTMS Journal of Economics* 4 (2):113–125.