How accurate are macroeconomic forecasts of Slovakian commercial banks?

Ádám Csápai[1]

DOI: 10.32725/978-80-7394-976-1.31

Abstract: In this paper, we examine the accuracy of the forecast models of the selected commercial banks in Slovakia. We focus on GDP and inflation forecasts in three different forecast horizons. The data is taken from an ongoing survey administered and regularly updated by the National Bank of Slovakia. The sample consists of nine of the largest commercial banks in Slovakia. Firstly, we see how well the banks compare to each other based on RMSE. Secondly, we estimate different baseline models, namely random walk and ARIMA models, for each horizon and compare each bank's accuracy to the accuracy of these baseline models. Data is taken from the OECD revisions database, Eurostat and the database of the National Bank of Slovakia. Thirdly, we take different types of averages of the individual banks' forecasts, as forecast averaging should increase accuracy based on contemporary forecasting scientific literature. We find that the models of these commercial banks are generally more accurate than our baseline models, except for one bank, which systematically provides less accurate forecasts in the survey. What is more, we find that individual models are capable of being more accurate than the averages for each and every horizon.

Keywords: forecast comparison, arima, random walk, commercial banks
JEL Classification: G32, G33, C35

1 Introduction

The aim of this paper is to find out whether there is such a macroeconomic model in some of the Slovak commercial banks, which can systematically predict the development of macroeconomic indicators, specifically GDP and inflation, at different time horizons more accurately than our benchmark models, models of other banks and averages of individual forecasts.

The paper is structured as follows. Firstly, we describe the sources of our data. Secondly, we characterize the survey conducted by the National Bank of Slovakia among commercial banks. Thirdly, we describe the horizons at which we compare the forecasting ability of the models. We also provide an overview of the data preparation and characterize the benchmark models, as well as the theory behind our research. In section 4 we present our results and the final section concludes.

2 Methods

2.1 Data description

To begin with, we collected data from three different sources. First, the source for obtaining forecasts of commercial banks was a survey. The survey is carried out monthly by the National Bank of Slovakia (hereinafter NBS). The data has been collected approximately in the last working week of the given month since 2007. In the survey, commercial banks report their forecasts of key macroeconomic indicators, such as the exchange rate, interest rates, balance of payments, inflation and GDP at constant or current prices. The set of included indicators has changed over time. In this commentary, we focus on inflation measured by HICP and GDP at constant prices. At the beginning, nine commercial banks sent their forecasts, the list of which is presented in Table 1. Later, one bank ceased operations and two banks stopped their participation in the survey. For this reason, we had at our disposal data from six banks for the whole period, but not all of these banks sent their forecasts every month. In case the monthly forecast was missing, we assumed that its value did not change compared to the previous month. In this way, we could consider the survey a reliable source of information, which we then compared to the actual values of GDP and inflation.

[1] University of Economics in Bratislava, Department of Operations Research and Econometrics, Dolnozemská cesta 1, 852 35 Bratislava, Slovak Republic, adam.csapai@euba.sk

Table 1 List of participating banks

List of participating banks		
Bank	First contribution	Last contribution
Poštová Banka (PABK)	2007M10	2020M06
ČSOB Banka (CSOB)	2007M01	2020M06
Komerčná Banka	2007M04	2008M06
ING Banka	2007M01	2012M09
Istrobanka	2007M01	2009M04
Slovenská Sporiteľňa (SLSP)	2007M01	2020M06
Tatrabanka (TATR)	2007M01	2020M06
Unicredit Banka (UNIB)	2007M01	2020M06
VÚB Banka (VUB)	2007M01	2020M06

Source: Own processing

Our second source of information were the NBS and OECD databases. We obtained the seasonally and calendar adjusted GDP time series and the inflation data from these. In the case of GDP, we used revised data so that we could compare the forecasts with the actual GDP values available at the time the banks sent us their forecasts.

The only issue with the data was presented by the different frequencies. As mentioned above, banks send their forecasts on a monthly basis, but the available GDP data is quarterly. Therefore, we interpolated the available GDP data points in such a way that we considered the same quarterly value as the real GDP value in each month of the given quarter.

We considered three different time horizons for the forecast exercises. In the first horizon we compared actual GDP growth data to the same quarter of the previous year. In the second forecast horizon, we analyzed the year-on-year GDP growth in the current year. In the third and final forecast horizon we forecast the year-on-year GDP growth for next year.

With regards to inflation, we obtained the time series from the Eurostat database. We measured inflation as the annual rate of change in the HICP. Monthly inflation data is readily available on Eurostat, so we did not have to interpolate the values as in the previous case with GDP. We also took into account three forecast time horizons. The first was the forecast of year-on-year price growth. The second was forecasting inflation at the end of the actual calendar year. The third and final forecast horizon represented the inflation forecast at the end of next year.

2.2 Benchmark models

In this chapter, we describe our model selection. To begin with, in addition to the banks' point estimates we obtained from the survey, we estimated two benchmark models of our own. Our first benchmark model is a simple random walk model, as suggested by Bjornland et al (2012). In this case, the forecast value from the previous period was equal to the last known value of the time series. This represents the simplest way to model a stochastic time series. We write the model as

$$Y_t = Y_{t-1} + u_t \qquad (1)$$

where the u_t are zero mean $E(u_t \mid Y_{t-1}, Y_{t-2}, \ldots) = 0$ i.i.d. errors. Given that

$$\begin{aligned} E(Y_t \mid Y_{t-1}, Y_{t-2} \ldots) &= E(Y_{t-1} \mid Y_{t-1}, Y_{t-2} \ldots) + E(u_t \mid Y_{t-1}, Y_{t-2} \ldots) \\ &= Y_{t-1} \end{aligned}$$

it is clear that we consider yesterday's observation Y_{t-1} to be the most accurate forecast for Y_t. This entails that we cannot predict the difference between Y_t and Y_{t-1}, and Y_t follows a path given by random steps u_t.

Our second type of benchmark models are ARIMA models, again suggested by Bjornland et al (2012). These models forecast the future path of the time series based on its historical variations. We write

$$y_t = \alpha + \sum_{j=1}^{p} \phi_j y_{t-j} + \sum_{j=0}^{q} \theta_j \varepsilon_{t-j}, \qquad (2)$$

where y_t gives us the target variable, while p and q represent the lag order of the autoregressive (AR) and moving average (MA) terms, respectively. We estimated these models in EViews, included a maximum of five AR and MA elements in (2) and used I(1) ARIMA models when necessary.

To continue with, Aiolfi et al. (2011) states that the combination of individual forecasting models can substantially improve forecast accuracy. Burgi (2015), **for example, argues that it is hard to achieve more accurate estimates than by combining the individual models using equal weights.** Because of this, we first combined the models by taking the arithmetic mean of the individual forecasts of commercial banks. In addition, Armstrong (2001) argues that a trimmed mean may also improve forecast accuracy. To calculate the trimmed mean, we excluded the lowest and highest forecast values for each single observation, thereby varying the model space over time. We then calculated the arithmetic mean of the remaining forecasts.

To sum up, we compared the forecast accuracy of the aforementioned benchmark models, individual models of commercial banks and the combination of these individual models. We measure forecast accuracy by the root mean squared error (hereinafter RMSE). RMSE is the square root of the mean of the square of all of the error. The lower the RMSE is, the more accurate is the forecast. The RMSE is formulated as

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (S_i - O_i{}^2)}$$

where S_i gives us the forecasted values of the variable, O_j represents the observations, and n stands for the number of observations available for analysis.

3 Results

In this chapter we describe our findings. In Table 2 we see the RMSE for each model in each period. The smallest RMSE in each column is highlighted in bold. Several interesting results stand out. Regarding averages in general, it appears that the arithmetic mean of all bank models could forecast GDP more accurately than the trimmed mean, regardless of the prediction horizon. The reason may be that we included only six banks in the analysis, as sufficient data was not available for the others. In our case, when we calculated the trimmed arithmetic mean, the highest and lowest observations for each time period were discarded. With this however, the Makridakis and Winkler (1983) rule might be violated, according to which it would be better to average at least 5 models. For the shortest horizon, the benchmark RMSEs are 1.71 for RW and 3.04 for ARIMA. Table 2 shows that all banks outperformed the ARIMA benchmark model, while RW forecasted the development of GDP better than the CSOB model. The VUB model most accurately forecasted the development of GDP. Apart from VUB, not even one of the banks was able to forecast GDP more accurately than the average of all six banks, which we expected based on the literature. Furthermore, it is also written in the literature that the model that is the best on one horizon lags behind the others on other horizons in its performance.

Table 2 GDP forecast

GDP forecast			
Model	Q-o-Q	Actual year	Next year
Arithmetic mean	0.9	0.76	3.14
Trimmed mean	0.98	0.79	3.19
ARIMA	3.04	3.19	3.53
RW (random walk	1.71	2.01	4.13
PABK	0.92	0.80	3.30
CSOB	2.11	2.20	3.19
SLSP	1.10	0.86	**3.00**
TATR	1.00	1.16	3.40
UNIB	0.97	**0.72**	3.04
VUB	0.81	0.77	3.01

Source: Own processing

The above stated assumption was true in our case, since on a yearly horizon, it was no longer the model from VUB, but the model from UNIB that had the most accurate prognostic ability. Although it was closely followed by VUB, which was in second place among the banks, but the RMSE of 0.77 was not enough in this case to overcome the arithmetic average of all six banks. Apart from UNIB, not a single bank could surpass this average. The result on the annual horizon was, therefore, similar to the result on the shortest horizon, since there were banks that could beat the arithmetic average, which finished in second place. Furthermore, each bank built a model that is more accurate than the benchmark ARIMA, but the model from CSOB cannot outperform the RW benchmark. This model has an RMSE of 2.20 compared to the RMSE of the RW model of 2.01. As for the longest horizon, RW with its RMSE of 4.53 was significantly behind even the ARIMA model, and these benchmark models have a larger RMSE than the models of individual banks. SLPS analysts were the most accurate in forecasting the development of GDP in this period. The model from TATR, which with its RMSE of 3.40, was the least accurate in this horizon. It is interesting that while only one bank had a smaller RMSE than the arithmetic mean at shorter horizons, at the longest horizon this applies to three banks. From this we can draw the conclusion that in our case the GDP predictions of these individual banks became more accurate relative to the average with an increasing length of forecast horizons. The question is whether the same banks had the same results in the case of inflation.

We see these results in Table 3. Here, too, we highlight the smallest RMSE in each column in bold. As for the means, the trimmed arithmetic mean was better only on the annual horizon, where it had an RMSE of 0.99 compared to the arithmetic mean with an RMSE of 1.00. As for the shortest horizon, the model from CSOB has the worst predictive ability with an RMSE of 1.36. Apart from this model, the model from VUB did not even surpass the ARIMA benchmark model. Among the benchmark models RW was again better than ARIMA, but both were outperformed by the UNIB, SLSP and TATR bank models, while the PABK model has the exact same RMSE. Another interesting fact is that on this horizon, none of the banks with their models exceeded the accuracy of the arithmetic mean, whose RMSE is 0.19.

Table 3 Inflation forecast

Inflation forecast			
Model	Q-o-Q	Actual year	Next year
Arithmetic mean	**0.19**	1.00	1.80
Trimmed mean	0.24	0.99	1.83
ARIMA	0.35	1.25	2.31
RW (random walk	0.32	0.92	2.24
PABK	0.32	1.07	1.89
CSOB	1.36	1.59	1.96
SLSP	0.22	1.05	1.84
TATR	0.23	1.02	1.73
UNIB	0.20	0.98	1.89
VUB	0.42	1.08	1.91

Source: Own processing

As for the annual horizon, the predictive power of the models generally deteriorated as expected. In this case, our benchmark RW model had the smallest RMSE, namely 0.92, which was not surpassed by any of the averages or any bank. Our next benchmark ARIMA model had an RMSE of 1.25. There was only one model whose prediction ability was worse than this, and that is the model from CSOB. The RMSE of this model was 1.59. We conclude the interpretation of the results by evaluating the last column, which represents the RMSE for the next year. On this horizon, the least accurate forecasts were produced by the benchmark RW and ARIMA models, whose RMSEs are 2.24 and 2.31, respectively. The model from TATR, whose RMSE was 1.73, had the best prognostic ability. CSOB once again had the largest RMSE among all banks, although this time it surpassed the benchmark, but its models were the least accurate at every horizon. This also applies to GDP forecasts. The reason for this systematic inaccuracy is unknown.

4 Conclusions

The aim of this paper was to find out whether there is such a macroeconomic model in some of the Slovak commercial banks which can systematically predict the development of macroeconomic indicators, specifically GDP and inflation, at different horizons more accurately than benchmark models, models of other banks and averages of individual forecasts. According to our results, we can draw the following conclusions. In Slovakia, there is no macroeconomic model in any of the commercial banks that could systematically outperform the arithmetic average of individual forecasts with its predictive ability. Nevertheless, some banks succeeded, but not consistently. On the other hand, CSOB models are many times less accurate than models of other banks, or even benchmark models. However, it is important to note that almost every bank can formulate better prognostic models than our benchmark models. We consider this to be important especially for lay people, who often read the forecasts of individual banks in the media, and those forecasts can influence their financial behavior. In the future, we plan to extend these results by calculating the weighted averages of individual forecasts using artificial intelligence methods.

References

Armstrong, J. S. (2001). Combining Forecasts. University of Pennsylvania Scholarly Commons, Marketing Papers 34, 1-21.

Aiolfi et al. (2012). Forecast Combinations. The Oxford Handbook of Economic Forecasting, 1-60, DOI: 10.1093/oxfordhb/9780195398649.013.0013

Bjornland, H. C. et al. (2012). Does Forecast Combination Improve Norges Bank Inflation Forecasts? Oxford Bulletin of Economics and Statistics, 74(2), 1-18, DOI: 10.1111/j.1468-0084.2011.00639.x

Bürgi, C. (2015). Can a subset of forecasters beat the simple average in the SPF? *RPF Working Paper*, 2015-001.

Makridakis, S., Winkler, R. (1983). Average of forecasts: some empirical results. *Management Science*, 29, 987-996.