EKONOMETRIA ECONOMETRICS 3(41) • 2013

ISSN 1507-3866

Martin Pavlik

INCOME TAX: A COMPARISON OF THE FORECASTING METHODOLOGIES

Abstract: The author focuses on income tax forecasting. He compares the different forecasting methodologies. He also compares his solution to the official annual forecasts in the Slovak Republic. He chose the quarters of years as the time units.

Keywords: income tax, forecasting, ARCH models, GARCH models, EGARCH models, Akaike information criteria.

1. Introduction

The short-term forecasting of tax revenue is an important part of managing the state's liquidity. Short-term forecasting, based on a monthly and quarterly basis, takes place in the Financial Directorate of the Slovak Republic. Long-term forecasting takes places in the Ministry of Finance. The key taxes are: VAT, Income tax, Corporate tax. We will focus on income tax since VAT was described in [Pavlík 2008].

We were forecasting the tax revenue of the Slovak Republic from 2004 to 2007. We developed a methodology for the forecasts [Pavlík 2011]. We have faced a problem with the short time series since the Slovak Republic was a young state with a new tax system. That is the reason why we used a set of data based on months or quarters of the year. We developed a methodology where we set as a main criteria the quality of the ex post forecast. We will compare this methodology to the methodology based on the Akaike information criteria and R^2 adjusted. The aim of the research is to develop an appropriate methodology for adequate forecasts and compare it to the forecasts which the authorities used.

2. Income tax

Income tax belongs to the main taxes of the Slovak Tax system. Income tax has two parts. Personal income tax and entrepreneurs' income tax. Entrepreneurs' income tax is paid by small businesses and is different from corporate tax which is paid by limited companies and public limited companies. Income tax is described in Income

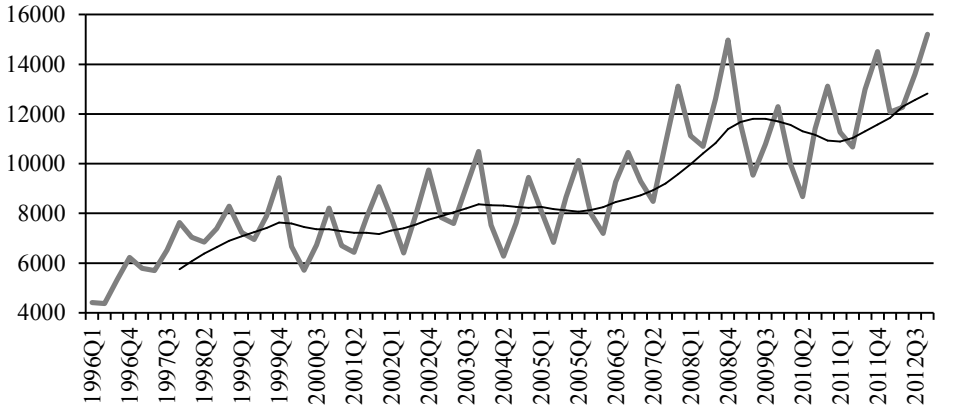


Figure 1. Personal income tax − gross yield with the trend line

Source: author's calculation.

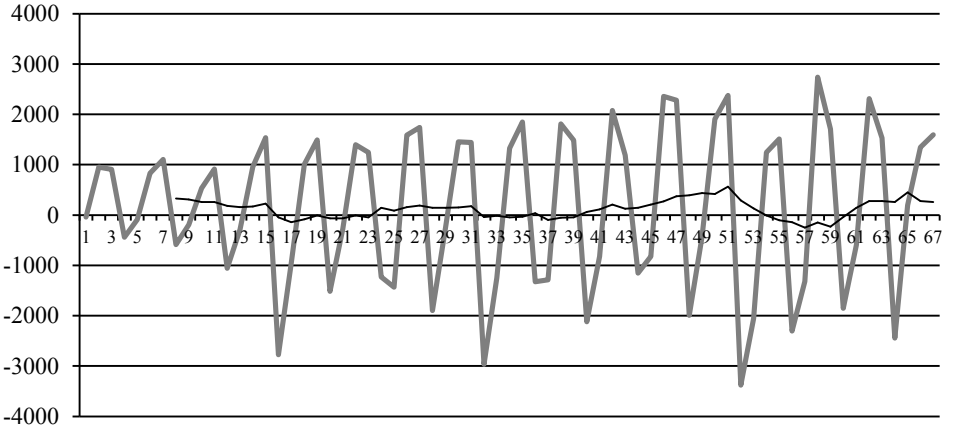


Figure 2. Personal income tax gross yield – the first difference with the trend line

Source: author's calculation.

Tax Act No. 595/2003. About 97 percent of the amount of the whole income tax is personal income tax. As anyone can see, Slovak small firms are undeveloped. Both taxes are forecasted separately. Both income taxes use their own model. We will focus on personal income tax. The tax is probably the easiest forecasted tax. Figure 1 shows the graph of the tax, the time unit is a quarter of a year.

Figure 1 shows that personal income tax represents an unstable process, because of a clearly increasing trend. We had to differentiate the time series. Figure 2 shows the differentiated time series.

3. The construction of the forecasting models

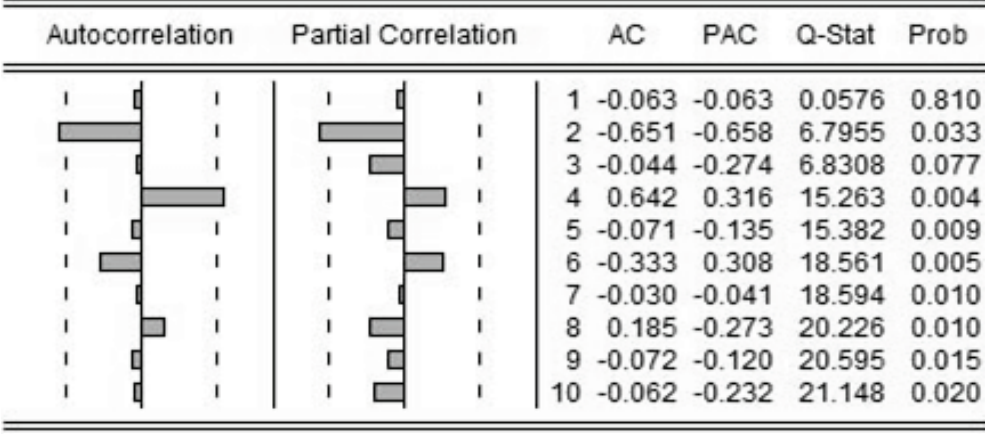
We have chosen the following forecasting models according to the correlogram.

1) $y = f(c, y_{t-4}, \varepsilon_t)$,
2) $y = f(c, y_{t-4}, y_{t-6}, \varepsilon_t)$,
3) $y = f(c, y_{t-4}, y_{t-8}, \varepsilon_t)$,
4) $y = f(c, y_{t-4}, y_{t-6}, y_{t-8}, \varepsilon_t)$.

We have not used either y_{t-2}, or moving average θ_{t-2}, because we have forecasted the whole year forecasts. We have also added the following moving averages variables to the models from (a) to (d)

a) ma(4),
b) ma(4), ma(6),
c) ma(4), ma(8),
d) ma(4), ma(6), ma(8).

So, we have tested models (1) to (4) without and with moving averages variables (a) to (d). All those models were tested in versions: ARIMA, ARCH(1), ARCH(2), GARCH(1,1), GARCH(2,1), GARCH(2,2), EGARCH(1,0), EGARCH(1,1). The number of asymmetric terms in EGARCH models was set from 1 to 3. We also tested models without the intercept. This means that we have tested 480 models.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.063	-0.063	0.0576	0.810
		2	-0.651	-0.658	6.7955	0.033
		3	-0.044	-0.274	6.8308	0.077
		4	0.642	0.316	15.263	0.004
		5	-0.071	-0.135	15.382	0.009
		6	-0.333	0.308	18.561	0.005
		7	-0.030	-0.041	18.594	0.010
		8	0.185	-0.273	20.226	0.010
		9	-0.072	-0.120	20.595	0.015
		10	-0.062	-0.232	21.148	0.020

Figure 3. Correlogram

Source: author's calculation.

Our original methodology was that we made a decision on which models would be used according to the quality of ex post forecast for the last two years. The model with the lowest root mean square for the last two years was chosen as the best one. We faced criticism for the methodology [Pavlík 2011]. Different criteria were recommended to us, especially, information criteria. We decided to rebuild our

methodology and replace the quality of the ex post forecast with the Akaike information criteria. The Akaike information criteria are widely used in econometrics nowadays. We compared the quality of the forecasts made with the Akaike information criteria with those made with the old methodology. What is more, we decided to compare the forecasts based on the Akaike information criteria with the forecasts based on the quality of ex post forecast from the last one to three years. We used the set of data from 1996 to 2012. Although the Slovak Republic was established in 1993, the older data are not available. We also added R^2 adjusted to make the research more complex.

Table 1. Personal income tax – gross yield in million SKK

Year		Reality	Forecast	Quality
2009	1q	11 600	13 129	1.13
	2q	9 543	12 472	1.31
	3q	10 788	14 790	1.37
	4q	12 297	17 234	1.40
2010	1q	9 991	9 955	1.00
	2q	8 673	8 471	0.98
	3q	11 409	10 837	0.95
	4q	13 119	12 767	0.97
2011	1q	11 268	10 952	0.97
	2q	10 676	9 740	0.91
	3q	12 988	11 942	0.92
	4q	14 516	13 939	0.96
2012	1q	12 073	12 186	1.01
	2q	12 267	10 766	0.88
	3q	13 613	13 300	0.98
	4q	15 208	15 259	1.00
RMS				**1 880**

Source: author's calculation.

Table 1 shows the quality of the forecasts of the "winning" models. The best model was chosen according to the quality of ex post forecast for the last year. This means that the model with the best forecasted results in 2008 was used as the model for 2009. The model with the best forecasted results in 2009 was used as the model for 2010. The model with the best forecasted results in 2010 was used as the model for 2011. The model with the best forecasted results in 2011 was used as the model for 2012. The *Quality* column shows the forecasts divided by the reality.

Table 2 shows the quality of the forecasts of the other "winning" models. The best model was chosen according to the quality of the ex post forecasts for the last two years. This means that the model with the best ex post forecast for 2007 and 2008 was used as the best model for 2009. The quality of the ex post forecast was

Table 2. Personal income tax − gross yield in million SKK

Year		Reality	Forecast	Quality
2009	1q	11 600	13 155	1.13
	2q	9 543	12 189	1.28
	3q	10 788	14 426	1.34
	4q	12 297	16 690	1.36
2010	1q	9 991	9 955	1.00
	2q	8 673	8 471	0.98
	3q	11 409	10 837	0.95
	4q	13 119	12 767	0.97
2011	1q	11 268	10 643	0.94
	2q	10 676	8 973	0.84
	3q	12 988	11 505	0.89
	4q	14 516	13 511	0.93
2012	1q	12 073	12 424	1.03
	2q	12 267	11 279	0.92
	3q	13 613	13 643	1.00
	4q	15 208	15 258	1.00
RMS				**1 769**

Source: author's calculation.

measured with the root mean square – RMS. The same methodology was used for 2010, 2011 and 2012.

Table 2 shows that the forecasts based on the quality of the two years ex-post forecasts are better than the forecasts based on the quality of one year forecasts. This is what the RMS and the column *Quality* show.

Table 3 shows the quality of the forecasts of the "winning" models based on the quality of the ex-post forecasts for the last three years. This means that the forecasts for 2009 were made according to the quality of ex post forecasts for 2006-2008. The same methodology was used for 2010, 2011 and 2012. Results described in Table 3 are very interesting. They are exactly the same as the results in Table 2, except 2012. Also the models which were used for the forecasts are exactly the same, except 2012. Changing the range of the ex post forecast from two to three years made almost no impact on the quality of the ex post forecasts for another year. This is what happened for one particular time series, which is the income tax in the Slovak Republic. It might be different for different time series.

We explored the forecasts for 2009-2012 using another methodology. We chose R^2 adjusted as the only criteria. This means that we have taken the sample 1996q1 to 2008q4 and calculated R^2 adjusted for all the 480 models. We found the best one and according to that model we made forecasts for 2009. We took the sample 1996q1 to 2009q4 and calculated R^2 adjusted for all the 480 models afterwards. We found the best model, which is the model with the highest R^2 adjusted and calculated forecasts

Table 3. Personal income tax – gross yield in million Skk

Year		Reality	Forecast	Quality
2009	1q	11 600	13 155	1.13
	2q	9 543	12 189	1.28
	3q	10 788	14 426	1.34
	4q	12 297	16 690	1.36
2010	1q	9 991	9 955	1.00
	2q	8 673	8 471	0.98
	3q	11 409	10 837	0.95
	4q	13 119	12 767	0.97
2011	1q	11 268	10 643	0.94
	2q	10 676	8 973	0.84
	3q	12 988	11 505	0.89
	4q	14 516	13 511	0.93
2012	1q	12 073	12 029	1.00
	2q	12 267	10 345	0.84
	3q	13 613	12 771	0.94
	4q	15 208	14 579	0.96
RMS				**1 833**

Source: author's calculation.

Table 4. Personal income tax – gross yield in million SKK

Year		Reality	Forecast	Quality
2009	1q	11 600	13 418	1.16
	2q	9 543	12 706	1.33
	3q	10 788	15 188	1.41
	4q	12 297	17 551	1.43
2010	1q	9 991	10 748	1.08
	2q	8 673	9 702	1.12
	3q	11 409	12 713	1.11
	4q	13 119	14 873	1.13
2011	1q	11 268	11 361	1.01
	2q	10 676	10 362	0.97
	3q	12 988	12 806	0.99
	4q	14 516	14 904	1.03
2012	1q	12 073	12 182	1.01
	2q	12 267	10 818	0.88
	3q	13 613	12 991	0.95
	4q	15 208	14 569	0.96
RMS				**2 090**

Source: author's calculation.

for 2010. We did the same also for 2011 and for 2012. Results are described in the Table 4.

Table 4 shows that the methodology based on the R^2 adjusted is worse than the methodology based on the quality of the ex-post forecast.

We also explored the last methodology which are the forecasts for 2009-2012 based on the Akaike information criteria.

Table 5. Personal income tax – gross yield in million Skk

Year		Reality	Forecast	Quality
2009	1q	11 600	13 150	1.13
	2q	9 543	11 867	1.24
	3q	10 788	14 357	1.33
	4q	12 297	16 139	1.31
2010	1q	9 991	10 017	1.00
	2q	8 673	8 422	0.97
	3q	11 409	10 642	0.93
	4q	13 119	12 408	0.95
2011	1q	11 268	10 728	0.95
	2q	10 676	9 106	0.85
	3q	12 988	11 563	0.89
	4q	14 516	13 348	0.92
2012	1q	12 073	12 182	1.01
	2q	12 267	10 818	0.88
	3q	13 613	12 991	0.95
	4q	15 208	14 569	0.96
RMS				**1 687**

Source: author's calculation.

The Akaike information criteria are different from R^2 adjusted. The higher the R^2 adjusted, the better it is. This is different from Akaike information criteria. The lower the criteria, the better it is. Table 5 shows the results.

Table 5 shows that the forecasts based on the Akaike information criteria are better compared with the forecasts based on the quality of the ex post forecasts. On comparing Tables 3 and 5, it can be seen that the difference is small. Another interesting fact can be seen when comparing tables 4 (R^2 adjusted) and 5. The "winning models" are the same for 2012.

Another important fact is the quality of the annual forecasts. The annual forecasts of income tax are important for the state and municipal budget, since some part of the income tax is the income of the municipal budgets. The quality of those forecasts is shown in Table 6.

Table 6. The quality of the annual forecasts in million Skk

Year	1 Year	2 Years	3 Years	R^2 adjusted	Akaike	FDF
2009	1.30	1.28	1.28	1.33	1.26	1.24
2010	0.97	0.97	0.97	1.11	0.96	1.54
2011	0.94	0.90	0.90	1.00	0.90	1.06
2012	0.97	0.99	0.94	0.95	0.95	0.98
RMS	**6 925**	**6 604**	**6 818**	**7 817**	**6 307**	**12 912**

Source: author's calculation.

It can be seen that all the methodologies provide about the same annual forecast quality and are better than the Financial Directorate forecasts. R^2 adjusted has the poorest forecast quality from all of the explored methodologies. The methodology which is based on the Akaike information criteria has the best forecast quality and demonstrates that it is justifiably widely used all over the world.

4. Comparison with the Financial Directorate solution

Table 6. and 7 also show the real solution which was applied in practice. It is shown in the FDF column. FDF means Financial Directorate Forecasts. It can be seen that FDF RMS is more than two times bigger than the solution which works with the Akaike information criteria. Our solution would have brought about a clear improvement.

An interesting fact which describes the connection between economics and politics is that 2010 was an election year, and this is probably the reason why the official forecast was so overestimated.

We used the currency which is not in use in the Slovak Republic anymore - the Slovak crown. The Slovak Republic gave up the Slovak crown in January 1st 2009. The Slovak Republic switched from the Slovak crown to the euro. The conversion rate was set at 30.126 Slovak crowns for 1 euro. So, we had time series in crowns and also in euros. We solved this problem by converting Euros to Slovak crowns with the conversion rate of 30.126. All the calculations were made in Eviews 6.0 and Excel 2007.

Table 7. Financial Directorate solution – annual forecasts in million Skk

Year	Reality	FDF	Quality
2009	44 228	54 759	1.24
2010	43 192	66 537	1.54
2011	49 448	52 505	1.06
2012	53 160	51 869	0.98

Source: author's calculation using data from www.drsr.sk.

5. Conclusion

The Akaike information criteria appear to be the best measure for measuring a model's quality. We used to use the quality of ex post forecasts as a measure for the model quality. Research shows that this is a good choice, because the Akaike information criteria brought just a small improvement. Another interesting fact which the research showed was that switching from two to three years ex post forecasts does not bring almost any improvement.

Our solution appears to be better than the solution, which the Financial Directorate uses nowadays. The facts clearly show the RMS Financial Directorate forecasts were overestimated in almost all the observed years, except 2012.

The aim of the research was carried out.

Literature

Davidson R., Mackinnon J.G., *Estimation and Inference in Econometrics,* Oxford University Press, New York 1993.

Granger C.W.J., *Forecasting in Business and Economics,* Academic Press, San Diego 1989.

Hamilton J.D., *Time Series Analysis*, Princeton University Press, Princeton 1994.

Hayashi F., *Econometrics*, Princeton, Princeton University Press 2000.

Income Tax Act No. 595/2003.

Lukáčiková A., Lukáčik M., *Ekonometrické modelovanie s aplikáciami*, EKONÓM, Bratislava 2008.

Pavlík M., *Forecasting of income tax in the Slovak Republic*, Prace Naukowe Universitetu Ekonomicznego, Ekonometria 31, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wroclaw 2011, pp. 125-131.

Pavlík M., *The Usage Of the Dummy Variable for VAT Forecasting of the Tax Administration in the Slovak Republic*, Prace Naukowe Universitetu Ekonomicznego we Wrocławiu, Ekonometria 21, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wroclaw 2008, pp. 40-54.

Vogelvang B., *Econometrics Theory and Applications with Eviews,* Financial Times Prentice Hall, 2005.

Surmanová K., *Modelling and forecasting of wages: Evidence from the Slovak Republic*, [in:] *Quantitative Methods in Economics: Multiple Criteria Decision Making XVI. Proceedings of the International Scientific Conference*, EKONÓM, Bratislava 2012. www.drsr.sk.

PODATEK DOCHODOWY: PORÓWNANIE METOD PROGNOZOWANIA

Streszczenie: Autor koncentruje się na prognozowaniu podatku dochodowego. Porównuje różne metody prognozowania, a także swoje rozwiązania dla oficjalnych rocznych prognoz w Republice Słowackiej. Jako jednostki czasu wybrał czwarte kwartały analizowanych lat.

Słowa kluczowe: podatek dochodowy, prognozowanie, modele ARCH, modele GARCH, modele EGARCH, kryterium informacyjne Akaikego.