Martin Pavlik, PhD.

FORECASTING OF INCOME TAX IN THE SLOVAK REPUBLIC

Abstract : The author focuses on forecasting income tax in cases where the time series is short. He uses the methodology which was described in his previous papers and applied on the monthly basis. He applies the methodology on an annual quarterly basis.

Key words: income tax, forecasting, ARIMA models, ARCH models, GARCH models, EGARCH models

1. Overview

Income tax is an important part of the Slovak tax assessment. The yield of the tax is 15% from the whole tax assessment. The structure of the tax assessment is shown in Figure 1.

The aim of the article is to construct forecasts for all the quarters of the year 2011.

Income tax is very specific, because it is the main part of the municipal budgets incomes. The yield of the tax is split into the national budget and municipal budgets. The history of the splitting and gross yield is shown in Table 1.

The table shows that from 2005, municipal budgets get 93.8% of gross income tax revenues. Every town has to manage its economy and the forecast of income tax revenue is important. The most important fact is the whole annual amount of the tax revenue. Forecasting on an annual basis is rather difficult in the Slovak Republic. The Slovak Republic is a new state which was established on January 1st 1993. This is the reason why the time series of income tax on an annual basis has been monitored just 17 times. That is why we have to develop models on a quarter of the year basis. Income tax time series is shown in Figure 2.

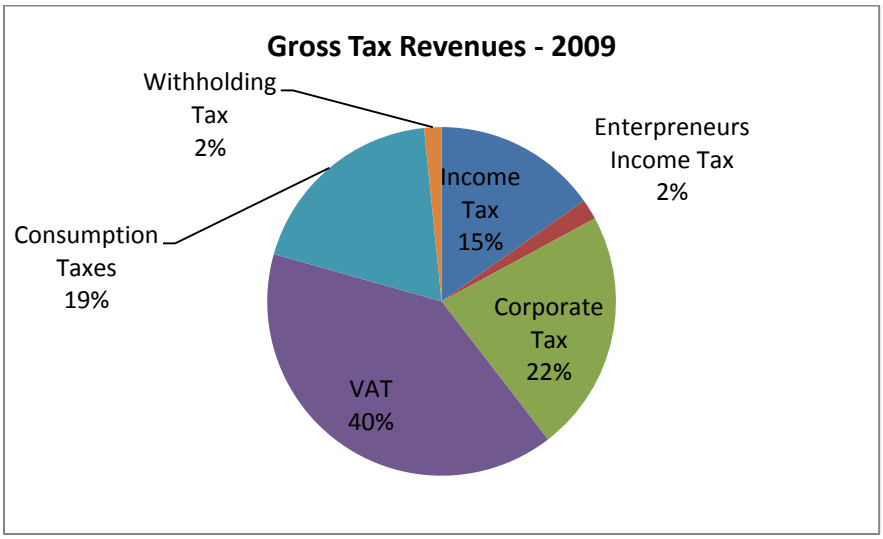

Figure 1. The structure of the Slovak Tax assessment revenues in 2009.
Source: own calculations using data from the Tax Directorate of the Slovak Republic

Table 1. Income tax – gross yield. 1 Euro = 30.126 Slovak Crowns

Year	Act	National budget	Municipal budgets	Gross yield in million	Currency
1995	58/1995	77.47%	22.53%	15 705	Skk
1996	304/1995	76.40%	23.60%	20 336	Skk
1997	386/1996	79.30%	20.70%	25 638	Skk
1998	375/1997	81.23%	18.77%	29 568	Skk
1999	63/1999	81%	19%	31 527	Skk
2000	372/1999	rest	6 440 million	27 320	Skk
2001	472/2000	rest	6 890 million	30 019	Skk
2002	586/2001	rest	7 300 million	31 984	Skk
2003	750/2002	rest	7 801 million	34 980	Skk
2004	598/2003	rest	8 463 million	30 854	Skk
2005	564/2004	6.20%	93.80%	33 743	Skk
2006	564/2004	6.20%	93.80%	35 763	Skk
2007	564/2004	6.20%	93.80%	41 729	Skk
2008	564/2004	6.20%	93.80%	49 400	Skk
2009	564/2004	6.20%	93.80%	1 468	Euro
2010	564/2004	6.20%	93.80%	1 434	Euro

Source: Tax Directorate of the Slovak Republic

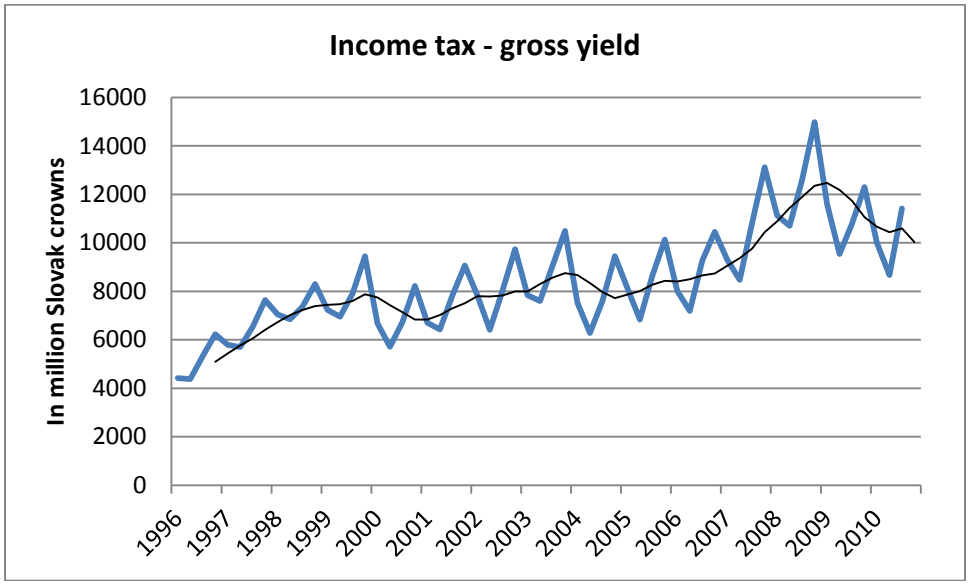

Figure 2. Income tax time series – quarter of the year gross yield
Source: own calculations using data from the Tax Directorate of the Slovak Republic

Figure 2 shows that income tax time series does not represent a stable process. This is the reason why we will use first difference time series. Table 2 shows correlograms of the quarter of the year series and the first difference in the quarter of the year time series.

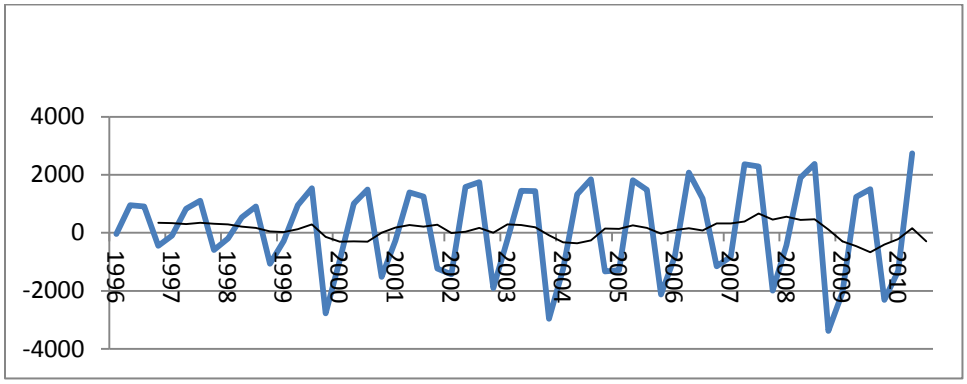

Figure 3. Income tax – first difference
Source: own calculations

Table 2 Correlograms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.698	0.698	30.230	0.000
		2	0.439	-0.093	42.420	0.000
		3	0.587	0.624	64.592	0.000
		4	0.726	0.189	99.122	0.000
		5	0.452	-0.421	112.71	0.000
		6	0.210	-0.129	115.71	0.000
		7	0.354	0.134	124.36	0.000
		8	0.472	0.043	140.10	0.000
		9	0.230	-0.098	143.92	0.000
		10	0.014	-0.109	143.93	0.000
		11	0.162	0.069	145.89	0.000
		12	0.302	0.106	152.85	0.000
		13	0.099	-0.036	153.61	0.000
		14	-0.064	0.050	153.94	0.000
		15	0.098	0.001	154.72	0.000
		16	0.267	0.152	160.70	0.000
		17	0.081	-0.130	161.26	0.000
		18	-0.079	-0.030	161.82	0.000
		19	0.054	-0.130	162.08	0.000
		20	0.186	-0.012	165.29	0.000
		21	0.016	0.006	165.31	0.000
		22	-0.127	0.060	166.87	0.000
		23	-0.018	-0.138	166.90	0.000
		24	0.090	0.012	167.74	0.000

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.031	-0.031	0.0602	0.806
		2	-0.803	-0.805	40.152	0.000
		3	-0.012	-0.235	40.160	0.000
		4	0.827	0.481	84.270	0.000
		5	-0.045	-0.074	84.403	0.000
		6	-0.781	-0.206	125.21	0.000
		7	0.000	-0.133	125.21	0.000
		8	0.736	-0.089	162.94	0.000
		9	-0.018	0.013	162.97	0.000
		10	-0.714	-0.048	199.88	0.000
		11	-0.006	-0.120	199.89	0.000
		12	0.687	-0.019	235.59	0.000
		13	-0.040	-0.216	235.71	0.000
		14	-0.628	0.025	266.88	0.000
		15	0.000	-0.091	266.88	0.000
		16	0.641	0.018	300.97	0.000
		17	-0.038	-0.064	301.09	0.000
		18	-0.568	-0.021	329.14	0.000
		19	-0.005	-0.129	329.14	0.000
		20	0.579	-0.022	359.85	0.000
		21	-0.026	-0.068	359.91	0.000
		22	-0.484	0.180	382.54	0.000
		23	0.006	0.020	382.54	0.000
		24	0.492	-0.009	407.35	0.000

2. Construction of the forecasting models

According to the correlogram we tested the following model formulas :

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) $y = f(c, y_{t-4}, y_{t-6}, \varepsilon_t)$
5) $y = f(c, y_{t-4}, y_{t-6}, y_{t-8}, \varepsilon_t)$

We also added moving averages variables ma(4) and ma(4), ma(6) and ma(4),ma(8) and ma(4), ma(6), ma(8). We also tested model formulas without intercept. We tested 40 model formulas altogether.

We tested models : 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.

According to this methodology, we tested 480 models. The purpose of modelling is the annual forecast for the year 2011. We ran annual ex post forecast for 2008 and then 2009. This means that we forecasted 8 quarters in two steps. We calculated RMS for each model. The whole calculation process was very quick. The old Intel Pentium 4 is an appropriate processor and the whole calculation should not last more than 5 minutes. The source codes were written in Eviews 6.0. The codes can be downloaded from www.mpavlik.net. The models with the smallest RMS were chosen and we calculated another ex post forecast for the quarters of the year 2010. Table 3 shows the top 5 “winning” models with the smallest RMS. Table 4 shows the ex post forecast results for quarters of the years 2008 and 2009 for all 5 “winning” models. Table 5 shows the ex post forecasts of the quarters of the year 2010. The Jarque-Bera test shows whether random errors (residuals) have normal distribution or not. The results can be found in Table 6.

The final step was the ex ante forecast for 2011. The ex ante forecasts are shown in Table 7.

Table 3. Results

Nr.	Model formula	Model	RMS
1	$y = f(y_{t-4}, y_{t-6}, \varepsilon_t, \varepsilon_{t-4}, \varepsilon_{t-8})$	EGARCH(1,1),asy=2	1 708
2	$y = f(c, y_{t-4}, y_{t-6}, \varepsilon_t, \varepsilon_{t-4}, \varepsilon_{t-6})$	EGARCH(1,0),asy=2	1 710
3	$y = f(c, y_{t-4}, y_{t-6}, \varepsilon_t, \varepsilon_{t-4})$	EGARCH(1,1),asy=3	1 735
4	$y = f(y_{t-4}, y_{t-6}, y_{t-8}, \varepsilon_t, \varepsilon_{t-4})$	EGARCH(1,0),asy=2	1 759
5	$y = f(y_{t-4}, y_{t-6}, \varepsilon_t, \varepsilon_{t-4})$	EGARCH(1,1),asy=2	1 773

Source: own calculations

Table 4. Ex post forecasts in million Slovak crowns

Year	Quarters	Reality	Model 1	Model 2	Model 3	Model 4	Model 5	for./reality Model 1
2008	1	11 126	11 074	11 250	10 899	11 083	11 255	1.00
	2	10 697	9 734	9 927	9 641	10 199	9 949	0.91
	3	12 600	11 841	12 960	11 889	12 204	12 028	0.94
	4	14 977	13 505	15 071	13 897	14 200	13 615	0.90
2009	1	11 600	12 607	12 878	12 510	12 839	12 846	1.09
	2	9 543	11 208	11 347	11 237	11 621	11 521	1.17
	3	10 788	13 424	13 614	13 449	13 653	13 619	1.24
	4	12 297	15 284	15 423	15 533	15 411	15 285	1.24
2008	annual	49 400	46 154	49 208	46 325	47 686	46 846	0.93
2009	annual	44 228	52 523	53 262	52 730	53 523	53 270	1.19

Source: own calculations

Table 5. Ex post forecasts of the quarters of the year 2010 in million Slovak crowns

Year	Quarters	Reality	Model 1	Model 2	Model 3	Model 4	Model 5
2010	1	9 991	9 955	9 674	9 605	10 045	9 846
	2	8 673	8 471	7 787	8 210	8 622	8 203
	3	11 409	10 837	9 849	10 556	11 009	10 407
	4	13 119	12 767	11 656	12 875	12 912	12 179
2010	annual	43 193	42 028	38 966	41 246	42 588	40 635
			for./reality	for./reality	for./reality	for./reality	for./reality
2010	1	1	1.00	0.97	0.96	1.01	0.99
	2	1	0.98	0.9	0.95	0.99	0.95
	3	1	0.95	0.86	0.93	0.96	0.91
	4	1	0.97	0.88	0.98	0.98	0.93
2010	annual	x	0.97	0.90	0.95	0.98	0.94

Source: own calculations

Table 6. Serial correlation and Jarque-Bera test

	Model 1	Model 2	Model 3	Model 4	Model 5
JB test	5.28	0.82	1.73	3.92	4.92
Probability	0.07	0.66	0.42	0.14	0.08
Serial correlation	No	Yes	No	No	No

Source: own calculations

Table 7. Ex ante forecast

		in million Skk		*in million Euros*	
	quarters	*Model 1*	*Model2*	*Model 1*	*Model 2*
2011 ex ante	1	10 626	10 821	353	359
	2	8 938	9 261	297	307
	3	11 472	11 934	381	396
	4	13 496	13 745	448	456
	annual	44 532	45 760	1 478	1 519

Source: own calculations

Table 6 also shows that there is almost no serial correlation in the winning models. The presence of the serial correlation was measured with the Q statistics. The values of the Q statistics were calculated according to the Ljung-Box test of the serial correlation, which is a modified version of the Box-Pierce test. The Ljung-Box test is defined as :

$$Q = n(n+2)\sum_{i=1}^{n}\frac{\hat{\rho}_i^2}{n-i}, \text{ where } \rho_i = \frac{\operatorname{cov}(e_t, e_{t-1})}{\sqrt{\operatorname{var}(e_t)}.\sqrt{\operatorname{var}(e_t - 1)}} \quad (1)$$

for i=1,2, …, p

Q statistics has χ^2 distribution with p degrees of freedom. [Vogelvang 2005, pp. 121-123]

3. Conclusion

Forecasting of tax revenues is not just the area of the Ministry of Finance or the Tax Headquarters, but it is very useful also in municipal scope, because it influences the municipal budgets very much. We focused on monthly basis forecasts in a previous article. It was shown that good forecasts can be achieved without using explanatory variables recommended by the

economic theory in the previous article. Those variables are not observed on a monthly basis. This time we focused on quarter of the year basis and although explanatory variables on the quarter of the year basis exist, we are trying to show that they are not necessary.

This methodology is useful in the Slovak Republic, because of the short time series on an annual basis, but it might be also useful in Poland, which experienced a communist system, and the time series from those times might not to be very suitable. We realized that quarter of the year basis is appropriate for annual forecasts because the time series contains approximately 60 observations. Reviewing the article is a process which takes some time and we already know 1q2011 results. The income tax revenue was 374 million euro for the 1q2011 which is 94% forecast.

References

ARLT J., ARLTOVÁ M. : *Finanční časové řady,* Praha, Grada, 2003
ARLT J., ARLTOVÁ M. : *Ekonomické časové řady,* Praha,Grada, 2006
DAVIDSON R., MACKINNON J. G. : *Estimation and Inference in Econometrics, New York*, Oxford University Press, 1993
GRANGER C. W. J. : *Forecasting in business and economics,* San Diego, Academic Press, 1989
HAMILTON J.D. : *Time series analysis,* Princeton, Princeton, 1994
HATRÁK M. : *Ekonometrické metódy I,* Bratislava, Ekonóm, 1995
HATRÁK M. : *Ekonometrické metódy II,* Bratislava, Ekonóm, 1995
HAYASHI F. : *Econometrics,* Princeton, Princeton University Press, 2000
KRÁLOVÁ J., NOVYSEDLÁK V. : *Prognóza daňových príjmov verejnej správy na roky 2010-2013 (jún 2010)*, www.finance.gov.sk
LUKÁČIKOVÁ A., LUKÁČIK M. : *Ekonometrické modelovanie s aplikáciami*, Bratislava, EKONÓM, 2008
PAVLÍK M. : *The Usage of Dummy Variable for VAT forecasting of the Tax Administration in the Slovak Republic,* Wroclaw, Ekonometria 21, 2008, p. 40-54
SURMANOVÁ K.: E*konometrické prognózovanie vývoja miezd vo vybraných krajinách EÚ*, Medzinárodný seminár mladých vedeckých pracovníkov Katedry ekonometrie FIS VŠE v Praze a Katedry operačného výskumu a ekonometrie FHI EU v Bratislave zborník : Praha 29.-30. november 2007., Bratislava, EKONÓM, 2008.
VOGELVANG B. : *Econometrics Theory and Applications with Eviews,* Financial Times Prentice Hall, 2005
www.mpavlik.net