PRACE NAUKOWE UNIWERSYTETU EKONOMICZNEGO WE WROCŁAWIU nr 84
Ekonometria 27 2009

Martin Pavlík
Tax Office of the Slovak Republic

NEW APPROACH TO THE SHORT-TERM FORECASTING IN THE TIME SERIES WITH THE STRUCTURAL SHOCK

Summary: The aim is to show a new approach in the short-term forecasting in case the time series contain structural shock. Structural shock appeared in the Slovak VAT time series. The author runs ex post forecasts based on the data from January 1996-December 2006. He explored thousands of models including the transformation and compared the new research with previous research already published and logarithmised time series using logarithms with a lot of different bases. He explored bases: 5, 10, 15, 20, 25 and natural logarithm, compared the quality of the ex post forecasts and discovered that decimal logarithm was the best choice for this particular time series, the increasing number of iterations could slightly improve the forecasts and transformation was not convergent and a lot of calculation was without indicating convergence even after 250 iterations.

Key words: short-term forecasting, structural shock, logarithmic function, transformation.

1. Introduction

The need for the basic research in the structural shock modelling appeared after the tax reform and the acceptance of the Slovak Republic for the membership of the European Union in 2004. The tax reform influenced tax revenues very much, especially VAT for the tax administration. We developed something what we call "transformation". The transformation precises the forecasts in case of the structural shock in the time the series appears. The aim of the article is to explore the transformation and the quality of the forecasts. The article is in fact part 2 of [Pavlík 2008].

2. Transformation – formula and the convergence

The formulas are:

$$y_t = \beta_0 + \beta_1 y_{t-12} + \gamma x_t + \varepsilon_t$$

x_t	0	before 7/2004
	1	7/2004 6/2005
	$1-\beta_1$	after 7/2005

$$y_t = \beta_0 + \beta_1 y_{t-12} + \beta_2 y_{t-24} + \gamma x_t + \varepsilon_t$$

x_t	0	before 7/2004
	1	7/2004 6/2005
	$1-\beta_1$	7/2005 6/2006
	$1-\beta_1-\beta_2$	after 7/2006

$$y_t = \beta_0 + \beta_1 y_{t-3} + \beta_2 y_{t-12} + \beta_3 y_{t-24} + \gamma x_t + \varepsilon_t$$

x_t	0	before 7/2004
	1	7/2004 9/2004
	$1-\beta_1$	10/2004 6/2005
	$1-\beta_1-\beta_2$	7/2005 6/2006
	$1-\beta_1-\beta_2-\beta_3$	after 7/2006

Structural shock appeared in July 2004. Observations x_t time series depends on the output of the calculation‘s process (β_1, β_2, β_3). This is the reason why the calculation of the x_t observations is an iterative process. Every iterative process needs a convergence criteria. We set the convergence criteria:

$$\left(\left|\beta_{1_t} - \beta_{1_{t-1}}\right| < 0{,}0001\right) \wedge \left(\left|\beta_{2_t} - \beta_{2_{t-1}}\right| < 0{,}0001\right) \wedge \left(\left|\beta_{3_t} - \beta_{3_{t-1}}\right| < 0{,}0001\right).$$

Maximum iterations count was 250. We added 1 to 3 explanatory variables to the 30 fundamental model formulas. We used ARIMA models, ARCH(1), ARCH(2), GARCH(1,1), GARCH(2,1) and GARCH(2,2). We applied transformation developed by us in two different ways:

1) we applied ARCH and GARCH models from the beginning (transf),

2) we applied LS, until convergence was achieved and ARCH or GARCH models were applied just in the final iteration (transf2).

We also varied the iteration process. The process called BEFORE means that every single iteration process starts from a starting point where x_t variable contains just 0 and 1. The process called AFTER means that ex post forecasts for every model starts from x_t variable where x_t contains just 0 and 1 just for January 2005. The rest of the ex post forecasts (February 2005-December 2006) start from the previous months convergence.

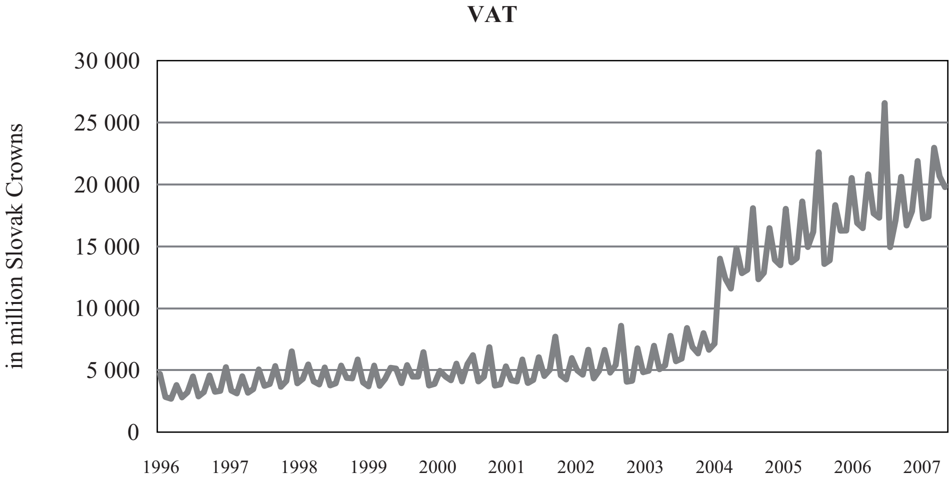

Fig. 1. Specific values of VAT for the Slovak tax administration

Source: own calculations using data from the Tax Directorate of the Slovak Republic.

The transformation covers the structural shock. The explored time series contains a structural shock which is shown in fig. 1.

3. Model formulas

We used 30 fundamental model formulas:

1: $y_t = \beta_0 + \beta_1 y_{t-3} + \varepsilon_t$

2: $y_t = \beta_0 + \beta_1 y_{t-12} + \varepsilon_t$

3: $y_t = \beta_0 + \beta_1 y_{t-3} + \beta_2 y_{t-12} + \varepsilon_t$

4: $y_t = \beta_0 + \beta_1 y_{t-3} + \beta_2 y_{t-12} + \beta_3 y_{t-24} + \varepsilon_t$

5: $y_t = \beta_0 + \beta_1 y_{t-12} + \beta_2 y_{t-24} + \varepsilon_t$

6: $y_t = \beta_0 + \beta_1 y_{t-3} + \varepsilon_t + \theta_1 \varepsilon_{t-12}$

7: $y_t = \beta_0 + \beta_1 y_{t-12} + \varepsilon_t + \theta_1 \varepsilon_{t-12}$

8: $y_t = \beta_0 + \beta_1 y_{t-3} + \beta_2 y_{t-12} + \varepsilon_t + \theta_1 \varepsilon_{t-12}$

9: $y_t = \beta_0 + \beta_1 y_{t-3} + \beta_2 y_{t-12} + \beta_3 y_{t-24} + \varepsilon_t + \theta_1 \varepsilon_{t-12}$

10: $y_t = \beta_0 + \beta_1 y_{t-12} + \beta_2 y_{t-24} + \varepsilon_t + \theta_1 \varepsilon_{t-12}$

11-20: The same as 1-10, but formulas are without β0

21: $y_t = \beta_0 + \beta_1 y_{t-3} + \varepsilon_t + \theta_1 \varepsilon_{t-12} + \theta_2 \varepsilon_{t-24}$

22: $y_t = \beta_0 + \beta_1 y_{t-12} + \varepsilon_t + \theta_1 \varepsilon_{t-12} + \theta_2 \varepsilon_{t-24}$

23: $y_t = \beta_0 + \beta_1 y_{t-3} + \beta_2 y_{t-12} + \varepsilon_t + \theta_1 \varepsilon_{t-12} + \theta_2 \varepsilon_{t-24}$

24: $y_t = \beta_0 + \beta_1 y_{t-3} + \beta_2 y_{t-12} + \beta_3 y_{t-24} + \varepsilon_t + \theta_1 \varepsilon_{t-12} + \theta_2 \varepsilon_{t-24}$

25: $y_t = \beta_0 + \beta_1 y_{t-12} + \beta_2 y_{t-24} + \varepsilon_t + \theta_1 \varepsilon_{t-12} + \theta_2 \varepsilon_{t-24}$
25-30: The same as 21-25, but formulas are without β_0.

Model formulas contained 1 to 3 additional x_t variables. The variables covered structural shock, the influence of the tax rate and seasonal tax payers. Both tax rate variable and seasonal tax rate variable were covered in two different ways. The calculations were extremely demanding and took couple of months at Pentium D 3,4 GHz. We ran ex post forecast from January 2005 to December 2006 in a way forecasts are made in practice. It means that when we made a forecast for January 2005, we used time series just for November 2004. The December 2004 is unavailable when forecast for January 2005 is made.

4. Measuring results

We also set the appropriate measure for measuring the quality of the forecasts. We used two measures: root mean square (RMS) and mean simulation error (MSE)

$$RMS = \sqrt{\left(\frac{1}{s}\sum_{t=1}^{s}\left(y_t^s - y_t\right)^2\right)},$$

$$MSE = \frac{1}{s}\sum_{t=1}^{s}\left(y_t^s - y_t\right).$$

We used MSE as a supplementary measure. We were inspired by two necessary properties of estimators: unbiasedness and efficiency.

- Unbiasedness:

$$E(\hat{\beta}) = \beta.$$

- Efficiency:

$$E(\hat{\theta} - E(\hat{\theta}))^2 \leq E(\tilde{\theta} - E(\tilde{\theta}))^2.$$

We suppose that $E(\hat{\beta})$ is a set from many $\hat{\beta}$, where $\hat{\beta}$ are a set from many selections of data sets

MSE is:

$$MSE = \frac{1}{s}\sum_{t=1}^{s}(y_t^p - y_t).$$

After the substitution $o_t = (y_t^p - y_t)$, we get a formula for the arithmetic mean:

$$MSE = \frac{1}{s}\sum_{t=1}^{s}(o_t).$$

We see here the analogy with the unbiased estimator. The difference is that $E(\hat{\beta})$ is β, whereas the desired MSE value is 0.

The RMS formula is:

$$\text{RMS} = \sqrt{\left(\frac{1}{s}\sum_{t=1}^{s}\left(y_t^p - y_t\right)^2\right)}.$$

We here see the analogy with efficiency. RMS formula and formula for efficiency are in fact the formulas for standard deviation. Both of them ought to be close zero.

The different view on RMS and MSE is:

Let us describe two models which cover the same fact. We ran an ex post forecast with following residuals.

Table 1.

Residuals	Model 1	–1	–1	–1	–1	–1	–1	Σ = –6
	Model 2	–1	1	–1	1	–1	1	Σ = 0

Source: own calculations.

Model 1 underestimates reality, model 2 oscillates. MSE of the first model is –1; MSE of the second model is 0. Both models have RMS = 1. It is obvious that model 2 has better forecasting results than model 1 and MSE shows it.

5. Results of the usage the different logarithm basis

We measured results with RMS and MSE. The results for the BEFORE process are shown in table 2 and 3. The tables show how quickly the convergence was achieved and they also show RMS min and MSE. RMS and MSE were calculated from ex post forecasts (January 2005 to December 2006).

Table 2. Before, transf process – the results

	Log basis	Iterations 1-49	50-99	100-149	150-199	200-249	250	RMS min	MSE
1	2	3	4	5	6	7	8	9	10
1 variable	5	3 073	70	47	23	13	1 094	1 119	–441
	15	3 056	60	50	25	14	1 115	1 119	–443

1	2	3	4	5	6	7	8	9	10
	20	3 045	63	34	34	24	1 120	1 119	–443
	25	3 031	81	42	36	26	1 104	1 119	–443
	ln	3 036	72	37	5	19	1 131	1 119	–441
	raw	2 708	137	63	53	20	1 339	1 353	24
	10	3 067	68	26	20	19	1 120	1 108	–523
2 variables	5	11 782	335	181	117	106	4 759	844	–127
	15	11 709	341	191	145	106	4 788	793	–108
	20	11 764	296	184	144	88	4 804	843	–95
	25	11 767	337	165	117	103	4 791	809	–124
	ln	11 815	353	186	122	97	4 707	838	–367
	raw	10 630	669	329	243	162	5 247	1 065	–219
	10	11 697	312	211	135	104	4 821	823	–147
3 variables	5	11 288	361	218	144	97	5 172	866	12
	15	11 284	362	189	156	104	5 185	913	–122
	20	11 395	377	176	153	105	5 074	863	–385
	25	11 476	322	198	127	112	5 045	906	–54
	ln	11 400	384	219	146	116	5 015	870	–27
	raw	10 336	782	352	244	191	5 375	1 074	–229
	10	11 304	375	202	142	107	5 150	875	–368

Source: own calculations.

Table 3. Before, transf2 process – the results

	Log basis	Iterations 1-49	50-99	100-149	150-199	200-249	250	RMS min	MSE
1 variable	5	3 455	5	0	0	5	135	1 049	–103
	15	3 455	10	5	0	0	130	1 050	–107
	20	3 450	5	0	0	5	140	1 050	–108
	25	3 450	10	5	5	0	130	1 061	–111
	ln	3 450	10	0	0	0	140	1 008	–633
	raw	3 385	35	10	10	0	160	1 267	–306
	10	3 455	10	0	0	0	135	1 015	–520
2 variables	5	13 520	25	5	10	15	825	779	–213
	15	13 515	30	10	15	15	815	778	–226
	20	13 535	20	10	0	10	825	773	–223
	25	13 495	45	15	15	15	815	763	–211
	ln	13 525	30	25	5	15	800	785	–375
	raw	13 545	120	25	35	20	655	963	–106
	10	13 515	30	25	15	25	790	756	–238
3 variables	5	13 195	75	25	5	5	1 095	899	–279
	15	13 215	45	40	15	20	1 065	871	–143
	20	13 215	70	5	20	30	1 060	883	–52
	25	13 205	75	25	0	10	1 085	893	0
	ln	13 180	45	25	20	15	1 115	869	–293
	raw	13 965	95	27	12	10	291	991	–60
	10	13 170	55	35	15	15	1 110	893	–134

Source: own calculations.

Tables 2 and 3 show that the best choice for the before process is the decadic logarithm and a model with two additional x_t variables with RMS = 756 and MSE = –238. The best model is the model where structural shock was modeled in a transf2 way.

The results for the AFTER process are shown in table 4 and 5. The tables show how quickly the convergence was achieved and they also show RMS min and MSE.

Table 4. After, transf process – the results

	Log basis	Iterations 1-49	50-99	100-149	150-199	200-249	250	RMS min	MSE
1 variable	5	3 094	86	29	31	23	1 057	1 119	–441
	15	3 058	83	41	20	19	1 099	1 118	–444
	20	3 032	91	37	31	27	1 102	1 118	–444
	25	3 048	90	48	32	22	1 080	1 119	–443
	Ln	3 076	83	40	26	16	1 079	1 119	–441
	Raw	2 665	183	78	42	34	1 318	1 368	–8
	10	3 063	79	48	19	15	1 096	1 118	–443
2 variables	5	11 750	383	219	147	90	4 691	874	–323
	15	11 660	394	198	133	94	4 801	837	–322
	20	11 695	368	196	139	88	4 794	811	–166
	25	11 713	387	184	125	89	4 782	823	–106
	Ln	11 789	447	180	138	107	4 619	822	–378
	Raw	10 463	777	359	238	172	5 271	1 065	–219
	10	11 682	405	177	126	105	4 785	862	–399
3 variables	5	11 208	442	225	162	110	5 133	893	–421
	15	11 210	460	208	131	106	5 165	897	–92
	20	11 260	461	228	155	110	5 066	854	–71
	25	11 390	423	197	129	93	5 048	910	–68
	Ln	11 354	478	234	149	104	4 961	866	5
	Raw	10 210	826	404	274	198	5 368	1 095	–176
	10	11 265	457	247	157	101	5 053	808	–59

Source: own calculations.

Table 5. After, transf2 process – the results

	Log basis	Iterations 1-49	50-99	100-149	150-199	200-249	250	RMS min	MSE
1 variable	5	3 435	20	10	5	5	125	1 049	–104
	15	3 435	15	10	0	5	135	1 049	–107
	20	3 435	15	10	0	5	135	1 050	–107
	25	3 420	25	20	10	5	120	1 060	–111
	ln	3 430	20	10	5	0	135	1 052	–123
	Raw	3 335	105	5	0	5	150	1 267	–363
	10	3 425	35	15	0	0	125	1 023	–646
2 variables	5	13 475	105	20	15	5	780	780	–207
	15	13 430	90	45	15	20	800	778	–226
	20	13 430	90	30	15	10	825	771	–133
	25	13 460	70	35	15	15	805	818	–172
	ln	13 495	80	10	30	5	780	783	–398
	Raw	13 395	295	30	20	40	620	963	–106
	10	13 460	80	30	5	10	815	757	–145
3 variables	5	13 100	105	25	20	30	1 120	921	–328
	15	13 110	115	40	30	10	1 095	911	–38
	20	13 115	110	20	30	10	1 115	894	–472
	25	13 110	115	40	20	25	1 090	942	84
	ln	13 100	95	25	10	25	1 145	952	–406
	Raw	13 415	230	45	50	50	610	986	–51
	10	13 090	110	35	25	5	1 135	917	–421

Source: own calculations.

Tables 4 and 5 show that the best choice for the after process is the decadic logarithm and a model with two additional x_t variables with RMS = 756 and MSE = –145. The best model is the model where structural shock was modeled in a transf2 way.

Tables 6-7 contain the best models according to RMS min.

Table 6. Before process, RMS min

	Log basis	Additional variables	RMS	MSE	Model formula	Model
transf2	10	2	756	–238	25	ARCH(2)
transf2	25	2	763	–211	25	ARCH(2)
transf2	20	2	773	–223	25	ARCH(2)
transf2	15	2	778	–226	25	ARCH(2)
transf2	5	2	779	–213	25	ARCH(2)
transf2	Ln	2	785	–375	10	GARCH(1,1)

Source: own calculations.

Table 7. After process, RMS min

	Log basis	Additional variables	Rms min	MSE	Model formula	Model
transf2	10	2	757	–145	10	ARCH(1)
transf2	20	2	771	–133	10	ARCH(1)
transf2	15	2	778	–226	25	ARCH(2)
transf2	5	2	780	–207	25	ARCH(2)
transf2	Ln	2	783	–398	10	GARCH(1,1)

Source: own calculations.

6. Ex post forecasts

The “winning” models from tables 6-7 are:

- Before process:

$$y_t = \beta_0 + \beta_1 y_{t-12} + \beta_2 y_{t-24} + \gamma_1 x_{t1} + \gamma_2 x_{t2} + \varepsilon_t + \theta_1 \varepsilon_{t-12} + \theta_2 \varepsilon_{t-24}$$

$$\varepsilon_t = \xi + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + w_t,$$

where: x_{t1} – describes the structural shock, x_{t2} – describes the influence of the tax rate.

- After process:

$$y_t = \beta_0 + \beta_1 y_{t-12} + \beta_2 y_{t-24} + \gamma_1 x_{t1} + \gamma_2 x_{t2} + \varepsilon_t + \theta_1 \varepsilon_{t-12}$$

$$\varepsilon_t = \xi + \alpha_1 \varepsilon_{t-1}^2 + w_t ,$$

where: x_{t1} − describes the structural shock, x_{t2} − describes the influence of the tax rate.

The "winning" models should generate the best forecasts. We tested them for the *ex post* 2007 forecast. And what is more, we tested all the "winning" models from table 6 and table 7 and set *ex post* 2007 forecasts. The results are shown in tables 8 and 9.

Table 8. Before process – *ex post* 2007 forecast (in million SKK)

2007	10	25	20	15	5	Ln	Reality
1	25 055	25 048	25 048	25 054	25 088	24 709	26 568
2	15 583	15 585	15 585	15 583	15 605	15 749	14 932
3	15 788	15 789	15 789	15 788	15 810	15 951	17 108
4	20 870	20 873	20 872	20 871	20 901	21 114	20 604
5	18 173	18 167	18 169	18 171	18 195	18 199	16 682
6	18 249	18 243	18 245	18 247	18 271	18 227	17 831
7	23 026	23 024	23 023	23 026	23 057	22 949	21 893
8	18 850	18 843	18 845	18 848	18 872	18 607	17 248
9	18 689	18 688	18 688	18 688	18 713	18 699	17 381
10	23 712	23 717	23 715	23 715	23 748	23 911	22 961
11	19 734	19 728	19 730	19 732	19 758	19 615	20 663
12	19 798	19 802	19 801	19 800	19 827	20 081	19 787
Geometrical average	1.019	0.982	1.019	1.019	1.020	1.020	
RMS	1 077	1 076	1 077	1 076	1 084	1 117	

Source: own calculations.

Geometrical average is the geometrical average of the values $\frac{forecast_i}{reality_i}$.

Table 9. After process – *ex post* 2007 forecast (in million SKK)

2007	10	20	15	5	Ln	Reality
1	24 153	24 146	25 054	25 088	24 709	26 568
2	15 493	15 491	15 583	15 605	15 749	14 932
3	15 624	15 623	15 788	15 810	15 951	17 108
4	20 772	20 768	20 871	20 901	21 114	20 604
5	17 912	17 910	18 171	18 195	18 199	16 682
6	17 958	17 957	18 247	18 271	18 227	17 831
7	22 293	22 287	23 026	23 057	22 949	21 893
8	18 379	18 377	18 848	18 872	18 607	17 248
9	18 240	18 238	18 688	18 713	18 699	17 381
10	22 922	22 916	23 715	23 748	23 911	22 961
11	19 216	19 213	19 732	19 758	19 615	20 663
12	19 284	19 280	19 800	19 827	20 081	19 787
Geometrical average	0.997	0.997	1.019	1.020	1.020	
RMS	1 097	1 098	1 076	1 084	1 117	

Source: own calculations.

7. The convergation of the transformation

Another task is to explore how the transformation converges. Maximum iterations count was set to 250 for every single calculation. If the convergation was not achieved within 250 iteration, the calculations stopped. We explored just “the winning” decadic logarithm models and raw data models.

We ran *ex post* forecast from January 2005-December 2006. Figure 2 shows how many models achieved convergence within borders on the x-axis. Figures 2 and 3 show that convergence is usually achieved within 50 iterations and if the convergence is not achieved within 50 iterations, than it is mostly not achieved within 250 iterations. Only few models achieved converegence within 50-249 iterations.

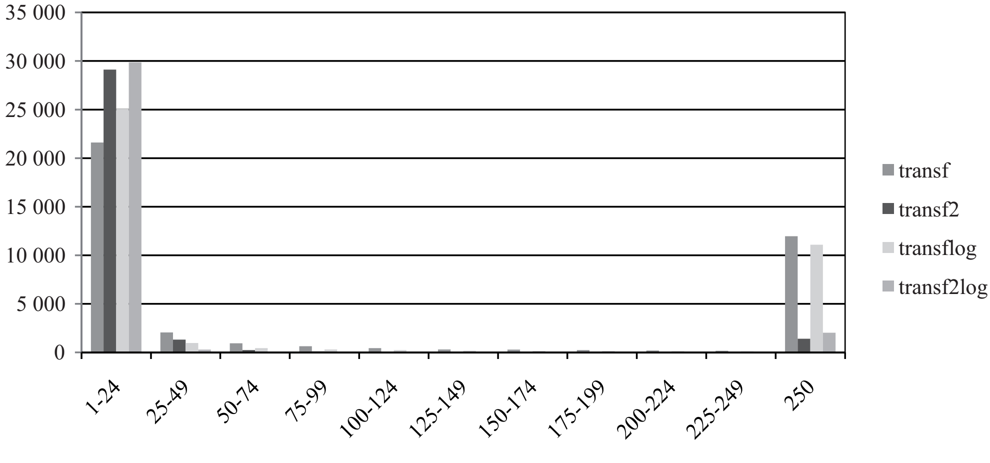

Fig. 2. BEFORE − together

Source: own calculations.

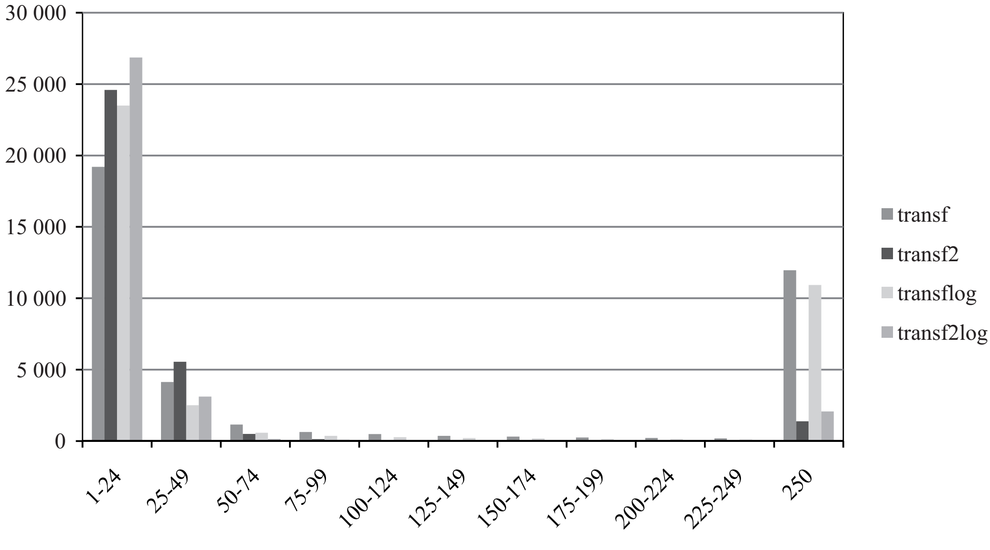

Fig. 3. AFTER − together

Source: own calculations.

BEFORE iteration process converges quicker than AFTER, but AFTER process leaves less calculations without achieving convergence (250 iterations column). It is 26 492 for BEFORE process and 26 346 for AFTER process. Tables 10 and 11 show exactly the same what fig. 2 and 3 show.

Table 10. BEFORE iterations process – specific figures

		Iterations count										
	Before	1-24	25-49	50-74	75-99	100-124	125-149	150-174	175-199	200-224	225-249	250
Together	Transf	21 613	2061	943	645	442	302	289	251	207	166	11 961
	transf2	29 120	1320	245	110	65	25	30	40	30	10	1 405
	transflog	25 091	977	450	305	249	190	155	142	119	111	11 091
	transf2log	29 850	290	65	30	35	25	20	10	30	10	2 035
1. variable	Transf	2 500	208	91	46	43	20	25	28	10	10	1 339
	transf2	3 195	190	15	20	10	0	5	5	0	0	160
	transflog	2 964	103	30	38	14	12	7	13	10	9	1 120
	transf2log	3435	20	10	0	0	0	0	0	0	0	135
2. variables	Transf	9 718	912	393	276	191	138	129	114	91	71	5 247
	transf2	12 875	670	80	40	25	0	15	20	15	5	655
	transflog	11 287	410	189	123	124	87	75	60	67	37	4 821
	transf2log	13 395	120	20	10	10	15	10	5	15	10	790
3. variables	Transf	9 395	941	459	323	208	144	135	109	106	85	5 375
	transf2	13 050	460	150	50	30	25	10	15	15	5	590
	transflog	10 840	464	231	144	111	91	73	69	42	65	5 150
	transf2log	13 020	150	35	20	25	10	10	5	15	0	1 110

Source: own calculations.

Table 11. AFTER iterations process – specific figures

		Iterations count										
	After	1-24	25-49	50-74	75-99	100-124	125-149	150-174	175-199	200-224	225-249	250
1	2	3	4	5	6	7	8	9	10	11	12	13
Together	Transf	19 206	4 132	1 150	636	483	358	305	249	217	187	11 957
	transf2	24 590	5 555	495	135	40	40	50	20	45	50	1 380
	transflog	23 502	2 508	577	364	268	204	169	133	118	103	10 934
	transf2log	26 865	3 110	140	85	50	30	10	20	15	0	2 075
1. variable	Transf	2 211	454	120	63	40	38	22	20	18	16	1 318
	transf2	2 585	750	80	25	5	0	0	0	0	5	150
	transflog	2 749	314	49	30	29	19	6	13	6	9	1 096
	transf2log	2 930	495	25	10	15	0	0	0	0	0	125

1	2	3	4	5	6	7	8	9	10	11	12	13
2. variables	Transf	8 623	1 840	511	266	207	152	135	103	91	81	5 271
	transf2	10 675	2 720	240	55	10	20	15	5	20	20	620
	transflog	10 568	1 114	242	163	97	80	72	54	62	43	4 785
	transf2log	11 945	1 515	30	50	15	15	5	0	10	0	815
3. 3variables	Transf	8 372	1 838	519	307	236	168	148	126	108	90	5 368
	transf2	11 330	2 085	175	55	25	20	35	15	25	25	610
	transflog	10 185	1 080	286	171	142	105	91	66	50	51	5 053
	transf2log	11 990	1 100	85	25	20	15	5	20	5	0	1 135

Source: own calculations.

Table 12. Models which did not achieve convergence (in %)

	Variables	in %	ARIMA	ARCH(1)	GARCH(1,1)	ARCH(2)	GARCH(2,1)	GARCH(2,2)
BEFORE	1	Transf	4,4	12,1	40,3	40,6	44,6	44,0
		transflog	3,8	18,9	33,3	28,5	37,4	33,8
	2	Transf	4,5	14,5	38,2	38,2	43,7	43,0
		transflog	5,5	19,3	36,6	30,3	40,2	35,5
	3	Transf	4,1	15,4	42,4	40,3	43,9	40,6
		transflog	7,7	18,6	39,3	34,0	40,7	38,5
AFTER	1	Transf	4,2	12,2	40,0	40,0	44,9	41,8
		transflog	3,5	17,2	32,5	26,8	36,9	35,3
	2	Transf	4,3	14,5	38,6	38,7	43,7	43,2
		transflog	5,7	19,4	36,8	30,0	38,5	35,7
	3	Transf	4,2	15,1	42,0	41,0	43,6	40,4
		transflog	7,9	18,5	37,5	33,4	40,9	37,2

Source: own calculations.

Another interesting question is the structure of the unconvergent iterations according to the sort of model (ARIMA, ARCH, GARCH). Table 12 gives information about the rate of models which did not achieve convergence according to the class of models. It can be seen that ARIMA models mostly achieved convergence, but ARCH models and GARCH models did not. ARCH(1) models achieved the best convergence results from ARCH and GARCH models class.

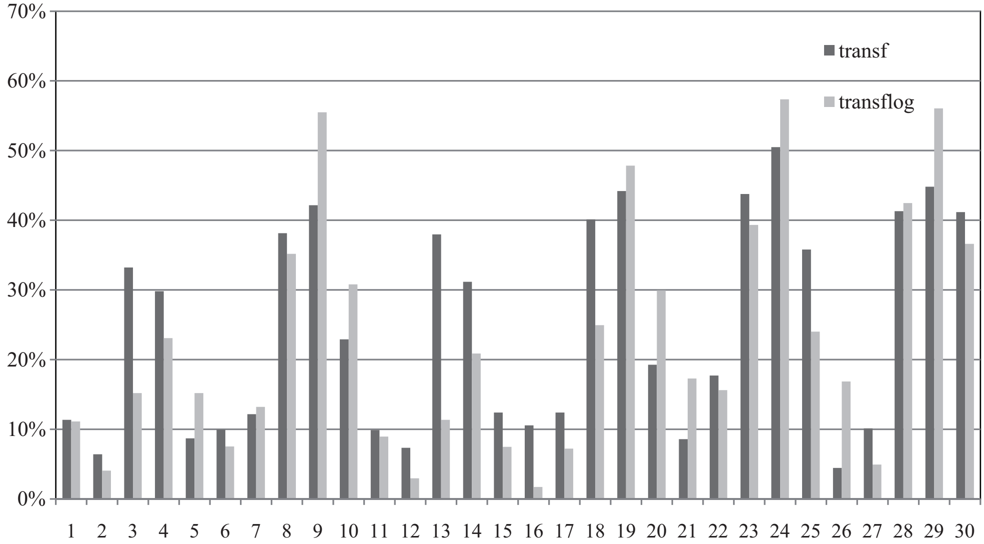

Fig. 4. Model formulas, which did not achieve convergence – BEFORE process

Source: own calculations.

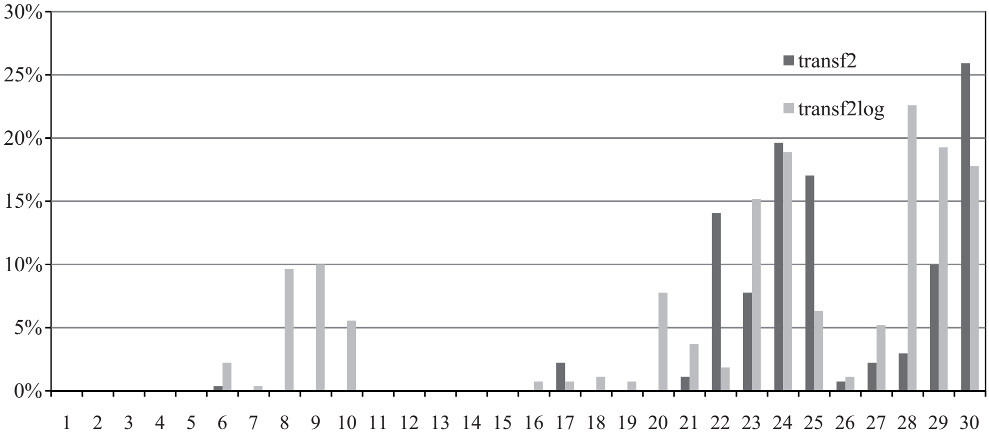

Fig. 5. Model formulas which did not achieve convergence – BEFORE process

Source: own calculations.

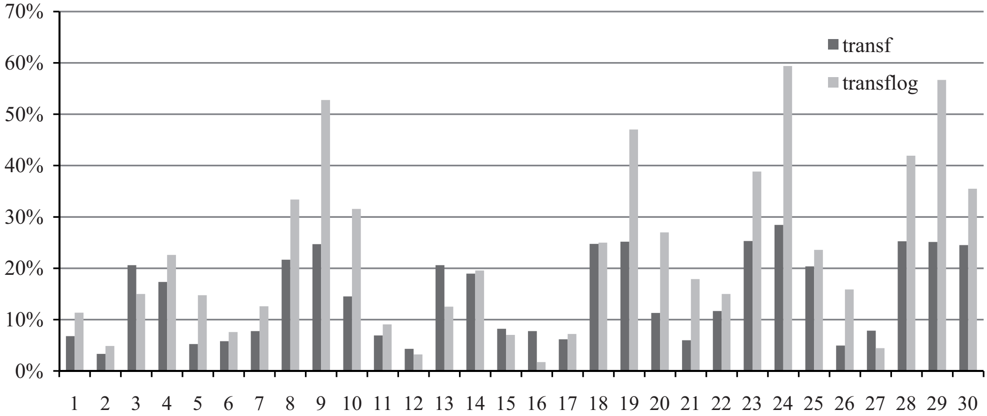

Fig. 6. Model formulas which did not achieve convergence – AFTER process

Source: own calculations.

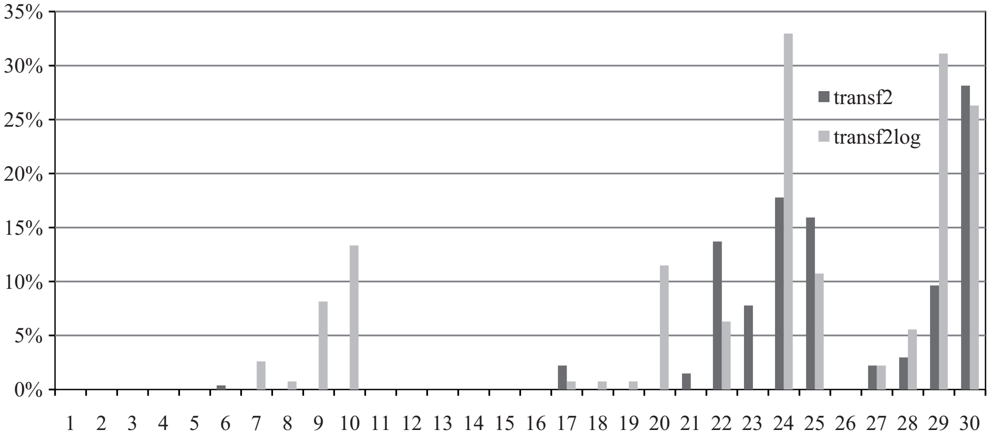

Fig. 7. Model formulas, which did not achieve convergence – AFTER process

Source: own calculations.

Another interesting question is which fundamental model formulas achieved convergence and which model formulas did not. This is what show fig. 4, 5, 6, 7.

8. Conclusion

The research shows that it is better to use different logarithms than natural logarithm. Because we used just one time series we cannot generalize the results. The paper shows that statistical researchers should also try different logarithms not just natural logarithm and the usage of different logarithm may result in significant improvement in the forecasts. The presented results could start a discussion which could lead into the progression in the econometric/statistic modelling.

We show that the transformation does not converge easily and the large amount of the calculations was left without achieving the convergence. The solution could be the rearrangement of the convergence criteria and the reduction of the number of lagged y_t.

References

Arlt J., Arltová M., *Finanční časové řady*, Grada, Praha 2003.
Arlt J., Arltová M., *Ekonomické časové řady*, Grada, Praha 2006.
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, Princeton 1994.
Hatrák M., *Ekonometrické metódy I*, Ekonóm, Bratislava 1995.
Hatrák M., *Ekonometrické metódy II*, Ekonóm, Bratislava 1995.
Hayashi F., *Econometrics*, Princeton University Press, Princeton 2000.
Krajčír Z., Králová J., Livermore S., *Prognózovanie dane z pridanej hodnoty*, www.finance.gov.sk.
Lukáčiková A., Lukáčik M., *Ekonometrické modelovanie s aplikáciami*, Ekonóm, Bratislava 2008.
Pavlík M., *The usage of dummy variable for VAT forecasting of the tax administration in the Slovak Republic,* Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 6 (1206), Ekonometria 21, J. Dziechciarz (red.), UE, Wrocław 2008, p. 40-54.
Pavlík M., *Odhadovanie daňových príjmov na dani z príjmov fyzických osôb zo závislej činnosti a dani z príjmov fyzických osôb z podnikania,* „Ekonomika Informatika" 2006,1, p. 129-140.
Vogelvang B., *Econometrics theory and applications with eviews*, „Financial Times" 2005.
www.mpavlik.net.

NOWE PODEJŚCIE DO PROGNOZOWANIA KRÓTKOTERMINOWEGO NA PODSTAWIE SZEREGÓW CZASOWYCH ZE ZMIANAMI STRUKTURALNYMI

Streszczenie: Celem artykułu jest wskazanie nowych podejść w prognozowaniu krótkoterminowym w przypadku wystąpienia w szeregach czasowych zmian strukturalnych. Zmiany strukturalne pojawiają się w szeregach czasowych związanych z obserwacjami podatku VAT w Słowacji. Na podstawie danych z okresu styczeń 1996-grudzień 2006 przeprowadzono prognozowanie *ex post*. Przeanalizowano wiele modeli, łącznie z ich transformacją. Za pomocą logarytmów o różnych podstawach przeprowadzono logarytmowanie obserwacji z szeregów czasowych. Porównano wyniki analizy przeprowadzonej gdy podstawami logarytmów były: 5, 10, 15, 20, 25 oraz logarytm naturalny. Wykryto, że logarytm dziesiętny jest najlepszym logarytmem w prowadzonej w ten sposób analizie. Wzrastająca liczba iteracji poprawia prognozę. Sprawdzono, że transformacja nie jest zbieżna i wiele obliczeń pozostaje bez wskazania zbieżności nawet po 250 iteracjach.