Forecasting Probability of Re-employment in Slovakia

Eva Rublíková[1], Martina Lubyova [2]

Abstract

Labour market can be viewed as a dynamic system of economically active and non – active persons flowing from one state to the other. In this paper we focus on studying the monthly time series of outflow rate from unemployment as a proxy for the aggregated probability of leaving the unemployment pool (aggregated re-employment probability) in conjecture with the disposable rate of unemployment in Slovakia during the period from January 2001 to December 2014 using a customized SARIMA model. We use regression models explaining the relationship between outflow rates and vacancy ratios to examine the elasticity of the vacancy ratios with respect to the outflow rates. Finally, we use ARIMA model to forecast disposable rate of unemployment as a benchmark for comparing the precision of forecasts for disposable rate of unemployment by the means of dynamic regression model.

Keywords: *unemployment outflows, vacancy ratio, SARIMA model, dynamic regression model, disposable rate of unemployment*

Abstrakt

Trh práce možno chápať ako na dynamický systém pozostávajúci z tokov ekonomicky aktívnych a neaktívnych osôb, ktorých status na trhu práce sa často mení. V tomto článku sa zameriavame na skúmanie tokov z nezamestnanosti. Konkrétne skúmame mesačné časové rady miery odtoku z nezamestnanosti, ktorá aproximuje agregovanú pravdepodobnosť opustenia nezamestnanosti a disponibilnú mieru nezamestnanosti na Slovensku počas obdobia od januára 2001 do decembra 2014, ktoré modelujeme pomocou modelu SARIMA. Ďalej

[1] Eva Rublikova, CSPS Slovak Academy of Sciences, Sancova 56, Bratislava. eva.rublikova@savba.sk
[2] Martina Lubyova, CSPS Slovak Academy of Science, Šancová 56, Bratislava. progluby@savba.sk

používame regresné modely vysvetľujúce vzťah medzi mierou odtoku z nezamestnanosti a počtom voľných pracovných miest, t.j. skúmame elasticitu. Napokon prognózujeme disponibilnú mieru nezamestnanosti pomocou modelu ARIMA v rámci dynamického regresného modelu.

Kľúčové slová: *odtok z nezamestnanosti, pomer voľných pracovných miest, model SARIMA, dynamický regresný model, disponibilná miera nezamestnanosti*

1. Time series analysis

In general, the Box-Jenkins (1976) methodology represents an important tool in the estimation of monthly time series models. A SARIMA (p, d, q)(P, D, Q)s model with seasonality can be represented by the following formula:

$$\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^s)^D y_t = K + \theta_q(B)\Theta_Q(B^S)a_t \qquad (1)$$

where:

y_t is the value of the time series in time *t*,

s is the number of seasons,

B is back shift operator (i.e. $By_t = y_{y-1}$ or $B^s y_t = y_{t-s}$),

D is the number of seasonal differences; d is the number of non seasonal differences,

$K = \mu(1-\phi_1-\phi_2-...-\phi_p)(1-\Phi_{1,s}-\Phi_{2,s}-...-\Phi_{P,s})$ with $\mu = E(y_t)$ is constant.

Non seasonal polynomials:

$\phi_p(B) = (1-\phi_1 B-\phi_2 B^2-...-\phi_p B^p)$ autoregressive of order *p*

$\theta_q(B) = (1-\theta_1 B-\theta_2 B^2-...-\theta_q B^q)$ moving average of order *q*

Seasonal polynomials:

$\Phi_P(B^S) = (1-\Phi_1 B^S-\Phi_2 B^{2S}-...-\Phi_P B^{PS})$ autoregressive of order *P*

$\Theta_Q(B^S) = (1-\Theta_1 B^S-\Theta_2 B^{2S}-...-\Theta_Q B^{QS})$ moving average of order *Q*

a_t is white noise[3].

The most frequently used model is SARIMA (0, 1, 1)(0, 1, 1)$_s$ which could be written for monthly data in the following form:

$$(1-B)(1-B^{12})y_t = (1-\theta_1 B)(1-\Phi_{1,12}B^{12})a_t \qquad (2)$$

Box and Jenkins popularized their method as a three-stage process aimed at selecting an appropriate model for the purpose of estimating and forecasting various univariate time series. During the **identification stage** of the procedure, the researcher visually examines the time plot of the series to find out whether the series is stationary or non-stationary (series meander without a constant long-run mean or variance). Augmented Dickey-Fuller (ADF) test is used to check the unit roots and the value d of non-seasonal differences to make the non-stationary series to the stationary one.[4] Then a comparison of the sample autocorrelation function and sample partial autocorrelation function with the autocorrelation functions of various theoretical SARMA processes may suggest several plausible models, which can be estimated.

During the **estimation stage**, each of the tentative SARIMA models is fit and various autoregressive coefficients ϕ_i or $\Phi_{i,s}$ and coefficients of moving averages θ_i or $\Theta_{i,s}$ are estimated using statistical software packages.

The third stage of the procedure involves **diagnostic checking**. A standard practice is to plot the residuals to look for outliers and evidence of periods in which the model does not fit the data well. If the variance of residuals is increasing, a logarithmic or Box-Cox transformation can be performed. The residuals from an estimated model have to be serially uncorrelated. To check for the correlation of residuals, one can use correlograms or Ljung-Box Q statistics.[5]
If time series is longer than 60 observations, we can expect some structural change. Using ADF test for time series with structural change may not lead to expressly given results. To answer the question whether the underlying generating process is stable, one can split the data into two sub-samples (the first one from the beginning up to the point of the change, the

[3] RUBLÍKOVÁ, E., PRÍHODOVÁ, I.: *Analýza vybraných časových radov - ARIMA modely*. Bratislava: Vydavateľstvo Ekonóm, Ekonomická univerzita v Bratislave, 2008.
[4] ARLT, J., ARLTOVÁ, M.: *Finanční časové řady*. Praha: Grada Publishing,, 2003.
[5] LJUNG, G. M., BOX, G.E.P. On a Measure of a Lack of Fit in Time Series Models. *Biometrika* 65 (2): p.297-303

second one following the change and to the end of the series). Fitting the same SARIMA model to each of the two sub-samples of the series can provide useful information concerning the assumption that the data-generating process is not changing. This idea could be verified by the means of *F*-test.

Suppose we are estimating a SARIMA model using a sample of T observations and the sum of squared residuals is denoted as SSR. Then we divide the T observations into two sub-samples with t_1 observations in the first sub-sample and $t_2 = T - t_1$ observations in the second one. Using each sub-sample to estimate the model SARIMA with the same coefficients and denote the sum of the squared residuals from each sub-sample model as SSR_1 and SSR_2 respectively, we can test the restriction that all coefficients are equal by the means of *F*-test of the following form:

$$F = \frac{(SSR - SSR_1 - SSR_2)/n}{(SSR_1 + SSR_2)/(T-2n)} \qquad (3)$$

where n is the number of parameters estimated ($n = p + P + q + Q + 1$) if an intercept is included. F has F-distribution with (n, T-$2n$) degrees of freedom. The larger the calculated value of F, the more restrictive is the assumption that the two sets of coefficients are equal. The null hypothesis with this assumption is rejected at the level α if $F > F_{\frac{\alpha}{2}}(n, T-2n)$.[6]

Another possibility of checking the fitness of the model is to use almost entire time series (for example fitting a model using 13 years of monthly data from 14 years), which is denoted as the Estimated period. Verified estimated model is then used to make out sample forecasts for the remaining so-called Validation period (i.e. the last year or the 14-th year period in our example). The mean statistics like RMSE, MAE, MAPE or Theil Inequality Coefficient computed on both samples provide useful indicators to compare the adequacy of some alternative models[7] whereas those models with poor mean statistics for out of sample forecasts errors (for the Validation period) should be eliminated.

[6] ENDERS, W.: *Applied Econometric Time Series*. Wiley Series in Probability and Mathematical Statistics, New-York, 1995. pp. 99
[7] E Views Users Guide, Quantitative Micro Software

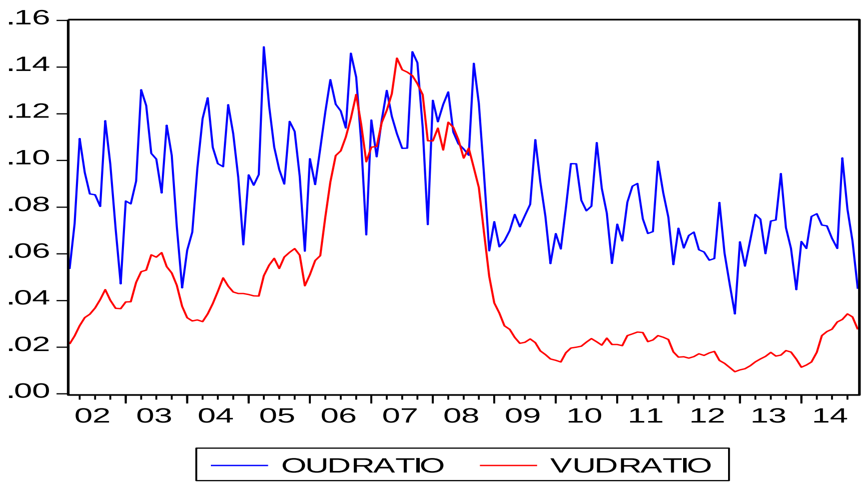
1.2 Model SARIMA(p, 1, q)(0, 1, Q)$_{12}$ for outflow rates and vacancy ratios

In this section we construct and examine seasonal ARIMA model for fitting the time series of monthly outflow rates (O/UD) and vacancy ratios (V/UD) in Slovakia using the monthly data beginning in January 2001 and ending in May 2015, which represents 173 observations. The data used are from the Centre of Labour, Social Affairs and Family of the Slovak Republic.

The outflow rates time series is defined as the ratio of outflow to disposable job applicants registered in the databases of labour offices, representing a proxy for the "aggregated probability of reemployment", labeled as OUDRATIO. The vacancy ratio is defined as the ratio of vacancies to disposable job applicants registered in the database of labour offices and is labeled as VUDRATIO. Both time series are depicted at Figure 1.

From the Figure 1 it can be seen that the economic crisis incurred change in both variables. During the period from January 2001 till December 2008, the time series OUDRATIO exhibits a slow upward trend with constant variability. Thus we assume that January 2009 is the point of change in the level of series. From that date until May 2015, the series is almost stationary. On the other hand, the time series of vacancy ratios (VUDRATIO) has a stochastic upward trend till June 2007, after that it turns to a stochastic downward trend till December 2008. From January 2009 till March 2015 the data of vacancy ratios seem to be stationary.

Figure 1 Monthly outflow rates and vacancy ratios in SR, January 2001 - May 2015



Source: Own processing based on the data from the Centre of Labour, Socia Affairs and Family of SR

In order to examine how the assumed change of the regime influences the precision of the forecasts for aggregated re-employment probability (OUDRATIO). First we will use our model for the whole period (173 observations, denoted as MODEL) and compare the precision of forecasts when the model is used for shorter data series from January 2009 till May 2015 (75 observations). To see whether the change at the end of 2008 is statistically significant at the level of 5 %, we use *F* test computed from the sum of squares of residuals of the same SARIMA model, which is estimated on the two sub-samples. The first sub-sample is described by MODEL 1 using data from January 2001 to December 2008. The second sub-sample is described by MODEL 2 using data from January 2009 to December 2014, respectively.

To compare the precision of the forecasts given by the three models we will use five observations (January 2015 - May 2015) as the Validation period to see how the accuracy of the short-term out of sample forecast errors could be changed.

SARIMA(0, 1, 0)(0, 1, 1)$_{12}$ model for outflow rates

We start by analyzing the time series of outflow rates, i.e. OUDRATIO_t for t = 1, 2, ... , 173. Since the time series is not stationary, we need to make it stationary by the means of transformation $z_t = (1 - B^{12})(1 - B)\log(OUDRATIO)_t$

We have identified the Box–Jenkins model for log(OUDRATIO) as SARIMA(0,1,0)x(0,1,1)$_{12}$ given by the following formula:

$$z_t = (1 - \Theta_{1,12} B^{12}) a_t .$$

The statistics concerning the estimated parameters of all the three models are summarized in Table 1.

Table 1 Estimated statistics of SARIMA models for aggregated probability to be re-employed in Slovakia

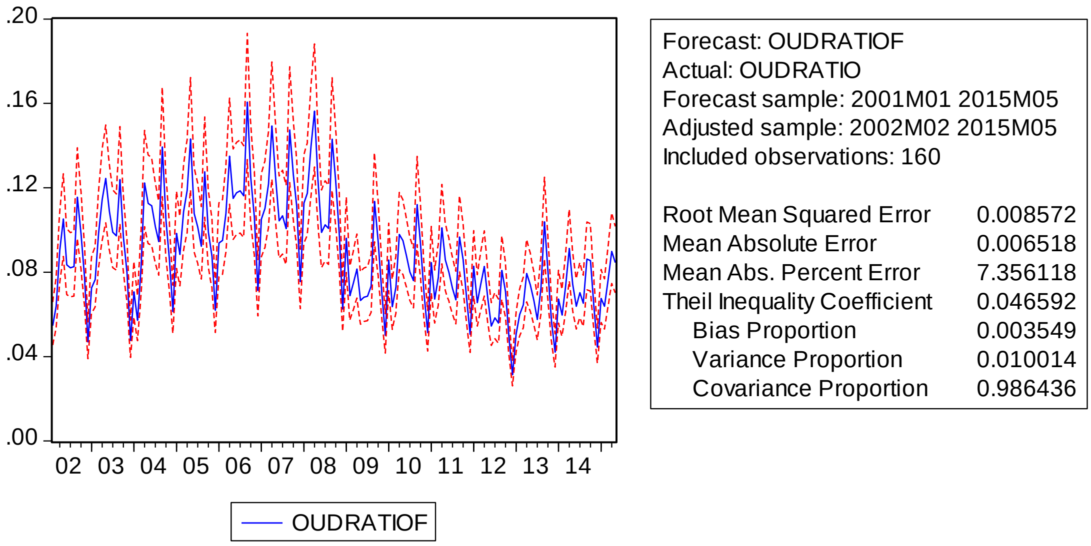
Statistics	MODEL	MODEL 1	MODEL 2
T (period)	173 (1/2001-5/2015)		77 (1/2009-5/2015)
Model	SARIMA(0,1,0)(0,1,1)12	SARIMA(0,1,0)(0,1,1)12	SARIMA(0,1,0)(0,1,1)12
Parameter	$\hat{\Theta}_{1,12} = -0.927205$		$\hat{\Theta}_{1,12} = -0.866766$
R-square	0.403952		0.503050
S.E. of Regress.	0.092603		0.086971
SSResidual	1.363490		0.574860
DW	2.11		1.99
Schwartz	–1.896		–2.003
Estimation period	**168 (1/2001-12/2014)**	**96 (1/2001-12/2008)**	**72 (1/2009-12/2014)**
Parameter	$\hat{\Theta}_{1,12} = -\ 0.920938$	$\hat{\Theta}_{1,12} = -\ 0.898499$	$\hat{\Theta}_{1,12} = -\ 0.869701$
R square	0.401039	0.4664	0.50344
S.E. of Regress.	0.009384	0.0857	0.088597
SSResidual	1.356056	0.602745	0.557303
DW	2.11	2.14	1.95
Schwartz	-1.87	-2.034	-1.96
RMSE	0.00868	0.009218	0.006417
MAE	0.00663	0.006838	0.004896
MAPE	7.46 %	6.49 %	6.67 %
Theil Inequality Coeff.	0.0047	0.0438	0.0434
Validat. Period	**5 (1/2015-5/2015)**	**5 (1/2009-5/2009)**	**5 (1/2015-5/2015)**
RMSE	0.002067	0.01145	0.003376
MAE	0.001765	0.01037	0.002975
MAPE	2.28 %	14.63 %	3.89 %
Theil Inequality Coeff.	0.01349	0.07764	0.0224

Source: Own computations

From Table 1 it is clear that the estimated seasonal moving average parameter is stable in all models (the value is about – 0.9). Mean statistics of residuals for Estimation period are slightly better for the MODEL 2 computed on time series after change. During the Validation period the mean forecast errors are smaller for the MODEL using time series from January 2001 till December 2014. *F*-test shows that the generating process has been changed significantly at 5 % due to the outside reasons. Statistics F = 28.04 is larger than the critical value $F_{0,05}(1,171) = 3.00$. In spite of this, we will use the entire SARIMA model as the best one.

Figure 2 shows how the model SARIMA (0, 1, 0)(0, 1, 1)12 fits real data during the whole period (January 2001 – May 2015) with 95 % confidence interval.

Figure 2 **95 % confidence interval for in sample forecasts of outflow rates (OUDRATIO) during the period from January 2001 till May 2015**



Source: Own computations

SARIMA(1, 1, 0)(0, 1, 1)12 model for vacancy ratio

In order to identify a Box–Jenkins model for vacancy ratio (VUDRATIO) during the whole period January 2001 till March 2015, we used sample autocorrelation and sample partial autocorrelation functions for stationary transformation series

$$z_t = (1-B)(1-B^{12})\log(VUDRATIO)_t \quad \text{t = 14, 15, …, 171.}$$

We assume that the model SARIMA (1, 1, 0)(0, 1, 1)12 can be provide an appropriate approximation of the generating process. The suggested model is of the form:

$$\phi_1(B)z_t = \Theta_{1,12}(B^{12})a_t$$

The estimated statistics describing all models are listed in Table 2 for the whole time series, and for the Estimation and Validation periods, respectively.

Table 2 Estimated statistics of SARIMA models for vacancy ratio in Slovakia

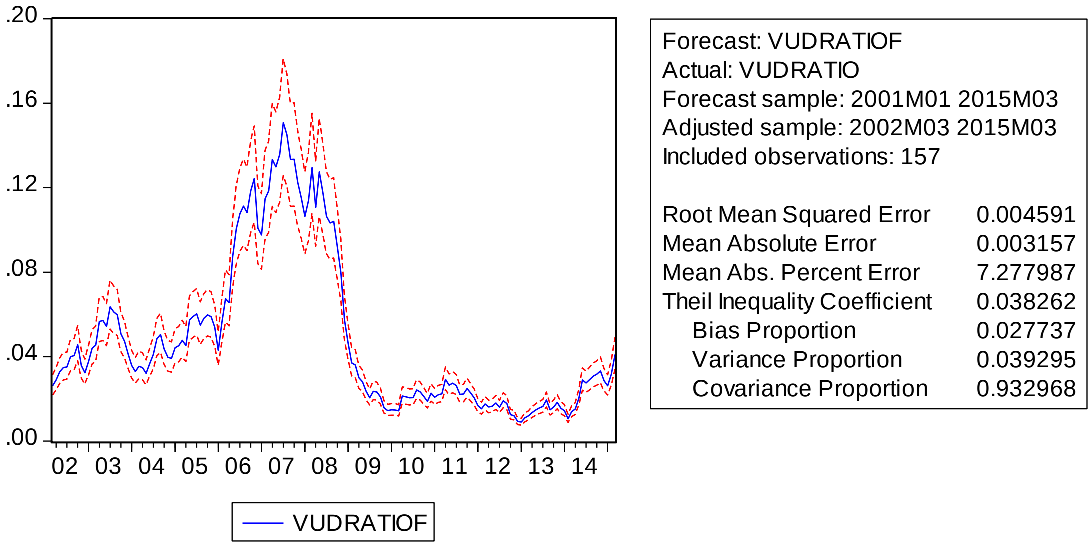
Statistics	MODEL	MODEL 1	MODEL 2
T (period)	171 (1/2001-3/2015)		75 (1/2009-3/2015)
Model	SARIMA(1,1,0)(0,1,1)12	SARIMA(1,1,0)(0,1,1)12	SARIMA(0,1,0)(0,1,1)12
Parameter	$\hat{\phi}_1 = 0.398456$ $\hat{\Theta}_{1,12} = -0.908523$		$\hat{\phi}_1 = 0.423678$ $\hat{\Theta}_{1,12} = -0.863803$
R square	0.5094		0.5016
S.E. of Regress.	0.090836		0.103746
SSResidual	1.278937		0.785815
DW	2.04		2.03
Schwartz	-1.91		-1.60
Estimation period	**168 (1/2001-12/2014)**	**96 (1/2001-12/2008)**	**72 (1/2009-12/2014)**
Parameter	$\hat{\phi}_1 = 0.3343$ $\hat{\Theta}_{1,12} = -0.9003$	$\hat{\phi}_1 = 0.345633$ $\hat{\Theta}_{1,12} = -0.89683$	$\hat{\phi}_1 = 0.352152$ $\hat{\Theta}_{1,12} = -0.854304$
R square	0.4957	0.501076	0.4937
S.E. of Regress.	0.4924	0.077438	0.102203
SSResidual	1.235791	0.479734	0.731188
DW	2.05	1.99	2.09
Schwartz	–1.92	-2.19	–1.63
RMSE	0.00458	0.005424	0.00206
MAE	0.00314	0.004327	0.00154
MAPE	7.12 %	6.48 %	7.65 %
Theil Inequality Coeff.	0.038	0.0034	0.0481
Validat. Period	**3 (1/2015-3/2015)**	**3 (1/2009-3/2009)**	**3 (1/2015-3/2015)**
RMSE	0.00493	0.004754	0.005025
MAE	0.00466	0.004375	0.000474
MAPE	12.17 %	12.93 %	12.43 %
Theil Inequality Coeff.	0.0692	0.0648	0.0706

Source: Own computations

From Table 2 we can conclude that in spite of large observed variability, all the three models estimated parameters wit very close values of mean statistics for residuals and forecasts errors. Therefore, we will use as the MODEL for forecasting the entire time series. The end of 2008 has not been proved as the point of the change, as the statistics F = 1,705 is not larger than the critical value $F_{0,05}(2,164) = 3.00$.

Figure 3 shows how well the model describes the entire time series with 95 % confidence interval for residuals and their mean measures during the entire period from January 2001 till March 2015.

Figure 3 **95 % confidence interval for in sample forecasts of vacancy ratios (VUDRATIO) during the period from January 2001 till March 2015**

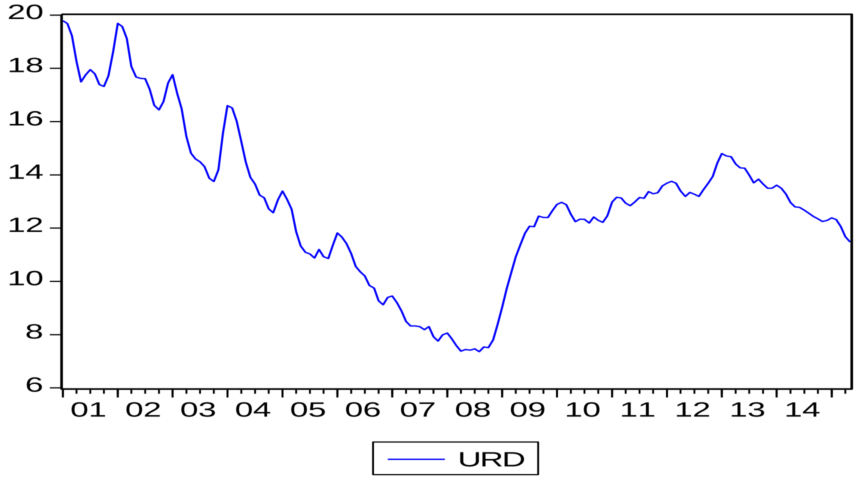


Source: Own computations

SARIMA(0, 2, 1)(0, 1, 1)$_{12}$ model for disposable rate of unemployment

In this section we are interested in SARIMA model for the disposable unemployment rate computed from the disposable registered job applicants in Slovakia to make short-term forecasts. The variable is abbreviated as URD. Figure 4 shows the development of disposable rate of unemployment during the period January 2001 till May 2015. At Figure 4 we can trace a stochastic downward trend till the end of 2008, while from January 2009 till May 2015 the series is again rather stationary. We can say that the data for the entire period exhibit non-stationary character with seasonality.

Figure 4 **Monthly time series of Disposable rate of Unemployment in Slovakia from January 2001 till May 2015**



Source: Own computations

On order to identify a suitable Box–Jenkins model for disposable rate of unemployment (URD) during the whole period from January 2001 till May 2015 we used sample autocorrelation function and sample partial autocorrelation function for the following stationary transformation series:

$$z_t = (1-B)^2(1-B^{12})\log(URD)_t \quad t = 15, 16, \ldots, 175.$$

We assume the model SARIMA (0, 2, 1)(0, 1, 1)12 could be appropriate. The suggested model has the following form:

$$z_t = \theta_1(B)\Theta_{1,12}(B^{12})a_t$$

By comparing the estimation results in Table 3 we can conclude that the best fitting model is MODEL 2 for the Estimation period from January 2009 till December 2014 that delivers very small differences when compared to the Validation period. So there is no good base to decide which of the Estimation period would be used to find out regression model for disposable rate of unemployment to get good forecasts. For this reason, we decided to use both models.

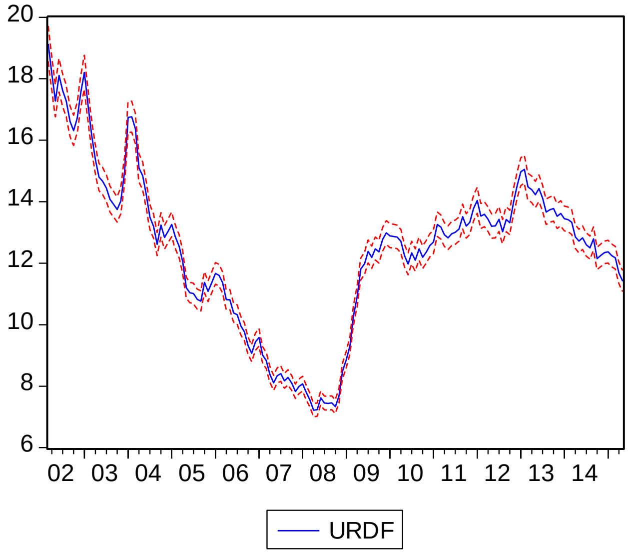
Figure 5 shows how well the models describe the entire time series with 95 % confidence interval for residuals and their mean measures during the entire period from January 2001 till March 2015.

Table 3 **Estimated statistics of SARIMA models for disposable rate of unemployment in Slovakia**

Statistics	MODEL	MODEL 1	MODEL 2
T (period)	173 (1/2001-5/2015)		77 (1/2009-5/2015)
Model	SARIMA(0,2,1)(0,1,1)12	SARIMA(0,2,1)(0,1,1)12	SARIMA(0,2,1)(0,1,1)12
Parameter	$\hat{\theta}_1 = -0.183695$ $\hat{\Theta}_{1,12} = -0.463827$		$\hat{\theta}_1 = -0.152096$ $\hat{\Theta}_{1,12} = -0.821971$
R square	0.185518		0.51914
S.E. of Regress.	0.015001		0.010802
SSResidual	0.035369		0.008751
DW	1.97		2.19
Schwartz	–5.51		–6.13
Estimation period	**168 (1/2001-12/2014)**	**96 (1/2001-12/2008)**	**72 (1/2009-12/2014)**
Parameter	$\hat{\theta}_1 = -0.18204$ $\hat{\Theta}_{1,12} = -0.463956$		$\hat{\theta}_1 = -0.149106$ $\hat{\Theta}_{1,12} = -0.824685$
R square	0.1871	0.136541	0.527744
S.E. of Regress.	0.01522	0.016483	0.011049
SSResidual	0.035211	0.021736	0.008546
DW	1.97	2.04	2.19
Schwartz	–5.48	-5.29	–6.08
RMSE	0.188482	0.204264	0.13957
MAE	0.137036	0.146363	0.108704
MAPE	1.12 %	1.224	0.8448 %
Theil Inequality Coeff.	0.007708	0.008173	0.005355
Validat. Period	**5 (1/2015-5/2015)**	**5 (1/2009-5/2009)**	**5 (1/2015-5/2015)**
RMSE	0.067749	0.275637	0.05263
MAE	0.054219	0.205435	0.04895
MAPE	0.4517 %	2.002 %	0.4079 %
Theil Inequality Coeff.	0.00283	0.01336	0.002196

Source: Own computations

Figure 5 **95 % confidence interval for in sample forecasts of disposable rate of unemployment (URD) in Slovakia during the period from January 2001 till May 2015**



Forecast: URDF
Actual: URD
Forecast sample: 2001M01 2015M05
Adjusted sample: 2002M03 2015M05
Included observations: 159

Root Mean Squared Error	0.185881
Mean Absolute Error	0.134442
Mean Abs. Percent Error	1.097167
Theil Inequality Coefficient	0.007318
Bias Proportion	0.000705
Variance Proportion	0.013259
Covariance Proportion	0.986036

Source: Own computations

2. Regression Model for outflow rates

In order to understand how vacancy ratios influence the outflow rates, we select only the last part of the entire time series, i.e. the period from January 2009 till March 2015 ($T_2 = 75$ observations) to avoid the large variability in the vacancy ratios. We build up a simple model of hiring function in its log-log form with autocorrelated residuals, as the seasonal dummy variables were not sufficient.

$$\log(OUDRATIO)_t = \beta_0 + \beta_1 \log(VUDRATIO)_t + a_t$$

$$\phi_1(B)a_t = \Theta_{1,12}(B)\varepsilon_t$$

where ε_t is white noise.

The estimation of the given model is:

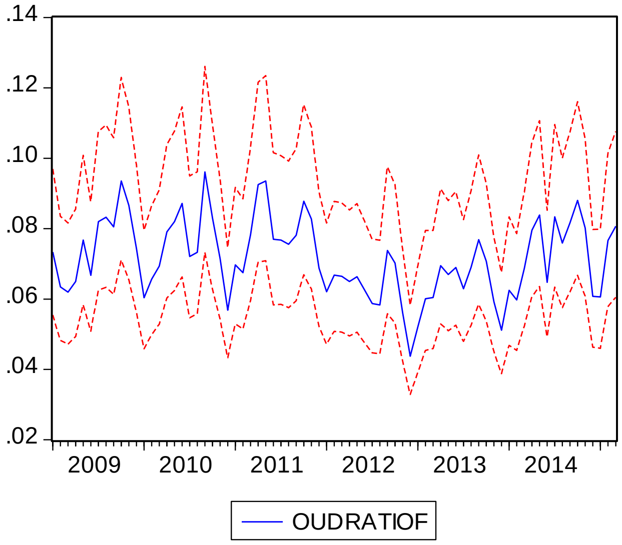
$$est\ \log(OUDRATIO)_t = -1.9788 + 0.17885 \log(VUDRATIO)_t$$
$$(0.332) \quad (0.085)$$

$$(1 - 0.4823B)\hat{a}_t = (1 - 0.8428B^{12}\varepsilon_t$$
$$(0.106) \qquad\qquad (0.031)$$

with $R^2 = 0.5892$; D-W=2.04. Residuals are normally distributed with J-B = 0.5842 (its P-value 0.7467).

The coefficient of 0.17885 can be interpreted as the constant percentage increase in outflow rates as a result of a 1 percent increase in vacancy ratio. Since the vacancy ratio elasticity value of 0.17885 is less than one, we can say that outflow rates are inelastic. Figure 6 shows the 95 % confidence interval for outflow rates together with the mean measures of residuals.

Figure 6 95 % confidence interval for outflow rates during the period from January 2009 till March 2015 and Mean measures of residuals



Forecast: OUDRATIOF	
Actual: OUDRATIO	
Forecast sample: 2009M01 2015M03	
Included observations: 75	
Root Mean Squared Error	0.009432
Mean Absolute Error	0.007489
Mean Abs. Percent Error	10.53530
Theil Inequality Coefficient	0.064529
Bias Proportion	0.004102
Variance Proportion	0.121865
Covariance Proportion	0.874032

Source: Own computation

Coefficient of determination is close to 60 percent, so the aggregate characteristic such as the vacancy ratio is not sufficient to express all the variability of outflow rates. Possibly there are other variables in the structural domain that significantly influence the results, such as the share of long term unemployment, skills mismatch. Unemployment rate with time lag and effective taxation of labour with time lag can be also used to improve the explanatory power of the model. Nowadays the policy makers try to influence the rate of unemployment directly by various aggregate instruments aimed at slowing down the job destruction and boosting up job creation. However, the future results may depend on addressing the underlying structural issues.

3. Regression model for disposable rate of unemployment with autocorrelated residuals

In the previous section we have seen that the mean measures of residuals in the sub-sample from January 2009 till May 2015 for disposable rate of unemployment showed better values in Estimation and Validation period, respectively. For this reason we are at first interested in the regression model for disposable rate of unemployment during the Estimation period from January 2009 to December 20014 while the Validation period will span from January 2015 to May 2015.

Table 4 summarizes the estimation results for the coefficients of the model with autoccorelated residuals of the following form

$$\log(URD)_t = c + \beta_1 \log(VUDRATIO)_t + \beta_2 time + \beta_3 time^2 + a_t$$

$$\phi_1(B)a_t = \theta_{10}(B)\varepsilon_t \,.$$

Table 4 Results of the estimated model for logarithm of disposable rate of unemployment in Slovakia

Dependent Variable: LOG(URD)
Method: Least Squares
Date: 07/28/15 Time: 09:17
Sample: 2009M01 2014M12
Included observations: 72
Convergence achieved after 22 iterations
Back-cast: 2008M03 2008M12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.642484	0.411426	1.561601	0.1232
LOG(VUDRATIO)	-0.069344	0.011895	-5.829599	0.0000
TIME	0.023705	0.006064	3.908869	0.0002
TIME^2	-8.32E-05	2.15E-05	-3.878853	0.0002
AR(1)	0.766686	0.052414	14.62742	0.0000
MA(7)	-0.468943	0.079330	-5.911272	0.0000
MA(10)	-0.509383	0.074708	-6.818342	0.0000

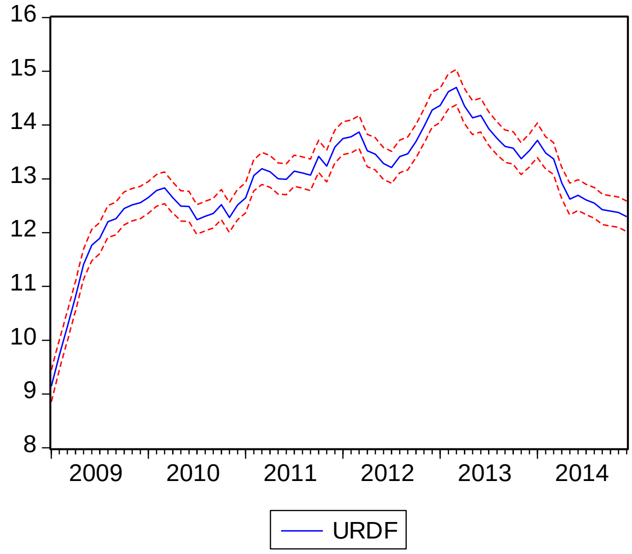
R-squared	0.984592	Mean dependent var	2.555863
Adjusted R-squared	0.983170	S.D. dependent var	0.084750
S.E. of regression	0.010995	Akaike info criterion	-6.090646
Sum squared resid.	0.007857	Schwarz criterion	-5.869304
Log likelihood	226.2633	F-statistic	692.2675
Durbin-Watson stat	1.603639	Prob(F-statistic)	0.000000

Inverted AR Roots	.77			
Inverted MA Roots	1.00	.70-.60i	.70+.60i	.32+.82i
	.32-.82i	-.27+.95i	-.27-.95i	-.82-.52i
	-.82+.52i	-.86		

Source: Own computation

It is interesting to note that a one-percent increase in the vacancy ratio would lower the disposable rate of unemployment by about seven percent. The autocorelated residuals express the missing variables and seasonality in data. The picture of the model fit is provided by Figure 7.

Figure 7 **Real data of disposable rate of unemployment in Slovakia with 95 % confidence interval for residuals and their in sample mean statistics.**



Forecast: URDF
Actual: URD
Forecast sample: 2009M01 2014M12
Included observations: 72

Root Mean Squared Error	0.136177
Mean Absolute Error	0.104826
Mean Abs. Percent Error	0.810963
Theil Inequality Coefficient	0.005253
Bias Proportion	0.005154
Variance Proportion	0.021616
Covariance Proportion	0.973230

Source: Own computation

Based on the properties depicted at Figure 7 we can see that out of sample statistics of forecast errors gives MAPE = 0.8109 % , which is close to 1 %.

Conclusions

The paper presents several models estimating the properties of the time series of unemployment outflow rates, vacancy ratios, and disposable rate of unemployment during the period from January 2001 to March (or May) 2015. We found out that despite the fact that the economic crises influenced the behaviour of the time series, the variables have a stationary path between January 2009 and March 2015. The question was whether the break observed at the end of 2008 could influence the precision of the forecasts. To answer this question we split the series into two sub-samples (before and after the break: sub-sample 1 covered observations from January 2001 till December 2008 and sub-sample 2 observations from January 2009 to December 2014) to find out whether estimated model is stable on both sub-samples and on the entire period. Our results show that the estimated parameters in the all three models are not significantly influenced by the change. Therefore, there is no reason to curtail some observations and to base the forecasts only on the sub-sample 2. Thus we suggest basing the forecasts on both the entire sample and the sub-sample 2, as alternative sources for forecasting.

Further we were interested in regression models that estimate the relationship between outflow rates and vacancy ratios. For thus purpose we used data series only from sub-sample 2 in order to avoid the relatively large variability in vacancy ratios before the break. The log-log hiring regression model delivered the coefficient of determination of 60 % which indicates that using the aggregate explanatory variables such is not sufficient. Government is looking for new policy in active and passive policy to reduce unemployment. Looking for other variables of structural character could improve the explanatory power of the model. Since we were interested in outflow ratios for disposable unemployed, we examined also how the disposable rate of unemployment is influenced by the vacancy ratios. We used Log-Log regression model for the sub-sample 2 in order to avoid the larger variability in vacancy ratios. We find that a one-percent increase in the vacancy ratio would lower the disposable rate of unemployment by about seven percent. The autoccoerelated residuals of the model could be expressed as ARMA(1,10) model to substitute missing variables and seasonality in data.

Acknowledgements:

The article was supported by the project VEGA 2/0160/13 “Financial stability and sustainability of economic growth of Slovakia in conditions of Global Economy” and VEGA Num. 2/0010/2014 “Institutional and technological changes in the context of European challenges”.

References

ARLT, J., ARLTOVÁ, M.: *Finanční časové řady.* Praha: Grada Publishing, 2003.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C.: *Time Series Analysis. Forecasting and Control*. Prentice Hall, Inc.1994.

ENDERS, W.: *Applied Econometric Time Series*. Wiley Series in Probability and Mathematical Statistics, New-York, 1995.

E Views Users Guide, Quantitative Micro Software

LJUNG, G. M., BOX, G.E.P.: On a Measure of a Lack of Fit in Time Series Models. *Biometrika* 65. 1978

Rublíková, E., Príhodová, I.: *Analýza vybraných časových radov –ARIMA modely*. Vydavateľstvo EKONÓM, Ekonomická univerzita Bratislava, 2008.