ESTIMATING ARIMA–ARCH MODEL RATE OF UNEMPLOYMENT IN SLOVAKIA

Eva Rublikova[1]
Martina Lubyova[2]

ABSTRACT

Slovak labour market is characterized by persistently high rate of unemployment. Unemployment is therefore a variable of utter importance for the Slovak economy. In this paper we examine the possibility to forecast the time series of the rate of unemployment in Slovakia by the use of technique that does not require the assumption of constant variance over time. The analysed data represent the monthly rate of disposable unemployment during the period January 1999 - May 2013. We examine whether the observed changing variability of the time series is statistically significant and could be described by appropriate ARIMA–ARCH model. We used the Lagrange multiplier test to determine whether autoregressive heteroscedastic effect is statistically significant at 5 % level of significance. The results of analysis confirmed our hypothesis about statistically significant changes in the variance of the time series. Thus we proposed a combination of ARIMA(0,1,2)(0,1,1)12 + GARCH(1,1) models that proved to provide good predictors for both the conditional mean and the conditional variance.

Keywords: *Unemployment rate, Time series analysis, Forecasting*
JEL codes: C53 Forecasting and prediction methods, J64 Unemployment: models, duration, incidence and job search

---

[1] Institute for Forecasting, Slovak Academy of Sciences, VEGA project 2/0160/13 “Financial stability and sustainable growth of Slovakia in the conditions of global economy”
[2] Institute for Forecasting, Slovak Academy of Sciences, VEGA project 2/0058/10 “Structural adaptation of a small opened economy”

ABSTRAKT v slovenskom jazyku

Slovenský trh práce je špecifický trvale vysokou mierou nezamestnanosti. Nezamestnanosť je preto veľmi dôležitým ukazovateľom slovenskej ekonomiky. V práci sa zaoberáme možnosťami prognózovania časového radu vývoja nezamestnanosti na metódou, ktorá nevyžaduje predpoklad o konštantnej variancii radu hodnôt v čase. Využívame časový rad mesačných údajov o disponibilnej miere nezamestnanosti na Slovensku za obdobie január 1999 – máj 2013. Testujeme predpoklad o štatisticky významnej zmene variability v čase a možnosť charakterizovať časový rad pomocou ARIMA–ARCH modelu. Na základe testu založeného na Langrangeovom multiplikátore zisťujeme štatisticky významnú prítomnosť autoregresívneho heteroskedastického efektu na hladine významnosti 5 %. Výsledky potvrdzujú náš predpoklad o štatisticky významných zmenách variability radu hodnôt v čase. Následne navrhujeme kombináciu modelov ARIMA(0,1,2)(0,1,1)12 + GARCH(1,1), ktoré poskytujú dobré predikcie pre kondicionálnu strednú hodnotu aj kondocionálnu varianciu skúmaného časového radu.

Kľúčové slová: *miera nezamestnanosti, analýza časových radov, prognózovanie*

JEL kódy: C53 Prognózovanie predikčné modely, J64 Nezamestnanosť: modely, dĺžka trvania, výskyt a hľadanie práce

Introduction

Slovak labour market is characterized by persistently high rate of unemployment. Unemployment is therefore a variable of utter importance for the Slovak economy. Predicting the values of unemployment can be used to forecast the future economic development and its related aspects, including its fiscal implications. The latter include, for example the direct outlays spent on the various types of labour market policies, notably payments of unemployment benefits and expenditures on active labour market measures. However, there are also other less direct links between the unemployment and fiscal policies. A classical example is the empirically observed relationship between the unemployment and rate of inflation – the so-called Phillips curve. The curve has a hyperbolic shape and illustrates an indirect relationship between the two variables. It has been widely accepted in the economic literature that the trade-off tends to be observed in the short run; however, there are better predictors of the unemployment rate, in particular in the long run. Without making any a priori statements about the causality of the relationship, we can reasonably assume that the inflation is in general more volatile than the unemployment rate. This is mainly because the inflation is dictated by the evolution of prices, which are considered to be a "jump" variable that can undergo remarkable changes practically on the spot. On the other hand, the rate of unemployment is determined by processes that are less volatile and characterised by certain degree of inertia. This is mainly because the underlying process is linked to the activities of people such as job search or training, which typically take longer time to get completed. Therefore, the rate of unemployment can represent a variable that is more suitable for the use of time-series based forecasting methods, which assume certain degree of inertia.

Traditionally, the national institutions providing forecast of economic variables are more concerned with financial monetary indicators rather than those of the social sphere. The Ministry of Finance and National Bank of Slovakia have been providing short-term to mid-term macroeconomic forecasts, including those of unemployment rate. The forecasts are based time series with quarterly or annual frequency with the predicted period of 2.5 to 3 years. In our paper we focus on the forecasts based on time series with monthly frequency that provides a shorter predicted period, however, it provides also different insights and qualities (i.e. better account of seasonality, the possibility to predict the expenditures on labour market measures on a more accurate basis, etc.). It is also noteworthy that LFS provides only quarterly data while the data on administrative unemployment are recorded with monthly frequency that could and should be used also for the forecasting purposes. We focus on the so-called disposable rate of unemployment that originate from administrative data and is based on the number of registered unemployed who are able to take up a suitable job, if offered to them.[3]

In addition to using a new source of data, we employ a more complex methodology that takes into consideration forecasting the level (conditional mean of the series), as well as its variability. The traditional Box-Jenkins methodology of ARIMA models is concentrated on forecasting time series in the form of conditional means, i.e. the point forecast of a time series at time t given information on the series up to time t – 1 is $E(y_t|y_{t-1}, y_{t-2}, ...)$ while implicitly assuming that the conditional variance remains constant. However, there are many time series for which this assumption is not valid. For example, many financial time series exhibit changes in variance over time, which are usually serially correlated with groups of highly volatile observations occurring together. If these changes are statistically significant and create autoregressive heteroscedastic effect, than we can create ARIMA – ARCH model, which was the first time described by Engle (1982). Both models, ARIMA(p, d, q)(P, D, Q)s for conditional mean and the ARCH(q) model for variance of the series, taken together can provide a better predictor for the future values of the series in the short-term horizon.

1. Autoregressive Moving Average Process for Modelling Conditional Mean

The analysis of long time series y_t, t = 1, 2, …, T could detect stochastic trend, stochastic seasonality and possible cycle in the series which could be described by the ARIMA(p, d, q)(P, D, Q)$_s$ that takes form

$$\phi_p(B)\Phi_P(B)(1-B)^d(1-B^s)^D y_t = \theta_q(B)\Theta_Q(B^s)a_t, \quad (1)$$

where

[3] The disposable unemployment rate is thus lower than the total unemployment rate, as the latter includes also persons who are not immediately available for work because of reasons such as sickness, maternity leave, etc.

B is back shift operator, $B^s y_t = y_{t-s}$,
$\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p)$
$\Phi_P(B^s) = (1 - \Phi_1 B^s - \Phi_2 B^{22} - ... - \Phi_P B^{Pp})$
$\theta_p(B) = (1 - \theta_1 B - \theta_2 B^2 - ... - \varsigma_q B^p)$
$\Theta_Q(B^s) = (1 - \Theta_1 B^s - \Theta_2 B^{22} - ... - \Theta_Q B^{Qs})$
d is the order of no seasonal differencing, not larger than 2,
D is the order of seasonal differencing , obviously 1,
B is back shift operator, $B^s y_t = y_{t-s}$,
s is length of seasonality.

We assume that y_t is stationary in mean and variance. If the series is not stationary in mean, the no seasonal and seasonal differences adjust the series from stochastic trend and stochastic seasonality very well. If the series is not stationary in variance, Box-Cox transformation is very useful tool to make it homoscedastic.

Generalized form of the stationary transformation is

$$z_t = (1 - B)^d (1 - B^s)^D y_t \qquad (2)$$

2. ARCH (q) model

Engle in 1982 proposed the model which allows the variance to depend upon the available information set. Assuming conditional normality, a general specification of the evolution of y_t would be

$$y_t | Y_{t-1} \sim N(g_t, h_t) \qquad (3)$$

where $Y_{t-1} = \{y_{t-s}, s \geq 1\}$, and where both g_t and h_t are functions of the past observations of the variable Y_{t-1}. So in general g_t could be expressed by appropriate AR(p) model, but h_t is defined as follows

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + ... + \alpha_q a_{t-q}^2 , \; a_t = y_t - g_t . \qquad (4)$$

with $\alpha_0 > 0$, $\alpha_i \geq 0$ for i = 1, 2, ..., q and $\sum_{i=1}^{q} \alpha_i < 1$.

Equations (3) and (4) together are known as an autoregressive conditional heteroscedasticity model or ARCH(q) model.

The errors a_t although serially uncorrelated through the white noise assumption, are not independent since they are related through their second moments.
To test heteroscedasticity in time series according to Lagrangian multiplier test, we perform the second order regression of the following form

$$\hat{a}_t^{\,2} = \alpha_0 + \alpha_1 \hat{a}_{t-1}^{\,2} + \alpha_2 \hat{a}_{t-2}^{\,2} + ... + \alpha_q \hat{a}_{t-q}^{\,2} + \nu_t \qquad (5)$$

where $\hat{a}_t$ are residuals of ARIMA(p, d, q)(P, D, Q)s model and ν_t is white noise.

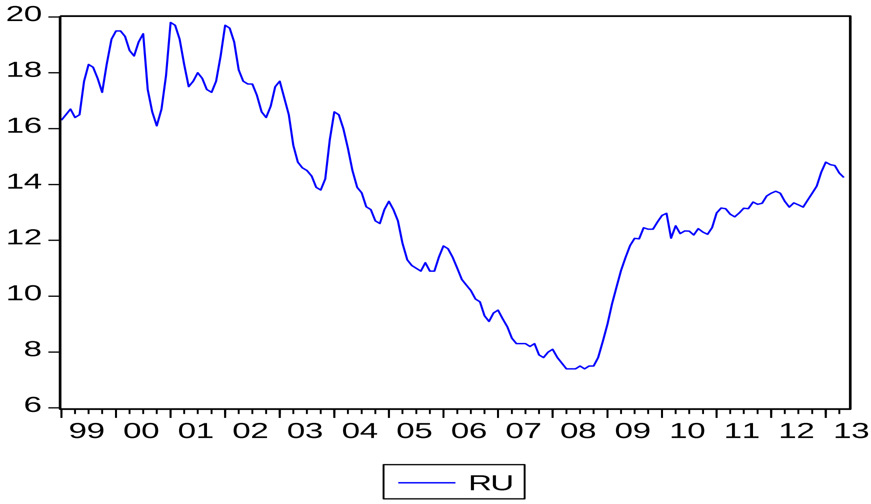
Lagrangian multiplier test statistic LM = T.R^2 has probability distribution $\chi^2(q)$. We reject null hypothesis that series has homoscedastic variability if LM is larger than α % critical value of χ^2 distribution with q degrees of freedom. Conclusion is that the ARCH (q) model of the form (4) is appropriate to estimate.

3. Estimating the conditional mean model for Disposable Rate of Unemployment

Time series of the rate of unemployment in SR during the period January 1999 till May 2013 has 173 observations computed from the disposable unemployed persons. The data are from the Ministry of Labour Social Affairs and Family of Slovakia. The analysed period is quite long time period to investigate whether the series contain not only seasonality but also cycle or changing variability in time.

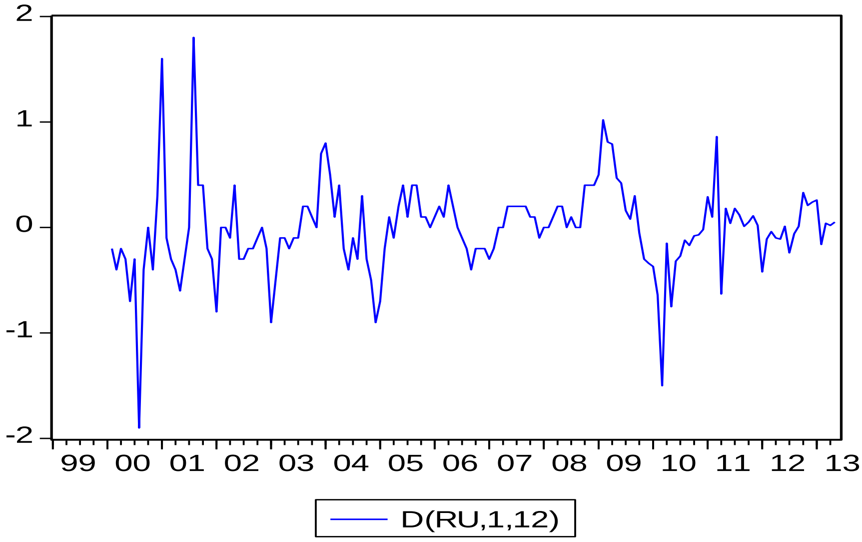
Figure 1 shows the development of disposable rate of unemployment (%) in Slovakia during the period from January 1999 till May 2013. The series is non stationary showing stochastic trend, stochastic seasonality and non constant variability. At first we will use seasonal and non seasonal differences to adjust disposable rate of unemployment from seasonality and trend. We assume, such transformation is enough to create stationary series $z_t = (1-B)(1-B^{12})DRU_t$ given on Figure 2, for which the preliminarily ARIMA model we will look for. Transformation *z* in the computer outputs is abbreviated as ddsDRU.

Figure 1 Disposable Rate of Unemployment (%) in Slovakia, Jan 1999 – May 2013



Source: Centre for Labour, Social Affairs and Family of SR

Figure 2 Adjusted rate of unemployment in SR, February 2000- May 2013



Source: Authors

Preliminarily identification of the model for adjusted rate of unemployment z_t was done by means of the analysis of its empirical autocorrelation (ACF) and partial autocorrelation (PACF) functions. Because the first two autocorrelation coefficients and also the autocorrelation coefficient of order twelve are statistically significant at 5 % level of significance ($r_1 = 0{,}455$, $r_2 = 0{,}331$ and $r_{12} = -0{,}408$) we identify ARMA(2, 0)(0, 1)12. Estimation results of the model is given in Table 1.

Autocorrelation function shows that residuals of the estimated model for transformation z = ddsDRU are non correlated, but their squares are partially autocorrelated of order 5, because $\hat{\phi}_{5,5} = 0{,}223$ is statistically significant at 5 % level of significance.

Table 1 Estimated model for z = ddsDRU

Dependent Variable: D(MN,1,12)
Method: Least Squares
Sample (adjusted): 2000M02 2013M05
Included observations: 160 after adjustments
Convergence achieved after 12 iterations
Backcast: 1998M12 2000M01

Variable	Coeff.	Std. Error	t-Statistic	Prob.
MA(1)	0.444932	0.075706	5.877088	0.0000
MA(2)	0.247538	0.075897	3.261503	0.0014
SMA(12)	-0.737280	0.046297	-15.92507	0.0000

R-squared	0.450504	Mean dependent var	-0.013375
Adjusted R-squared	0.443504	S.D. dependent var	0.425223
S.E. of regression	0.317211	Akaike info criterion	0.560073
Sum squared resid	15.79779	Schwarz criterion	0.617733
Log likelihood	-41.8058	Durbin-Watson stat	1.938387

Inverted MA Roots	.97	.84+.49i	.84-.49i	.49-.84i
	.49+.84i	-.00-.97i	-.00+.97i	-.22+.45i
	-.22-.45i	-.49-.84i	-.49+.84i	-.84+.49i
	-.84-.49i	-.97		

Source: Authors´ calculations

Figure 3 Actual and fitted values, residuals of the model ARMA(0,2)(0,1)12 for ddsDRU

Source: Authors

Figure 3 that plots the actual and fitted values of the series and the residuals of the estimated model gives reason to assume that ARCH effect is present in the data. Therefore, we test its presence by the Lagrangean multiplier test. The test output is presented in Table 2. It can be seen that ARCH effect of order 5 is statistically significant and that ARCH effect is statistically significant at the level of 6 %. Given that 5 parameters represents rather high number for ARCH model, we will estimate GARCH (1,1) model instead. Estimation results are depicted in Table 3.

Table 2 LM test for residuals of the model ARMA(0,2)(0,1)12 of ddsRU

ARCH Test:

F-statistic	2.216731	Probability	0.055549
Obs*R-squared	10.73168	Probability	0.056967

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Sample (adjusted): 2000M07 2013M05
Included observations: 155 after adjustments

Variable	Coeff.	Std. Error	t-Statistic	Prob.
C	0.077858	0.026185	2.973346	0.0034
RESID^2(-1)	0.119537	0.079820	1.497585	0.1364
RESID^2(-2)	-0.02080	0.080339	-0.258838	0.7961
RESID^2(-3)	-0.05541	0.080207	-0.690845	0.4907
RESID^2(-4)	-0.03842	0.080299	-0.478459	0.6330
RESID^2(-5)	0.225557	0.079798	2.826592	0.0054

R-squared	0.069237	Mean dependent var	0.101236
Adjusted R-squared	0.038003	S.D. dependent var	0.250734
S.E. of regression	0.245923	Akaike info criterion	0.070345
Sum squared resid	9.011247	Schwarz criterion	0.188155
Log likelihood	0.548277	F-statistic	2.216731
Durbin-Watson stat	2.004780	Prob(F-statistic)	0.055549

Source: Authors

Table 3 Estimation of the model ARMA(0,2)(0,1)+GARCH(1,1) for ddsDRU

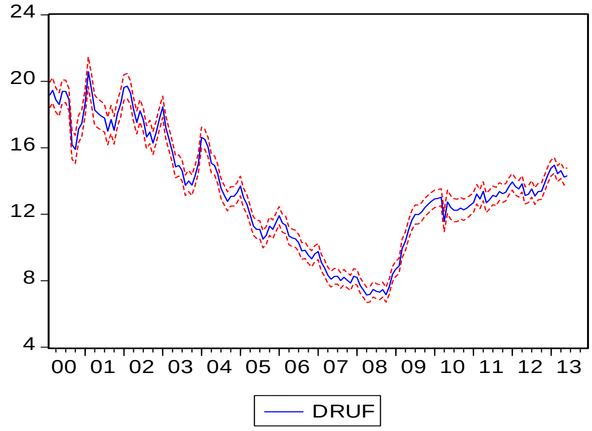
Dependent Variable: D(RU,1,12)
Method: ML – ARCH (Marquardt) - Normal distribution
Sample (adjusted): 2000M02 2013M05
Included observations: 160 after adjustments
Convergence achieved after 50 iterations
MA backcast: 1998M12 2000M01, Variance backcast: ON
GARCH = C(4) + C(5)*RESID(-1)^2 + C(6)*GARCH(-1)

	Coeff.	Std. Error	z-Statistic	Prob.
MA(1)	0.519390	0.057018	9.109287	0.0000
MA(2)	0.325511	0.072995	4.459352	0.0000
SMA(12)	-0.60800	0.046187	-13.14648	0.0000
		Variance Equation		
C	0.003073	0.001861	1.651715	0.0986
RESID(-1)^2	0.036849	0.018777	1.962444	0.0497
GARCH(-1)	0.918628	0.036682	25.04333	0.0000
R-squared	0.429525	Mean dependent var		- 0.013375
Adjusted R-squared	0.411003	S.D. dependent var		0.425223
S.E. of regression	0.326343	Akaike info criterion		0.526449
Sum squared resid	16.40093	Schwarz criterion		0.641768
Log likelihood	- 36.11594	Durbin-Watson stat.		2.163646
Inverted MA Roots	.96	.83-.48i	.83+.48i	.48+.83i
	.48-.83i	.00+.96i	-.00-.96i	-.26+.51i
	-.26-.51i	-.48+.83i	-.48-.83i	-.83-.48i
	-.83+.48i	-.96		

Source: Authors

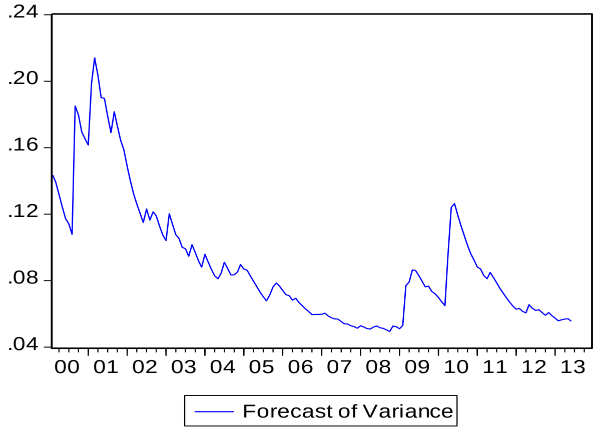
The results presented in Tale 3 are plotted at Figure 4, which shows the fitted values with the 95 % confidence interval for disposable rate of unemployment given by the model ARIMA(0,1,2)(0,1,1)12 together with the predicted values of variance given by the model GARCH(1,1). The mean characteristics of goodness of fit show RMSE = 0,32 percentage points, MAE = 0,22 percentage points, and MAPE = 1,65 per cent. Theil coefficient of inequality shows that it could be used for forecasting.

Figure 4 Fitted values with 95 % confidence interval and fitted values of variance given by the models ARIMA(0,1,2)(0,1,1)12 + GARCH(1,1) together with the measures of goodness of fit



Forecast: MNF
Actual: MN
Forecast sample: 1999M01 2013M12
Adjusted sample: 2000M02 2013M12
Included observations: 160

Root Mean Squared Error	0.320165
Mean Absolute Error	0.221344
Mean Abs. Percent Error	1.685475
Theil Inequality Coefficient	0.011679
Bias Proportion	0.000588
Variance Proportion	0.038505
Covariance Proportion	0.960908



The forecast of disposable rate of unemployment is precise for one month ahead, because of simple model with 2 parameters in model mean and 1 parameter in model

GARCH. The forecast for June 2013 given by the model is 14,32 (mean) ± 0,23 (standard deviation). Actual value of disposable rate of unemployment in June 2013 was 14,25 %.

Conclusions

The findings of our analysis showed that the series of the rate of unemployment is heteroscedastic. We found the appropriate model to estimate generalized conditional autoregressive model of order GARCH(1, 1), which allows for the estimation of both the mean and the standard deviation for the time series. This finding will help us to forecast not only the conditional mean of the series, but also its conditional standard deviation. The best length of horizon is one month ahead.

We already have experience with forecasting of the rate of unemployment with during the period from January 2001 till September 2013 by the ARIMA(0,2,1)(0,1,1)12 model. The forecasts are computed for the 6 months ahead and we have found out that the precision of such forecasts is the best only for one month ahead. Because 95 % confidence interval of the mean is very wide for larger horizon, we conclude that it is possible to forecast conditional mean and conditional variance only one month ahead and to update the forecasts permanently with similar technique.

Literature

Arlt, J., Arltová, M.: Finanční časové řady. Praha: GRADA Publishing, 2003

Bollerslev, T.: Generalized autoregressive conditional heteroscedasticity. In: Journal of Econometrics 31, (1986), s. 307-327.

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

Bulletin Prognostického ústavu SAV, at www.prog.sav.sk

Engle, R. F.: Autoregressive conditioinal heteroscedasticity with estimates for the variance of United Kingdom Inflation. Econometrica 52 (1982), s. 987-1007.

ILO Short Term Indicators of Labour Market Database, at www.ilo.org

Lubyová, M. and E. Rublíková (2012), Modelovanie a prognózy vývoja vybraných ukazovateľov trhu práce v SR, Prognostické práce, 4, 2012, č. 2

Národná banka Slovenska, Strednodobá predikcia 3Q 2012, NBS Bratislava

OECD Working Group on Short Term Economic Prospects, at www.oecd.org

OECD (2012), Economic Survey of the Slovak Republic, December 2012, OECD Publishing

Radvanský, M., Kvetan, V. a L. Pániková (2010), Makroekonomická prognóza vývoja slovenskej ekonomiky so zameraním na vývoj dopytu po práci. Marek Radvanský, Vladimír Kvetan, Lucia Pániková. In Expertízne štúdie EÚ SAV [online], 2010, č. 7, s. 1-107. ISSN 1337-0812.
http://www.ekonom.sav.sk/uploads/journals/ES07.pdf

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

Workie Tiruneh Menbere, Štefánik Miroslav, Bujňáková Tatiana, Horvát Peter, Lichner Ivan: Predvidanie potrieb trhu prace v SR: Teoreticke vychodiska a empiricke vysledky, EKONOM, 2012 (v tlači), výstupy online na

http://www.ekonom.sav.sk/sk/projekty/predvidanie-potrieb-slovenskeho-trhu-prace-v-strednodobom-horizonte-do-roku-2025-p129