LABOUR MARKET MODELLING IN THE CZECH REPUBLIC

Radmila Stoklasová

Abstract: *The article deals with labour market modelling in the Czech Republic using regression model. The regression equation shows the behaviour of the unsatisfied labour supply determined by education and age structure of the unemployed, registered length of unemployment, available jobs average wage. In the second part of article is used prediction by Box-Jenkinson methodology. On the basis of the verified model a prediction for the first quarter year of 2015 is worked out and its success assessed. Population ageing will considerably influence economic and social structure of the population. The labour market will change completely. The future development of labour market will face a challenge of reduction of the negative initial demographic conditions. The main problem of the labour market in the Czech Republic is the exclusion of all groups of population due to the lack of flexible jobs, mainly part-time jobs (for example students, mothers with children and elderly people).*

Keywords: *Labour market modeling, Labour market supply, Regression model, Correlation analysis, Box-Jenkinson methodology.*

JEL Classification: *C32, C51.*

Introduction

In the first part of this article will be followed an investigation of the past and present development behaviour of labour supply using regression analysis. Pissarides estimates model for employment, real wages and Beveridge curve, we can see in [8]. For estimation of equation in this paper was used the following time series [9]: the number of unemployed registered at the labour office, the number of registered unemployed according to the educational structure, the number of registered unemployed according to the age structure, the number of registered unemployed according to the length of unemployment, the number official free jobs at the labour offices and average nominal wage.

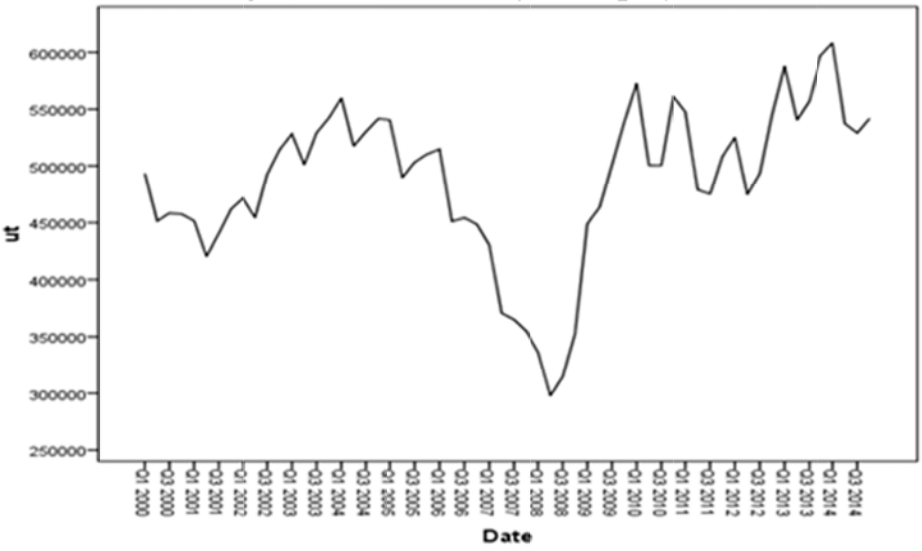
The second part is related to analysis time series the number of unemployed registered at the labour office using by Box-Jenkins methodology. This part consists of identifying the model and its validation of model quality. Data that has been used can be found on the Internet at [7] and [10]. Calculations are conducted using SPSS statistical software like in [6] and [11].

The third part is devoted to the second equation model. The second equation model the labour demand influenced by demanded education structure, registered available jobs, gross domestic product (GDP), the unemployed and average wage.

1 The first equation model for the labour market

In the next figure shows the course of the time series the number of unemployed registered at the labour office. The minimum is 297 880 unemployed and it was at the second quarter 2008 and the maximum is 608 315 unemployed and it was at the first quarter 2014.

Fig. 1: The number of unemployed



Source: [7]

From the regression model we can express following linear regression function:

$$u_t = f(u_Z; u_s; u_v; u_1; u_2; u_3; u_k; u_d; v_t, w_t), \tag{1}$$

where:

u_t the number of unemployed registered at the labour office,

$u_{EDU,t}$ the number of registered unemployed according to the educational structure (u_Z, u_s, u_v - primary, secondary and university degree education),

$u_{AGE,t}$ the number of registered unemployed according to the age structure (u_1, u_2, u_3 – till 19 years old, from 20 to 49 years old, more than 50 years old),

$u_{TIME,t}$ the number of registered unemployed according to the length of unemployment (u_k , u_d - time up to 1 year and long time unemployed lasting more than 1 year),

v_t official free jobs at labour offices, w_t average nominal wage.

For estimating labour market model the statistical software SPSS has been used. In deciding on the best set of explanatory variables for each regression model we follow the stepwise regression [3]. In this method we proceed by introducing the acceptable explanatory variables and rejecting them at a time. The decision to odd or drop a variable is made on the basis of the contribution of that variable to the explained sum of squares, as judged by the F test for independent variable entering or removing at each step according to the tolerance criterion.

The estimated regression relation:

$$u_t = -1423 + 0.019u_v + 0.998u_2 + 1.002u_3 + 0.996u_1 + 0.002u_d . \tag{2}$$

The follow table shows that all regression coefficients are statistically significant on the 5% significance level.

Tab. 1: Regression coefficients

	B	Std.Error	t	Sig.
Constant	-1423	52.092	-27.33	0.008
u_v	0.019	0.001	19.230	0.000
u_2	0.998	0.000	4248.760	0.000
u_3	1.002	0.000	5147.320	0.000
u_1	0.996	0.001	996.454	0.000
u_d	0.002	0.000	5.787	0.000

Source: [own calculations]

Based on this model, we can estimate the number of unemployed in the first quarter of 2015. The actual values are:

$$u_v = 63050; \quad u_2 = 297996; \quad u_3 = 212083; \quad u_1 = 15236; \quad u_d = 230941.$$

After substituting into the regression equation we get u_t = 525319 and the actual number of unemployed in the first quarter of 2015 was 525315, we can see [7].

The assumptions regression model are:

a) Correctly specified model,

b) the mean value of the residual component is equal to 0,

c) the residual component has a constant variance,

d) the individual components of the residual vector are uncorrelated,

e) residual component has a normal distribution.

The following table confirms the fulfillment of assumption d).

Tab. 2: Autocorrelations of residual component

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic		
			Value	df	Sig.[b]
1	0.299	0.119	27.211	1	0,000
2	0.281	0.111	27.118	2	0,000
3	0.195	0.114	28.224	3	0,000
4	-0.215	0.134	31.512	4	0,000
5	-0.075	0.132	32.176	5	0,000
6	-0.275	0.125	36.742	6	0,000
7	-0.242	0.118	40.689	7	0,000
8	-0.175	0.111	42.959	8	0,000

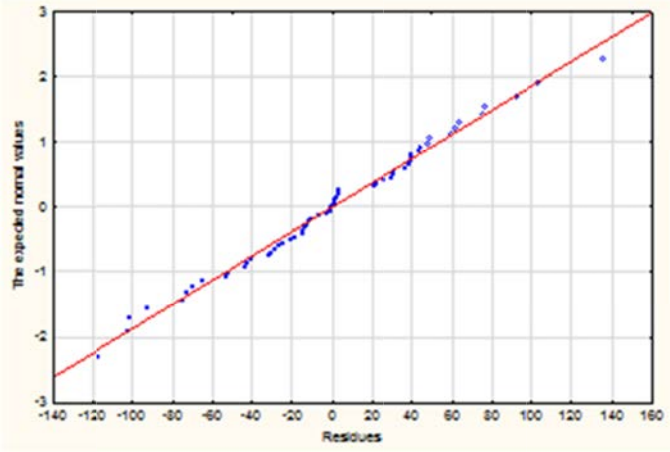
a. The underlying process assumed is independence (white noise).
b. Based on the asymptotic chi-square approximation.

Source: [own calculations]

The assumption has (that the residual component is not correlated) been verified Durbin-Watson test, DW = 1.95.

The normal distribution of random component is confirmed in the following graph and result Shapiro-Wilk test. Residue: SW-W = 0.9902; p = 0.9130. The test is not statistically significant. There are no statistically significant deviations from normality.

Fig. 2: Normal p-graph of residues



Source: [own calculations]

The following table shows the level of dependence of individual variables.

Tab. 3: Pearson Correlation

	u_z	u_s	u_v	
u_t	0.686**	0.306*	0.488**	
	u_1	u_2	u_3	
u_t	0.264*	0.783**	0.615**	
	u_k	u_d	v_t	w_t
u_t	0.837**	0.695**	-0.837**	0.184

Source: [own calculations]

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

The largest values of the correlation coefficient are calculated for variables:

$$u_k = 0.837;\ u_2 = 0.783;\ u_d = 0.695.$$

This means that the number of unemployed depends mainly on the length of unemployment, see [2], [4]. Indirect dependence logically based on the number of unemployed and the number of vacancies: $v_t = 0.837$. The only variable was not shown to the dependence of the average wage [5].

2 Box-Jenkins methodology

Behavior and prediction of time series is calculated by using Box-Jenkins methodology (ARIMA model). ARIMA (*p, d, q*) model is a complex linear model composed of three sub-sections: AR – Autoregressive, I - Integrative, MA - Moving Average. These models are extremely flexible, they are relatively difficult to calculate and to understand the results, are demanding quality and number of measured data (assumed to be at least 50 measurements or observations). The investigated time series values has 60 observations, so that the ARIMA model can be used, see [1].

In the SPSS output in the ARIMA model for each parameter there are calculated their significance. According to these values it is possible to determine whether the parameters in the model to include or not. When calculating the number of models that describe the behavior of the series about the same, then we select a model in which the AIC (Akaik´s information criterion), respectively SBC (Schwartz-Bayesian criterion) is minimal and Log Likelihood maximal. Finally, verify that residual component is white noise, i.e. the sequence of random variables normally distributed with zero mean and constant standard deviation, see [1].

The ARIMA model assumes interdependence among neighboring values. If the process contains even seasonal fluctuations, we can also expect dependencies between variables in different seasons.

This process is called **SARIMA (*p,d,q*) (*P,D,Q*)$_s$** , where

p is order of the process AR, *q* is order of the process MA, *d* is order of the difference, *P* is order of the seasonal process AR, *Q* is order of the seasonal process MA, *D* is order of the seasonal difference, *s* is the length of the seasonal period.

Model can be expressed by the formula:

$$\Phi_P\left(B^s\right)\varphi_p(B)(1-B)^d\left(1-B^s\right)^D y_t = \theta_q(B)\Theta_Q\left(B^s\right)a_t . \tag{3}$$

2.1 Identification of the model

Based on the autocorrelation functions we can chose models: SARIMA (1,1,0) $(1,1,0)_4$ or model SARIMA (1,1,0) $(0,1,0)_4$. Because Akaik‘s information criterion as well as Schwartz-Bayes criterion are less for the model SARIMA (1,1,0) $(1,1,0)_4$, we chose that model.

SPSS output for model SARIMA (1,1,0) $(1,1,0)_4$

FINAL PARAMETERS:

Log likelihood	-629.79572
AIC	1265.5914
SBC	1271.6134

Variables in the Model:

	B	SEB	T-RATIO	APPROX. PROB.
AR1	0.591	0.11853	4.9897870	0.00000714
SAR1	-0.254	0.14234	-1.7909735	0.04811875
CONSTANT	-394.11	6042.75173	-0.0652203	0.94824866

Both coefficients are statistically significant at a significance level of 5%.

General form of the model is:

$$\Phi_1\left(B^4\right)\varphi_1(B)(1-B)^1\left(1-B^4\right)^1 y_t = a_t . \tag{4}$$

After adjustment equation and substituting the estimated values into (4) se obtain the following equation, which describes the dynamics of the investigated time series:

$$y_t = 1.591y_{t-1} - 0.591y_{t-2} + 0.746y_{t-4} - 1.186y_{t-5} - 0.741y_{t-6} + \\ + 0.254y_{t-8} - 0.41y_{t-9} + 0.15y_{t-10} + a_t \quad (5)$$

2.2 Diagnostic model control

The assumption has (that the residual component is not correlated) been verified Durbin-Watson test, DW = 1.72. The next table shows values of autocorrelations of residual component and the assumption about not correlated of residual has been verified.

Tab. 4: Autocorrelations of residual component

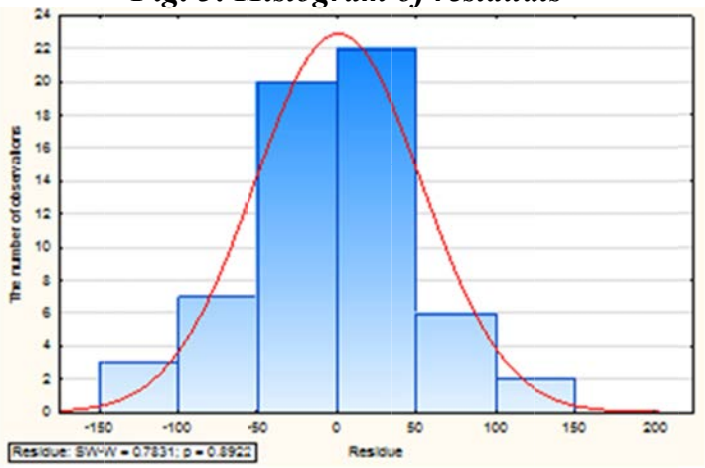
Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic		
			Value	df	Sig.[b]
1	0.311	0.126	23.566	1	0,000
2	0.231	0.125	26.998	2	0,000
3	0.185	0.124	29.244	3	0,000
4	0.205	0.123	32.048	4	0,000
5	-0.065	0.122	32.336	5	0,000
6	-0.285	0.120	37.933	6	0,000
7	-0.240	0.119	41.987	7	0,000
8	-0.155	0.118	43.701	8	0,000
9	-0.258	0.117	48.549	9	0,000
10	-0.212	0.116	55.777	10	0,000

a. The underlying process assumed is independence (white noise).
b. Based on the asymptotic chi-square approximation.

Source: [own calculations]

The normal distribution of random component is confirmed in the following graph and result Shapiro-Wilk test. Residue: SW-W = 0.7831; p = 0.8922. The test is not statistically significant. There are no statistically significant deviations from normality.

Fig. 3: Histogram of residuals

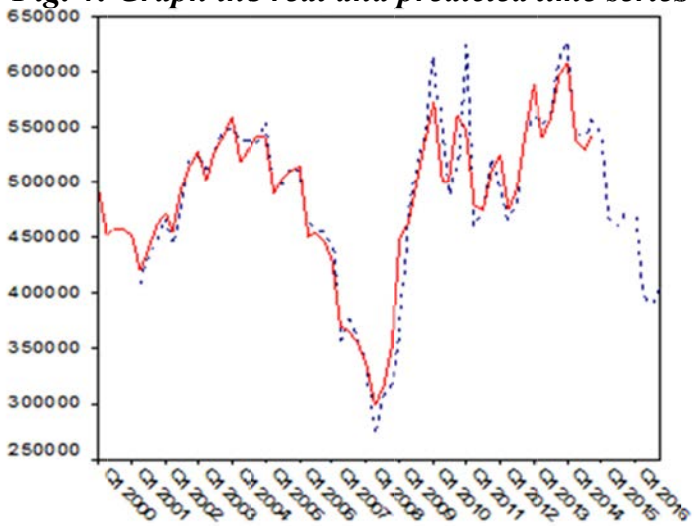


Source: [own calculations]

2.3 Short – time prediction by estimated model

Graph shows the real and predicted time series.

Fig. 4: Graph the real and predicted time series



Source: [own calculations]

Point and interval forecasts (95%) of the number unemployed are shown in Table.

Tab. 5: Prediction of the number unemployed

Quarter	Point prediction	Interval prediction
1.Q/2015	543 283	(496 235 ; 590 331)
2.Q/2015	467 371	(378 249 ; 556 493)
3.Q/2015	458 957	(329 833 ; 588 081)
4.Q/2015	474 527	(308 275 ; 640 780)
1.Q/2016	475 837	(253 742 ; 697 932)
2.Q/2016	399 381	(117 367 ; 681 395)
3.Q/2016	389 807	(49 287 ; 730 327)
4.Q/2016	403 738	(7 678 ; 799 797)

Source: [own calculations]

In 2015 we can expect the highest number of the unemployed during the first quarter. The lowest number of the unemployed is expected in the third quarter 2016. The reason may be seasonal work in construction.

3 The second equation

The second equation is related to the labour demand that is limited to unsatisfied labour demand registered at the labour offices. This unsatisfied labour demand is expressed by free jobs or its change is determinated by the educational structure of free jobs, by the regional economy output, registered unsatisfied labour supply and average wage. Regression exponential equation can be re-written as linear:

$$\ln v_t = g\left(\ln v_Z ; \ln v_s ; \ln v_v ; \ln v_{GDP,t} ; \ln v_{GDP,t-1} ; \ln u_t ; \ \ln w_t\right), \qquad (6)$$

where:

$v_{EDU,t}$ the number of officially free jobs according to the required educational structure v_z, v_s, v_v - primary, secondary and university degree education),

$v_{GDP,t}$ gross domestic product in CZK.

Concerning demand (6) the positive influence of educational structure of free jobs is supposed. The economic growth expressed by the change of GDP should encourage regional economy and stimulate higher number of free jobs (positive relation). Based on Beveridge´s curve we suppose negative influence of variable ln*u* on ln*v*. If we regard the nominal wage growth in the wider context than the average wage change should positively influence the change of free jobs number.

Tab. 6: Regression coefficients

	B	Std.Error	t	Sig.
Constant	2.191	0.293	7.478	0.008
ln v_z	0.614	0.063	9.746	0.000
ln v_s	0.312	0.075	4.16	0.000
ln $v_{GDP,\ t-1}$	0.421	0.068	6.191	0.000

Source: [own calculations]

$$\ln v_t = 2.191 + 0.614 \ln v_z + 0.312 \ln v_s + 0.421 \ln_{GDP,t-1}; \quad R^2_{adj} = 0.981; \quad D = 2.015 . \quad (7)$$

The last estimate of the equation for the labour demand in the Czech Republic can be statistically and econometrically successfully verified and may be used for the prediction.

Tab. 7: Prediction of the unsatisfied labour demand in the Czech Republic

Quarter	Actual value	Point prediction	Interval prediction
1.Q/ 2015	66 111	68 389	(63 176; 73 121)
2.Q/ 2015	91 140	90 286	(86 102; 94 987)
3.Q /2015	103 477	101 357	(95 648; 107 142)

Source: [own calculations]

The final summary of the success of the prediction of the unsatisfied labour demand expressed by the number of available jobs shows that the actual values of available jobs registered at job centres are inside the 95% confidential interval.

Conclusion

Statistical verification has been successful in both cases. From the viewpoint of economic interpretation the results of the estimated labour market model in the Czech Republic in 2000 – 2014 can be summed up as follows: The change in the number of the unemployed has been influenced by the change in the age structure of applicants – namely between 20 and 49 years of age and more than 50 years old. The change of the unsatisfied labour demand has been influenced by education structure demanded for available jobs, principally for workers with basic and secondary education.

With unemployment facing many countries, therefore it is important to capture the trend in unemployment. Box-Jenkins model used is a highly flexible, so it is highly likely, that unemployment will develop according to the above model if there is absence of government intervention that would significantly change this trend.

Acknowledgement

This paper was supported by the Ministry of Education, Youth and Sports Czech Republic within the Institutional Support for Long-term Development of a Research Organization in 2015.

References

[1] ARLT, J., ARLTOVÁ, M. *Ekonomické časové řady*. Praha: Professional Publishing, 2009. ISBN 978-80-86946-85-6.

[2] HANČLOVÁ, J. A Multivariate Labour Market Model in the Czech Republic. In: *Proceedings of the International Conference on Quantitative Methods in Economics*. Ekonomická univerzita Bratislava, 2000, 43-52. ISBN 80-245-0057-4.

[3] HANČLOVÁ, J. Vliv vybraných charakteristik struktury trhu práce na vývoj nezaměstnanosti v ČR. *Informační bulletin České Statistické Společnosti*, Česká statistická společnost, 2000, 3-10.

[4] HANČLOVÁ, J. *Modelování a klasifikace regionálních trhů práce*. Ostrava: VŠB – Technická univerzita, 2002. ISBN 80-248-0220-1.

[5] JAHODA, M., LAZARSFELD, P., ZEISEL, H. *Marienthal: Sociologie komunity nezaměstnaných*. Brno: Masarykova univerzita, 2013. ISBN 978-80-210-6226-9

[6] MONTGOMERY, A. L., ZARNOWITZ, V., TSAY, R. S., & TIAO, G. C. Forecasting the US unemployment rate. *Journal of the American Statistical Association,* 1998, 478-493.

[7] Ministerstvo práce a sociálních věcí. Informace o počtu nezaměstnaných v České republice, duben 2015. Available from WWW: <http://www.portal.mpsv.cz/sz/stat/nz/qrt>

[8] PISARIDES, C.: The Labour Market and the Euro. *Cyprus Economic Policy Review 2*. 2008. 3-9. ISSN 1450-4561.

[9] POŠTA, V., MACÁKOVÁ, L., PAVELKA, T. *Strukturální míra nezaměstnanosti v ČR*. Praha: Management Press, 2015.ISBN 978-80-7261-296-3.

[10] *Statistická ročenka trhu práce v České republice 2014*. Praha: Ministerstvo práce a sociálních věcí – odbor trhu práce, 2015. ISBN 978-80-7421-093-8.

[11] STOKLASOVÁ, R. Model of the unemployment rate in the Czech Republic. *In Proceedings of 30th International Conference on Mathematical Methods in Economics*. Silesian University in Opava, School of Business Administration in Karviná, 2012, 836-841. ISBN 978-80-7248-779-0.

Contact Address

Mgr. Radmila Stoklasová, Ph.D.
School of Business Administration in Karviná, Silesian University in Opava,
Department Informatics and Mathematics
University Square 1934/3, 733 40, Karviná, Czech Republic
Email: stoklasova@opf.slu.cz
Phone number: +420 596 398 283

Received: 22. 08. 2015
Reviewed: 22. 09. 2015, 24. 09. 2015, 23. 10. 2015
Approved for publication: 17. 12. 2015