ANALYSIS OF THE RELATIONSHIP BETWEEN UNEMPLOYMENT RATE, EDUCATION LEVEL AND SELECTED ECONOMIC INDICATORS IN THE CZECH REPUBLIC 2005-2017

Dagmar Blatná

Abstract

The unemployed are defined as jobless people who are actively seeking employment, currently available for work. The aim of the present paper is to analyse the development of the unemployment rate in groups of people broken down by educational attainment and other selected socio-economic factors in the Czech Republic over the period 2005–2017. Data were drawn from regular Eurostat reports entitled "Unemployment rates of the population aged 25–64 by educational attainment level"; the categorization of educational attainment being based on the International Standard Classification of Education (below upper secondary, upper secondary non-tertiary and tertiary [ISCED 0-2, 3-4 and 5-8, respectively]). The main focus of the study was on finding regression relations between the unemployment rate time series. The regression analysis was carried out applying multivariate non-stationary time series methodology. The unemployment rates of tertiary school graduates proved to be significantly the lowest of all the groups examined throughout the whole monitored period.

Key words: unemployment rate, educational attainment level, time series regression analysis, stationarity, non-stationarity

JEL Code: E24, C32

Introduction

The unemployment rate is one of the most frequently observed labour market and socio-economic indicators, measuring the percentage of unemployed 25–64-year-old persons in the labour force (i.e. people who are employed or seeking a job). The unemployed are generally defined as people without jobs who make efforts to find employment and are currently available for work. They are the persons who (a) had no employment during the survey

reference week, (b) are currently available for work (within the fortnight after the reference week), (c) are actively looking for a job (i.e. have taken specific steps over a four-week period to seek [self-]employment) or who have found a job to start working later (within at most three months). The unemployment rate by educational attainment level is based on the EU Labour Force Survey data. According to Eurostat explanation, this indicator shows the "probability" of being without a job for those who would like to have one, broken-down by educational stage reached, providing a measure of difficulties that people with different levels of education face in the labour market.

In line with the International Standard Classification of Education (ISCED), educational attainment falls into nine stages. For analytical purposes, they are grouped into the following three categories:

ISCED levels 0-2: less than primary, primary and lower secondary education,
ISCED levels 3 and 4: upper secondary and post-secondary non-tertiary education and
ISCED levels 5-8: tertiary education.

The objective of the present paper is to study the development of the unemployment rate by the three levels of education and other social and economic factors in the Czech Republic over the period 2005–2017, focusing on its time series regression analysis.

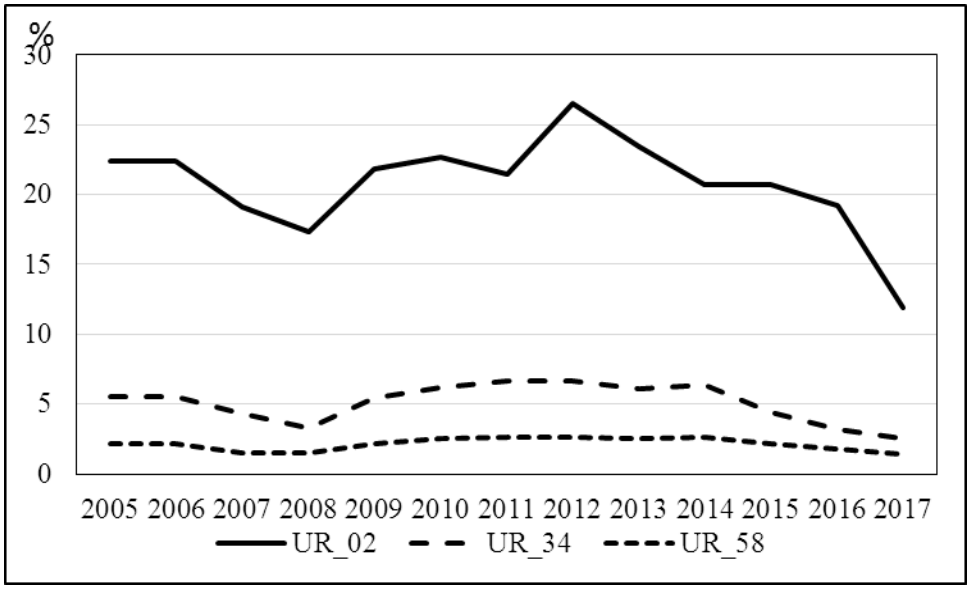
1 An analysis of unemployment rates by educational attainment in the Czech Republic (2005–2017)

The analysis is based on the 2005–2017 EU Labour Force Survey data and indicator definitions adopted from the Eurostat database. Statistical calculations were done using EViews 9, Statgraphics and Excel software packages.

As can be seen from Figure 1, the employment rates vary according to the level of education, Table 1 presenting the basic statistics. It is obvious that the unemployment rate of people who attained only lower secondary education (UR_02) was the highest throughout the whole research period.

Statistical significance was tested by the non-parametric Kruskall-Wallis test of the equality of medians. The ANOVA test could not have been used because of the inequality of variances (Bartlett's test = 2.8290, p-value = 0.0001).

Fig. 1: Unemployment rates of population aged 25–64 by educational attainment in the Czech Republic (2005–2017)



Source: Eurostat data, author's own elaboration

Tab. 1: Unemployment rates by educational attainment – basic statistics

Statistics	UR_02	UR_34	UR_58
Average.	20.7385	5.0923	2.1385
Median	21.5	5.5	2.2
Standard deviation	3.4923	1.4192	0.4464
Variance	12.1959	2.0141	0.1992
Skewness	-1.1808	-0.6095	0.6330
Kurtosis	2.9684	-0.9427	-1.1026
Minimum	11.9	2.5	1.4
Maximum	26.5	6.7	2.6
Range	14.6	4.2	1.2
Coefficient of variation (%)	0.1684	0.2787	0.2087
Jarque-Bera test of normality / p-value	3.5806/0.1664	6.2167/0.5443	1.3904/0.4990

Source: data Eurostat, author's own calculations

2 Methodology

Using time series data during the regression model estimation, it is necessary to know whether the variables are stationary (time series with an autoregressive model of zero order I(0)) or non-stationary (time series generated by an autoregressive process of order one I(1)).

In the regression analysis, the time series of both variables must have the same order of integration. If the series are not of the same order, they do not show any statistical relationships. A classical regression model can be employed when the time series analysed are

zero-order stationary ones. The relation between two non-stationary time series exists only if their linear combination is stationary. In this case, there is a common stochastic trend in both time series and the so-called cointegration regression occurs, the model representing a long-term relationship.

When the test of a non-systematic component of the linear combination of two non-stationary time series does not reject the non-stationarity, only a spurious regression occurs, indicating the relationships even when there are none; for details, see Arlt a Arltová (2009), Hušek (2007), Elliot at al. (1996), Granger and Newbold (1974) and Phillips (1987). When the non-systematic component is auto-correlated, the Autoregressive Distributed Lag (ADL) model with time-shifted variables is used for estimation; see Arlt and Arltová (2009), Hušek (2007).

For testing (non-)stationarity, the unit root tests of the autoregressive parameter ϕ_1 are carried out. In the present paper, the Augmented Dickey-Fuller (ADF) test with or without a drift and trend term, respectively, is applied, taking into account the possibility that the process is higher-order autoregressive. The ADF tests verify the null hypothesis H_0: $\phi_1 = 1$ for non-stationary I(1) time series against H_1:$|\phi_1| < 1$ for stationary I(0) time series; for more details, see Arlt and Arltová (2009), Hušek (2007), Caner and Kilian (2001), Dickey and Fuller (1979, 1981), Elliot *at al.* (1996) or Phillips (1987).

Validation of the calculated regression model is performed using diagnostic tests of the model's non-systematic component. Testing for the error autocorrelation in the regression model, Breusch-Godfrey Serial Correlation LM test was conducted with the null hypothesis that there is no serial correlation of any order up to p; see Breusch and Godfrey (1986). To verify the normality, the Jarque-Bera test is undertaken. This is a goodness-of-fit test of whether the sample data show the skewness and kurtosis matching a normal distribution; see, e.g. Jarque and Bera (1980). Autoregressive conditional homoscedasticity in the residuals was verified by ARCH(1) LM test; see Engle (1982).

3 Unit root test of unemployment rates and other selected indicators

The regression analysis of multivariate time series was done, unemployment rates (UR) representing the dependent variables. Table 2 lists URs and some socio-economic indicators selected as explanatory variables, displaying the results of the ADF test and relevant p-values.

Tab. 2: Unit root test of unemployment rates and selected indicators

Indicator	Abbr.	t_{ADF}	p-value	Stationarity/ Non-stationarity
Unemployment rate by education 0-2	UR_02	-0.9872	0.2711	N
Unemployment rate by education 3-4	UR_34	-1.4296	0.5293	N
Unemployment rate by education 5-9	UR_58	-1.7899	0.3484	N
Long-term unemployment rate	LUR	-1.9929	0.0482	S
Employment rate	ER	1.0098	0.9049	N
Employment growth	EG	-1.7847	0.0716	N
Part-time employment rate	PTER	-7.7446	0.0004	S
Real GDP per capita growth rate	GDPG	-1.9848	0.0460	S
Social benefits	SB	-1.4541	0.5210	N
GERD	GERD	-3.9023	0.0576	N
Gross households saving rate	HSR	-1.5284	0.4857	N
Inflation rate (HICP)	IR	-7.1010	0.0014	S
Average salary growth	SG	-0.5044	0.4762	N
Minimum wage growth	MWG	-2.8321	0.2227	N
Households with broadband access	IA	-6.4647	0.0006	S
People living in households with very low work intensity	LWI	-3.6671	0.0777	N

Source: Eurostat data, author's own calculations

The Augmented Dickey-Fuller test identified non-stationarity of unemployment rate time series by educational attainment level for all groups. In the regression analysis, the time series of variables must be of the same order integrated process. It is therefore obvious that there is no relationship between the unemployment rate in all groups divided by the educational level and indicators whose time series are stationary – i.e. the rates of inflation, part-time employment, long-term unemployment and real per capita GDP growth, and the share of households with broadband access, respectively. Thus, only non-stationary indicators were included in the regression analysis as explanatory variables.

4 Time series of unemployment rates and selected indicators: a regression analysis (2005–2017)

4.1 The unemployment rate of people with lower secondary education

The ADF test does not reject the non-normality of unemployment rate time series for people with lower secondary education or less (UR_02). Several regression models for UR_02 as the dependent variable having been calculated, the most statistically appropriate one with the selected indicators whose time series are non-stationary – including SB and LWI indicators (i.e. social benefit and the share of people living in households with very low work intensity) as explanatory variables – are presented in Table 3. This model explains 69 per cent of unemployment rate variability of people who attained only lower secondary level of education. Table 4 shows the outcomes of diagnostic tests confirming that the non-systematic

component has a normal distribution (Jarque-Bera test JB = 0.2349, p = 0.8892), is homoscedastic (ARCH = 1.4637, p = 0.2542) and not autocorrelated (Breuch-Godfray LM test F = 2.7513, p = 0.1232). As the ADF test rejects non-stationarity (t = -3.3717, p = 0.0029), the time series are cointegrated, the model representing a long-term relationship.

Tab. 3: Regression model UR_02 ~ SB + LWI

Variable	Coefficient	Std. error	t-statistic	Prob.
C	-57.6947	16.6670	-3.4616	0.0061
SB	4.8937	1.0945	4.4710	0.0012
LWI	2.4419	0.6619	3.6891	0.0042
R-squared	0.6908	F test	F= 11.1727	Prob(F) = 0.0028
Augmented Dickey-Fuller test statistic		t_{ADF}=-5.1828	p-value = 0.0001	

Source: Eurostat data, author's own calculations

Tab. 4: Diagnostic tests of UR_02 ~ SB + LWI model's non-systematic component

Test	Test statistic	Prob.	p-value
Breusch-Godfrey serial correlation LM test	2.7513	Prob. F(2,8)	0.1232
Normality test: Jarque-Bera	0.2349		0.8892
Heteroskedasticity test: ARCH test	1.4637	Prob. F(1,10)	0.2542

Source: Eurostat data, author's own calculations

4.2 The unemployment rate of people with upper- and post-secondary education

The ADF test does not reject the non-normality of unemployment rate time series for people with the upper secondary (ISCED 3) and post-secondary (ISCED 4) level of education (UR_34). The best regression model containing SB and ER (social benefits and employment rate) explanatory variables explains 77.46 per cent of unemployment rate variability of people who completed their secondary education. As it can be seen in Table 5, the unemployment rate of this group rises with increasing social benefits and declines with the growing employment rate. The time series are cointegrated, the model representing a long-term relationship because of the stationarity of the non-systematic component.

Tab. 5: Regression model UR_34~ SB + ER

Variable	Coefficient	Std. error	t-statistic	Prob.
C	17.8669	8.3985	2.1274	0.0593
SB	1.0464	0.3304	3.1669	0.0100
ER	-0.3546	0.0879	-4.0325	0.0024
R-squared	0.7746	F-test	F= 17.1851	Prob(F) = 0.0006
Augmented Dickey-Fuller test statistic		t_{ADF} = -2.2357	p-value = 0.0299	

Source: Eurostat data, author's own calculations

In terms of statistics, this model is acceptable, diagnostic tests proving that the non-systematic component is normally distributed (Jarque-Bera test t = 0.686, p = 0.8317), homoscedastic (ARCH = 0.0007, p = 0.9791) and not autocorrelated (Breuch-Godfray LM test F = 1.5813, p = 0.2638). Diagnostic test results yielded are presented in Table 6 below.

Tab. 6: Diagnostic tests of UR_34 ~ SB + ER model's non-systematic component

Test	Test statistic	Prob.	p-value
Breusch-Godfrey serial correlation LM test	1.5813	Prob. F(2,8)	0.2638
Normality test: Jarque-Bera	0.3686		0.8317
Heteroskedasticity test: ARCH test	0.0007	Prob. F(1,10)	0.9791

Source: Eurostat data, author's own calculations

4.3 The unemployment rate of people with tertiary education

The last group comprises unemployed people with tertiary education (ISCED levels 5-8, UR_58). Table 7 shows the regression model with AGS and ER explanatory variables (i.e. average salary growth and employment rate, respectively). The model explains 81.60 per cent of UR_58 time series dynamics, the F-test being significant (F = 22.4756, p = 0.0002). This model is acceptable from the statistical point of view. For diagnostic test outcomes, see Table 8.

The ADF test confirms that the non-systematic component (t = -4.0747, p = 0.0007) is stationary. This model expresses cointegration, thus also suggesting a long-term relationship. The unemployment rate of people who have tertiary education decreases with the overall employment decline and the increasing growth of average salary.

Tab. 7: Regression model UR_58 ~ ASG + EG

Variable	Coefficient	Std. error	t-statistic	Prob.
C	7.7951	1.7727	4.3973	0.0013
ASG	-0.1436	0.0265	-5.4201	0.0003
EG	-0.0696	0.0246	-2.8290	0.0179
R-squared	0.8180	F-test	F= 22.4756	Prob(F) = 0.0002
Augmented Dickey-Fuller test statistic		t_{ADF} = -4.0747	p-value = 0.0007	

Source: Eurostat data, author's own calculations

Tab. 8: Diagnostic tests of UR_58 ~ ASG + EG model's non-systematic component

Test	Test statistic	Prob.	p-value
Breusch-Godfrey serial correlation LM test	0.2882	Prob. F(2,8)	0.7571
Normality test: Jarque-Bera	0.5373		0.7644
Heteroskedasticity test: ARCH test	0.2518	Prob. F(1,10)	0.6267

Source: Eurostat data, author's own calculations

Conclusion

The unemployment rate is one of the most frequently reported labour market indicators of economic and social trends. The principal goal of the paper was to study the 2005–2017 unemployment rate of the Czech Republic's population divided according to educational attainment. The categorization of educational attainment is based on the International Standard Classification of Education (ISCED), grouping people together in groups of (a) below upper secondary, (b) upper secondary non-tertiary and (c) tertiary levels of education attained (ISCED 0-2, 3-4 and 5-8, respectively). The analysis showed that the unemployment rate is significantly dependent on the educational attainment level, UR_02 being the highest throughout the 2005–2017 period.

The Augmented Dickey-Fuller test identified the non-stationarity of unemployment rate time series for all groups of different educational levels. It can be concluded that there is no relationship between the unemployment rate in the groups broken down by the level of education and indicators whose time series are stationary – namely the rates of inflation, part-time employment, long-term unemployment and real GDP per capita growth, and the share of households with broadband access.

Several regression models of unemployment rates as dependent variables were created. The present study offers the best statistically significant models with selected indicators whose time series are non-stationary. The unemployment rate of people with both lower and upper secondary education (UR_02 and UR_34, respectively) increases with the growing social benefits and declines with an increasing employment rate. The unemployment rate of those who attained tertiary education (UR_58), on the other hand, decreases with a declining overall employment rate and with higher average salary growth. All the models presented imply cointegration, thus also suggesting a long-term relationship.

Acknowledgment

The paper was institutionally supported within the long-term research scheme of the Faculty of Informatics and Statistics, University of Economics, Prague.

References

Arlt,J., Arltová, M. (2009). *Ekonomické časové řady*. Praha: Professional Publishing. ISBN 978-80-247-1319-9.

Breusch, T., Goldfray, L.G. (1986). Data transformation Tests. *Economic Journal, 96*,

pp. 47-58.

Caner, M., Kilian, L. (2001). Size Distortions of Tests of the Null Hypothesis of Stationarity: Evidence and Implications for the PPP Debate. *Journal of International Money and Finance, 20,* pp. 639-657.

Dickey, D., Fuller, W. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association, 74*, pp. 427-431.

Dickey, D., Fuller, W. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, *Econometrica*, 49, 1057-1072.

Elliot, G., Rothenberg, J., Stock, J.H. (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica,* 64, 813-836

Engle, R.F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica 50,* 987-1007.

Granger, C. W. J., Newbold, P. (1974). Spurious Regression in Econometrics. *Journal of Econometrics*, 2(2), s. 111–120. ISSN 0304-4076

Hušek, R. (2007). *Ekonometrická analýza*. Praha: Oeconomica. ISBN 978-80-245-1300- 3.

Jarque, C., Bera, A. (1980). Efficient Tests for Normality, Heteroscedasticity, and Serial Independence of Regression Residuals, *Economics Letters 6,* pp.255-259.

Phillips, P.C.B. (1987). Time Series Regression with a Unit Root. *Econometrica 55*, pp. 277-301.

Contact

Dagmar Blatná
University of Economisc, Prague
W. Churchill sq. 4
130 67 Prague 3
Czech Republic
blatna@vse.cz