



Munich Personal RePEc Archive

A logit model for the estimation of the educational level influence on unemployment in Romania

Bogdan Oancea and Richard Pospisil and Raluca Dragoescu

University of Bucharest, Romania, Palacky University of Olomouc,
The Czech Republic, NSI Romania

7 June 2016

Online at <https://mpra.ub.uni-muenchen.de/81719/>

MPRA Paper No. 81719, posted 2 October 2017 21:05 UTC

A LOGIT MODEL FOR THE ESTIMATION OF THE EDUCATIONAL LEVEL INFLUENCE ON UNEMPLOYMENT IN ROMANIA)

BOGDAN OANCEA, RICHARD POSPÍŠIL, RALUCA MARIANA DRĂGOESCU

University of Bucharest - Faculty of Administration and Business - Department of Economic and Administrative sciences, Palacky University - Faculty of Arts - Department of Applied Economics, National Statistics Institute of Romania)

Abstract: Education is one of the main determinants of the unemployment level in all EU countries. In this paper we used a logit model to estimate the effect of the educational level on the unemployment in Romania using data recorded at the Population and Housing Census 2011. Besides the educational level we also used other socio-demographic variables recorded at the Census like gender, marital status, residential area. Data processing was achieved using R software system and since the data set used for model estimation was very large we used special techniques suited for big data processing. The results showed that the lowest odds ratio to be unemployed was recorded for population with tertiary education which is consistent with other studies at international level and with the official statistics data, but our study indicates that tertiary education has a greater impact on unemployment in Romania than in other EU countries.

Keywords: *educational level, unemployment, logit, higher education*

JEL classification: *I20, J24*

1 Introduction

The massification of the higher education sector all over the world has led to a substantial increase in the number of people with a very high level of competence trying to enter the labor market (OECD, 2015) which could make the task of finding an appropriate job to become difficult and eventually it may lead to an increasing unemployment rate. In this paper we investigated the influence of the educational level and especially of the higher education on the unemployment in Romania using the data from the 2011 Census.

There are many studies in the international literature that highlight a strong link between the educational level and unemployment rate such as Spence (1981), Mincer (1994) or Winkelmann (1996) to mention only a few of them. Among the first theoretical models that explain the low unemployment among higher education graduates we can mention those proposed by Nickell (1973) and Arrow (1973). They point out that higher education leads to the accumulation of human capital with high productivity and employers have an interest to maintain the productivity at a high level, so employees with higher education are less likely to be sent to unemployment when companies make staff cuts. Also, the signaling theory shows that higher education graduates have innate abilities more developed than the rest of the population and an academic title indicates these skills, employers being interested in having such people as employees.

We studied the influence of the educational level as well as other socio-demographic variables on the unemployment through a logit model. For higher education graduates we also investigated how the unemployment depends on the field of study.

2 Literature review

The relationship between the educational level and unemployment is an interesting research topic in labor economics. Altbeker and Storme (2013) carried out a study for South Africa and showed how different shocks or changes in economic conditions affect people differently in that they have less effect on the unemployment among people with higher education than among people without higher education. Using data for US Daly et al. (2007) also showed that the unemployment rate for higher education graduates is considerably lower than for people with only secondary education and the difference between the two categories was kept since 1978.

Nunez et al. (2010) examined the impact of higher education graduation and of the field of study on unemployment in Europe using data from the statistical survey "Labor Force Survey". Using an M-logit model, the authors analyzed the effect of the education on short term and long term unemployment using variables such as the marital status, gender, age, educational level and the field of study. Their results showed that a higher education diploma increases the odds of being employed on short-term and also lead to lower odds of being long-term unemployed. The authors also analyzed the changes in the probability of being unemployed by country and found that graduates of higher education in Belgium, Ireland and the UK have the lowest chances of being unemployed on the short term while graduates from Germany, Italy and Ireland have the lowest probability of being long-term unemployed. Considering the fields of study, the authors showed that most fields of study provide approximately the same probability of employment. Fields of study such as physics, chemistry, mathematics, statistics or informatics present the same probability of employment. The study fields that have the lowest unemployment probability are medicine, engineering and the sciences of education.

The effect of the public policies in the higher education area was investigated by Plumper (2007). He showed that public policies can be an effective tool available to the government in the fight against unemployment. Increasing the number of students lead to a decrease of the pressure on the labor market on short-term which has an effect on unemployment on short-term but the author warns about long-term effects of this type of policies: a decrease in the quality of education and an unsustainable growth in the number students. Garrouste (2010) used a binary response model and investigated the relationship between education and long term unemployment in EU countries showing that the probability of long-term unemployment decreases with the educational level.

In Romania, Danacica (2008) studied the influence of the gender, age and educational level on the probability of being hired or re-hired for 2002-2006 period, for a single county, showing that the educational level positively influences the likelihood that an unemployed person finds a job.

3 The unemployment and the educational level in Romania and EU

Figure 1 shows the evolution of the unemployment rate in Romania for three levels of education, during 2004-2013, for people aged over 15 years. It can easily be observed that every year the unemployment rate for higher education graduates has lower values compared with other educational

levels, which is consistent with international figures. The unemployment rate for higher education graduates registered a minimum value in 2008 reaching 2.5%.

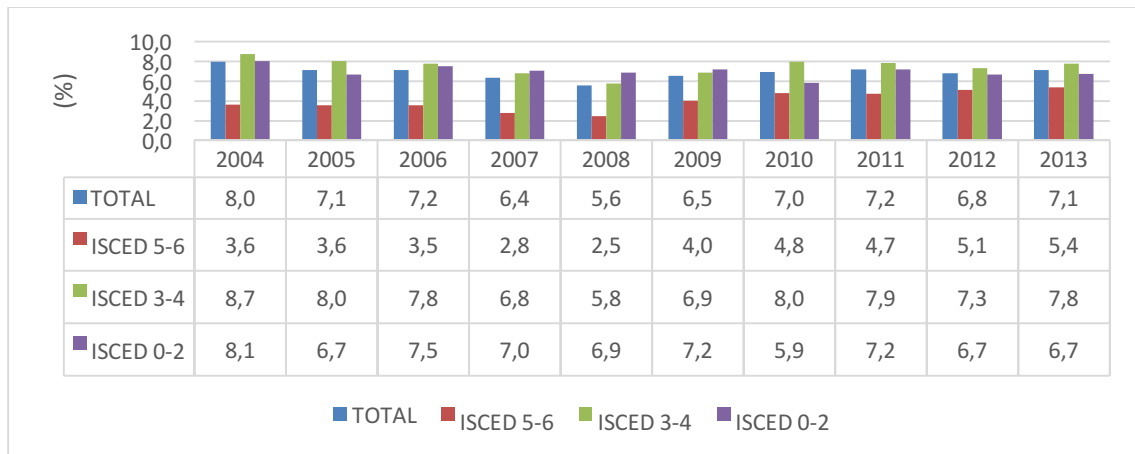
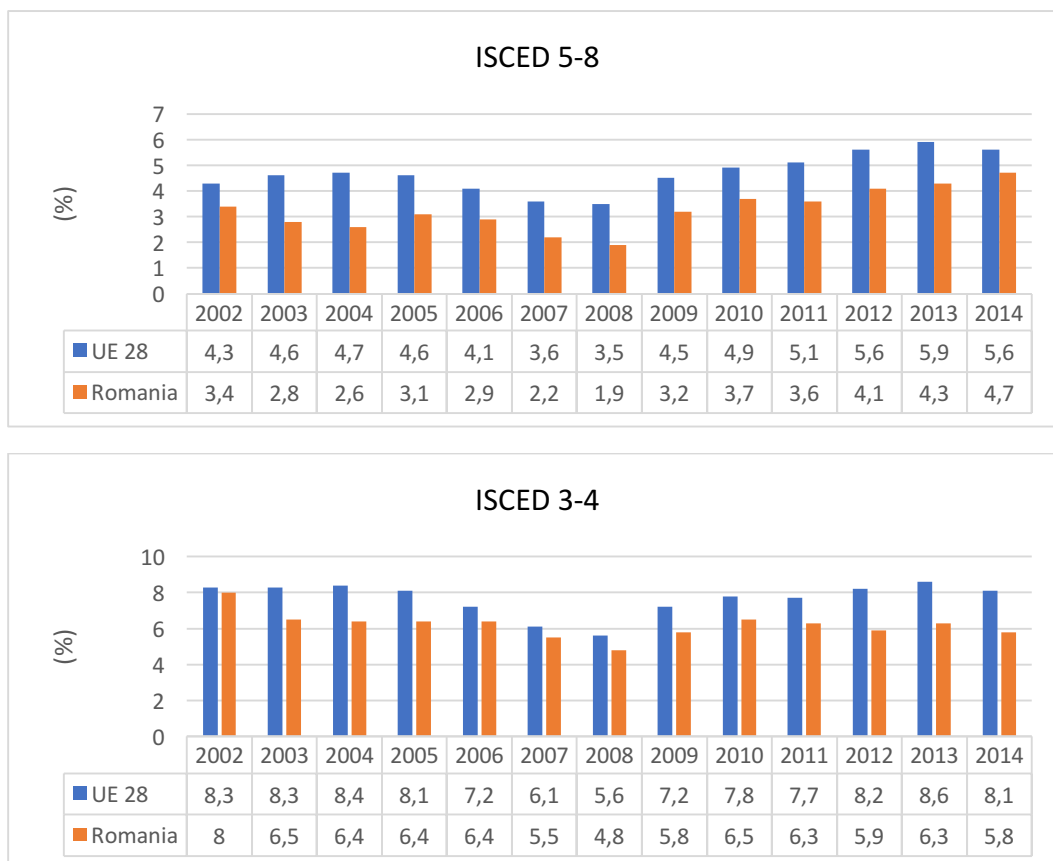


Figure 1. Unemployment rate on educational levels during 2004-2013

Figure 2 presents the evolution of the unemployment by the educational level for Romania and the average value for EU28 countries, considering the population aged between 25 and 64 years.



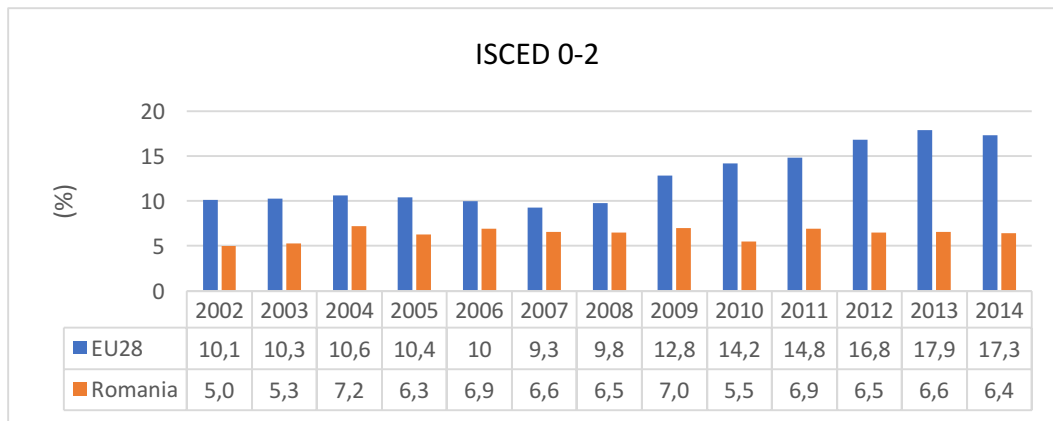


Figure 2. The evolution of unemployment by the levels of education for Romania and EU28 countries

Romania recorded unemployment rates lower than the average of the EU28 countries for all educational levels. The value of the unemployment rate for people with higher education reached a minimal value in 2008 for Romania (1.9%) and EU28 (3.5%), after this year having an upward trend.

Figure 3 shows the unemployment rate for people with higher education in 2005, 2010 and 2014 for all EU countries.

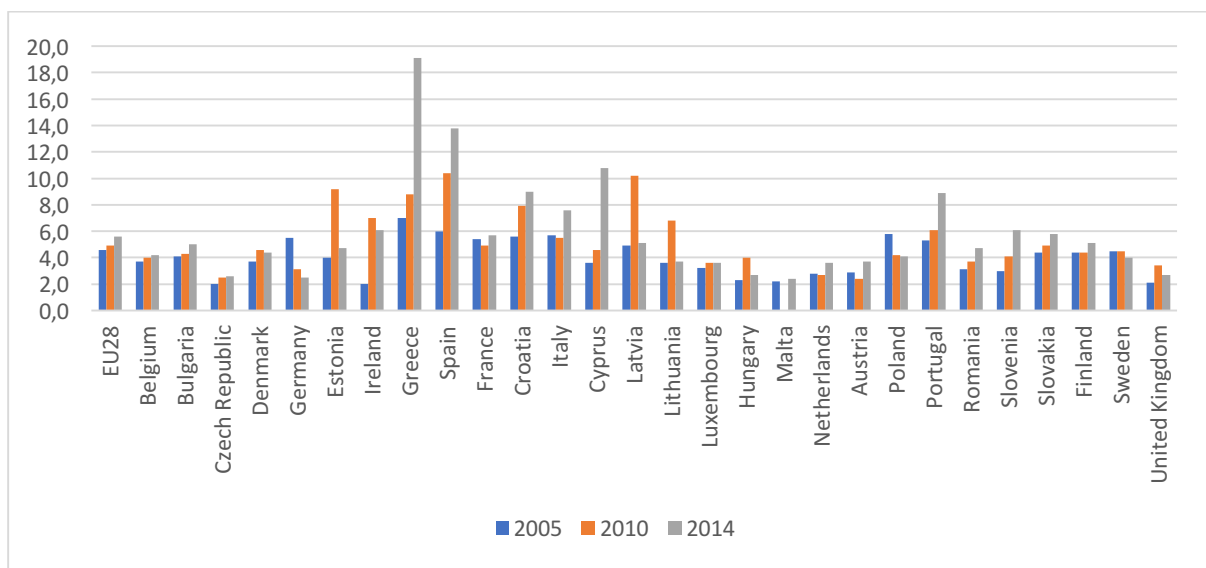


Figure 3. The unemployment rate for higher education graduates aged between 25 and 64 years in EU countries

Romania is among the countries with the lowest unemployment rate for people with higher education compared with other EU member countries. Greece, Spain, Cyprus, Latvia and Portugal are at the opposite side, recording high unemployment rates among people with tertiary education. In 2014, the unemployment rate for higher education graduates in Greece was 19.1% while in Spain it was 13.8%. In most of the countries, the unemployment rate for people with higher education graduates increased from 2005 to 2014.

4 The econometric model

We studied the influence of the educational level as well as other socio-demographic variables on unemployment through a logit model (Wooldridge, 2008) with the dependent variable being the occupational status registered at the 2011 Census. The logistic regression allows one to predict the values of a dichotomous variable Y which takes only two values, 0 or 1, depending on a set of explanatory variables, which can be quantitative or categorical variables. The logistic regression does not work with probabilities, but with odds. For people with higher education we also investigated how the unemployment depends on the field of study.

We used the variables presented in table 1 (Drăgoescu, 2015), derived either directly from data recorded at the 2011 Census or by processing these data. Besides the variables directly related to the educational level, we used a number of socio-demographic variables: gender, marital status, residence (urban/rural), and a binary variable that indicate whether the person lives in a big city (a city with a total population greater than 150,000). The study included population aged between 25 and 64 years and the data set comprised 7,652,044 records.

Table 1. Variables used in studying the influence of educational level on unemployment

Variable	Avg.	Std. Dev.	Description
UNEMPL	0.06	0.23	1-unemployed, 0 otherwise
GENDER	0.55	0.49	Gender: 1-M, 0-F
MARRIED	0.76	0.45	Marital status 1-married, 0-widow, single or divorced
BIG_TOWN	0.27	0.44	1 for persons living in cities with more than 150,000 inhabitants, 0 otherwise
URBAN	0.59	0.49	1 for persons living in urban areas, 0 otherwise
EDU_CAT	3.86	1.48	Categorical variable indicating a person with: 1-primary education 2-gymnasium 3-vocational school 4-high school 5-post-secondary school 6-higher education

HAS_GYM	0.19		0.39		1-gymnasium, 0 otherwise
HAS_PROF	0.18		0.38		1-vocational studies, 0 otherwise
HAS_HS	0.31		0.46		1-high school, 0 otherwise
HAS_POST	0.04		0.19		1-post-secondary studies, 0 otherwise
HAS_HE	0.24		0.43		1-tertiary education, 0 otherwise
	Total population	Higher education graduates	Total population	Higher education graduates	
HE_TECH	0.07	0.30	0.26	0.46	1-tertiary education in the technical field, 0 otherwise
HE_UNIV	0.06	0.26	0.24	0.43	1-tertiary education in the field of education and natural sciences, 0 otherwise
HE_EC	0.07	0.28	0.25	0.45	1-tertiary education in the field of social sciences and economics, 0 otherwise
HE_LAW	0.02	0.09	0.15	0.29	1-tertiary education in the field of law, 0 otherwise
HE_MED	0.014	0.06	0.12	0.24	1-tertiary education in the field of medicine, 0 otherwise

All computations were performed using the R software system. The probability of being unemployed by the educational level was studied using several logit models. The first model considered the gender, marital status, residence and the categorical variable EDU_CAT, whose benchmark was set to EDU_CAT=1, as explanatory variables. The results of the model estimations indicate the odds ratio of being unemployed for a person with gymnasium, vocational school, high school or university degree compared with the reference category. The dependent variable was UNEMPL. The equation of the model is given by (1) and the results of the estimation are shown in table 2.

$$UNEMPL_i = b_0 + b_1 GENDER_i + b_2 MARRIED_i + b_3 URBAN_i + b_4 BIG_TOWN_i + b_5 EDU_CAT_i + \varepsilon_i \quad (1)$$

Table 2. The estimations of model (1)

	Estimate	Std. err.	p-value
INTERCEPT	-2.67	0.0087	<2e-16

GENDER	0.41	0.0033	<2e-16
MARRIED	-0.63	0.0032	<2e-16
URBAN	0.67	0.0039	<2e-16
BIG_TOWN	-0.12	0.0039	<2e-16
EDU_CAT2	-0.13	0.0089	<2e-16
EDU_CAT3	-0.14	0.0089	<2e-16
EDU_CAT4	-0.28	0.0088	<2e-16
EDU_CAT5	-0.73	0.0125	<2e-16
EDU_CAT6	-0.77	0.0093	<2e-16

All coefficients of the equation have significant values at 1% significance level. These coefficients indicate the change of the log odds of being unemployed at a unit increase of the predictor variable.

The coefficient of the GENDER variable shows that the odds ratio of being unemployed for men comparing to women is 1.5, the odds of unemployment being 50% higher for men than for women. This is in agreement with official statistics data for Romania but contradicts other studies carried out at European level showing that women have a higher probability of unemployment (Nunez, 2010; Garrouste et al., 2010).

The marital status influences the unemployment, married people having fewer chances to be unemployed compared to the unmarried people. In this case the odds ratio has a value of 0.53, the odds of unemployment being 47% lower for married persons compared to the rest of people.

Residence in urban areas increases the unemployment, the odds ratio of being unemployed for those living in urban areas compared to those living in rural areas is 1.9, meaning that the odds of unemployment is about 90% higher in urban areas than in rural areas which is confirmed by official statistics of Romania. The explanation comes from the fact that in rural areas most people perform subsistence farming activities, thus they are not considered unemployed. However, people living in cities with population greater than 150,000 inhabitants are less likely to be unemployed. The odds ratio of being unemployed for the population in large cities compared with the rest of the population is 0.81.

The effect of the educational level on unemployment is studied through EDU_CAT categorical variable. The reference value is EDU_CAT=1. The coefficients of the EDU_CAT2 ... EDU_CAT6 give the change in the log odds of being unemployed for people with different educational levels compared to the reference level. For example, the change of the log odds of being unemployed for people with secondary education compared to the reference category is -0.28, i.e. the odds ratio of being unemployed for people with secondary education compared to those with only primary school is 0.75. The educational level influences the odds of being unemployed, people with tertiary education having the lowest odds of unemployment compared with other categories of persons. The coefficients

obtained for people with tertiary education in Romania's case are higher than the coefficients obtained in other similar studies in Europe (Nunez, 2010).

We tested the significance of the overall effect of the educational level using the WALD test on the significance of the EDU_CAT variable. Chi-squared statistics for five degrees of freedom is 21640.1 and p-value<0.001, indicating that the educational level effect is statistically significant.

The goodness of fit of our model was tested using the significance test described in (Hosmer and Lemeshow, 2000). This test checks whether the model that uses the predictor variables is significantly better than a model with only a constant, called the null model. The statistics of this test is distributed according to a Chi-squared law with a number of degrees of freedom equals to the difference between the degrees of freedom of the model with explanatory variables and those of the null model. The statistics of the test is 105,456.5 for nine degrees of freedom with p<0.001, which indicates that our model is statistically significant.

Next, we built a model considering the educational level through HAS_PRI HAS_GIM, HAS_PROF, HAS_HS, HAS_POST, HAS_HE variables. The model is given by:

$$UNEMPL_i = b_0 + b_1GENDER_i + b_2MARRIED_i + b_3URBAN_i + b_4BIG_TOWN_i + b_5HAS_PRI_i + b_6HAS_GIM_i + b_7HAS_PROF_i + b_8HAS_HS_i + b_9HAS_POST_i + b_{10}HAS_HE_i + \varepsilon_i \quad (2)$$

Table 3. The estimations of model (2)

	Estimate	Std. err.	p-value
INTERCEPT	-1.98	0.0141	<2.22e-16
GENDER	0.41	0.0033	<2.22e-16
MARRIED	-0.63	0.0032	<2.22e-16
URBAN	0.68	0.0039	<2.22e-16
BIG_TOWN	-0.13	0.0038	<2.22e-16
HAS_PRI	-0.70	0.0161	<2.22e-16
HAS_GYM	-0.83	0.0143	<2.22e-16
HAS_PROF	-0.84	0.0143	<2.22e-16
HAS_HS	-0.98	0.0142	<2.22e-16
HAS_POST	-1.44	0.0169	<2.22e-16
HAS_HE	-1.46	0.0145	<2.22e-16

We tested the significance of the model and obtained a statistics equals to 110,748.65 for 10 degrees of freedom, p<0.0001, i.e. our model is statistically significant.

GENDER, MARRIED, URBAN and BIG_TOWN have coefficients approximately equal with those of equation (1), their interpretation being identical. The coefficients of the dummy variables show that the log odds of being unemployed decreases with an increased educational level, values ranging from -0.63 for HAS_PRI to -1.46 for HAS_HE variable. Table 4 presents the exponentials of the coefficients of the dummy variables corresponding to the educational level, these values representing the odds ratio of being unemployed for the people with a certain level of education than those who do not have that level. These values show that the odds of being unemployed decrease as educational levels increase, people with higher education having the lowest unemployment odds. The odds ratio of being unemployed for higher education graduates compared to the rest of people is 0.23 meaning that the odds of being unemployed for a higher education graduate are 77% lower.

We applied a WALD procedure to test whether there is a statistically significant difference between the coefficient of the HAS_HE variable and HAS_POST variable coefficient. We obtained a statistics value equals to 6.3 for one degree of freedom, and p=0.01, i.e. there is a significant difference between higher education and the educational level immediately below in terms of influence on the unemployment.

Table 4. The exponential of the coefficients of model (2)

Variable	e^{b_i}
HAS_PRI	0.49
HAS_GIM	0.44
HAS_PROF	0.43
HAS_HS	0.37
HAS_POST	0.24
HAS_HE	0.23

The last model that we estimated considered only persons with higher education and we assessed the field of study influence on unemployment. The model is:

$$UNEMPL_i = b_0 + b_1HE_TECH_i + b_2HE_UNIV_i + b_3HE_EC_i + b_4HE_LAW_i + b_5HE_MED_i + \varepsilon_i \tag{3}$$

Table 5. The estimations of model (3)

	Estimate	Std. err.	p-value
INTERCEPT	-2.63	0.0260	<2.22e-16
HE_TECH	-0.46	0.0268	<2.22e-16
HE_UNIV	-0.40	0.0269	<2.22e-16

HE_EC	-0.39	0.0268	<2.22e-16
HE_LAW	-0.51	0.0287	<2.22e-16
HE_MED	-0.90	0.0317	<2.22e-16

Table 6. The exponential of the coefficients of model (3)

Variable	e^{b_i}
HE_TECH	0.63
HE_UNIV	0.67
HE_EC	0.68
HE_LAW	0.60
HE_MED	0.41

The values from table 6 show that the field of study presenting the lowest odds of unemployment is medicine, the odds ratio of unemployment for a person with higher education in medicine compared to the rest of the graduates being 0.41. This result is consistent with other studies: Nunez (2010) shows that for 15 EU countries the odds ratio of being unemployed on short term as well as on long term has the lowest value for graduates from medicine schools. It follows in descending order the law, technical and education and natural sciences fields of education. The biggest odds of being unemployed is recorded for the graduates of the social and economics field of study. The results can be explained by the fact that social and economic sciences has the largest share in the total number of students and the labor market can not absorb all graduates. On the other hand, the medicine requires large personal investment which makes the number of people undergoing such studies to be low and the probability of employment at the end of studies to be very high. The significance of the model (3) was tested with a similar procedure and the statistics of the test is 1145.7 for 5 degrees of freedom with $p < 0.001$, which means that the model is significant.

5 Conclusions

In this paper we studied the relationship between unemployment and the educational level using a series of logit models. The models that we estimated show that education influences the odds of being unemployed, increasing levels of education being correlated with decreasing odds of unemployment. Higher education graduates have the lowest unemployment odds compared with other educational levels. Regarding the field of study, the smallest odds of unemployment are recorded for medicine while social and economic sciences present the biggest odds of unemployment.

Literature:

Altbeker, A., & Storme, E. (2013). Graduate unemployment in South Africa: A much exaggerated problem, The Centre for Development and Enterprise, Johannesburg.

Arrow, K. (1973). Higher education as a filter. *Journal of Public Economics*, 2, 193-216.

Daly, M. C., Osborne, J., & Valletta, R. G. (2007). Educational Attainment, Unemployment, and Wage Inflation, *Economic Review*, Federal Reserve Bank of San Francisco, 49-61.

Dănănică, D. E., & Babucea, A. G. (2008). Risk of employment – a logistic regression approach, *Analele Universității "Constantin Brâncuși" from Târgu Jiu, Seria Economie*, 1, 123-136.

Drăgoescu, R.M., (2015). Analiza prin metode cantitative a evoluției și tendințelor în învățământul superior din România, Ph.D. Thesis, The Bucharest University of Economic Studies.

Garrouste, C., Kozovska, K., & Perez, E.A. (2010). Education and long term unemployment. JRC Scientific and Technical Reports, European Commission.

Hosmer, D. & Lemeshow, S. (2000). *Applied Logistic Regression*, Second Edition. New York: John Wiley & Sons, Inc.

Mincer, J. (1994). Human Capital: A Review. In Clark Kerr et al. (eds) *Labour economic and Industrial Relations: Markets and Institutions*. Cambridge: Harvard University Press.

Nickell, S. (1973). Education and lifetime patterns of unemployment. *Journal of Political Economy*, 87(5), 117-131.

Nunez, I., & Livanos, I. (2010). Higher education and unemployment in Europe: an analysis of the academic subject and national effects. *Higher Education*, 59(4), 475-487.

Plümper, T., & Schneider, C. J. (2007). Too much to die, too little to live: unemployment, higher education policies and university budgets in Germany. *Journal of European Public Policy*, 14(4), 631-653.

OECD. (2015). *Education at a Glance 2015: OECD Indicators*. OECD Publishing, Paris.

Spence, M. (1981). Signaling, screening, and information. In S. Rose (Ed.) *Studies in Labour Markets*. Chicago: University of Chicago Press.

Winkelmann, R. (1996). Employment prospects and skills acquisition of apprenticeship trained workers in Germany. *Industrial and Labour Relations Review*, 49, 658-672.

Wooldridge, J. M. (2008). *Introductory Econometrics*. South Western College.