

## **Determinants of Earnings Inequality - Case Study for Romania**

**Larisa Stanila, PhD**

Researcher

National Scientific Research Institute for Labour and Social Protection

6-8 Povernei Str., District 1

Bucharest, 010643

Romania

**Maria Denisa Vasilescu**

**Amalia Cristescu, PhD**

**Madalina Ecaterina Popescu, PhD**

Researcher

National Scientific Research Institute for Labour and Social Protection

6-8 Povernei Str., District 1

Bucharest, 010643

PhD student

Bucharest University of Economic Studies

6 Romana Square, 010374

Romania

### **Abstract**

*In this paper we study the main determinants of wage inequality in 2012 in Romania through a Mincer wage equation. An ordered logit model was therefore estimated, based on the following explanatory variables of wage inequalities: gender, age, residence area, education and major sectors of activity. The main results are consistent with the human capital theory and with the findings of Mincer's followers. As expected, people with more education receive higher earnings, confirming that education pays off. We also found a positive correlation between age and earnings, meaning that with age, it increases the likelihood of a person to earn more. Regarding the gender inequality, our findings indicate that women in Romania are less likely to earn more than men and when it comes to residence area, the individuals in urban areas have higher chances in earning more, as compared to those employed in the rural area.*

**Keywords:** net earnings, Mincer equation, earnings inequality, ordered logit model, Romania

### **1. Introduction**

In this paper we study the main determinants of earnings inequality in Romania by elaborating a Mincer wage equation using ordered logit model, based on the following explanatory variables of earnings inequalities: gender, age, residence area, education and major sectors of activity. For that, we used micro-data for the year 2012 from the Romanian Household Labour Force Survey.

The choice for the selected explanatory variables was based on data availability grounds and was also inspired by the list of Mincer's variables from the wage inequality model (1958, 1974), that was later extended with other supplementary variables, including gender, ethnicity and union membership, in more recent studies (Polachek, 2007).

The human capital theory states that increased skills and knowledge should lead to a growth in labour productivity. This means that education and training are the key factors of an individual economic performance that should normally lead to higher earnings.

The framework of wage inequality is mostly based on the Mincer's model of earnings (1974), which regresses the log earnings on linear term in years of schooling and linear and quadratic terms in years of labour market experience. According to Mincer, people with more education receive higher earnings and the ratio of earnings for persons with education levels differing by a fixed number of years is roughly constant across schooling levels.

Mincer's model captures not only how the labour market rewards depends on attributes as schooling or work experience, but also the rate of return to schooling, which can be interpreted as an interest rate to human capital investments. Although there were studies focusing on estimating the rate of return to schooling in the wage equation (Heckman, Lochner and Todd, 2003), recent studies of the Mincer model no longer provide valid estimates for returns to schooling.

The Mincerian wage model has however been used in many studies in order to see the differences between earnings of different racial groups, differentials due to working in public or private sector, wage gaps between males and females. For example, Fournier and Koske (2012) analysed the determinants of earnings inequality for a cross-section of 32 countries using household survey data, by considering the following variables: working hours, gender, age and age squared and the highest education level attained. They found that women are less likely to be employed than men and those who work typically earn less than their male counterparts. Moreover, work experience plays a larger role in lower-paid jobs when considering that the age terms capture the individual's work experience and working hours lead to higher earnings inequality among the working population. The empirical evidence also indicated that policies to increase upper-secondary graduation rates should reduce earnings inequality.

Mincer's human capital wage model has also been used in some studies to investigate the impact of union status of workers on their wages (Kahn, 1998).

In terms of openness, Wood (1995) concluded that trade openness, FDI and outsourcing play significant roles in explaining the dispersion of wage structure. In contrast, Chen and Kuan (2006) used micro data to examine the impacts of globalization on the wage inequality in Taiwan's manufacturing industry and agreed that trade has limited ability to explain the variation of wage differentials.

Lemieux (2006) studied the growth in wage inequality in the United States and other advanced industrialized countries and presented evidence from quantile regressions for a human capital model with heterogeneous returns to show that most of the increase in wage inequality between 1973 and 2005 is due to a dramatic increase in the return to post-secondary education.

In another study, Buchinsky (2001) focused on the changes in the returns to education for the US female workers over the period 1967 - 1990, by applying the Mincerian model with quantile regression. The results indicated that returns to education differ significantly over time and across age groups and these changes are not uniform over the quantiles of the conditional wage distribution.

Pereira and Martins (2000) address the impact of education upon wage inequality by drawing on evidence from fifteen European countries, during the period 1980-1995. They focused on within-educational-levels wage inequality and found that, in most countries, dispersion in earnings increases with educational levels and that education is a risky investment.

Tansel and Bircan (2011) investigated the male wage inequality over the period 1994-2002 in Turkey by estimating Mincerian wage equations using OLS and quantile regression techniques. Education contributed to higher wage inequality through both within and between dimensions, public employment and living in urban areas all contributed to lower male wage inequality through the within dimension, while working for a large firm generated higher male wage inequality in 1994 and lower male wage inequality in 2002.

Although the wage inequality was studied less often in the developing countries, there are several studies in the field, such as Monsueto, Machado and Golgher (2006) who studied the wage inequality in Brazil and Tello, Ramos and Artís (2012) for the case of Mexico.

It is known that income inequality in Brazil is generally considered one of the worst in the world. Having this into consideration, Monsueto et al. (2006) analysed a possible improvement in the income distribution in favour of low-income workers, focusing on the reduction in income gaps between racial groups within the poorest 25%, based on data from the National Household Survey for 1987, 1995 and 2001.

On the other hand, Tello et al.(2012) conducted an empirical analysis of the relationship between wage inequality, employment structure and returns to education in urban areas of Mexico during the past two decades (1987-2008). They found that changes in wage inequality have been driven mainly by variations in educational wage premia and that changes in employment structure, including occupation and firm size, have played a vital role. Although it is claimed that raising education serves to reduce wage inequality, their analysis suggested that in Mexico increases in educational levels do not necessarily result in a more equitable wage distribution.

As highlighted from the vast literature and from the numerous empirical studies in the field, one can without doubt conclude that finding the main determinants of wage inequality in order to be able to elaborate proper policies to reduce wage inequality is a permanent concern of research, with high importance. Our paper is therefore focusing on these issues and is structured as following: section 2 presents a statistical analysis of the main wage inequalities in Romania, section 3 describes the econometric analysis, while section 4 concludes.

## 2. Statistical Analysis

The data used in this paper come from the Romanian Household Labour Force Survey (RLFS). The survey covers the whole country, all the surveyed households are private and the participation is voluntary. Data are collected with identical questionnaires through the whole year, which are divided into two parts: a dwelling and an individual questionnaire. The individual questionnaire is addressed only to the household members aged 15 and over and contains, among others, questions related to: professional status, place of work, occupation, main and secondary activity, hours worked, willingness of another job, training.

In our study, we selected only the employed individuals. The main variable of interest is the monthly (take home) pay from main job, the individuals being assigned into 10 groups (corresponding to the earnings deciles).

As personal characteristics of the individuals we used the gender, the age and the residence area (urban/rural). Even if the survey collects the exact age of the individuals, we decided to use intervals: 15-24 years, 25-34 years, 35-54 years, 55-64 years and over 65 years.

The human capital of the individual was evaluated using its education. Four possibilities were taken into consideration: no education (ISCED 0); primary and lower secondary education (ISCED 1-2); upper secondary and post-secondary non-tertiary education (ISCED3-4); first and second stage of tertiary education (ISCED 5-6).

Regarding the economic characteristics of the individuals we included in our study the sector of activity (NACE Rev 2), grouped into 5 categories:

- Agriculture (A. Agriculture, forestry and fishing)
- Industry (B. Mining and quarrying; C. Manufacturing; D. Electricity, gas, steam and air conditioning supply; E. Water supply, sewerage, waste management and remediation activities)
- Construction (F. Construction)
- Services of the business economy (G. Wholesale and retail trade, repair of motor vehicles and motorcycles; H. Transportation and storage; I. Accommodation and food service activities; J. Information and communication; K. Financial and insurance activities; L. Real estate activities; M. Professional, scientific and technical activities; N. Administrative and support service activities)
- Public services (O. Public administration and defence, compulsory social security; P. Education; Q. Human health and social work activities; R. Arts, entertainment and recreation; S. Other service activities).

The analysis was conducted for the year 2012. After applying our selection on the RLFS, the database used in the analysis contains entries for 62146 individuals.

Figure 1 highlights the distribution of net earnings by level of education and age groups. It can be seen that the lowest level of earnings is predominant in the group of workers with no education (ISCED 0), aged between 55-64 years. For the category of basic education (ISCED 1-2) the percentage of those with low earnings is situated at an almost constant level (about 45%) for all five categories of age. Interestingly, low wages are also present in the categories of workers with secondary education (ISCED 3-4) and higher education (ISCED 5-6), even though in a lower percentage and decreasing with age. For the employees with no education, the hope for a median or superior level of salary is very poor. It can be seen that only the experience rendered by age and maybe some favourable circumstances are likely to offer a chance to earn a higher level of salary (7.69% of those with no education aged 35-54 have earnings situated in the top two deciles).

This segment of workers with no education may be formed of the young generation of the '90s which benefited from favourable professional paths. It might also be formed of freelancers or even small business owners who also have an employee status.

The ISCED 1-2 education group is dominated by low levels of earnings, especially in the case of workers aged 65 and over (52.17%). Average wages have a low percentage for all levels of education (between 4.35% and 10.63%) and high wages are not present in the case of elder employees, reaching a maximum of 7.49% for young people. The higher percentage of young people with average and high levels of earnings shows that they tend to choose better-paid fields of activity (they have a greater capacity of professional orientation and are better adapted to the market economy - educated after 1990).

For those with an ISCED 3-4 education level, high earnings are predominant at 55-64 years (15.07%) and especially at the age of 65 and over (22.73%); this category is formed of employees who entered the labour market during the period in which the percentage of employees with higher education was low (therefore one could get a well-paid job even with an upper secondary education level).

As expected, in the group of employees with ISCED 5-6 education there is a high level of earnings for all the categories of age, with an increasing trend for the greater ages due to work experience and professional promotions within the companies (better paid jobs, management positions etc.). Thus, the highest percentages of employees with high levels of earnings are found among those with ISCED 5-6 education and over 35 years old (over 45%).

The analysis based on economic activities (figure 2) indicates a polarization of net earnings since there are sectors of activity in which higher levels of salary are predominant (Public services - 22.61%; Construction - 19.40%) and sectors of activity in which lower levels of salary are predominant (Agriculture - 23.87%; Services of the business economy - 22.54%; Industry - 21.33%). The distribution by economic activities reveals more significant levels of median earnings in Construction (11.48%), Services of the business economy (10.80%) and Agriculture (10.59%). The most balanced distribution of the earnings profiles seems to be in Construction and more or less in Services of the business economy. But while in Construction high salaries dominate, in the Services of the business economy field low salaries are the dominant ones. In the Construction sector, the higher level of earnings can be explained by pay rises which are given in case of harder working conditions and longer working hours.

### 3. Econometric Analysis

#### 3.1. Methodology

According to Agresti (2010), the logit model is a single-period classification model for which estimates are based on a maximum likelihood function in order to determine the conditional probability of an individual belonging to a category based on the independent variables. The logit model describes the relationship between the binary variable  $Y$ , which takes the values 1 or 0, and  $k$  explanatory variables  $x_1, x_2, \dots, x_k$ . Starting from the linear probability model  $\Pr(y = 1|x) = x\beta + \varepsilon$ , if the probabilities are restricted to be in the  $[0,1]$  range, the following model is obtained:  $\Omega(x) = \frac{\Pr(y=1|x)}{\Pr(y=0|x)} = \frac{\Pr(y=1|x)}{1-\Pr(y=1|x)}$ .

The logit model is derived by applying logarithms to these probabilities and thus obtaining  $\ln\Omega(x) = \beta_0 + \sum_{j=1}^k \beta_j x_{i,j}$ . This equation is known under the name of logit form of the model, where  $\ln\Omega(x)$  is the logarithm of the probability for an individual to be part of one category or another in relation to the explanatory variables  $x_{i,1}, x_{i,2}, \dots, x_{i,k}$ .

In this paper, the dependent variable is the decile of monthly take home pay, therefore a categorical variable taking values from 1 to 10. Although having this codification can become tempting to apply a linear regression, one should keep in mind that if a dependent variable is ordinal the classical linear regression model assumptions are violated, leading to incorrect results (Long and Freese, 2006). Thus, one of the models that can be applied in such cases is the ordered logit non-linear model, which can be viewed as an extension of the logit binary model. The ordered logit model estimates the probability of an individual belonging to a category versus a smaller or a higher category (Hosmer et al., 2003), as the values of the dependent variable are ordered:

$$\Omega_{\leq m | > m}(x) = \frac{\Pr(y \leq m | x)}{\Pr(y > m | x)}$$

One of the characteristics of the ordered logistic regression is that the relationship between each pair of outcome categories is the same, that is the ordered logistic regression implies that the coefficients describing the relationship between the lowest versus all higher categories of the outcome remain unchanged as those that describe the relationship between the next lowest category and all higher categories, etc. This hypothesis is called the proportional odds assumption or the parallel regression assumption (Long and Freese, 2006). Because of this assumption, there is only one set of coefficients (i.e. only one model estimated).

### **3.2. Estimation Results**

In order to evaluate the earnings inequality we applied an ordered logit model. After estimating such a model, one can evaluate if the coefficients are statistically significant, namely if the associated variable significantly affects the dependent variable. This assessment can be made based on the associated probabilities reported in the initial estimation, or based on a Wald test. In this particular case, both the p-values and the Wald test indicate that each predictor significantly affects the probability of an individual to be found in the higher earnings decile. Moreover, the Wald test for multiple coefficients is also indicating that the hypothesis that all the variables are simultaneously equal to zero can be rejected at the 0.01 level.

We decided not to report the ordered log-odds (logit) regression coefficients, but the odds ratios, as they clearly indicate the probability for higher earnings compared with lower earnings, for each considered variable (figure 3). The first obvious thing is that positive coefficients increase the chances that the subject will be observed in a higher category and negative coefficients increase the chances the subject will be observed in a lower category. Regarding the gender inequality, we came at the conclusion that the odds of having higher earnings (being in a higher decile) are 134.8% bigger for men compared with women.

When it comes to residency, the persons living in urban areas have a 44.8% probability of having higher take home pay, compared to people from the rural area.

Another important conclusion is the confirmation of a positive correlation between age and earnings: with age, the likelihood that a person earns more increases.

For the young employees, 15 to 24 years, the odds of being in a higher earnings decile are 26.3% smaller than the employees aged 25 to 34 years. The persons in the age group 35-54 years have a 29.6% higher probability to earn more than the young adults (25-34 years). And this probability increases with age. The employees aged between 55 - 64 years have the odds of being in a higher decile with 43.1% bigger than the group 25-34 years.

Finally, the employees aged 65 or over have 57.6% higher chances to earn more compared to the young adults (25-34 years). When analysing the individuals in the 65+ age group we see that 44% of them are highly educated (ISCED 5-6) and 37% have secondary education (ISCED 3-4). Moreover, 41% of elderly workers are Professionals and the majority (63%) are concentrated into three categories of occupations: Professionals, Managers and Technicians and associate professionals. Therefore, highly educated people working in prestigious occupations continue to be active on the labour market even after 65 years, the legal age for pension, and to be considerably remunerated.

Education is another important factor influencing the distribution of earnings. We see that the individuals with no education have 83.3% smaller chances to be in higher earnings deciles, compared to the individuals with ISCED 3-4 level of education. Also, the odds of having higher take home pay are 63% smaller for the employees with only ISCED 1-2 level of education, than for the employees with secondary education. When considering the employees with tertiary education (ISCED 5-6), the results showed a very high probability for these individuals to be in the top deciles of earnings. The chances for them to have higher take-home pay are 356.8% bigger compared to the employees with secondary education. Therefore, it seems that the education really pays off.

When considering the sector of activity, we can state that the employees working in Agriculture have 9% more chances of being in a higher earnings decile than the individuals working in Public services. Also, the probability to have higher take home pay is 8.9% higher for the individuals working in Industry, compared with those from Public services. In Construction are the highest chances to be in the top deciles of earnings, as the probability for an employee working in this sector to earn more than one from Public services is 24.8% higher. The employees from Services of the business economy have the odds of being in a higher decile with 8.7% smaller than the employees from Public services.

#### 4. Conclusions

Our analysis focused on the study of the main determinants of earnings inequality in 2012 in Romania and consisted in elaborating a Mincer wage equation using ordered logit model. The following explanatory variables of earnings inequalities were taken into consideration: gender, age, residence area, education and major sectors of activity.

The main findings that have emerged from our empirical study for the case of Romania are consistent with the human capital theory and with the findings of Mincer's followers, in all aspects concerning the impact of gender, age, residence area, education upon wage inequality.

As expected, people with more education receive higher earnings. More precisely, individuals with no education have smaller chances to be in higher earnings deciles, compared to the individuals with ISCED 3-4 level of education. Also, the odds of having higher take home pay are smaller for the employees with only ISCED 1-2 level of education, than for the employees with secondary education. When considering the employees with tertiary education, the results showed a very high probability for these individuals to be in the top deciles of earnings, confirming that education pays off.

Our results also confirm the positive correlation between age and earnings, meaning that with age, it increases the likelihood of a person to earn more.

Regarding the gender inequality, our findings indicate that women are less likely to earn more than their male counterparts. Moreover, when it comes to residence area, the individuals living in urban areas turned out to have higher probability of earning more, as compared to those employed in the rural area.

When considering the sector of activity, we can state that in Construction are the highest chances to be in the top deciles of earnings, according to the probability for an employee working in this sector to earn more than one from Public services, while Services of the business economy is the only sector in which an employee has smaller chances to earn more.

The paper hopes to make a significant contribution by using micro-data in order to identify the determinants of earnings inequality for the case of Romania, which might give an insight for the policy makers in order to reduce earnings inequality.

#### References

- Agresti, A. (2010). *Analysis of Ordinal Categorical Data*. (2nd ed.), New York: Wiley Inc.
- Buchinsky, M. (2001). Quantile regression with sample selection: Estimating women's return to education in the U.S. *Empirical Economics*, 26 (1), 87-113
- Chen, C.L., & Kuan, C.M. (2006). The Impacts of Globalization on Wage Inequality in Taiwan Manufacturing: Evidence from Quantile Regression Analyses
- Fournier J.M., & Koske, I. (2012). The Determinants of Earnings Inequality – Evidence From Quantile Regressions. *OECD Journal: Economic Studies*, 7-36
- Heckman, J.J., Lochner, L.J., & Todd, P.E. (2003). Fifty Years of Mincer Earnings Regressions. NBER Working Papers, 9732.
- Hosmer, D.W., Lemeshow, S., & Sturdivant, R.X. (2013). *Applied Logistic Regression*. (3rd ed). New York: Wiley Inc., (chapter 8).
- Kahn, L.M. (1998). Collective Bargaining and the Interindustry Wage Structure: International Evidence. *Economica*, New Series, 65 (260), 507-534.
- Lemieux, T. (2006). Post-Secondary Education and Increasing Wage Inequality, NBER Working Paper No. 12077
- Long, J.S., & Freese, J. (2006). *Regression models for categorical dependent variables using Stata*, (2nd ed.), Stata Press.
- Mincer, J. (1958). Investment in Human Capital and the Personal Income Distribution. *Journal of Political Economy*, Vol. 66, 4, 281-302.
- Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: National Bureau of Economic Research.
- Monsueto, S.E., Machado, A.F., & Golgher, A.B. (2006). Earning inequalities in Brazil: quantile regressions and the decomposition approach, *CEPAL REVIEW* 9 0

Pereira, P.T., & Martins, P.S. (2000). Does Education Reduce Wage Inequality? Quantile Regressions Evidence from Fifteen European Countries. IZA Discussion Paper No. 120

Polachek, S.W. (2007). Earnings Over the Lifecycle: The Mincer Earnings Function and Its Applications, IZA Discussion Papers, No. 3181, Institute for the Study of Labor (IZA).

Tansel, A., & Bircan, F. (2011). Wage Inequality and Returns to Education in Turkey: A Quantile Regression Analysis, Working Paper 1102

Tello, C., Ramos, R., & Artís, M. (2012). Changes in Wage Structure in Mexico Going Beyond the Mean: An Analysis of Differences in Distribution, IZA Discussion Paper No. 6576, 1987-2008

Wood, A. (1995). How Trade Hurt Unskilled workers. Journal of Economic Perspectives, 9: 57-80.

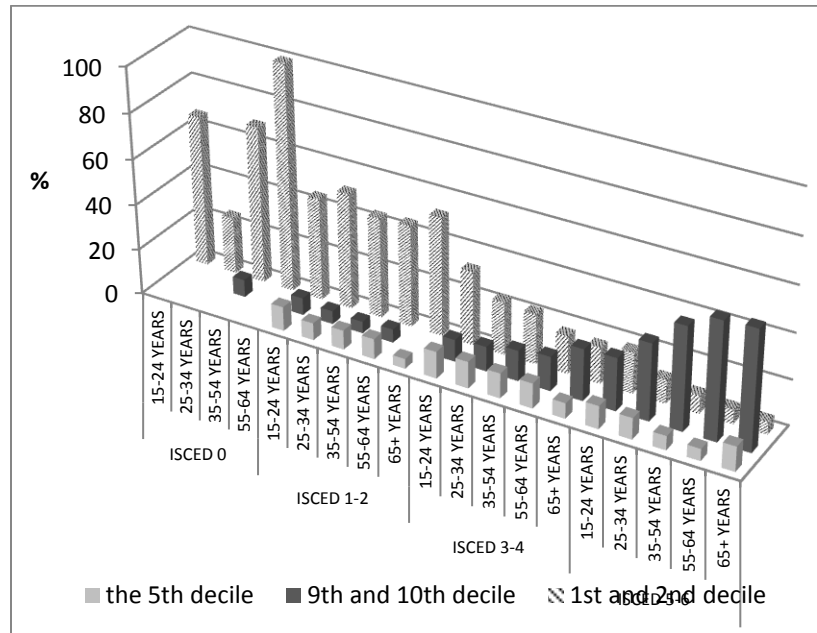


Figure 1: Relation between Decile of Net Earnings, Age and Level of Education

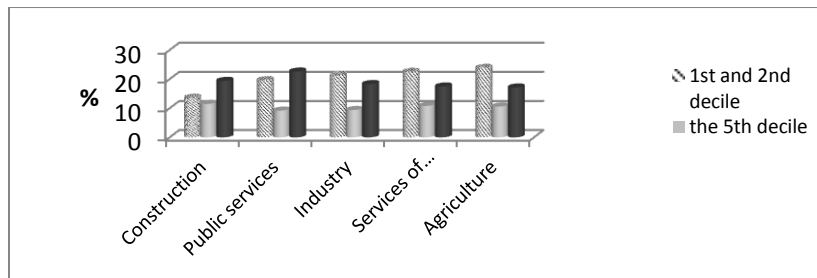


Figure 2: Decile of Net Earnings by Economic Activity

dec	b	z	P> z	%	%StdX	SDofX
male	0.85376	56.295	0.000	134.8	52.9	0.4972
urban	0.36995	22.761	0.000	44.8	18.1	0.4504
v15_24	-0.30477	-8.673	0.000	-26.3	-6.5	0.2198
v35_54	0.25957	14.591	0.000	29.6	13.5	0.4894
v55_64	0.35858	14.258	0.000	43.1	12.6	0.3314
v65	0.45470	2.753	0.006	57.6	2.0	0.0441
iscsed0	-1.79189	-3.766	0.000	-83.3	-3.2	0.0184
iscsed12	-0.99526	-35.516	0.000	-63.0	-23.1	0.2644
iscsed56	1.51915	78.921	0.000	356.8	89.7	0.4214
agric	0.08598	1.960	0.050	9.0	1.5	0.1785
industry	0.08539	4.121	0.000	8.9	4.1	0.4679
constr	0.22138	7.385	0.000	24.8	6.3	0.2764
bus_serv	-0.09080	-4.500	0.000	-8.7	-4.2	0.4744

Figure 3 Estimation Results