Theoretical and Applied Economics
Volume XVIII (2011), No. 2(555), pp. 207-218

The Impact of Investments and Gross Value Added upon Earnings*

Larisa APARASCHIVEI
National Scientific Research Institute for Labour and Social Protection,
ap_larisa@yahoo.com
Maria Denisa VASILESCU
National Scientific Research Institute for Labour and Social Protection,
antonie_denisa@yahoo.com
Nicolae CĂTĂNICIU
National Scientific Research Institute for Labour and Social Protection,
ncataniciu@incsmps.ro

Abstract. *In this paper we tried to capture the impact of investments and gross value added, but also the impact of the employment on the average wage. The analysis refers to the period 1998-2008 and we are using data on the activities of the Romanian economy. The results of this study confirm the negative influence of the employment, being consistent with the theory. Also, the impact of investments and that of gross value added came out to be positive and significant.*

Keywords: employment; wages; added value; investments.

JEL Codes: E24, J30, J64.
REL Codes: 8G, 12F, 12I.

---

* Ideas in this article were presented at the Symposium „The global crisis and reconstruction of economics?", 5-6 November 2010, Faculty of Economics, Bucharest Academy of Economic Studies.

1. Introduction

The purpose of this paper is to estimate the impact of investments, gross value added and employed population upon the average earnings in Romania, during the period 1998-2008. We used in this analysis 11 activities of the Romanian economy.

This paper contributes to the vast literature on earnings determination. The approaches, however, are very different. For example, Riahi-Belkaoui (1999) empirically tested a net value added-earnings policy model. The model was estimated annually for nonfinancial firms over the period 1976 – 1995, the results being consistent with the role of net value added and the previous level of earnings in the determination of earnings.

Concerning Romania, there is a recent study on earnings determination that tries to capture the dependence of the average wage on the unemployment rate, the number of immigrants and emigrants, the proportion of employees working in industry from the total employees, the proportion of employees working in financial and insurance activities from the total employees and the proportion of part-time employed population from the total employed population. The analysis was conducted over a period of nine years (2000-2008), using a panel data for the 42 counties of Romania**.** The most important results of the panel data estimation indicated a negative impact of the unemployment rate on wage and a significant positive effect of the immigrants on the average wage.

The paper is organized as follows: Section 2 describes the data and presents a descriptive analysis, Section 3 presents the econometric framework used for this study, while Section 4 presents the econometric model. Section 5 concludes.

2. Data analysis

The variables used in this paper are: the average monthly wage (*lrwage*), the employed population (*lpopoc*), the gross value added (*vab*) and the investments (*invest*). In order to ensure the comparability of the data for the econometric analysis, we used the consumer price index to deflate the average monthly wage. All the variables were used in the log form.

The analysis refers to the period 1998-2008. We used a panel data for 11 activities of the Romanian economy, NACE Rev 1. The source of our data was the Romanian National Institute of Statistics (the online database TEMPO).

Analyzing the evolution of the monthly average wage in Romania, for the period 1998-2008, one can observe that the wages have had an upward trend for the entire period. The highest wages are earned in financial intermediation activity, followed by public administration and defence, transport, storage and communication and education. All these activities have the average wage greater than the national average value. The smallest wages in Romania are earned by workers in hotels and restaurants activity.

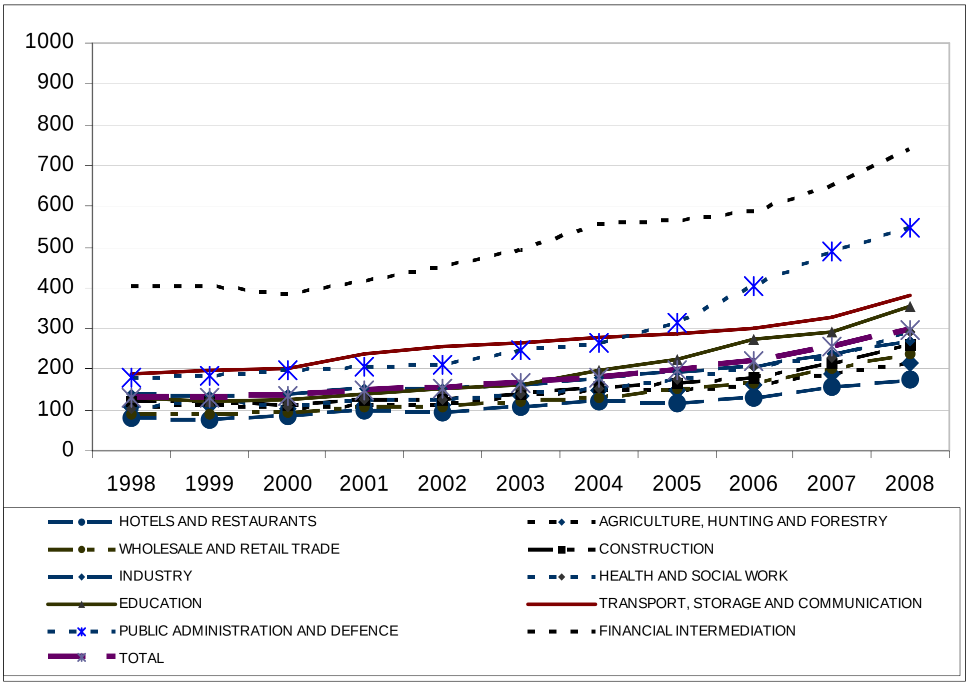

Figure 1. *The evolution of the real average wage, 1998-2008*

Regarding the investments it is very clear that the industry received the largest investments. In 2008, the value of these investments was around 31,632.5 millions RON (5,336.21 millions Ron in 1998 prices), more than two times higher than the next value, registered for wholesale and retail trade activity: 14,438.2 millions RON (2,435.63 millions RON in 1998 prices). Close to the investments in wholesale and retail trade for 2008 we find also the transport, storage and communication activity and the construction activity.

The activities for which the level of investments is low are: education, health and social work, hotels and restaurants and financial intermediation.

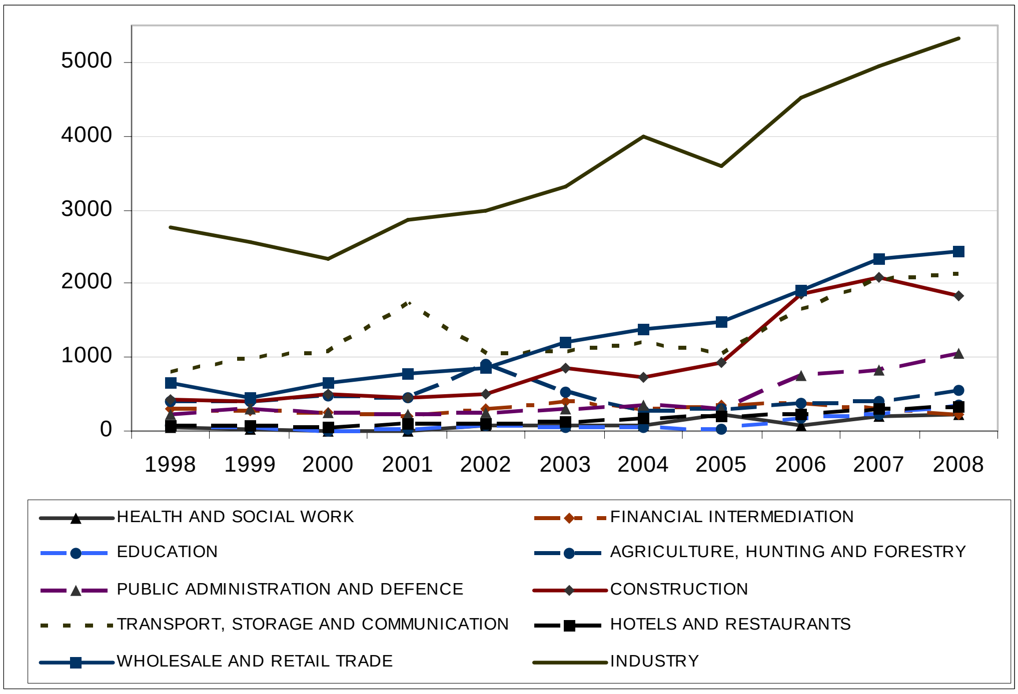

Figure 2. *The evolution of the net investments, 1998-2008*

The Gross Value Added (GVA**)** is a measure of the value of goods and services produced in an area, industry or sector of an economy. In the national accounts, GVA is output minus intermediate consumption; it is a balancing item of the national production account.

In Romania, the gross value added is calculated for each of the activities of the national economy. The highest GVA is registered for industry, which is a normal situation. Thus, in 2008, the GVA for industry was 114,873.29 million RON (19378.41 in 1998 prices). The next values are registered for transport, storage and communication activity, commerce activity and construction activity, which benefit also from considerable investments during this period of time.

The activities that generated the lowest GVA were hotels and restaurants and financial intermediation.

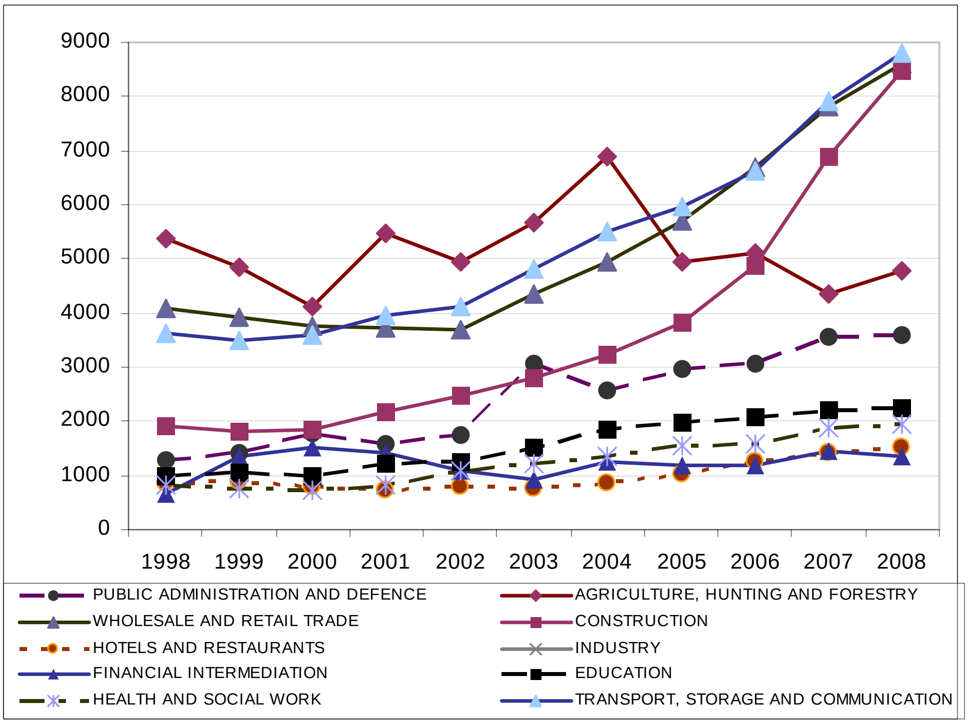

Figure 3. *The evolution of the gross value added, 1998-2008*

Employment includes, according to the methodology of "Household labour force survey", all the persons aged 15 years and over who carried out an economic activity producing goods or services of at least one hour during the reference period (the week previous to the recording) in order to get income as salaries, payment in kind or other benefits.

From the statistical data we can observe that the employed population in agriculture, hunting and forestry is declining since 1998. The same situation was registered at a different scale for industry. In 1998 were employed 2.6 million persons in industry, compared to 4.3 million persons in agriculture. In 2008, the situation was a little bit changed: in industry were employed 2.2 million persons and in agriculture were 2.7 million persons. Both activities suffered a reduction of employed population, but the decline in agriculture is more obvious.

On the other hand, in wholesale and retail trade and construction we can observe a slight increase of employment.

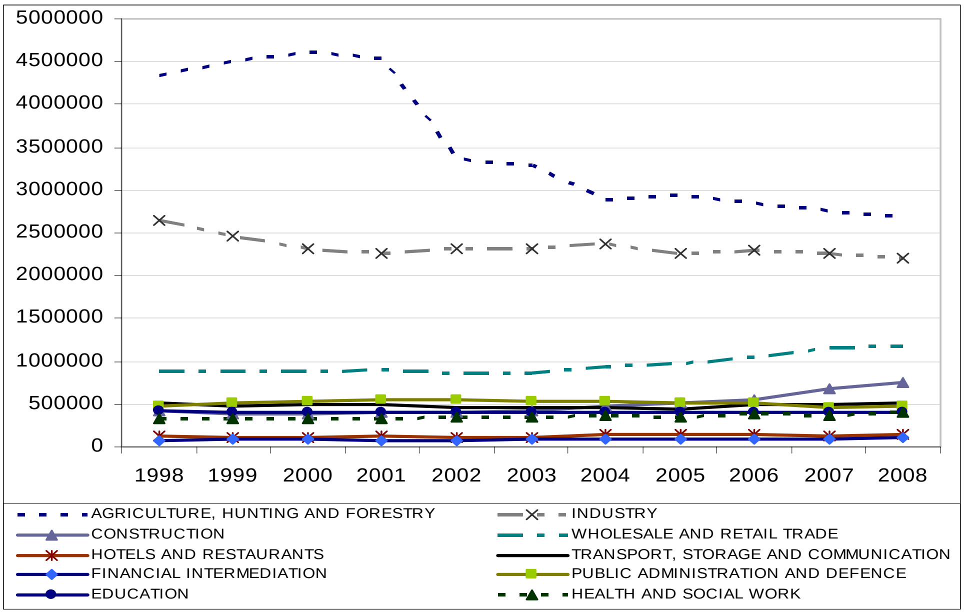

Figure 4. *The evolution of the employed population, 1998-2008*

3. Methodology

A panel data regression differs from a regular time-series or cross-section regression in that it has a double subscript on its variables. The *i* subscript denotes the cross-section dimension and *t* denotes the time-series dimension (Baltagi, 2008).

$$y_{it} = a + X_{it} \times b + u_{it} \ , \qquad i = 1, \ldots, N; \ t = 1, \ldots, T \qquad (1)$$

One of the main motivations behind pooling a time series of cross-sections is to widen the database in order to get better and more reliable estimates of the parameters of the model. The simplest poolability test has its null hypothesis the OLS model: $y_{it} = b' \times X_{it} + \alpha_i + \varepsilon_{it}$ and as its alternative the fixed-effects (FE) model: $y_{it} = a + b' \times X_{it} + \alpha_i + \varepsilon_{it}$. In other words, we test for the presence of individual effects. In Stata, if we run the *xtreg* command with the *fe* option, we obtain at the bottom of the output the F-test that all α_i=0. If we reject the null hypothesis it also means that the OLS estimates are biased and inconsistent.

Most of the panel data applications use a one-way error component model for the disturbances, with: $u_{it} = \alpha_i + \varepsilon_{it}$. There are several different linear models for panel data. The fundamental distinction between fixed-effects and random-effects models consists in the fact that, in the fixed-effects (FE) model, the α_i are permitted to be correlated with the regressors x_{it}, while continuing to assume that x_{it} is uncorrelated with the idiosyncratic error ε_{it}. On the other hand, in the random-effects (RE) model, it is assumed that α_i is purely random, which is a stronger assumption implying that α_i is uncorrelated with the regressors (Baum, 2001).

In order to decide whether a RE or a FE model is more appropriate, we can run a Hausman test. Its principle can be applied to all hypothesis testing problems, in which two different estimators are available. In the concrete case of panel models, we know that the FE estimator is consistent in the RE model as well as in the FE model. In the FE model it is even efficient, while in the RE model it has good asymptotic properties. By contrast, the RE–GLS estimator cannot be used in the FE model, while it is efficient by construction in the RE model (Kunst, 2009). In Stata, the Hausman test statistic can be properly computed based upon the contrast between the RE estimator and fixed effects (FE).

The most commonly used estimator for a FE model is the within estimator which eliminates the fixed-effect by mean-differencing. Because the within estimator provides a consistent estimate of the FE model, it is often called the FE estimator. It is also consistent under the RE model, but alternative estimators are more efficient. The fixed-effects α_i can be eliminated by subtraction of the corresponding model for individual means $\bar{y}_i = \bar{x}_i' \times b + \bar{\varepsilon}_{it}$ leading to the within model or mean-difference model:

$$(y_{it} - \bar{y}_i) = (x_{it} - \bar{x}_i)'b + (\varepsilon_{it} - \bar{\varepsilon}_i) \quad (2)$$

The within estimator is the OLS estimator of this model. Because α_i has been eliminated, OLS leads to consistent estimates of b even if α_i is correlated with x_{it} as is the case in the FE model. This result is a great advantage of panel data.

In Stata, the within estimator is computed by using the xtreg command with the fe option. The default standard errors assume that after controlling for αi, the error εit is independent and identically distributed (i.i.d) (Cameron, 2009).

Also, the command xtreg, fe estimates this model assuming that the regression disturbances are homoskedastic with the same variance across time and individuals. This may be a restrictive assumption for panels. When heteroskedasticity is present the standard errors of the estimates will be biased and we should compute robust standard errors correcting for the possible presence of heteroskedasticity. The most likely deviation from homoskedastic errors in the context of pooled cross-section time-series data (or panel data) is likely to be error variances specific to the cross-sectional unit. When the error process is homoskedastic within cross-sectional units, but its variance differs across units we have the so called groupwise heteroskedasticity. The xttest3 Stata command calculates a modified Wald statistic for groupwise heteroskedasticity in the residuals of a fixed-effect regression model (Baum, 2001).

The next step is the need to identify serial correlation in the idiosyncratic error term of the panel-data model. While a number of tests for serial correlation in panel-data models have been proposed, a test discussed by Wooldridge (2002) is very attractive because it requires relatively few assumptions and is easy to implement (Drukker, 2003). This test is implemented in Stata by David Drukker under the name xtserial. The command xtserial performs a Wald test, where the null hypothesis is no first order autocorrelation.

Stata has a long tradition of providing the option to estimate standard errors that are "robust" to certain violations of the underlying econometric model. The Stata program xtscc, implemented by Daniel Hoechle (2007), estimates pooled OLS and fixed effects (within) regression models with Driscoll and Kraay standard errors. The error structure is assumed to be heteroskedastic, autocorrelated up to some lag and possibly correlated between the groups (panels).

4. Econometric analysis

We considered the following general form of the linear regression model:

$$lrwage_{it} = a + b_1 \times lpopoc_{it} + b_2 \times lvab_{it} + b_3 \times linvest_{it} + u_{it} \qquad (3)$$

We also included in the model the first lag of the variables and we tested them for significance.

From the variance decomposition table presented in Figure 5 we can see that all the variables have within variation, which means that the FE estimation of the panel data model is appropriate. Moreover, we observe that for wages (*lrwage)* and gross value added (*lvab)* the within variation is bigger than the between variation. For investments (*linvest)* we obtained a value of 1.21 for the within variation, slightly less than the value of the between variation (1.27).

Variable		Mean	Std. Dev.	Min	Max	Observations	
lrwage	overall	6.194321	1.034195	3.906741	8.311286	N =	121
	between		.4342396	5.675914	7.205543	n =	11
	within		.9469467	4.316736	7.609135	T =	11
lpopoc	overall	13.08824	1.06279	11.1881	15.34306	N =	121
	between		1.103376	11.36246	15.05194	n =	11
	within		.1163521	12.84133	13.53121	T =	11
lvab	overall	8.832773	1.251682	6.031292	11.57601	N =	121
	between		.8334634	7.849986	10.44833	n =	11
	within		.964332	6.785973	10.53331	T =	11
linvest	overall	7.009771	1.720073	1.736255	10.28637	N =	121
	between		1.27865	5.025507	9.127047	n =	11
	within		1.208283	3.720519	9.552464	T =	11

Figure 5. *Variance decomposition*

A starting point for estimating the model is a pooled OLS regression. But we must know if pooling the data is the solution in our case. So, a poolability test is needed. The result obtained in Stata tells us to reject the null hypothesis that all α_i are zero (Figure 6). This also means that the OLS estimator is biased and inconsistent and we accept the presence of the individual effects.

```
Fixed-effects (within) regression               Number of obs      =       110
Group variable: id                              Number of groups   =        11

R-sq:  within  = 0.9735                         Obs per group: min =        10
       between = 0.0831                                        avg =      10.0
       overall = 0.6948                                        max =        10

                                                F(4,95)            =    872.43
corr(u_i, Xb)  = -0.2919                        Prob > F           =    0.0000
```

lrwage	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
lpopoc	-.4825027	.124931	-3.86	0.000	-.7305222	-.2344833
lvab	.3287556	.1069955	3.07	0.003	.1163426	.5411685
l1_vab	.5036203	.0954148	5.28	0.000	.3141979	.6930427
l1_invest	.0675143	.0308443	2.19	0.031	.0062806	.128748
_cons	4.870886	1.593512	3.06	0.003	1.707366	8.034407
sigma_u	.52491895					
sigma_e	.1396495					
rho	.93390096	(fraction of variance due to u_i)				

```
F test that all u_i=0:     F(10, 95) =     79.78             Prob > F = 0.0000
```

Figure 6. *Fixed-effects (within) regression*

The next step is to decide whether we have a fixed-effects model or a random-effects one, using the Hausman test. The probability is 0.0133, indicating that our model is more suitable for a fixed-effects estimation.

```
              ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       f            r           Difference          S.E.
-------------+----------------------------------------------------------------
      lpopoc |   -.4825027    -.5881681        .1056654        .0836841
        lvab |    .3287556     .3363458       -.0075903            .
     l1_vab  |    .5036203      .501312        .0023083         .008759
  l1_invest  |    .0675143     .0650788        .0024355        .0053139
------------------------------------------------------------------------------
                       b = consistent under Ho and Ha; obtained from xtreg
        B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(4) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       12.63
                Prob>chi2 =      0.0133
                (V_b-V_B is not positive definite)
```

Figure 7. *The Hausman test*

Concluding that we have a fixed-effects model, we continue with the estimation of our model using the within estimator. The estimated standard deviation of α_i (sigma_u) is 0.52, much bigger than the standard deviation of ε_{it} (sigma_e) which is 0.14, suggesting that the individual-specific component of the error is more important than the idiosyncratic error (Figure 6).

When performing both the modified Wald test for groupwise heteroskedasticity in the FE model, implemented in Stata by Christopher Baum, and the serial correlation test proposed by David Drukker, it resulted that the errors are both autocorrelated and heteroskedastic.

```
Modified Wald test for groupwise heteroskedasticity
in fixed effect regression model

H0: sigma(i)^2 = sigma^2 for all i

chi2 (11)  =      595.89
Prob>chi2 =       0.0000


Wooldridge test for autocorrelation in panel data
H0: no first-order autocorrelation
    F(  1,      10) =      81.592
           Prob > F =      0.0000
```

Figure 8. *Tests for heteroskedasticity and autocorrelation of the errors*

To ensure the validity of the statistical results, we performed a fixed-effects (within) regression with Driscoll and Kraay standard errors. The error structure is assumed to be heteroskedastic, autocorrelated up to some lag and possibly correlated between the groups. The author of this Stata command is Daniel Hoechle.

```
Regression with Driscoll-Kraay standard errors   Number of obs     =       110
Method: Fixed-effects regression                 Number of groups  =        11
Group variable (i): id                           F( 4,      10)    =   3772.79
maximum lag: 2                                   Prob > F          =    0.0000
                                                 within R-squared  =    0.9735
```

lrwage	Coef.	Drisc/Kraay Std. Err.	t	P>|t|	[95% Conf.	Interval]
lpopoc	-.4825027	.0999547	-4.83	0.001	-.7052158	-.2597897
lvab	.3287556	.0904774	3.63	0.005	.1271593	.5303519
l1_vab	.5036203	.068718	7.33	0.000	.350507	.6567336
l1_invest	.0675143	.0308076	2.19	0.053	-.0011293	.1361579
_cons	4.870886	1.358993	3.58	0.005	1.842861	7.898911

Figure 9. *The robust fixed-effects regression*

The resulted econometric model is:

$$lrwage_{it} = -0.48 \times \underset{(0.09)^{**}}{l\,popoc_{it}} + 0.33 \times \underset{(0.09)^{**}}{lvab_{it}} + 0.5 \times \underset{(0.06)^{**}}{l1\_vab_{it}} + 0.07 \times \underset{(0.03)^{**}}{l1\_invest_{it}} + \underset{(1.36)^{**}}{4.87} \quad (4)$$

where between brackets are the Driscoll and Kraay standard errors, while the levels of statistical significance are: ()** for the 0.05 level and ()* for the 0.1 level.

In the estimation process, we also tested the significance of lagged variables. Therefore, the final form of our model includes the first order lag of investments (*l1_invest)* and that of the gross value added (*l1_vab).*

As expected, the employed population has a significant negative coefficient, indicating that an increase with 10% leads to a decrease by 4.8% of the earnings.

Regarding the gross value added, we found a strong and positive contemporaneous influence on the earnings; also, the value of GVA from previous year has a positive impact even stronger than the present value. Thus, we obtained a cumulative effect of 0.83, meaning that a 10% increase of GVA induces a 8.3% increase of wages. The gross value added can be used in order to calculate the productivity. Therefore the GVA can be seen as an indicator of efficiency. It is very important for an economy to have a good, healthy relation between productivity and wages.

Analyzing the impact of investments on earnings, as expected, we found an one year delay between when the investments are made and the moment when their influence can be noticed. The coefficient is statistically significant and the impact is positive, but not very important as value. An increase by 10% of investments will lead to only 0.67% increase of earnings.

5. Conclusions

This paper tried to capture the impact of investments and gross value added, but also the impact of the employment on the average wage.

We made first a descriptive analysis of the average earnings, the investments, the gross value added and the employment. Therefore, we saw that the highest earnings are obtained in financial intermediation activity, which has only around 2% employees of the total employees. Also, this activity received only few investments during this period. On the other hand, the industry, which in Romania has a long tradition, is the activity that has benefited from most part of the investments over the analyzed period. Moreover, the industry is the activity that generates the most considerable gross value added in the Romanian economy.

From the econometric analysis we concluded that the investments and the GVA have a positive effect, while the employed persons have a negative impact. Another aspect of our results is that the investments have an one year delay impact upon average earnings, as expected. Regarding the influence of the gross value added, in our model we captured both the contemporary and the one year delayed component, resulting a cumulated effect of 0.83. As for the employment, after estimating the equation, we obtained a negative effect, but relatively strong (0.48). This is a normal condition, according to the economic theory.

We consider that the work presented here enriches the research on earnings from our Romania and we are confident that our future analysis will produce even more interesting and important results, given the economic crisis that already affects Romania for two years.

References

Baltagi, B.H. (2008). *Econometric Analysis of Panel Data,* John Wiley & Sons Ltd.

Christopher F. Baum, „Residual diagnostics for cross-section time series regression models", *The Stata Journal*, vol. 1, no. 1, 2001, pp. 101-104

Cameron, A.C., Trivedi, P.K. (2009). *Microeconometrics Using Stata*, Stata Press

Drukker, D.M., „Testing for serial correlation in linear panel-data models", *The Stata Journal*, vol. 3, no. 2, 2003, pp. 168-177

Hoechle, D., „Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence", *The Stata Journal*, vol. 7, no. 3, 2007, pp. 281-312

Kunst, R.M., „Econometric Methods for Panel Data – Part II", 2009, available online at http://homepage.univie.ac.at/robert.kunst/panels2e.pdf

Riahi-Belkaoui, A., „Net Value Added and Earnings Determination", *Review of Quantitative Finance and Accounting*, vol. 13, 1999, pp. 393-399

Vasilescu, Maria Denisa, Aparaschivei, Larisa, Andreica, Mădălina Ecaterina, Cătăniciu, N., „The determinants of average wage: evidence from a panel of Romanian counties", accepted for publication, 2010