Theoretical and Applied Economics
Volume XIX (2012), No. 7(572), pp. 13-24

Analysis of the Romanian employment rate. A panel data approach

Larisa APARASCHIVEI
Bucharest Academy of Economic Studies
ap_larisa@yahoo.com

Abstract. *This paper examines the evolution of the employment rate in Romania. I employ a panel data approach, considering the 42 counties, in order to explore the relationship between labour productivity, average earnings, investments and employment. The results reveal a positive impact of the average wage and gross investments and a negative impact of the labour productivity on employment rate.*

Keywords: employment; average earnings; labour productivity; panel data; regional analysis.

JEL Codes: C23, E24, J21.
REL Codes: 12G, 12I.

1. Introduction

This paper examines the evolution of the employment in Romania. The analysis is conducted over the 42 counties in Romania, for the period 2000-2008. I use a panel data approach to explore the relationship between labour productivity, average earnings, investments and employment.

Technology is important in explaining income levels across countries. Recent work has shown, however, that the major sources of technical change leading to productivity growth in OECD countries are not domestic, but abroad (Keller, 2004). One of the sources of technical change is the level of the investments, especially the foreign direct investment (Montes-Rojas, Santamaria, 2007). Therefore, if the level of the investments is growing, the labour productivity might increase, leading to different changes at the labour force market level: wages can go up and the employment down.

There is a vast literature regarding the sources of labour productivity and concerning the effects on the employment rate. A number of studies – Gali (1999), Francis and Ramey (2002), and Basu, Fernald and Kimball (2004) – have reported that favourable technology shock may reduce total hours worked in the short run. On the other hand, Mollick and Cabral (2009) examined the effects of labour productivity and total factor productivity on employment across Mexican manufacturing industries, employing panel data methods. They observed that productivity exerts a procyclical, positive effect on employment. Chang and Hong (2006) test whether technological improvement of an industry raises or lowers employment in US manufacturing; they discovered that the effect varies vastly across industries (there are far more industries in which both employment and hours per worker increase in the short run).

For the Romanian case, Aparaschivei, Vasilescu and Pirciog (2011) have tested the effect of labour productivity on employment considering 12 major activities in Romanian economy. Their results suggested a negative impact of the labour productivity on employment.

The rest of this paper is organized as follows: Section 2 explains the methodology used and Section 3 describes the data and presents the empirical results. Section 4 presents the main conclusions of this paper.

2. Methodology

The econometric analysis is based on panel data estimation, using the Stata software. A panel data regression has the form:

$$y_{it} = \alpha_i + x_{it}' \times \beta + \varepsilon_{it} \qquad i=1...N,\ t=1...\ T \qquad (1)$$

The i subscript denotes the cross-section dimension and t denotes the time-series dimension. Most of the panel data application utilizes a one-way error component model for the disturbances (Baltagi, 2008):

$$u_{it} = \alpha_i + \varepsilon_{it} \quad (2)$$

There are several different linear models for panel data. The individual fixed effects may be either assumed to be correlated with the right hand side variables (fixed effects model: FEM) or be incorporated into the error term (random effects model: REM) and assumed uncorrelated with the explanatory variables (Baum, 2001).

The first step when working with panel data is to test whether the data series can be estimated through a panel data model or through a pooled OLS. Baltagi (2008) considers that the question is “To pool or not to pool the data?”

A very simple poolability test has the null hypothesis the OLS model and the alternative hypothesis the FE model. The test practically tests the presence of the fixed effects.

The next step would be to decide whether a FE model or a RE model is more appropriate. The decision between the two models can be made based on different tests, economic reasons and/or information criteria. Baltagi suggests all of these methods; hence one can estimate both models and choose between them according to the information criteria and/or based on economic arguments. When one cannot consider the observations to be random draws from a large population – for example, if the data refers to states or provinces – it often makes sense to think of the individual effect as parameters to estimate, in which case one should use fixed effects methods (Wooldridge, 2002).

Considering the case of the fixed effects model, the estimator that is mostly used is called the within estimator. It performs OLS on the mean-differenced data. Because all the observations of the mean-difference of a time-invariant variable are zero, using a time-invariant variable is not recommended.

The fixed-effects α_i can be eliminated by subtraction of the corresponding model for individual means $\bar{y}_i = \bar{x}_i' \times \beta + \bar{\varepsilon}_{it}$ leading to the within model or mean-difference model:

$$(y_{it} - \bar{y}_i) = (x_{it} - \bar{x}_i)' \times \beta + (\varepsilon_{it} - \bar{\varepsilon}_i) \quad (3)$$

The within estimator is the OLS estimator of this model. Because the fixed effects have been eliminated, OLS leads to consistent estimates of b even if α_i is correlated with x_{it} as is the case here.

This result is a great advantage of panel data. Consistent estimation is possible even with endogenous regressors, provided that x_{it} is correlated only with the time-invariant component of the error, α_i, and not with the time-varying component of the error, ε_{it}.

For the random-effects model, the α_i from (1) is incorporated into the error term and assumed uncorrelated with the explanatory variables. Considering this assumption, and the relations (1) and (2) we have:

$$y_{it} = x_{it}' \times \beta + u_{it} \qquad i = 1...N, t = 1...T \qquad (4)$$

Because the α_i is incorporated in u_i in each time period, we might say that we have to deal with autocorrelation of the error. Therefore the general least square method is used for the estimation of a RE model.

An advantage of the RE model is that it allows the use of explanatory variables that are constant over time, but a great disadvantage is that if the FE model would be more appropriate those estimates would be inconsistent.

The default standard errors assume that, after controlling for the fixed effects α_i, the error ε_{it} is independent and identically distributed (i.i.d) (Cameron, Trivedi, 2009).

Also, the model is estimated assuming the homoskedasticity of the residuals. When heteroskedasticity is present the standard errors of the estimates will be biased and one should compute robust standard errors correcting for the possible presence of heteroskedasticity. The most likely deviation from homoskedastic errors in the context of 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.

Another problem is the serial correlation of the idiosyncratic error term in a panel-data regression. Wooldridge proposed a test for checking the autocorrelation of the residuals.

In order to account for these problems, one should estimate the regression model using robust standard errors. Some authors have provided a number of tests in order to identify the problems encountered (Drukker, 2003, Baum, 2001, Green, 2000). Thus, the *xtserial* command in STATA tests for the autocorrelation of the residuals and the *xttest3* command tests for the homoskedasticity.

For the fixed effect model, the *xtscc* command in STATA, implemented by Hoechle (2007), corrects the error structure of a model, assuming that the errors are heteroskedastic, autocorrelated up to some lag and possibly correlated between the groups (panels).

For the random effects model, the STATA command *xtgls* fits panel-data linear models by using feasible generalized least squares. This command allows estimation in the presence of AR(1) autocorrelation within panels and cross-sectional correlation and heteroskedasticity across panels.

3. Empirical results

The data

The data used in this study are: the employment rate (*empl*), the gross average earnings (*earn*), the gross investments (*invest*) and the labour productivity (*prod*). The source of the data is the National Institute of Statistics, the main collector and publisher of statistical data in Romania.

The labour productivity is considered here as the ratio of gross domestic product to labour. The data expressed in Lei – the Romanian currency – were deflated with the consumer price index (2000=100). The time period ranges from 2000 to 2008. The data refers to the 42 counties in Romania.

Descriptive analysis

Figure 1 presents the gross investments and labour productivity for each county in 2008, against the employment rate.

As can be noticed, Bucharest receives the most investments in Romania. Also, in Bucharest the labour productivity and the employment rate are relatively high. Ilfov is placed second, but far behind the capital city. As a conclusion, I cannot say that in counties with high labour productivity or high investments level, the employment level is also high.

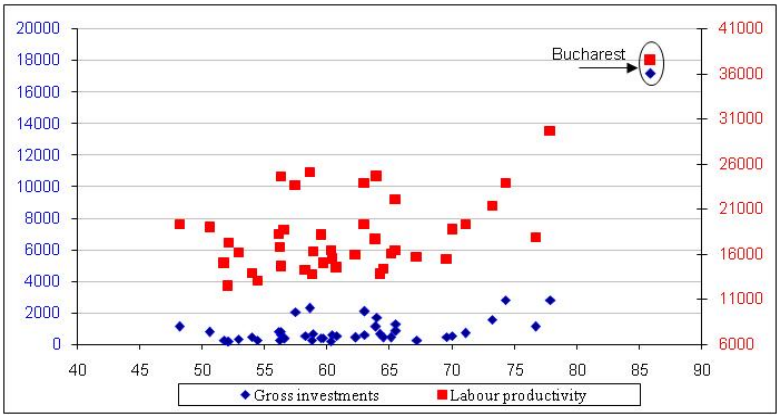

Figure 1. *The gross investments and labour productivity against the employment rate, county level, 2008*

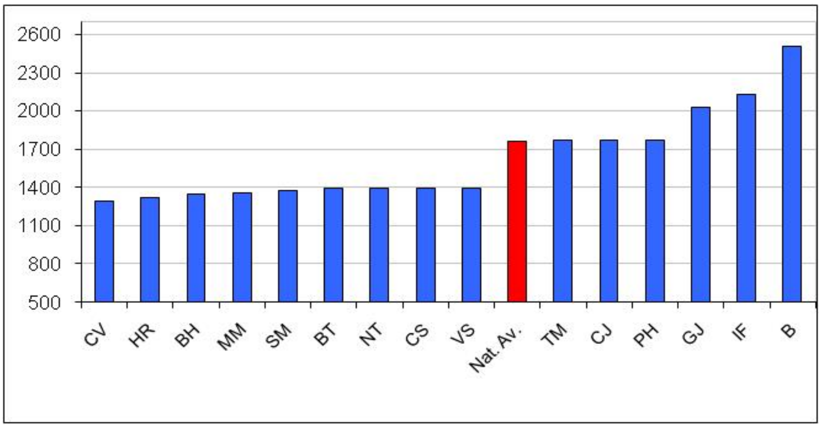

Figure 2. *The level of the average wage in 2008, current prices*

In Figure 2 I plotted the wages in current prices, sorting the counties and highlighting the national average wage. As expected, in Bucharest are earned the highest wages. In Timis, Cluj, Prahova, Gorj and Ilfov the wages are above the national average, which was 1,761 Lei in 2008. The smallest wage is registered in Covasna, being 50% lower than the average wage in Bucharest.

In Figure 3 I represented the level of wages (current prices) against the employment rate. It is easy to observe that, once again, the capital city and Ilfov County have high wages and employment rates.

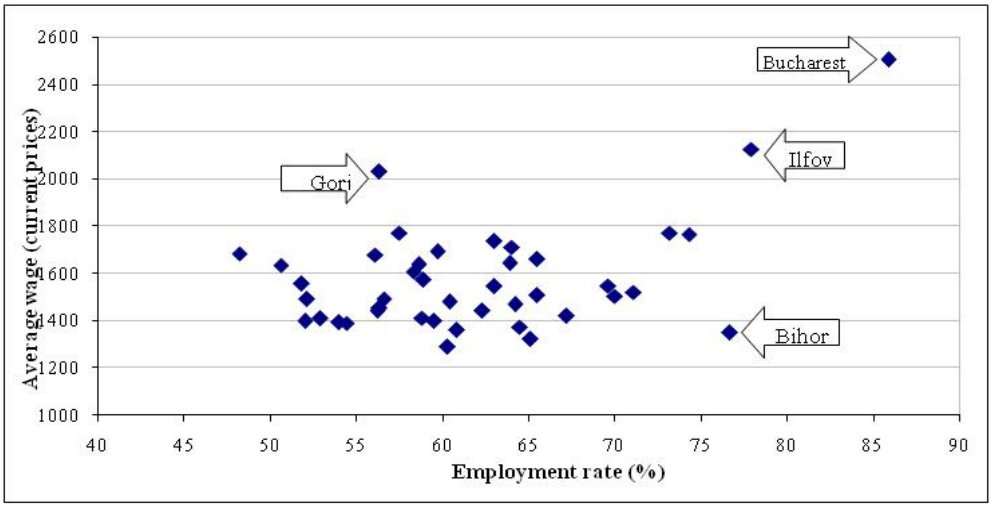

Figure 3. *The average wage (current prices) against the employment rate, 2008*

Furthermore, for counties like Gorj and Bihor I encountered a different situation: on one hand, in Gorj the employment rate is low, but the wages are high, on the other hand in Bihor the employment rate is high, but the wages are low. Once again, I cannot say if the relation between these two variables is positive or negative.

Table 1 presents the growth rates for the variables used in the analysis. For the variables expressed in Lei I used real values. One can easily observe that the employment rates did not change much in the analysed period, but I must say that in Romania the effects of the economic and financial crisis started to feel from 2009. Under these conditions, the largest decrease in the employment rate was registered in Botosani, with 19% and the highest increase was registered in Bucharest, of 47%.

Concerning the average monthly gross earnings, the variations are relatively similar between counties. The smallest increase was 78% registered in Hunedoara County and the highest in Iasi (151%).

Labour productivity growth has values between 64% (in Bucharest) and 163% (in Calarasi County). It might be surprising the fact that although in 2008 Bucharest registered the highest productivity, in terms of growth rate the capital city registered the lowest one in the analyzed period.

Regarding the evolution of the gross investments, they dropped by no more than 30% (Dolj County) and increased up to 622% (Ilfov). In Bucharest they grew by only 30%.

Table 1

Growth rates, 2000-2008

County	Employment rate	Average real wage	Gross investments	Labour productivity
Alba	-0.06	1.21	0.83	1.58
Arad	0.05	1.16	-0.08	0.94
Arges	-0.10	1.29	2.17	1.88
Bacau	-0.12	1.26	1.23	1.26
Bihor	-0.04	1.02	1.27	1.32
Bistrita-Nasaud	0.03	1.10	3.03	1.04
Botosani	-0.19	1.29	1.35	1.46
Braila	-0.10	1.23	-0.15	1.59
Brasov	-0.01	1.07	3.21	1.11
Buzau	-0.07	1.10	2.09	1.23
Bucharest	0.47	1.28	0.30	0.64
Calarasi	-0.12	1.34	1.51	1.63
Caras-Severin	-0.11	1.11	2.21	1.37
Cluj	0.10	1.30	3.54	0.97
Constanta	0.09	0.91	1.27	1.01
Covasna	-0.05	1.02	-0.02	0.66
Dambovita	-0.13	1.07	2.83	1.34
Dolj	-0.08	1.02	-0.30	1.55

County	Employment rate	Average real wage	Gross investments	Labour productivity
Galati	-0.12	0.84	1.06	1.15
Giurgiu	-0.17	1.22	4.14	1.59
Gorj	-0.15	1.07	0.02	1.50
Harghita	-0.08	1.09	1.57	0.93
Hunedoara	0.00	0.78	0.26	1.28
Ialomita	-0.03	1.06	2.06	1.12
Iasi	-0.10	1.51	1.64	1.38
Ilfov	0.20	1.23	6.22	1.32
Maramures	-0.09	1.10	2.43	1.32
Mehedinti	-0.10	1.05	3.60	1.37
Mures	-0.06	1.09	1.43	0.92
Neamt	-0.17	1.16	1.20	1.09
Olt	-0.11	0.96	2.87	0.98
Prahova	0.01	1.13	2.78	1.47
Salaj	-0.09	1.14	2.81	1.41
Satu Mare	-0.08	1.19	2.33	1.00
Sibiu	0.06	1.31	4.31	1.41
Suceava	-0.15	1.29	1.98	1.23
Teleorman	-0.16	1.01	2.62	1.32
Timis	0.07	1.48	2.61	1.37
Tulcea	-0.08	1.20	1.08	1.50
Valcea	-0.07	1.09	1.20	0.99
Vaslui	-0.18	1.18	2.39	1.59
Vrancea	-0.15	1.18	1.69	0.94

Econometric analysis

The relation estimated in this analysis has the following form:

$$empl_{it} = c + \beta_1 \times learn_{it} + \beta_2 \times linvest_{it} + \beta_3 \times lprod_{it} + \alpha_i + \varepsilon_{it} \qquad (5)$$

where:

$empl_{it}$ = the employment rate for county *i* in year *t*;

$learn_{it}$ = the natural logarithm of the average real gross earnings for county *i* in year *t*;

$linvest_{it}$ = the natural logarithm of the real gross investments for county *i* in year *t*;

$lprod_{it}$ = the natural logarithm of the labour productivity (real values) for county *i* in year *t*;

c = the constant;

α_i = individual effects;

ε_{it} = the error of the model.

As Baltagi (2008) mentioned, the question is whether to pool or not to pool the data. So, a poolability test is necessary. The results initially obtained in Stata suggests the rejection of the null hypothesis that all α_i are zero (the OLS estimator is biased and inconsistent and I must accept the presence of the individual effects and therefore a panel data estimation is better than a pooled OLS).

The next step was to run a Hausman test in order to decide whether I have a random-effects model or a fixed-effects one. The probability is less than 0.01, so I reject the null hypothesis that individual effect are random and that RE provides consistent estimates. Although the Hausman test (Figure 4) suggests the use of the FE model for this data set, I will estimate both type of models (FE and RE model). The decision to estimate both models is based on Baltagi's suggestion to also use the information criteria in order to choose between FE and RE models.

```
                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       f            r          Difference          S.E.
-------------+----------------------------------------------------------------
       learn |    14.65966     10.89781        3.761848        .6851313
     linvest |    1.018033     1.741926       -.7238936        .1596067
       lprod |   -18.62437    -15.77012       -2.854251        .5082089
------------------------------------------------------------------------------
                           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(3) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       76.14
                Prob>chi2 =      0.0000
                (V_b-V_B is not positive definite)
```

Figure 4. *The Hausman test*

The standard error component model assumes that the regression disturbances are not serially correlated and are homoskedastic. I tested both the FE and the RE model and the results are presented in Table 2.

Table 2

Results of the autocorrelation and heteroskedasticity tests

Models/Tests	**Autocorrelation** **H_0: no autocorrelation**	**Heteroskedasticity** **H_0: homoskedasticity**
The FE model	F(1, 41) = 63.861 Prob > F = 0.0000	chi2 (42) = 1855.86 Prob>chi2 = 0.0000
The RE model	F(1, 41) = 63.861 Prob > F = 0.0000	chi2 (42) = 9332.87 Prob>chi2 = 0.0000

As can be seen from Table 2, the errors of both models suffer from heteroskedasticity and serial correlation. To ensure the validity of the results, I must obtain robust estimations.

For the FE model, the STATA command *xtscc, fe* performs fixed-effects 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). For the RE model, the command is *xtgls, panels (correlated) corr (ar1)* (the error structure is assumed heteroskedastic with cross-sectional correlation and common AR (1) coefficient for all panels).

The robust estimation results are presented in Table 3.

Table 3

Estimations through FE and GLS

$empl_{it} = cons + b_1 \times learn_{it} + b_2 \times linvest_{it} + b_3 \times lprod_{it} + \alpha_i + \varepsilon_{it}$		
Models	**FE model**	**RE model**
b_1	14.66* *(5.06)*	-0.30[a] *(-0.15)*
b_2	1.02*** *(1.78)*	1.16* *(3.58)*
b_3	-18.62* (3.04)	-2.74** *(-2.08)*
cons	129.96* *(5.93)*	67.28* *(5.79)*
AIC	1823.772	2471.209
SC	1839.511	2486.948

Note: The fixed effects term (α_i) is included in the estimation but is omitted in the table. The symbols *, ** and *** refer to levels of significance of 1%, 5% and 10%; "a" denotes no significance. In parenthesis are reported the t-statistics for the FE model and z-statistics for the RE model.

For the FE model all the coefficients are statistically significant at 1% level, except the coefficient of the investments, which is significant at 10% level. For the RE model, the coefficient of the investments and the constant are significant at 1% level, the coefficient of the labour productivity is significant at 5% level. The coefficient of the average wage is not statistically significant.

The Hausman test suggested that the FE model is the appropriate model in this case, but because Baltagi recommends attention when choosing between the two models, I decided to choose between them according to the information criteria. The analysis confirmed the Hausman test result. Moreover, I cannot consider that the observations are random draws from a large population and therefore the fixed effects model is the best model to use.

Considering the FE model, I obtained that if the average wage will increase by 10%, the employment rate will also increase, but with only 1.46%. The relation between the two variables indicates that when wages are going up, more individuals are willing to search for a job and eventually get employed.

The impact of the gross investments upon the employment rate is also positive, but far smaller than that of the wages. Thus, if the investments increase by 10%, the employment rate will increase with only 0.1%.

Regarding the effect of the labour productivity, this effect is negative, in other words increases in productivity lead to labour-saving decisions as in Galí (the countercyclical productivity case). The coefficient obtained through the model was -18.62, so if the labour productivity will increase by 10%, the employment rate will decrease by 1.86%. Price rigidity prevents demand from changing (even if the marginal costs gets lower due to productivity gains) therefore firms tend to produce the same output with less labour.

4. Conclusions

This paper analyzes the relation between the employment rate and average wage, gross investments and labour productivity, considering the 42 counties in Romania. The period analyzed in this study is 2000-2008.

I decided to estimate a panel data model and the decision between the FE and the RE model was made based on the Hausman test and information criteria. In the end, the FE model was the model that fit the data properly.

The econometric results showed a positive impact of the average wage and gross investments and a negative impact of the labour productivity on employment rate, considering the panel data used.

I hope to make a contribution by using data from an emerging country to investigate the relationship between productivity, wages, investments and the employment rate.

References

Aparaschivei, L., Vasilescu, D., Pirciog, S., "The Role of Labor Productivity in the Evolution of Romanian Employment", *Proceedings of the 5th International Conference on Applied Mathematics, Simulation, Modelling*, 2011, pp. 96-99

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

Basu, S., Fernald, J., Kimball, M., "Are technology improvements contractionary?", *Working Paper 10592*, 2004

Baum, C.F., „Residual diagnostics for cross-section time series regression models", *The Stata Journal*, Vol. 1, 2001, pp. 101-104

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

Chang, Y., Hong, J.H., „Do technological improvements in the manufacturing sector raise or lower employment?", *American Economic Review*, Vol. 96, 2006, pp. 352-368

Drukker, D.M., „Testing for serial correlation in linear panel-data models", *The Stata Journal*, Vol. 3, 2003, pp. 168-177

Francis, N., Ramey, V., „Is The Technology-Driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited”, *Discussion Paper*, 2002

Galí, J., „Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?”, *American Economic Review*, Vol. 89, 1999, pp. 249-271

Green, W. (2000). *Econometric Analysis*, Prentice – Hal

Hoechle, D., „Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence”, *The Stata Journal*, Vol. 7, 2007, pp. 281-312

Keller, W., “International technology diffusion”, *Journal of Economic Literature*, Vol. 42, 2004, pp. 752-782

Mollick, A.V., Cabral, R., „Productivity effects on Mexican manufacturing Employment”, *The North American Journal of Economics and Finance*, Vol. 20, 2009, pp. 66-81

Montes-Rojas, G., Santamaría, M., „Sources of productivity growth: Evidence from the Mexican manufacturing sector”, *The North American Journal of Economics and Finance*, Vol. 18, 2007, pp. 263-278

J. M. Wooldridge, (2002). *Introductory econometrics – A modern approach*, Second Edition, South-Western College Pub, pp. 501-528