Earnings analysis: A panel data approach for the E.U. members

Denisa Vasilescu, Madalina E. Andreica, Larisa Aparaschivei, and Nicolae Cataniciu

***Abstract* —** The aim of this paper consists in quantifying the impact of several macroeconomic variables upon earnings, such as the GDP per capita, the unemployment rate, the employment rate and the foreign direct investment. The analysis was conducted over the period 1998-2009, using data for the EU-27 countries. We first applied a Hierarchical cluster analysis, which allowed grouping the E.U. members into two main clusters. Based on the two clusters formed, two panel data models were estimated.

***Keywords* —** average earnings, cluster analysis, EU-27, panel data.

I. Introduction

The aim of this paper consists in quantifying the impact of several macroeconomic variables upon earnings, such as the GDP per capita, the unemployment rate, the employment rate and the foreign direct investment. The analysis was conducted over the period 1998-2009, using data for the EU-27 countries.

The earnings are a permanent concern for both economists and policymakers. Of great importance are the factors that influence the dynamics of average earnings. Such factors include: the GDP per capita, the unemployment rate, the employment rate, the migration and the foreign direct investment.

The GDP per capita is often considered an indicator of economic well-being, so we expect that its impact on earnings to be a positive and significant one.

Foreign direct investment (FDI) is mostly seen as a driver for economic development as it may bring capital, technology, management know-how, jobs and access to new markets. Therefore, policymakers have tended to emphasize the benefits that FDI can bring to host economies, particularly in developing countries.

Between unemployment and earnings the classic theory states that there is a negative correlation. Blanchflower and Oswald [6] found an empirical regularity of a robust negative correlation between wages and log unemployment for a wide range of different countries and datasets. For the case of Great Britain, Cameron and Muellbauer [7] also observed a negative long-run effect for log unemployment on log earnings for full-time men.

The empirical evidence in support of positive wage spillovers as a result of FDI is relatively limited [17]. For instance, Aitken et al. [3] found no evidence of positive wage spillovers from FDI to domestic firms in Mexico and Venezuela, even though foreign-owned plants pay substantially higher wages. Moreover, Zhao [16] found that when labor-management bargaining is industry-wide FDI reduces the negotiated wage as well as the union employment and the competitive wage. But, if labor-management bargaining is firm-specific and unionization is industry-wide, then the above effects of FDI are substantially reduced.

Several other recent studies, however, found evidence of positive spillovers concentrating on the wage effects of FDI through its impact on labor demand and supply.

In general, the empirical evidence indicated a small and positive foreign wage premia in developed economies and potentially larger foreign wage premia in developing countries. For example, Driffield and Girma [9] found that FDI has a large positive effect on wages in domestic firms in UK electronics industry through its impact on labor demand and a small positive effect through its impact on labor supply.

Among the international studies some also quantified the determinants of wages upon European developing countries [1], [2], [13], [14], [18].

The paper is organized as follows: Section II describes the data, Section III presents the econometric framework used for this study, whereas Section IV presents the econometric results. The last section concludes.

II. Data Description

The variables used in this study are: the annual net earnings expressed in Purchasing Power Standards (*earnings*), the proportion of the net inflows of the foreign direct investment from the GDP (*fdi*), the GDP per capita expressed in Purchasing Power Standards (*gdp*), the unemployment rate (*unempl*) and the proportion of part-time employed population from the total employed population (*empl_ppt*).

The analysis was conducted over a period of twelve years (1998-2009), using macroeconomic data for all 27 members of the European Union. The main sources of our data were the Worldbank and the Eurostat databases.

Based on the fact that each of the 27 members has distinctive patterns, we decided to start our analysis with a basic data description. In order to see how each of the 27 European Union members is being situated when considering macroeconomic performances, we decided to build two bi-dimensional spaces.

The first one, presented in fig. 1, is described by the intersection of the level of the annual net earnings expressed in Purchasing Power Standards and the GDP per capita level.

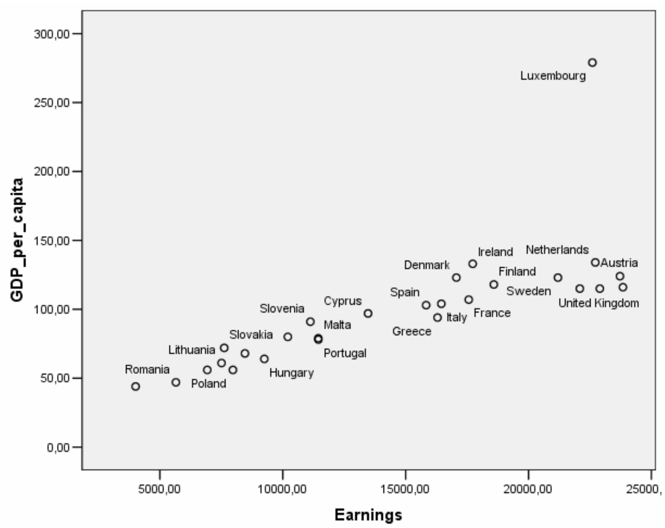

Fig. 1 The bi-dimensional Earnings - GDP per capita representation of the U.E. members

The second representation is described by the intersection of the level of the annual net earnings expressed in Purchasing Power Standards and the unemployment rate (see fig.2).

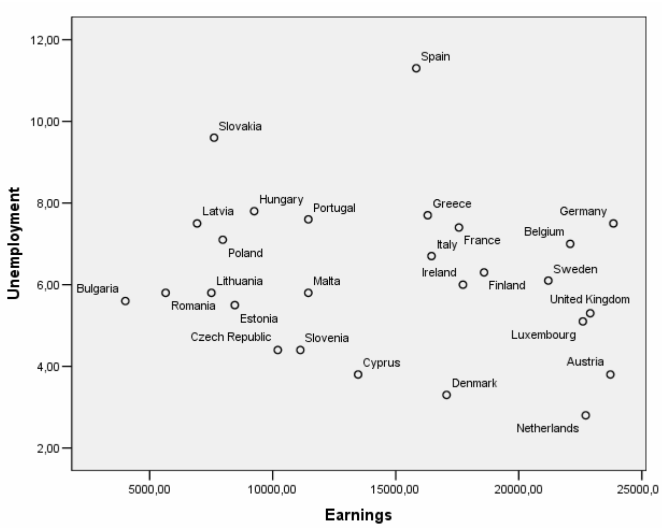

Fig. 2 The bi-dimensional Earnings - Unemployment rate representation of the U.E. members

Since from both fig.1 and fig. 2 one can easily notice that the E.U members tend to form at least two distinctive groups based on their macroeconomic performance, we decided to classify the E.U. members into two groups. In order to do that we applied a Hierarchical cluster analysis, based on a method of unsupervised learning that allows assigning a set of observations into subsets (called *clusters*) so that observations in the same cluster are similar. The cluster technique was built on the between groups linkage cluster method, whereas the intervals were calculated using the squared Euclidean distance. Based on the dendrogram (fig. 3), we notice that the 27 European Union members can be easily assigned into two main clusters, as following:

- **Cluster 1**: Malta, Portugal, Slovenia, Czech Republic, Hungary, Cyprus, Lithuania, Slovakia, Poland, Estonia, Latvia, Bulgaria and Romania;
- **Cluster 2**: Greece, Italy, Spain, France, Ireland, Denmark, Finland, Austria, Germany, Netherlands, UK, Luxemburg, Belgium and Sweden.

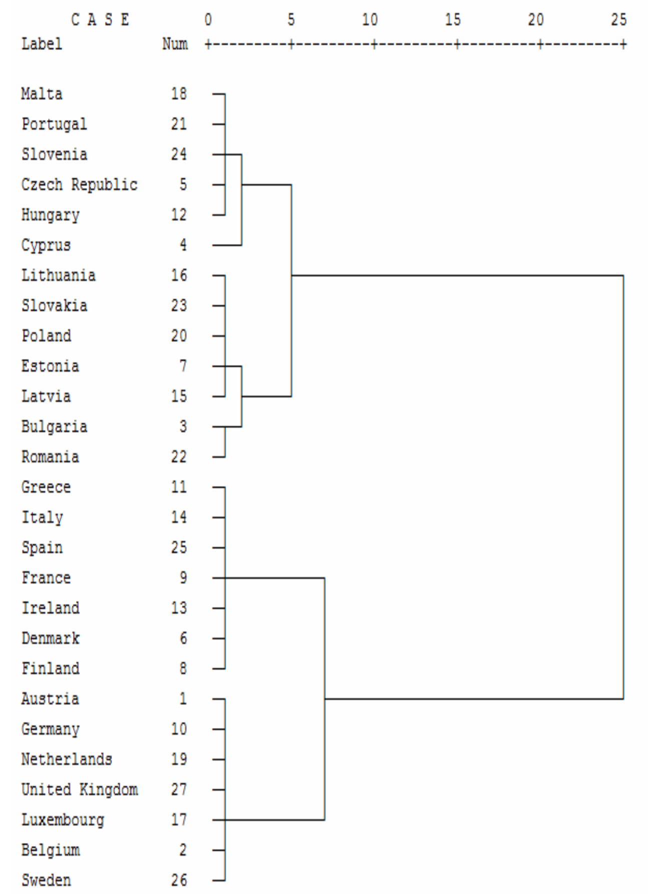

Fig. 3 The Dendrogram

III. MODELS AND METHODOLOGIES

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

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

Most of the panel data applications utilize a one-way error component model for the disturbances [4] with:

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

There are several different linear models for panel data. The main 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 [5].

The decision between the two models can be made based on different tests, from which the Hausman test is the most simple to use. The basic idea of this test starts with the hypothesis 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, whereas in the RE model it has good asymptotic properties. By contrast, the RE–GLS estimator cannot be used in the FE model, since it is efficient by construction in the RE model. The violation of the assumption $E\alpha = 0$ for the regression model leads to an inconsistency [12]. Therefore, a rejection of the null is often considered as an adoption of the fixed effects model and nonrejection as an adoption of the random effects model. And yet, Baltagi [4] recommends attention in choosing between FE and RE models based only on this test.

The most commonly used estimator for a FE model is the within estimator, which eliminates the fixed-effect by mean-differencing. 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, leading to the within model which can be estimated with the OLS method. The default standard errors assume that after controlling for α_i, the error ε_{it} is independent and identically distributed (i.i.d) [8].

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 common 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. Once again, if the errors are correlated and the estimation does not take into account this fact, the estimates will be biased. Wooldridge [15] 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 and estimation procedures in order to identify and solve this kind of problems [10], [11].

IV. EXPERIMENTAL RESULTS

In order to equalize the measuring units we decided to use natural logarithm values for average earnings and GDP per capita, as the other variables are percentage data.

The general form of the earnings equation that was estimated as a panel data is the following:

$$l\_earnings_{it} = cons + b_1 * fdi_{it} + b_2 * \lg dp\_pps_{it} + b_3 * unempl_{it} + \\ + b_3 * empl\_ppt_{it} + \alpha_i + \varepsilon_{it} \quad (3)$$

where the dependent variable was considered to be the average earnings in natural logarithm values (*l_earnings*) and the explanatory variables were: the proportion of the net inflows of the foreign direct investment from the GDP (*fdi*), the log of the GDP per capita expressed in Purchasing Power Standards (*lgdp_pps*), the unemployment rate (*unempl*) and the proportion of part-time employed population from the total employed population (*empl_ppt*). In this equation i stands for each of the European Union countries, whereas t stands for the years 1998-2009.

We also tested the statistical significance of the first lag of each variable.

Based on the two groups formed using cluster method, section A will handle the econometric model of the first cluster, whereas section B presents the econometric model of the second cluster.

A. Cluster 1

After estimating the general form of the earnings equation, we first noticed that for this group of countries the coefficients of both the proportion of part-time employed population from the total employed population and the foreign direct investment were not statistically significant. Therefore, we excluded these variables from our future estimations. Second, we decided to use the first lag of the unemployment rate.

Once we decided that we have a panel data model and not a pooled OLS, based on a poolability test, we had to choose between the FE and RE model. For that, we applied the Hausman test which suggested that a FE model would be more appropriate to describe our data. The fact that there are only 13 countries in this cluster (being impossible to say that the observations are randomly drawn from a large population) makes the decision upon the FE model more reliable.

Concluding that we have a FE model, we then checked for the presence of heteroskedasticity and autocorrelation of the errors. For this purpose we performed both the modified Wald test for groupwise heteroskedasticity, as well as the serial correlation test implemented in Stata by David Drukker. The conclusion was that a robust estimation, which takes into consideration the presence of both heteroskedasticity and autocorrelation of the residuals, is required. In order to do this we used the Driscoll and Kraay standard errors.

The robust panel data model describing the average earnings equation of the first cluster is presented in fig. 4.

```
Regression with Driscoll-Kraay standard errors   Number of obs     =       133
Method: Fixed-effects regression                 Number of groups  =        13
Group variable (i): id                           F(  2,      12)   =    646.05
maximum lag: .                                   Prob > F          =    0.0000
                                                 within R-squared  =    0.8337
```

l_earnings	Coef.	Drisc/Kraay Std. Err.	t	P>|t|	[95% Conf.	Interval]
lgdp_pps	.836383	.0375559	22.27	0.000	.7545558	.9182103
lag_unempl	-.0042162	.001426	-2.96	0.012	-.0073231	-.0011093
_cons	.8995568	.3607241	2.49	0.028	.1136065	1.685507

Fig. 4 The robust fixed-effects earnings regression for cluster 1

The econometric results indicated that the GDP per capita has the highest impact upon the average earnings. A 1% increase in the GDP per capita leads to an increase of 0.84% of the average earnings, keeping all other variables constant.

ISBN: 978-1-61804-016-9

Secondly, the influence of the unemployment rate is negative, showing that an increase of 1% of the unemployment leads to a decrease of 0.42% of the earnings. This outcome is consistent with the classic theory, which asserts the presence of a negative relationship between the unemployment and earnings.

B. *Cluster 2*

For this second group of countries, the variable that was excluded from the model is the proportion of part-time employed population from the total employed population because the coefficient obtained was not statistically significant. The final form of the model includes the GDP per capita and the first lag of the FDI and unemployment rate, as explanatory variables.

Similarly with the econometric analysis for the first cluster, we ran the Hausman test, the modified Wald test and the serial correlation test, concluding that a FE model is more appropriate than a RE model and that, once again, we had to take into account the serial autocorrelation and heteroscedasticity of the errors.

The robust panel data model is presented in fig.5.

```
Regression with Driscoll-Kraay standard errors   Number of obs     =       152
Method: Fixed-effects regression                 Number of groups  =        14
Group variable (i): id                           F(  3,      13)   =     57.93
maximum lag: .                                   Prob > F          =    0.0000
                                                 within R-squared  =    0.7623
```

l_earnings	Coef.	Drisc/Kraay Std. Err.	t	P>|t|	[95% Conf.	Interval]
lgdp_pps	.8884502	.0893788	9.94	0.000	.695359	1.081541
lag_fdi	.000387	.0000731	5.30	0.000	.0002291	.0005449
lag_unempl	-.0045513	.0022196	-2.05	0.061	-.0093465	.0002439
_cons	.7373641	.916338	0.80	0.435	-1.242264	2.716992

Fig. 5 The robust fixed-effects earnings regression for cluster 2

The econometric results indicated that the variable that influences the most the average earnings is the GDP per capita. A 1% increase in the GDP per capita leads to an increase of 0.89% of the average earnings.

Secondly, the first lag of the foreign direct investment has also a positive effect on earnings, but far smaller than that of the GDP. In this case, an increase of 1% in FDI will lead to only 0.039% increase of earnings. Our result is consistent with the empirical evidence that indicates a small and positive foreign wage premia in developed economies.

Thirdly, the influence of the first lag of the unemployment rate is negative. If the unemployment increases by 1%, the earnings will decrease by 0.46%. Similarly with the model obtained for the first cluster, the resulted coefficient confirms the classic theory that states that there is a negative correlation between these two variables.

V. CONCLUSIONS

This paper analyzes the determinants of the average earnings through a panel data of the 27 European Union countries. The variables used are the unemployment rate, the foreign direct investment, the GDP per capita in Purchasing Power Standards and the proportion of part-time employed population from the total employed population.

Considering our basic statistical analysis of the macroeconomic performance of all the 27 European Union countries (see Section 2), we noticed that the E.U members tend to form at least two distinctive groups. For this reason we decided to apply a Hierarchical cluster analysis, which allowed grouping the E.U. members into two main clusters.

Two panel data models were estimated, considering the two clusters formed. The GDP per capita turned out to be the factor that mostly influences the average earnings.

Moreover, the results of our analysis are consistent with other empirical studies asserting that the relation between the foreign direct investment and the earnings is positive. Based on our econometric models we can also say that an increase of the unemployment rate will lead to a decrease of the average earnings.

REFERENCES

[1] M. E. Andreica, M. D. Antonie, A. Cristescu, and N. Cătăniciu, „A Panel Data Analysis of the Romanian Labour Market", in *Proc of the 5th WSEAS International Conference on Economy and Management Transformation*, *Timisoara, Romania,* pp. 406-411, 2010.

[2] L. Aparaschivei, M. D. Vasilescu, and N. Cătăniciu, *The impact of investments and gross value added upon earnings*, Theoretical and Applied Economics, vol. 18, pp. 207-218, 2011.

[3] B. Aitken, A.E. Harrison and R. Lipsey, "Wages and Foreign Ownership: A Comparative Study of Mexico, Venezuela, and the United States", *Journal of International Economics*, Vol. 40, No. 3/4, pp. 345-371, 1996.

[4] B. H. Baltagi,*"Econometric Analysis of Panel Data",* John Wiley & Sons Ltd, 2008.

[5] C. F. Baum, "Residual diagnostics for cross-section time series regression models", *The Stata Journal*, vol. 1, pp. 101–104, 2001.

[6] D. G. Blanchflower, A. J. Oswald, "An Introduction to the Wage Curve, Journal of Economic Perspectives", *American Economic Association*, Vol. 9, pp 153-67, 1995.

[7] G. Cameron., J. Muellbauer, "Earnings, unemployment, and housing in Britain", *Journal of Applied Econometrics*, Vol. 16, pp. 203–220, 2001.

[8] A. C. Cameron, and P.K. Trivedi, "*Microeconometrics Using Stata*", Stata Press, 2009.

[9] N. Driffield and S. Girma, "Regional Foreign Direct Investment and Wage Spillovers: Plant Level Evidence", *Oxford Bulletin of Economics and Statistics*, vol. 65, issue 4, pages 453-474, 2003.

[10] D. M. Drukker, "Testing for serial correlation in linear panel-data models", *The Stata Journal*, vol. 3, pp. 168–177, 2003.

[11] D. Hoechle, "Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence", *The Stata Journal*, vol. 7, pp. 281-312, 2007.

[12] R. M. Kunst, (2009). Econometric Methods for Panel Data – Part II,.Available: http://homepage.univie.ac.at/robert.kunst/panels2e.pdf

[13] E. Militaru, C. Stroe, and S. Cojanu, "Labor market participation, social benefits and wages in Romania", *International Journal of Mathematical Models and Methods in Applied Sciences*, Issue 2, vol. 5, 2011

[14] M. D. Vasilescu, L. Aparaschivei, M. E. Andreica, and N. Cătăniciu, The determinants of average wage: evidence from a panel of Romanian counties, in *Proc of the 5th International Conference on Applied Statistics, Bucharest, Romania,* 2010.

[15] J. M. Wooldridge, *Introductory econometrics – A modern approach*, Second Edition, 2002, pp. 501-528.

[16] L. Zhao, "The Impact of Foreign Direct Investment on Wages and Employment", *Oxford Economic Papers*, Vol. 50, 1998.

[17] *The Impact of Foreign Direct Investment on Wages and Working Conditions*, OECD-ILO Conference On Corporate Social Responsibility, Employment and Industrial Relations: Promoting Responsible Business Conduct in a Globalising Economy, 23-24 June, OECD Conference Centre, Paris, France, 2008.

[18] *Wage Determinants in Belarus: Labor Productivity and Wage Policy,* German Economic Team in Belarus, 2006. Available: http://research.by/pdf/pp2006e04.pdf.

ISBN: 978-1-61804-016-9