Factors influencing employment in the EU

Denisa Vasilescu, Larisa Aparaschivei, and Speranța Pîrciog

Abstract—This article examines some of the key factors that influence employment, focusing mainly on the demand side of the labor market. We use panel data for the 27 members of the European Union to estimate the impact of foreign direct investment, trade openness and labor productivity on employment. The results indicate that these factors have a positive impact on the employment rate.

Keywords — employment, EU27, panel data, robust estimation.

I. INTRODUCTION

The impact of globalization on employment is a central issue of contemporary political economy. From the point of view of workers in developed countries, globalization is often seen as a threat as traditional industrial jobs disappear or are relocated around the globe. On the other hand, increased employment, as an outcome of globalization, is seen in developing countries as a major contribution to reducing poverty [12]. The impact of globalization on labor markets and the mechanisms through which closer integration with the global economy may lead to job creation are still a subject of debate.

There are a variety of ways in which globalization affects labor: the most important ones are through increased trade, foreign direct investment (FDI) and international technology transfer.

Trade liberalization is associated with both job destruction and job creation. In the long run, we would expect that the population of both emerging and developed economies to benefit from trade through higher living standards and more product choice. Also, the efficiency gains due to trade liberalization are expected to generate positive employment effects, either in terms of quantity or quality of jobs or a combination of both. In the short run, however, some adjustment costs could be related to distributional effects associated with sectoral reallocation of labor. Such adjustment costs may arise from frictional unemployment, associated with sectoral reallocation of displaced workers and any associated need for retraining; and public policies that impede the mobility of labor by slowing-down the transfer of resources from declining to expanding activities. The net employment effect in the short run depends mainly on country specific factors, such as the functioning of the labor market.

The economic literature has produced a large number of empirical studies analyzing the effects of trade on labor market outcomes. Nevertheless, so far no clear message emerges from the literature.

Most studies find a relatively small impact of trade liberalization on employment. These include Marquez and Pagés-Serra [17] for Latin American and the Caribbean and Levinsohn [16] for Chile.

In a more recent study, Peluffo [19] analyzes the impact of increased trade exposure on plants’ productivity and labor market outcomes in Uruguay finding that trade liberalization seems to increase total factor productivity, decreases employment namely for unskilled workers and reduces the gap between white and blue collar wages.

Despite various potential theoretical benefits of FDI, the majority of the empirical research has focused on its influence on economic growth. However, a highly desired outcome of attracting FDI is the potential for creating employment opportunities in the host country.

Fu and Balasubramanyam [9] studied the role of FDI in employment determination in China. They found a strong linkage between FDI and employment, as well as FDI and exports. The authors estimate that a 1% increase in FDI raises employment growth by about 3%.

The empirical results of Craigwell [6] suggest that an increase in FDI in the sample of Caribbean countries leads to higher employment. Also, FDI inflows have resulted in the transfer of new technologies and the development of specialized knowledge and skills for local managers.

Traditional Keynesian theory emphasizes the central role of demand-side factors such as monetary, fiscal, and investment shocks in macroeconomic fluctuations [13]. In contrast, real business cycle (RBC) theory puts forth technology shocks as the main drivers of business cycles. A major prediction of the RBC theory is a high positive correlation between productivity and employment. The underlying idea is that a positive technology shock increases both productivity and demand for labor, which, in turn, increases employment [3].

For the case of the United States there are studies that show no correlation or often negative correlation between productivity and employment, relation that is more consistent with the sticky prices of Keynesian models. Price rigidity prevents demand from changing in the face of lower marginal costs due to productivity gains, leading firms to produce the same output with less labor. Gali [10] finds that productivity reduced hours worked in the US as well as in all other G-7 economies except Japan. Considering the EU-15 countries, Dew-Becker and Gordon [7] make a provocative contribution to the productivity debate, arguing that there is a strong negative tradeoff between productivity and employment growth.

On the other hand, there are studies that contradict the negative effect of the productivity on employment/hours
worked. Chang and Hong [5] argue based on the aggregation of 458 four-digit U.S. manufacturing industries for the period 1958–1996. They show that technological improvements lead to increases in employment in most U.S. industries. For the 25 Mexican manufacturing industries, Mollick and Cabral [18] examine the effects of labor productivity and total factor productivity on employment using panel data methods. This panel data approach indicates that increases in TFP and VA/L ratios have had positive effects on employment.

Kim, Lim and Park [14] investigate the relationship between productivity and employment in Republic of Korea using structural vector autoregression (VAR) models. Productivity-enhancing technology shocks significantly increase hours worked, which also support the real business cycle theory.

II. DATA

We used three explanatory variables in order to analyze the employment. The basic specification of the empirical estimation equation is:

\[ \text{empl}_t = b_0 + b_1 \times \text{fdi}_t + b_2 \times \text{trade}_t + b_3 \times \text{gdp}_t + \alpha_t + \varepsilon_t \]

The dependent variable is the employment rate (empl). Fdi represents the foreign direct investment stocks as a percentage of GDP. Trade openness (trade) is defined and measured as total imports plus total exports over GDP. For the labor productivity we used as proxy the ratio between the gross domestic product and the employed persons (prod), considering the prices of 2000.

The analysis was conducted over the period 1999-2009, using data for all 27 members of the European Union. The main sources of the data were the Eurostat and the UNCTAD databases.

In figure 1 we represented the employment rate for the EU countries, figures of 2009. One can easily observe that the highest employment rates are recorded in Netherlands and Denmark (over 75%). On the other side, Malta and Hungary have the smallest employment rates (around 55%).

![Fig. 1 Employment rate in EU countries, 2009](image1)

![Fig. 2 FDI, trade openness and productivity for the EU members in 2009, sorted by the employment rate](image2)
Although Malta and Hungary have the lowest employment rates in EU, in terms of FDI as percentage of GDP, these countries are among the first (figure 2). Higher values of FDI (as % of GDP) are encountered only in Luxembourg and Belgium. Countries like Greece, Italy, Germany have registered smaller proportion of FDI from GDP, which may come as a surprise.

In figure 2 we also represented the trade openness. Once more, Hungary is among the first countries regarding this indicator. Only Belgium and Slovakia are more open to trade. Once again, countries like Greece, Spain, Italy, France, Great Britain have the smallest ratio between the commercial flows and GDP, meaning that these countries are less open to trade.

In terms of labor productivity, Malta and Hungary are among the countries less productive (Romania, Poland, Lithuania, Slovakia, Latvia, Bulgaria, Estonia, Czech Republic). Netherlands, Denmark and Sweden, countries with the highest employment rate in EU are in the middle of the distribution of labor productivity. The most productive country is Luxembourg, far ahead of countries like Great Britain, France, Austria, Belgium. Ireland comes also ahead of these countries, placing second after Luxembourg.

III. 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:

\[ y_{it} = \alpha_i + x_{it}'\beta + \epsilon_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T \]  

(1)

The subscript denotes the cross-section dimension and \( t \) denotes the time-series dimension. Most of the panel data applications utilize a one-way error component model for the disturbances, with: \( u_{it} = \alpha_i + \epsilon_{it} \) [1].

There are several different linear models for panel data. The fundamental distinction is that between fixed-effects and random-effects models. In the fixed-effects (FE) model, the \( \alpha_i \) are permitted to be correlated with the regressors \( x_{it} \), while continuing to assume that \( x_{it} \) is uncorrelated with the idiosyncratic error \( \epsilon_{it} \). In the random-effects (RE) model, it is assumed that \( \alpha_i \) is purely random a stronger assumption implying that \( \alpha_i \) is uncorrelated with the regressors [4].

**Test for poolability of the data**

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 and as its alternative the FE model [15]. In other words, we test for the presence of individual effects. Formally, we write \( H_0: \alpha_i = 0, \quad i = 1, \ldots, N \). We consider the \( F \) statistics according to the construction principle:

\[ F_{1-wop} = \frac{(ESS_{u} - ESS_{a})/(N - 1)}{ESS_{a} / ((T - 1)N - K)} \]

where \( ESS_{u} \) denotes the residual sum of squares under the null hypothesis, \( ESS_{a} \) the residual sum of squares under the alternative. Under \( H_0 \), the statistic \( F_{1-wop} \) is distributed as \( F \) with \( (N-1, (T-1)N-K) \) degrees of freedom. The two sums of squares evolve as intermediate results from OLS and from FE estimation.

In Stata, we run the `xtreg` command with the `fe` option and we obtain at the bottom of the output the \( F \)-test that all \( \alpha_i = 0 \). If we reject the null hypothesis it also means that the OLS estimates suffer from an omission variables problem and they are biased and inconsistent.

**The Hausman test**

The Hausman principle can be applied to all hypothesis testing problems, in which two different estimators are available, the first of which \( \hat{b} \) is efficient under the null hypothesis, however inconsistent under the alternative, while the other estimator \( \tilde{b} \) is consistent under both hypotheses, possibly without attaining efficiency under any hypothesis.

Hausman suggested the statistic \( m = q (\operatorname{var} q)^{-1} q \), where \( \operatorname{var} q = \operatorname{var} \hat{b} - \operatorname{var} \tilde{b} \) follows from the known properties of both estimators under the null hypothesis and from uncorrelatedness. The statistic \( m \) is distributed as \( \chi^2 \) under the null hypothesis, with degrees of freedom corresponding to the dimension of \( b \).

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, 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. The inconsistency of the RE estimator in the FE model follows from the fact that, as \( T \to \infty \), the individual fixed effects \( \alpha_i \) are not estimated but are viewed as realizations of random variables with mean zero. The violation of the assumption \( E\alpha = 0 \) for the regression model leads to an inconsistency [15].

In Stata, the Hausman test statistic can be properly computed based upon the contrast between the RE estimator and fixed effects (FE).

**Estimators for the fixed-effects model**

Estimators of the parameters \( b \) of the FE model must remove the fixed-effects \( \alpha_i \). The within estimator eliminates the fixed-effect by mean-differencing. It performs OLS on the mean-differenced data. Because all the observations of the mean-difference of a time-invariant variable are zero, we cannot estimate the coefficient of a time-invariant variable.

The fixed-effects \( \alpha_i \) can be eliminated by subtraction of the corresponding model for individual means \( \bar{y}_i = \bar{x}_i'b + \bar{\epsilon}_i \), leading to the within model or mean-difference model:

\[ (\bar{y}_i - \bar{\epsilon}_i) = (x_i - \bar{x}_i)\hat{b} + (\epsilon_i - \bar{\epsilon}_i) \]

(2)

The within estimator is the OLS estimator of this model.
Because $a_t$ has been eliminated, OLS leads to consistent estimates of $b$ even if $a_t$ is correlated with $x_t$, as is the case in the FE model. This result is a great advantage of panel data. Consistent estimation is possible even with endogenous regressors, provided that $x_t$ is correlated only with the time-invariant component of the error, $a_t$, and not with the time-varying component of the error, $e_{it}$.

Stata fits the model:

\[(y_{it} - \overline{y_t}) = a + (x_{it} - \overline{x_t})b + (e_{it} - \overline{e_t})\]

where, for example, $\overline{y_t} = (1/N)\Sigma y_{it}$ is the grand mean of $y_{it}$. This parameterization has the advantage of providing an intercept estimate, the average of the individual effects $a_t$, while yielding the same slope estimate $b$ as that from the within model. 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 $a_t$, the error $e_{it}$ is independent and identically distributed (i.i.d).

**Heteroskedasticity**

The standard error component given by equation (1) assumes 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 fixed-effects regression model estimated by `xtreg, fe` invokes the OLS estimator under the classical assumptions that the error process is independently and identically distributed [2]. The most likely deviation from homoskedasticity 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. The null hypothesis specifies that $\sigma_i^2 = \sigma^2$ for $i = 1, \ldots, N_p$, where $N_p$ is the number of cross-sectional units. The resulting test statistic is distributed Chi-squared under the null hypothesis of homoskedasticity.

**Serial correlation**

Because serial correlation in linear panel-data models biases the standard errors and causes the results to be less efficient, researchers need to identify serial correlation in the idiosyncratic error term in a panel-data model. While a number of tests for serial correlation in panel-data models have been proposed, a new test discussed by Wooldridge (2002) is very attractive because it requires relatively few assumptions and is easy to implement [8].

Wooldridge’s method uses the residuals from a regression in first-differences. Note that first-differencing the data removes the individual-level effect, the term based on the time-invariant covariates and the constant.

Wooldridge’s procedure begins by regressing $\Delta y_{it}$ on $\Delta X_{it}$ and obtaining the residuals $\hat{e}_{it}$. Central to this procedure is Wooldridge’s observation that, if the $e_{it}$ are not serially correlated, then $\text{Corr}(\Delta y_{it}, \Delta y_{i,t-1}) = -0.5$. Given this observation, the procedure regresses the residuals $\hat{e}_{it}$ from the regression with first-differenced variables on their lags and tests that the coefficient on the lagged residuals is equal to $-0.5$. To account for the within-panel correlation in the regression of $\hat{e}_{it}$ on $\hat{e}_{i,t-1}$, the VCE is adjusted for clustering at the panel level. Since `cluster()` implies robust, this test is also robust to conditional heteroskedasticity.

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.

**Driscoll and Kraay estimator**

Standard error estimates of commonly applied covariance matrix estimation techniques are biased and hence statistical inference that is based on such standard errors is invalid. Fortunately, Driscoll and Kraay (1998) propose a nonparametric covariance matrix estimator which produces heteroskedasticity consistent standard errors that are robust to very general forms of spatial and temporal dependence.

The Stata program `xtcvc`, implemented by Daniel Hoechle, 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 [11].

**IV. RESULTS**

In the econometric estimation we used the log values of labor productivity.

The first step 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 $a_i$ are zero. This also means that the OLS estimator is biased and inconsistent and we accept the presence of the individual effects.

**Table 1**

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>(E)</th>
<th>(B-E)</th>
<th>(S-E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_{it}$</td>
<td>0.019346</td>
<td>0.019346</td>
<td>0.000000</td>
</tr>
<tr>
<td>$x_{it}$</td>
<td>0.028929</td>
<td>0.028929</td>
<td>-0.000000</td>
</tr>
<tr>
<td>$z_{it}$</td>
<td>1.212605</td>
<td>1.212605</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

Tests for differences in coefficients not systematic

$H_0: \alpha = \beta$ and $\Delta \beta = \Delta \beta \epsilon \text{ is rejected}$.

$\text{Prob} \geq 0.0000$

**Fig. 3 The Hausman test**

To deal with the presence of unobserved heterogeneity across countries we used a fixed effect estimator which eliminates the country-specific effect. Moreover, the Hausman test suggested that a FE model would be more appropriate to the data (figure 3).
When performing both the modified Wald test for groupwise heteroskedasticity in the FE model and the serial correlation test it resulted that the errors are both autocorrelated and heteroskedastic (figure 4).

\[
\text{Wooldridge test for autocorrelation in panel data} \\
\text{H}_0: \text{no first-order autocorrelation} \\
F(1, 26) = 118.634 \\
\text{Pr} > F = 0.0000
\]

\[
\text{Modified Wald test for groupwise heteroskedasticity} \\
in \text{fixed effect regression model} \\
\text{H}_0: \sigma_i^2(G)_2 = \sigma_i^2 \text{ for all } i \\
\chi^2(27) = 404.13 \\
\text{Pr} > \chi^2 = 0.0000
\]

Fig. 4 Tests for heteroskedasticity and autocorrelation of the errors.

Considering all above, we estimated a robust FE regression model with Driscoll and Kraay standard errors to account for the error structure that is assumed to be heteroskedastic, autocorrelated up to some lag and possibly correlated between the groups.

The obtained model is:

\[
\text{empl}_t = -12.6689 + 0.0197 \text{fdi}_t + 0.0269 \text{fdi}_t + 7.1146 \text{gdp}_t - 7.1146 \text{gdp}_t \\
(2.8192)^* (0.0082)^* (0.0148)^* (2.6541)^* \\
\]

where between brackets are the Driscoll-Kraay standard errors and the symbols * and ** refer to levels of significance of 5% and 10%; a denotes no significance.

As can be noticed, the most important effect on employment is that of the labor productivity. If the labor productivity will increase by 10%, the employment rate will also increase by 0.7111%. This outcome is consistent with the RBC theory, which states that a positive technology shock increases both productivity and demand for labor, and therefore the employment will increase.

As theory suggests, trade openness is likely to generate positive employment effects, either in terms of quantity or quality of jobs or a combination of both. Although our results suggest that trade openness increases employment in a quantitative way, the size of this increase is relatively small: an increase of trade openness by 10% will generate an increase in employment rate by 0.269%.

The FDI effect upon employment among the EU members is statistically significant but not very strong, a 10% increase of the FDI leading to an increase of the employment rate by only 0.197%.

V. CONCLUSION

This paper analyzes some factors that influence the employment, using a panel data of the 27 European Union countries for the period 1999-2009. The explanatory variables used in this study are the labor productivity, the trade openness and the foreign direct investment.

The most influencing factor turned out to be the labor productivity, which was considered here to be the ratio between the GDP and the employed persons.

Consistent with many other empirical studies, we found that foreign direct investment has a positive influence on employment.

As for trade openness, this multi-country analysis indicated a positive impact on employment, but we must keep in mind that country-specific factors are important, and the value of any broad generalization on the link between trade openness and employment is therefore undermined.

This suggests that it would be interesting to approach country-specific studies in the search for answers regarding the key factors that influence the employment.

REFERENCES