THE DYNAMICS OF TRADE AT THE EUROPEAN UNION LEVEL. CASE STUDY – UNITED KINGDOM

Maria-Anca CRAIU

Romanian Academy, National Institute for Economic Research „Costin C.Kiritescu", Romania, marycraiu@yahoo.com

Abstract*: This paper analyzes the dynamics of trade at EU level, using panel data for the 28 Member States (MS). The relationship between GDP, population level, imports and exports for the period 2010-2017 is analyzed. Afterwords we move to a particular case, namely the United Kingdom using data series for the period 2008 - 2017and the same variables as above. In this case, the relationship between GDP, exports and population level will be analyzed.In both cases, the data source is Eurostat.*

Keywords: Trade, Economic development, Imports, Exports, Panel data.

JEL Classification Codes: C01, C23, F1, F15.

1. INTRODUCTION

Trade[1] is an important factor for economic development. Three major benefits of opening up to trade are mentioned in a European Commission working document (2010):

1) economic growth: the completion of all free trade negotiations would add more than 0.5% to EU GDP, and making progress on services and regulatory issues with major trading partners could push the value mentioned above to over 1% of EU GDP;
2) consumer benefits: trade brings a variety of goods and services to consumers and companies, at low prices. Only the consumer benefits are estimated at a level of 600 euros per year;
3) employment: 7.2% of EU employment depends directly or indirectly on exports. If all the effects of trade are taken into account (exports, imports, productivity, income effects, etc.), about 18% of the EU workforce depends on trade performance.

The degree of trade openness is a very important lever for the development of a country.

A large number of studies tried to estimate the impact of trade on economic development and growth. So taken into account the data published by Eurostat between 2008-2017, the paper analysis the importance of trade at EU level first and afterwards at UK level, trying to estimate the main correlations among variables and their influence on the dynamics of trade.

2. METHODOLOGY

2.1.Estimation of the econometric model using panel data at the level of European Union Member States

The econometric analysis is based on the estimation of a regression on panel data[2] in Eviews. A regression on panel data differs from a simple cross-sectional regression or one that uses time series in that it has a double index on its variables.

The panel has a two-dimensional structure - it contains data with a time dimension and an individual dimension.

Each observation (yit) coressponds to a certain unit (individual)[3] *i*, i = 1, …, N, at a certain moment of time *t*, t = 1, …, Ti, where *N* is the number of recorded units and *T*i is the number of time records made for unit *i*.

$Yit = \alpha i + Xit \times \beta + Uit$, i = 1, …, N; t = 1, …, T, (1) where the index *i* shows the cross-sectional dimension and the index *t* is the temporal one (Baltagi, 2008).

Panel data sets used in economic research have several advantages over cross-sectional data or time series. Usually, panel data provides more data for a parameter, increasing the degrees of freedom and reducing collinearity of the explanatory variables resulting in an increase of the efficiency of the econometric estimates (Hsiao, 2003).

The panel data series can be represented as follows:

Units (i)	**Time (t)**	**yit**	**x1it**	**x2it**	**x3it**	**…**	**xkit**
1	1	y1,1	x1;1,1	x2;1,1	x3;1,1	…	xk;1,1
1	2	y1,2	x1;1,2	x2;1,2	x3;1,2	…	xk;1,2
⋮	⋮	⋮	⋮	⋮	⋮	…	⋮
1	T	y1,T	x1;1,T	x2;1,T	x3;1,T	…	xk;1,T
2	1	y2,1	x1;2,1	x2;2,1	x3;2,1	…	xk;2,1
2	2	y2,2	x1,2,2	x2,2,2	x3,2,2	…	xk,2,2
⋮	⋮	⋮	⋮	⋮	⋮	…	⋮
2	T	y2,T	x1;2,T	x2;2,T	x3;2,T	…	xk;2,T
…	…	…	…	…	…	…	…
N	1	yN,1	x1;N,1	x2;N,1	x3;N,1	…	xk;N,1
N	2	yN,2	x1;N,2	x2;N,2	x3;N,2	…	xk;N,2
⋮	⋮	⋮	⋮	⋮	⋮	…	⋮
N	T	yN,T	x1;N,T	x2;N,T	x3;N,T	…	xk;N,T

where Yit is a recording of unit i, at a time t and with explanatory variables k (x1it, x2it…., xkit etc).

A common panel data regression model is in the form of:

$$Yit = a + bxit + \varepsilon it$$

, where Y is the dependent variable, x independent variable, a și b coefficients, i și t indexes for individuals (in our case the member states) and time.

The error εit is very important for the analysis. The hypothesis we consider in the case of the error term determine whether we are talking about fixed effects or random effects.

In the case of fixed effects models, εit is assumed to vary non-stochasticly on i or t thus causing this model to be analogous to the one-dimensional dummy variable model. In the case of random effects models, εit is assumed to vary stochastically on i or t, requiring special treatment of the error variance matrix.

The individual effects can be considered to be correlated with the variables on the right side of an equation (the fixed effects model - FEM) or they can be incorporated into the error term (the random effects model - REM) and not correlated with the explanatory variables. (Baum, 2001).

[1] Albu, 2013 demonstrates that FDI has also a direct impact on economic development

[2] Data are arranged in overlapping time series

[3] In our case the 28 Member States

When estimating a data panel, the first step is to determine whether the regression is a regular panel model or an usual regression. Baltagi (2008) considers that the main question is "To pool or not to pool the data?". The simplest test for data grouping has as the null hypothesis the usual regression model and the fixed effects model as an alternative hypothesis. In other words, the presence of individual effects is tested. The next step would be to decide which model is most suitable: one with fixed effects or one with random effects[4]. This decision can be taken based on tests, economic reasons and/or informational criteria. Baltagi suggests using all these methods; therefore, both models can be estimated and then the decision can be made taking into account informational criteria and/or based on economic arguments.

The analysis of panel data can be done through 3 approaches/methods[5]:

- Pooled OLS Regression Model
- Fixed Effect (LSDV model)
- Random Effect

We take into account 28 member states and 4 variables: GDP, population, imports and exports and the data range from 2010 to 2017, so we have 224 observations.

We want to test the relationship between exports and GDP, population, imports.

After estimating the 3 models, we can decide which model is more suitable. Thus the Hausman Test will be used.

- *Pooled OLS Regression Model*

The main problem with this model is that **it does not distinguish among countries** (we consider that all are the same), in other words we deny the heterogeneity or individuality that could exist among the 28 member states.

In this case we estimated the equation:

exports = c(1) + c(2)*gdp + c(3)*population + c(4)*imports

Dependent Variable: EXPORTS
Method: Panel Least Squares
Date: 05/19/18 Time: 09:37
Sample: 2010 2017
Periods included: 8
Cross-sections included: 28
Total panel (balanced) observations: 224

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-6004.917	3525.586	-1.703239	0.0899
GDP	-0.087485	0.018049	-4.847171	0.0000
POPULATION	-0.085829	0.363617	-0.236042	0.8136
IMPORTS	1.333059	0.033891	39.33315	0.0000

R-squared	0.985070	Mean dependent var	217845.0
Adjusted R-squared	0.984867	S.D. dependent var	283765.2
S.E. of regression	34907.94	Akaike info criterion	23.77651
Sum squared resid	2.68E+11	Schwarz criterion	23.83743
Log likelihood	-2658.969	Hannan-Quinn criter.	23.80110
F-statistic	4838.613	Durbin-Watson stat	0.077962
Prob(F-statistic)	0.000000		

Figure 1. Data results from running the *Pooled OLS Regression Model*

Source:own calculations using EViews program and based on Eurostat data

[4] Advice on the topic regarding the question on which model to be chosen is plentiful including Greene, 2012 and Wooldridge, 2010.

[5] For Panel Data Analysis and methods used see also Maddala,2001

From the data interpretation as well as by taking into account the p value for a 5% confidence level, we notice that among the independent variables, GDP and IMPORTS are significant.

- *Fixed Effect (LSDV model-Least squares dummy variables estimators) – Fixed individual effects*

Here we can talk about heterogeneity or individuality. So the 28 MS are individualized and each intercept may vary between them **but not over time**.

For the fixed effects model, the most used estimator is called the "within" estimator. The ordinary least squares (OLS) method is applied to the model obtained by eliminating the individual mean values, thus eliminating the fixed effects. Because this method eliminates the variables constant over time, their use in this model is not recommended. The fixed effects were eliminated by this method, which is why the OLS method leads to consistent estimates of the coefficients. A great advantage of panels is that consistent estimates can be obtained even with endogenous regressors, as long as they are correlated only with that component of the error that is constant over time.

Dependent Variable: EXPORTS
Method: Panel Least Squares
Date: 05/19/18 Time: 09:53
Sample: 2010 2017
Periods included: 8
Cross-sections included: 28
Total panel (balanced) observations: 224

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-59012.86	71439.18	-0.826057	0.4098
GDP	-0.006217	0.028850	-0.215503	0.8296
POPULATION	1.475217	4.407324	0.334719	0.7382
IMPORTS	1.253845	0.055311	22.66904	0.0000

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.998560	Mean dependent var	217845.0
Adjusted R-squared	0.998337	S.D. dependent var	283765.2
S.E. of regression	11573.52	Akaike info criterion	21.67866
Sum squared resid	2.59E+10	Schwarz criterion	22.15080
Log likelihood	-2397.010	Hannan-Quinn criter.	21.86924
F-statistic	4462.162	Durbin-Watson stat	0.770302
Prob(F-statistic)	0.000000		

Figure 2. Data results from running the OLS method with fixed effects
Source:own calculations using EViews program and based on Eurostat data

In this case we obtain as semnificant independent variables - IMPORTS.

- *Random Effect- Random individual effects*

If in the case of Pooled OLS Regression Model, all the individual effects are completely ignored, in this case this is solved by implementing a specific intercept in the model, which is supposed to be random. This implies complete exogenity of the model, which can be tested with the Hausmann test.

For the random effects model, the term αi from relation (1) is incorporated into the error term and is assumed not to be correlated with the explanatory variables. Given this hypothesis, the model represents a random effects model:

$$yit = xit '\times\beta + uit \; i=1,...,N,t =1,...,T \quad (2)$$

Since αi is embedded in ui for every time period, it can be said that we are dealing with the autocorrelation of errors. Therefore, the general method of OLS is based on the estimation of a random effects model. An advantage of the random effects (RE) model is that it allows the use of the explanatory variable that is constant over time; a major disadvantage, however, is that if the fixed effects model was more appropriate, the estimates obtained by the RE model would not be consistent.

Dependent Variable: EXPORTS
Method: Panel EGLS (Cross-section random effects)
Date: 05/19/18 Time: 09:59
Sample: 2010 2017
Periods included: 8
Cross-sections included: 28
Total panel (balanced) observations: 224
Swamy and Arora estimator of component variances

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	374.8258	8857.875	0.042316	0.9663
GDP	-0.007129	0.021975	-0.324419	0.7459
POPULATION	-1.908916	0.505865	-3.773570	0.0002
IMPORTS	1.266012	0.048001	26.37459	0.0000

Effects Specification

	S.D.	Rho
Cross-section random	34982.75	0.9013
Idiosyncratic random	11573.52	0.0987

Weighted Statistics

R-squared	0.945424	Mean dependent var	25308.32
Adjusted R-squared	0.944680	S.D. dependent var	49459.31
S.E. of regression	11632.93	Sum squared resid	2.98E+10
F-statistic	1270.364	Durbin-Watson stat	0.685260
Prob(F-statistic)	0.000000		

Unweighted Statistics

R-squared	0.983131	Mean dependent var	217845.0
Sum squared resid	3.03E+11	Durbin-Watson stat	0.067352

Figure 3. Data results from running the model using the random effects
Source:own calculations using EViews program and based on Eurostat data

In this case we obtain as semnificant independent variables POPULATION and IMPORTS.

Hausman Test:
We perform this test for the previous regression, the one where we considered random effect at the estimated equation. We have as working hypothesis:

Null Hypothesis: Random effects model is suitable
Alternative hypothesis: The fixed-effects model is suitable

Correlated Random Effects - Hausman Test
Equation: Untitled
Test cross-section random effects

Test Summary	Chi-Sq. Statistic	Chi-Sq. d.f.	Prob.
Cross-section random	5.264514	3	0.1534

Figure 4. Results obtained from running the Hausman Test
Source:own calculations using EViews program and based on Eurostat data

Taking into account the p-value (15,3%>5%), the null hypothesis cannot be rejected.
Thus, if we return to the random effects model (random effects) and analyze the coefficients of the significant independent variables, we notice that the population (POPULATION) has a negative coefficient and the imports (IMPORTS) a positive one, thus the fluctuations of these 2 variables produce changes in the balance of exports.

2.2. Estimation of the econometric model using data series. Case study - United Kingdom

We use as data exports, imports, GDP, population over the period of time 2008-2017. Initially we opened the data series for exports, GDP and population as a data group and by using the graphs it can be noticed that exports are following the GDP trend.
As the number of observation is limited, the accuracy of the results is relative.

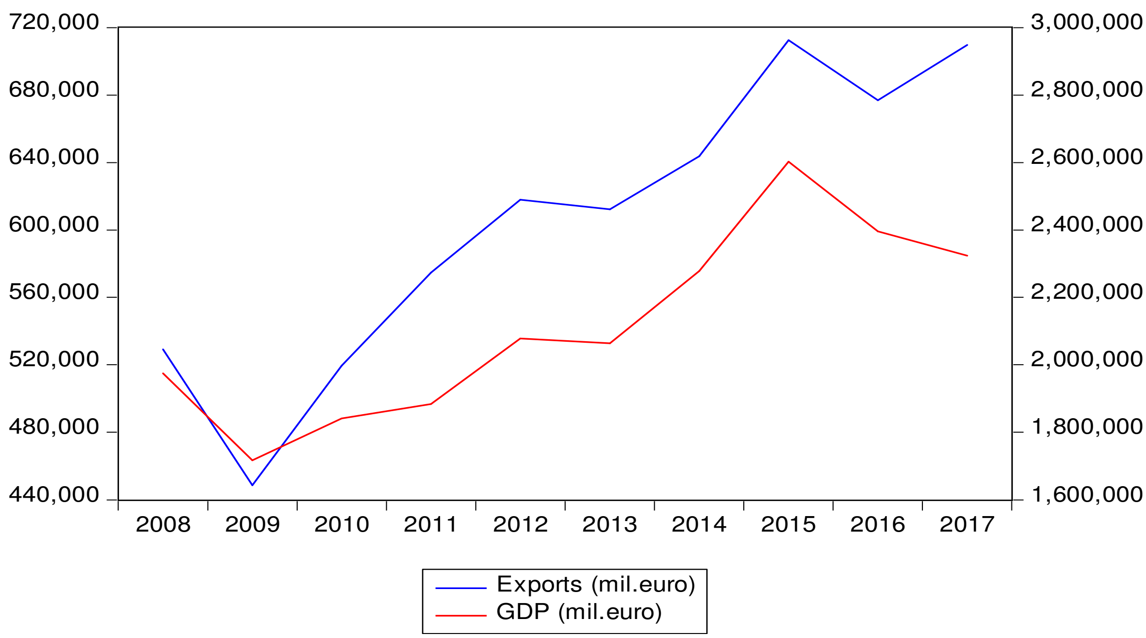

Figure 5. The evolution of Exports and GDP in UK during 2008-2017
Source:own calculations using EViews program and based on Eurostat data

The regression equation applied: exportsuk = c(1) + c(2)*gdpuk + c(3)*populationuk

Dependent Variable: EXPORTSUK
Method: Least Squares
Date: 05/19/18 Time: 19:11
Sample: 2008 2017
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1620523.	590116.3	-2.746108	0.0287
GDPUK	0.163320	0.055078	2.965235	0.0209
POPULATIONUK	29.39644	10.72650	2.740544	0.0289

R-squared	0.939035	Mean dependent var	604512.0
Adjusted R-squared	0.921617	S.D. dependent var	86789.72
S.E. of regression	24298.53	Akaike info criterion	23.27754
Sum squared resid	4.13E+09	Schwarz criterion	23.36832
Log likelihood	-113.3877	Hannan-Quinn criter.	23.17796
F-statistic	53.91021	Durbin-Watson stat	1.677694
Prob(F-statistic)	0.000056		

Figure 6. Results of the model obtained from running the regression (UK)
Source:own calculations using EViews program and based on Eurostat data

Both GDP and population are significant independent variables that together with the constant from the model (which also includes the factors not specified by the model) help explaining the dependent variable (exports = EXPORTSUK) in a proportion of almost 94%.

From a statistical point of view it is a good model because:

1) R^2 = 93% of the variation of exports compared to the average can be explained by the model, the rest of 7% is due to the influence of other factors that are not included in the model;

2) P value for GDPUK (2.09%) and POPULATIONUK (2.8%) are less than 5%, so the 2 independent variables are significant and **individually** influence the dependent variable;

3) The P-value of F-statistic that represents the **jointly** influence of the 2 dependent variables is 0.0056%, therefor significant;

4) The sign of the coefficients should follow the economic theories, expectations or intuition.

In the case of GDP[6] for an increase of one unit, exports increase by 0.16 units, and for the population[7] for an increase of one unit, exports increase by 29.3 units.

exportsuk = -1620522.93234 + 0.163319637514 * gdpuk + 29.3964395142 * populationuk

5) The residuals must follow certain characteristics which will be detailed later.
From the graph below, it can be noticed that the most obvious differences are in 2009, 2011, 2012 and 2016.

[6] *GDP = CF + FBCF + VS + (E - I) where: CF = final consumption; FBCF = gross fixed capital formation; VS = stock change; E = value of exports; I = value of imports.*
[7] A large population means higher production, possibly in those branches where the country can have a competitive advantage, so it is in a position to export those goods and/or services

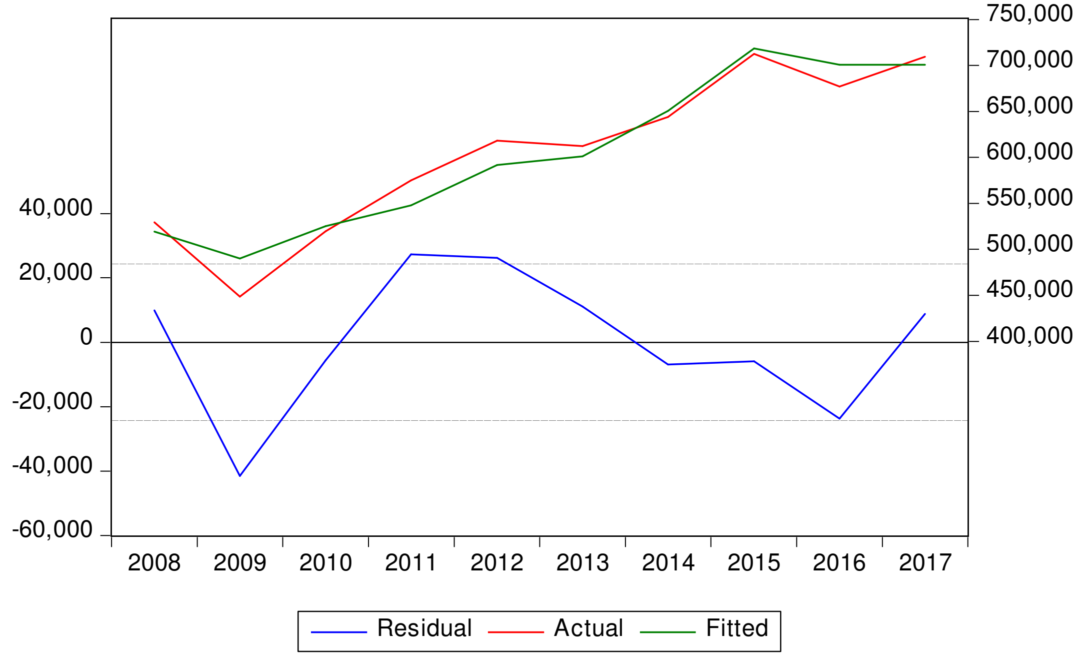

Figure 7. Actual and fitted values for the dependent variable of the model, along with the residuals

Source:own calculations using EViews program and based on Eurostat data

Possible explanations would be that UK officially entered into recession in the last quarter of 2008, for the first time after 1991. Only in the third quarter of 2009 a positive GDP growth rate (+ 0.2%) determined the exit from the recession and the beginning of a period of economic recovery. However, as expected, due to the financial turmoil of the euro area, as well as other shocks on the international market, the recovery of the British economy has proved extremely slow, with prospects of reaching again the economic levels precedings 2008 only after the year 2014.[8]

The explanation for 2016 could be the referendum itself, by which the UK decided to leave the European Union and the uncertainty about future relations with the EU made Britan's economy vulnerable.

Also, in the case of residuals, certain criteria must be met in order to state that the model is suitable (fitted):

5.1. There should be no serial correlation/autocorrelation. This will be tested by the Serial Correlation LM TEST[9] - (Breusch-Godfrey Test) which has as working hypothesis:

Null hypothesis: No correlation between residuals
Alternative hypothesis: There is a correlation between residuals.

[8] MINISTERUL ECONOMIEI - Biroul de Promovare Comercial Economică Londra - Îndrumar de afaceri Regatul Unit al Marii Britanii și Irlandei de Nord, 2014 (Ministry of Economy, Bureau of Economic Trade Promotion - *Business guide for United Kingdom of Great Britain and Northern Ireland)*

[9] Because the sample is not too big we use only 1 lag

Breusch-Godfrey Serial Correlation LM Test:
Null hypothesis: No serial correlation at up to 1 lag

F-statistic	0.124812	Prob. F(1,6)	0.7360
Obs*R-squared	0.203780	Prob. Chi-Square(1)	0.6517

Figure 8. Results obtained from running the Breusch-Godfrey Test

Source:own calculations using EViews program and based on Eurostat data

The p value (65%)>5%, so we cannot reject the null hypothesis stating that there is no correlation between residuals.

5.2. There is no heteroskedasticity – so in this case to test it, we use among the dedicated tests, Breusch-Pagan-Godfrey Test which has as working hypothesis:

Null hypothesis: Homoskedasticity in the case of residuals
Alternative hypothesis: Heteroskedasticity in the case of residuals

Heteroskedasticity Test: Breusch-Pagan-Godfrey
Null hypothesis: Homoskedasticity

F-statistic	1.532553	Prob. F(2,7)	0.2805
Obs*R-squared	3.045279	Prob. Chi-Square(2)	0.2181
Scaled explained SS	1.154503	Prob. Chi-Square(2)	0.5614

Figure 9. Results obtained from running the Breusch-Pagan-Godfrey Test

Source:own calculations using EViews program and based on Eurostat data

The p value (21,8%)>5%, so we cannot reject the null hypothesis stating that there is homoskedasticity (so no heteroskedasticity in the model).

5.3. The residuals follow a normal distribution - the Jarque-Bera test is used[10]

Jarque-Bera represents a statistic of a test to verify the hypothesis that the series is normally distributed. The statistic is a measure of the distance between the Skewness and Kurtosis indicators of the analyzed series compared to the normal distribution and is calculated as follows:

$$JB = \left[\frac{N}{6} \quad S^2 + \frac{(K-3)^2}{4} \right] \tag{2}$$

where S is the Skewness indicator and K is the Kurtosis indicator. Under the null hypothesis of a normal distribution, the Jarque-Bera statistic is distributed χ^2 with 2 degrees of freedom.

The reported probability is the probability that the Jarque Bera statistic will exceed (in absolute value), the value observed under the null hypothesis - a small value of this probability indicates the rejection of the null hypothesis meaning that the series has not a normal distribution.

[10] See Jula et al. 2018 for examples

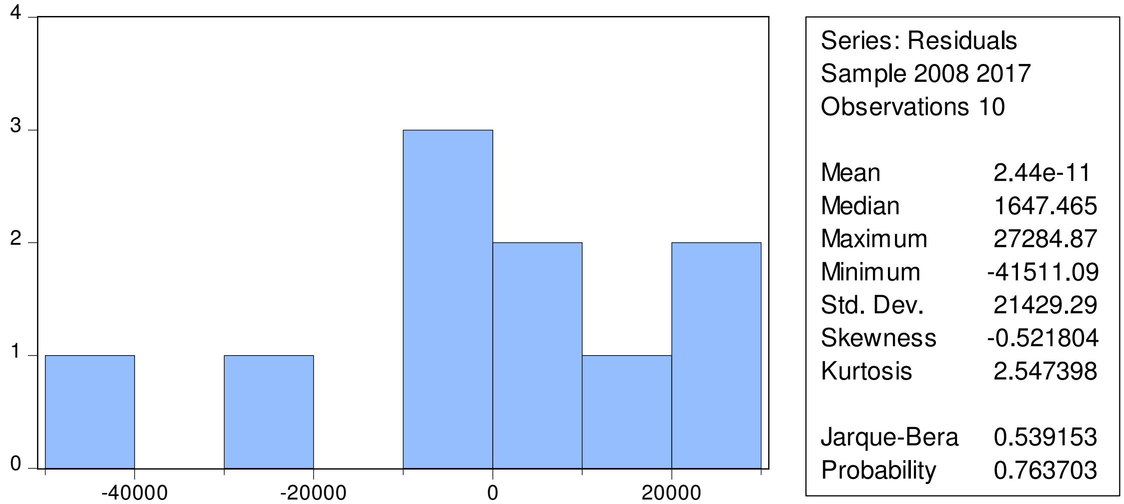

Figure 10. Results obtained from running the Jarque-Bera Test

Source:own calculations using EViews program and based on Eurostat data

P value >5%, so we cannot reject the null hypothesis that states that residuals have a normal distribution.

3. CONCLUSIONS AND REMARKS

The paper follows the evolution of important macroeconomic indicators for the EU, analyzing the changes caused by the crisis in recent years and trying to find out which variables affect exports.

Initially a panel data type model was used, taking into account the 28 EU Member States.

The analysis of time series allows to study the evolution in time of the relations between variables, without taking into account the individual characteristics of the elements that determine those evolutions. Data analysis in cross-sectional structures studies the effect of individual diversity, but does not take into account the dynamic behavior of entities. Here are some advantages of using panel data (Baltagi, 2005):

- Allows the analysis of the individual dynamics and the effects generated by the ordering in time of the events (separation of individual and/or time effects).
- Allows control of the heterogeneity induced by time-invariant variables, or across transverse structures, whether or not the respective variables are observable (latent).
- The models allow, for example, a better examination of the dynamics of the adjustments and allow the study of more complex behaviors.
- From an econometric perspective, the number of observations, therefore the degrees of freedom increase, which leads to the increase of the tests power, the consistency and efficiency of the estimators and the reduction level of collinearity between the variables.

Also, the distortion induced by the aggregation of activities, companies or individuals is reduced or eliminated.

The main problems that arise regarding the use of panel data in econometric models are (Baltagi, 2005):

- Issues related to panel structure design, data collection and management. These problems mainly concern: sampling errors or unrelated sampling errors.
- Short dimension of time series in the panel. Typically, panel data covers a short period for each unit (individual). Even though, from the technical point of view, the asymptotic properties of the statistics are based on the number of units in the sample, this characteristic generates additional difficulties in the case of econometric models with qualitative dependent variables in the panel (Baltagi, 2005).
- Dependency issues may occur between cross-sections (series) of the panel.

The econometric model was estimated including all variables initially considered: GDP, population, imports and exports. Depending on the model applied, significant variables were highlighted. We concluded that the right model is the random effects model where we obtained as significant variables only imports and population.

In the case study (it was considered the United Kingdom) where we used data series, the regression model applied proved to be a good statistical model, leading to the explanation of the dependent variable (ie UK exports) through the variation of GDP and population. The model also highlighted the influence of the financial crisis period of 2008, which emerged in the real estate and banking sector in the US and quickly spread to markets around the world.

The result was a global recession and almost an unprecedented global trade contraction, with a strong focus on demand in those key sectors, especially for long-term investment and investment goods, which accounted a great part of the international and EU trade.

REFERENCES

1. Albu, L. L., 2013, *Foreign Trade and FDI as Main Factors of Growth in the EU*, Journal for Economic Forecasting, vol. 0(2), pp. 7-17.
2. Baltagi *B.H.,2005, Econometric Analysis of Panel Data,*Third Edition, Wiley, ISBN 0-470-01456-3, pp.1-29,53-61,79-104,124-128,237-253
3. Baltagi B.H.,2008, *Econometric Analysis of Panel Data,* Fourth Edition,Wiley, ISBN-10: 0470518863,ISBN-13: 978-0470518861, pp.1-22,35-42,63-81
4. Baum, C.F., 2001, *Residual diagnostics for cross-section time series regression models*, The Stata Journal, 1, Number 1, pp. 101–104
5. Dobrescu, E, 2000, *Macromedels of the Romanian Transition Economy*, Expert Publishing House
6. European Commission, 2010, *"Trade as a driver of prosperity",* http://register.consilium.europa.eu/doc/srv?l=EN&t=PDF&gc=true&sc=false&f=ST%2016183%202010%20ADD%202
7. Greene, W.H. 2012, *Econometric Analysis*, Seventh Edition, Prentice Hall, ISBN 0-13-139538-6, https://www.academia.edu/37844952/Greene_2012_Econometric_Analysis_7th_Edition

8. Hsiao C., 2003, *Analysis of Panel Data*, 2nd Edition, Cambridge University Press

9. Jula, D., Jula, N.M, 2018, *Econometrie (Econometrics)*, Ed.Mustang ,ISBN 978-606-652-132-1, pp.17-51,115-135,149-154,243-327

10. Jula N., M., Jula D., 2017, *Modelare economică: Modele econometrice și de optimizare (Economic modeling: econometric and optimization models)*, Mustang Publishing House

11. Maddala, G. S. (2001), *Introduction to Econometrics* (Third ed.). New York: Wiley, ISBN10 0471497282, ISBN13 9780471497288

12. MINISTERUL ECONOMIEI Biroul de Promovare Comercial Economica (Ministry of Economy, Bureau of Economic Trade Promotion*)*, Londra, Editor (nu are, e scris de cineva din interior) Red. Cecilia Deaconescu – Ministru Consilier 1- Î*ndrumar de afaceri Regatul Unit al Marii Britanii și Irlandei de Nord* (*Business guide for United Kingdom of Great Britain and Northern Ireland*)- 2014

13. Wooldridge J.M., 2010, *Econometric Analysis of Cross-Section and Panel Data*, 2nd Edition, MIT Press, https://jrvargas.files.wordpress.com/2011/01/wooldridge_j-_2002_econometric_analysis_of_cross_section_and_panel_data.pdf