Correlation between Total Unemployment and Expenses with Unemployed Social Protection in Romania

Gabriela – Mariana Ionescu[1]

Abstract: The *goal* of the study is to estimate the correlation between two important macroeconomic variables in the labour market: the total unemployment rate over the active population (as dependent variable), respectively the rate of social protection expenditures with the unemployment over GDP (as independent or causal variable) in Romania. In order to achieve the goal, the following *objectives* were achieved: a) the formation of time series and the statistical analysis for the two variables; b) designing a mono-factorial linear econometric model; c) estimation of the econometric model. The *empirical data* used are quarterly data between January 1997 and December 2016. The *data quality* was provided by using the following sources: EUROSTAT, respectively the Romanian Statistical Yearbook. The main scientific results of the study are as follows: 1) statistically the time series of the two variables are stationary only at the level of the first order differences, the series of the independent variable is non-normal, and the dependent variable is quasi-normal; 2) the estimation of the quantitative model shows that there is a direct correlation (though weak, about 3%) between the dependent and the independent variable. At the same time, there are no influences outside the independent variable (the free term of the first degree equation is approximately - 0.002). In conclusion, it can be said that the variation of the total unemployment rate in the active population is direct, albeit relatively weak, influenced by the change in the rate of social protection expenditures in GDP.

Keywords: unemployment total rate; social protection expenditures; stationarity; first order differences; seasonality.

JEL Classification: C10; J64; O11

1. Introduction

To achieve the goal of the study two macroeconomic variables which are working on the labour market are taken into consideration, namely: a) the total unemployment rate over the active population (as dependent variable); b) the rate of social protection expenditures with the unemployment over GDP (as independent or causal variable) in Romania.

The correlation analysis between the two variables analysed in this paper considers two aspects. Firstly, the unemployed existence (those registered at the National

[1] The School of Advanced Studies of the Romanian Academy (*SCOSAAR*), Romania, Address: Calea Victoriei 125, Sector 1, 010071 Bucharest , Romania, Corresponding author:

Agency for Employment of Romania) directly leads to the social protection costs formation of the unemployed. Carrying out these costs has the effect (at least an in intention or expectation) or reducing unemployment (Mankiw, 2015), for example by finding jobs as a result of retraining funded by these costs.

Secondly, the unemployment benefit, as a component of social protection expenditure for the unemployed, may encourage the unemployment maintenance (by creating moral hazard). Therefore, it is obvious that the two variables are correlated.

The independent variable X of the mono-factorial linear regression model, the rate of the social protection expenditure for the unemployment refers to the total expenditure related to the unemployment, calculated as ratio between social protection costs of the unemployment and GDP.

The dependent variable Y of the mono-factorial regression model, thus the unemployment rate refers to the total unemployment (both short and long term), calculated as ratio between unemployed number and active population. The data series includes quarterly registration expressed as a percentage (data source being EUROSTAT).

2. The Topic Importance/Relevance

The economic product distribution in society is a fundamental issue, and also a questionable problem of economic discourse, both in the conceptual and methodological aspect (Dinga, E., 2011), and under the social policy too. The common research does not visit this issue from the perspective of the correlation between the total rate of unemployment over active population on the one hand, and the rate of social protection expenditures over GDP. However, it is possible that the unemployment benefits exert a behavioural impact on the kinematics of the total rate of unemployment. In such a context, the study is aimed at to examine a possible correlation (maybe even a possible causality) between the evoked variables.

3. Short Statistical Analysis of the Time Series

The (kinematic) evolution of the two variables, the independent variable X, and the dependent variable Y, for the analysed period shows that, in the case of the independent variable X, the quarterly values of social protection expenditure for the unemployed, as a GDP share, has a long-term downward trend. At the same time, there is a strong oscillation on the medium and short term. This oscillation can be interpreted either as a result of the change in the unemployment rate, or as an effect of budgetary policy decisions.

The quarterly values of the unemployment rate (Statistical yearbook of Roumania, 2017), as a share of the number of unemployed people over the active population, exhibit two behaviour patterns: a) permanent increases (although there are short-term oscillators) on the 1997 – 2002 period, respectively between 2008 – 2013; b) permanent decreases (although there are short-term oscillators) on the 2002 – 2008 period, respectively between 2013 – 2016. On 2008 there is an extremely sharp reduction in the unemployment rate, the result of substantial economic growth of that year (starting next year, with the international financial crisis in Romania, unemployment is rising).

From the perspective of the simultaneous evolution of the two variables, it can be noticed that on the 1997 – 2002 period the two variables evolve in a contradictory way (while the unemployment rate is steadily rising, the social expenditure rate with the unemployment protection decreases, except for the beginning of the period, where increases and decreases alternate). Starting with 2012, both variables decrease almost concerted as shown in figure 1.

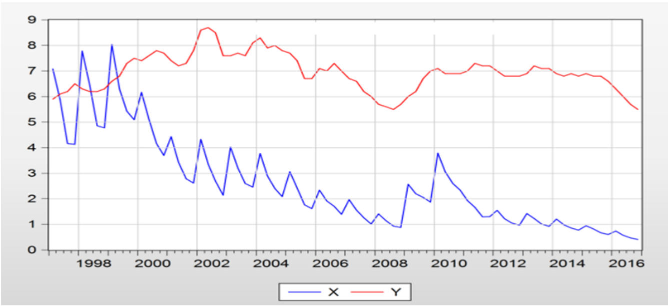
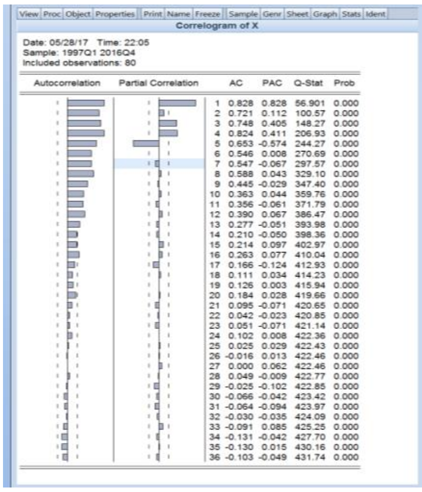


Figure 1. Simultaneous kinematics of the two variables

Source: author's work

Stationary analysis for the independent variable X, for observed values (level values), was done using the correlogram, and Dickey-Fuller test (Books, C., 2014). From the correlogram graph (which is decreasing depending on the number of lags considered), as shown in figure 2a, can be concluded that the series of the independent variable X at the observed values (level values) is non-stationary. In order to check rigorously the stationary/non-stationary character of the series, the unit root test was done. It can be observed (figure 2b), that the calculated value of the test (-3.001) is greater than the tabulated values for all three levels of significance (1%: - 4.09; 5%: - 3.47; 10%: - 3.16).

Consequently, the time series of the independent variable X for the observed values (level values) is non-stationary (the null hypothesis is accepted).



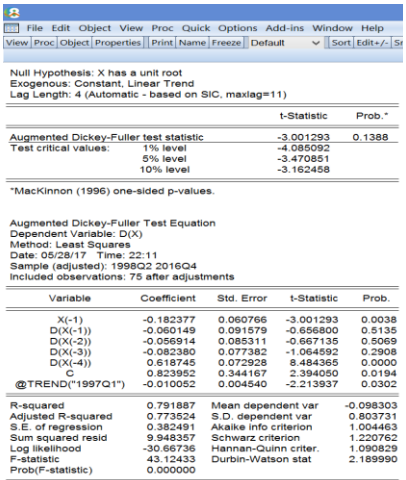
View Proc Object Properties Print Name Freeze Sample Genr Sheet Graph Stats Ident

Correlogram of X

Date: 05/28/17 Time: 22:05
Sample: 1997Q1 2016Q4
Included observations: 80

	AC	PAC	Q-Stat	Prob
1	0.828	0.828	56.901	0.000
2	0.721	0.112	100.57	0.000
3	0.748	0.405	148.27	0.000
4	0.824	0.411	206.93	0.000
5	0.653	-0.574	244.27	0.000
6	0.546	0.008	270.69	0.000
7	0.547	-0.067	297.57	0.000
8	0.588	0.043	329.10	0.000
9	0.445	-0.029	347.40	0.000
10	0.363	0.044	359.76	0.000
11	0.356	-0.061	371.79	0.000
12	0.390	0.067	386.47	0.000
13	0.277	-0.051	393.98	0.000
14	0.210	-0.050	398.36	0.000
15	0.214	0.097	402.97	0.000
16	0.263	0.077	410.04	0.000
17	0.166	-0.124	412.93	0.000
18	0.111	0.034	414.23	0.000
19	0.126	0.003	415.94	0.000
20	0.184	0.028	419.66	0.000
21	0.095	-0.071	420.65	0.000
22	0.042	-0.023	420.85	0.000
23	0.051	-0.071	421.14	0.000
24	0.102	0.008	422.36	0.000
25	0.025	0.029	422.43	0.000
26	-0.016	0.013	422.46	0.000
27	0.000	0.062	422.46	0.000
28	0.049	-0.009	422.77	0.000
29	-0.025	-0.102	422.85	0.000
30	-0.066	-0.042	423.42	0.000
31	-0.064	-0.094	423.97	0.000
32	-0.030	-0.035	424.09	0.000
33	-0.091	0.085	425.25	0.000
34	-0.131	-0.042	427.70	0.000
35	-0.130	0.015	430.16	0.000
36	-0.103	-0.049	431.74	0.000

a)



File Edit Object View Proc Quick Options Add-ins Window Help
View Proc Object Properties Print Name Freeze Default Sort Edit+/- Sm

Null Hypothesis: X has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 4 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.001293	0.1388
Test critical values:	1% level	-4.085092	
	5% level	-3.470851	
	10% level	-3.162458	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(X)
Method: Least Squares
Date: 05/28/17 Time: 22:11
Sample (adjusted): 1998Q2 2016Q4
Included observations: 75 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X(-1)	-0.182377	0.060766	-3.001293	0.0038
D(X(-1))	-0.060149	0.091579	-0.656800	0.5135
D(X(-2))	-0.056914	0.085311	-0.667135	0.5069
D(X(-3))	-0.082380	0.077382	-1.064592	0.2908
D(X(-4))	0.618745	0.072928	8.484365	0.0000
C	0.823952	0.344167	2.394050	0.0194
@TREND("1997Q1")	-0.010052	0.004540	-2.213937	0.0302

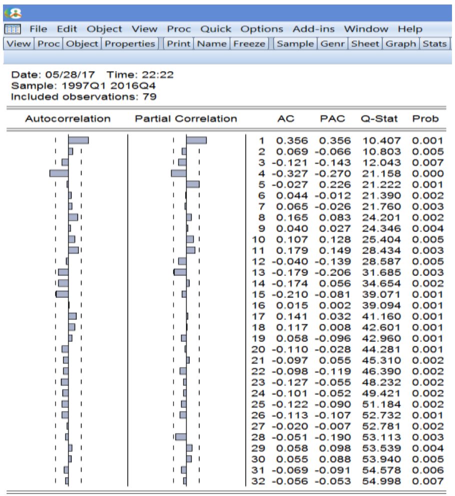
R-squared	0.791887	Mean dependent var	-0.098303
Adjusted R-squared	0.773524	S.D. dependent var	0.803731
S.E. of regression	0.382491	Akaike info criterion	1.004463
Sum squared resid	9.948357	Schwarz criterion	1.220762
Log likelihood	-30.66736	Hannan-Quinn criter.	1.090829
F-statistic	43.12433	Durbin-Watson stat	2.189990
Prob(F-statistic)	0.000000		

b)

Figure 2. The correlogram and Dickey – Fuller test for the independent variable X – registered values

Source: author's work

The stationarity of the independent variable X was then analysed as first-order differences of the level values $D(X)$. From the correlogram graph (which is relatively constant depending on the number of lags considered), we can conclude that the time series of the independent variable $D(X)$ is stationary (figure 3a). In order to check rigorously the stationary/non-stationary character of the series the unit root test was done. It can be observed from figure 3b that the calculated value of the test (-3.48) is greater than the tabulated value for the significance level 1%: - 4.08, but is lower that the tabulated value for the significance level 5%: - 3.47, as well as the tabulated value for the significance level 10%: - 3.16. Consequently, the time series of the independent variable $D(X)$ is stationary (the null hypothesis is rejected) only for significance levels equal or greater than 5%.

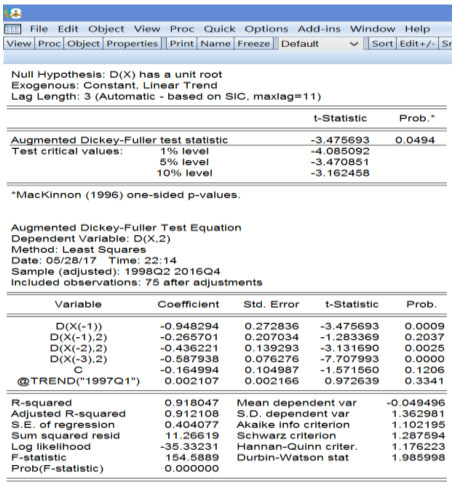


File Edit Object View Proc Quick Options Add-ins Window Help
View Proc Object Properties Print Name Freeze Sample Genr Sheet Graph Stats

Date: 05/28/17 Time: 22:22
Sample: 1997Q1 2016Q4
Included observations: 79

	AC	PAC	Q-Stat	Prob
1	0.356	0.356	10.407	0.001
2	0.069	-0.066	10.803	0.005
3	-0.121	-0.143	12.043	0.007
4	-0.327	-0.270	21.158	0.000
5	-0.027	0.226	21.222	0.001
6	0.044	-0.012	21.390	0.002
7	0.065	-0.026	21.760	0.003
8	0.165	0.083	24.201	0.002
9	0.040	0.027	24.346	0.004
10	0.107	0.128	25.404	0.005
11	0.179	0.149	28.434	0.003
12	-0.040	-0.139	28.587	0.005
13	-0.179	-0.206	31.685	0.003
14	-0.174	0.056	34.654	0.002
15	-0.210	-0.081	39.071	0.001
16	0.015	0.002	39.094	0.001
17	0.141	0.032	41.160	0.001
18	0.117	0.008	42.601	0.001
19	0.058	-0.096	42.960	0.001
20	-0.110	-0.028	44.281	0.001
21	-0.097	0.055	45.310	0.002
22	-0.098	-0.119	46.390	0.002
23	-0.127	-0.055	48.232	0.002
24	-0.101	-0.052	49.421	0.002
25	-0.122	-0.090	51.184	0.002
26	-0.113	-0.107	52.732	0.001
27	-0.020	-0.007	52.781	0.002
28	-0.051	-0.190	53.113	0.003
29	0.058	0.098	53.539	0.004
30	0.055	0.088	53.940	0.005
31	-0.069	-0.091	54.578	0.006
32	-0.056	-0.053	54.998	0.007

a)



File Edit Object View Proc Quick Options Add-ins Window Help
View Proc Object Properties Print Name Freeze Default Sort Edit+/- Sr

Null Hypothesis: D(X) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 3 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.475693	0.0494
Test critical values:	1% level	-4.085092	
	5% level	-3.470851	
	10% level	-3.162458	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(X,2)
Method: Least Squares
Date: 05/28/17 Time: 22:14
Sample (adjusted): 1998Q2 2016Q4
Included observations: 75 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(X(-1))	-0.948294	0.272836	-3.475693	0.0009
D(X(-1),2)	-0.265701	0.207034	-1.283369	0.2037
D(X(-2),2)	-0.436221	0.139293	-3.131690	0.0025
D(X(-3),2)	-0.587938	0.076276	-7.707993	0.0000
C	-0.164994	0.104987	-1.571560	0.1206
@TREND("1997Q1")	0.002107	0.002166	0.972639	0.3341

R-squared	0.918047	Mean dependent var	-0.049496
Adjusted R-squared	0.912108	S.D. dependent var	1.362981
S.E. of regression	0.404077	Akaike info criterion	1.102195
Sum squared resid	11.26619	Schwarz criterion	1.287594
Log likelihood	-35.33231	Hannan-Quinn criter.	1.176223
F-statistic	154.5889	Durbin-Watson stat	1.985998
Prob(F-statistic)	0.000000		

b)

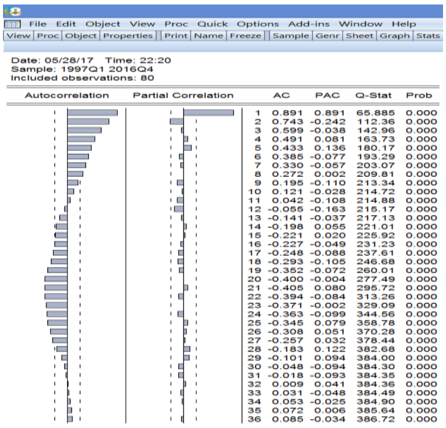
Figure 3. The correlogram and Dickey – Fuller test for the independent variable $D(X)$ – first-order differences

Source: author's work

The analysis of the dependent variable Y stationarity, for observed values (level values), as shown in figure 4a and the correlogram graph (which is, at the same time, decreasing but also oscillating depending on the number of lags considered) shows that the series of the dependent variable Y at the level of observed values (level values) is non-stationary. To verify rigorously the stationary/non-stationary character of the series, the unit root test was done.

As seen in figure 4b, the calculated value of the test (-2.57) is greater than the tabulated values for all three levels of significance (1%: - 4.08; 5%: - 3.47; 10%: - 3.16). Consequently, the time series of the dependent variable Y for the observed values (level values) is non-stationary (the null hypothesis is accepted).

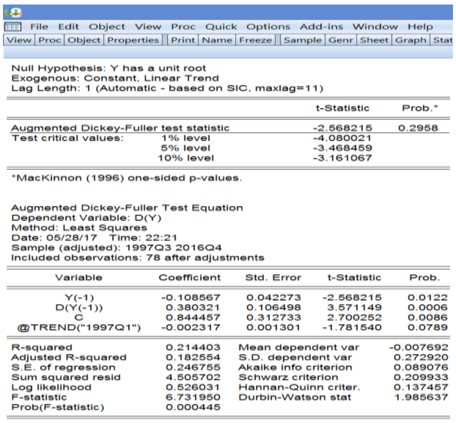
In this case, the time series stationarity will be checked in the case of quarterly values calculated as first-order differences $D(Y)$ of the level values.



Date: 05/28/17 Time: 22:20
Sample: 1997Q1 2016Q4
Included observations: 80

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.891	0.891	65.885	0.000
		2	0.743	-0.242	112.36	0.000
		3	0.599	-0.038	142.96	0.000
		4	0.491	0.081	163.73	0.000
		5	0.433	0.136	180.17	0.000
		6	0.385	-0.077	193.29	0.000
		7	0.330	-0.057	203.07	0.000
		8	0.272	0.002	209.81	0.000
		9	0.195	-0.110	213.34	0.000
		10	0.121	-0.028	214.72	0.000
		11	0.042	-0.108	214.88	0.000
		12	-0.055	-0.163	215.17	0.000
		13	-0.141	-0.037	217.13	0.000
		14	-0.198	0.055	221.01	0.000
		15	-0.221	0.020	225.92	0.000
		16	-0.227	-0.049	231.23	0.000
		17	-0.248	-0.088	237.61	0.000
		18	-0.293	-0.105	246.68	0.000
		19	-0.352	-0.072	260.01	0.000
		20	-0.400	-0.004	277.49	0.000
		21	-0.405	0.080	295.72	0.000
		22	-0.394	-0.084	313.26	0.000
		23	-0.371	-0.002	329.09	0.000
		24	-0.363	-0.099	344.56	0.000
		25	-0.345	0.079	358.78	0.000
		26	-0.308	0.051	370.28	0.000
		27	-0.257	0.032	378.44	0.000
		28	-0.183	0.122	382.68	0.000
		29	-0.101	0.094	384.00	0.000
		30	-0.048	-0.094	384.30	0.000
		31	-0.018	-0.093	384.35	0.000
		32	0.009	0.041	384.36	0.000
		33	0.031	-0.048	384.49	0.000
		34	0.053	-0.025	384.90	0.000
		35	0.072	0.006	385.64	0.000
		36	0.085	-0.034	386.72	0.000

a)



Null Hypothesis: Y has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.568215	0.2958
Test critical values:	1% level	-4.080021	
	5% level	-3.468459	
	10% level	-3.161067	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(Y)
Method: Least Squares
Date: 05/28/17 Time: 22:21
Sample (adjusted): 1997Q3 2016Q4
Included observations: 78 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y(-1)	-0.108567	0.042273	-2.568215	0.0122
D(Y(-1))	0.380321	0.106498	3.571149	0.0006
C	0.844457	0.312733	2.700252	0.0086
@TREND("1997Q1")	-0.002317	0.001301	-1.781540	0.0789

R-squared	0.214403	Mean dependent var	-0.007692
Adjusted R-squared	0.182554	S.D. dependent var	0.272920
S.E. of regression	0.246755	Akaike info criterion	0.089076
Sum squared resid	4.505702	Schwarz criterion	0.209933
Log likelihood	0.526031	Hannan-Quinn criter.	0.137457
F-statistic	6.731950	Durbin-Watson stat	1.985637
Prob(F-statistic)	0.000445		

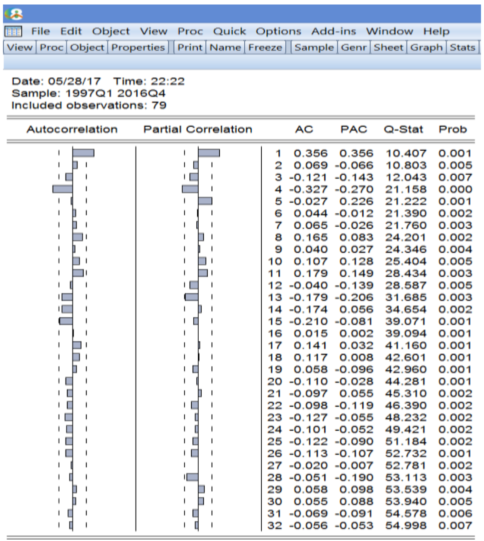
b)

Figure 4. The correlogram and Dickey – Fuller test for the dependent variable *Y* – registered values

Source: author's work

The dependent variable $D(Y)$ stationarity is highlighted by the correlogram graph (which is relatively constant depending on the number of lags considered) as seen in figure 5a, so the series of the dependent variable $D(Y)$ is stationary. In order to check rigorously the stationary/non-stationary character of the series the unit root test was done.

The Dickey – Fuller test data shows that the calculated value of the test (-6.12) is lower than the tabulated values for all three levels of significance 1%: - 4.08; 5% - 3.47; 10% - 3.16. Consequently, the time series of the variable $D(Y)$ is stationary (the null hypothesis is rejected) for any of the levels of tabulated significance.

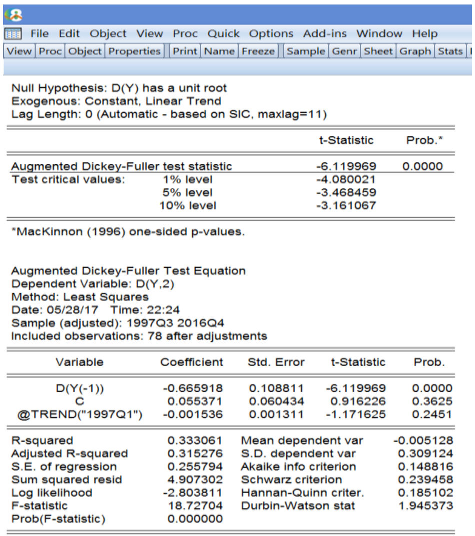


File Edit Object View Proc Quick Options Add-ins Window Help
View Proc Object Properties Print Name Freeze Sample Genr Sheet Graph Stats

Date: 05/28/17 Time: 22:22
Sample: 1997Q1 2016Q4
Included observations: 79

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.356	0.356	10.407	0.001
		2	0.069	-0.066	10.803	0.005
		3	-0.121	-0.143	12.043	0.007
		4	-0.327	-0.270	21.158	0.000
		5	-0.027	0.226	21.222	0.001
		6	0.044	-0.012	21.390	0.002
		7	0.065	-0.026	21.760	0.003
		8	0.165	0.083	24.201	0.002
		9	0.040	0.027	24.346	0.004
		10	0.107	0.128	25.404	0.005
		11	0.179	0.149	28.434	0.003
		12	-0.040	-0.139	28.587	0.005
		13	-0.179	-0.206	31.685	0.003
		14	-0.174	0.056	34.654	0.002
		15	-0.210	-0.081	39.071	0.001
		16	0.015	0.002	39.094	0.001
		17	0.141	0.032	41.160	0.001
		18	0.117	0.008	42.601	0.001
		19	0.058	-0.096	42.960	0.001
		20	-0.110	-0.028	44.281	0.001
		21	-0.097	0.055	45.310	0.002
		22	-0.098	-0.119	46.390	0.002
		23	-0.127	-0.055	48.232	0.002
		24	-0.101	-0.052	49.421	0.002
		25	-0.122	-0.090	51.184	0.002
		26	-0.113	-0.107	52.732	0.001
		27	-0.020	-0.007	52.781	0.002
		28	-0.051	-0.190	53.113	0.003
		29	0.058	0.098	53.539	0.004
		30	0.055	0.088	53.940	0.005
		31	-0.069	-0.091	54.578	0.006
		32	-0.056	-0.053	54.998	0.007

a)



File Edit Object View Proc Quick Options Add-ins Window Help
View Proc Object Properties Print Name Freeze Sample Genr Sheet Graph Stats

Null Hypothesis: D(Y) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-6.119969	0.0000
Test critical values:	1% level	-4.080021	
	5% level	-3.468459	
	10% level	-3.161067	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(Y,2)
Method: Least Squares
Date: 05/28/17 Time: 22:24
Sample (adjusted): 1997Q3 2016Q4
Included observations: 78 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(Y(-1))	-0.665918	0.108811	-6.119969	0.0000
C	0.055371	0.060434	0.916226	0.3625
@TREND("1997Q1")	-0.001536	0.001311	-1.171625	0.2451

R-squared	0.333061	Mean dependent var	-0.005128
Adjusted R-squared	0.315276	S.D. dependent var	0.309124
S.E. of regression	0.255794	Akaike info criterion	0.148816
Sum squared resid	4.907302	Schwarz criterion	0.239458
Log likelihood	-2.803811	Hannan-Quinn criter.	0.185102
F-statistic	18.72704	Durbin-Watson stat	1.945373
Prob(F-statistic)	0.000000		

b)

Figure 5. The correlogram and Dickey – Fuller test for the dependent variable $D(Y)$ – first-order differences

Source: author's work

The paper presents the statistical analysis of the two-stationary series, $D(X)$ and $D(Y)$. As seen in figure 6, the histogram of the independent variable $D(X)$ shows that the series does not have a normal distribution, the average of the share of the rate of social protection expenditure of the unemployed over GDP is -0.08% during the whole period. As for the statistical distribution of the series, of the Skewness indicator value, i.e. the asymmetry coefficient (1.71) it is found that it is not a normal distribution. As for the value of the indicator showing the flattening of the distribution (Kurtosis: 7.46) the value greater than 3 indicates a leptokurtotic distribution. The value of the Jarque–Bera test is 103.91 (which indicates, by comparing the difference between the asymmetry coefficient and the flattening coefficient against the normal distribution, the degree of normality of the distribution) has a null probability associated, so the hypothesis of the normal distribution of the time series $D(X)$ is rejected.

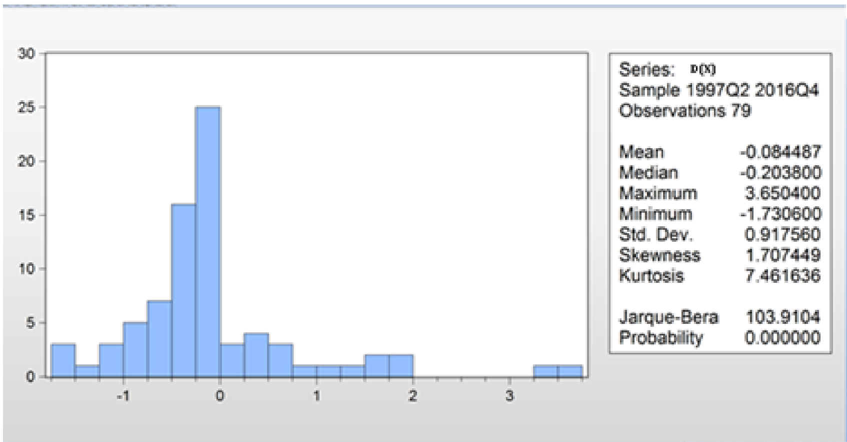
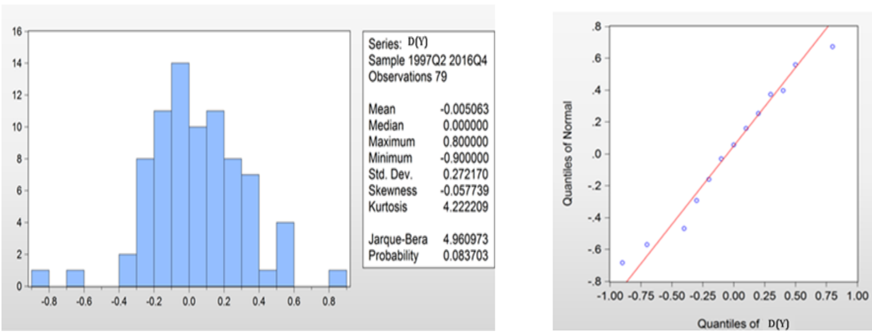


Figure 6. The histogram and the quantiles graph for the variable $D(X)$

Source: author's work

The statistical analysis of the dependent variable $D(Y)$, as seen in figure 7a shows that the average of the total unemployment rate series over the active population is -0.005% during the whole period. As for the statistical distribution of the series, of the Skewness indicator value, i.e. the asymmetry coefficient (0.27) it is observed that it is a relatively close distribution of the normal distribution. As for the value of the indicator which shows the flattening of the distribution (Kurtosis: 4.22) the value greater than 3 indicates a leptokurtotic distribution. The value of the Jarque–Bera test: 4.96, has a very low probability associated, so the hypothesis of the normal distribution of time series $D(Y)$ is rejected.

The quantiles graph of the dependent variable $D(Y)$ (figure 7b) reveals an overlap of the distribution of the residual quantiles with the normal distribution.



a) b)

Figure 7. The histogram and the quantiles graph for the variable $D(Y)$

Source: author's work

It is observed that the quarterly averages of the independent variable $D(X)$ differ significantly among them, which indicates a strong seasonality of social protection expenditure for the unemployed as seen in figure 8. The same phenomenon has also been encountered in the analysis of seasonality on observed values (level values) which means that the seasonality attribute of this time series has more deeper causes than those involved in the statistical calculation.

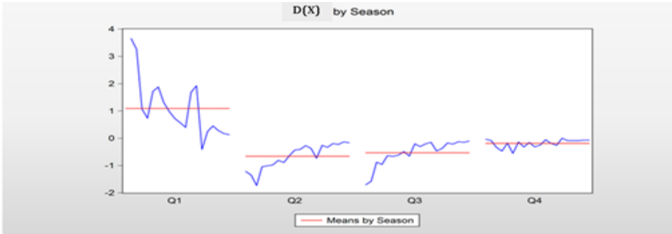


Figure 8. Seasonality of the time series D(X), together with the quarterly average indication

Source: author's work

The quarterly averages of the dependent variable $D(Y)$, as seen in figure 9, differ very little among them, keeping close the same value, which indicates an absence of seasonality over the unemployment rate. The same phenomenon has also been encountered in the analysis of seasonality on observed values (level values) which means that the seasonality attribute is not proper to this time series.

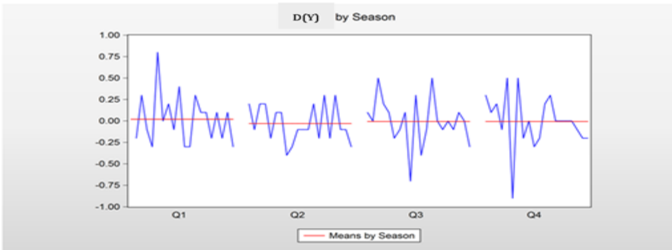


Figure 9. Seasonality of the time series $D(Y)$, together with the quarterly average indication

Source: author's work

The long-term trend (the Hodrick – Prescott trend) for the two variables analysed in the current paper, respectively the independent variable $D(X)$, the rate of social protection expenditure of the unemployed over GDP, and the dependent variable $D(Y)$, the total unemployment rate over the active population as seen in figure 10, and figure 11 show the evolution of the two-time series in relation to the time axis.

By eliminating the cyclicality (seasonality) variation of the independent variable $D(X)$, using the Hodrick – Prescott filter, there is a flat de-seasoned trend (zero slope in relation to the time axis), during the analysed period, having linear type pattern.

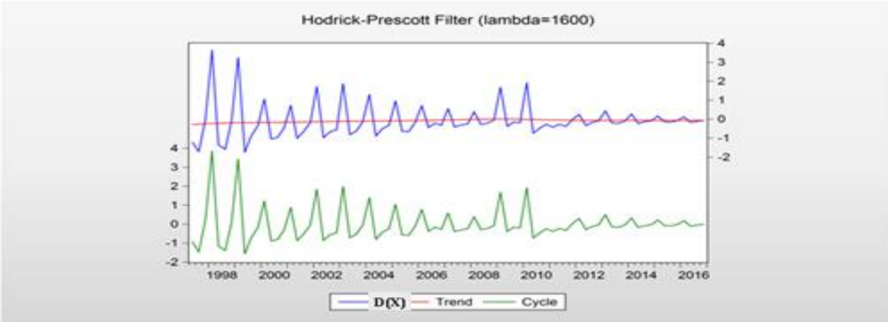


Figure 10. The long-term trend of the expenditure rate change with the unemployed social protection

Source: author's work

By eliminating the cyclicality (seasonality) variation of the dependent variable $D(Y)$, using the Hodrick – Prescott filter, there is a much more flattened deseasonalized trend than in the case of recorded values (level values), during the analysed period, having a quasi-sinusoidal pattern.

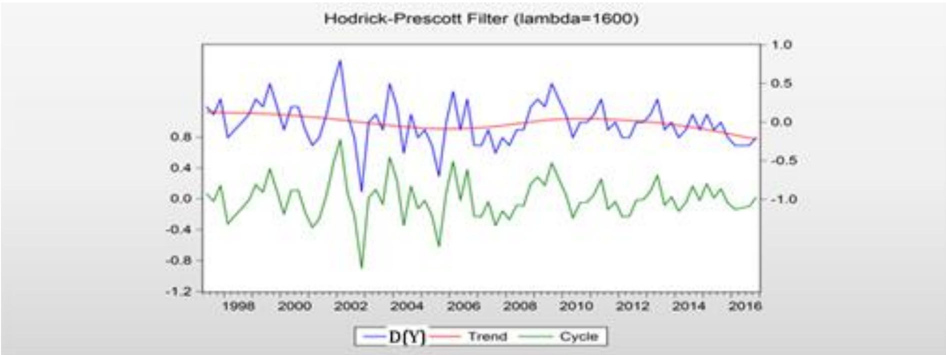


Figure 11. The long-term trend of the unemployment rate variation

Source: author's work

4. A Mono-Factorial Linear Regression Model

The linearity property of the correlation between the two variables involved in the study is suggested by the „points cloud" showed in the figure 12:

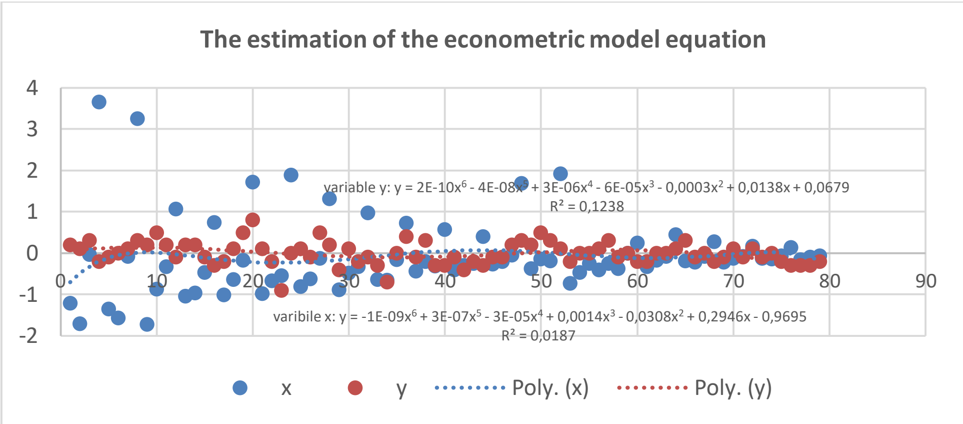


Figure 12. The „points cloud" of the variables involved

Source: author's work

By approximation:

$$x(t) = 0.2946 \cdot t - 0.9695$$

$$y(t) = 0.0138 \cdot t + 0.0679$$

By eliminating of the variable t between the two equations, gives the following equation relating x and y :

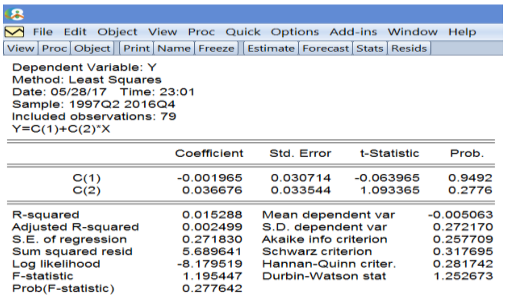
$$y = 0.046843 \cdot x + 0.113314 + u_t$$

So, a linear relation between the dependent variable y and the independent variable x is considered:

$$y_t = c(1) + c(2) \cdot x_t + u_t$$

where $y \equiv D(Y)$, respectively $x \equiv D(X)$.

Parameters estimation:



File Edit Object View Proc Quick Options Add-ins Window Help

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: Y
Method: Least Squares
Date: 05/28/17 Time: 23:01
Sample: 1997Q2 2016Q4
Included observations: 79
Y=C(1)+C(2)*X

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.001965	0.030714	-0.063965	0.9492
C(2)	0.036676	0.033544	1.093365	0.2776

R-squared	0.015288	Mean dependent var	-0.005063
Adjusted R-squared	0.002499	S.D. dependent var	0.272170
S.E. of regression	0.271830	Akaike info criterion	0.257709
Sum squared resid	5.689641	Schwarz criterion	0.317695
Log likelihood	-8.179519	Hannan-Quinn criter.	0.281742
F-statistic	1.195447	Durbin-Watson stat	1.252673
Prob(F-statistic)	0.277642		

Figure 13. The estimation of the econometric model parameters

Source: author's work

Final equation: $y_t = -0.001965 + 0.036676 \cdot x_t$

The hypothesis verification of the linear regression model show that the average of errors is null $(E(u_t) = 0)$: - 0.000000000000000913, as seen in figure 14 of the histogram graph.

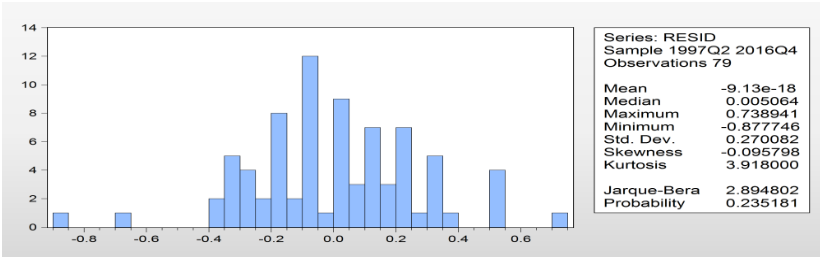
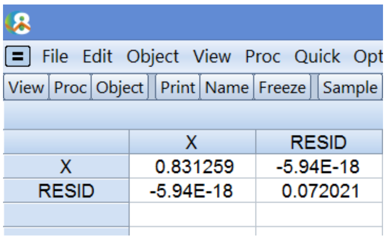


Figure 14. The histogram of the linear regression model errors

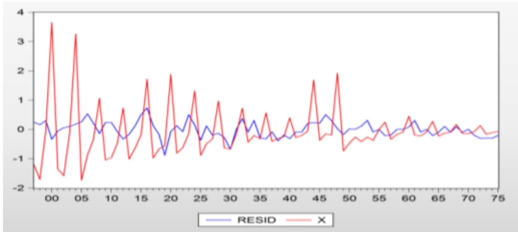
Source: author's work

The errors are linearly dependent on each other and are positively correlated $cov(u_i, u_j) = 0)$, error spreading is constant and finite $var(u_t) = ct$ (homoscedasticity property), there are no correlations between the residual and the independent variable $cov(u_t, x_t) = 0)$, and the series of errors is normally distributed $u_t \sim N(0; \sigma^2)$, $u_t \sim N(0; 0.073)$.



	X	RESID
X	0.831259	-5.94E-18
RESID	-5.94E-18	0.072021

a)

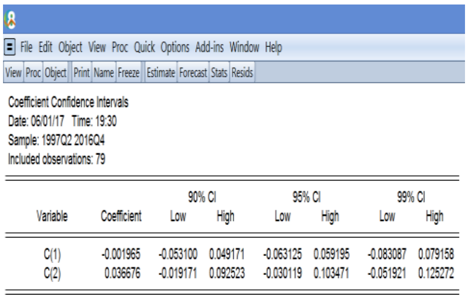


b)

Figure 15. Correlations between the residual and the independent variable

Source: author's work

The confidence intervals for estimated coefficients were determined, as seen in figure 16, and the ellipse of the confidence intervals of the estimated values for the regression equation coefficients is presented in figure 17.



Coefficient Confidence Intervals
Date: 06/01/17 Time: 19:30
Sample: 1997Q2 2016Q4
Included observations: 79

		90% CI		95% CI		99% CI	
Variable	Coefficient	Low	High	Low	High	Low	High
C(1)	-0.001965	-0.053100	0.049171	-0.063125	0.059195	-0.083087	0.079158
C(2)	0.036676	-0.019171	0.092523	-0.030119	0.103471	-0.051921	0.125272

Figure 16. The confidence intervals for the estimated coefficients

Source: author's work

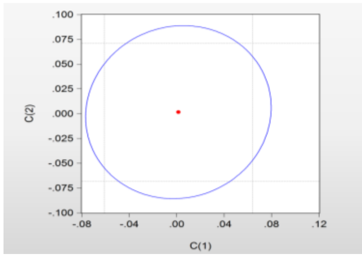


Figure 17. The ellipse of confidence intervals of estimated values for the regression equation coefficients

Source: author's work

5. Conclusions

Regarding the *first objective* of the study (the formation of time series and the statistical analysis for the two variables), the time series of the two variables are stationary only for the first-order differences (in the case of the independent variable, for significance levels ≥ 5%), for the statistical distribution of the independent variable $D(X)$ the econometric model has highlighted the fact that there is no normal distribution, while the statistical distribution of the dependent variable $D(Y)$ is near normal. The analysis of seasonality showed that the independent variable $D(X)$ is strongly seasonalized (as a result of quarterly GDP seasonality), while the dependent variable $D(Y)$ does not show seasonality at the level of quarters, thus the trend of the independent variable $D(X)$ is almost horizontal, while the dependent variable $D(Y)$ exhibits a weak sinusoidal trend (much more flattened than the constant trend to level values). Regarding *the second objective* of the study (designing a mono-factorial linear econometric model), we have found (by „points cloud" of the two data series) that two series of data are linearly correlated by means of a mono-factorial linear regression model. So, we could establish an econometric mono-factorial model. Regarding *the third objective* of the study (estimation of the econometric model), we have obtainted a direct proportional relationship between the dependent variable $D(Y)$ and the independent one $D(X)$ as follows: $y_t = -0.001965 + 0.036676 \cdot x_t$, where $y_t \rightarrow D(Y)_t$, respectively $x_t \rightarrow D(X)_t$.

The main scientific results of the study are as follows: 1) statistically the time series of the two variables are stationary only at the level of the first order differences, the series of the independent variable is non-normal, and the dependent variable is quasi-normal; 2) the estimation of the quantitative model shows that there is a direct correlation (though weak, about 3%) between the dependent and the independent variable. At the same time, there are no influences outside the independent variable (the free term of the first degree equation is approximately - 0.002). In conclusion, it can be said that the variation of the total unemployment rate in the active population is direct, albeit relatively weak, influenced by the change in the rate of social protection expenditures in GDP.

6. References

Dinga, E. (2011). *Economics II. Macroeconomics, Cours notes,* pp. 111 – 113.

Jula, D. & Jula, N. (2017). *Economic modeling. Econometric and optimization models.* Bucharest: The Publishing house Mustang.

Brooks, C. (2014). *Introductory Econometrics for Finance*, 3 rd Edition, Cambridge University Press;

Mankiw, G. (2015). *Principles of Economics*. Cengage Learning.

***EUROSTAT. Web page. Retrieved from http://ec.europa.eu/eurostat/data/database. Date: 02.06.2017.

***STATISTICAL YEARBOOK OF ROMANIA. Web page. Retrieved from www.insse.ro/cms/ro/tags/anuarul-statistic-al-romaniei, date 02.06.2017.