Bul. Ştiinţ. Univ. Baia Mare, Ser. B,
Matematică-Informatică, Vol. XVIII(2002), Nr. 2, 141 - 146

THE MODELLING OF THE UNEMPLOYMENT IN ROMANIA

Alina BĂRBULESCU

Abstract. The aim of this paper is to study the existence of a Okun's law, for the Romanian economy, after 1990.

MSC 2000:11M27
Keywords: autocorrelation, ACF, PACF

In 1962, Okun deduced the following low, for the USA economy:

$$RGNP_{t/t-1} = 3\% - (RU_t - RU_{t-1}), \tag{1}$$

where $RGNP_{t/t-1}$ is the GNP (gross national product) growth rate in the period t beside $t-1$, RU_t is the unemployment instalment in the year t and 3% is the GNP trend augmentation. Below, we shall analyze the existence of an Okun - type law, for the Romanian economy, between 1990 - 2000, using the data given in the *table 1.*

Table 1. The unemployment instalment and the GNP rate between 1990 and 2000

year	'90	'91	'92	'93	'94	'95	'96	'97	'98	'99	2000
t	0	1	2	3	4	5	6	7	8	9	10
RU_t	0	3.4	8.2	10.4	10.9	9.5	6.6	8.9	10.4	11.8	10.5
$RGNP_t$	-5.6	-12.9	-8.8	1.5	3.9	6.9	4.5	-6.1	-4.8	-2.3	1.6
RU_t-RU_{t-1}	-	3.4	4.8	2.2	0.5	-1.4	-2.9	2.3	1.5	1.4	-1.3

If there exists an Okun - type law for the Romanian economy, its form is:

$$RGNP_{t/t-1} = -0.538 - 0.384(RU_t - RU_{t-1});$$

but the standard deviation is big (2.862), the residual have big values and the corre-lation coefficient of the residuals is 0.658. Thus, the model is not valid.

In [3] had been shown that the trend of GNP augmentation between 1980 - 1989 was around 1.4% and it was suggested the existence of a lag - time relation between RGNP and RU, in the period 1990 -2000. Thus, we search a relation:

$$RGNP_{t/t-1} = a + b(RU_{t+1} - RU_t) \tag{2}$$

The values determined for the coefficients a and b are: $a = 0.392, b = -0.277$.

The corresponding residuals are big and the trend value is not 1.4%. The mean of the squared errors is 1.6477 and the correlation coefficient is 0.7533 (big enough). Thus, the model is not valid.

Another idea is to try to determine the dependence between the GNP and the unemployment rate as an equation of the following types:

$$RU_{t+1} = aRU_t + bRGNP_{t/t-1} + c, t = 1, ..., 10 \quad (3)$$

$$RU_{t+1} = aRU_t + b/RGNP_{t/t-1} + c, t = 1, ..., 10 \quad (4)$$

We shall discuss only the first model. If we denote by: $y_t = RU_{t+1}, x_t = RGNP_{t/t-1}, t = 1, ..., 10$, ε_t (ε) - the residual variable and

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{10} \end{pmatrix}, X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{10} \end{pmatrix}, Z = \begin{pmatrix} y_0 & x_1 & 1 \\ y_1 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ y_{10} & x_{10} & 1 \end{pmatrix}, A = \begin{pmatrix} a \\ b \\ c \end{pmatrix},$$

then, the model can be written:

$$y_t = ay_{t-1} + bx_t + c + \varepsilon_t, t = 1, ..., 9 \Leftrightarrow Y = Z \cdot A + \varepsilon, \quad (5)$$

The moment $t = 0$ is the year 1990; it was supposed that $y_0 = 0$.

I. The estimation of the coefficients a and b.

From (5) it results that:

$$\widehat{Y} = Z\widehat{A} \Rightarrow \widehat{A} = (Z'Z)^{-1}(Z'Y) \Leftrightarrow \widehat{A} = \begin{pmatrix} 0.66777 \\ -0.1915 \\ 3.2476 \end{pmatrix},$$

where Z' is the transposed of $Z, \widehat{Y}$ - an estimation of Y and $\widehat{A}$ - an estimation of A. The variances of the estimators can be determined using the relations ([5]):

$$\widehat{\sigma}_{\varepsilon}^2 = \frac{1}{n-p}\sum_{i=1}^{n} \widehat{\varepsilon}_i^{\,2} = 1.59024,$$

$$\widehat{\Omega}_{\widehat{a}} = \widehat{\sigma}_{\varepsilon}^2 (Z'Z)^{-1} = \begin{pmatrix} 0.01598 & -0.0042 & -0.13816 \\ -0.0042 & 0.00552 & 0.047068 \\ -0.13816 & 0.047068 & 1.379589 \end{pmatrix},$$

$$\widehat{\sigma}_{\widehat{a}}^2 = 0.01598, \ \widehat{\sigma}_{\widehat{b}}^2 = 0.00552.$$

where: n is the data number , p - the parameters number, $\widehat{\varepsilon}_i$ is the difference between the data and them estimations.

II. The validity test for the estimators of the coefficients a and b.

First, we shall test the hypothesis: $H_0 : a = 0$ against $H_1 : a \neq 0$, at the significance level $\alpha = 0.05$.

If $\left|\frac{\widehat{a}}{\widehat{\sigma_{\widehat{a}}}}\right| \geq t_{n-p,1-\alpha/2}$, then H_0 is rejected, where $t_{1-\alpha/2}$ is the point in the t tables for degrees of freedom $n-p$. Since

$$\left|\frac{\widehat{a}}{\widehat{\sigma_{\widehat{a}}}}\right| = \frac{0.66777}{0.12641} = 5.2825 \geq t_{7,0.975} = 2.365,$$

we accept the hypothesis H_1.

Analogous, to test the hypothesis: $H_0 : b = 0$ against $H_1 : b \neq 0$. H_0 will be rejected because

$$\left|\frac{\widehat{b}}{\widehat{\sigma_{\widehat{b}}}}\right| = 2.581 \geq t_{7,0.975} = 2.365.$$

To test the significance of all the parameters, we use the F - test. If

$$F_c = \frac{n-p}{p} \cdot \frac{\left(\widehat{Y}-\overline{\widehat{Y}}\right)'\left(\widehat{Y}-\overline{\widehat{Y}}\right)}{\widehat{\varepsilon}'\widehat{\varepsilon}}.$$

is greater than $F_{p,n-p,\alpha}$, the hypothesis that the parameters are zero is rejected, where $F_{p,n-p,\alpha}$ is the point in the F tables for degrees of freedom p and $n-p$.

In our case:

$$F_c = 9.3219 > F_{3,7,0.05} = 5.89.$$

Thus, we accept the hypothesis of nonvanishing parameters.

III. The intensity of the relation among the model's variables is studied using the de-termination coefficient, R^2 and the "modified" determination coefficient, $\overline{R^2}$:

$$R^2 = \frac{\widehat{A}'(Z-\overline{Z})(Y-\overline{Y})}{(Y-\overline{Y})'(Y-\overline{Y})} = 0.7998, \quad \overline{R^2} = 1 - \frac{n-1}{n-p}(1-R^2) = 0.77478,$$

where $\overline{Z}$ is the mean value of the elements of Z, n is the sample volume and p,the number of the explanatory variables.

The obtained value is big enough. In order to be sure that the influence of the endogenous variable on the exogenous one is important, the significance test of $\overline{R^2}$ is necessary.

Consider: $H_0 : \overline{R^2} = 0$ and $H_1 : \overline{R^2} \neq 0$, $F_c = \frac{n-p-1}{p} \cdot \frac{\overline{R^2}}{1-\overline{R^2}}$.

If $F_c > F_{p,n-p-1,\alpha}$, the hypothesis H_1 will be accepted.

In our case, $p = 3$, $n = 10$, $F_c = 6.88027 > 4.757 = F_{3,6,0.05}$ and H_0 is rejected.

Thus, the influence of the endogenous variables on the exogenous one is significantly.

IV. The tests on the residual variable.

1. The normality test. We shall use the Lilliefors test.

Consider the selection $(x_1, x_2, \ldots, x_n)$, $\overline{x}$ - the selection mean, s^2 - the selection variance, $z_i = \dfrac{x_i - \overline{x}}{s}$, $z_1 \leq z_2 \leq \ldots \leq z_n$ the values of z_i increasing ordered,

$$F_0(z_i) = \Phi(z_i),\ i = \overline{1,n},$$

$$F_n(z_i) = \begin{cases} 0\ ,\ i \leq 0 \\ \frac{i}{n}\ ,\ 1 \leq i \leq n-1 \\ 1\ ,\ i = n \end{cases}$$

$$D_n = \max_{i=\overline{1,n}} |F_0(z_i) - F_n(z_i)|$$

where Φ is the normal distribution function.

The hypothesis about the normality of the residual variable is accepted, at the significance level α, if $D_n \leq D_{n,1-\alpha}$. The values $D_{n,1-\alpha}$ are given in the Lilliefors table.

Let consider: $H_0 : \varepsilon \sim N(0,\ 1.236854)$ and $\alpha = 0.05$. Since

$$D_{10} = 0.1247 < 0.258 = D_{10,\ 0.95},$$

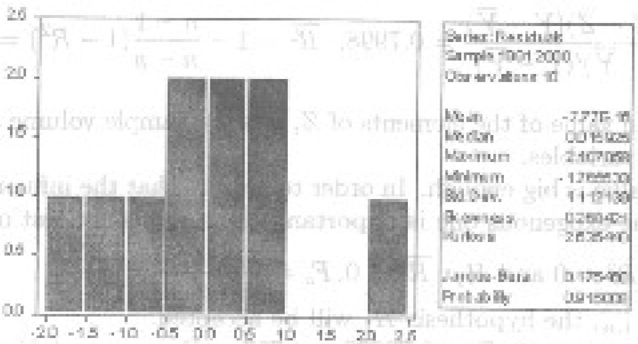
we accept the hypothesis H_0.

The Jarque - Bera test can also be used to prove the normality of the residue.

$JB = 0.916008 < 4.61 = \chi^2(2)$, with the probability 0.916008.

It results that the residual variable is normal.

Table2. The Jarque – Bera test



2. The homoscedasticity test. To use the Cochran test, we form two samples of residuals, with the volumes: $n_1 = n_2 = 5$. For these: $s_1^2 = 0.4898$, $s_2^2 = 2.2931$.

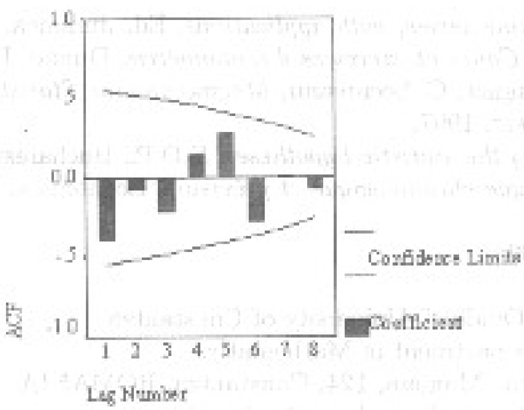
$$Q \overset{def}{=} \frac{\max(s_1^2, s_2^2)}{s_1^2 + s_2^2} = 0.824 < 0.9057 = Q_{0.05,\,2,\,4}$$

Then, the homoscedasticity hypothesis is accepted.

Table.3.ACF and PACF

<table>
<tr><th>Autocorrelation</th><th>Partial autocorrelation</th><th>Lag</th><th>ACF</th><th>PACF</th><th>Q-Stat</th></tr>
<tr><td>.***| . |</td><td>.***| . |</td><td>1</td><td>-0.389</td><td>-0.389</td><td>2.0180</td></tr>
<tr><td>. *| . |</td><td>. **| . |</td><td>2</td><td>-0.071</td><td>-0.262</td><td>2.0935</td></tr>
<tr><td>. **| . |</td><td>.***| . |</td><td>3</td><td>-0.210</td><td>-0.439</td><td>2.8503</td></tr>
<tr><td>. |* . |</td><td>. **| . |</td><td>4</td><td>0.144</td><td>-0.295</td><td>3.2632</td></tr>
<tr><td>|** . |</td><td>. |* . |</td><td>5</td><td>0.274</td><td>0.157</td><td>5.0671</td></tr>
<tr><td>**| . |</td><td>. *| . |</td><td>6</td><td>-0.269</td><td>-0.117</td><td>7.2447</td></tr>
<tr><td>. | . |</td><td>. | . |</td><td>7</td><td>0.011</td><td>-0.051</td><td>7.2491</td></tr>
<tr><td>. *| . |</td><td>. *| . |</td><td>8</td><td>-0.077</td><td>-0.061</td><td>7.6061</td></tr>
</table>

Fig.1.The autocorrelation function (ACF)

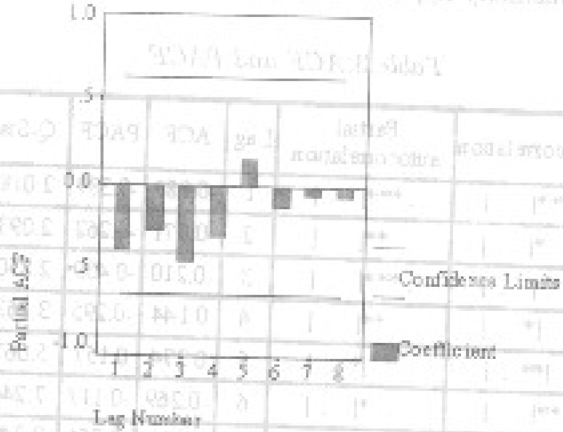


3.The independence test. To prove the independence of the residuals, we use the autocorrelation and the partial autocorrelation functions (*the figures 1* and *2*).

In the *table 3* are given the values of ACF and PACF. These are inside the confidence interval and, by consequence, they can be considered zero. In the 6^{th} row of this table appear the values of the Box - Ljung statistic (Q - Stat). All are less than $\chi^2(8)$.

Thus,we accept the hypothesis that the residuals form a white noise.

Fig.2. The partial autocorrelation function (PACF)



The model (3) is valid. It can be improved if the value of the unemployment instalment in 1996 is considered to be aberrant.

REFERENCES

[1] A.Barbulescu, *Time series, with applications*, Ed. Junimea, Iasi, 2002

[2] J. Bourbonnais, *Cours et exercices d'économétrie*, Dunod, Paris, 1993.

[3] I. Capanu, P. Wagner, C. Secareanu, *Macroeconomic Statistics*, Ed. Economica, Bucharest, 1997.

[4] V. Craiu, *Testing the statistic hypotheses*, E.D.P., Bucharest, 1972

[5] M. Vaté, *Statistique chronologique et prévision*, Economica, Paris, 1993

Received: 19.10.2002

"Ovidius" University of Constantza,
Department of Mathematics
Bd. Mamaia, 124, Constantza, ROMANIA
E-mail:abarbulescu@univ-ovidius.ro