The International College of Economics and Finance
Econometrics-2021-2022.
Home assignment 8. Autocorrelated Disturbance Term.
To be submitted by February, 13, 23:55

1. [30 marks] A researcher has a data on total consumer expenditures on air flights, *AIR*, and gross disposable personal income, *DPI*, both measured in billions of US dollars (1984-2008), and price index for air flights, *P*, for the same period. She gets the following regression equations (standard errors in parentheses):

$$\text{(OLS)} \quad LGAIR_t = \underset{(0.68)}{-12.75} + \underset{(0.10)}{2.06}\, LGDPI_t \qquad R^2 = 0.47, \quad DW = 0.31 \tag{1}$$

$$\text{(C.O.)} \quad LGAIR_t = \underset{(5.89)}{-7.45} + \underset{(0.84)}{1.28}\, LGDPI_t \qquad R^2 = 0.98, \quad DW = 1.40 \tag{2}$$

$$\text{(OLS)} \quad LGAIR_t = \underset{(0.40)}{-9.58} + \underset{(0.05)}{2.28}\, LGDPI_t - \underset{(0.09)}{0.99}\, LGP_t \qquad R^2 = 0.99, \quad DW = 1.46 \tag{3}$$

$$\text{(C.O.)} \quad LGAIR_t = \underset{(0.54)}{-9.45} + \underset{(0.07)}{2.24}\, LGDPI_t - \underset{(0.11)}{0.97}\, LGP_t \qquad R^2 = 0.99, \quad DW = 1.88 \tag{4}$$

where OLS means usual least squares method, C-O – iterative procedure by Cochrane-Orcutt (iterative AR(1) transformation), and variables *LGAIR*, *LGDPI* and *LGP* are logarithms of the variables *AIR*, *DPI* and *P*, correspondingly.

1.1. [10 marks] □ Why one can suggest the presence of autocorrelation in some of the equations listed above? Show your work.
□ What are the possible explanations of this phenomena? What are possible consequences of the suggested autocorrelation?

1.2. [10 marks] □ Why the autoregressive transformation allows in some cases to remove autocorrelation in the residuals of the regression model? In which cases the autocorrelation transformation has no effect?
□ Compare four equations paying attention to the values of their coefficients, standard errors and other characteristics of their statistical quality.

1.3. [10 marks] A colleague of the researcher gave him an advice to use lagged variable as the best and simple tool to make DW statistics acceptable. The corresponding equation is

$$LGAIR_t = \underset{(2.04)}{0.64} - \underset{(0.33)}{0.08}\, LGDPI_t + \underset{(0.15)}{0.97}\, LGAIR_{t-1} \qquad R^2 = 0.98, \quad DW = 1.33 \tag{5}$$

Comment the advice of the researcher's colleague. Was his advice helpful? What is your final conclusion on the analysis of all regressions?

2. [20 marks] A researcher has annual data on *LIFE*, aggregate consumer expenditure on life insurance, *DPI*, aggregate disposable personal income, and *PRELLIFE*, a price index for the cost of life insurance relative to general inflation, for the United States for the period 1959–1994. *LIFE* and *DPI* are measured in US$ billion. *PRELLIFE* is an index number series with 1992=100. She defines *LGLIFE*, *LGDPI*, and *LGPRLIFE* as the natural logarithms of *LIFE*, *DPI*, and *PRELLIFE*, respectively. She fits the regressions shown in columns (1) – (4) of the table, each with *LGLIFE* as the dependent variable. (Standard errors in parentheses; OLS = ordinary least squares; AR(1) is a specification appropriate when the disturbance term follows a first-order autoregressive process; *RSS* = residual sum of squares; *d* = Durbin–Watson *d* statistic; $\hat{\rho}$ is the estimate of the autoregresive parameter in a first-order autoregressive process; (–1) means the lagged value of a variable.)
(see the table on the next page)

2.1. [10 marks] □ Discuss whether specification (1) is an adequate representation of the data.
□ Discuss whether specification (2) is an adequate representation of the data. (*Note*: *As part of the answer to this question, you will need to perform a Common Factor test, and you will need to explain why you are performing this test.*)
□ Discuss whether specification (3) is an adequate representation of the data.

2.2. [10 marks] □ Discuss whether the specification (4) is an adequate representation of the data.

□ At the seminar a commentator points out that in specification (4) neither *LGDPI* nor *LGPRLIFE* have significant coefficients. What is the economic meaning of this econometric phenomenon? She suggested to drop these variables from equation. As it happens, the researcher has considered this specification, and the results are shown as specification (5) in the table. What would be your answer to the commentator?

	(1) OLS	**(2)** AR(1)	**(3)** OLS	**(4)** OLS	**(5)** OLS
LGDPI	1.37 (0.10)	1.41 (0.25)	0.42 (0.60)	0.28 (0.17)	–
LGPRLIFE	– 0.67 (0.35)	– 0.78 (0.50)	– 0.59 (0.51)	– 0.26 (0.21)	–
LGLIFE(–1)	–	–	0.82 (0.10)	0.79 (0.09)	0.98 (0.02)
LGDPI(–1)	–	–	– 0.15 (0.61)	–	–
LGPRLIFE(–1)	–	–	0.38 (0.53)	–	–
Constant	– 4.39 (0.88)	– 4.20 (1.69)	– 0.50 (0.72)	– 0.51 (0.70)	0.12 (0.08)
R^2	0.958	0.985	0.986	0.986	0.984
RSS	0.2417	0.0799	0.0719	0.0732	0.0843
D	0.36	1.85	2.02	1.92	2.05
$\hat{\rho}$	–	0.82 (0.11)	–	–	–

In answering the practical questions, no general theoretical/mathematical explanations are acceptable, nor any information on the topic that is not directly related to the question.

3. [50 marks] Use data file **ha08_dataNN.wf1** The purpose of this exercise is to estimate and compare different models describing the relationship between total expenditure by US residents on physicians DOC_t and disposable personal income DPI_t, and the relative price index for physicians services $PRELDOC_t$, as well as several other factors on the base of annual data for 1959-2003. Full description of variables can be found in the file **HA08 Data Description.pdf**. The main focus should be on the problems associated with autocorrelation of residuals and methods for eliminating autocorrelation and mitigating its effects. In your analysis, ignore the problems associated with non-stationarity of the time series.

3.1. [10 marks]. □ Investigate the effect of equation specification on its quality characteristics, in particular the presence of autocorrelation. Consider the following equations **LS DOC C DPI; LS DOC C DPI PRELDOC; LS LGDOC C LGDPI; LS LGDOC C LGDPI LGPRDOC.** What signs of autocorrelation can be seen here? What changes when the specification is changed? (*use DW test when needed*)

□ Discuss the economic and econometric meaning of the estimated values of the equation coefficients separately and in comparison with each other.

3.2. [10 marks]. □ Add the AR(1) autoregressive transformation to the equations with one and two explanatory variables discussed in 3.1. What conclusions can be drawn from a comparative analysis of the estimated equations?

□ Repeat the same steps with the MA(1) transformation.

3.3. [10 marks]. □ Give a printout of the data obtained by repeated iterative autoregressive transformation of the variables **LGDOC C LGDPI LGPRDOC** (the first five rows of the data table are sufficient). Calculate a regression equation from these data and compare it with the original equation **LS LGDOC C LGDPI LGPRDOC AR(1)**.

□ An autoregressive transformation results in the loss of the first observation. Show both theoretically and practically that attempting to recover the first observation by mechanical addition from the original data leads to the problem of heteroscedasticity.

3.4. [10 marks]. □ Examine the logarithmic ADL(1,0) simple and multiple regression models by performing the Durbin h-test for them.

□ Examine them using the Breusch-Godfrey test manually and automatically.

3.5. [10 marks]. □ Conduct a common factor test for logarithmic ADL(1,1) simple and multiple regression models. What are your conclusions?.

If you have any questions please ask at **vladimir.tcherniak@gmail.com**