Econometrics – 2021-2022
Midterm Exam, March 31, 2022.
(One session 2 hours without break)

Suggested Solutions

SECTION A

Answer **ALL** questions from this section (questions **1-2**).

Question 1 (30 points) A student runs logarithmic regressions of the total expenditures of USA citizens on beer B (billions of dollars) on price index PB for the period 1996-2020 (100% corresponds to the level of the year 2020).

$$\ln \hat{B}_t = \underset{(0.09)}{1.56} + \underset{(0.02)}{0.21} \ln PB_t \qquad R^2 = 0.82 \quad d = 0.76 \qquad \text{(1-OLS)}$$

$$\ln \hat{B}_t = \underset{(4.84)}{5.18} - \underset{(0.20)}{0.28} \ln PB_t \qquad R^2 = 0.92 \quad d = 2.61 \qquad \text{(2-AR(1))}$$

(a) (12 points) □ Can equation (1) be considered as satisfactory?

Solution: Specification of the equation (1) is certainly wrong – the equation does not include important economic variables. So its coefficents are biased (the coefficient of $\ln PB_t$ has wrong sign), all tests are invalid.

The coefficient 0.21 can be interpreted as price elasticity of beer consumption, it means that with an increase in the price index by one percentage point the consumption of beer rises by 0.21%. Although this coefficient is significant, in the absence of important explanatory variables this equation is unsatisfactory. Durbin-Watson statistic equal to 0.76 indicates on the presence of positive autocorrelation: $0.76 < 1.05 = d_{crit}(1\%, 25, \textit{number of parameters} = 2, \textit{lower})$.

[6 marks]

□ What is the autoregressive transformation AR(1)? Show mathematically how autoregressive transformation works. Was this transformation successful (equation 2)?

Solution: Assuming 1-st order autocorrelation $u_t = \rho u_{t-1} + \varepsilon_t$, where $E\varepsilon_t = 0$, $E(\varepsilon_t^2) = \sigma_\varepsilon^2$, $E(\varepsilon_t \varepsilon_s) = 0, s \neq t$, we can subtract from initial equation

$$\ln B_t = \beta_1 + \beta_2 PB_t + \rho u_{t-1} + \varepsilon_t$$

the lagged equation multiplied by ρ:

$$\rho \ln B_{t-1} = \rho\beta_1 + \beta_2 \rho PB_{t-1} + \rho u_{t-1}$$

If ρ is defined correctly the disturbance term becomes $\rho u_{t-1} + \varepsilon_t - \rho u_{t-1} = \varepsilon_t$:

$$\ln B_t - \rho \ln B_{t-1} = \beta_1(1-\rho) + \beta_2(PB_t - \rho PB_{t-1}) + \varepsilon_t.$$

otherwise the autoregressive transformation can be applied repeatedly, refining each time the value of ρ from the Darbin-Watson statistics of the preceding equation $\hat{\rho} = 1 - d/2$. After repeating autoregressive transformation $d = 2.61$, that points to the negative autocorrelation.

Comparing $4 - d = 4 - 2.61 = 1.39$ with the corresponding critical value from the Durbin-Watson table $d_{crit}(5\%, 25, \textit{number of param.} = 2, \textit{lower}) = 1.30 < 1.39 < 1.46 = d_{crit}(5\%, 25, \textit{number of param.} = 2, \textit{upper})$ – we obtain dark zone, no conclusion.

[6 marks]

(b) (18 points) □ The teacher advised the student to include into equation more variables: $\ln DPI_t$ − logarithm of the gross disposable personal income (billions of dollars) and $TIME_t$ (which is equal 1 in 1996):

$$\ln \hat{B}_t = \underset{(1.74)}{7.23} - \underset{(0.11)}{0.37} \ln PB_t + \underset{(0.22)}{0.55} \ln DPI_t + \underset{(0.012)}{0.049} TIME_t \qquad R^2 = 0.93, \quad d = 1.26 \qquad \text{(3-OLS)}$$

Give interpretation to the coefficients of equation (3) and compare them with those of (1) and (2). Was the teacher's advice helpful?

Solution: The coefficient -0.37 is price elasticity of beer consumption evaluated under assumption of constant income and time, now it has correct sign. The coefficient 0.049 can be interpreted as follows: each year (keeping prices and income constant) the consumption of beer increases by $0.049 \cdot 100 = 4.9$ percent.
Omiting relevant variables in (1) could cause the autocorrelation of the disturbance term. Now DW statistic is bigger but not enough to cope with the problem of autocorrelation. In fact
$d_{crit}(5\%, 25, \mathit{number\ of\ param.} = 4\ \mathit{lower}) = 1.10 < 1.26 < 1.66 = d_{crit}(5\%, 25, \mathit{number\ of\ param.} = 4, \mathit{upper})$, – dark zone, so the problem remains unsolved.
[6 marks]

□ Additionally the student decided to do Breusch-Godfrey test for the first order autocorrelation getting value of the test statistic equal to 3.3088. Explain what is Breusch-Godfrey test and help the student interpret the test results.

Solution: Breusch-Godfrey test involves the evaluation of the auxiliary equation

$$. \ \ln B_t = \beta_1 + \beta_2 PB_t + \beta_3 \ln DPI_t + \beta_4 TIME_t + RESID_{t-1} + u_t$$

where $RESID_{t-1}$ are lagged residuals of equation (3). Under null hypothesis of no autocorrelation the statistics $T \cdot R^2$ (T is the number of observation, R^2 is R-squared of auxiliary equation) has χ^2-distribution with 1 degree of freedom (the number of lagged residuals in the auxiliary equation – here 1. As $\chi^2_{crit}(5\%, df = 1) = 3.8415 > 3.3088$ the null hypothesis of no autocorrelation cannot be rejected.
[6 marks]

□ A student friend said that beer consumption can be explained mainly by the habit of drinking beer. So she tries one more equation with the lagged variable

$$\ln \hat{B}_t = \underset{(0.51)}{0.62} - \underset{(0.10)}{0.08} \ln PB_t + \underset{(0.07)}{0.90} \ln B_{t-1} \qquad R^2 = 0.90, \quad d = 2.18 \qquad \text{(4-OLS)}$$

Was friend's advice helpful?

Solution: One should use h-statistic $h = (1 - d/2)\sqrt{\dfrac{n}{1 - n \cdot \mathrm{Var}(b)}}$ which is here $h = (1 - \dfrac{2.18}{2})\sqrt{\dfrac{24}{1 - 24 \cdot (0.06)^2}} = -0.461$. Using normal distribution $z_{crit}(5\%) = -1.96$ we come to the conclusion that hypothesis of no autocorrelation cannot be rejected. The slope coefficient is significant so the idea of strong influence of habits of drinking beer looks reasonable.
[6 marks]

Question 2. (30 points) The student of the course 'Elements of Econometrics' decided to investigate the influence of income DPI_t and relative prices on home electronics $PREL_{it}$ on the expenditures on home electronics EL_{it}. Different types of electronics (TV, computers, phones etc. – 7 items in total), considered in 1977-1921 (45 years) form a panel. The model under investigation is

$$\log(EL_{it}) = \beta_1 + \beta_2 \log(DPI_t) + \beta_3 \log(PREL_{it}) + u_{it};$$

The student runs three alternative approaches to the evaluation of this model: 1) pooled OLS regression, 2) fixed effects panel regression model, and 3) random effects panel regression model. The results of the estimation are shown below

$$\begin{array}{ll} \log(EL)_{it} = \underset{(2.47)}{4.28} + \underset{(0.24)}{0.63}\log(DPI_t) - \underset{(0.15)}{0.83}\log(PREL_{it}) + e_{it} \quad DW = 0.02 & R^2 = 0.11;\ RSS = 514.9 \end{array} \quad \text{(1 - Pooled)}$$

$$\begin{array}{ll} \log(EL)_{it} = \underset{(0.46)}{3.81} + \underset{(0.05)}{0.62}\log(DPI_t) - \underset{(0.05)}{0.92}\log(PREL_{it}) + e_{it} \quad DW = 1.05 & R^2 = 0.98;\ RSS = 12.5 \end{array} \quad \text{(2 - Fixed)}$$

$$\begin{array}{ll} \log(EL)_{it} = \underset{(0.73)}{3.18} + \underset{(0.04)}{0.65}\log(DPI_t) - \underset{(0.06)}{0.89}\log(PREL_{it}) + e_{it} \quad DW = 1.04 & R^2 = 0.73;\ RSS = 12.7 \end{array} \quad \text{(3 - Random)}$$

(you can assume that the LSDV approach is used when evaluating the fixed effects model)

(a) (12 marks) □ The student is worried that none of estimated equations (1)-(3) contained information about unobserved heterogeneity which, as he remembers from the lecture, is one of the main features of panel data. Explain to the student the differences between the three methods, stating the reasons for the absence of unobserved heterogeneity term in the results.

Solution: General structure of panel model is as follows

$$Y_{ij} = \beta_0 + \sum_{k=1}^{K} \beta_k X_{ijk} + \alpha_i + u_{ij}$$

α_i is referred to as unobserved heterogeneity term. In fact $\alpha_i = \sum_{p=1}^{s} \gamma_p Z_{pi}$ where Z_p – are some unobservable factors.

In pooled model unobserved heterogeneity α_i is ignored that can lead to omitted variable bias.

In fixed effect panel models the unobserved heterogeneity α_i is attributed to fixed differences between cross sectional objects (types of electronics in our case). A set of dummy variables D_i is defined, where D_i is equal to 1 in the case of an observation relating to an individual i and 0 otherwise. The model can be written as

$$Y_{it} = \sum_{k=1}^{7} \beta_k X_{itk} + \sum_{i=1}^{7} \alpha_i D_i + u_{it}.$$

The intercept is not included in the model to avoid dummy variable trap.

In random effect panel models the unobserved heterogeneity α_i is a part of disturbance term, which is assumed not be correlated with explanatory variables.

$$Y_{it} = \beta_0 + \sum_{k=1}^{7} \beta_k X_{itk} + (\alpha_i + u_{it})$$

[8 marks]

□ Help the student interpret the coefficients of the regression (2). The printout of regression (2) contains data on fixed effects (see table to the right). Explain to the student what are fixed effects and how this information can be used in the interpretation of the regression results.

EL Type	Fixed Effect
1	1.228209
2	-1.886449
3	2.008554
4	0.636180
5	-1.041255
6	-0.207966
7	-0.737273

Solution: The coefficients of variables are correspondingly income and relative price elasticities of expenditures on medical services (other factor fixed), they are all reasonable from economic point of view.
Fixed effects allow us to estimate how the regression plane shifts for each type of electronics (to get the regression equation for a particular type, just add the value of the fixed effect to the LSDV equation constant, for example, for type 1 the constant is equal to $3.81+1.23=5.04$, assuming that income and price elasticities are the same for all types of electronins.
[4 marks]

(b) (18 marks) □ Help the student to choose between models (1-pooled) and (2-fixed). Do appropriate test stating null hypothesis and giving explanation to the paremeters used, and make a conclusion.
Solution: This test evaluates the joint significance of the variables used in the LSDV method, and is based on the null hypotheses that there is no unbserved heteroscedasticity in the data.
To answer this question we need to compare RSS_{POOLED} with RSS_{LSDV}, or R^2_{POOLED} with R^2_{LSDV} using F-test:

$$F=\frac{(RSS_{POOLED}-RSS_{LSDV})/(7-1)}{RSS_{LSDV}/(7\cdot 45-7-2)}=\frac{(514.9-12.5)/6}{12.5/306}=2049.8$$

or $F=\dfrac{(R^2_{LSDV}-R^2_{POOLED})/(7-1)}{(1-R^2_{LSDV})/(7\cdot 45-7-2)}=\dfrac{(0.98-0.11)/6}{(1-0.98)/306}$=2218.5

(7 - number of crosssection objects – types of electronics), 45 - number of years, 2 - number of variables) while $F^{1\%}_{crit}(6,300)=2.86$ $\Rightarrow$ significant, there is significant unobserved heterogeneity, so we choose fixed panel regression.
[6 marks]

□ A friend of the student said that it would be better to use some other method for fixed effects regression instead of LSDV. Comment on this proposal.

Solution: The difference and within groups methods can also be used (brief standard mathematical description expected – omitted here). In terms of the number of degrees of freedom used, all three methods are equivalent.
As the Durbin-Watson statistics in the equation (1-pooled) DW=0.02 shows, the problem of autocorrelation does exist. In regression (2) the problem is not fixed DW=1.05. Using the first difference method can help neutralize positive autocorrelation by introducing a moving average type random term that causes negative autocorrelation. However, this method has its drawbacks (it is based on the small changes of the variables, we lose parameters that are constant over time).
[6 marks]

□ Help the student to choose between equation (2), and equation (3) using Hausman test. The value of test statistic for the Hausman test when comparing equations (2) and (3) was evaluated as 0.136. Comment on the idea of the test, assumptions used and results. What is the final choice of the regression and what are its advantages and risks.

Solution: Darbin-Wu-Hausman (DWH) test is used to choose between fixed and random effects. It compares estimates of coefficients obtained by two alternative models. Under H_0 (absence of endogeneity expressed in correlation of a random term with explanatory variables and so there is no difference between coefficients obtained by two alternative models) both fixed effect and random effect models provide us with consistent estimates. The test is based on chi-square statistic with degrees of freedom equal to the number of variables in the equation under consideration (2 in our case). The estimated value of test statistic is 0.136. The critical value for 5% is $\chi^2_{crit}(2)=5.99$, so null hypothesis is not rejected, there is no difference in results. Random effect panel model should be chosen as the most efficient (advantage). The risk is that if there is a correlation between the random term and the explanatory variables, inconsistent estimates of the coefficients of the equation will be obtained.
[6 marks]

SECTION B
Answer only ONE question from this section (question 3 OR question 4)

Question 3.
(a) (15 points) □ What are the conditions for a series to be covariance stationary? What is non-stationary time series? Why it is important for time series to be stationary?

Solution:
For a process x_t to be stationary, its expected value, variance, and covariance of x_t and x_{t+s} are independent of time *t*. Violation of any of these conditions makes time series non-stationary. The importance of the stationarity condition is determined by the fact that regressions built on non-stationary series turn out to be spurious.
[3 marks]

□ Let $x_t = x_{t-1} + \varepsilon_t$, where ε_t is i.i.d. with mean zero and variance σ^2 for $t = 1, ..., T$. Is the time series x_t stationary?

Solution:
This is random walk without drifrt. Lagging $x_{t-1} = x_{t-2} + \varepsilon_{t-1}$ and substituting obtain $x_t = x_{t-2} + \varepsilon_t + \varepsilon_{t-1}$ and so on so finally $x_t = x_0 + \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + ... + \varepsilon_1$.
$Ex_t = Ex_0 + E\varepsilon_t + E\varepsilon_{t-1} + E\varepsilon_{t-2} + ... + E\varepsilon_1 = Ex_0$ – does not depend on time. But
$\operatorname{var} x_t = \operatorname{var} x_0 + \operatorname{var} \varepsilon_t + \operatorname{var} \varepsilon_{t-1} + \operatorname{var} \varepsilon_{t-2} + ... + \operatorname{var} \varepsilon_1 = \operatorname{var} x_0 + t\sigma^2$ –depends on time, so the time series is non stationary.
[6 marks]

□ Let $x_t = \lambda^2 \varepsilon_t + \lambda \varepsilon_{t-1} + \varepsilon_{t-2}$, where ε_t is i.i.d. with mean zero and variance σ^2 for $t = 1, ..., T$. Is the time series x_t stationary?

Solution:
$Ex_t = \lambda^2 E\varepsilon_t + \lambda E\varepsilon_{t-1} + E\varepsilon_{t-2} = 0$ – does not depend on time.
$\operatorname{var} x_t = \lambda^2 \operatorname{var} \varepsilon_t + \lambda \operatorname{var} \varepsilon_{t-1} + \operatorname{var} \varepsilon_{t-2} = (\lambda^2 + \lambda + 1)\sigma^2$ – does not depend on time.

$\operatorname{cov}(x_t, x_{t-1}) = \operatorname{cov}(\lambda^2 \varepsilon_t + \lambda \varepsilon_{t-1} + \varepsilon_{t-2}, \lambda^2 \varepsilon_{t-1} + \lambda \varepsilon_{t-2} + \varepsilon_{t-3}) = \operatorname{cov}(\lambda \varepsilon_{t-1}, \lambda^2 \varepsilon_{t-1}) + \operatorname{cov}(\varepsilon_{t-2}, \lambda \varepsilon_{t-2}) = (\lambda^3 + \lambda)\sigma^2$
$\operatorname{cov}(x_t, x_{t-2}) = \operatorname{cov}(\lambda^2 \varepsilon_t + \lambda \varepsilon_{t-1} + \varepsilon_{t-2}, \lambda^2 \varepsilon_{t-2} + \lambda \varepsilon_{t-3} + \varepsilon_{t-4}) = \operatorname{cov}(\varepsilon_{t-2}, \lambda^2 \varepsilon_{t-2}) = \lambda^2 \sigma^2$
For $s > 2$ $\operatorname{cov}(x_t, x_{t-s}) = \operatorname{cov}(\lambda^2 \varepsilon_t + \lambda \varepsilon_{t-1} + \varepsilon_{t-2}, \lambda^2 \varepsilon_{t-s} + \lambda \varepsilon_{t-s-1} + \varepsilon_{t-s-2}) = 0$ – all covariances does not depend on time. The time series x_t is stationary.
[6 marks]

(b) (12 points) □ Now consider the following time series model for $\{x_t\}_{t=1}^T$ that combines two time series above:

$$x_t = x_{t-1} + \lambda^2 \varepsilon_t + \lambda \varepsilon_{t-1} + \varepsilon_{t-2},$$

where ε_t is i.i.d. with mean zero and variance σ^2, for $t = 1, ..., T$, and $x_0 = 0$
Show that x_t can be represented as

$$x_t = x_1 + \lambda^2 \varepsilon_t + (\lambda + \lambda^2)\varepsilon_{t-1} + (1 + \lambda + \lambda^2)(\varepsilon_{t-2} + ... + \varepsilon_2) + (\lambda + 1)\varepsilon_1 + \varepsilon_0$$

Explain your work.

Solution:
Lagging one period

$$x_{t-1} = x_{t-2} + \lambda^2 \varepsilon_{t-1} + \lambda \varepsilon_{t-2} + \varepsilon_{t-3}$$

and then substituting back to the model, until time *t* = 2, we have:

$$x_t=(x_{t-2}+\lambda^2\varepsilon_{t-1}+\lambda\varepsilon_{t-2}+\varepsilon_{t-3})+(\lambda^2\varepsilon_t+\lambda\varepsilon_{t-1}+\varepsilon_{t-2})=x_{t-2}+\lambda^2\varepsilon_t+(\lambda^2+\lambda)\varepsilon_{t-1}+(\lambda+1)\varepsilon_{t-2}+\varepsilon_{t-3}$$

Lagging again and substituting

$$x_{t-2}=x_{t-3}+\lambda^2\varepsilon_{t-2}+\lambda\varepsilon_{t-3}+\varepsilon_{t-4}$$

$$x_t=x_{t-3}+\lambda^2\varepsilon_t+(\lambda^2+\lambda)\varepsilon_{t-1}+(\lambda^2+\lambda+1)\varepsilon_{t-2}+(\lambda+1)\varepsilon_{t-3}+\varepsilon_{t-4}$$

Repeating again

$$x_{t-3}=x_{t-4}+\lambda^2\varepsilon_{t-3}+\lambda\varepsilon_{t-4}+\varepsilon_{t-5}$$

$$x_t=x_{t-4}+\lambda^2\varepsilon_t+(\lambda^2+\lambda)\varepsilon_{t-1}+(\lambda^2+\lambda+1)\varepsilon_{t-2}+(\lambda^2+\lambda+1)\varepsilon_{t-3}+(\lambda+1)\varepsilon_{t-4}+\varepsilon_{t-5}$$

Now it becomes clear that finally

$$x_t=x_1+\lambda^2\varepsilon_t+(\lambda+\lambda^2)\varepsilon_{t-1}+(1+\lambda+\lambda^2)(\varepsilon_{t-2}+...+\varepsilon_2)+(\lambda+1)\varepsilon_1+\varepsilon_0.$$

[12 marks]

(c) (13 points) □ Discuss whether there are any values of λ so that x_t is covariance stationary

Solution:

Expectation $Ex_t=Ex_1+\lambda^2E\varepsilon_t+(\lambda+\lambda^2)E\varepsilon_{t-1}+(1+\lambda+\lambda^2)(E\varepsilon_{t-2}+...+E\varepsilon_2)+(\lambda+1)E\varepsilon_1+E\varepsilon_0=x_1$ and so does not depend on time.

The variance of x_t is: $\mathrm{Var}(x_t)=\sigma^2[\lambda^4+(\lambda+1)^2\lambda^2+(\lambda^2+\lambda+1)^2(t-3)+1]$. For the variance to be independent of *t*, we must have $\lambda^2+\lambda+1=0$. There is no such λ because $\lambda^2+\lambda+1=\left(\lambda+\frac{1}{2}\right)^2+\frac{3}{4}>0$ for all λ.

As one of the conditions of stationarity is violated we do not need in analizing covariances.

[13 marks]

Question 4.

Let us consider the relationship between the natural logarithm of GDP, GDP_t, and the lagged long-term interest rates $rate_t$, and $rate_{t-1}$.

(a) (10 points) □ Assume that GDP_t and $rate_t$ are difference stationary, or they are integrated of order 1: $GDP_t \in I(1)$, $rate_t \in I(1)$. What does it mean? How to test that a time series is $I(1)$?

Solution: If a nonstationary process can be transformed into a stationary process by differencing then the series is said to be *difference stationary*.

To test whether a time series x_t is $I(1)$ one should apply Dickey-Fuller test to x_t and if it is not stationary to repeat the test for Δx_t.

[4 marks]

□ What is cointegration? How to test whether time series GDP_t and $rate_t$ are cointegrated.

Solution: Definition: Several time series are called cointegrated if

1) They all are of the same order of integration (say I(1))
2) There is a linear combination of them that it is stationary

The second condition is equivalent to the following: the residuals of regression of one of them on the others are stationary.

So to test time series GDP_t and $rate_t$ for cointegration is is sufficient to do the following:

1) Check whether GDP_t and $rate_t$ are of the same order of integration (we are given $GDP_t \in I(1)$, $x_t \in I(1)$**).**

2) Run regression estimated by OLS: $GDP_t = \alpha_0 + \alpha_1 rate_t + u_t$ to get its residuals e_t

3) Run new regression of the type $\Delta\hat{e}_t = \theta e_{t-1} + \sum \theta_i \Delta e_{t-i}$ and test θ using the ADF test.

If the null $H_0 : \theta = 1$ is not rejected $\Rightarrow$ e_t is non-stationary. This means that GDP_t and $rate_t$ are not cointegrated. For GDP_t and $rate_t$ to be cointegrated the null has to be rejected.

[6 marks]

(b) (10 points) □ If GDP_t and $rate_t$ are not co-integrated, what are the properties of an OLS estimator in the regression of GDP_t on $rate_t$?

Solution:

Technically, there exists no linear combination of the variables that becomes stationary *I*(0).

It means that of the two variables are not co-integrated, the difference between GDP_t and $rate_t$ could become very large, with no tendency for them to come back together.

The regression result, if any, is spurious, and it tells us nothing meaningful about the relationship between the variables.

The OLS estimator will not be consistent, the R^2 is superficially high. The error term is likely *I*(1) and the statistical inference becomes invalid. Any significant parameter is meaningless.

[5 marks]

□ Discuss the implications of the existence of cointegration between GDP_t and $rate_t$ on the short-run and long-run relationship of two variables.

Solution:

Two *I*(1) processes are co-integrated when there is a long-term relationship between these two.

Co-integration tells us nothing about the short-run dynamics between the two variables, but it indicates that there must be some short-term forces that maintain the long-term relationship.

This implies that it should be possible to construct a more comprehensive model that combines the short-run and long-run dynamics of the relationship. For example, we can conduct an Error Correction Model that describes the relationship.

[5 marks]

(c) (20 points) Consider the following regression model, where ε_t has mean zero and is uncorrelated with GDP_{t-1}, GDP_{t-2}, $rate_t$, and $rate_{t-1}$, where $GDP_t \in I(1)$, $rate_t \in I(1)$ and are conintegrated:

$$GDP_t = \alpha + \beta_1 GDP_{t-1} + \beta_2 GDP_{t-2} + \gamma_1 rate_t + \gamma_2 rate_{t-1} + \varepsilon_t \qquad (1)$$

□ Derive the Error Correction Model (ECM) representation of the equation, and discuss the long-run equilibrium relation for GDP_t and $rate_t$.

Solution:

We derive the ECM by adding and subtracting the lagged term of GDP_t:

$$GDP_t = \alpha + \beta_1 GDP_{t-1} + \beta_2 GDP_{t-2} + \gamma_1 rate_t + \gamma_2 rate_{t-1} + \varepsilon_t,$$

$$GDP_t - GDP_{t-1} = \alpha + (\beta_1 - 1) GDP_{t-1} + \beta_2 GDP_{t-2} + \gamma_1 rate_t + \gamma_2 rate_{t-1} + \varepsilon_t,$$

$$GDP_t - GDP_{t-1} = \alpha + (\beta_1 + \beta_2 - 1) GDP_{t-1} + \beta_2 (GDP_{t-2} - GDP_{t-1}) +$$
$$+ \gamma_1 (rate_t - rate_{t-1}) + (\gamma_2 + \gamma_1) rate_{t-1} + \varepsilon_t.$$

Simplify the expression, we have the following ECM model:

$$\Delta GDP_t = -\beta_2 \Delta GDP_{t-1} + \gamma_1 \Delta rate_t + (\beta_1 + \beta_2 - 1)\left(GDP_{t-1} - \frac{\gamma_2 + \gamma_1}{1 - \beta_1 - \beta_2} rate_{t-1} - \frac{\alpha}{1 - \beta_1 - \beta_2} \right) + \varepsilon_t.$$

[12 marks]

□ Interpret the parameters of the model and discuss the properties of the OLS estimator of the ECM in comparison with the model (1), as well as in comparison with the model built only on differences.

Solution:

The error correction term, which measures the disequilibrium at time $t-1$, when $\Delta GDP_t = \Delta rate_t = 0$, is:

$$GDP_{t-1} - \frac{\gamma_2 + \gamma_1}{1 - \beta_1 - \beta_2} rate_{t-1} - \frac{\alpha}{1 - \beta_1 - \beta_2}.$$

The parameter $(\beta_1 + \beta_2 - 1)$ indicates how quickly the variable GDP_t will error correct in the long run: if GDP_{t-1} is larger than the equilibrium value, then $\Delta GDP_t < 0$ to bring GDP_t back to equilibrium.

The parameters $-\beta_2$ and γ_1 are the direct effects of a change in the past *GDP* level and the current interest rate has on the current *GDP* level.

The error term ε_t is now a white noise (because of the co-integration), an OLS estimator will be valid and super-consistent.

Model (1) built on the non-stationary series GDP_t and $rate_t$ with a high probability turns out to be sputious, the conclusions obtained with its help cannot be taken into account.

The difference model $\Delta GDP_t = -\beta_2 \Delta GDP_{t-1} + \gamma_1 \Delta rate_t + \varepsilon_t$ uses stationary series ΔGDP_t and $\Delta rate_t$,, but takes into account only short-term changes (myopia), which makes the forecast unreliable.

All elements of the ECM model (differences of variables and residuals) are stationary, while combining short-term and long-term information improves the quality of the forecast.

The error term ε_t is now a white noise (because of the co-integration), an OLS estimator will be valid and super-consistent.

[8 marks]