Econometrics – 2019-2020.
Training exam. March 27.
Solution

<u>General instructions.</u> Candidates should answer SIX of the following SEVEN questions: all 4 questions of the Section A and any 2 of the questions from Section B (questions 5-7). The weight of the Section A is 60% of the exam; three other questions from the Section B add 20% each. You are advised to divide your time accordingly. Structure your answers in accordance with the structure of the questions. When testing hypotheses state clearly null and alternative hypotheses provide critical value(s) used for the test, mentioning degrees of freedom, the significance level chosen for the test and the assumptions for the test to be valid.

SECTION A

(1 hour 50 minutes + 10 minutes for reading)

Answer **ALL** questions from this section (questions **1-4**).
Each question in this section bears **15 marks**

Question 1 The relationship between the US citizens' expenditure on local transport *LOCT* in 1989-2013 (in billions of dollars) and aggregate personal disposable income *DPI* (also in billions of dollars) is studied, this relationship is described with the following equation:

$$\hat{LOCT_t} = 5.31 - 0.28 \ln DPI_t$$
$$(s.e.) \qquad (1.02)\ (0.15)$$

(a) (5 marks) □ Perform t-tests for significance of both coefficients of the model.

Solution: For the slope t-statistic is $t = \dfrac{-0.28}{0.15} = -1.87$ while $t_{crit}(5\%, 23) = 2.069$, so the slope coefficient is insignificant; for intercept $t = \dfrac{5.31}{1.02} = 5.21$ while $t_{crit}(1\%, 23) = 2.807$ so it is significant at 1% level.

□ Give interpretation to both coefficients.

Solution: Formally the slope coefficient -0.28 shows that if disposable personal income increases by 1% the expenditure on local transport decreases by $\dfrac{0.28}{100} = 0.0028$ billions of dollars $= 2.8$ millions of dollars.
But it is insignificant, so for the equation

$$LOCT_t = \beta_1 + \beta_2 \ln DPI_t + u_t$$

the null hypothesis $H_0 : \beta_2 = 0$ is not rejected, so the equation reduces to $LOCT_t = \beta_1 + u_t$ where β_1 can be interpreted as the average expenditures on local transport independently of DPI_t.

(b) (5 marks) □ In the equation above, there is no information about the R-square. Is it possible to restore its value from the available information?

Solution: For the simple linear regression t-test and F-test are equivalent so according F-test the equation is olso insignificant. Moreover $F = t^2 = (1.87)^2 = 3.5$, so solving equation $\dfrac{R^2}{1-R^2}(25-2) = 3.5$ find $R^2 = 0.132$.

□ Is R-square high enough to make the equation significant?

Solution: Insignificance of the equation follows from the insignificance of slope (in simple regression model F-test and two-sided t-test are equivalent).

(c) (5 marks) At the seminar one of the participants remarked that coefficient of $\log DPI_t$ is expected to be negative as people tend to use personal cars instead of local transport as their income rises. The other participant objected to him saying that just the opposite is true (the coefficient of $\log DPI_t$ should be positive) as local transport is used mainly by elderly people who prefer not to use cars and are quite sensitive to changes in income.

□ How both suggestions change your conclusion on significance of the coefficients of the model?

Solution: If the pair of hypotheses $H_0 : \beta_2 = 0; H_a : \beta_2 < 0$ is used with one sided critical value $t_{crit}(5\%, 23, one\ sided) = -1.714$ the slope coefficient becomes significant.
If the pair of hypotheses $H_0 : \beta_2 = 0; H_a : \beta_2 > 0$ is used the coefficient is obviously insignificant because the observed value contradicts the alternative.

Question 2 Recently, the Central Bank of Russian Federation revoked the licenses of several Russian banks, some other banks are considered as problematic and some of them are at risk of revocation of license in the near future. In this connection customers and entrepreneurs are interested in the estimating the state of the bank to select reliable banks. Preliminary expert analysis showed that such variables can be useful:

X_1 - percentage of the amount in foreign currency liabilities in the balance sheet of the bank;

X_2 - share of total loans in the amount of operating assets, and

X_3 - the ratio of the balance sheet profit to balance losses.

A survey was conducted of 100 largest banks (in terms of value of liabilities and capital), which resulted in the following data: ($X_{1i}, X_{2i}, X_{3i}, Y_i, i = 1,...,100$), where $Y_i = 1$, if the Central Bank attributes this bank to the category of problematic banks, and $Y_i = 0$ otherwise. Suppose you are invited as a consultant to build the regression model Y_i as a function of X_{1i}, X_{2i}, X_{3i}.

(a) (5 marks) □ Explain why linear regression model $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$ estimated using OLS is not appropriate in this case. Give some mathematical details.

Solution: Y_i is a binary choice variable, it can be interpreted as probability that the bank will be considered as problematic. The linear probability model

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

estimated by OLS, is not appropriate here as there are 4 main disadvantages of the linear probability model:
1) possible predicted values can be outside the range of probability [0; 1];
2) it assumes constant marginal effect of each factor;
3) heteroscedasticity is present caused by discontinuous disturbance term, and so Gauss-Markov conditions are violated, so OLS produces inefficient estimates;
4) the distribution of the disturbance term is not normal, so the useful tests of coefficients are not valid.

(b) (5 marks) □ Explain why probit model can solve some problems indicated in (a) providing the rationale of the proposed form of the function, and explaining why the obtained estimates have acceptable econometric properties.

Solution: One of the best solutions is to construct the binary choice model $Y_i = F(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i)$ where $F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{x^2}{2}} dx$ (probit model).

The most important properties of this model are

1) predicted values $\hat{Y}_i$ belong to the range [0; 1];
2) the marginal effect is not constant, it is maximal for $z = 0$ and is negligible for extreme values of z;
3) For estimation of the models the maximum likelihood estimation method is used that provides asymptotically efficient estimates.
4) The estimates obtained by the maximum likelihood method prove to be efficient even under conditions of inevitable heteroscedasticity.

□ How to estimate that some factor, say, X_3, is significant on the basis of the proposed model?

Solution: The new system of tests based on the values of logarithmic likelihood function, is used instead of tests based on RSS in OLS estimation. To estimate that some factor, say, X_3, is significant, it is sufficient to evaluate its asymptotic z-values (instead of conventional t-statistics) (provided by the software), and compare it with the critical values of standard normal distribution. All tests under consideration are asymptotic so the sample is supposed to be large enough (100 in our case).

□ How to estimate that a group of factors, for example, X_1, X_2, is significant on the basis of the same model?

Solution: To evaluate the significance of the group of two X_1, X_2 (or more) variables we need some analog of F-test. It is so called LR-test based on the statistic $LR = 2(\log L_U - \log L_R)$, where L_U is likelihood function of the unrestricted model, evaluated for the function including all variables

$$Y_i = F(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i)$$

and L_R is the value of likelihood function evaluated for the restricted model ($H_0 : \beta_1 = \beta_2 = 0$):

$$Y_i = F(\beta_0 + \beta_1 X_{3i} + u_i).$$

LR-statistic has χ^2-distribution with the degrees of freedom equal to the number of restrictions (2 in our case).

(c) (5 marks) Suppose some characteristic, say X_3 has changed while maintaining other explanatory variables unchanged.

□ How to evaluate the effect of this change on the probability of revocation of license in the near future?

Solution: We can estimate marginal effect of some factor say X_3 approximately using the defferential of the function $Y = F(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u)$. For the probit model the marginal effect is evaluated as $\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \cdot \beta_3 dX_{3i}$ where z is evaluated in the initial point (X_{1i}, X_{2i}, X_{3i}): $z = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i}$ and dX_{3i} is a small change of X_{3i}..

Bonus answer: Alternative methods allowing to estimate the effect accurately (even in the case of large changes), you need to calculate the value of the probability after the change of the factor $X_3 + \Delta X_3$ and subtract from it the value of the probability calculated before the change of the factor X_3:

$$\Delta Y = F(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_3 + \Delta X_3 + u) - F(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_3 + u)$$

□ Is it possible to evaluate maximum marginal effect of some factor, say X_3 (while keeping X_1, X_2 fixed)? How to find the value of X_3, which provides this maximum, considering X_1, X_2 known?

Solution: As it is said above the maximum effect in both cases achieved at $z = 0$, so to find the value of X_3 for which the maximum effect achieved under fixed X_1, X_2. To do this we have to solve the equation $\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 = 0$ for X_3.

Question 3 Let the multiple linear regression equation be

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i; \quad i = 1, 2, ..., N.$$

(a) (5 marks) □ Outline briefly, how you would test $\beta_2 = \beta_3 = 0$.

□ Outline briefly, how you would test $\beta_1 = 0$.

Solution: To test $\begin{pmatrix} H_0 : \beta_2 = \beta_3 = 0, \\ H_a : otherwise \end{pmatrix}$ we need to know R_1^2 from equation

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \text{ (1).}$$

Then we run standard F-test using statistics $F = \dfrac{R_1^2/2}{(1-R_1^2)/(N-3)}$.(instead of all tests based on R^2 one can use equivalent tests based on RSS.

(b) (5 marks) □ Outline briefly, how you would test $\beta_2 = \beta_3$.

□ Outline briefly, how you would test $\beta_2 = \beta_3$ against one-sided alternative $\beta_2 > \beta_3$

Solution: To test $\begin{pmatrix} H_0 : \beta_2 = \beta_3, \\ H_a : otherwise \end{pmatrix}$ we can run additionally restricted regression

$$Y_i = \alpha_1 + \alpha_2 (X_{2i} + X_{3i}) + u_i \text{ (2)}$$

and finding R_2^2 compare it with R_1^2 using F-test $F = \dfrac{(R_1^2 - R_2^2)}{(1-R_1^2)/(N-3)}$.

Alternative way is to use standard t-test for coefficient $\beta_2 - \beta_3$ of the equation

$$Y_i = \beta_1 + (\beta_2 - \beta_3) X_{2i} + \beta_3 (X_{3i} + X_{2i}) + u_i \text{ (3)}$$

Testing one sided pair of hypotheses $\begin{pmatrix} H_0 : \beta_2 = \beta_3, \\ H_a : \beta_2 > \beta_3 \end{pmatrix}$ is possible only on the basis of equation (3), but now one should use one-sided t-test.

(c) (5 marks) □ What assumptions are necessary to all of these tests were valid.

Solution: All these tests require the disturbance term of equations (1-3) to satisfy Gauss-Markov conditions (so the equations to be valid). Additionally disturbance terms should be distributed normally as it provides that statistics $t = \dfrac{\hat{\beta}}{s.e.(\hat{\beta})}$ have t-distribution for all coefficients of equations (1-3) and also the statistics $F = \dfrac{R_1^2/2}{(1-R_1^2)/(N-3)}$ and $F = \dfrac{(R_1^2 - R_2^2)}{(1-R_1^2)/(N-3)}$ have F-distribution with corresponding degrees of freedom.

To apply one-tailed test in c3 we should assume that $\beta_2 < \beta_3$ is impossible, so this case can be excluded.

Question 4. Suppose that the time series $HOUS_t$ (Housing), WAT_t (Water), GAS_t (Gas), $FUEL_t$ (Fuel oil), KIT_t (Kitchen appliances), TAB_t (Tableware) form the panel (as observations for particular units of some general type of good ($GOOD_{it}$) related to maintaining a home). DPI_t is the disposable personal income and $PRGOOD_{it}$ is the relative price index for corresponding $GOOD_{it}$. $t = 1, 2, ..., T$, $T = 40$, $i = 1, 2, ..., n$, $n = 6$
Let the model under investigation be

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

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.

(a) (5 marks) □ Help the student to choose between different approaches: fixed effects panel regression model, and random effects panel regression model. Indicate corresponding test, null hypothesis, type of distribution of the test statistic, the number of degrees of freedom and the decision rule (which alternative is chosen if null hypothesis is rejected and which if it is not).

Solution: In fixed effect panel models unobserved heterogeneity α_i is attributed to fixed differences between cross sectional objects.
In random effect panel models unobserved heterogeneity α_i as a part of disturbance term, which is assumed not be correlated with DPI_t and $PRGOOD_{it}$.

Darbin-Wu-Hausman (DWH) test is used to choose between fixed and random effects. It is based on using chi-square statistics with degrees of freedom equal to the number of variables in the equation under consideration (2 in our case). It compares estimates of coefficients obtained by two alternative models. Under H_0 that there is no difference between coefficients obtained by two alternative models – both fixed effect and random effect models provide us with consistent estimates.

□ What are advantages and risks of each choice?

Solution: If H_0 is not rejected we choose in this case random effect models as it retains in disturbance term all unobserved heterogeneity, there is no reduction of degrees of freedom typical for fixed effects models.

If H_0 is rejected, so there are essential differences between coefficients we choose fixed effects model, because rejecting of H_0 means that main assumption of independence of the disturbance term from regressors is violated so using random effects model we are under risk of getting inconsistent estimates of parameters.

Bonus answer: Note that fixed effect model drops intercept (all factors that are constant in time) and includes a set of additional dummy variables (LSDV model) so it involves estimating an additional set of coefficients, that leads to reduction of degrees of freedom and so sometimes we observe more insignificant coefficients (for example in first difference or in within group versions of fixed effect models).

(b) (5 marks) Any model of panel data is based on the assumption that unobserved heterogeneity is present in the data under consideration.

□ Explain what is unobserved heterogeneity, describing general structure of the panel data model.

Solution: General structure of the panel data model is

$$Y_{it} = \beta_0 + \sum_{j=1}^{k} \beta_j X_{jit} + \alpha_i + u_{it}$$

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.

□ How to test whether there is a significant difference between pooled regression and separate regressions?

Solution: Conventional Chow test could be performed on the base of comparison of pooled regression with 6 separate regressions for $HOUS_t$, WAT_t, GAS_t, $FUEL_t$, KIT_t, and TAB_t under H_0 that there is no difference between separate regressions. If H_0 is rejected then some panel data model should be applied.

□ How to test for the presence of unobserved heterogeneity if random effect model was chosen in (a)?

Solution: If the random effect model is chosen on the base of DWH test, it is possible that ordinary least squares is even better in efficiency if there is in fact no unobserved heterogeneity and so there no random effects at all. There are some tests for this purpose, for example Breush-Pagan test based on Lagrange Multiplier approach. It also uses chi-square distribution with degrees of freedom equal to 1 under H_0 of the absence of random effects.

(c) (5 marks) □ How to test for the presence of unobserved heterogeneity if fixed effect model was chosen in (a)? Give some details for the test statistic, degrees of freedom and used critical values. Do the test and make the conclusion.

Solution: To answer this question is simpler in case of LDDV fixed effects model of the type

$$\log(GOOD_{it}) = \beta_2 \log(DPI_t) + \beta_3 \log(PRGOOD_{it}) + \sum_{i=1}^{6} \gamma_i D_i + u_{it}. \qquad (1)$$

where D_i are dummies corresponding different goods under consideration. To take into account that the elasticities of $PRGOOD_{it}$ could be different for different goods, it is possible to run pooled regression (based on the assumption of equal elasticities) and fixed effect panel regression (that allows models for different goods be different), and then perform an F-test:

$$F = \frac{(RSS_{POOLED} - RSS_{LSDV})/(n-1)}{RSS_{LSDV}/(n \cdot T - n - K)} = \frac{(RSS_{POOLED} - RSS_{LSDV})/(6-1)}{RSS_{LSDV}/(6 \cdot 40 - 6 - 2)} = \frac{(RSS_{POOLED} - RSS_{LSDV})/5}{RSS_{LSDV}/232}$$

or $F = \dfrac{(R^2_{LSDV} - R^2_{POOLED})/(n-1)}{(1-R^2_{LSDV})/(n \cdot T - n - K)} = \dfrac{(R^2_{LSDV} - R^2_{POOLED})/(6-1)}{(1-R^2_{LSDV})/(6 \cdot 40 - 6 - 2)} = \dfrac{(R^2_{LSDV} - R^2_{POOLED})/5}{(1-R^2_{LSDV})/232}$

Here RSS_{LSDV} and R^2_{LSDV} - the values of RSS and R^2 for LSDV fixed effect model, RSS_{POOLED} and R^2_{POOLED}- the values of RSS and R^2 for pooled model, $n = 6$ - number of crosssection objects, $T = 40$ - number of time periods, $K = 2$ - number of variables.

The rule for degrees of freedom follows general principles of F-test and is clear from the formula above. If H_0 is not rejected there is no significant differences between elasticities.

Question 5. A researcher is considering the following alternative regression models for the expenditures on tobacco y_t (in billions of euro) subject to the relative price index of tobacco x_t using annual data (1990-2013 - **24** observations) for Germany:

$$y_t = \lambda + \mu x_t + u_t \qquad (1)$$

$$\Delta y_t = \gamma + \delta \Delta x_t + v_t \qquad (2)$$

$$y_t = \beta_1 + \beta_2 x_t + \beta_3 x_{t-1} + \beta_4 y_{t-1} + w_t \qquad (3)$$

where $\Delta y_t = y_t - y_{t-1}$, $\Delta x_t = x_t - x_{t-1}$, and u_t, v_t, and w_t are disturbance terms.

He obtained the following results estimating these equations

$$\hat{y}_t = \underset{(3.86)}{15.6} - \underset{(0.04)}{0.03} x_t \qquad R^2 = 0.03 \quad DW = 0.11 \quad RSS = 22.4 \qquad (1')$$

$$\hat{\Delta y}_t = \underset{(0.06)}{0.13} - \underset{(0.02)}{0.06} \Delta x_t \qquad R^2 = 0.29 \quad DW = 3.02 \quad RSS = 1.95 \qquad (2')$$

$$\hat{y}_t = \underset{(1.82)}{0.59} - \underset{(0.05)}{0.03} x_t + \underset{(0.05)}{0.06} x_{t-1} + \underset{(0.09)}{0.91} y_{t-1} \qquad R^2 = 0.92 \quad DW = 3.08 \quad RSS = 1.75 \qquad (3')$$

(a) (5 marks) □ Test the first equation for autocorrelation.

Solution: In the case of (1) $DW = 0.11$ indicates on the positive autocorrelation as $DW_L^{1\%}(24) = 1.04$.

□ Test the second equation for autocorrelation.

Solution: In the case of (2) as $DW = 3.02$ is greater than 2, compare $4 - 3.02 = 0.98$ with $DW_L^{1\%}(24) = 1.04$ and conclude that there is a negative autocorrelation.

□ Test the third equation for autocorrelation.

Solution: In the case of (3) lagged dependent variable is present in the list of regressors so DW-test cannot be used, use Durbin *h*-test instead:

$$h = (1 - 0.5 \cdot DW)\sqrt{\frac{n}{1 - n \cdot (s.e.(b_{Y_{t-1}}))^2}} = (1 - 0.5 \cdot 3.08)\sqrt{\frac{23}{1 - 23 \cdot (0.09)^2}} = 8.19$$

As $8.19 > 2.58 = N_{crit}^{1\%}$ we conclude that null hypothesis of no negative autocorrelation is rejected.

(b) (5 marks) To test the model (1) for autocorrelation using Common Factor Test the researcher estimated the restricted version of the equation (3) using the restriction $\beta_3 = -\beta_2 \cdot \beta_4$

$$\hat{y}_t = \underset{(1.03)}{0.54} - \underset{(0.02)}{0.05} x_t - \underset{(-)}{0.05} x_{t-1} + \underset{(0.06)}{0.92} y_{t-1} \qquad R^2 = 0.92 \quad DW = 3.00 \quad RSS = 1.81 \qquad (3R')$$

□ Show that equation (1) under assumption of the first order autocorrelation $u_t = \rho u_{t-1} + w_t$ can be transformed into equation

$$y_t = \beta_1 + \beta_2 x_t - \beta_2 \beta_4 x_{t-1} + \beta_4 y_{t-1} + w_t \qquad (3R)$$

what is restricted version of equation (3) under restriction $\beta_3 = -\beta_2 \cdot \beta_4$.

Solution: Rewrite equation $y_t = \lambda + \mu x_t + u_t$ in standard form

$$y_t = \beta_1 + \beta_2 x_t + u_t \qquad (*)$$

where $u_t = \rho u_{t-1} + w_t$, so

$$y_t = \beta_1 + \beta_2 x_t + \rho u_{t-1} + w_t \qquad (**)$$

Assuming ρ is known subtract lagged equation (*)

$$y_{t-1} = \beta_1 + \beta_2 x_{t-1} + u_{t-1}$$

multiplied by ρ from (**) to obtain

$$y_t - \rho y_{t-1} = \beta_1(1-\rho) + \beta_2(x_t - \rho x_{t-1}) + w_t$$

Or

$$y_t = \beta_1(1-\rho) + \beta_2 x_t - \beta_2 \rho x_{t-1} + \rho y_{t-1} + w_t$$

This is a special form for the model

$$y_t = \beta_1 + \beta_2 x_t + \beta_3 x_{t-1} + \beta_4 y_{t-1} + w_t$$

under nonlinear restriction $\beta_3 = -\beta_2 \beta_4$

□ Explain the logic of Common Factor Test? Do Common Factor test, giving test statistic and decision rule.

Equation (3R') is obtained under assumption of the first order autocorrelation $u_t = \rho u_{t-1} + w_t$ and is restricted version of equation (3'). So if in comparison of equations (3R') and (3') the restriction will not be rejected the assumption of the first order autocorrelation $u_t = \rho u_{t-1} + w_t$ also will not be rejected. Test statistic is $\chi^2 = n \cdot \ln\left(\frac{RSS_R}{RSS_U}\right)$ has χ^2-distribution with 1 degree of freedom (one restriction in our case).

$$\chi^2 = 23 \cdot \ln\left(\frac{1.81}{1.75}\right) = 0.77 < 3.84 = \chi^2_{crit}(1, 5\%)$$

□ What is your conclusion? What is interpretation of the coefficinent β_4 of the variable y_{t-1} in (3R)?

Solution: The restriction $\beta_3 = -\beta_2 \cdot \beta_4$ and so the assumption of the first order autocorrelation $u_t = \rho u_{t-1} + w_t$ is not rejected and so we choose restricted version (3R'). The value 0.92 corresponding coefficient $\beta_4 = \rho$ can be interpreted as autocorrelation coefficient.

(c) (5 marks) Let (1) is the correct specification and w_t satisfies the Gauss-Markov conditions. Taking into account your conclusions in a)

□ Explain what problems, if any, would be encountered if ordinary least squares were used to fit (1) ?

Solution: Under positive autocorrelation detected in **a)** estimators of coefficients of (1) will be unbiased, consistent but inefficient.

□ Explain what problems, if any, would be encountered if ordinary least squares were used to fit (2) ?

Solution: Almost the same in previous question on equation (1). But now additional MA(1) type autocorelation emerges as disturbance term v_t will be now of the type $u_t - u_{t-1}$. As we could see from (a) the original model suffered from strong positive autocorrelation, so negative MA(1) autocorrelation can help to cope with it. On the other hand, the difference can be significantly lower than the absolute values, so the disturbance term becomes relatively large, which leads to a decrease in the accuracy of estimations.
Bonus answer. The probable non-stationarity of the initial series should also be taken into account.

□ Explain what problems, if any, would be encountered if ordinary least squares were used to fit (3) ?
Можно сюда вставить теоретический вывол

Solution: □ In (3) lagged variable y_{t-1} is present in the right side of equation so under negative autocorrelation detected in (a) OLS estimates will be inconsistent.

(d) (5 marks) □ Show that (1) and (2) are restricted versions of (3), what are the restrictions?

Solution: In the case of (1), the restrictions are $\beta_3 = \beta_4 = 0$.

Eq. (2) $\Delta y_t = \gamma + \delta \Delta x_t + v_t$ or $y_t - y_{t-1} = \gamma + \delta(x_t - x_{t-1}) + v_t$ and $y_t = \gamma + \delta x_t - \delta x_{t-1} - y_{t-1} + v_t$ so

So in the case of (2), the restrictions are $\beta_4 = -1$ and $\beta_3 = -\beta_2$.

□ Test these restrictions? Which of these tests could be done using both R^2 and RSS and which can be done using only one of these indicators?

Solution: In the case of (1) both tests give correct results $F = \dfrac{(0.92 - 0.03)/2}{(1 - 0.92)/(23 - 4)} = 106$ and $F = \dfrac{(22.4 - 1.75)/2}{1.75/(23 - 4)} = 112$ - significant as $F_{crit}^{1\%}(2{,}19) = 5.93$.

In the case of (2) only test based on RSS is correct $F = \dfrac{(1.95 - 1.76)/2}{1.76/(23 - 4)} = 1.1$ - insignificant as $F_{crit}^{5\%}(2{,}19) = 3.52$

Question 6. A researcher is considering the following alternative regression models for the expenditures on tobacco y_t (in billions of euro) subject to the relative price index of tobacco x_t using annual data (1990-2013 - 24 observations) for Germany:

$$y_t = \lambda + \mu x_t + w_t \tag{1}$$

$$\Delta y_t = \gamma + \delta \Delta x_t + v_t \tag{2}$$

where $\Delta y_t = y_t - y_{t-1}$, $\Delta x_t = x_t - x_{t-1}$, and u_t, v_t, and w_t are disturbance terms.
He got the following results of the estimation of these equations

$$\hat{y}_t = 15.6 - 0.03 x_t \qquad R^2 = 0.03$$
$$(3.86)\ (0.04) \quad DW = 0.11 \quad RSS = 22.4 \tag{1'}$$

$$\Delta\hat{y}_t = 0.13 - 0.06 \Delta x_t \qquad R^2 = 0.29$$
$$(0.06)\ (0.02) \quad DW = 3.02 \quad RSS = 1.95 \tag{2'}$$

(a) (5 marks) The colleague advised the researcher to check the data series for stationarity by telling him about the computer (Monte Carlo) experiments of Newbold-Granger, and said that his estimated regression (1') is spurious regression.

□ What is covariance stationarity, why is it important in econometrics?

Solution: A process $\{y_t\}$ is (covariance) stationary if its mean and variance exist and do not depend on time and its covariance is a function of distance in time only (not location). Violation of either of these requirements renders the process non-stationary.

□ What is the Monte Carlo experiment? What are Newbold-Granger's experiments, what are their most important results?

Monte-Carlo experiments allow you to use a computer to repeatedly reproduce a number of conditions and then analyze the results of their consequences for the evaluation of econometric models.
Spurious regression was first demonstrated by Granger and Newbold who showed, using Monte Carlo techniques, that a regression involving two non-stationary series could give rise to spurious results in that the t-statistics over-rejected the null hypothesis of a zero coefficient for two independent random walk series.

The researcher argued that the equation (1') is insignificant, so the conclusions of the experiments of Newbold-Granger cannot be applied to it, and therefore, regression (1') is not spurious.

□ Who is right in assessing equation (1') with respect to the property 'spurious' – the researcher or his colleague?

Solution: Equation (1') is really insignificant. But it doesn't mean anything and doesn't have any evidencing power, because spurious regression can be both significant and insignificant. In order to be able to talk about spurious regression it is necessary to perform tests for stationarity of time series used in it.

(b) (5 marks) The researcher decided to test the data series for stationarity by constructing the following Dickey-Fuller equations

$$\hat{\Delta y}_t = \underset{(2.17)}{5.48} - \underset{(0.20)}{0.50} y_{t-1} + \underset{(0.03)}{0.07} t \qquad R^2 = 0.25 \qquad (3')$$

$$\hat{\Delta x}_t = \underset{(8.09)}{5.48} - \underset{(0.09)}{0.05} x_{t-1} + \underset{(0.25)}{0.46} \Delta x_{t-1} - \underset{(0.08)}{0.08} t \qquad R^2 = 0.43 \qquad (5')$$

$$\hat{\Delta^2 y}_t = -\underset{(0.23)}{1.15} \Delta y_{t-1} \qquad R^2 = 0.58 \qquad (4')$$

$$\hat{\Delta^2 x}_t = -\underset{(0.18)}{2.16} \Delta x_{t-1} \qquad R^2 = 0.19 \qquad (6')$$

□ Interpret the results of the evaluation of Dickey-Fuller equations. Which of the time series are stationary and which are non-stationary?

Solution: Using t-ADF table.

For equation (3') $t_{y_{t-1}} = \dfrac{b}{s.e.(b)} = \dfrac{-0.50}{0.2} = -2.5 > -3.62 = t^{5\%}_{ADF,\,crit}(23,\text{ trend in model})$ - expenditures on tobacco y_t is non-stationary, there is also significant trend.

For equation (4') $t_{\Delta y_{t-1}} = \dfrac{b}{s.e.(b)} = \dfrac{-1.15}{0.23} = -5 < -3.77 = t^{1\%}_{ADF,\,crit}(22,\text{ no trend})$ - their differences Δy_t are stationary.

Equation (5') uses additional lag and shows that prices x_t is obviously non-stationary but has no trend.

t-statistic in equation (6') indicates that Δx_t is stationary.

□ What is the difference between the specification of equation (3') and the specification of equation (5'), does each of these specifications have any advantages?

Solution: The test is sensitive to the presence of serial correlation in the error term so we need to take steps to remove the effects of this serial correlation - this is done by including lagged values of y_t in the regression. On the other hand equation (5') has reduced number of degrees of freedom, so the estimators become less efficient.

□ What does "best specification" in the Dickey-Fuller equation mean? On what basis can one choose the best specification of the equation?

Solution: The optimum lag length could be chosen on the base of some information criteria (Akaike or Schwarz). They balance the benefits of additional lags and the resulting loss of degrees of freedom.

(c) (5 marks) Explore Dickey-Fuller equation models (3) and (5) mathematically. Explain the logic of Dickey-Fuller test.

□ Derive theoretical model corresponding time series model (3'). Give null and alternative hypothesis and decision rule.

Solution: The standard test for a unit root is due to Dickey and Fuller and is based on the model $y_t = \beta_1 + \beta_2 y_{t-1} + \gamma t + u_t$ which can be re-written as $\Delta y_t = \beta_1 + (1-\beta_2) y_{t-1} + \gamma t + u_t$ where $\Delta y_t = y_t - y_{t-1}$. The null hypothesis for stationarity is $H_0 : \beta_2 - 1 = 0$, $H_A : \beta_2 - 1 < 0$. We cannot use the standard t-test procedure in this case because the distribution of the t-statistic is not a t-distribution so critical values have been computed by Dickey and Fuller using Monte Carlo techniques.

□ Derive theoretical Dickey-Fuller equation model corresponding time series model (5').

Solution: we start now from the model $y_t = \beta_1 + \beta_2 y_{t-1} + \beta_3 y_{t-2} + \gamma t + u_t$ to control for AR(1) serial correlation. This is more easily tested by using the model $\Delta y_t = \beta_1 + (1-\beta_2-\beta_3) y_{t-1} - \beta_3 \Delta y_{t-2} + \gamma t + u_t$ with null hypothesis $H_0 : 1 - \beta_1 - \beta_2 = 0$. Once again using Dickey-Fuller tables.

(d) (5 marks) The researcher noticed that the chaotic dynamics of relative tobacco prices reminds both random walk and negative autocorrelation, and decided to use for modeling an equation

$$x_t = x_{t-1} - \beta x_{t-2} + u_t, \quad 0 < \beta < 1 \qquad (7),$$

where u_t has zero mean $Eu_t = 0$, constant variance $\operatorname{var} u_t = \sigma^2$ and not correlated with u_{t-k} and x_{t-k}, $k > 0$.

□ Help the researcher to decide whether this time series is stationary or non-stationary

Hint: Use difference Dickey-Fuller equation to explore time series.

Solution: (a) The series $x_t = x_{t-1} - \beta x_{t-2} + u_t$, $0 < \beta < 1$, can be considered as a special case of the model $x_t = \alpha x_{t-1} + \beta x_{t-2} + u_t$, where $\alpha = 1$. To apply Dickey-Fuller approach to this series one should rearrange terms like this

$x_t - x_{t-1} = \alpha x_{t-1} - x_{t-1} + \beta x_{t-2} + u_t$,

$x_t - x_{t-1} = \alpha x_{t-1} - x_{t-1} + \beta x_{t-1} - \beta x_{t-1} + \beta x_{t-2} + u_t$

and

$\Delta x_t = (\alpha + \beta - 1) x_{t-1} - \beta \Delta x_{t-1} + u_t$

Now substitute $\alpha = 1$

$\Delta x_t = (1 - \beta - 1) x_{t-1} - \beta \Delta x_{t-1} + u_t = -\beta x_{t-1} + \beta \Delta x_{t-1} + u_t$

As $0 < \beta < 1$ the time series will be classified as stationary (necessary condition of stationarity holds).

Note: in fact this time series is asimptotically stationary, but for simplicity we will say further that it is approximately stationary.

Question 7.

A researcher is considering the following alternative regression models for the expenditures on tobacco y_t (in billions of dollars) subject to the relative price index of tobacco x_t using annual data (1971-2015 - **45** observations) this time for USA:

$$y_t = \lambda + \mu x_t + w_t \quad (1)$$

$$\Delta y_t = \gamma + \delta \Delta x_t + v_t \quad (2)$$

where $\Delta y_t = y_t - y_{t-1}$, $\Delta x_t = x_t - x_{t-1}$, and u_t, v_t, and w_t are disturbance terms.

He got the following results of the estimation of these equations

$$\hat{y}_t = \underset{(2.59)}{114.1} - \underset{(0.04)}{0.38} x_t \qquad R^2 = 0.61 \quad DW = 0.16 \quad RSS = 1781.3 \quad (1')$$

$$\hat{\Delta y}_t = \underset{(0.43)}{0.29} - \underset{(0.11)}{0.36} \Delta x_t \qquad R^2 = 0.21 \quad DW = 2.27 \quad RSS = 281.9 \quad (2')$$

(a) (5 marks) You are given that the time series y_t and x_t are difference stationary, or they are integrated of order 1: $y_t \in I(1)$, $x_t \in I(1)$.

□ Explain what does it mean that time series is difference stationary and trend stationary.

Solution: If after removing the trend from a nonstationary series (detrending) the resulting variable becomes stationary, then the variable is called *trend stationary*.
If a nonstationary process can be transformed into a stationary process by differencing then the series is said to be *difference stationary*.

□ Give the examples of difference stationary and trend stationary time series. Show that the time series you have given posess the specified properties.

Solution: Let: $X_t = \alpha_0 + \alpha_1 t + u_t$, where $\mathrm{E}(u_t) = 0$, $\mathrm{var}(u_t) = \sigma^2$ and $\mathrm{E}(u_t u_{t-s}) = 0$ for all *s* and *t*. Let: $Z_t = X_t - \alpha_1 t = \alpha_0 + u_t$ (detrending).

$$\mathrm{E}(Z_t) = \mathrm{E}(\alpha_0 + u_t) = \alpha_0,$$

$$\mathrm{var}(Z_t) = \mathrm{Var}(\alpha_0 + u_t) = \sigma^2,$$

$$\mathrm{cov}(Z_t, Z_{t-s}) = \mathrm{E}[(Z_t - \mathrm{E}(Z_t))(Z_{t-s} - \mathrm{E}(Z_{t-s}))] = \mathrm{E}(u_t u_{t-s}) = 0.$$

This means that Z_t has constant mean and variance for all *t*, and covariance is zero for all $s > 0$. It implies that the series is trend stationary.
Bonus answer: Alternatively for practical purposes the time variable could be included into the regression model.
Solution: Let X_t be a random walk with a drift: $X_t = \beta_0 + X_{t-1} + \varepsilon_t$, where $\mathrm{E}(\varepsilon_t) = 0$, $\mathrm{var}(\varepsilon_t) = \sigma^2$ and $\mathrm{E}(\varepsilon_t \varepsilon_s) = 0$ for all *s* and *t*, $s \neq t$. Subtract X_{t-1} from both sides of equation to get: $\Delta X_t = X_t - X_{t-1} = \beta_0 + \varepsilon_t$.
Checking the same properties we conclude that ΔX_t is stationary. This implies that X_t is difference stationary.
Bonus answer: The trend stationary time series instead of detrending can be differenced.

□ What is the main difference between two types of time series under consideration.

Solution: It is important to know whether a variable is difference stationary or trend stationary because for difference stationary variables shocks have a permanent effect whereas for trend stationary variables shocks are transitory.

(b) (5 marks) □ What is cointegration? How to test whether time series y_t and x_t are cointegrated (you are given again that $y_t \in I(1)$, $x_t \in I(1)$.

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 such that it is stationary

The second condition is equivalent to the following: the residuals of regression of one of them on others are stationary. So to test time series y_t and x_t for cointegration is is sufficient to do the following:

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

2) Run regression estimated by OLS: $y_t = \alpha_0 + \alpha_1 x_t + u_t$ 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.

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

□ What are relative advantages and disadvantages of the models (1) and (2)?

Let z_t are residuals of equation (1'). To test them for stationarity the researcher runs the regression

$$\hat{\Delta z}_t = -0.118 z_{t-1} \qquad R^2 = 0.07$$
$$(0.058)$$

□ What is the meaning of this equation? What is your conclusion from the test (**use the following critical values for ADF t-statistic** $t_{ADF}^{1\%} = -2.62$, $t_{ADF}^{5\%} = -1.95$, $t_{ADF}^{10\%} = -1.61$)?

Solution: This is Dickey-Fuller equation of the type $\Delta z_t = (\beta - 1) z_{t-1} + u_t$ to test $H_0 : \beta = 1$ against $H_a : \beta < 1$ Evaluate $t = \dfrac{b}{s.e.(b)} = \dfrac{-1.118}{0.058} = -2.03$. As $-2.62 < t = -2.03 < -1.94$ null hypothesis of non-stationarity is rejected at 5%. It means that y_t and x_t are cointegrated.

(c) (5 marks) Consider a simple $ADL(1,1)$ model

$$y_t = \alpha_1 + \alpha_2 y_{t-1} + \alpha_3 x_t + \alpha_4 x_{t-1} + u_t \tag{3}$$

where y_t and x_t are $I(1)$.

□ Derive error correction model based on (3).

Solution: One of the possible solutions is following. Rewrite (3) as

$$y_t - y_{t-1} = \alpha_1 + \alpha_2 y_{t-1} - y_{t-1} + \alpha_3 x_t - \alpha_3 x_{t-1} + \alpha_3 x_{t-1} + \alpha_4 x_{t-1} + u_t$$

$$\Delta y_t = \alpha_1 - (1-\alpha_2) y_{t-1} + \alpha_3 \Delta x_t + (\alpha_3 + \alpha_4) x_{t-1} + u_t$$

$$\Delta y_t = \alpha_3 \Delta x_t - (1-\alpha_2)\left[y_{t-1} - \frac{\alpha_1}{(1-\alpha_2)} - \frac{(\alpha_3+\alpha_4)}{(1-\alpha_2)} x_{t-1} \right] + u_t$$

$$\Delta y_t = \alpha_3 \Delta x_t - (1-\alpha_2)[y_{t-1} - \beta_1 - \beta_2 x_{t-1}] + u_t$$

or,
$$\Delta y_t = \alpha_3 \Delta x_t - \pi[y_{t-1} - \beta_1 - \beta_2 x_{t-1}] + u_t$$

where
$$\pi = (1-\alpha_2); \;\; \beta_1 = \frac{\alpha_1}{(1-\alpha_2)} \text{ and } \beta_2 = \frac{(\alpha_3+\alpha_4)}{(1-\alpha_2)}.$$

This equation is the Error Correction Model.

The point of rearranging the ADL(1,1) model in this way is that, although Y_t and X_t are both I(1), all of the terms in the regression equation are I(0) and hence the model may be fitted using least squares in the standard way.
Bonus answer: Of course, the β parameters are not known and the cointegrating term is unobservable. One way of overcoming this problem, known as the Engle–Granger two-step procedure, is to use the values of the parameters estimated in the cointegrating regression to compute the cointegrating term. It can be demonstrated that the estimators of the coefficients of the fitted equation will have the same properties asymptotically as if the true values had been used.

(d) (5 marks) Let model (4') is the result of estimation error correction model, obtained in (c) for years 1960-1996

$$\hat{\Delta y_t} = -0.795\Delta x_t - 0.128 Z_{t-1} \qquad R^2 = 0.32 \qquad (4')$$
$$(0.217) \qquad (0.061)$$

□ Give interpretation to the estimated model. Explain how error correction mechanism works.

Solution: The first term corresponds to the difference model (short run behavior), the second term is the error-correction mechanism (describing long run behaviour), both coefficients are significant. The size of the adjustment is proportional to the discrepancy. The effect of this term is to reduce the discrepancy between y_{t-1} and its cointegrating level. The error-correction mechanism allows you to overcome the short-sightedness of the model. At the same time, the correction value is relatively small (-0.128). It means that each year only 12.8 of discrepancy is covered by an adjustment using a cointegration relationship.

Let variable yf_t be forecast obtained using by error correction model and ydf_t be forecast obtained using difference model (2).

□ Comment the following graph, demonstrating predictive power of different models

Solution: Tobacco expenditure has a complicated dynamics, with a rise at the beginning of the period followed by a decline in subsequent years. However, the difference model generally provides a more accurate forecast, while the error correction model overestimates the impact of relative prices on tobacco use decline. This may be due to the poor quality of the cointegration relationship: it uses a linear model, while the dynamics of tobacco expenditure is non-linear.