The International College of Economics and Finance
Econometrics – 2019-2019.
Final exam 2019 May 24.

General instructions. Candidates should answer 6 of the following 7 questions: all questions of the Section A and any two of the questions from Section B (questions 5-7). The weight of the Section A is 60% of the exam; two 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 always state clearly null and alternative hypotheses provide critical value used for test, mentioning degrees of freedom and the significance level chosen for the test.

SECTION A

Answer **ALL** questions from this section (questions **1-4**).
НОВЫЕ ВОПРОСЫ

Question 1.

The researcher obtains data on household annual expenditure on science-fiction books *B*, and annual household income, *Y*, for 100 Russian citizens in 2018. He hypothesizes that *B* is positively related to *Y* and also positively related to the average years of schooling of the citizens, *S*, by the relationship

$$\log B = \beta_1 + \beta_2 \log Y + \beta_3 \log S + u \qquad \text{(A)}$$

where *u* is a disturbance term that satisfies the Gauss–Markov conditions. He also considers the possibility that log *B* may be determined by log *Y* alone:

$$\log B = \alpha_1 + \alpha_2 \log Y + u \qquad \text{(B)}$$

It may be assumed that *Y* and *S* are both nonstochastic. In the sample the correlation between log *Y* and log *S* is 0.86. He performs the following regressions: (1) log *B* on both log *Y* and log *S*, and (2) log *B* on log *Y* only, with the results shown below (standard errors in parentheses):

$$\hat{\log} B = -6.89 + 1.10 \log Y + 0.59 \log S \quad R^2 = 0.29 \qquad (1)$$
$$\qquad (2.28) \quad (0.65) \qquad (0.35)$$

$$\hat{\log} B = -3.37 + 2.10 \log Y \quad R^2 = 0.27 \qquad (2)$$
$$\qquad (0.89) \quad (0.35)$$

(a) □ Give interpretations to the coefficients of regressions (1) and (2).
□ Test the significance of the coefficients (both taken alone and together for the multiple regression).
□ Regression (2) is restricted version of the regression (1). What is the restriction? Test the restriction.

Slope coefficient 2.10 in regression (2) shows income elasticity of the expenditures on books (obviously overestimated). Its counterpart in the equation (1) 1.10 shows income elasticity of the expenditures on books under assumption that $\log S$ is kept constant. The coefficient 0.59 shows schooling elasticity of the expenditures on books keeping income constant.
Slope coefficient 2.10 in (2) is significant at any significance level.
Slope coefficients 1.10 and 0.95 in (2) both are significant at 5% significance level ($t_{\log(Y)} = \dfrac{1.10}{0.65} = 1.69 > 1.66$ and $t_{\log(S)} = \dfrac{0.59}{0.35} = 1.68 > 1.66 = t_{crit}^{5\%,one-sided}(100)$) as we are told that we are entitled to do one sided tests. Both coefficients taken together are significant as $F = \dfrac{0.29/2}{(1-0.29)/(100-3)} = 19.81 > 4.82 = F_{crit.}^{1\%}(2,100)$.
The restriction is $H_0 : \beta_3 = 0$ against $H_0 : \beta_3 > 0$, it is rejected as it was shown above.

(b) □ Assuming that (A_1) is the correct specification, explain, using covariance and variance properties, whether you would expect the coefficient of log Y to be greater in regression (2) than in (1).

□ Based on the data available and using formula for the omitted variable bias evaluate the ratio of $\dfrac{\text{Var}(\log S)}{\text{Var}(\log Y)}$.

Let a_2 be the OLS estimate of α_2 in equation (2).

$$\log B = \alpha_1 + \alpha_2 \log Y + u \qquad \text{(B)}$$

The formula for the omitted variable bias is

$$E(a_2) = \beta_2 + \beta_3 \frac{\text{Cov}(\log S, \log Y)}{\text{Var}(\log Y)}$$

It is said that sample correlation between log Y and log S is 0.86, so it is positive. The sign of the coefficient β_3^1 is not known but in equation (1) its estimate 0.59 is positive and significant at 5% level as it was shown in (a), so the null $H_0: \beta_3 \le 0$ can be rejected at 5% level. So the assumption $\beta_3 > 0$ does not contradict the data, using this we may conclude that bias $\beta_3 \dfrac{\text{Cov}(\log S, \log Y)}{\text{Var}(\log Y)}$ is positive. So we would expect an upward bias in regression (2).

$$E(a_2) = \beta_2 + \beta_3 \frac{\text{Cov}(\log S, \log Y)}{\text{Var}(\log Y)} = \beta_2 + \beta_3 \cdot corr(\log Y, \log S) \cdot \sqrt{\frac{\text{Var}(\log S)}{\text{Var}(\log Y)}}.$$

For the estimates, $2.10 = 1.10 + 0.59 \cdot 0.86 \cdot \sqrt{\dfrac{\text{Var}(\log S)}{\text{Var}(\log Y)}}$, so $\sqrt{\dfrac{\text{Var}(\log S)}{\text{Var}(\log Y)}} \approx 4$ and so $\dfrac{\text{Var}(\log S)}{\text{Var}(\log Y)} \approx 16$.

(c) □ When sending the estimation results to a referee, the researcher forgot to include the determination coefficients of both regressions. Nevertheless, the referee was able to find out the values of R^2 for both equations using other data from equations above. How could he do this?

From (2): *t*-statistic for slope coefficient equals 6; so *F*-statistic equals 6^2=36. $F = R^2(n-k_2)/(1-R^2) = 36; (n-k_2) = 98$, so $R^2 \approx 0.27$.
Then, F-statistic for adding log S as an explanatory variable also equals to the squared t-statistic of its coefficient: $F(1;97) = (R_1^2 - R_2^2)(n-k_1)/(1-R_1^2)$. From (1): *t*-statistic for slope coefficient of log S equals 1.69; so *F*-statistic equals $1.69^2 = 2.84$. Hence, $F(1;97) = 2.84 = (R_1^2 - 0.27) \cdot 97/(1-R_1^2)$, so $R_1^2 \approx 0.29$.

Question 2.

1. Consider the linear regression model

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 Y_{t-1} + u_t, \; t = 1, ..., T,$$

where the errors u_t are distributed independently of the regressors X_t and $|\beta_2| < 1$.

(a) □ The student suspect that the mean zero errors exhibit autocorrelation, so he decided to use Durbin-Watson test to check his suspicions. Comment on student ideas.

□ What tests would you suggest in case of the first order autocorrelation? Describe briefly the logic of Durbin h-test and its application? What are its relative advantages and disadvantages?

The use of the Durbin-Watson test is incorrect as lagged explanatory variable is included into equation. In addition Durbin-Watson test is only applicable to detection of the autocorrelation of the first order.

Durbin test is based on the *h*-statistics: $h = \hat{\rho} \cdot \sqrt{\dfrac{n}{1 - n \cdot S_{\beta_3}^{\;2}}}$. Here $\hat{\rho}$ is the supposed autocorrelation coefficient (in equation $u_t = \rho u_{t-1} + \varepsilon_t$) and could be approximately estimated using usual Durbin-Watson d-statistics: $\hat{\rho} = (1 - 0.5d)$. $S_{\beta_3}^{\;2}$ in the formula for *h*-statistics means the square of standard error for coefficient of lagged

dependent variable Y_{t-1}, and number of observations is evaluated taking into account that using lagged variables diminiches total number of observation (by one in our case). Under assumption of no serial correlation *h*-statistics follows asymptotically the standard normal *z*-distribution, so usual critical value could be used (1.96 for significance level 0.05, for example). Quite a rare problem with this statistics is that expression under square root $1-n\cdot S_{\beta_3}{}^2$ could be negative (then usual d-statistics could be studied with some reservations or Breush-Godfrey test).
Unlike conventional Durbin-Watson test this is large sample test.

(b) □ What test would you suggest in case if the order of autocorrelation is unknown? Describe briefly the logic of this test? What are its relative advantages and disadvantages?

To detect autocorrelation of any specified order the Breusch-Godfrey test can be proposed. In case of the autocorrelation of the first order $u_t = \rho u_{t-1} + \varepsilon_t$ it is enough to run original regression, memorize its residuals $\hat{u}_t$ and then run the following auxiliary regression:

$$\hat{u}_t = \gamma_0 + \gamma_1 X_t + \gamma_2 Y_{t-1} + \gamma_3 \hat{u}_{t-1} + v_t .$$

Then a pair of hypotheses $H_0 : \gamma_3 = 0$, $H_a : \gamma_3 \neq 0$ ahould be tested using conventional t-test.
For the second order autocorrelation $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t$ the auxiliary equation will be the following

$$\hat{u}_t = \gamma_0 + \gamma_1 X_t + \gamma_2 Y_{t-1} + \gamma_3 \hat{u}_{t-1} + \gamma_4 \hat{u}_{t-2} + v_t .$$

F test should be used to evaluate the joint significance of the coefficients on the lagged residuals $H_0 : \gamma_3 = \gamma_4 = 0$.
Alternatively Lagrange multiplier statistic can be used: nR^2 is distributed as $\chi^2(1)$ when testing for first-order autocorrelation, and nR^2 is distributed as $\chi^2(2)$ when testing for the second-order autocorrelation, etc.

Main advantages of this test
1. No limitation to the structure of the model (it is possible to use it with for regression with lagged explanatory variables).
2. Exact critical values (no uncertainty zone)
3. Always applicable (no need to extract square root)
4. It is possible to investigate also autocorrelation of the higher order: second, third and so on… (see below).
5. This test remains valid also for autocorrelation caused by moving average processes
6. It can be performed automatically in EViews 8.0 (see below)

Main disadvantage of this test
Only large sample allowed

(c) □ Assume that u_t follows an $AR(1)$ process. Discuss, for the given model, the consequences for the ordinary least squares estimator.

□ Derive the formula for $\mathrm{Cov}(Y_{t-1}, u_t) = \dfrac{\rho\sigma_u^2}{1-\beta_2\rho}$ assuming covariance stationarity of the time series Y_t.

The model is

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 Y_{t-1} + u_t ,\ t = 1, \ldots, T ,$$

$AR(1)$ process is $u_t = \rho u_{t-1} + \varepsilon_t$ so wi can see that in the presence of the lagged endogenous variable Y_{t-1} this variable correlates with $u_t = \rho u_{t-1} + \varepsilon_t$: $\mathrm{Cov}(Y_{t-1}, u_t) \neq 0$. This yields inconsistency of OLS estimators.

Now evaluate $\mathrm{Cov}(Y_{t-1}, u_t)$,

$$\mathrm{Cov}(Y_{t-1}, u_t) = \mathrm{Cov}(\beta_0 + \beta_1 X_{t-1} + \beta_2 Y_{t-2} + u_{t-1},\ \rho u_{t-1} + \varepsilon_t) =$$

$$= \beta_1 \rho \mathrm{Cov}(X_{t-1}, u_{t-1}) + \beta_2 \rho \mathrm{Cov}(Y_{t-2}, u_{t-1}) + \rho \mathrm{Cov}(u_{t-1}, u_{t-1}) + \mathrm{Cov}(\beta_0 + \beta_1 X_{t-1} + \beta_2 Y_{t-2} + u_{t-1}, \varepsilon_t)$$

Since u_t is uncorrelated with X_{t-1}, Y_{t-2} and u_{t-1} and since the errors u_t are distributed independently of the regressors, only two terms are not zero

$$\mathrm{Cov}(Y_{t-1}, u_t) = \beta_2 \rho \mathrm{Cov}(Y_{t-2}, u_{t-1}) + \rho \mathrm{Cov}(u_{t-1}, u_{t-1}).$$

Using the fact that $\mathrm{Cov}(Y_{t-2}, u_{t-1}) = \mathrm{Cov}(Y_{t-1}, u_t)$ by covariance stationarity we get:

$$(1 - \beta_2 \rho) \mathrm{Cov}(Y_{t-1}, u_t) = \rho \sigma_u^2.$$

And finally

$$\mathrm{Cov}(Y_{t-1}, u_t) = \frac{\rho \sigma_u^2}{1 - \beta_2 \rho}$$

Question 3.

For the population of men who grew up with disadvantaged backgrounds, let *poverty* be a dummy variable equal to one if a man is currently living below the poverty line, and zero otherwise. The variable *age* is age and *educ* is total years of schooling. Let *vocat* be an indicator equal to unity if a man's high school offered vocational training. Using a random sample of 850 men, you obtain

$$\widehat{\Pr(poverty = 1 | educ, age, vocat)} = \Lambda(0.453 - 0.016 age - 0.087 educ - 0.049 vocat),$$

where $\Lambda(z) = \dfrac{1}{1 + \exp(-z)} = \dfrac{\exp(z)}{1 + \exp(z)}$ is the logit function.

(a) It is argued that using the logit regression model is better than using the linear probability model when explaining the binary variable *poverty*.

□ Discuss the benefits/drawback of using the logit regression model when trying to explain a binary variable.

Solution:

The logit model has two main advantages over the linear probability model (LPM): predicted probabilities are restricted to lie in $[0, 1]$ and MLE is (asymptotically) efficient whereas OLS (LPM) will be inefficient given the inherent presence of heteroskedasticity.. Additionally the marginal effects of factors are not constant in logit model unlike LPM where they do not depend on the point chosen.

The main drawbacks of the logit model relative to the linear probability model are that the coefficients cannot be directly interpreted as the marginal effects of the regressor(s) of interest and it is also computationally more complicated. The estimations of marginal effects obtained by the direct calculation and by the derivative method can be slightly different.

(b) For a 40-year old man, with 12 years of education, what is the estimated effect of having vocational training available in high school on the probability of currently living in poverty?

□ Use direct comparison of the two values of logit functions to evaluate this effect.

The logistic function has the following expression

$$\Lambda(z) = \frac{1}{1 + \exp(-z)} = \frac{\exp(z)}{1 + \exp(z)}.$$

where $z = 0.453 - 0.016 age - 0.087 educ - 0.049 vocat$.

For a 40-year old man with 12 years of education with vocational training the estimated probability of living in poverty is given by:

$$\widehat{\Pr(y_i = 1 | age_i = 40, educ_i = 12, vocat_i = 1)} = \frac{1}{1 + \exp(-z_1)} \approx 0.218,$$

where $z_1 = 0.453 - 0.016 \times 40 - 0.087 \times 12 - 0.049 \times 1 \approx -1.28$.

The estimated probability of living in poverty for the same man without the vocational training is given by:

$$\widehat{\Pr(y_i = 1 | age_i = 40, educ_i = 12, vocat_i = 0)} = \frac{\exp(z_2)}{1 + \exp(z_2)} \approx 0.226,$$

where $z_2 = 0.453 - 0.016\times 40 - 0.087\times 12 - 0.049\times 0 \approx -1.23$.
the difference between the two probabilities is $0.218 - 0.226 = -0.008$. Therefore, for a 40-year old man with 12 years of education having vocational training in high school decreases the probability of living in poverty by 0.8 percentage points.

(c) □ Now do the same estimation of marginal effect of vocational education as in (b) using derivatives.
□ Compare it with the maximum possible marginal effect.

The derivative for the logistic function is $\dfrac{\partial\Lambda(z)}{\partial z} = \dfrac{d}{dz}\left(\dfrac{1}{1+\exp(-z)}\right) = \dfrac{\exp(-z)}{(1+\exp(-z))^2} = \dfrac{\exp(z)}{(1+\exp(z))^2}$.
For a 40-year old man with 12 years of education without vocational training the estimated marginal effect of training is

$$\frac{\partial\Lambda(z)}{\partial E} = \frac{\partial\Lambda(z)}{\partial z}\cdot\frac{dz}{dE} = \frac{\exp(-z)}{(1+\exp(-z))^2}\cdot\beta_E = \frac{\exp(-z_2)}{(1+\exp(-z_2))^2}\cdot(-0.049) = 0.175\cdot(-0.049) = -0.0086,$$

where $z_2 = 0.453 - 0.016\times 40 - 0.087\times 12 - 0.049\times 0 \approx -1.23$.
Therefore, for a 40-year old man with 12 years of education having vocational training in high school decreases the probability of living in poverty by 0.86 percentage points, the result is close to that obtained in (b).
The maximum possible effect is achieved at the point $z = 0$:

$$\left.\frac{\partial\Lambda(z)}{\partial E}\right|_{\max} = \frac{\exp(0)}{(1+\exp(0))^2}\cdot(-0.049) = 0.25\cdot(-0.049) = -0.012$$

The difference is considerable: the current marginal effect is 71.7% of maximum

Question 4.
The researcher investigate US jewelry market using simple dynamic models, describing total expenditure on jewelry J_t in USA for 1993-2017 (annual data) and the influence of the prices P_t (nominal price index) on jewelry. First explanation of market dynamics is based on the adaptive expectation model

$$J_t = \beta_1 + \beta_2 P^e_{t+1} + u_t \qquad \textbf{(AE)},$$

where P^e_{t+1} stands for expectations of the value of P for the future period $t+1$, the dynamics of that is described by the following adaptive expectations equation $P^e_{t+1} - P^e_t = \lambda(P_t - P^e_t)$.
Second explanation of market dynamics uses partial adjustment model

$$J^*_t = \beta_1 + \beta_2 P_t + u_t \qquad \textbf{(PA)}.$$

where J^*_t represents 'target' or 'desired' level of expenditure on jewlery, while actual level J_t is subject the adjustment process $J_t - J_{t-1} = \lambda(J^*_t - J_{t-1})$.

To estimate these dynamic models she started from the regression of the ADL(1,0) type

$$\hat{J}_t = \underset{(0.18)}{0.36} - \underset{(0.003)}{0.003}P_t + \underset{(0.060)}{1.052}J_{t-1} \qquad R^2 = 0.980 \quad DW = 1.81 \quad RSS = 1.27 \qquad (1)$$

After additionally added lagged prices the equation of the type ADL(1,1) looks as

$$\hat{J}_t = \underset{(0.159)}{0.204} - \underset{(0.007)}{0.023}P_t + \underset{(0.006)}{0.02}P_{t-1} + \underset{(0.051)}{1.083}J_{t-1} \qquad R^2 = 0.987 \quad DW = 2.19 \quad RSS = 0.842 \qquad (2)$$

She also estimated a restricted version of the same regression

$$\hat{J}_t = \underset{(0.191)}{0.308} - \underset{(0.001)}{0.0009}P_t + 0.0009\cdot 1.022 P_{t-1} + \underset{(0.059)}{1.022}J_{t-1} \qquad R^2 = 0.979 \quad DW = 1.75 \quad RSS = 1.332 \qquad (3)$$

(a) □ How the parameters of the partial adjustment model (PA) could be evaluated on the basis of the estimation results (1-3).
□ Find estimates of parameters β_1 β_2 and λ.

□ What are the properties of estimated coefficient?

Solution and marking

a) The partial adjustment model can be evaluated on the basis of regression (1), where the term with lagged dependent variable is present. In principle if to substitute the regression (PA1) $J_t^* = \beta_1 + \beta_2 P_t + u_t$ into equation of adjustment (PA2) $J_t - J_{t-1} = \lambda(J_t^* - J_{t-1})$, one can get $J_t - J_{t-1} = \lambda(\beta_1 + \beta_2 P_t + u_t - J_{t-1})$ and so $J_t = \lambda\beta_1 + \lambda\beta_2 P_t + \lambda u_t + (1-\lambda)J_{t-1} + \lambda u_t$. Considering (1) as partial adjustment model one can evaluate $1-\lambda = 1.052$, so $\lambda = -0.052$ then $\lambda\beta_2 = (-0.052)\cdot\beta_2 = -0.003$ so $\beta_2 = 0.058$. By analogy $\lambda\beta_1 = (-0.052)\cdot\beta_1 = 0.36$ then $\beta_1 = -6.92$ (*it was not expected that students to present this math, the general idea would be enough*).

It is essentially to note that if the disturbance term is not subject autocorrelation OLS estimators of the regression equation in the partial adjustment model would be consistent (the disturbance term of equation (PA1) is merely the same as disturbance term in (2)). It could be mentioned that although both slope coefficients of the regression equation (2) have correct sigh, the coefficient of prices is insignificant, so some other models should be considered.

[Total 5 marks for a].

(b) □ How the adaptive expectation model (AE1-AE2) could be evaluated on the basis of the estimation results.
□ What are the properties of estimated coefficient?

Solution and marking

b) The adaptive expectation model is also fitted by equation (2), now using Koyck transformation:

First step – substitution of expectation mechanism into regression

$$P_{t+1}^e - P_t^e = \lambda(P_t - P_t^e) \Rightarrow P_{t+1}^e = P_t^e + \lambda P_t - \lambda P_t^e = \lambda P_t + (1-\lambda)P_t^e$$

$$J_t = \beta_1 + \beta_2 P_{t+1}^e + u_t = \beta_1 + \beta_2(\lambda P_t + (1-\lambda)P_t^e) + u_t$$

$$J_t = \beta_1 + \beta_2\lambda P_t + (1-\lambda)\beta_2 P_t^e + u_t \qquad (*)$$

Second step –lagging regression

$$J_{t-1} = \beta_1 + \beta_2 P_t^e + u_{t-1} \Rightarrow \beta_2 P_t^e = \beta_1 - J_{t-1} + u_{t-1}$$

and substituting it into transformed (*)

$$J_t = \beta_1 + \beta_2\lambda P_t + (1-\lambda)(\beta_1 - J_{t-1} + u_t) + u_t$$

$J_t = \beta_1 + (1-\lambda)\cdot\beta_1 + \beta_2\lambda P_t - (1-\lambda)J_{t-1} + u_t + (1-\lambda)\cdot u_{t-1}$ (*this math deserves bonus mark*).

But now the disturbance term after Koyck transformation is subject the autocorrelation (the disturbance term follows autocorrelation process of the moving average type). So the OLS estimates of the regression (2) and would be inconsistent even if u_t is not subject autocorrelation.

(c) □ Do common factor test and make your decision.
□ What is interpretation of the coefficient for lagged dependent variable?
□ What is final conclusion from the whole analysis?

Solution and marking

The common factor test relates to the ADL(1, 1) models (2) and (3),

$$J_t = \beta_1 + \beta_2 P_t + \beta_3 P_{t-1} + \beta_4 J_{t-1} + u_t \qquad (2^*)$$

$$J_t = \beta_1 + \beta_2 P_t - \beta_2\cdot\beta_4 P_{t-1} + \beta_4 J_{t-1} + u_t \qquad (3^*)$$

which has the same set of variables but the model (3) uses the restriction $\beta_3 = -\beta_2\cdot\beta_4$.

The test statistics is $\chi^2 = n\cdot\ln(RSS_{restricted}/RSS_{unrestricted})$, where n should be adjusted according to the lag structure. Using data provided we get $\chi^2 = 24\cdot\ln(1.332/0.842) = 11.01$. As chi-square critical at 0.1% significance (here df=1, as the number of restrictions is 1) is 10.828 the null hypothesis of the validity of the restriction is rejected at any reasonable significance level. So unrestricted version of ADL(1, 1) model (equation (2)) should be chosen for estimation and analysis. The interpretation of the coefficient for lagged dependent variable is something like influence of the former consumption of jewelry items on current consumption. We would have interpreted this coefficient as autocorrelation coefficient if the restriction were valid so we could have assumed the autocorrelated disturbance term.

SECTION B

Answer **TWO** questions from 5-7.

Question 5.

The following results were obtained in an investigation of the agricultural production function of two groups of farmers: Transylvanian natives and Arcadian immigrants.

Transylvanian farmers: $$\log \hat{Q} = \underset{(0.30)}{0.50 + 0.65 \log L} + \underset{(0.14)}{0.30 \log K}, \qquad N = 20, \quad R^2 = 0.90, \quad TSS = 1{,}000 \qquad (1)$$

Arcadian farmers: $$\log \hat{Q} = \underset{(0.31)}{0.70 + 0.75 \log L} + \underset{(0.15)}{0.35 \log K}, \qquad N = 30, \quad R^2 = 0.80, \quad TSS = 1{,}500 \qquad (2)$$

All farmers: $$\log \hat{Q} = 0.65 + 0.68 \log L + 0.33 \log K, \qquad N = 50, \quad R^2 = 0.70, \quad RSS = 600 \qquad (3)$$

standard errors are in parentheses, N denotes the sample size, TSS the total sum of squares, RSS the residual sum of squares.

(a) □ Test whether the two ethnical groups obey different production functions based on available information.

To study whether there is a significant difference between two functions one can to perform a Chow test. First one should evaluate the values of RSS for Transylvanian and Arcadian farmers. As $R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$ one can get $0.9 = 1 - \frac{RSS_{Trans}}{1{,}000}$, so $RSS_{Trans} = 100$, analogously $0.8 = 1 - \frac{RSS_{Arc}}{1{,}500}$, so $RSS_{Arc} = 300$. Now there are all data for Chow test $F = \frac{(RSS_{Total} - (RSS_{Trans} + RSS_{Arc}))/k}{(RSS_{Trans} + RSS_{Arc})/(n+m-2k)}$, where $n = 20$ - number of Transylvanian farmers, $m = 30$ - number of Arcadian farmers, $k = 3$ - number of parameters estimated, so $F = \frac{(600-(100+300))/3}{(100+300)/(20+30-6)} = 7.33$, so the difference is significant as $F^{1\%}_{crit}(3, 44) < F^{1\%}_{crit}(3, 40) = 4.31$.

(b) □ Describe how you would use dummy variables to answer the question in (a).

□ What additional information is needed for this?

□ What is relation of this approach to the Chow test, and what are the advantages and disadvantages of the two approaches?

Another approach to evaluate the difference between two production functions is introduction of the full set of dummies. It includes one shift dummy D_i (equal 0 for Transylvanian farmers and 1 for Arcadian), and two slope dummies $D_i \log K_i$ and $D_i \log L_i$. The equation

$$\log Y_i = \gamma + \gamma' D_i + \alpha \log L_i + \alpha' D_i \log L_i + \beta \log K_i + \beta' D_i \log K_i + u_i \qquad (*)$$

should be estimated for the pooled sample of all farmers belonging to both ethnic group. The equation

$$\log Y_i = \gamma + \alpha \log L_i + \beta \log K_i + u_i \qquad (**)$$

refers now to Transylvanian farmers, while $\log Y_i = (\gamma + \gamma') + (\alpha + \alpha') \log L_i + (\beta + \beta') \log K_i + u_i$ relates to the Arcadian ones. The difference between two production functions could be evaluated using F test on the explanatory power of a set of variables, interacted with the dummy variable indicating farmer's type:

$$F = \frac{(RSS(**) - RSS(*))/3}{RSS(*)/(50-6)}$$

Two tests are equivalent, they gives precisely the same value of F. Chow test is well known and popular in applied Econometrics as it is usually available in many computer programs of econometrics analysis.
On the other hand one should be noticed that dummy approach gives more detailed information on the model. Combining F test with t tests for dummies it is possible to reveal some variables responsible for the difference of two production functions.

(c) □ Suggest some test to decide whether the Transylvanian farmers experience constant returns to scale.
□ What additional information is needed for this? How to get it?

Constant returns to scale corresponds to the restriction $\alpha+\beta=1$ for the regression (standard Cobb-Douglas production function) for Transylvanian farmers

$$\log Y_i = \gamma + \alpha \log L_i + \beta \log K_i + u_i \qquad \textbf{(U)}$$

with slope coefficients that represent correspondingly the labor and capital elasticities. This is unrestricted equation with $RSS_U^{Trans}=100$ (from (a)). To use F-test, he we need to know also the value of RSS_R^{Trans} from estimating equation restricted equation

$$\log Y_i = \gamma + (1-\beta)\log L_i + \beta \log K_i + u_i$$

or
$$\log Y_i = \gamma + \log L_i - \beta \log L_i + \beta \log K_i + u_i$$

or
$$\log Y_i - \log L_i = \gamma + \beta(\log K_i - \log L_i) + u_i$$

or rather
$$\log \frac{Y_i}{L_i} = \gamma + \beta \log \frac{K_i}{L_i} + u_i \qquad \textbf{(R)}$$

Now do F test $\dfrac{(RSS_R^{Trans} - RSS_U^{Trans})/1}{RSS_U^{Trans}/(20-3)} \sim F(1,17)$

(d) □ Suggest some test to decide whether the Arcadian farmers experience increasing returns to scale.
□ What additional information is needed for this? How to get it?

Unlike the conventional two sided alternative that allows to use F-test, here we need in one-sided test. Appropriate test could be constructed on the base of conventional t-test.
To do this we should first reparameterize the original equation for Arcadian farmers

$$\log Y_i = \gamma + \alpha \log L_i + \beta \log K_i + u_i$$
$$\log Y_i = \gamma + (\alpha+\beta-1)\log L_i - \beta \log L_i + \log L_i + \beta \log K_i + u_i$$
$$\log Y_i - \log L_i = \gamma + (\alpha+\beta-1)\log L_i + \beta(\log K_i - \log L_i) + u_i$$

or
$$\log \frac{Y_i}{L_i} = \gamma + (\alpha+\beta-1)\log L_i + \beta \log \frac{K_i}{L_i} + u_i$$

Running this equation we get t-statistic for the estimated coefficient $(\alpha+\beta-1)$: $t_{\alpha+\beta-1}$. Now do one-sided t-test for the pair of hypotheses $H_0:(\alpha+\beta-1)=0$, $H_a:(\alpha+\beta-1)>0$.

Question 6.
An OLS regression of housing expenditures y_t on income x_t and x_{t-1} (both in hundreds of dollars) evaluated using monthly data gives the following results (with the standard errors given in parentheses)

$$\hat{y}_t = \underset{(2.30)}{8.88} + \underset{(2.11)}{5.17}x_t - \underset{(3.01)}{3.18}x_{t-1}; \quad R^2 = 0.095, \; T = 209. \qquad \textbf{(1)}$$

(a) □ What is the estimate of the short-run effect of x_t on y_t? Interpret these estimate. Is it significant?
□ Test the hypothesis that a one dollar increase in *x* results in a five dollars increase in *y* in the same month.
□ What about ten-fold increase in *y* in the same month?
□ Under what assumptions is this test valid?

Solution:

Short-run effect: this is coefficient at variable x_t: y_t increases by 5.07 hundreds of dollars as x_t increases by 1 hundred of dollars. It is significant at 5% significance level as $t=\frac{5.17}{2.30}=2.25>1.97=t_{crit}^{5\%}(200)$

to test the hypothesis of a five-fold effect of change in x one should evaluate slightly modified statistic $t=\frac{5.17-5}{2.30}=0.07<1.96$ so this hypothesis cannot be rejected, while for ten-fold effect null hypothesis is rejected $t=\frac{|5.17-10|}{2.30}=|-2.1|>1.97=t_{crit}^{5\%}(200)$. The assumptions for these tests are validity of GMC plus normality.

(b) □ Derive the expression for the long run effect of x_t on y_t?

□ How to test that it is significant?

To derive long run effect first write theoretical equation under consideration.

$$y_t=\beta_0+\beta_1 x_t+\beta_2 x_{t-1}+u_t \quad \textbf{(*)}$$

In the long run all variables are equal to their equilibrium values $\bar{x}, \bar{y}$, so

$$\bar{y}=\beta_0+\beta_1\bar{x}+\beta_2\bar{x}+u_t$$

or

$$\bar{y}=\beta_0+(\beta_1+\beta_2)\bar{x}+u_t$$

So long-run effect is $\beta_1+\beta_2$: **it shows the effect a permanent change in *x* with one hundred dollars has on *y* after one month period (the last lagged response) has passed**: increases with $5.07-3.18=1.89$ hundreds of dollars.

To evaluate the significance of this effect we need to reparameterise equation (*) to obtain the standard error of $\hat{\beta}_1+\hat{\beta}_2$ directly:

$$y_t=\beta_0+\beta_1 x_t+\beta_2 x_{t-1}+u_t,$$
$$y_t=\beta_0+\beta_1 x_t+\beta_2 x_t-\beta_2 x_t+\beta_2 x_{t-1}+u_t,$$
$$y_t=\beta_0+(\beta_1+\beta_2)x_t-\beta_2\Delta x_t+u_t.$$

Regress y_t on a constant, x_t and Δx_t and the standard error of the estimated coefficient on x_t is the standard error of the long-run effect, so the standard t test can be done $t=\frac{\widehat{\beta_1+\beta_2}}{s.e.(\widehat{\beta_1+\beta_2})}$.

(c) Let e_t be the OLS residuals from the above regression. An OLS regression of e_t on e_{t-1} yields

$$\underset{(2.12)}{e_t=0.55}+\underset{(0.18)}{0.44e_{t-1}}+\underset{(2.18)}{2.16x_t}-\underset{(.99)}{1.09x_{t-1}};\ R^2=0.175,\ T=208. \qquad \textbf{(2)}$$

□ Using this result, test for evidence of autocorrelation, clearly indicating the null and alternative hypotheses, the test statistic, rejection rule and assumptions underlying the test. What name do we give this test?

Solution:

The test equation tells us that we have to perform the Breusch-Godfrey test for first-order autocorrelation.

Let ρ denote the coefficient associated with e_{t-1}. The pair of hypotheses - $H_0:\rho=0$ (No autocorrelation) vs. $H_1:\rho\neq 0$ (Autocorrelation) - can be tested either using conventional t test: $t=\frac{0.44}{0.18}=2.44>1.97$ (H_0 is rejected) or using Lagrange multiplier test statistic: $LM=nR^2=28\times 0.175\approx 4.9$. Under H_0, $LM\overset{asympt}{\sim}\chi_1^2$ we reject H_0 in favour of autocorrelation since $LM>3.84=\chi_1^{5\%}$.

(d) □ As it is known under the presence of autocorrelation conventional t-test becomes invalid. Indicate how you proceed with your test in this case.

The inclusion of the lagged dependent variable into the set of regressors can help to cope with autocorrelation.

$$\hat{y}_t = \underset{(2.30)}{6.84} + \underset{(2.11)}{2.12} x_t - \underset{(3.01)}{1.54} x_{t-1} + 0.79 y_{t-1}; \; R^2 = 0.23, \; T = 209. \quad \textbf{(3)}$$

□ How this changes the value and interpretation of the long run effect of x on y?

The usual standard errors will be invalidated so we need to use HAC (Heteroscedasticity and Autocorrelation Consistent) standard errors instead (so called **robust standard errors**, for example in the form of Newey-West). Given the robust standard errors, one could compute the test statistic robust to autocorrelation as $t^{robust} = \dfrac{\hat{\theta}}{se(\hat{\theta})^{HAC}}$ and conclude the long-run effect is statistically significant at the 5% significance level if $|t^{robust}| > 1.96$.

To derive long run effect first write theoretical equation under consideration.

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \beta_3 y_{t-1} + u_t \; \textbf{(*)}$$

In the long run all variables are equal to their equilibrium values $\bar{x}, \bar{y}$, so

$$\bar{y} = \beta_0 + \beta_1 \bar{x} + \beta_2 \bar{x} + \beta_3 \bar{y} + u_t$$

or

$$(1-\beta_3)\bar{y} = \beta_0 + (\beta_1 + \beta_2)\bar{x} + u_t$$

and so

$$\bar{y} = \frac{\beta_0}{1-\beta_3} + \frac{\beta_1 + \beta_2}{1-\beta_3}\bar{x} + u_t$$

So long-run effect is $\dfrac{\beta_1 + \beta_2}{1-\beta_3}$: **it shows the effect a permanent change in *x* with one hundred dollars has on *y* after an infinite number of one month periods has passed**: increases with $\dfrac{2.12 - 1.54}{1 - 0.79} = 2.76$ hundreds of dollars.

Question 7.

Consider the model

$$y_t = \alpha + \beta t + \varepsilon_t, \; t = 1, ..., T, \quad (1)$$

$$\varepsilon_t = \rho \varepsilon_{t-1} + v_t,$$

and v_t is an i.i.d. $(0, \sigma^2)$ innovation which is independent of the past. Let $|\rho| \le 1$.

(a) □ What name do we give the y_t process given above? Provide the condition(s) that ensures that ε_t is stationary.

□ Discuss what we mean by the concept of stationarity (more precisely 'covariance stationarity').

The process

$$y_t = \alpha + \beta t + \varepsilon_t, \; t = 1, ..., T, \quad (1)$$

is certainly time trend, which is non-stationary. Its disturbance term ε_t follows autocorrelation process $\varepsilon_t = \rho \varepsilon_{t-1} + v_t$ if $|\rho| < 1$ (stationary process), but it turns into a random walk $\varepsilon_t = \varepsilon_{t-1} + v_t$ if $|\rho| = 1$ (non-stationary process).

The time series y_t is called covariance stationary if three conditions holds

1) Ey_t does not depent on t
2) $\text{var}(y_t)$ does not depent on t
3) $\text{cov}(y_t; y_{t+s})$ does not depent on t but can change with the change in s.

(b) □ It will be important to distinguish between the above process for y_t, being 'trend stationary' as opposed to 'difference stationary'. Explain these concepts clearly. Why is it important to distinguish between these two types of non-stationarity?

□ Show that under the condition $|\rho|<1$ y_t is trend stationary.

□ Show that if $\rho=1$ then y_t is difference stationary.

A process $\{y_t\}_{t=-\infty}^{\infty}$ is said to be difference stationary if its first-difference is stationary ($\Delta y_t \sim I(0)$). It is said to be trend stationary if the process $y_t - \beta t$ is stationary. It is important to distinguish between the two because the source of non-stationarity has different implications on how we proceed to obtain a stationary time-series to use for regression analysis and for statistical inference.

If $|\rho|<1$ then ε_t is stationary, since $y_t - \beta t = \alpha + \varepsilon_t$ is also stationary. Therefore, y_t is trend stationary under $|\rho|<1$.

If $\rho=1$ then the process $\varepsilon_t = \varepsilon_{t-1} + v_t$ is random walk, and so ε_t is difference stationary ($\Delta\varepsilon_t \sim I(0)$). Taking first differences of $y_t = \alpha + \beta t + \varepsilon_t$ we obtain

$$y_t = \alpha + \beta t + \varepsilon_t$$
$$y_{t-1} = \alpha + \beta(t-1) + \varepsilon_{t-1}$$
$$\Delta y_t = \beta + \Delta\varepsilon_t,$$

since β is just a constant we have that $\Delta y_t \sim I(0)$, that is, y_t is difference stationary.

(c) □ Show that you can rewrite the above model in the following form

$$\Delta y_t = \gamma_1 + \gamma_2 t + \gamma_3 y_{t-1} + v_t. \qquad (2)$$

Clearly indicate the relation between $(\gamma_1, \gamma_2, \gamma_3)$ and (α, β, ρ).

Solution:

Lagging

$$y_t = \alpha + \beta t + \varepsilon_t,\ t = 1, ..., T,\quad \varepsilon_t = \rho\varepsilon_{t-1} + v_t \qquad (1)$$

by one period and multiplying both sides by ρ we obtain:

$$\rho y_{t-1} = \rho\alpha + \rho\beta(t-1) + \rho\varepsilon_{t-1}.$$

Subtracting the above from (1) yields:

$$y_t - \rho y_{t-1} = (1-\rho)\alpha + \beta t - \rho\beta t + \rho\beta + \varepsilon_t - \rho\varepsilon_{t-1}.$$

Let $v_t \equiv \varepsilon_t - \rho\varepsilon_{t-1}$ and rearrange:

$$y_t = ((1-\rho)\alpha + \rho\beta) + \beta(1-\rho)t + \rho y_{t-1} + v_t.$$

Finally, subtract y_{t-1} on both sides to obtain:

$$\Delta y_t = \underbrace{(1-\rho)\alpha + \rho\beta}_{\gamma_1} + \underbrace{\beta(1-\rho)}_{\gamma_2}t + \underbrace{(\rho-1)}_{\gamma_3}y_{t-1} + v_t.$$

(d) □ How to use Dickey-Fuller test to decide whether the process

$$y_t = \alpha + \beta t + \varepsilon_t,\ t = 1, ..., T,\ \text{where } \varepsilon_t = \rho\varepsilon_{t-1} + v_t \qquad (1)$$

is trend stationary or difference stationary?

□ What additional problems emerge when v_t exhibits autocorrelation? What solution do you suggest we adopt?

For this purpose one can use representation

$$\Delta y_t = \gamma_1 + \gamma_2 t + \gamma_3 y_{t-1} + v_t \qquad (2)$$

from **(c).**

To test for difference stationarity one should test null hypothesis $H_0:\gamma_3=0$ against $H_1:\gamma_3<0$ using ADF t test.

To test for trend stationarity one should test null hypothesis $H_0:\gamma_2=0$ against $H_1:\gamma_2\neq 0$ using conventional t test.

If v_t exhibits autocorrelation it must be eliminated before running the test regression otherwise the Dickey-Fuller test will not be valid. To eliminate it we should include the lags of Δy_t in the test equation (2).

$$\Delta y_t=\gamma_1+\gamma_2 t+\gamma_3 y_{t-1}+\gamma_4\Delta y_{t-1}+v_t \qquad (2)$$

This test is then known as the augmented Dickey-Fuller test.