Econometrics – 2022-2023. First Semester Exam, December 29, 2022

Part 1. (30 minutes). In each of 12 multiple choice tests indicate the correct answer.

1. Which of the following statements is true?

1) If the calculated value of *F* statistic is higher than the critical value, we reject the alternative hypothesis in favor of the null hypothesis.
2) The *F* statistic is always nonnegative as *SSRr* is never smaller than *SSRur*.
3) Degrees of freedom of a restricted model is always less than the degrees of freedom of an unrestricted model.
4) The *F* statistic is more flexible than the t statistic to test a hypothesis with a single restriction.
5) None of the above.

2. In a regression model, if variance of the dependent variable, *Y*, conditional on an explanatory variable, *X*, is not constant, then

1) the t statistics are invalid and confidence intervals are valid for small sample sizes;
2) the t statistics are valid and confidence intervals are invalid for small sample sizes;
3) the t statistics and confidence intervals are valid no matter how large the sample size is;
4) the t statistics and confidence intervals are both invalid no matter how large the sample size is;
5) The OLS estimators are biased, and hence no need to discuss t statistics and confidence intervals.

3. In econometrics, simultaneity bias arises when:

1) strictly exogenous explanatory variables determine the dependent variable through a step-by-step process.
2) the disturbance term is correlated with the dependent variable.
3) one or more of the explanatory variables is jointly determined with the dependent variable.
4) heteroscedasticity is present in the model.
5) There is correlation between some explanatory variables.

4. For the Model $Y_i = \beta_1+\beta_2 X_i + u$ (X_i are non-stochastic, the Model A assumptions satisfied) the estimator $b = \dfrac{\sum_{i=2}^{n}(Y_i - Y_{i-1})}{\sum_{i=2}^{n}(X_i - X_{i-1})}$ is generally speaking:

1) unbiased and efficient estimator of β_2;
2) unbiased but inefficient estimator of β_2;
3) biased estimator of β_2;
4) non-linear estimator of β_2;
5) non-stochastic.

5. For the sample of 55 observations, functions (1) and (2) were estimated:

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$ (1)

$Y = \beta_0 + \beta_1 (X_1 - X_2) + u$ (2)

The R^2 (determination coefficients) for these models are 0.9 in (1) and 0.7 in (2) respectively. F – statistic for testing the hypothesis $\beta_1=\beta_2$ in (1) equals

1) 6.7; 2) 8.2; 3) 30; 4) 104; 5) You can not test this hypothesis using (1) and (2).

6.The function of expenditures for cosmetics depending on disposable personal income has been estimated using OLS, for a representative sample of people:

$$Y=\beta_0+\beta_1 D_1+\beta_2 X+\beta_3 X(1-D_2)+u$$

where Y is expenditure for cosmetics, X is disposable personal income,
$D_1 = 1$ for females and 0 for males, $D_2 = 1$ for males and 0 for females.
For this regression the following is correct:

1) The estimates of intercept are the same for male and female subsamples, while the estimates of slope coefficient, generally speaking, differ for them;
2) The estimates of slope coefficient are the same for male and female subsamples, while the estimates of intercept, generally speaking, differ for them;
3) Both intercepts and slope coefficients estimated, generally speaking, differ for male and female subsamples;
4) Both intercepts and slope coefficients estimated are the same for male and female subsamples;
5) The combination of intercept and slope dummies is incorrect, and the model can not be estimated.

7. If you have estimated the parameters of the following model using the OLS directly (Gauss-Markov conditions satisfied),

$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + (\beta_2(1+\beta_3))x_3 + u,$ then:

1) you can get an unbiased estimate of β_3;
2) you can not get an unbiased estimate of β_3, but can easily get a consistent estimate of it;
3) you can not get an unbiased, or biased but consistent estimate of β_3;
4) you can not get any estimate of β_3;
5) all the above statements are incorrect.

8. If the OLS is used in simple regression model in the case of heteroscedasticity, the population variance of slope coefficient is $\mathrm{var}(b_2) = \dfrac{\sum_{i=1}^{n} x_i^2 \sigma_i^2}{(\sum_{i=1}^{n} x_i^2)^2}$ (1). The formula for homoscedasticity case is $\mathrm{var}(b_2) = \dfrac{\sigma^2}{\sum_{i=1}^{n} x_i^2}$ (2). Let $\sigma_i^2=\sigma^2 k_i$, where k_i are unknown non-negative weights $(\Sigma k_i=1)$.

Then:

1) The expression (1) is always greater than (2);
2) The expression (1) is always less than (2);
3) The expression (1) is greater or equal to (2);
4) The expression (1) is less or equal to (2);
5) The expression (1) can be greater, less or equal to (2), depending on the nature of relationship between σ_i and x_i.

9. In the regression model $y = \alpha + \beta x + u$ (where the disturbance term u satisfies Gauss-Markov conditions and is normally distributed) the explanatory variable x includes random measurement errors (which are independent, normally distributed, homoscedastic, not autocorrelated, with zero expected values), ($\beta<0$), and negative mean value of x. In this case, when estimating the model using OLS, for large samples

1) the estimator of α will be biased upwards;
2) the estimator of α will be biased downwards;
3) the estimator of α will be unbiased;
4) the estimator of α may be biased upwards or downwards;
5) the OLS estimator of α does not exist.

10. For the simultaneous equations model with 7 equations, 7 endogenous variables and 7 exogenous variables, the following statement is true:

1) with that number of potential instruments, any equation is identified in the model;
2) the equation in the model is identified if and only if only exogenous variables are available on its right side;
3) the number of potential instruments is insufficient to make all the equations identified;
4) no equation can be overidentified in the model;
5) None of the above.

11. Economic model is described by the following simultaneous equations:

$$(1) \qquad y_1 = \delta + \tau y_2 + \pi x_2 + u_2$$
$$(2) \qquad y_2 = \alpha + \pi y_1 + \gamma x_1 + \varphi x_2 + u_1$$

where y_1 and y_2 are endogenous variables, x_1 and x_2 are stochastic exogenous variables, u_1 and u_2 are disturbance terms satisfying Gauss-Markov conditions. Indicate the correct statement:

1) you may apply TSLS in (1), but not in (2);
2) you may apply TSLS in (2), but not in (1);
3) you may apply TSLS in both (1) and (2);
4) you may not apply TSLS in either (1) or (2);
5) TSLS is not needed since the OLS provides consistent estimates in (1) and (2).

12. The model with the dependent variable P_i (monthly pension), as a function of Work Experience WE_i and the average earnings $EARN_i$ is being considered:

$$P_i = \beta_1 + \beta_2 WE_i + \beta_3 EARN_i + u_i$$

The value of pension is restricted by the values P_U and P_L from the top and from the bottom, but there are no actual observations in the sample with $P=P_U$ or $P=P_L$. The student decided to estimate Tobit model for this sample. Please indicate the **correct** statement among the following ones:

1) The Tobit estimators of the model coefficients are biased and inconsistent;
2) The Tobit estimators of the model coefficients are biased but consistent;
3) The Tobit model estimates will be the same as of the OLS here;
4) The Tobit model may not be estimated for this sample;
5) None of the above.