The International College of Economics and Finance
Econometrics-2021-2022
Home assignment 5. Stochastic Regressors. Simultaneous Equations.
To be submitted by December 12, 23:55

1. [30 marks] The following simultaneous equations model is considered:

$$Y = \beta_1 + \beta_2 X + u \quad \textbf{(1)}$$

$$X = \alpha_1 + \alpha_2 Y + v \quad \textbf{(2)}$$

where X and Y are endogenous variables, and u and v are identically and independently distributed disturbance terms with zero means. The sample consists of n observations (X_i, Y_i).

1.1. [10 marks] □ Derive reduced form system of equations for the system above.
□ Using reduced form system show that in equations (1)-(2) Gauss-Markov conditions (GMC) are violated

1.2. [10 marks] □ Show that OLS estimator $\hat{\alpha}_2^{OLS}$ of α_2 is inconsistent.

1.3. [10 marks] □ What can be said on the identification of the second equation?
An additional instrument is introduced into the equation system an exogenous variable Z that correlates with a variable Y but does not correlate with a random term v of the second equation.

$$Y = \beta_1 + \beta_2 X + \beta_3 Z + u \quad \textbf{(1*)}$$

$$X = \alpha_1 + \alpha_2 Y + v \quad \textbf{(2)}$$

□ Show that the instrumental variable estimator $\hat{\alpha}_2^{IV}$ based on the instrument Z is consistent.
□ The researcher decides to use two-stage least squares (TSLS) hoping to obtain a more efficient estimator of α_2. First he fits OLS regression

$$\hat{Y} = h_1 + h_2 Z \quad \textbf{(3)}$$

saves the fitted values, and uses them as an instrument for *Y* in equation (2). Demonstrate that obtained TSLS estimator $\hat{\alpha}_2^{TSLS}$ is the same as $\hat{\alpha}_2^{IV}$.

2. [20 marks] During the pandemic, traditional cinemas suffer losses and companies providing films for viewing on the Internet (like Netflix) thrive. A researcher is interested in investigating the expendures on internet films using cross-section data for 43 countries assuming that expenditure on films, *q*, is related to total consumer expenditure, *z*, by the relationship

$$q = \alpha + \beta z + v$$

where *v* is a disturbance term which satisfies the Gauss-Markov conditions. Both variables *q* and *z* are measured with error, and the researcher believe that any error in the estimation of *q* affects the estimate of *z* by the same amount: $y_i = q_i + w_i$ and $x_i = z_i + w_i$ where y_i is the estimated value of q_i, x_i is the estimated value of z_i, and w_i is the measurement error affecting both variables in observation *i*. It is assumed that the expected value of *w* is zero and that *v* and *w* are distributed independently of *z* and of each other. Note since expenditure on films is a component of total consumer expenditure, β will lie between 0 and 1.

2.1. [10 marks] □ Derive an expression for the large-sample bias in the estimate of β when Ordinary Least Squares is used to regress y_i on x_i, and determine its sign if this is possible.

2.2. [10 marks] □ The researcher is worried of the fact that the analysis could be affected by positive correlation of *w* with *z,* as observations with large *z* tend to have larger measurement errors *w*. Comment.
□ Trying to overcome consequences of bias caused by measurement errors the researcher decided to use disposable personal income, *I* as an instrument for total consumer expenditure, *z*, assuming that *I* correlates

with *z* but not correlates with *v* and *w*. Comment providing necessary proofs, taking into account that consumer expenditures on films, *q*, still are under measurement errors *w* .

3. [50 marks] It is known that people who are more successful in learning tend to study longer (as measured by the indicator S_i), and there is also the opposite effect, which is that people who are more inclined to continue learning exhibit higher abilities as measured by the indicator $ASVABC_i$. Based on this, a model of simultaneous equations has been developed. A dummy variable $MALE_i$ was also included in the second equation because it was assumed that learning outcomes are partially dependent on the gender of the respondent.

$$S_i = \beta_1 + \beta_2 ASVABC_i + u_i \quad \textbf{(1)}$$

$$ASVABC_i = \alpha_1 + \alpha_2 S_i + \alpha_3 MALE_i + v_i \quad \textbf{(2)}$$

In all parts of this question full mathematical reasoning is not expected and no additional points will be given for that. Demonstration and comparison of the calculation results with brief comments are welcome.

3.1. [10 marks]. Discuss briefly the situation with identification of both equations. What are the consequences of their estimation by OLS? Obtain OLS estimates of the slope coefficients α_2 and β_2 using **your data set ha05_data__**. Discuss possible direction of bias of one of the slope coefficients' estimators. Is it possible to see the bias of estimates from the estimated equations?

3.2. [10 marks] □ Perform the Instrumental Variables estimation of the coefficient β_2 and compare it with the corresponding OLS estimate. Comment the statisical properties of obtained IV estimator? If the instrument used is found to be weak, then try to find a more suitable instrument. Please provide a justification for your choice and compare the results.

3.3. [10 marks] A new system of equations can be proposed

$$S_i = \beta_1 + \beta_2 ASVABC_i + \beta_3 SM_i + \beta_4 SF_i + \beta_4 POV78_i + u_i \quad \textbf{(1*)}$$

$$ASVABC_i = \alpha_1 + \alpha_2 S_i + \alpha_3 MALE_i + v_i \quad \textbf{(2)}$$

Perform the TSLS estimation of coefficients α_2 and β_2 of equations (1*) and (2). Explain the reasons for combining different instruments into the one superinstrument.

3.4. [10 marks] Perform Hausman test (Davidson-McKinnon modification based on the instruments above) to test for the endogeneity of a variable S_i caused by the interaction of variables S_i and $ASVABC_i$. Conclude.

3.5. [10 marks] It has been suggested that a variable $ASVABC_i$ (student test score) measures the ability and knowledge of students with the error caused by student nervousness on any exam and test. Sometimes excitement in the exam helps to show the best result, in other cases it has negative effect, this factor can be interpreted as a measurement error.

□ How the presence of a measurement error in a variable $ASVABC_i$ influence the estimation of the coefficient β_2 of equation (1)?

□ Design a Monte Carlo experiment that demonstrates the effect of the measurement error of a variable $ASVABC_i$ on the estimation of the coefficient β_2 of equation (1) (give a general description of the principle and scheme of the method)?

□ Conduct a small Monte Carlo experiment according to the scheme you designed (no more than 5 iterations), demonstrating the technique of the Monte Carlo experiment, choosing the appropriate values for the parameters of your scheme. (*A full-range experiment using programming and repeating thousands and millions of times is not intended here and does not give any additional points*).

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