Econometrics 1 – Honours

Economics 468

Thursday December 10, 2020

Examiner: Russell Davidson

Instructions:

Your completed exam should be uploaded to myCourses by 18.30 on December 12. Please submit two files per student: one, which should be a PDF file, with your written answers, and another, which can be a simple text file, with your computer code. These files must be all your own work. You may make use of whatever non-human resources you wish, but you must not ask for or receive any help from any other person.

All students in this course have the right to submit in English or in French any written work that is to be graded.

Tou(te)s les étudiant(e)s qui suivent ce cours on le droit de soumettre tout travail écrit en français ou en anglais.

Academic Integrity statement [approved by Senate on 29 January 2003]:

McGill University values academic integrity. Therefore all students must understand the meaning and consequences of cheating, plagiarism and other academic offences under the Code of Student Conduct and Disciplinary Procedures.

L'université McGill attache une haute importance à l'honnêteté académique. Il incombe par conséquent à tou(te)s les étudiant(e)s de comprendre ce que l'on entend par tricherie, plagiat et autres infractions académiques, ainsi que les conséquences que peuvent avoir de telles actions, selon le Code de conduite de l'étudiant(e) et des procédures disciplinaires.

This exam comprises six pages, including the cover page

Answer all questions in this exam.

Faites tous les exercices de cet examen.

1. Consider the following classical normal linear model:

$$\boldsymbol{y} = \beta_0\boldsymbol{\iota} + \beta_1\boldsymbol{x}_1 + \beta_2\boldsymbol{x}_2 + \beta_3\boldsymbol{x}_3 + \boldsymbol{u}, \quad \boldsymbol{u} \sim \mathrm{N}(\boldsymbol{0}, \sigma^2\mathbf{I}).$$

Here are the results from estimating this model for a given data set by ordinary least squares (OLS).

```
Ordinary Least Squares:

Variable     Parameter estimate  Standard error  T statistic

constant        20.565309          71.387472       0.288080
x1               2.683988           0.615305       4.362045
x2              -2.903900           0.411412      -7.058367
x3               1.516683           0.461700       3.285000

Number of observations = 10   Number of estimated parameters = 4
Mean of dependent variable = 136.541558
Sum of squared residuals = 7486.980957
Explained sum of squares = 309162.149792
Estimate of residual variance
  (with d.f. correction) = 1247.830160
Mean of squared residuals = 748.698096
Standard error of regression = 35.324640
R squared (uncentred) = 0.976356  (centred) = 0.942502

Estimated covariance matrix:

5096.171209  -39.175469  -1.724230  24.047773
 -39.175469    0.378600   0.012932  -0.107232
  -1.724230    0.012932   0.169260  -0.008771
  24.047773   -0.107232  -0.008771   0.213166
```

For each of the four estimated coefficients, β_i, $i = 1, 2, 3, 4$, compute the P value for the null hypothesis that $\beta_i = 0$ against the alternative that $\beta_i \neq 0$.

Construct a two-tailed confidence interval for the variance parameter σ^2 at confidence level 95%. Carefully explain your reasoning. Then construct a one-tailed confidence interval for σ^2, open out to plus infinity, at the same confidence level.

2. Use the following data-generating process (DGP)

$$\boldsymbol{y} = \boldsymbol{x}\beta + \sigma_u\boldsymbol{u}$$
$$\boldsymbol{x} = \boldsymbol{w}\pi + \sigma_v\boldsymbol{v},$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ are bivariate normal, both with expectation 0 and variance 1, and such that $\mathrm{E}(u_t v_t) = \rho$ for all $t = 1, \ldots, n$, in order to generate at least 1000 sets of simulated data for $\boldsymbol{x}$ and $\boldsymbol{y}$ with sample size $n = 10$, $\sigma_u = \sigma_v = 1$, $\beta = 0$, $\pi = 1$, and $\rho = 0.5$. For the exogenous instrument $\boldsymbol{w}$, use independent drawings from the standard normal distribution, and then rescale $\boldsymbol{w}$ so that $\boldsymbol{w}^\top\boldsymbol{w}$ is equal to n.

For each simulated data set, compute the simple IV estimator

$$\hat{\beta}_{\text{IV}} = (\boldsymbol{w}^\top\boldsymbol{x})^{-1}\boldsymbol{w}^\top\boldsymbol{y}.$$

Then draw the graph of the empirical distribution function (EDF) of the realizations of the estimator on the same plot as the cumulative distribution function (CDF) of the normal distribution with expectation zero and variance $\sigma_u^2/n\pi^2$. Explain why this is an appropriate way to compare the finite-sample and asymptotic distributions of the estimator. In addition, for each simulated data set, compute the ordinary least squares (OLS) estimator, and plot the EDF of the realizations of this estimator on the same axes as the EDF of the realizations of the IV estimator.

Include one more instrument, $\boldsymbol{z}$, in the simulations, generated in the same way as $\boldsymbol{w}$, independently of anything else in the simulation. Continue to use the same DGP for $\boldsymbol{y}$ and $\boldsymbol{x}$, but replace the simple IV estimator by the generalized one, based on the two instruments $\boldsymbol{w}$ and $\boldsymbol{z}$. See if you can verify the theoretical prediction that the overidentified estimator computed with two instruments is more biased, but has thinner tails, than the just identified estimator.

3. Consider the linear regression model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}, \quad \text{E}(\boldsymbol{u} \,|\, \boldsymbol{X}) = \boldsymbol{0}, \quad \text{E}(\boldsymbol{u}\boldsymbol{u}^\top) = \boldsymbol{\Omega}. \tag{1}$$

The generalized least-squares (GLS) estimator is

$$\hat{\boldsymbol{\beta}}_{\text{GLS}} = (\boldsymbol{X}^\top\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{\Omega}^{-1}\boldsymbol{y},$$

and its covariance matrix is $(\boldsymbol{X}^\top\boldsymbol{\Omega}^{-1}\boldsymbol{X})^{-1}$. Consider an instrumental-variables (IV) estimator defined by the estimating equations

$$\boldsymbol{W}^\top(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\text{IV}}) = \boldsymbol{0},$$

where both $\boldsymbol{X}$ and $\boldsymbol{W}$ are $n \times k$ matrices of exogenous variables. What is the covariance matrix of $\hat{\boldsymbol{\beta}}_{\text{IV}}$ under the data-generating process (DGP) given by (1)?

Show that the difference between the precision matrix of the GLS estimator and the precision matrix of the IV estimator is a positive semi-definite matrix.

4. Let the n–vector $\boldsymbol{y}$ be a vector of mutually independent realizations from the uniform distribution on the interval $[\beta_1, \beta_2]$, usually denoted by $\text{U}(\beta_1, \beta_2)$. Thus, $y_t \sim \text{U}(\beta_1, \beta_2)$ for $t = 1, \ldots, n$. A suitable estimator of the parameter β_1 is

$$\hat{\beta}_1 = \min_{1 \le t \le n}(y_t).$$

Suppose that the true values of the parameters are $\beta_1 = 0$ and $\beta_2 = 1$. Show that the CDF of $\hat{\beta}_1$ is

$$F(\beta) \equiv \Pr(\hat{\beta}_1 \le \beta) = 1 - (1-\beta)^n.$$

Use this result to show that the scaled estimation error $n(\hat{\beta}_1 - \beta_1)$, which in this case is just $n\hat{\beta}_1$, is asymptotically exponentially distributed with parameter $\theta = 1$.

Note: The CDF of the exponential distribution is

$$F(y, \theta) = 1 - \mathrm{e}^{-\theta y}, \quad y > 0, \ \ \theta > 0. \tag{2}$$

(**Hint:** The limit as $n \to \infty$ of $(1 + x/n)^n$, for arbitrary real x, is e^x.)

Show that, for arbitrary true values β_1 and β_2, with $\beta_2 > \beta_1$, the asymptotic distribution of $n(\hat{\beta}_1 - \beta_1)$ is characterized by the density (2) with $\theta = (\beta_2 - \beta_1)^{-1}$.

5. In Section 3.6 of the revised textbook, it is shown that, if the linear regression model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}, \quad \boldsymbol{y} \text{ an } n \times 1 \text{ vector}, \ \boldsymbol{X} \text{ an } n \times k \text{ matrix},$$

is estimated by OLS with observation t omitted from the sample, the parameter estimates for $\boldsymbol{\beta}$ are the same as those given by the regression

$$\boldsymbol{y} = \boldsymbol{X\beta} + \alpha\boldsymbol{e}_t + \boldsymbol{u}, \tag{3}$$

where $\boldsymbol{e}_t$ is a vector all of whose elements are zero except for observation t, for which it is one. Show that the t^{th} residual from running regression (3) is 0.

The usual notation is that h_t is the diagonal element $(\boldsymbol{P_X})_{tt}$ of the "hat matrix" $\boldsymbol{P_X}$, which is the orthogonal projection matrix on to the linear span $\mathcal{S}(\boldsymbol{X})$ of the columns of $\boldsymbol{X}$. Show that the leverage measure h_t is the square of the cosine of the angle between the unit basis vector $\boldsymbol{e}_t$ and its projection on to the span $\mathcal{S}(\boldsymbol{X})$ of the regressors.

Suppose that the matrix $\boldsymbol{X}$ is of dimensions 150×5 and has full rank. Let $\boldsymbol{P_X}$ be the matrix that projects on to $\mathcal{S}(\boldsymbol{X})$ and let $\boldsymbol{M_X} = \mathbf{I} - \boldsymbol{P_X}$. What is $\mathrm{Tr}(\boldsymbol{P_X})$ (the **trace** of $\boldsymbol{P_X}$)? What is $\mathrm{Tr}(\boldsymbol{M_X})$? What would these be if $\boldsymbol{X}$ did not have full rank but instead had rank 3?

6. The file

https://russell-davidson.arts.mcgill.ca/e468/e468.dec20.dat

contains 4266 observations arranged in five columns. The first is just the observation number. The next three columns are indicator (dummy) variables. Each indicates which of three groups the observation belongs to. Observation t of the fifth column, y_t, is the income of the person indexed by t.

Using the data in the file, run the regression

$$y_t = \beta_1 d_{1t} + \beta_2 d_{2t} + \beta_3 d_{3t} + u_t,$$

The d_{it} are the observations on the dummy variables. For each t, one and only one of these is equal to 1. Then test the null hypothesis of homoskedasticity: $\mathrm{E}(u_t^2) = \sigma^2$, against the heteroskedastic alternative that

$$\mathrm{E}(u_t^2) = \gamma_1 d_{1t} + \gamma_2 d_{2t} + \gamma_3 d_{3t}.$$

Report P values for F and nR^2 tests.

Obtain a bootstrap P value for this test. Explain carefully how you set up the bootstrap DGP, and to what extent it satisfies the two Golden Rules of bootstrapping.

7. Consider the following set of three linear regressions:

$$\begin{aligned}
\boldsymbol{y}_1 &= \boldsymbol{x}_{11}\beta_1 + \boldsymbol{x}_{12}\beta_2 + \boldsymbol{u}_1;\\
\boldsymbol{y}_2 &= \boldsymbol{x}_{23}\gamma_1 + \boldsymbol{x}_{24}\gamma_2 + \boldsymbol{u}_2;\\
\boldsymbol{y}_3 &= \boldsymbol{x}_{35}\delta_1 + \boldsymbol{x}_{36}\delta_2 + \boldsymbol{u}_3.
\end{aligned}$$

Here, $\boldsymbol{y}_i$ is an n_i–vector, $i = 1, 2, 3$ et $\boldsymbol{x}_{ij}$ is also an n_i–vector, $i = 1, 2, 3$, $j = 1, \ldots, 6$. If these three regressions are stacked into one big regression with $n = n_1 + n_2 + n_3$ observations, we have

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \boldsymbol{y}_3 \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_{11} & \boldsymbol{x}_{12} & \mathbf{O} & \mathbf{O} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \boldsymbol{x}_{23} & \boldsymbol{x}_{24} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \mathbf{O} & \boldsymbol{x}_{35} & \boldsymbol{x}_{36} \end{bmatrix} \boldsymbol{\theta} + \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \boldsymbol{u}_3 \end{bmatrix}$$

What are the components of the parameter vector $\boldsymbol{\theta}$? Here are the results from running the stacked regression:

```
Ordinary Least Squares:

Variable   Parameter estimate  Standard error  T statistic

X1           0.8844155           0.0674323      13.1156061
X2           1.9587237           0.1322871      14.8066116
X3           0.9680135           0.2517097       3.8457543
X4          -1.8890838           0.0953457     -19.8129985
X5          -1.3769082           2.4581225      -0.5601463
X6           3.2760938           1.7775009       1.8430898

Number of observations = 300
Number of regressors = 6
Sum of squared residuals = 1173754.2783340
Explained sum of squares = 17805601.6102454
Estimate of residual variance (with d.f. correction) = 3992.3614909
R squared (uncentred) = 0.9381563
```

Suppose that $\boldsymbol{u}_i \sim \mathrm{N}(\mathbf{0}, \sigma^2\mathbf{I})$. Test the hypotheses that $\beta_1 = \gamma_1$ and that $\beta_2 = \delta_2$. For each test, explain how you computed a test statistic, and what its distribution should be under the null hypothesis. Next, it was desired to test the joint hypothesis constituted by these two equality restrictions. To this end, a regression that respects this joint hypothesis was set up, and gave the following results:

```
Ordinary Least Squares:

Variable   Parameter estimate  Standard error  T statistic

Z1          0.8871208           0.0648717       13.6749992
Z2          1.9565546           0.1285511       15.2200578
Z3         -1.8926384           0.0945299      -20.0215927
Z4          0.4470136           0.1933415        2.3120416

Number of observations = 300
Number of regressors = 4
Sum of squared residuals = 1176376.9482211
Explained sum of squares = 17802978.9403583
Estimate of residual variance (with d.f. correction) = 3974.2464467
R squared (uncentred) = 0.9380181
```

Explain how the four regressors, `Z1`, `Z2`, `Z31`, and `Z4` were constructed. What is the result of the joint hypothesis test? What test statistic did you use, and what is its null distribution?