Econometrics

Economics 662 D1

Your completed exam should be uploaded to myCourses by 13.00 on December 11. 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.

Please take note of the McGill Academic Integrity Statement that was written out in full on the midterm exam.

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.

This exam comprises 6 pages, including the cover page

1. Consider two linear regressions, one restricted and the other unrestricted:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}, \text{ and}$$
$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \boldsymbol{u}.$$

Show that, in the case of mutually orthogonal regressors, with $\boldsymbol{X}^\top\boldsymbol{Z} = \mathbf{O}$, the estimates of $\boldsymbol{\beta}$ from the two regressions are identical.

If $\boldsymbol{A}$ is a symmetric positive definite $k \times k$ matrix, then $\mathbf{I} - \boldsymbol{A}$ is positive definite if and only if $\boldsymbol{A}^{-1} - \mathbf{I}$ is positive definite, where $\mathbf{I}$ is the $k \times k$ identity matrix. Prove this result by considering the quadratic form $\boldsymbol{x}^\top(\mathbf{I} - \boldsymbol{A})\boldsymbol{x}$ and expressing $\boldsymbol{x}$ as $\boldsymbol{R}^{-1}\boldsymbol{z}$, where $\boldsymbol{R}$ is a symmetric matrix such that $\boldsymbol{A} = \boldsymbol{R}^2$.

Extend the above result to show that, if $\boldsymbol{A}$ and $\boldsymbol{B}$ are symmetric positive definite matrices of the same dimensions, then $\boldsymbol{A} - \boldsymbol{B}$ is positive definite if and only if $\boldsymbol{B}^{-1} - \boldsymbol{A}^{-1}$ is positive definite.

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 IV estimator

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

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. It frequently happens that, in a linear regression model, the matrix $\boldsymbol{X}$ of regressors can be partitioned into two groups of regressors, giving rise to two matrices, $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, of dimensions $n \times k_1$ and $n \times k_2$ respectively. The $n \times (k_1 + k_2)$ matrix $\boldsymbol{X}$ is just $[\boldsymbol{X}_1\ \boldsymbol{X}_2]$. Denote by $\boldsymbol{P_X}$, $\boldsymbol{P}_1$, and $\boldsymbol{P}_2$ the $n \times n$ orthogonal projection matrices on to the linear spans of the columns of $\boldsymbol{X}$, $\boldsymbol{X}_1$, and $\boldsymbol{X}_2$ respectively. Show that $\boldsymbol{P_X P}_1 = \boldsymbol{P}_1 \boldsymbol{P_X}$, and use this result to show that $\boldsymbol{P_X} - \boldsymbol{P}_1$ is an orthogonal projection matrix, that is, it is symmetric and idempotent.

Show that $\boldsymbol{P_X} - \boldsymbol{P}_1$ is in fact equal to $\boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}$, the orthogonal projection on to the span of the columns of the matrix $\boldsymbol{M}_1\boldsymbol{X}_2$, where $\boldsymbol{M}_1 = \mathbf{I} - \boldsymbol{P}_1$. What is the dimension of the image of $\boldsymbol{P_X} - \boldsymbol{P}_1$?

4. In the linear regression

$$\boldsymbol{y} = \boldsymbol{Z\beta}_1 + \boldsymbol{Y\beta}_2 + \boldsymbol{u}, \tag{1}$$

it is supposed that the variables in the $n \times k_1$ matrix $\boldsymbol{Z}$ are exogenous or predetermined, while those in the $n \times k_2$ matrix $\boldsymbol{Y}$ may be correlated with the disturbances in the vector $\boldsymbol{u}$. An $n \times l$ matrix $\boldsymbol{W}$, $l \geq k \equiv k_1 + k_2$, of exogenous or predetermined instrumental variables is available, with k_1 of its columns being those of $\boldsymbol{Z}$.

An easy way of performing a Durbin-Wu-Hausman (DWH) test is to run the testing regression

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{P_W Y\delta} + \boldsymbol{u}, \tag{2}$$

where $\boldsymbol{X} = [\boldsymbol{Z}\ \boldsymbol{Y}]$, and $\boldsymbol{\beta} = [\boldsymbol{\beta}_1 \vdots \boldsymbol{\beta}_2]$; see section 8.7 of the revised textbook. The test takes the form of any test of the (artificial) null hypothesis that $\boldsymbol{\delta} = \mathbf{0}$.

Show that if, instead of (2), the regression

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{M_W Y\delta} + \boldsymbol{u} \tag{3}$$

is run, the sum of squared residuals, and also therefore the F statistic for the hypothesis that $\boldsymbol{\delta} = \mathbf{0}$, is identical to that from (2). Show further that the (OLS) estimator of $\boldsymbol{\beta}$ from (3) is numerically identical to the estimator obtained when (1) is estimated by instrumental variables with $\boldsymbol{W}$ as the matrix of instruments.

Using seasonally adjusted quarterly data for the logarithm of the real money supply, m_t, real GDP, y_t, and the 3-month Treasury Bill rate, r_t, for Canada for the period 1967:1 to 1998:4, an econometrician ran the regression

$$m_t = \beta_1 + \beta_2 r_t + \beta_3 y_t + \beta_4 m_{t-1} + \beta_5 m_{t-2} + u_t. \tag{4}$$

The results were as follows:

```
Ordinary Least Squares:

Variable    Parameter estimate  Standard error  T statistic

constant       -0.258645           0.082851      -3.121806
r              -0.003783           0.000728      -5.196030
y               0.068132           0.012987       5.246303
lag(1,m)        1.404989           0.076522      18.360598
lag(2,m)       -0.459048           0.072441      -6.336865

Number of observations = 124   Number of estimated parameters = 5
Mean of dependent variable = 11.253704
Sum of squared residuals = 0.033514
Explained sum of squares = 15717.143072
Estimate of residual variance
  (with d.f. correction) = 0.000282
Mean of squared residuals = 0.000270
Standard error of regression = 0.016782
R squared (uncentred) = 0.999998  (centred) = 0.997440
```

A doubt arose as to the exogeneity of the variable r_t, and so it was decided to perform a DWH test of the null hypothesis of the exogeneity of r_t. To this end, the first two lags of r_t were used as instrumental variables, on the assumption that, as they are predetermined at time t, they are valid instruments. First, r_t was regressed on the regressors in (4), other than r_t itself, along with the two lags of r_t. The residuals from this regression were then used as an additional regressor (`Mwr`) in (4), with the following results:

```
Ordinary Least Squares:

Variable    Parameter estimate  Standard error  T statistic

constant       -0.263146           0.084122      -3.128136
r              -0.003933           0.000844      -4.659799
y               0.070097           0.014166       4.948203
lag(1,m)        1.396969           0.080072      17.446394
lag(2,m)       -0.452827           0.074799      -6.053892
Mwr             0.000598           0.001687       0.354242

Number of observations = 124   Number of estimated parameters = 6
Mean of dependent variable = 11.253704
Sum of squared residuals = 0.033478
Explained sum of squares = 15717.143108
Estimate of residual variance
  (with d.f. correction) = 0.000284
Mean of squared residuals = 0.000270
Standard error of regression = 0.016844
R squared (uncentred) = 0.999998  (centred) = 0.997443
```

What test statistic could the econometrician validly use for the DWH test? How many degrees of freedom should it have? Can you give its numerical value? Do you think it likely that the null would be rejected?

It has been emphasised that whenever an IV estimation is done of an over-identified model, the over-identifying restrictions should be tested, for instance, by use of a Sargan test. How many degrees of freedom would this test have in the present case?

The residuals of the IV estimation (of which the results are not given here) were regressed on the full set of instruments, with the following results:

```
Ordinary Least Squares:

Variable    Parameter estimate  Standard error  T statistic

constant        0.043901           0.079645       0.551209
lag(1,r)       -0.003946           0.001355      -2.910885
lag(2,r)        0.005334           0.001352       3.945743
y              -0.020157           0.013434      -1.500417
lag(1,m)        0.042306           0.075254       0.562174
lag(2,m)       -0.023524           0.070469      -0.333818

Number of observations = 124   Number of estimated parameters = 6
Mean of dependent variable = 0.000000
Sum of squared residuals = 0.029497
Explained sum of squares = 0.004029
Estimate of residual variance
  (with d.f. correction) = 0.000250
Mean of squared residuals = 0.000238
Standard error of regression = 0.015811
R squared (uncentred) = 0.120174  (centred) = 0.120174
```

Obtain the test statistic for at least one version of the Sargan test. Do you think it likely that the null hypothesis according to which the over-identifying restrictions are satisfied would be rejected? If so, how could you interpret this result?

5. In the file

https://russell-davidson.arts.mcgill.ca/e662/e662.dec20.dat

there are 100 observations on four variables, $\boldsymbol{y}$, $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, and $\boldsymbol{x}_3$. Regress $\boldsymbol{y}$ on a constant, $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, and $\boldsymbol{x}_3$.

Test the null hypothesis of the homoskedasticity of the disturbances in this regression against an alternative of heteroskedasticity, using as explanatory variables for the skedastic function the squares of the three non-constant regressors. Compute first an asymptotic F statistic, and then compute the corresponding P value.

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.

6. Consider a linear regression in which the dependent variable $\boldsymbol{y}$ is regressed on a constant and a non-constant regressor $\boldsymbol{x}$:

$$\boldsymbol{y} = \alpha\boldsymbol{\iota} + \beta\boldsymbol{x} + \boldsymbol{u}. \tag{5}$$

(As usual, $\boldsymbol{\iota}$ is a vector of ones.) Let $\boldsymbol{w}$ denote the variable $\boldsymbol{y}$ centred, so that $w_t = y_t - \bar{y}$, where $\bar{y}$ is the mean of the elements of the vector $\boldsymbol{y}$. Similarly, let $\boldsymbol{z}$ be $\boldsymbol{x}$ centred. In which ones of the following regressions:

$$\begin{aligned}
\boldsymbol{w} &= \beta\boldsymbol{z} + \boldsymbol{u},\\
\boldsymbol{w} &= \alpha\boldsymbol{\iota} + \beta\boldsymbol{z} + \boldsymbol{u},\\
\boldsymbol{y} &= \alpha\boldsymbol{\iota} + \beta\boldsymbol{z} + \boldsymbol{u},\\
\boldsymbol{y} &= \beta\boldsymbol{z} + \boldsymbol{u},\\
\boldsymbol{y} &= \beta\boldsymbol{x} + \boldsymbol{u}.
\end{aligned}$$

is the estimate of β equal to the one found from regression (5)? Explain why for each of the five regressions. In the two cases in which α is estimated, how is the estimate for α related to that from (5)? Why?

7. The disturbances of the linear regression

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u} \tag{6}$$

follow an MA(1) process (moving average of order one), as follows:

$$\begin{aligned}
u_1 &= \varepsilon_1;\\
u_t &= \varepsilon_t - \alpha\varepsilon_{t-1}, \quad t > 1,
\end{aligned}$$

where the ε_t are white-noise innovations. Show that the covariance matrix of the vector $\boldsymbol{u}$ is

$$\boldsymbol{\Omega} \equiv \sigma^2 \begin{bmatrix} 1 & -\alpha & 0 & 0 & \dots & 0\\ -\alpha & 1+\alpha^2 & -\alpha & 0 & \dots & 0\\ 0 & -\alpha & 1+\alpha^2 & -\alpha & \dots & 0\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & 0 & \dots & 1+\alpha^2 \end{bmatrix}.$$

where σ^2 is the variance of the ε_t.

Next show that $\boldsymbol{\Omega} = \boldsymbol{A}\boldsymbol{A}^\top$, where the matrix $\boldsymbol{A}$ is such that all of its diagonal elements are equal to 1, the other elements being zero except for those of the diagonal directly below the principal diagonal, which are all equal to α. Explicitly,

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0\\ -\alpha & 1 & \dots & 0 & 0\\ 0 & -\alpha & \dots & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \dots & -\alpha & 1 \end{bmatrix}.$$

If one wishes to use generalised least squares (GLS) to estimate the regression (6), one needs a matrix $\boldsymbol{\Psi}$ such that $\boldsymbol{\Psi}\boldsymbol{\Psi}^\top = \boldsymbol{\Omega}^{-1}$. Show that a possible choice is $\boldsymbol{\Psi} = (\boldsymbol{A}^\top)^{-1}$.