Queen's University
School of Graduate Studies and Research
Department of Economics

Economics 850 **Econometrics I** **Fall, 2016**

Professor James MacKinnon

Final Examination

December 15, 2016. Time Limit: 3 hours

Please answer any **four (4)** of the following **six (6)** questions. Each question has **four parts** and is worth 25% of the final mark. Before deciding which questions to answer, it would be a very good idea to read every question carefully.

Note: A table with some critical values of the χ^2 distribution appears at the end of the examination.

1. Suppose you estimate the linear regression model

$$\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{u} \tag{1}$$

by ordinary least squares, when the data are actually generated by the DGP

$$\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1^0 + \boldsymbol{u}, \quad \mathrm{E}(\boldsymbol{u}\boldsymbol{u}^\top) = \boldsymbol{\Omega}. \tag{2}$$

There are n observations, and $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ have k_1 and k_2 columns, respectively, with $k = k_1 + k_2$. The $n \times n$ matrix $\boldsymbol{\Omega}$ is known to be diagonal.

a) Under standard assumptions, what is the asymptotic distribution of $\hat{\boldsymbol{\beta}}_1$, the OLS estimator of $\boldsymbol{\beta}_1$ from regression (1)? State the asymptotic distribution formally in terms of a random vector that is $O_p(1)$.

b) Consider the matrices

$$\boldsymbol{S}_{11} \equiv \operatorname*{plim}_{n\to\infty} \frac{1}{n}\boldsymbol{X}_1^\top\boldsymbol{X}_1, \quad \boldsymbol{S}_{22} \equiv \operatorname*{plim}_{n\to\infty} \frac{1}{n}\boldsymbol{X}_2^\top\boldsymbol{X}_2, \text{ and } \boldsymbol{S}_{12} \equiv \operatorname*{plim}_{n\to\infty} \frac{1}{n}\boldsymbol{X}_1^\top\boldsymbol{X}_2.$$

What assumptions must you make about these matrices in order to obtain the result you stated in part a)?

c) In practice, how would you estimate the covariance matrix of $\hat{\boldsymbol{\beta}}_1$? How would you estimate the covariance matrix of $\tilde{\boldsymbol{\beta}}_1$, the OLS estimator of $\boldsymbol{\beta}_1$ subject to the restriction that $\boldsymbol{\beta}_2 = \boldsymbol{0}$?

d) Explain how you would test the hypothesis that $\boldsymbol{\beta}_2 = \boldsymbol{0}$ using an asymptotic test. Write down the test statistic you would use and state what distribution it would follow asymptotically under the null.

Continued on next page ...

2. You have a sample of 109,813 individuals who live in one of 12 jurisdictions. Let y_{gi} denote the value of the dependent variable for individual i in jurisdiction g and $\boldsymbol{X}_{gi}$ denote a row vector of values of exogenous variables, which include a constant term but do not include jurisdiction fixed effects. You are interested in the effect of a treatment dummy z_{gi}, which is 1 for 11,235 observations and 0 for the remaining 98,578 observations. Only five jurisdictions are "treated," so that the value of z_{gi} is 0 for every observation in 7 of the 12 jurisdictions.

a) Suppose you run an OLS regression of the y_{gi} on the $\boldsymbol{X}_{gi}$ and on z_{gi}. The heteroskedasticity-robust t statistic for the coefficient on z_{gi} to equal 0 is 8.74, and the cluster-robust one (clustering at the jurisdiction level) is 5.93. What would you conclude?

b) Would it make sense to add jurisdiction fixed effects to this regression? How many of them could you add? How would you test the joint hypothesis that all the fixed effects are zero?

c) In the regression with fixed effects, the heteroskedasticity-robust t statistic for the coefficient on z_{gi} to equal 0 is 3.64, and the cluster-robust one is 2.73. What would you conclude now?

d) A referee asks you to compute a bootstrap P value for the hypothesis that the coefficient on z_{gi} equals 0 in the regression of part c). Which test statistic would you bootstrap? Briefly explain how you would do so.

3. You have two different cross-section datasets, with n_1 and n_2 observations, respectively. For each of them, you wish to estimate a model of the form

$$y_{ti} = \beta_{1i} + \beta_{2i} x_{ti}^{\gamma} + u_{ti}, \quad i = 1, 2. \tag{3}$$

Notice that the parameter γ is constrained to be the same for both sets of data, but the other parameters are allowed to be different.

a) Explain how you would estimate γ efficiently and construct a .95 asymptotic confidence interval for it under the strong assumption that the disturbances u_{ti} are distributed as $\text{IID}(0, \sigma^2)$.

b) Explain how you would estimate γ efficiently under the weaker assumption that $u_{ti} \sim \text{IID}(0, \sigma_i^2)$, $i = 1, 2$. How would you construct a .95 asymptotic confidence interval for γ in this case?

c) Now suppose that $u_{ti} \sim \text{IID}(0, \omega_{ti}^2)$, where the ω_{ti} are unknown. Could you still use the estimation method of part a)? Could you still use the estimation method of part b)? Explain how you would construct a .95 asymptotic confidence interval for γ based on one of these sets of estimates.

d) Briefly explain how you would construct a .95 bootstrap confidence interval for γ under the assumptions of part c). Do so for the same estimation method you used to construct the asymptotic confidence interval.

Continued on next page . . .

4. Consider the linear simultaneous equations model

$$\boldsymbol{y}_1 = \beta \boldsymbol{y}_2 + \boldsymbol{Z\gamma} + \boldsymbol{u}, \tag{4}$$
$$\boldsymbol{y}_2 = \boldsymbol{W\pi} + \boldsymbol{v}, \tag{5}$$

where $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are n–vectors of observations on endogenous variables, $\boldsymbol{Z}$ and $\boldsymbol{W}$ are $n \times k$ and $n \times l$ matrices of observations on exogenous or predetermined variables, with every column of $\boldsymbol{Z}$ belonging to $\mathcal{S}(\boldsymbol{W})$, and $\boldsymbol{u}$ and $\boldsymbol{v}$ are n–vectors of disturbances. Assume that $\mathrm{E}(u_t v_s) = 0$ for $t \neq s$ and $\mathrm{E}(u_t v_t) = \rho\sigma_u\sigma_v$, where u_t and v_s are elements of $\boldsymbol{u}$ and $\boldsymbol{v}$. The rest of the notation should be obvious.

a) Under what circumstances will OLS estimation of equation (4) yield a consistent estimate of β? Explain how you could test whether $\hat\beta_{\mathrm{OLS}}$ is consistent.

b) The generalized IV estimator for equation (4) can be obtained by minimizing a certain criterion function. Write down this function explicitly. Then write down $\hat\beta_{\mathrm{IV}}$, the generalized IV estimator of β.

c) Suppose that $n = 1536$, $k = 4$, $l = 8$, $\hat\sigma_u = 5.23$, and the minimized value of the criterion function from part b) is 183.85. In this case, how many overidentifying restrictions are there? Would you reject the hypothesis that they are satisfied at the .05 level? Explain briefly.

d) Explain how you could use equation (5) to check whether the instruments are weak by calculating a certain test statistic. How many degrees of freedom would this test statistic have under the assumptions of part c)? What would you conclude about the reliability of conventional inference based on $\hat\beta_{\mathrm{IV}}$ if the test statistic were 2.71? What would you conclude if the test statistic were 15.62? Explain briefly.

5. Consider the dynamic linear regression model

$$y_t = \beta_1 + \beta_2 x_t + \beta_3 y_{t-1} + u_t, \quad u_t \sim \mathrm{IID}(0, \sigma^2), \tag{6}$$

where there are n observations, x_t is an exogenous variable, and it is assumed that $|\beta_3| < 1$.

a) Are the OLS estimates of β_1 and β_2 biased or unbiased? Are they consistent or inconsistent? Explain carefully.

b) Would your answer to part a) change if the IID assumption were replaced by the assumption that the u_t follow the AR(1) process

$$u_t = \rho u_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{IID}(0, \sigma_\varepsilon^2), \tag{7}$$

with $|\rho| < 1$? Explain why or why not.

Continued on next page ...

c) Suppose you estimate the model (6) by OLS and then regress the OLS residuals $\hat{u}_t$ on themselves lagged once. What would you conclude if the t statistic for the only coefficient in this regression were 1.42? What would you conclude if it were 2.63?

d) Explain briefly how you would obtain consistent and asymptotically efficient estimates of the model consisting of equations (6) and (7), and how you would estimate the standard errors of the parameter estimates $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$, and $\hat{\rho}$.

6. This question concerns the univariate nonlinear regression model

$$y_t = x_t(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) + u_t, \quad u_t \sim \text{IID}(0, \sigma^2), \tag{8}$$

where $\boldsymbol{\beta}_1$ is a k_1–vector and $\boldsymbol{\beta}_2$ is a k_2–vector of unknown parameters, and there are n observations.

a) Suppose this model is estimated by nonlinear least squares. Briefly describe three different ways that you could test the hypothesis $\boldsymbol{\beta}_2 = \mathbf{0}$. One of them should require estimation only under the null hypothesis, and one of them should require estimation only under the alternative hypothesis.

b) Two of the procedures you described in part a) can readily be modified to allow for heteroskedasticity of unknown form. Which two are they? Briefly explain how to modify one of the two procedures.

c) Suppose this model is estimated unrestrictedly by instrumental variables using the $n \times m$ matrix of instruments $\boldsymbol{W}$, with $m = k_1 + k_2$. What criterion function must be minimized to obtain generalized IV estimates? What do you know about the minimized value of this criterion function? Explain.

d) How would your answer to part c) change if $m = k_1 + k_2 + 3$? How many over-identifying restrictions are there for the unrestricted model? How many are there for the restricted model? How could you test these over-identifying restrictions in each case?

Table 1. Some Critical Values of the χ^2 Distribution

D.F. / Level	.10	.05	.025	.01
1	2.706	3.841	5.024	6.635
2	4.605	5.991	7.378	9.210
3	6.251	7.815	9.348	11.345
4	7.779	9.488	11.143	13.277
5	9.236	11.070	12.833	15.086
6	10.645	12.592	14.449	16.812