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

**Economics 850** | **Econometrics I** | **Fall, 2023**

**Professor James MacKinnon**

# Final Examination

December 18, 2023. Time: 3 hours

**Notes:** The examination is in two parts. Please answer the only question in Part I and three (3) questions from Part II. Tables with some critical values of the $\chi^2$ and Student's $t$ distributions appear at the end of the examination.

**Part I.** Please answer the following question, which is worth 28% of the final mark.

**1.** Consider the linear regression model

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + u_i,$$

where there are 87 observations, the regressors are assumed to be exogenous, and the four $\beta_j$ parameters are unknown.

a) Write down $\hat{\beta}_4$, the OLS estimate of $\beta_4$, using a compact and readable notation. **Hint:** You may wish to define some vectors and matrices.

b) Explain how you would estimate the standard error of $\hat{\beta}_4$ and obtain a 95% confidence interval for $\beta_4$ under the assumption that the $u_i$ are independently and identically distributed. Would this interval be exact without making additional assumptions?

c) Now suppose that $\mathrm{E}(u_i^2) = \exp(\gamma_0 + \gamma_1 x_{3i})$, with the $\gamma_j$ unknown. Would the confidence interval from part b) be valid, even asymptotically? Are there any restrictions on the $\gamma_j$ under which it would be (asymptotically) valid? Are there circumstances in which it would be asymptotically valid even without restrictions on the $\gamma_i$?

d) Briefly explain two different ways to obtain the standard error of $\hat{\beta}_4$ under the assumptions of part c). How would you use each of these standard errors to form a confidence interval? Would these intervals be asymptotically valid? Would they be valid for the actual sample in this case? Which of the two intervals would you expect to be longer?

**Part II.** Please answer three (3) of the following four (4) questions. Each question has four (4) parts and is worth 24% of the final mark.

Continued on next page . . .

**2.** Consider the linear regression model

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i,$$

which is to be estimated using a sample of 73 observations. The regressors $x_{2i}$ and $x_{3i}$ are assumed to be exogenous. You are interested in the parameter $\gamma \equiv \beta_2/\beta_3$.

a) Explain how you would obtain an estimate $\hat{\gamma}$ and an asymptotically valid standard error $s(\hat{\gamma})$ analytically (i.e., without doing any simulations) under the assumption that $\mathrm{E}(u_i u_j) = 0$ for $i \neq j$ and $\mathrm{E}(u_i^2 \,|\, \boldsymbol{X}_i) = \sigma_i^2$. Do not waste time doing the algebra.

b) Explain how you would perform a bootstrap test of the hypothesis that $\gamma = 3$ under the assumptions of part a). Would your bootstrap DGP incorporate any restrictions? Explain precisely how it would work.

c) Explain how you would obtain a bootstrap standard error $s^*(\hat{\gamma})$ under the assumptions of part a) and how you could use that standard error to form a 95% confidence interval. Would your bootstrap DGP incorporate any restrictions? Explain precisely how it would work.

d) Explain how you would construct a 95% studentized bootstrap confidence interval for $\gamma$ under the assumptions of part a). Would your bootstrap DGP incorporate any restrictions? Explain precisely how it would work.

**3.** Consider the linear regression model

$$y_i = \boldsymbol{X}_i \boldsymbol{\beta} + u_i,$$

where the $u_i$ are independently but not identically distributed. Specifically, it is assumed that $\mathrm{E}(u_i \,|\, \boldsymbol{X}_i) = 0$, and $\mathrm{E}(u_i^2 \,|\, \boldsymbol{X}_i) = \sigma_i^2$.

a) Write the OLS estimator $\hat{\boldsymbol{\beta}}$ as a function of $\boldsymbol{\beta}_0$, the true value of $\boldsymbol{\beta}$, $\boldsymbol{X}$, and $\boldsymbol{u}$, where $\boldsymbol{X}$ has typical row $\boldsymbol{X}_i$, and $\boldsymbol{u}$ has typical element $u_i$. The expression you obtain should have two terms, only one of which is stochastic. Of what order in the sample size $N$ is the stochastic term under standard assumptions? Explain.

b) Under the assumptions made so far, will $\hat{\boldsymbol{\beta}}$ be unbiased? Will it be consistent? What can you say about its asymptotic distribution? Do you need to make any additional assumptions to answer these questions? Explain briefly.

c) Suppose now that the $u_i$ are not assumed to be independent. Of course, the $\boldsymbol{X}_i$ were never assumed to be independent. Do your answers to part b) change in any way as a result of these new assumptions?

d) Consider the special case in which

$$u_i = v + w_i, \quad v \sim \mathrm{N}(0, \omega^2), \;\; \mathrm{E}(w_i w_j) = 0 \;\text{ for all } i, j.$$

Continued on next page ...

What can you say about the stochastic term in the expression for $\hat{\boldsymbol{\beta}}$ that you obtained in part a)? Is $\hat{\boldsymbol{\beta}}$ consistent in this case? Explain.

**4.** You are given a sample of 20,781 stroke patients treated at 122 hospitals located in 21 communities, along with data on the characteristics of the patients and the hospitals. Let $y_{hi}$ denote the number of days in hospital for patient $i$ in hospital $h$, and let $\boldsymbol{X}_{hi}$ be a vector of patient and hospital characteristics. Some hospitals claim to base their operations on "patient-centred care," and others do not. The variable $z_{hi}$ is equal to 1 if hospital $h$ made such a claim when patient $i$ was treated, and equal to 0 otherwise.

Suppose you regress the $y_{hi}$ on the $z_{hi}$, the $\boldsymbol{X}_{hi}$, and a vector of community fixed effects. The coefficient on $z$ is $\beta$. The OLS estimate of $\beta$, say $\hat{\beta}$, is $-0.734$, which suggests that stroke patients are discharged about 3/4 of a day earlier in hospitals with patient-centred care.

a) You compute the standard error of $\hat{\beta}$ in several different ways. The $HC_1$, $HC_2$, and $HC_3$ standard errors are 0.112, 0.116, and 0.122, respectively. When you cluster by hospital, the $CV_1$ standard error is 0.226, the $CV_2$ standard error is 0.245, and the $CV_3$ standard error is 0.268. Which of these standard errors seems most believable? Explain why, and also explain precisely how to compute this standard error.

b) In view of the number of hospitals, it may seem odd that the various cluster-robust standard errors vary so much. What is the most likely explanation for this? What quantities could you compute and study in order to provide evidence for this explanation?

c) After you submit your paper on patient-centred care to a journal, a referee argues that you should have clustered by community instead of by hospital. You therefore compute three more cluster-robust standard errors, this time clustering by community. What would you do if these were a little smaller than the ones you obtained originally? What would you do if they were a little larger? What would you do if they were 30% to 40% larger? Would it make sense to use another method of inference to verify the results when you cluster by community?

d) Suppose the cluster-robust standard error that you prefer when you cluster by community is 0.366. Use this standard error and the table at end of the examination to obtain a 95% confidence interval for $\beta$. Does this interval suggest that the hypothesis $\beta = 0$ can be rejected at the .05 level?

Continued on next page ...

**5.** Consider the linear simultaneous equations model

$$\boldsymbol{y}_1 = \beta \boldsymbol{y}_2 + \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{u}, \tag{1}$$
$$\boldsymbol{y}_2 = \boldsymbol{W}\boldsymbol{\pi} + \boldsymbol{v} = \boldsymbol{W}_1\boldsymbol{\pi}_1 + \boldsymbol{Z}\boldsymbol{\pi}_2 + \boldsymbol{v}, \tag{2}$$

where $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are $n$–vectors of observations on endogenous variables, $\boldsymbol{Z}$ is an $n \times k$ matrix of observations on exogenous or predetermined variables, $\boldsymbol{W} \equiv [\boldsymbol{W}_1 \;\; \boldsymbol{Z}]$ is an $n \times l$ matrix of observations on exogenous or predetermined variables, and $\boldsymbol{u}$ and $\boldsymbol{v}$ are $n$–vectors of homoskedastic 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 $\boldsymbol{v}$, and $\sigma_u^2$ and $\sigma_v^2$ are their variances. The rest of the notation should be obvious.

a) In general, will the OLS estimate of $\beta$ be unbiased? Will it be consistent? Is there any special case in which it will be consistent? Explain carefully.

b) Explain how to compute the generalized IV estimator $\hat{\beta}_{\mathrm{IV}}$. What can you say about the mean squared error of this estimator relative to the mean squared error of the OLS estimator $\hat{\beta}_{\mathrm{OLS}}$? What are the key features of the data-generating process (DGP) that explain the relationship between the two MSE values? Explain.

c) How many overidentifying restrictions are there? Explain how to test them using an asymptotic test. What would you conclude if $n = 350$, $k = 12$, $l = 15$, and the test statistic (in $\chi^2$ form) were 6.46? Note that there is a table of the $\chi^2$ distribution at the end of the examination.

d) Suppose the $F$ statistic for $\boldsymbol{\pi}_1 = \boldsymbol{0}$ in equation (2) were 2.24. Would you expect the standard error of $\hat{\beta}_{\mathrm{IV}}$ to be similar to the standard error of $\hat{\beta}_{\mathrm{OLS}}$ in this case? Would you expect the finite-sample distribution of $\hat{\beta}_{\mathrm{IV}}$ to be well approximated by its asymptotic distribution? Would your answers to either or both of these questions change if the value of this $F$ statistic were 17.58?

Continued on next page ...

**Table 1. Upper-Tail Critical Values of the $\chi^2$ 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 |

**Table 2. Two-Tail Critical Values of the Student's $t$ Distribution**

| D.F. / Level | .10 | .05 | .025 | .01 |
|---|---|---|---|---|
| 10 | 1.812 | 2.228 | 2.634 | 3.169 |
| 15 | 1.753 | 2.131 | 2.490 | 2.947 |
| 16 | 1.746 | 2.120 | 2.473 | 2.921 |
| 17 | 1.740 | 2.110 | 2.458 | 2.898 |
| 18 | 1.734 | 2.101 | 2.445 | 2.878 |
| 19 | 1.729 | 2.093 | 2.433 | 2.861 |
| 20 | 1.725 | 2.086 | 2.423 | 2.845 |
| 21 | 1.721 | 2.080 | 2.414 | 2.831 |