G020: Mock Midterm Exam in Econometrics

2013-2014

Answer ALL questions. **Time allowed:** 1:50 hours

Question 1 (34 points)

Consider the linear regression model

$$y_i = \beta_1 + f_i \beta_2 + u_i,$$

where y_i is the height (in some units) of individual i, and f_i is a gender dummy, which takes values $f_i = 1$ for females and $f_i = 0$ for males. Let $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2)'$ be the OLS estimator of $\beta = (\beta_1, \beta_2)'$. In a concrete sample we observe $n_f = 100$ females and $n_m = 300$ males. The sample size is $n = n_f + n_m = 400$. The average of y_i in the female subsample is $\overline{y}_f = 5$. The average of y_i in the male subsample is $\overline{y}_m = 6$. Let $x_i = (1, f_i)$, and let y be the $n \times 1$ vector with entries y_i, and X be the $n \times 2$ matrix with rows x_i.

(a) Somebody proposes the alternative model specification $y_i = \gamma_1 + m_i \gamma_2 + \epsilon_i$, where $m_i = 1 - f_i$ is a male dummy. Let $\hat{\gamma} = (\hat{\gamma}_1, \hat{\gamma}_2)'$ be the corresponding OLS estimator. Find an expression for $\hat{\gamma}$ in terms of $\hat{\beta}$.

(b) Calculate the 2×2 matrix $X'X$, the 2×1 vector $X'y$, the 2×2 matrix $(X'X)^{-1}$, and the OLS estimator $\hat{\beta}$.

(c) Assume that errors are homoscedastic with $\mathrm{Var}(u_i | f_i) = 3$. Calculate the estimated variance-covariance matrix of $\hat{\beta}$ and the estimated standard error of $\hat{\beta}_2$.

For the following subquestion assume that you calculated the estimator $\hat{\beta}_2 = -2$ and that you calculated the estimator for the standard error of $\hat{\beta}_2$ to be equal to 0.4. (These are not the actual numbers you should have obtained above.)

(d) Consider the null hypothesis $H_0 : \beta_2 \geq 0$. Calculate the t-test statistics for testing H_0. Can you reject H_0 at 5% significance level using a large sample t-test? What is the critical value for this test? (Hint: the 97.5% quantile of $\mathcal{N}(0,1)$ is 1.96. The 95% quantile of $\mathcal{N}(0,1)$ is 1.64, the 90% quantile of $\mathcal{N}(0,1)$ is 1.28)

(e) Let $\tilde{y}_i = y_i - \frac{1}{n}\sum_{i=1}^n y_i$ and $\tilde{f}_i = f_i - \frac{1}{n}\sum_{i=1}^n f_i$. These are demeaned versions of the variables y_i and f_i. Let $\hat{\delta}$ be the OLS estimator obtained from regressing $\tilde{y}_i$ on $\tilde{f}_i$. What is the relationship between $\hat{\delta}$ and $\hat{\beta}_2$? Explain your answer (but no proof required).

Question 2 (33 points)

We observe two variables y_i and x_i such that the vector (y_i, x_i) is independent and identically distributed across observations $i = 1, \dots, n$. We assume that

$$\mathbb{E}(y_i \,|\, x_i) = \beta, \qquad \mathrm{Var}(y_i \,|\, x_i) = \gamma\, x_i, \qquad \mu_x = \mathbb{E}(x_i),$$

where β, γ and μ_x are unknown scalar parameters. We assume that $\gamma > 0$, $\mu_x > 0$, $\mathrm{Var}(x_i) > 0$, that $x_i > 0$ holds for all i, and that $\mathbb{E}(1/x_i)$ exists. This implies that $\mathbb{E}(1/x_i) > 1/\mu_x$. (the last statement follows by Jensen's inequality)

(a) Define $u_i = y_i - \beta$ and $\sigma^2 = \mathbb{E}(u_i^2)$. Find an expression for σ^2 in terms of the parameters γ and μ_x.

(b) Show that the estimator $\hat{\beta}_1 = \frac{1}{n}\sum_{i=1}^n y_i$ satisfies $\sqrt{n}(\hat{\beta}_1 - \beta) \Rightarrow \mathcal{N}(0, \Sigma_1)$ as $n \to \infty$. Find an expression for Σ_1. How is Σ_1 related to σ^2?

(c) Write down a consistent estimator for Σ_1. (no proof required)

(d) Consider the estimator

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n w_i y_i}{\sum_{i=1}^n w_i}, \qquad w_i = \frac{1}{x_i}.$$

Show that $\sqrt{n}(\hat{\beta}_2 - \beta) \Rightarrow \mathcal{N}(0, \Sigma_2)$ as $n \to \infty$. Find an expression for Σ_2. Compare Σ_1 and Σ_2. Which of the two estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ is asymptotically more efficient?

(e) Explain how Σ_2 can be consistently estimated. (no proof required)

Question 3 (33 points)

Consider the linear regression model with two regressors x_i and w_i^*

$$y_i = \beta_1 x_i + \beta_2 w_i^* + u_i,$$

where y_i and x_i are observed, but $\underline{w_i^* \text{ is unobserved}}$. Instead of w_i^* we only observe a variable w_i, which satisfies

$$w_i = w_i^* + v_i,$$

where v_i is unobserved measurement error. In addition, we observe one instrumental variable z_i. We make the following assumptions:

(i) Cross-sectional sampling: the vector $(y_i, x_i, w_i^*, w_i, z_i)$ is independent and identically distributed across $i = 1, \ldots, n$.

(ii) All variables are mean zero: $\mathbb{E}y_i = \mathbb{E}x_i = \mathbb{E}w_i^* = \mathbb{E}w_i = \mathbb{E}z_i = 0$.

(iii) Exogeneity: $\mathbb{E}(u_i \,|\, x_i, w_i, w_i^*, z_i) = 0$.

(iv) Classical measurement error: $\mathbb{E}(v_i \,|\, x_i, w_i^*) = 0$.

(v) Non-Collinearity: The 2×2 matrix $\mathbb{E}\left[(x_i, w_i)'(x_i, w_i)\right]$ is invertible.

We also assume that $\beta_2 \neq 0$, $\mathbb{E}(v_i^2) > 0$, and $\mathbb{E}(x_i w_i) \neq 0$. The main object of interest, which we want to estimate, is the parameter β_1. Answer the following questions:

(a) Consider the OLS estimator $\hat\beta_1$ obtained from regressing y_i on only x_i. What is the probability limit of $\hat\beta_1$ as $n \to \infty$? Are the above assumptions sufficient to guarantee that $\hat\beta_1$ is consistent for β_1?

(b) Now also assume that $\mathbb{E}(w_i^* z_i) = 0$ and $\mathbb{E}(x_i z_i) \neq 0$. Explain how β_1 can be consistently estimated under these additional assumptions. Provide an explicit formula for an estimator that is consistent for β_1 under these assumptions. (no proof required)

(c) Consider the OLS estimator $\hat\beta^* = (\hat\beta_1^*, \hat\beta_2^*)'$ from regressing y_i on x_i and w_i, i.e. from estimating the model $y_i = \beta_1^* x_i + \beta_2^* w_i + \epsilon_i$. What is the probability limit of $\hat\beta^*$ as $n \to \infty$? (Your answer can include an inverse matrix, which need not be further evaluated, yet.)

(d) Is the estimator $\hat\beta_1^*$ from part (c) consistent for β_1? Prove your answer. (For this you now need to evaluate your answer in (c) further.)

G020: Mock Midterm Exam in Econometrics

2014-2015

Answer ALL questions. **Time allowed:** 2:00 hours

Question 1 (34 points)

Consider the linear regression model

$$y_i = \beta_1 + w_i\beta_2 + u_i,$$

where y_i is the outcome variable, w_i is a scalar regressor, and u_i is the unobserved error term. We assume that $\mathbb{E}(u_i|w_i) = 0$ and $\mathbb{E}(u_i^2|w_i) = \sigma^2 > 0$. For an observed iid sample of (y_i, w_i) with sample size $n = 100$ one computes

$$\frac{1}{n}\sum_{i=1}^{n} y_i = 2, \qquad \frac{1}{n}\sum_{i=1}^{n} w_i = 3, \qquad \frac{1}{n}\sum_{i=1}^{n} y_i w_i = 1, \qquad \frac{1}{n}\sum_{i=1}^{n} w_i^2 = 10.$$

Let $y = (y_1, y_2, \ldots, y_n)'$, and let X be the $n \times 2$ matrix with rows $x_i = (1, w_i)$, $i = 1, \ldots, n$.

(a) Calculate the 2×2 matrix $X'X$, the 2×1 vector $X'y$, the 2×2 matrix $(X'X)^{-1}$, and the OLS estimator $\widehat{\beta} = (\widehat{\beta}_1, \widehat{\beta}_2)'$.

(b) Show that the OLS estimator $\widehat{\beta}$ is unbiased for $\beta = (\beta_1, \beta_2)'$.

(c) Calculate the OLS estimator for β_2 that is obtained from regressing y_i on w_i only (without including a constant). Is the resulting estimator unbiased for β_2? (no proof required, just a short explanation)

(d) Do you have enough information to calculate a consistent estimator for σ? Explain why or why not.

In the following assume that $\sigma^2 = 25$.

(e) Calculate the standard error of the OLS estimator $\widehat{\beta}_2$ obtained in (a). Calculate the t-test statistics for testing $H_0 : \beta_2 \leq 4$. Can you reject H_0 at 5% significance level using a large sample t-test? What is the critical value for this test? (The 97.5% quantile of $\mathcal{N}(0, 1)$ is 1.96. The 95% quantile of $\mathcal{N}(0, 1)$ is 1.64, the 90% quantile of $\mathcal{N}(0, 1)$ is 1.28)

Question 2 (33 points)

Let y_i and $x_i > 0$ be scalar random variables, and assume that

$$\mathbb{E}(y_i|x_i) = \beta, \qquad \mathrm{Var}(y_i|x_i) = \sigma^2 + \gamma x_i, \qquad \mathrm{Var}(x_i) > 0,$$

where β and $\sigma > 0$ and $\gamma > 0$ are unknown scalar parameters. These assumptions imply that $\mathbb{E}(y_i) = \beta$ and $\mathrm{Var}(y_i) = \sigma^2 + \gamma\mathbb{E}(x_i)$. We observe a random sample of y_i, x_i, with $i = 1, \ldots, n$. Let $\widehat{\beta} = \frac{1}{n}\sum_{i=1}^{n} y_i$.

(a) Show that $\sqrt{n}(\widehat{\beta} - \beta) \Rightarrow \mathcal{N}(0, \Sigma)$, and provide a formula for $\Sigma = \mathrm{AsyVar}(\sqrt{n}\,\widehat{\beta})$.

(b) Describe how σ^2 and γ can be consistently estimated (but no consistency proof required). We denote the resulting consistent estimators by $\widehat{\sigma}^2$ and $\widehat{\gamma}$ in the following.

(c) Consider $\widehat{\Sigma}_1 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\left(y_i - \widehat{\beta}\right)^2$ and $\widehat{\Sigma}_2 = \widehat{\sigma}^2 + \widehat{\gamma}\,\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i$. Are these consistent estimators for Σ? (explain your answer, but no proof required)

(d) Use $\widehat{\sigma}^2$ and $\widehat{\gamma}$ to provide an explicit formula for a consistent estimator $\widehat{\beta}_2$ of β such that $\mathrm{AsyVar}(\sqrt{n}\,\widehat{\beta}_2) < \mathrm{AsyVar}(\sqrt{n}\,\widehat{\beta})$. (no proof required, just provide $\widehat{\beta}_2$)

(e) Assume that the above condition on $\mathrm{Var}(y_i|x_i)$ is changed to $\mathrm{Var}(y_i|x_i) = \exp(x_i)$. In that case, is your $\widehat{\beta}_2$ from question (d) still consistent for β? Explain your answer, but no proof required.

Question 3 (33 points)

Consider the regression model

$$y_i = \beta_1 + \beta_2 w_i + u_i,$$

where y_i and w_i are observed scalar random variables, and u_i is an unobserved error term. We also observe a scalar instrumental variable v_i. Assume that (w_i, v_i, u_i), $i = 1, \ldots, n$, are iid draws from a population such that $\mathbb{E}(u_i|v_i) = 0$ and $\mathbb{E}(u_i^2|w_i, v_i) = \sigma^2$. Let $\beta = (\beta_1, \beta_2)'$, $x_i = (1, w_i)$ and $z_i = (1, v_i)$. Assume that the 2×2 matrices $\mathbb{E}(z_i' x_i)$ and $\mathbb{E}(z_i' z_i)$ and the corresponding sample analogs are all invertible.
Define $\widehat{x}_i = z_i \left(\sum_{i=1}^n z_i' z_i\right)^{-1} \sum_{i=1}^n z_i' x_i$, and consider the estimators

$$\widehat{\beta}_{\text{2SLS}} = \left(\sum_{i=1}^n \widehat{x}_i' \widehat{x}_i\right)^{-1} \sum_{i=1}^n \widehat{x}_i' y_i. \qquad \widehat{\beta}_{\text{IV}} = \left(\sum_{i=1}^n z_i' x_i\right)^{-1} \sum_{i=1}^n z_i' y_i.$$

(a) With the above definitions, is it true that $\widehat{\beta}_{\text{2SLS}} = \widehat{\beta}_{\text{IV}}$? Prove your answer.

(b) Prove that $\widehat{\beta}_{\text{IV}}$ consistent for β.

(c) Consistency of an IV estimator requires a relevance condition. Which of our above assumptions guarantees relevance of the instrument v_i? Explain.

(d) It can be shown that $\sqrt{n}(\widehat{\beta}_{\text{2SLS}} - \beta) \Rightarrow \mathcal{N}(0, \Sigma)$, as $n \to \infty$, with asymptotic variance-covariance given by $\Sigma = \sigma^2 \left(\frac{1}{n} \sum_{i=1}^n \widehat{x}_i' \widehat{x}_i\right)^{-1}$. Note that the question assumes homoscedasticity. Furthermore, in the second step of 2SLS we run an OLS regression of y_i on $\widehat{x}_i$, and the corresponding naive estimator for the OLS asymptotic variance-covariance matrix under homoscedasticity reads

$$\widehat{\Sigma} = \widehat{\sigma}^2 \left(\frac{1}{n} \sum_{i=1}^n \widehat{x}_i' \widehat{x}_i\right)^{-1}, \qquad \widehat{\sigma}^2 = \frac{1}{n-K} \sum_{i=1}^n \left(y_i - \widehat{x}_i \widehat{\beta}_{\text{2SLS}}\right)^2.$$

Is this naive estimator $\widehat{\Sigma}$ consistent for Σ? Explain your answer, but no proof required.

G020: Mock Midterm Exam in Econometrics

2015-2016

Answer ALL questions. **Time allowed:** 2:00 hours

Question 1 (30 points)

Consider the linear regression model with a constant regressor and one additional regressor

$$y_i = \beta_1 + w_i \beta_2 + u_i.$$

Assume that all variables have finite second moments, and that the errors u_i have mean zero, are independent of w_i, and have unknown variance $\sigma^2 = \mathbb{E}(u_i^2)$. We observe a random sample (y_i, w_i), $i = 1, \ldots, n$, of $n = 81$ observations. For this observed sample we find

$$\frac{1}{n}\sum_{i=1}^n w_i = 2, \quad \frac{1}{n}\sum_{i=1}^n y_i = 2, \quad \frac{1}{n}\sum_{i=1}^n w_i^2 = 5, \quad \frac{1}{n}\sum_{i=1}^n y_i w_i = 5, \quad \frac{1}{n}\sum_{i=1}^n y_i^2 = 30. \quad (1)$$

(a) Use the information in (1) to calculate the OLS estimates for β_1 and β_2.

(b) Three different estimators for σ^2 are given by

$$\widehat{\sigma}_A^2 = \frac{1}{n}\sum_{i=1}^n \widehat{u}_i^2, \qquad \widehat{\sigma}_B^2 = \frac{1}{n-1}\sum_{i=1}^n \widehat{u}_i^2, \qquad \widehat{\sigma}_C^2 = \frac{1}{n-2}\sum_{i=1}^n \widehat{u}_i^2,$$

where $\widehat{u}_i = y_i - \widehat{\beta}_1 - w_i \widehat{\beta}_2$. Which of these estimators for σ^2 is consistent? Which of these estimators for σ^2 is unbiased? No proof required.

(c) Use your result in (a) and the information in (1) to calculate $\widehat{\sigma}_A^2 = \frac{1}{n}\sum_{i=1}^n \widehat{u}_i^2$.

(d) Use your result in (c) and the information in (1) to calculate standard errors for $\widehat{\beta}_1$ and $\widehat{\beta}_2$.

Question 2 (35 points)

We observe a random sample (y_i, x_i, z_i), $i = 1, \ldots, n$, where n is the sample size, and y_i, x_i and z_i are three scalar variables with finite second moments. For simplicity we assume that $\mathbb{E}y_i = \mathbb{E}x_i = \mathbb{E}z_i = 0$. We consider the model

$$y_i = x_i\beta + u_i \ ,$$

where u_i is a mean zero error term and β is the parameter of interest. We also define $\gamma_{xx} = \mathrm{Var}(x_i)$, $\gamma_{xy} = \mathrm{Cov}(x_i, y_i)$, $\gamma_{xz} = \mathrm{Cov}(x_i, z_i)$ and $\gamma_{yz} = \mathrm{Cov}(y_i, z_i)$. We assume that $\gamma_{xx} > 0$.

(a) Write down consistent estimators for γ_{xx}, γ_{xy}, γ_{xz}, and γ_{yz}. Prove for one of these estimators that it is indeed consistent.

(b) Assume that x_i is exogenous, i.e. that $\mathbb{E}(x_i u_i) = 0$. Use this assumption to express β as a function of γ_{xx} and γ_{xy}. Use this expression for β and your result in (a) to provide a consistent estimator for β, which we denote by $\widehat{\beta}$. Prove that $\widehat{\beta}$ is indeed consistent.

(c) Under the assumptions in (b), show that $\widehat{\beta}$ is asymptotically normally distributed, i.e. show that $\sqrt{n}(\widehat{\beta} - \beta) \Rightarrow \mathcal{N}(0, \Sigma)$ as $n \to \infty$. Provide a formula for the asymptotic variance Σ.

(d) Assume that the result in (c) holds, that is, $\sqrt{n}(\widehat{\beta} - \beta) \Rightarrow \mathcal{N}(0, \Sigma)$ as $n \to \infty$. Assume furthermore that for a sample of size $n = 120$ we have calculated $\widehat{\beta} = 2$, and we have also calculated a consistent estimator for Σ as $\widehat{\Sigma} = 1.2$. Use this information to provide an asymptotically valid 95% confidence interval for β. (Hint: The 97.5% quantile of $\mathcal{N}(0, 1)$ is 1.96. The 95% quantile of $\mathcal{N}(0, 1)$ is 1.64.)

(e) Assume that x_i is endogenous, i.e. $\mathbb{E}(x_i u_i) \neq 0$, but that the instrument z_i satisfies the exclusion restriction $\mathbb{E}(z_i u_i) = 0$. What second assumption on z_i is required if we want to use z_i as an instrumental variable to estimate β? Use these two assumptions on z_i to express β as a function of γ_{xz} and γ_{yz}. Use this expression for β and your result in (a) to write down a consistent estimator for β. Show that this estimator is indeed consistent.

Question 3 (35 points)

The data generating process for the three scalar variables w_i, z_i and u_i is given by

$$\begin{pmatrix} w_i \\ z_i \\ u_i \end{pmatrix} \sim iid\ \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{wz} & 0 \\ \rho_{wz} & 1 & \rho_{zu} \\ 0 & \rho_{zu} & 1 \end{pmatrix}\right],$$

where $|\rho_{wz}| < 1$ and $|\rho_{zu}| < 1$. The scalar outcome variable y_i is generated from the model $y_i = w_i\,\beta_1 + z_i\,\beta_2 + u_i$, where β_1 and β_2 are two unknown parameters. For simplicity we do not include a constant into the model. We observe y_i, w_i and z_i for a sample of observations $i = 1, \ldots, n$, while u_i is an unobserved error term. Define $\beta = (\beta_1, \beta_2)'$ and $x_i = (w_i, z_i)$. The OLS estimator for β reads

$$\widehat{\beta} = \begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix} = \left(\sum_{i=1}^{n} x_i' x_i\right)^{-1} \sum_{i=1}^{n} x_i' y_i.$$

(a) Assume that $\rho_{zu} = 0$. In this case, is $\widehat{\beta}$ a consistent estimator for β? Prove your answer.

(b) For your answer in (a), explain why the two conditions $\rho_{zu} = 0$ and $|\rho_{wz}| < 1$ are important. Would there be a problem with $\widehat{\beta}$ if $\rho_{wz} = 1$?

(c) Assume that $\rho_{zu} \neq 0$ and $\rho_{wz} = 0$. In this case, is $\widehat{\beta}_1$ a consistent estimator for β_1? Prove your answer.

(d) Assume that $\rho_{zu} \neq 0$ and $\rho_{wz} \neq 0$. Show that $\widehat{\beta}_1 \to_p \beta_1^*$ as $n \to \infty$, where β_1^* is some constant. Find an expression for β_1^* as a function of β_1, ρ_{wz} and ρ_{zu}. Is $\widehat{\beta}_1$ a consistent estimator for β_1?

(e) Assume that the data generating process for (y_i, w_i, z_i, u_i) is unchanged and that $\rho_{zu} = 0$. However, we now estimate a regression model that also includes a constant, i.e. $\widehat{\beta} = (\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2)'$ is obtained by applying OLS to $y_i = \beta_0 + w_i\,\beta_1 + z_i\,\beta_2 + u_i$. Is the resulting estimator consistent for β_1 and β_2? No proof required.

G020: Mock Midterm Exam in Econometrics

2016-2017

Answer ALL questions. **Time allowed:** 2:00 hours

Question 1 (30 points)

Consider the linear regression model with a constant regressor and one additional regressor

$$y_i = \beta_1 + z_i\beta_2 + u_i.$$

Assume that all variables have finite second moments, and that the errors u_i have mean zero, are independent of z_i, and have variance $\sigma^2 = \mathbb{E}(u_i^2) = 25$, which for simplicity is assumed to be known. We observe a random sample (y_i, z_i), $i = 1, \ldots, n$, of $n = 100$ observations. For this observed sample we find

$$\sum_{i=1}^{n} z_i = 200, \qquad \sum_{i=1}^{n} y_i = 200, \qquad \sum_{i=1}^{n} z_i^2 = 500, \qquad \sum_{i=1}^{n} y_i z_i = 500.$$

(a) Calculate the OLS estimates for β_1 and β_2.

(b) Calculate an estimator for the standard deviation of $\widehat{\beta}_2$.

For the following subquestion assume that you calculated $\hat{\beta}_2 = 10$ and $\widehat{\text{std}}(\hat{\beta}_2) = 5$. (these are not the numbers that you should have actually obtained).

(c) Test the null hypothesis $H_0 : \beta_2 = 0$ using a two-sided t-test. Can you reject H_0 at 95% confidence level? (Hint: The 97.5% quantile of $\mathcal{N}(0,1)$ is 1.96. The 95% quantile of $\mathcal{N}(0,1)$ is 1.64.)

For the following two subquestions we define $w_i = 3z_i - 2$.

(d) Calculate the OLS estimator obtained from regressing y_i on $(1, w_i)$, that is, the OLS estimator for the model $y_i = \gamma_1 + w_i\gamma_2 + u_i$.

(e) Calculate the OLS estimator obtained from regressing y_i on $(1, w_i, z_i)$, that is, the OLS estimator for the model $y_i = \gamma_1 + w_i\gamma_2 + z_i\gamma_3 + u_i$.

Question 2 (35 points)

Consider a linear regression model with a single regressor

$$y_i = x_i\beta + u_i,$$

where β is the true value of the regression coefficient. We assume that (x_i, u_i) are independent and identically distributed across observation $i = 1, \ldots n$, and that $\mathbb{E}x_i = 0$, $\mathbb{E}x_i^2 = 1$, $\mathbb{E}x_i^3 = 0$, $\mathbb{E}x_i^4 = 3$, $\mathbb{E}\left(\frac{x_i^2}{1+x_i^2}\right) = 1/3$, $\mathbb{E}(u_i|x_i) = 0$ and $\mathbb{E}(u_i^2|x_i) = 1 + x_i^2$. The OLS and the weighted least squares (WLS) estimator for β are given by

$$\hat{\beta}_{\text{OLS}} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}, \qquad \hat{\beta}_{\text{WLS}} = \frac{\sum_{i=1}^n w_i x_i y_i}{\sum_{i=1}^n w_i x_i^2},$$

where the w_i are non-negative weights, which depend on x_i only. We assume that $\mathbb{E}(w_i x_i^2)$ and $\mathbb{E}[(w_i x_i u_i)^2]$ exist.

(a) Show that as $n \to \infty$ we have $\sqrt{n}(\hat{\beta}_{\text{WLS}} - \beta) \Rightarrow \mathcal{N}(0, \Sigma_{\text{WLS}})$ and derive a formula for the asymptotic variance Σ_{WLS}.

(b) The OLS estimator is a special case of $\hat{\beta}_{\text{WLS}}$ with $w_i = 1$, i.e. your results in (a) also shows that $\sqrt{n}(\hat{\beta}_{\text{OLS}} - \beta) \Rightarrow \mathcal{N}(0, \Sigma_{\text{OLS}})$ as $n \to \infty$. Calculate the asymptotic variance Σ_{OLS} (this should just be a number).

(c) To obtain the exact value of Σ_{OLS} in (b) you used that $\mathbb{E}(u_i^2|x_i) = 1 + x_i^2$. In practice the functional form of $\mathbb{E}(u_i^2|x_i)$ and thus Σ_{OLS} are unknown. Provide a formula for a consistent estimator of Σ_{OLS} that could be used in practice.

(d) Given that we know $\mathbb{E}(u_i^2|x_i) = 1 + x_i^2$, what are the optimal weights w_i, which minimize Σ_{WLS} (you do not need to show that these weights are optimal)? What is the value of Σ_{WLS} that is obtained for these optimal weights? Compare the optimal asymptotic variance Σ_{WLS} and the OLS asymptotic variance Σ_{OLS}.

(e) Is $\hat{\beta}_{\text{WLS}}$ a consistent estimator for β if we choose $w_i = 1 + |x_i|$, which are not the optimal weights from part (d)? Explain your answer, but no proof required.

Question 3 (35 points)

The data generating process for the three scalar variables x_i, z_i and u_i is given by

$$\begin{pmatrix} x_i \\ z_i \\ u_i \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{xz} & \rho_{xu} \\ \rho_{xz} & 1 & \rho_{zu} \\ \rho_{xu} & \rho_{zu} & 1 \end{pmatrix}\right].$$

where the correlations satisfy $|\rho_{xz}| < 1$, $|\rho_{xu}| < 1$ and $|\rho_{zu}| < 1$. The scalar outcome variable y_i is generated from the model $y_i = x_i \beta + u_i$, where β is an unknown scalar parameter. For simplicity we do not include a constant into the model. We observe y_i, x_i and z_i for an iid sample of observations $i = 1, \ldots, n$, while u_i is an unobserved error term. The OLS and IV estimators for β are given by

$$\widehat{\beta}_{\mathrm{OLS}} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}, \qquad \widehat{\beta}_{\mathrm{IV}} = \frac{\sum_{i=1}^n z_i y_i}{\sum_{i=1}^n z_i x_i}.$$

(a) Assume $\rho_{xz} \neq 0$. Show that $\widehat{\beta}_{\mathrm{OLS}} \to_p \beta^*_{\mathrm{OLS}}$ and $\widehat{\beta}_{\mathrm{IV}} \to_p \beta^*_{\mathrm{IV}}$ as $n \to \infty$, and find expressions for β^*_{OLS} and β^*_{IV} in terms of β, ρ_{xz}, ρ_{xu} and ρ_{zu}.

(b) Under what condition on ρ_{xu} is $\widehat{\beta}_{\mathrm{OLS}}$ consistent for β? Interpret this condition.

(c) Under what condition on ρ_{xz} and ρ_{zu} is $\widehat{\beta}_{\mathrm{IV}}$ consistent for β? Interpret those conditions.

(d) Assume that your consistency conditions in (c) hold. Show that $\sqrt{n}(\widehat{\beta}_{\mathrm{IV}} - \beta) \Rightarrow \mathcal{N}(0, \mathrm{AsyVar}(\sqrt{n}\widehat{\beta}_{\mathrm{IV}}))$, as $n \to \infty$, and find an expression for $\mathrm{AsyVar}(\sqrt{n}\widehat{\beta}_{\mathrm{IV}})$.

(e) Assume that your conditions in (b) and (c) hold, so that both estimators are consistent. Would you recommend to use $\widehat{\beta}_{\mathrm{OLS}}$ or $\widehat{\beta}_{\mathrm{IV}}$ in that case? Explain your answer.

Econometrics: Mock Midterm Exam in Econometrics

2017-2018

Answer ALL questions. **Time allowed:** 2:00 hours

[This is a take-home exam. Please hand in your solutions on **Thursday, Nov 23, before 5pm**, into the coursework box of your tutorial teacher in Drayton house.]

Question 1 (30 points)

Consider the linear regression model

$$y_i = \beta_1 + \beta_2 \, d_i + u_i,$$

where y_i is the subjective reported happiness of individual i (measured on some linear scale), d_i is a dummy variable which indicates whether individual i owns a dog ($d_i = 1$) or not ($d_i = 0$), and u_i is an unobserved error term. We assume that $\mathbb{E}(u_i) = 0$ and $\mathbb{E}(u_i d_i) = 0$. We observe an i.i.d. sample of (y_i, d_i), $i = 1, \ldots, n$. In total we observe $n = 100$ individuals, of which $n_0 = \sum_{i=1}^n (1 - d_i) = 80$ do not own a dog, and $n_1 = \sum_{i=1}^n d_i = 20$ do own a dog. The average reported happiness in the subpopulation with $d_i = 0$ is $n_0^{-1} \sum_{i=1}^n (1-d_i) y_i = 3$, and in the subpopulation with $d_i = 1$ is $n_1^{-1} \sum_{i=1}^n d_i y_i = 8$. We also calculate the sample average of squared happiness to be $n^{-1} \sum_{i=1}^n y_i^2 = 24$.

(a) Let $x_i = (1, d_i)$ and $\beta = (\beta_1, \beta_2)'$. Calculate $\sum_{i=1}^n x_i' \, x_i$, and $\sum_{i=1}^n x_i' \, y_i$, and the OLS estimator

$$\widehat{\beta} = \left(\sum_{i=1}^n x_i' \, x_i \right)^{-1} \sum_{i=1}^n x_i' \, y_i.$$

(b) Calculate the OLS estimator obtained from regressing d_i on a constant and y_i, that is, the OLS estimator $\widehat{\gamma} = (\widehat{\gamma}_1, \widehat{\gamma}_2)'$ for the regression model $d_i = \gamma_1 + \gamma_2 \, y_i + \epsilon_i$.

(c) We now also assume homoscedasticity, with $\mathbb{E}(u_i^2 | d_i) = 4$. Use this information to calculate the estimator for the OLS standard error of $\widehat{\beta}_2$.

For the following subquestion assume that you calculated $\widehat{\beta}_2 = 1$ and $\widehat{\mathrm{std}}(\widehat{\beta}_2) = 0.1$ (these are not the numbers that you should have actually obtained).

(d) Using a large sample t-test, can you reject the null hypothesis $H_0 : \beta_2 \leq 0$ against the alternative hypothesis $H_0 : \beta_2 > 0$ at 95% confidence level? (Hint: The 97.5% quantile of $\mathcal{N}(0,1)$ is 1.96, the 95% quantile of $\mathcal{N}(0,1)$ is 1.64, the 90% quantile of $\mathcal{N}(0,1)$ is 1.28.)

Question 2 (35 points)

The data generating process for the three variables w_i, z_i and u_i is given by

$$\begin{pmatrix} w_i \\ z_i \\ u_i \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{WZ} & 0 \\ \rho_{WZ} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right],$$

where $|\rho_{WZ}| < 1$. The outcome variable y_i is generated from the model

$$y_i = w_i \, \beta_1 + z_i \, \beta_2 + u_i,$$

where β_1 and β_2 are two unknown parameters. For simplicity we do not include a constant into the model. We observe y_i, w_i and z_i for an i.i.d. sample of observations $i = 1, \ldots, n$, while u_i is an unobserved error term. Define $\beta = (\beta_1, \beta_2)'$ and $x_i = (w_i, z_i)$. We consider the estimators

$$\widehat{\beta} = \begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix} = \left(\sum_{i=1}^n x_i' \, x_i\right)^{-1} \sum_{i=1}^n x_i' \, y_i \,, \qquad \widehat{\beta}_1^* = \frac{\sum_{i=1}^n y_i \, w_i}{\sum_{i=1}^n w_i^2} \,.$$

(a) Show that $\sqrt{n}(\widehat{\beta} - \beta) \Rightarrow \mathcal{N}(0, V)$, as $n \to \infty$, and provide an expression for the 2×2 asymptotic covariance matrix V as a function of ρ_{WZ} only. (Hint: Notice that the assumptions above guarantee homoscedasticity of the error u_i.)

(b) Would there be a problem with the estimator $\widehat{\beta}$ if our assumption $|\rho_{WZ}| < 1$ would be violated and we would have $\rho_{WZ} = 1$? Explain your answer.

(c) Under what conditions on β_2 and ρ_{WZ} is the estimator $\widehat{\beta}_1^*$ a consistent for β_1? Show that under those conditions we have $\sqrt{n}(\widehat{\beta}_1^* - \beta_1) \Rightarrow \mathcal{N}(0, V_*)$, as $n \to \infty$, and provide an expression for the asymptotic variance V_* that depends only on model parameters.

(d) Assume that $\beta_2 \neq 0$ and $\rho_{WZ} \neq 0$. In that case, taking into account consistency and asymptotic efficiency, which of the two estimators $\widehat{\beta}_1$ and $\widehat{\beta}_1^*$ would you recommend to use to estimate β_1?

(e) Assume that $\beta_2 = 0$ and $\rho_{WZ} \neq 0$. In that case, taking into account consistency and asymptotic efficiency, which of the two estimators $\widehat{\beta}_1$ and $\widehat{\beta}_1^*$ would you recommend to use to estimate β_1?

(f) Assume that $\beta_2 \neq 0$ and $\rho_{WZ} = 0$. In that case, taking into account consistency and asymptotic efficiency, which of the two estimators $\widehat{\beta}_1$ and $\widehat{\beta}_1^*$ would you recommend to use to estimate β_1?

Question 3 (35 points)

We observe an i.i.d. sample (y_i, x_i, z_i), $i = 1, \ldots, n$, where n is the sample size, and y_i, x_i and z_i are three variables with finite second moments. For simplicity we assume that $\mathbb{E}y_i = \mathbb{E}x_i = \mathbb{E}z_i = 0$. We consider the model

$$y_i = x_i \beta + u_i \ ,$$

where u_i is a mean zero error term, and β is the parameter of interest. We assume that $\mathbb{E}(z_i u_i) = 0$ and $\mathbb{E}(x_i z_i) \neq 0$. We consider the estimator

$$\widehat{\beta} = \frac{\sum_{i=1}^n z_i \, y_i}{\sum_{i=1}^n z_i \, x_i} \ .$$

(a) Show that $\widehat{\beta}$ is a consistent estimator for β.

(b) Show that $\sqrt{n}(\widehat{\beta} - \beta) \Rightarrow \mathcal{N}(0, V)$, as $n \to \infty$, and find an expression for the asymptotic variance V.

(c) What property of the instrumental variable z_i is described by our assumption $\mathbb{E}(z_i u_i) = 0$? What property of the instrumental variable z_i is described by our assumption $\mathbb{E}(x_i z_i) \neq 0$?

(d) For a sample with $n = 100$ observations we calculate

$$\frac{1}{n}\sum_{i=1}^n z_i y_i = 0.2, \qquad \frac{1}{n}\sum_{i=1}^n z_i x_i = 0.4, \qquad \frac{1}{n}\sum_{i=1}^n z_i^2 x_i^2 = 8,$$
$$\frac{1}{n}\sum_{i=1}^n z_i^2 x_i y_i = 1, \qquad \frac{1}{n}\sum_{i=1}^n z_i^2 y_i^2 = 3.$$

Use this information and your result in (a) to calculate an asymptotically valid 95% confidence interval for β. Notice that we do not assume homoscedasticity in this question. (Hint: The 97.5% quantile of $\mathcal{N}(0,1)$ is 1.96, the 95% quantile of $\mathcal{N}(0,1)$ is 1.64, the 90% quantile of $\mathcal{N}(0,1)$ is 1.28.)

(e) Formulate a homoscedasticity assumption on u_i. Show that under your assumption we can simplify the asymptotic variance V in (a) such that the variance of u_i, but not u_i itself, appears in the resulting expression for V.

G020: Mock Midterm Exam in Econometrics
SOLUTION

2013-2014

Question 1

(a) We must have $\hat{\beta}_1 + f_i\hat{\beta}_2 = \hat{\gamma}_1 + m_i\hat{\gamma}_2$ for all possible values of f_i and m_i. Since $f_i = 1 - m_i$ we obtain $\hat{\beta}_1 + \hat{\beta}_2 - m_i\hat{\beta}_2 = \hat{\gamma}_1 + m_i\hat{\gamma}_2$, i.e. we have $\hat{\gamma}_1 = \hat{\beta}_1 + \hat{\beta}_2$ and $\hat{\gamma}_2 = -\hat{\beta}_2$.

(b) $X'X = \begin{pmatrix} n_f + n_m & n_f \\ n_f & n_f \end{pmatrix} = \begin{pmatrix} 400 & 100 \\ 100 & 100 \end{pmatrix}$,

$X'y = \begin{pmatrix} n_f\overline{y}_f + n_m\overline{y}_m \\ n_f\overline{y}_f \end{pmatrix} = \begin{pmatrix} 2300 \\ 500 \end{pmatrix}$,

$(X'X)^{-1} = \frac{1}{30{,}000}\begin{pmatrix} 100 & -100 \\ -100 & 400 \end{pmatrix} = \begin{pmatrix} 1/300 & -1/300 \\ -1/300 & 1/75 \end{pmatrix}$,

$\hat{\beta} = (X'X)^{-1}X'y = \begin{pmatrix} 6 \\ -1 \end{pmatrix}$.

(c) $\widehat{\mathrm{Var}}(\hat{\beta}) = \mathrm{Var}(u_i|f_i)\ (X'X)^{-1} = \begin{pmatrix} 1/100 & -1/100 \\ -1/100 & 1/25 \end{pmatrix}$.

Therefore $\widehat{\mathrm{std}}(\hat{\beta}_2) = \sqrt{\widehat{\mathrm{Var}}(\hat{\beta}_2)} = 1/5 = 0.2$.

(d) We have $t = \frac{-2}{0.4} = -5$. Since $t < -1.64$ we reject the null hypothesis $H_0 : \beta_2 \geq 0$ at 5% significance level.

(e) We have $\hat{\delta} = \hat{\beta}_2$. This is an example of a <u>partitioned regression</u>: $\tilde{y}$ is the part of y that is not explained by the constant, and $\tilde{f}$ is the part of f that is not explained by the constant. Thus, regressing $\tilde{y}$ on $\tilde{f}$ gives the same result that we get by regressing y on f and a constant.

Question 2

(a) Applying the law of iterated expectations we find that

$$\sigma^2 = \mathbb{E}\left(u_i^2\right) = \mathbb{E}\left[\mathbb{E}(u_i^2|x_i)\right] = \mathbb{E}\left[\mathrm{Var}(u_i|x_i)\right] = \mathbb{E}\left[\mathrm{Var}(y_i|x_i)\right] = \mathbb{E}\left(\gamma x_i\right)$$
$$= \gamma\mu_x.$$

(b) As $n \to \infty$ we have

$$\sqrt{n}(\hat{\beta}_1 - \beta) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} u_i \Rightarrow \mathcal{N}(0, \sigma^2),$$

where in the first step we used the definition of $\hat{\beta}_1$ and u_i, and in the second step we applied the CLT. Thus, we have $\Sigma_1 = \sigma^2 = \gamma\mu_x$.

(c) Using the residuals $\hat{u}_i = y_i - \hat{\beta}_1$ we can estimate $\Sigma_1 = \sigma^2$ consistently with the following estimator

$$\hat{\Sigma}_1 = \frac{1}{n}\sum_{i=1}^{n} \hat{u}_i^2.$$

(d) We can write

$$\sqrt{n}\left(\hat{\beta}_2 - \beta\right) = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} w_i u_i}{\frac{1}{n}\sum_{i=1}^{n} w_i}.$$

For the numerator we can apply the CLT to find $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} w_i u_i \Rightarrow \mathcal{N}(0, \mathrm{Var}(w_i u_i))$ as $n \to \infty$. For the denominator we can apply the WLLN to find $\frac{1}{n}\sum_{i=1}^{n} w_i \to_p \mathbb{E}w_i$ as $n \to \infty$. Combining these results and applying Slutsky's theorem we obtain

$$\sqrt{n}\left(\hat{\beta}_2 - \beta\right) \Rightarrow \mathcal{N}\left(0,\ \frac{\mathrm{Var}(w_i u_i)}{\left(\mathbb{E}w_i\right)^2}\right).$$

So far we have not used the concrete form of the weights. Using $w_i = 1/x_i$ we find

$$\mathrm{Var}(w_i u_i) = \mathbb{E}\left(w_i^2 u_i^2\right) = \mathbb{E}\left(\frac{1}{x_i^2}\mathbb{E}(u_i^2|x_i)\right) = \mathbb{E}\left(\frac{1}{x_i^2}\mathrm{Var}(y_i|x_i)\right) = \mathbb{E}\left(\frac{\gamma}{x_i}\right) = \gamma\mathbb{E}\left(1/x_i\right),$$

where in the second step we also used the law of iterated expectations. Combining this with $\mathbb{E}w_i = \mathbb{E}(1/x_i)$ we find for the asymptotic variance of $\hat{\beta}_2$ that

$$\Sigma_2 = \frac{\mathrm{Var}(w_i u_i)}{\left(\mathbb{E}w_i\right)^2} = \frac{\gamma}{\mathbb{E}(1/x_i)}.$$

As argued in the question, we know that Jensen's inequality implies $\mathbb{E}(1/x_i) > 1/\mu_x$ and therefore $1/[\mathbb{E}(1/x_i)] < \mu_x$. This shows that

$$\Sigma_2 < \Sigma_1,$$

i.e. $\hat{\beta}_2$ is asymptotically more efficient than $\hat{\beta}_1$.

(e) We can estimate $\Sigma_2 = \frac{\gamma}{\mathbb{E}(1/x_i)}$ consistently via

$$\hat{\Sigma}_2 = \hat{\gamma}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}\right)^{-1}.$$

However, we still need to provide a consistent estimator $\hat{\gamma}$ for γ. There are at least two possibilities for this:

(1) We know $\sigma^2 = \gamma\mu_x$, so we can estimate γ via $\hat{\gamma} = \hat{\sigma}^2/\hat{\mu}_x$, where $\hat{\sigma}^2$ is provided above and $\hat{\mu}_x = \frac{1}{n}\sum_{i=1}^{n} x_i$.

(2) Alternatively, we can use the residuals $\hat{u}_i$ defined above and run a regression of $\hat{u}_i^2$ on x_i. The corresponding OLS estimator will be a consistent estimator for γ.

Question 3

(a) The OLS estimator reads

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}.$$

By applying the WLLN (for numerator and denominator) and the CMT we find as $n \to \infty$

$$\hat{\beta}_1 = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i y_i}{\frac{1}{n}\sum_{i=1}^{n} x_i^2} \to_p \frac{\mathbb{E}(x_i y_i)}{\mathbb{E}(x_i^2)}.$$

Using the model and the assumptions we obtain

$$\frac{\mathbb{E}(x_i y_i)}{\mathbb{E}(x_i^2)} = \beta_1 + \beta_2 \frac{\mathbb{E}(x_i w_i^*)}{\mathbb{E}(x_i^2)}$$

Thus, as $n \to \infty$ we have $\hat{\beta}_1 \to_p \beta_1 + \beta_2 \frac{\mathbb{E}(x_i w_i^*)}{\mathbb{E}(x_i^2)} \neq \beta_1$. We find that there is an omitted variable bias and that $\hat{\beta}_1$ is <u>not</u> consistent.

(b) Since w_i^* is unobserved we can consider $\beta_2 w_i^* + u_i$ as the effective error term of the model. The additional assumption $\mathbb{E}(w_i^* z_i) = 0$ guarantees that z_i is exogenous with respect to the effective error $\beta_2 w_i^* + u_i$. The assumption $\mathbb{E}(x_i z_i) \neq 0$ guarantees that z_i is a relevant instrument for estimation of β_1. Under these additional assumptions it is natural to employ the following instrumental variables estimator:

$$\hat{\beta}_{\mathrm{IV}} = \frac{\sum_{i=1}^n z_i y_i}{\sum_{i=1}^n z_i x_i}.$$

The question does not ask to show consistency of $\hat{\beta}_{\mathrm{IV}}$, but for completeness we provide the proof here: By applying the WLLN (for numerator and denominator) and the CMT we find as $n \to \infty$

$$\hat{\beta}_{\mathrm{IV}} = \frac{\frac{1}{n}\sum_{i=1}^n z_i y_i}{\frac{1}{n}\sum_{i=1}^n z_i x_i} \to_p \frac{\mathbb{E}(z_i y_i)}{\mathbb{E}(z_i x_i)}.$$

Using the model and the assumptions we obtain

$$\frac{\mathbb{E}(z_i y_i)}{\mathbb{E}(z_i x_i)} = \beta_1 \frac{\mathbb{E}(z_i x_i)}{\mathbb{E}(z_i x_i)} + \beta_2 \frac{\mathbb{E}(z_i w_i^*)}{\mathbb{E}(z_i x_i)} + \frac{\mathbb{E}(z_i u_i)}{\mathbb{E}(z_i x_i)} = \beta_1.$$

Thus, as $n \to \infty$ we have $\hat{\beta}_{\mathrm{IV}} \to_p \beta_1$, i.e. $\hat{\beta}_{\mathrm{IV}}$ is consistent for β_1.

(c) This OLS estimator reads

$$\hat{\beta}^* = \left[\sum_{i=1}^n (x_i, w_i)'(x_i, w_i)\right]^{-1} \sum_{i=1}^n (x_i, w_i)' y_i$$

By applying the WLLN (for "numerator" and "denominator" matrix) and the CMT we find that as $n \to \infty$

$$\begin{aligned}\hat{\beta}^* &= \left[\frac{1}{n}\sum_{i=1}^n (x_i, w_i)'(x_i, w_i)\right]^{-1} \frac{1}{n}\sum_{i=1}^n (x_i, w_i)' y_i \\ &\to_p \left\{\mathbb{E}\left[(x_i, w_i)'(x_i, w_i)\right]\right\}^{-1} \mathbb{E}\left[(x_i, w_i)' y_i\right].\end{aligned}$$

The question is not very specific regarding how much this probability limit should be evaluated further, so in one could stop here (and would get full points).

However, in the spirit of what was done in (a) it appears natural to also plug the model for y_i into this probability limit. We can write the model as $y_i = (x_i, w_i^*)(\beta_1, \beta_2)' + u_i$. Using this we obtain

$$\hat{\beta}^* \to_p \left\{\mathbb{E}\left[(x_i, w_i)'(x_i, w_i)\right]\right\}^{-1} \mathbb{E}\left[(x_i, w_i)'(x_i, w_i^*)\right] (\beta_1, \beta_2)'.$$

By using the relationship $w_i = w_i^* + v_i$ and the assumptions on v_i we find $\mathbb{E}(x_i w_i^*) = \mathbb{E}(x_i w_i)$ and $\mathbb{E}(w_i w_i^*) = \mathbb{E}[(w_i^*)^2]$ and $\mathbb{E}[w_i^2] = \mathbb{E}[(w_i^*)^2] + \mathbb{E}(v_i^2)$. This also implies that $\mathbb{E}(w_i w_i^*) = \mathbb{E}(w_i^2) - \mathbb{E}(v_i^2)$. Using this we obtain as $n \to \infty$

$$
\begin{aligned}
\hat{\beta}^* \to_p & \begin{pmatrix} \mathbb{E}(x_i^2) & \mathbb{E}(x_i w_i) \\ \mathbb{E}(x_i w_i) & \mathbb{E}(w_i^2) \end{pmatrix}^{-1} \begin{pmatrix} \mathbb{E}(x_i^2) & \mathbb{E}(x_i w_i^*) \\ \mathbb{E}(x_i w_i) & \mathbb{E}(w_i w_i^*) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \\
&= \begin{pmatrix} \mathbb{E}(x_i^2) & \mathbb{E}(x_i w_i) \\ \mathbb{E}(x_i w_i) & \mathbb{E}(w_i^2) \end{pmatrix}^{-1} \begin{pmatrix} \mathbb{E}(x_i^2) & \mathbb{E}(x_i w_i) \\ \mathbb{E}(x_i w_i) & \mathbb{E}(w_i^2) - \mathbb{E}(v_i^2) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \\
&= \begin{pmatrix} \mathbb{E}(x_i^2) & \mathbb{E}(x_i w_i) \\ \mathbb{E}(x_i w_i) & \mathbb{E}(w_i^2) \end{pmatrix}^{-1} \begin{pmatrix} \mathbb{E}(x_i^2) & \mathbb{E}(x_i w_i) \\ \mathbb{E}(x_i w_i) & \mathbb{E}(w_i^2) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \\
&\quad - \begin{pmatrix} \mathbb{E}(x_i^2) & \mathbb{E}(x_i w_i) \\ \mathbb{E}(x_i w_i) & \mathbb{E}(w_i^2) \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{E}(v_i^2) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \\
&= \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} - \begin{pmatrix} \mathbb{E}(x_i^2) & \mathbb{E}(x_i w_i) \\ \mathbb{E}(x_i w_i) & \mathbb{E}(w_i^2) \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{E}(v_i^2) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}.
\end{aligned}
$$

(d) To answer the question of whether $\hat{\beta}_1^*$ is consistent for β_1 we need to evaluate the inverse of the 2×2 matrix, which appears in the last expression for the probability limit of $\hat{\beta}^*$. We find

$$
\begin{pmatrix} \mathbb{E}(x_i^2) & \mathbb{E}(x_i w_i) \\ \mathbb{E}(x_i w_i) & \mathbb{E}(w_i^2) \end{pmatrix}^{-1} = \frac{1}{\mathbb{E}(x_i^2)\mathbb{E}(w_i^2) - \mathbb{E}(x_i w_i)^2} \begin{pmatrix} \mathbb{E}(w_i^2) & -\mathbb{E}(x_i w_i) \\ -\mathbb{E}(x_i w_i) & \mathbb{E}(x_i^2) \end{pmatrix}.
$$

Using this we find that

$$
\begin{aligned}
& \begin{pmatrix} \mathbb{E}(x_i^2) & \mathbb{E}(x_i w_i) \\ \mathbb{E}(x_i w_i) & \mathbb{E}(w_i^2) \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{E}(v_i^2) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \\
&= \frac{1}{\mathbb{E}(x_i^2)\mathbb{E}(w_i^2) - \mathbb{E}(x_i w_i)^2} \begin{pmatrix} -\mathbb{E}(x_i w_i)\mathbb{E}(v_i^2)\beta_2 \\ \mathbb{E}(x_i^2)\mathbb{E}(v_i^2)\beta_2 \end{pmatrix}.
\end{aligned}
$$

Thus, we have

$$
\hat{\beta}_1 \to_p \beta_1 + \frac{\mathbb{E}(x_i w_i)\mathbb{E}(v_i^2)}{\mathbb{E}(x_i^2)\mathbb{E}(w_i^2) - \mathbb{E}(x_i w_i)^2} \beta_2,
$$

as $n \to \infty$. This shows that $\hat{\beta}_1$ is <u>not</u> consistent for β_1.

G020: Mock Midterm Exam in Econometrics

Solutions

2014-2015

Answer ALL questions. **Time allowed:** 2:00 hours

Question 1

(a) We calculate

$$X'X = \sum_{i=1}^{n} x_i' x_i = \begin{pmatrix} \sum_{i=1}^{n} 1 & \sum_{i=1}^{n} w_i \\ \sum_{i=1}^{n} w_i & \sum_{i=1}^{n} w_i^2 \end{pmatrix} = 100 \begin{pmatrix} 1 & 3 \\ 3 & 10 \end{pmatrix},$$

$$X'y = \sum_{i=1}^{n} x_i' y_i = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} w_i y_i \end{pmatrix} = 100 \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

$$(X'X)^{-1} = \frac{1}{100} \begin{pmatrix} 10 & -3 \\ -3 & 1 \end{pmatrix},$$

$$\widehat{\beta} = \begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix} = (X'X)^{-1} X'y = \begin{pmatrix} 10 & -3 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 17 \\ -5 \end{pmatrix}.$$

(b) Plugging the model $y = X\beta + u$ into the definition $\widehat{\beta} = (X'X)^{-1}X'y$ gives $\widehat{\beta} = \beta + (X'X)^{-1}X'u$. Taking expectations of this equality, conditional on X, gives $\mathbb{E}(\widehat{\beta}|X) = \beta + (X'X)^{-1}X'\mathbb{E}(u|X)$. By assumption we have $\mathbb{E}(u|X) = \mathbb{E}(u|w) = 0$, and therefore $\mathbb{E}(\widehat{\beta}|X) = \beta$. Applying the law of iterated expectations then also gives $\mathbb{E}(\widehat{\beta}) = \beta$ for the unconditional expectation, so we have shown that $\widehat{\beta}$ is unbiased for β.

(c) When regressing y_i on w_i only we obtain the OLS estimator

$$\widetilde{\beta}_2 = \frac{\sum_{i=1}^{n} y_i w_i}{\sum_{i=1}^{n} w_i^2} = \frac{1}{10}.$$

This estimator is not unbiased, because the constant is a relevant omitted variable, causing omitted variable bias in $\widetilde{\beta}_2$. (The estimator $\widetilde{\beta}_2$ would only be unbiased if $\mathbb{E}(w_i) = 0$, but this is not assumed anywhere in the question.)

(d) One possible consistent estimator for σ^2 is given by $\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (y_i - \widehat{\beta}_1 - w_i\widehat{\beta}_2)^2$. In order to evaluate this expression we also need $\frac{1}{n}\sum_{i=1}^n y_i^2$ (this term appears when multiplying out the square in the expression for $\widehat{\sigma}^2$), but the question provides no information on $\frac{1}{n}\sum_{i=1}^n y_i^2$. We therefore do **not** have enough information to calculate a consistent estimator for σ.

However, some students might interpret the question differently, because it is stated that y_i and x_i are observed, and from those observations we can, of course, construct $\widehat{\sigma}^2$. This answer is also correct.

(e) The question assumes homoscedasticity, so that we can use the formula $\widehat{\mathrm{Var}}(\widehat{\beta}) = \sigma^2(X'X)^{-1}$. Using $\sigma^2 = 25$ and the result for $(X'X)^{-1}$ in (a) we thus find that $\widehat{\mathrm{Var}}(\widehat{\beta}_2) = \sigma^2(1/100) = 25/100 = 1/4$, and therefore $\widehat{\mathrm{std}}(\widehat{\beta}_2) = \sqrt{1/4} = 1/2$.

The t-test statistics for testing $H_0 : \beta_2 \leq 4$ thus reads $t = \frac{\widehat{\beta}_2 - 4}{\widehat{\mathrm{std}}(\widehat{\beta}_2)} = \frac{-9}{1/2} = -18$.

The critical value is 1.64, i.e. we would reject H_0 at 95% confidence level if $t > 1.64$. Since $-18 < 1.64$ we **cannot reject H_0 at 95% confidence level**.

Question 2

(a) By the CLT we have as $n \to \infty$

$$\sqrt{n}\left(\widehat{\beta} - \beta\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^n (y_i - \beta) \Rightarrow \mathcal{N}(0, \Sigma),$$

where $\Sigma = \mathrm{Var}(y_i - \beta) = \mathrm{Var}(y_i) = \sigma^2 + \gamma\mathbb{E}(x_i)$.

(b) Define the residuals $\widehat{u}_i = y_i - \widehat{\beta}$, with $\widehat{\beta}$ as in (a). We can then estimate σ^2 and γ by regressing $\widehat{u}_i^2$ on a constant (with coefficient σ^2) and x_i (with coefficient γ). The corresponding OLS estimator $(\widehat{\sigma}^2, \widehat{\gamma})$ will be consistent for (σ^2, γ).

Alternatively, we can also regress $y_i^2 - \widehat{\beta}^2$ on a constant (with coefficient σ^2) and x_i (with coefficient γ). This also gives consistent estimates for (σ^2, γ) here.

(c) The first estimator $\widehat{\Sigma}_1 = \frac{1}{n}\sum_{i=1}^n (y_i - \widehat{\beta})^2$ completely ignores the assumption that $\mathrm{Var}(y_i) = \sigma^2 + \gamma\mathbb{E}(x_i)$ and the observations for x_i, and simply estimates $\Sigma = \mathrm{Var}(y_i)$ via its sample analog. It is well-known from the lecture that this indeed gives a consistent estimator, i.e. we have $\widehat{\Sigma}_1 \to_p \Sigma$ as $n \to \infty$.

The second estimator uses the fact that $\Sigma = \sigma^2 + \gamma \mathbb{E}(x_i)$, together with the consistency results $\widehat{\sigma}^2 \to_p \sigma^2$, $\widehat{\gamma} \to_p \gamma$ and (according to the WLLN) $\frac{1}{n}\sum_{i=1}^n x_i \to \mathbb{E}(x_i)$, as $n \to \infty$. The CMT then implies that as $\widehat{\Sigma}_2 \to_p \Sigma$ as $n \to \infty$.

Thus, both estimators are consistent for Σ.

(d) The feasible weighted least squares (FWLS) estimator for β reads

$$\widehat{\beta}_2 = \widehat{\beta}_{\text{FWLS}} = \frac{\sum_{i=1}^n \widehat{w}_i y_i}{\sum_{i=1}^n \widehat{w}_i},$$

where the estimated weights $\widehat{w}_i$ should be chosen as

$$\widehat{w}_i = \frac{\lambda}{\widehat{\text{Var}}(y_i|x_i)} = \frac{\lambda}{\widehat{\sigma}^2 + \widehat{\gamma} x_i},$$

with arbitrary constant $\lambda > 0$ (e.g. simply $\lambda = 1$). The idea is that observations with larger variance should be discounted more (i.e. given less weight). This is exactly the FWLS estimator discussed in the lecture, for the special case of only one constant regressor. We know that if the variance of y_i (which equals the variance of $u_i = y_i - \beta$) is correctly specified, then the FWLS estimator has a smaller asymptotic variance than the OLS estimator.

(e) Yes, $\widehat{\beta}_2$ is still consistent in that case, in the same way that the OLS estimator is still consistent under heteroscedasticity. However, $\widehat{\beta}_2$ will not be efficient anymore: an estimator with smaller asymptotic variance could be obtained by changing the weights $\widehat{w}_i$.

Question 3

(a) Yes, in the exactly identified case we have $\widehat{\beta}_{2\text{SLS}} = \widehat{\beta}_{\text{IV}}$, namely by using the definition of $\widehat{x}_i$ we find

$$\begin{aligned}
\widehat{\beta}_{2\text{SLS}} &= \left(\sum_{i=1}^n \widehat{x}_i' \, \widehat{x}_i \right)^{-1} \sum_{i=1}^n \widehat{x}_i' y_i \\
&= \left(B' A^{-1} A A^{-1} B \right)^{-1} B' A^{-1} \sum_{i=1}^n z_i' \, y_i \\
&= \left(B' A^{-1} B \right)^{-1} B' A^{-1} \sum_{i=1}^n z_i' \, y_i,
\end{aligned}$$

where $A = \sum_{i=1}^{n} z_i' \, z_i$ and $B = \sum_{i=1}^{n} z_i' \, x_i$. Since A and B are square (2×2) matrices we simply obtain that

$$\begin{aligned}
\widehat{\beta}_{\text{2SLS}} &= B^{-1} A (B')^{-1} B' A^{-1} \sum_{i=1}^{n} z_i' \, y_i \\
&= B^{-1} A A^{-1} \sum_{i=1}^{n} z_i' \, y_i \\
&= B^{-1} \sum_{i=1}^{n} z_i' \, y_i \\
&= \left(\sum_{i=1}^{n} z_i' \, x_i \right)^{-1} \sum_{i=1}^{n} z_i' \, y_i = \widehat{\beta}_{\text{IV}}.
\end{aligned}$$

(b) Using the model $y_i = x_i \beta + u_i$ we find

$$\widehat{\beta}_{\text{IV}} = \beta + \left(\sum_{i=1}^{n} z_i' \, x_i \right)^{-1} \sum_{i=1}^{n} z_i' \, u_i.$$

Applying the WLLN twice, and afterwards the CMT, we obtain as $n \to \infty$

$$\widehat{\beta}_{\text{IV}} = \beta + \left(\frac{1}{n} \sum_{i=1}^{n} z_i' \, x_i \right)^{-1} \frac{1}{n} \sum_{i=1}^{n} z_i' \, u_i \;\; \to_p \;\; \beta + [\mathbb{E}(z_i' \, x_i)]^{-1} \mathbb{E}(z_i' \, u_i) = \beta,$$

where in the last step we used that by assumption $\mathbb{E}(z_i' \, x_i)$ is invertible and $\mathbb{E}(z_i' \, u_i) = 0$ (the last statement is implied by the assumption $\mathbb{E}(u_i | v_i) = 0$).

(c) An important assumption used in the consistency proof in (b) was that $\mathbb{E}(z_i' \, x_i)$ is invertible. This is the relevance condition, which guarantees that v_i is a relevant instrument for w_i. Note that invertibility of $\mathbb{E}(z_i' \, x_i)$ is equivalent to $\det[\mathbb{E}(z_i' \, x_i)] \neq 0$, and we have $\det[\mathbb{E}(z_i' \, x_i)] = \text{Cov}(w_i, v_i)$, i.e. in the current simple case the relevance condition can simply be expressed as $\text{Cov}(w_i, v_i) \neq 0$.

(d) Clearly $\widehat{\Sigma} = \widehat{\sigma}^2 \left(\frac{1}{n} \sum_{i=1}^{n} \widehat{x}_i' \, \widehat{x}_i \right)^{-1}$ would be consistent for $\Sigma = \sigma^2 \left(\frac{1}{n} \sum_{i=1}^{n} \widehat{x}_i' \, \widehat{x}_i \right)^{-1}$ if $\widehat{\sigma}^2$ would be consistent for σ^2. A consistent estimator for σ^2 would be given by $\frac{1}{n-K} \sum_{i=1}^{n} \widehat{u}_i^2$, with $\widehat{u}_i = y_i - x_i \widehat{\beta}_{\text{2SLS}}$. Thus, the only difference to the naive 2nd stage OLS variance estimator in the question is that there $\widehat{x}_i$ (the regressor in the 2nd stage of 2SLS) is used instead of x_i. This, however, makes a difference, because $\widehat{x}_i$ is not consistent for x_i. We have $y_i - \widehat{x}_i \widehat{\beta}_{\text{2SLS}} = \widehat{u}_i + \widehat{u}_i^{\text{extra}}$, where the extra component reads $\widehat{u}_i^{\text{extra}} = (x_i - \widehat{x}_i) \widehat{\beta}_{\text{2SLS}}$. This extra component results in inconsistency of $\widehat{\sigma}^2$.

For example, in the special case where w_i is exogenous (but different from v_i) we will asymptotically have that $\widehat{u}_i$ and $\widehat{u}_i^{\text{extra}}$ are independent, and $\widehat{\sigma}^2$ will converge to $\sigma^2 + \beta' \text{Var}(x_i - \widehat{x}_i)\beta > \sigma^2$, implying that $\widehat{\Sigma}$ overestimates the variance of the 2SLS estimator.

G020: Mock Midterm Exam in Econometrics

Solutions

2015-2016

Question 1

(a) For $x_i = (1, w_i)$ we have

$$\frac{1}{n}\sum_{i=1}^{n} x_i' x_i = \begin{pmatrix} 1 & \frac{1}{n}\sum_{i=1}^{n} w_i \\ \frac{1}{n}\sum_{i=1}^{n} w_i & \frac{1}{n}\sum_{i=1}^{n} w_i^2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix},$$

$$\frac{1}{n}\sum_{i=1}^{n} x_i' y_i = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n} y_i \\ \frac{1}{n}\sum_{i=1}^{n} w_i y_i \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}.$$

We thus find $\det\left(\frac{1}{n}\sum_{i=1}^{n} x_i' x_i\right) = 1$ and

$$\left(\frac{1}{n}\sum_{i=1}^{n} x_i' x_i\right)^{-1} = \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix},$$

$$\widehat{\beta} = \left(\frac{1}{n}\sum_{i=1}^{n} x_i' x_i\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} x_i' y_i\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

i.e. we have $\widehat{\beta}_1 = 0$ and $\widehat{\beta}_2 = 1$.

(b)
- All three estimators $\widehat{\sigma}_A^2$, $\widehat{\sigma}_B^2$, $\widehat{\sigma}_C^2$ are consistent. Notice that consistency of one of these estimators implies consistency for all of them, because $\frac{n}{n-K} \to 1$ as $n \to \infty$ for any fixed K.
- Only $\widehat{\sigma}_C^2$ is unbiased, because the degree of freedom correction for $K = 2$ regressors requires to change the prefactor $1/n$ into $1/(n-K) = 1/(n-2)$ to guarantee that the estimator for σ^2 is unbiased.

(c) We have $\widehat{\sigma}_A^2 = \frac{1}{n}\sum_{i=1}^{n} \widehat{u}_i^2$, where $\widehat{u}_i = y_i - \widehat{\beta}_1 - w_i \widehat{\beta}_2$. Therefore

$$\begin{aligned} \widehat{\sigma}_A^2 &= \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{\beta}_1 - w_i\widehat{\beta}_2)^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - w_i)^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}(y_i^2 - 2y_i w_i + w_i^2) \\ &= 30 - 2\cdot 5 + 5 = 25. \end{aligned}$$

We found $\widehat{\sigma}_A^2 = 25$.

(Comment: an alternative way of calculating $\widehat{\sigma}^2$ is to use the formula $\frac{1}{n}\sum_{i=1}^n \widehat{u}_i^2 = \frac{1}{n}\sum_{i=1}^n y_i^2 - \frac{1}{n}\sum_{i=1}^n (\widehat{\beta}_1 + w_i\widehat{\beta}_2)^2$, which is true due to the FOC of the OLS problem: $\sum_{i=1}^n \widehat{u}_i = 0$ and $\sum_{i=1}^n \widehat{u}_i w_i = 0$.)

(d) We have

$$\widehat{\mathrm{Var}}(\widehat{\beta}) = \widehat{\sigma}_A^2 \left(\sum_{i=1}^n x_i' x_i\right)^{-1} = \frac{\widehat{\sigma}_A^2}{n}\left(\frac{1}{n}\sum_{i=1}^n x_i' x_i\right)^{-1}$$
$$= \frac{25}{81}\begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 125/81 & -50/81 \\ -50/81 & 25/81 \end{pmatrix}.$$

Thus, $\widehat{\mathrm{std}}(\widehat{\beta}_1) = \sqrt{125/81} = 1.242$, and $\widehat{\mathrm{std}}(\widehat{\beta}_2) = \sqrt{25/81} = 5/9 = 0.556$.

Question 2

(a) Estimators for $\gamma_{xx} = \mathbb{E}(x_i^2)$, $\gamma_{xy} = \mathbb{E}(x_i y_i)$, $\gamma_{xz} = \mathbb{E}(x_i z_i)$ and $\gamma_{yz} = \mathbb{E}(y_i z_i)$ are given by the sample analogs

$$\widehat{\gamma}_{xx} = \frac{1}{n}\sum_{i=1}^n x_i^2, \quad \widehat{\gamma}_{xy} = \frac{1}{n}\sum_{i=1}^n x_i y_i, \quad \widehat{\gamma}_{xz} = \frac{1}{n}\sum_{i=1}^n x_i z_i, \quad \widehat{\gamma}_{yz} = \frac{1}{n}\sum_{i=1}^n y_i z_i.$$

These estimators are consistent, for example, the WLLN guarantees that as $n \to \infty$ we have $\widehat{\gamma}_{xx} = \frac{1}{n}\sum_{i=1}^n x_i^2 \to_p \mathbb{E}(x_i^2) = \gamma_{xx}$.

(b) From the model we obtain $u_i = y_i - x_i\beta$. We thus have

$$0 = \mathbb{E}(x_i u_i) = \mathbb{E}[x_i(y_i - x_i\beta)] = \mathbb{E}(x_i y_i) - \mathbb{E}(x_i^2)\beta$$
$$= \gamma_{xy} - \gamma_{xx}\beta.$$

Solving this equation for β gives $\beta = \frac{\gamma_{xy}}{\gamma_{xx}}$. A natural estimator for β is therefore given by

$$\widehat{\beta} = \frac{\widehat{\gamma}_{xy}}{\widehat{\gamma}_{xx}} = \frac{\frac{1}{n}\sum_{i=1}^n x_i y_i}{\frac{1}{n}\sum_{i=1}^n x_i^2} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}.$$

It turns out that this is simply the OLS estimator. We already know consistency of $\widehat{\gamma}_{xy}$ and $\widehat{\gamma}_{xx}$ from (a). Applying this, the CMT, and the assumption that $\gamma_{xx} > 0$ we obtain that as $n \to \infty$ we have

$$\widehat{\beta} = \frac{\widehat{\gamma}_{xy}}{\widehat{\gamma}_{xx}} \to_p \frac{\gamma_{xy}}{\gamma_{xx}} = \beta,$$

i.e. $\widehat{\beta}$ is indeed consistent.

(c) As $n \to \infty$ we have

$$\sqrt{n}(\widehat{\beta} - \beta) = \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i u_i}{\frac{1}{n} \sum_{i=1}^{n} x_i^2} \Rightarrow \frac{\mathcal{N}[0, \mathbb{E}(x_i^2 u_i^2)]}{\gamma_{xx}} \sim \mathcal{N}(0, \Sigma),$$

where

$$\Sigma = \frac{\mathbb{E}(x_i^2 u_i^2)}{\gamma_{xx}^2}.$$

Here, we have first used the definition of $\widehat{\beta}$ and the model, and then applied the WLLN, CLT and Slutsky's theorem.

(d) We first calculate the standard error of $\widehat{\beta}$,

$$\mathrm{sde}(\widehat{\beta}) = \sqrt{\widehat{\Sigma}/n} = \sqrt{1.2/120} = \sqrt{0.01} = 0.1$$

A 95% confidence interval is then given by

$$\left[\widehat{\beta} - 1.96 \times \mathrm{sde}(\widehat{\beta}), \widehat{\beta} + 1.96 \times \mathrm{sde}(\widehat{\beta})\right] = [2 - 1.96 \times 0.1, 2 + 1.96 \times 0.1] \approx [1.8, 2.2]$$

(e) If we want to use z_i as an instrument, then we also need the <u>relevance assumption</u> $\mathbb{E}(x_i z_i) \neq 0$, i.e. $\underline{\gamma_{xz} \neq 0}$. From the exclusion restriction $\mathbb{E}(z_i u_i) = 0$ we find, analogously to the answer in (b), that

$$\begin{aligned} 0 &= \mathbb{E}(z_i u_i) = \mathbb{E}[z_i(y_i - x_i\beta)] = \mathbb{E}(z_i y_i) - \mathbb{E}(z_i x_i)\beta \\ &= \gamma_{yz} - \gamma_{xz}\beta. \end{aligned}$$

Together with $\gamma_{xz} \neq 0$ this implies that $\beta = \frac{\gamma_{yz}}{\gamma_{xz}}$. A consistent estimator for β is then given by

$$\widehat{\beta}_{\mathrm{IV}} = \frac{\gamma_{yz}}{\gamma_{xz}} = \frac{\frac{1}{n} \sum_{i=1}^{n} z_i y_i}{\frac{1}{n} \sum_{i=1}^{n} z_i x_i} = \frac{\sum_{i=1}^{n} z_i y_i}{\sum_{i=1}^{n} z_i x_i}.$$

We already know consistency of $\widehat{\gamma}_{yz}$ and $\widehat{\gamma}_{xz}$ from (a). Applying this, the CMT, and the assumption that $\gamma_{xz} \neq 0$ we obtain that as $n \to \infty$ we have

$$\widehat{\beta}_{\mathrm{IV}} = \frac{\widehat{\gamma}_{yz}}{\widehat{\gamma}_{xz}} \to_p \frac{\gamma_{yz}}{\gamma_{xz}} = \beta,$$

i.e. $\widehat{\beta}_{\mathrm{IV}}$ is indeed consistent.

Question 3

(a) Consider $\rho_{zu} = 0$. We first use the model for y_i to obtain

$$\widehat{\beta} - \beta = \left(\frac{1}{n} \sum_{i=1}^{n} x_i' x_i \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} x_i' u_i \right).$$

By applying the weak law of large numbers (WLLN) and the continuous mapping theorem (CMT) we obtain as $n \to \infty$

$$\widehat{\beta} - \beta \to_p \left[\mathbb{E}(x_i' x_i) \right]^{-1} \mathbb{E}(x_i' u_i).$$

The condition $|\rho_{wz}| < 1$ guarantees that $\mathbb{E}(x_i' x_i)$ is invertible. The condition $\rho_{zu} = 0$ guarantees that $\mathbb{E}(x_i' u_i) = 0$. Therefore $\left[\mathbb{E}(x_i' x_i) \right]^{-1} \mathbb{E}(x_i' u_i) = 0$. We have thus shown that $\widehat{\beta} \to_p \beta$ as $n \to \infty$, i.e. $\underline{\widehat{\beta} \text{ is consistent for } \beta}$.

(b) The condition $\rho_{zu} = 0$ guarantees that x_i is <u>exogenous</u>. The condition $|\rho_{wz}| < 1$ guarantees that x_i is <u>non-collinear</u>.

If $\rho_{wz} = 1$, then $w_i = z_i$, so that the two regressors in $x_i = (w_i, z_i)$ are identical. The inverse of $\sum_{i=1}^{n} x_i' x_i$ and thus also $\widehat{\beta}$ are <u>not well-defined</u> in that case.

(c) Consider $\rho_{zu} \neq 0$ and $\rho_{wz} = 0$. We have already derived $\widehat{\beta} - \beta \to_p \left[\mathbb{E}(x_i' x_i) \right]^{-1} \mathbb{E}(x_i' u_i)$, as $n \to \infty$. This is still valid here. In the current case we have

$$\left[\mathbb{E}(x_i' x_i) \right]^{-1} \mathbb{E}(x_i' u_i) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ \rho_{zu} \end{pmatrix} = \begin{pmatrix} 0 \\ \rho_{zu} \end{pmatrix}.$$

This shows that still $\widehat{\beta}_1 - \beta_1 \to_p 0$, i.e. $\widehat{\beta}_1$, i.e. $\underline{\widehat{\beta}_1 \text{ is consistent for } \beta_1}$.

(this was not asked, but we also showed that $\widehat{\beta}_2 - \beta_2 \to_p \rho_{zu} \neq 0$, i.e. $\widehat{\beta}_2$ is not consistent)

(d) We have already derived $\widehat{\beta} - \beta \to_p \left[\mathbb{E}(x_i' x_i) \right]^{-1} \mathbb{E}(x_i' u_i)$, as $n \to \infty$. This is still valid here. We have

$$\begin{aligned}
\left[\mathbb{E}(x_i' x_i) \right]^{-1} \mathbb{E}(x_i' u_i) &= \begin{pmatrix} 1 & \rho_{wz} \\ \rho_{wz} & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ \rho_{zu} \end{pmatrix} \\
&= \frac{1}{1 - \rho_{wz}^2} \begin{pmatrix} 1 & -\rho_{wz} \\ -\rho_{wz} & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \rho_{zu} \end{pmatrix} \\
&= \frac{1}{1 - \rho_{wz}^2} \begin{pmatrix} -\rho_{wz} \rho_{zu} \\ \rho_{zu} \end{pmatrix}.
\end{aligned}$$

Thus, we have shown that $\widehat{\beta}_1 \to_p \beta_1^*$ as $n \to \infty$, where

$$\beta_1^* = \beta_1 - \frac{\rho_{wz}\rho_{zu}}{1 - \rho_{wz}^2}.$$

For $\rho_{zu} \neq 0$ and $\rho_{wz} \neq 0$ we have $\beta_1^* \neq \beta$, i.e. $\underline{\widehat{\beta}_1 \text{ is NOT consistent for } \beta_1.}$

(e) When including a constant, then all regressors are still non-collinear and exogenous, i.e. $\widehat{\beta}$ is still consistent.

SOLUTIONS

G020: Mock Midterm Exam in Econometrics

2016-2017

Question 1

(a) For $x_i = (1, z_i)$ we have

$$\sum_{i=1}^{n} x_i' x_i = \begin{pmatrix} n & \sum_{i=1}^n z_i \\ \sum_{i=1}^n z_i & \sum_{i=1}^n z_i^2 \end{pmatrix} = \begin{pmatrix} 100 & 200 \\ 200 & 500 \end{pmatrix},$$

$$\sum_{i=1}^{n} x_i' y_i = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n z_i y_i \end{pmatrix} = \begin{pmatrix} 200 \\ 500 \end{pmatrix}.$$

We thus find $\det\left(\sum_{i=1}^n x_i' x_i\right) = 10.000$ and

$$\left(\sum_{i=1}^{n} x_i' x_i\right)^{-1} = \frac{1}{10.000} \begin{pmatrix} 500 & -200 \\ -200 & 100 \end{pmatrix} = \frac{1}{100} \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix},$$

$$\widehat{\beta} = \left(\sum_{i=1}^{n} x_i' x_i\right)^{-1} \left(\sum_{i=1}^{n} x_i' y_i\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

i.e. we have $\widehat{\beta}_1 = 0$ and $\widehat{\beta}_2 = 1$.

(b) We have

$$\begin{aligned} \widehat{\mathrm{Var}}(\widehat{\beta}) &= \sigma^2 \left(\sum_{i=1}^{n} x_i' x_i\right)^{-1} \\ &= \frac{25}{100} \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 5/4 & -1/2 \\ -1/2 & 1/4 \end{pmatrix}. \end{aligned}$$

Thus, $\widehat{\mathrm{std}}(\widehat{\beta}_2) = \sqrt{1/4} = 1/2$.

(c) $t = \frac{\widehat{\beta}_2 - 0}{\widehat{\mathrm{std}}(\widehat{\beta}_2)} = \frac{10}{5} = 2$. Since $|t| \geq 1.96$ we can reject H_0 at 95% confidence level.

(d) We want

$$\hat{\beta}_1 + z_i\hat{\beta}_2 = \hat{\gamma}_1 + w_i\hat{\gamma}_2.$$

Plugging in $w_i = 3z_i - 2$ we obtain

$$\hat{\beta}_1 + z_i\hat{\beta}_2 = \hat{\gamma}_1 - 2\hat{\gamma}_2 + z_i(3\hat{\gamma}_2).$$

This needs to hold for any value of z_i, which can only be the case if we have

$$\hat{\beta}_2 = 3\hat{\gamma}_2, \qquad \hat{\beta}_1 = \hat{\gamma}_1 - 2\hat{\gamma}_2.$$

Solving for $\hat{\gamma}_1$ and $\hat{\gamma}_2$ gives

$$\hat{\gamma}_2 = \frac{\hat{\beta}_2}{3} = \frac{1}{3}, \qquad \hat{\gamma}_1 = \hat{\beta}_1 + 2\hat{\gamma}_2 = \hat{\beta}_1 + \frac{2\hat{\beta}_2}{3} = \frac{2}{3}.$$

(e) The regressors $(1, w_i, z_i)$ are perfectly collinear, because we have $(1, w_i, z_i)(2, 1, -3)' = 0$ according to the definition of $w_i = 3z_i - 2$. Thus, the OLS estimator from regressing any variables on $(1, w_i, z_i)$ is not well-defined.

Question 2

(a) We have

$$\sqrt{n}\left(\hat{\beta}_{\mathrm{WLS}}-\beta\right)=\frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}w_i x_i u_i}{\frac{1}{n}\sum_{i=1}^{n}w_i x_i^2}.$$

For the numerator of this expression we have $\frac{1}{n}\sum_{i=1}^{n}w_i x_i^2 \to_p \mathbb{E}(w_i x_i^2)$ as $n\to\infty$ by the WLLN. For the denominator we have $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}w_i x_i u_i \Rightarrow \mathcal{N}(0,\mathbb{E}[(w_i x_i u_i)^2])$ by the CLT, and since $\mathbb{E}(w_i x_i u_i)=0$ by assumption. Combining these results and applying Slutsky's theorem thus gives

$$\sqrt{n}\left(\hat{\beta}_{\mathrm{WLS}}-\beta\right)\Rightarrow\mathcal{N}\left(0,\frac{\mathbb{E}[(w_i x_i u_i)^2]}{[\mathbb{E}(w_i x_i^2)]^2}\right),$$

as $n\to\infty$. For the asymptotic variance we thus have

$$\Sigma_{\mathrm{WLS}}=\frac{\mathbb{E}[(w_i x_i u_i)^2]}{[\mathbb{E}(w_i x_i^2)]^2}=\frac{\mathbb{E}[w_i^2 x_i^2\mathbb{E}(u_i^2|x_i)]}{[\mathbb{E}(w_i x_i^2)]^2}=\frac{\mathbb{E}[w_i^2 x_i^2(1+x_i^2)]}{[\mathbb{E}(w_i x_i^2)]^2},$$

where we used the law of iterated expectations and $\mathbb{E}(u_i^2|x_i)=1+x_i^2$.

(b) For $w_i=1$ we obtain the OLS estimator, whose asymptotic variance thus reads

$$\Sigma_{\mathrm{OLS}}=\frac{\mathbb{E}[x_i^2(1+x_i^2)]}{[\mathbb{E}(x_i^2)]^2}=\frac{\mathbb{E}(x_i^2)+\mathbb{E}(x_i^4)}{[\mathbb{E}(x_i^2)]^2}=\frac{1+3}{1^2}=4.$$

(c) The error term u_i is heteroscedastic here, i.e. one needs to use the heterscedasticity robust variance estimator for $\Sigma_{\mathrm{OLS}}=\frac{\mathbb{E}[(x_i u_i)^2]}{[\mathbb{E}(x_i^2)]^2}$. This estimator is obtained from the last formula by replacing expectation by sample means and error u_i by residuals $\hat{u}_i=y_i-x_i\hat{\beta}_{\mathrm{OLS}}$, i.e.

$$\hat{\Sigma}_{\mathrm{OLS}}=\frac{\frac{1}{n}\sum_{i=1}^{n}x_i^2\hat{u}_i^2}{\left(\frac{1}{n}\sum_{i=1}^{n}x_i^2\right)^2}.$$

(d) From the lecture we know that the optimal weight are given by $w_i=\frac{\lambda}{\mathbb{E}(u_i^2|x_i)}=\frac{\lambda}{1+x_i^2}$, where $\lambda>0$ is an arbitrary scalar factor, which in the following we choose as $\lambda=1$. We know that for the optimal weights, there should be a cancelation between the denominator and the numerator in Σ_{WLS}, namely we find

$$\Sigma_{\mathrm{WLS}}=\frac{\mathbb{E}[w_i^2 x_i^2(1+x_i^2)]}{[\mathbb{E}(w_i x_i^2)]^2}=\frac{1}{\mathbb{E}(w_i x_i^2)}=\left[\mathbb{E}\left(\frac{x_i^2}{1+x_i^2}\right)\right]^{-1}=\left(\frac{1}{3}\right)^{-1}=3.$$

Thus, for the optimal weights we have $\Sigma_{\mathrm{WLS}}<\Sigma_{\mathrm{OLS}}$, i.e. the WLS estimator is more efficient than the OLS estimator.

(e) Yes, $\hat{\beta}_{\text{WLS}}$ is consistent even if we choose weights that are not optimal. In particular, our asymptotic normality proof in part (a) did not require the weights to be optimal, and asymptotic normality implies consistency.

Question 3

(a) We have as $n \to \infty$

$$\widehat{\beta}_{\text{OLS}} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i y_i}{\frac{1}{n}\sum_{i=1}^{n} x_i^2} \to_p \frac{\mathbb{E}(x_i y_i)}{\mathbb{E}(x_i^2)} = \beta_{\text{OLS}}^*,$$
$$\widehat{\beta}_{\text{IV}} = \frac{\frac{1}{n}\sum_{i=1}^{n} z_i y_i}{\frac{1}{n}\sum_{i=1}^{n} z_i x_i} \to_p \frac{\mathbb{E}(z_i y_i)}{\mathbb{E}(z_i x_i)} = \beta_{\text{IV}}^*,$$

where we have used the WLLN in both the numerators and the denominators, and we also applied the CMT. By also using the model for y_i and the data generating process for x_i, z_i, u_i we obtain

$$\beta_{\text{OLS}}^* = \frac{\mathbb{E}(x_i y_i)}{\mathbb{E}(x_i^2)} = \frac{\beta\mathbb{E}(x_i^2) + \mathbb{E}(x_i u_i)}{\mathbb{E}(x_i^2)} = \beta + \rho_{xu},$$
$$\beta_{\text{IV}}^* = \frac{\mathbb{E}(z_i y_i)}{\mathbb{E}(z_i x_i)} = \frac{\beta\mathbb{E}(z_i x_i) + \mathbb{E}(z_i u_i)}{\mathbb{E}(z_i x_i)} = \beta + \frac{\rho_{zu}}{\rho_{xz}}.$$

(b) If $\rho_{xu} = 0$, then $\beta_{\text{OLS}}^* = \beta$, so that $\widehat{\beta}_{\text{OLS}}$ is consistent for β. The condition $\rho_{xu} = 0$ is just the standard exogeneity condition $\mathbb{E}(x_i u_i) = 0$, which guarantees that x_i is uncorrelated with u_i.

(c)
- If $\rho_{zu} = 0$, then $\beta_{\text{IV}}^* = \beta$, so that $\widehat{\beta}_{\text{IV}}$ is consistent for β.
- The condition $\rho_{zu} = 0$ states that z_i is exogenous (uncorrelated with the error u_i), which is also called the exclusion restriction.
- The condition $\rho_{xz} \neq 0$ guarantees that z_i is correlated with x_i, i.e. that the instrument z_i is relevant for the endogenous regressor x_i.

(d) Using the model we find

$$\sqrt{n}\left(\widehat{\beta}_{\text{IV}} - \beta\right) = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i u_i}{\frac{1}{n}\sum_{i=1}^{n} z_i x_i}$$

Since we assume $\rho_{zu} = \mathbb{E}(z_i u_i) = 0$ we can apply the CLT to find that as $n \to \infty$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i u_i \Rightarrow \mathcal{N}(0, \text{Var}(z_i u_i)).$$

As above, by the WLLN we have $\frac{1}{n}\sum_{i=1}^{n} z_i x_i \to_p \rho_{xz}$. Applying Slutsky's theorem we thus find as $n \to \infty$

$$\sqrt{n}\left(\widehat{\beta}_{\text{IV}} - \beta\right) \Rightarrow \frac{\mathcal{N}(0, \text{Var}(z_i u_i))}{\rho_{xz}} = \mathcal{N}\left(0, \frac{\text{Var}(z_i u_i)}{\rho_{xz}^2}\right),$$

i.e.

$$\text{AsyVar}(\sqrt{n}\widehat{\beta}_{\text{IV}}) = \frac{\text{Var}(z_i u_i)}{\rho_{xz}^2}.$$

Note that $\text{Var}(z_i u_i) = \mathbb{E}[(z_i u_i)^2]$.

It is ok if students stop here. But, since z_i and u_i are normally distributed and uncorrelated they are also independent, so we can further evaluate $\text{Var}(z_i u_i) = \mathbb{E}[(z_i u_i)^2] = \mathbb{E}(z_i^2 u_i^2) = \mathbb{E}(z_i^2)\mathbb{E}(u_i^2) = 1$. We thus find the simplified final result

$$\text{AsyVar}(\sqrt{n}\widehat{\beta}_{\text{IV}}) = \frac{1}{\rho_{xz}^2}.$$

An intermediate level of simplification, which is also correct, would use homoscedasticity and $\sigma_u^2 = 1$ to obtain $\text{AsyVar}(\sqrt{n}\widehat{\beta}_{\text{IV}}) = \frac{\mathbb{E}(z_i^2)}{\rho_{xz}^2}$.

(e) If both estimators are consistent, then we prefer to use $\widehat{\beta}_{\text{OLS}}$, because it has a smaller (asymptotic) variance, that is, the OLS estimator is more efficient here.

SOLUTIONS

Econometrics: Mock Midterm Exam in Econometrics

2017-2018

Question 1

(a) Using that $d_i^2 = d_i$ we find

$$\sum_{i=1}^{n} x_i' x_i = \sum_{i=1}^{n} \begin{pmatrix} 1 \\ d_i \end{pmatrix} (1, d_i) = \sum_{i=1}^{n} \begin{pmatrix} 1 & d_i \\ d_i & d_i^2 \end{pmatrix} = \sum_{i=1}^{n} \begin{pmatrix} 1 & d_i \\ d_i & d_i \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^{n} d_i \\ \sum_{i=1}^{n} d_i & \sum_{i=1}^{n} d_i \end{pmatrix}$$

$$= \begin{pmatrix} n & n_1 \\ n_1 & n_1 \end{pmatrix} = \begin{pmatrix} 100 & 20 \\ 20 & 20 \end{pmatrix}.$$

Using that $\sum_{i=1}^{n} y_i = n_0 \left[\frac{1}{n_0} \sum_{i=1}^{n} (1 - d_i) y_i \right] + n_1 \left[\frac{1}{n_1} \sum_{i=1}^{n} d_i y_i \right]$ we find

$$\sum_{i=1}^{n} x_i' y_i = \sum_{i=1}^{n} \begin{pmatrix} 1 \\ d_i \end{pmatrix} y_i = \sum_{i=1}^{n} \begin{pmatrix} y_i \\ d_i y_i \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} d_i y_i \end{pmatrix}$$

$$= \begin{pmatrix} n_0 \left[\frac{1}{n_0} \sum_{i=1}^{n} (1 - d_i) y_i \right] + n_1 \left[\frac{1}{n_1} \sum_{i=1}^{n} d_i y_i \right] \\ n_1 \left[\frac{1}{n_1} \sum_{i=1}^{n} d_i y_i \right] \end{pmatrix} = \begin{pmatrix} 80 \times 3 + 20 \times 8 \\ 20 \times 8 \end{pmatrix}$$

$$= \begin{pmatrix} 400 \\ 160 \end{pmatrix}.$$

We thus find

$$\widehat{\beta} = \begin{pmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix} = \left(\sum_{i=1}^{n} x_i' x_i \right)^{-1} \sum_{i=1}^{n} x_i' y_i = \begin{pmatrix} 100 & 20 \\ 20 & 20 \end{pmatrix}^{-1} \begin{pmatrix} 400 \\ 160 \end{pmatrix}$$

$$= \frac{1}{1.600} \begin{pmatrix} 20 & -20 \\ -20 & 100 \end{pmatrix} \begin{pmatrix} 400 \\ 160 \end{pmatrix}$$

$$= \begin{pmatrix} 3 \\ 5 \end{pmatrix}.$$

(b) Define $z_i = (1, y_i)$. We calculate

$$\sum_{i=1}^{n} z_i' z_i = \sum_{i=1}^{n} \begin{pmatrix} 1 & y_i \\ y_i & y_i^2 \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} y_i & \sum_{i=1}^{n} y_i^2 \end{pmatrix} = \begin{pmatrix} 100 & 400 \\ 400 & 2400 \end{pmatrix},$$

and

$$\sum_{i=1}^{n} z_i' d_i = \begin{pmatrix} \sum_{i=1}^{n} d_i \\ \sum_{i=1}^{n} y_i d_i \end{pmatrix} = \begin{pmatrix} 20 \\ 160 \end{pmatrix}.$$

The OLS estimator for the regression model $d_i = \gamma_1 + \gamma_2\, y_i + \epsilon_i$ thus reads

$$\begin{aligned}
\widehat{\gamma} = \begin{pmatrix} \widehat{\gamma}_1 \\ \widehat{\gamma}_2 \end{pmatrix} &= \left(\sum_{i=1}^{n} z_i' z_i \right)^{-1} \sum_{i=1}^{n} z_i' d_i = \begin{pmatrix} 100 & 400 \\ 400 & 2400 \end{pmatrix}^{-1} \begin{pmatrix} 20 \\ 160 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 4 \\ 4 & 24 \end{pmatrix}^{-1} \begin{pmatrix} 0.2 \\ 1.6 \end{pmatrix} = \frac{1}{8} \begin{pmatrix} 24 & -4 \\ -4 & 1 \end{pmatrix} \begin{pmatrix} 0.2 \\ 1.6 \end{pmatrix} \\
&= \begin{pmatrix} -0.2 \\ 0.1 \end{pmatrix}.
\end{aligned}$$

(c) Under homoscedasticity with known $\sigma^2 = \mathbb{E}(u_i^2 | d_i) = 4$ the OLS estimator for the variance of $\widehat{\beta} = \binom{\widehat{\beta}_1}{\widehat{\beta}_2}$ is given by

$$\widehat{\mathrm{Var}}(\widehat{\beta}) = \sigma^2 \left(\sum_{i=1}^{n} x_i' x_i \right)^{-1} = 4\, \frac{1}{1.600} \begin{pmatrix} 20 & -20 \\ -20 & 100 \end{pmatrix} = \begin{pmatrix} 0.05 & -0.05 \\ -0.05 & 0.25 \end{pmatrix}.$$

We thus have

$$\widehat{\mathrm{std}}(\widehat{\beta}_2) = \sqrt{\widehat{\mathrm{Var}}(\widehat{\beta}_2)} = \sqrt{0.25} = 0.5.$$

(d) The t-test statistics reads

$$t = \frac{\widehat{\beta}_2}{\widehat{\mathrm{std}}(\widehat{\beta}_2)} = \frac{1}{0.1} = 10.$$

We can reject $H_0 : \beta_2 \leq 0$ in a one-sided t-test at 95% confidence level if $t > 1.64$. Since indeed $10 > 1.64$ we **do reject** H_0.

Question 2

(a) We first use the model for y_i to obtain

$$\sqrt{n} \left(\widehat{\beta} - \beta \right) = \left(\frac{1}{n} \sum_{i=1}^{n} x_i' x_i \right)^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i' u_i \right).$$

As $n \to \infty$, we apply the weak law of large numbers (WLLN) to obtain

$$\frac{1}{n}\sum_{i=1}^{n} x_i' x_i \to_p \mathbb{E}(x_i' x_i) = \begin{pmatrix} 1 & \rho_{WZ} \\ \rho_{WZ} & 1 \end{pmatrix},$$

and the central limit theorem (CLT) to obtain

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i' u_i \Rightarrow \mathcal{N}\left(0, \mathbb{E}(u_i^2\, x_i' x_i)\right) \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{WZ} \\ \rho_{WZ} & 1 \end{pmatrix}\right),$$

where in the last step we used that the DGP in the question implies homoscedasticity, $\mathbb{E}(u_i^2 | x_i) = 1$, so that $\mathbb{E}(u_i^2\, x_i' x_i) = \mathbb{E}(x_i' x_i)$. Combining the above results and applying Slutsky's theorem we find

$$\sqrt{n}(\widehat{\beta} - \beta) \Rightarrow \mathcal{N}(0, V),$$

with

$$V = [\mathbb{E}(x_i' x_i)]^{-1}\, \mathbb{E}(u_i^2\, x_i' x_i)\, [\mathbb{E}(x_i' x_i)]^{-1} = \begin{pmatrix} 1 & \rho_{WZ} \\ \rho_{WZ} & 1 \end{pmatrix}^{-1}$$
$$= \frac{1}{1-\rho_{WZ}^2}\begin{pmatrix} 1 & -\rho_{WZ} \\ -\rho_{WZ} & 1 \end{pmatrix}.$$

(b) Yes, if $\rho_{WZ} = 1$ that would mean that w_i and z_i would be perfectly correlated, implying that $w_i = z_i$. We then have a collinearity problem: The matrix $(\sum_{i=1}^{n} x_i'\, x_i)$ is not invertible, and $\widehat{\beta}$ is not well-defined.

(c) By the WLLN and continuous mapping theorem (CMT) we have, as $n \to \infty$,

$$\widehat{\beta}_1^* = \frac{\frac{1}{n}\sum_{i=1}^{n} y_i\, w_i}{\frac{1}{n}\sum_{i=1}^{n} w_i^2} \to_p \frac{\mathbb{E} y_i\, w_i}{\mathbb{E} w_i^2} = \beta_1 + \beta_2 \frac{\mathbb{E} z_i\, w_i}{\mathbb{E} w_i^2} = \beta_1 + \beta_2\, \rho_{WZ},$$

where we also use the model for y_i. Thus, $\widehat{\beta}_1^*$ is consistent iff

$$\beta_2\, \rho_{WZ} = 0,$$

or equivalently iff

$$\beta_2 = 0 \qquad \text{or} \qquad \rho_{WZ} = 0.$$

Under that assumption we can apply the WLLN, CLT and Slutsky's theorem, as in part (a), to find that

$$\sqrt{n}(\widehat{\beta}_1^* - \beta_1) \Rightarrow \mathcal{N}(0, V_*),$$

where using homoscedasticity of the effective error term $\epsilon_i = z_i \beta_2 + u_i$ we have

$$V_* = \frac{\mathrm{Var}(z_i \beta_2 + u_i)}{\mathbb{E}w_i^2} = (\beta_2)^2 + 1.$$

(d) For $\beta_2 \neq 0$ and $\rho_{WZ} \neq 0$ we have (as shown above) that

- $\widehat{\beta}_1$ is consistent.
- $\widehat{\beta}_1^*$ is inconsistent.

We would therefore recommend to use $\widehat{\beta}_1$.

(e) For $\beta_2 = 0$ and $\rho_{WZ} \neq 0$ both $\widehat{\beta}_1$ and $\widehat{\beta}_1^*$ are consistent, but we find (from part (a) and (b))

$$\mathrm{AsyVar}(\widehat{\beta}_1) = V_{11} = \frac{1}{1-\rho_{WZ}^2} > 1, \qquad \mathrm{AsyVar}(\widehat{\beta}_1^*) = V_* = (\beta_2)^2 + 1 = 1,$$

and therefore

$$\mathrm{AsyVar}(\widehat{\beta}_1) > \mathrm{AsyVar}(\widehat{\beta}_1^*).$$

We would therefore recommend to use $\widehat{\beta}_1^*$.

(f) For $\beta_2 \neq 0$ and $\rho_{WZ} = 0$ both $\widehat{\beta}_1$ and $\widehat{\beta}_1^*$ are consistent, but we find (from part (a) and (b))

$$\mathrm{AsyVar}(\widehat{\beta}_1) = V_{11} = \frac{1}{1-\rho_{WZ}^2} = 1, \qquad \mathrm{AsyVar}(\widehat{\beta}_1^*) = V_* = (\beta_2)^2 + 1 > 1,$$

and therefore

$$\mathrm{AsyVar}(\widehat{\beta}_1) < \mathrm{AsyVar}(\widehat{\beta}_1^*).$$

We would therefore recommend to use $\widehat{\beta}_1$.

Question 3

(a) We have as, $n \to \infty$,

$$\widehat{\beta} = \frac{\frac{1}{n}\sum_{i=1}^{n} z_i y_i}{\frac{1}{n}\sum_{i=1}^{n} z_i x_i} \to_p \frac{\mathbb{E}(z_i y_i)}{\mathbb{E}(z_i x_i)} = \beta + \frac{\mathbb{E}(z_i u_i)}{\mathbb{E}(z_i x_i)} = \beta,$$

where in the first step we used the WLLN in both the numerator and the denominator, and we also applied the CMT, and afterwards we used the model assumptions.

(b) Using the model we find

$$\sqrt{n}\left(\widehat{\beta} - \beta\right) = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i u_i}{\frac{1}{n}\sum_{i=1}^{n} z_i x_i}.$$

We apply the CLT to find that, as $n \to \infty$,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i u_i \Rightarrow \mathcal{N}(0, \mathrm{Var}(z_i u_i)).$$

By the WLLN we have $\frac{1}{n}\sum_{i=1}^{n} z_i x_i \to_p \mathbb{E}(z_i x_i)$. Applying Slutsky's theorem we thus find, as $n \to \infty$,

$$\sqrt{n}\left(\widehat{\beta} - \beta\right) \Rightarrow \frac{\mathcal{N}(0, \mathrm{Var}(z_i u_i))}{\rho_{xz}} = \mathcal{N}\left(0, \frac{\mathrm{Var}(z_i u_i)}{[\mathbb{E}(z_i x_i)]^2}\right).$$

Thus,

$$V = \mathrm{AsyVar}(\sqrt{n}\widehat{\beta}) = \frac{\mathrm{Var}(z_i u_i)}{[\mathbb{E}(z_i x_i)]^2} = \frac{\mathbb{E}[(z_i u_i)^2]}{[\mathbb{E}(z_i x_i)]^2}.$$

(c)
- The condition $\mathbb{E}(z_i u_i) = 0$ states that z_i is exogenous (uncorrelated with the error u_i), which is also called the exclusion restriction.
- The condition $\mathbb{E}(z_i x_i) \neq 0$ guarantees that z_i is correlated with x_i, i.e. that the instrument z_i is relevant for the endogenous regressor x_i.

(e) Using the information on the observed sample we calculate

$$\widehat{\beta} = \frac{\frac{1}{n}\sum_{i=1}^{n} z_i y_i}{\frac{1}{n}\sum_{i=1}^{n} z_i x_i} = \frac{0.2}{0.4} = 0.5,$$

$$\widehat{\mathbb{E}}(z_i x_i) = \frac{1}{n}\sum_{i=1}^{n} z_i x_i = 0.4.$$

Define the residuals $\widehat{u}_i = y_i - x_i\widehat{\beta}$. An estimator for $\mathrm{Var}(z_i u_i) = \mathbb{E}[(z_i u_i)^2]$ is then given by

$$\begin{aligned}
\widehat{\mathbb{E}}[(z_i u_i)^2] &= \frac{1}{n}\sum_{i=1}^{n}(z_i\widehat{u}_i)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} z_i^2(y_i - x_i\widehat{\beta})^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} z_i^2 y_i^2 - 2\widehat{\beta}\frac{1}{n}\sum_{i=1}^{n} z_i^2 x_i y_i + \widehat{\beta}^2\frac{1}{n}\sum_{i=1}^{n} z_i^2 x_i^2 \\
&= 3 - 2 \times 0.5 \times 1 + 0.5^2 \times 8 = 4.
\end{aligned}$$

A consistent estimator for the asymptotic variance is thus given by

$$\widehat{\mathrm{AsyVar}}(\sqrt{n}\widehat{\beta}) = \frac{\widehat{\mathbb{E}}[(z_i u_i)^2]}{[\widehat{\mathbb{E}}(z_i x_i)]^2} = \frac{4}{(0.4)^2} = 25.$$

An estimator for the variance of $\widehat{\beta}$ is thus given by

$$\widehat{\mathrm{Var}}(\widehat{\beta}) = \frac{1}{n}\widehat{\mathrm{AsyVar}}(\sqrt{n}\widehat{\beta}) = \frac{25}{100} = 0.25,$$

implying that $\widehat{\mathrm{std}}(\widehat{\beta}) = 0.5$. A 95% confidence interval for β is thus given by

$$\begin{aligned}
\mathrm{CI}_{95\%} &= \left[\widehat{\beta} - 1.96 \times \widehat{\mathrm{std}}(\widehat{\beta}), \widehat{\beta} + 1.96 \times \widehat{\mathrm{std}}(\widehat{\beta})\right] = [0.5 - 1.96 \times 0.5, 0.5 + 1.96 \times 0.5] \\
&= [-0.48, 1.48].
\end{aligned}$$

(e) If we assume that $\mathbb{E}(u_i^2|z_i) = \sigma^2$, then by applying the law of iterated expectations (LIE) we find

$$\mathrm{AsyVar}(\sqrt{n}\widehat{\beta}) = \frac{\mathbb{E}[(z_i u_i)^2]}{[\mathbb{E}(z_i x_i)]^2} = \frac{\sigma^2\,\mathbb{E}[z_i^2]}{[\mathbb{E}(z_i x_i)]^2}.$$