Economics 403 Michaelmas Test

Oliver Linton Name ______________________

Instructions. Answer one question from part A and one question from part B. Points for each question are given in the margin. Explain all your answers.

Part A

1. Researcher A regresses y on a constant and on x, obtaining ordinary least squares (OLS) estimates a (intercept) and b (slope) and residual vector e. Researcher B, however, regresses y on a constant, on x, and on z, obtaining OLS estimates $\widehat{\alpha}$, $\widehat{\beta}$, and $\widehat{\gamma}$ as the respective coefficient estimates, and $\widehat{\varepsilon}$ as residuals. Explain in detail under what circumstances the following could be true.

   (a) $b = \widehat{\beta}$ [5]

   (b) $\sum_{i=1}^{n} \widehat{\varepsilon}_i^2 \leq \sum_{i=1}^{n} \widehat{e}_i^2$ [5]

   (c) b is statistically significant at the 5% level yet $\widehat{\beta}$ is not. [5]

   (d) $\widehat{\beta}$ is statistically significant at the 5% level yet b is not. [5]

2. Consider the multiple regression equation:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i, i = 1, 2, \ldots, 25.$$

   Show how you could test the following hypotheses using only unconstrained least squares procedures. Take a significance level of 5% and give the exact critical region involving the test statistic and the critical value from the appropriate distribution. You can assume that the usual regression assumptions including normality hold.

   (a) $\mathbf{H}_0$: $\beta_1/\beta_2 = 1$, $\beta_3 = -2$ vs. $\mathbf{H}_\text{A}$: $\beta_1/\beta_2 \neq 1$, $\beta_3 \neq -2$. [8]

   (b) $\mathbf{H}_0$: $\beta_1 + \beta_2 + \beta_3 = 0$ vs. $\mathbf{H}_\text{A}$: $\beta_1 + \beta_2 + \beta_3 < 0$. [6]

   (c) $\mathbf{H}_0$: $\beta_1 = \beta_2 = \beta_3 = 1$ vs. $\mathbf{H}_\text{A}$: $\beta_1 = \beta_2 = \beta_3 \neq 1$. [6]

Part B

3. Let $y_i = \alpha + \beta x_i + \varepsilon_i$, $i = 1, 2, .., n$, where the usual assumptions hold. However, some data are missing, due to careless data recording. In fact, complete data on y is available, while x is only present for the first $n_1 = n/2$ observations. Let

$$\begin{pmatrix} \widehat{\alpha}_1 \\ \widehat{\beta}_1 \end{pmatrix} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y}_1, \mathbf{y}_1 = (y_1, \ldots, y_{n_1})', \mathbf{X}_1 = \begin{pmatrix} 1 & \ldots & 1 \\ x_1 & \ldots & x_{n_1} \end{pmatrix}',$$

be the OLS estimates of α and β based on the first n_1 observations. Now consider the augmented estimator

$$\begin{pmatrix} \widehat{\alpha}^* \\ \widehat{\beta}^* \end{pmatrix} = (\mathbf{X}_*'\mathbf{X}_*)^{-1}\mathbf{X}_*'\mathbf{y}, \mathbf{y} = (y_1, \ldots, y_n)', \mathbf{X}_* = \begin{pmatrix} 1 & . & . & . & . & 1 \\ x_1 & . & x_{n_1} & \overline{x}_1 & . & \overline{x}_1 \end{pmatrix}',$$

where $\overline{x}_1 = n_1^{-1}\sum_{i=1}^{n_1} x_i$ and $\mathbf{X}_*$ is n by 2.

(a) Derive the bias, if any, of the above estimators. [8]

(b) Show that $\mathrm{var}(\widehat{\alpha}^*) \leq \mathrm{var}(\widehat{\alpha}_1)$, while $\mathrm{var}(\widehat{\beta}_1) = \mathrm{var}(\widehat{\beta}^*)$. [12]
Hint: you may find it convenient to rewrite $\widehat{\beta}^*$ in deviation from means form.

4. Consider the functional errors in variables model $y_i^* = \beta x_i^*$, where both y^* and x^* are measured with error, i.e. $y_i = y_i^* + \varepsilon_i$ and $x_i = x_i^* + u_i$, where ε_i and u_i are mutually independent and both mean zero, while x_i^* are fixed in repeated samples. The ratio of sample means estimator $\overline{\beta} = \sum_{i=1}^n y_i / \sum_{i=1}^n x_i$ is proposed.

(a) Show that $\overline{\beta}$ is an instrumental variables estimator giving the instruments. [4]

(b) Prove that $\overline{\beta} \xrightarrow{P} \beta$, and derive the limiting distribution of $\sqrt{n}(\overline{\beta} - \beta)$ (You can assume that $n^{-1}\sum_{i=1}^n x_i^* \to \mu^* \neq 0$). [10]

(c) Provide consistent confidence intervals for β OR define a test of the hypothesis that $\beta = 0$. [6]