AAEC/ECON 5126 MIDTERM EXAM: SOLUTIONS

SPRING 2020 / INSTRUCTOR: KLAUS MOELTNER

You have until ***Thu, March 26, midnight****, to complete this exam. Please create a* ***single pdf with all text and equations typed in LaTeX****, and submit it via Canvas. Q1 asks for some graphs - you can create those in Powerpoint or some other simple graphing software and add it to your pdf, OR write it on paper and take picture with your cell phone, and add the picture as a separate file when you submit the midterm (or figure out how to add it to your pdf). I would like to try Canvas for your first attempt to submit the exam. If this gives you trouble, you can also send it to me via e-mail attachment. I do not want you to worry about how to submit.*

1. Question I (15 points)

Consider the CLRM for a full sample of n observations:

$$\mathbf{y} = \beta_0\mathbf{i} + \beta_1\mathbf{d} + \beta_2\mathbf{x} + \beta_3\left(\mathbf{x} * \mathbf{d}\right) + \boldsymbol{\epsilon}, \tag{1}$$

where all the usual assumptions hold, $\mathbf{i}$ is an n by 1 vector of "1"s, $\mathbf{d}$ is an indicator variable that takes the value of "1" for person i if that person belongs to a specific category (e.g. "female"), and a value of "0" otherwise, and $\mathbf{x}$ is a continuous explanatory variable. The last term, $\mathbf{x} * \mathbf{d}$ is an interaction term of $\mathbf{x}$ and $\mathbf{d}$, with elements:

$$\tilde{\mathbf{x}} = \mathbf{x} * \mathbf{d} = \begin{bmatrix} x_1 * d_1 \\ x_2 * d_2 \\ \vdots \\ x_n * d_n \end{bmatrix}$$

Assume that n_1 of the n individuals belong to the indexed category for whom $d_i = 1$ (with $0 < n_1 < n$).

(I) *(4 pts.)* For a single observation, show the form of: $E\left(y_i|d_i = 0, x_i\right)$, $E\left(y_i|d_i = 1, x_i\right)$, and $E\left(y_i|d_i = 1, x_i\right) - E\left(y_i|d_i = 0, x_i\right)$. Is the marginal effect ("slope") of x_i on $E\left(y_i|x_i, d_i\right)$ the same for both groups or not? How about the intercept?

(II) *(4 pts.)* Draw a graph for each of the following situations, with x on the x-axis, and y on the y-axis. Each graph should show two lines, one for $E\left(y|d = 0, x\right)$, and one for $E\left(y|d = 1, x\right)$. Throughout assume that $\beta_0, \beta_2 > 0$:

(a) $\beta_1 > 0$ and $\beta_3 > 0$

Date: March 19, 2020.

(b) $\beta_1 < 0$ and $\beta_3 > 0$
(c) $\beta_1 > 0$ and $\beta_3 < 0$
(d) $\beta_1 < 0$ and $\beta_3 < 0$

(III) *(4 pts.)* Consider the following changes to your data:

(a) Assume somebody eliminates all cases for which $d_i = 1$ from the sample. Could you still estimate the model in (1)? Why or why not? Can you show a simpler model that could still be estimated?
(b) Assume, instead, that you have all the data, but $\mathbf{x}$ is also a binary (0/1) indicator with some 0's and some 1's. What problem could this cause for the estimation of β_3, especially in a small sample?

(IV) *(3 pts.)* Going back to the original model in (1) with all the original assumptions, assume you want to perform the following hypothesis tests:

(a) The intercepts are the same for both groups.
(b) The slopes are the same for both groups.
(c) Intercepts and slopes are the same for both groups.

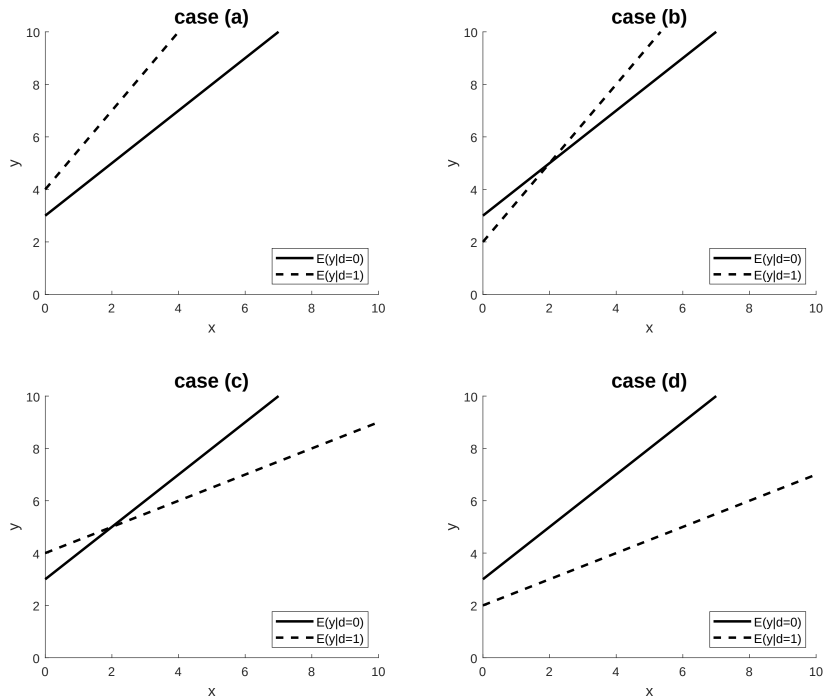
For *each test* do the following:
(i) Write the null hypothesis in mathematical notation (i.e. in terms of the individual β's),
(ii) Derive the $\mathbf{R}$-matrix and the $\mathbf{q}$-vector for the test $\mathbf{R}\boldsymbol{\beta} - \mathbf{q} = \mathbf{0}$,
(iii) Denote the number of restrictions for the corresponding F-test.

Solutions:

(I)

$$E\left(y_i|d_i = 0, x_i\right) = \beta_0 + \beta_2 * x_i$$
$$E\left(y_i|d_i = 1, x_i\right) = \left(\beta_0 + \beta_1\right) + \left(\beta_2 + \beta_3\right) * x_i$$
$$E\left(y_i|d_i = 1, x_i\right) - E\left(y_i|d_i = 0, x_i\right) = \beta_1 + \beta_3 * x_i$$

So in this case both the slopes (marginal effect of x_i) and the intercepts differ between the two groups.



(II)

(III) (a) No, since both $\mathbf{d}$ and $\mathbf{x} * \mathbf{d}$ would produce all zeros, violating the full-rank condition (A2) of the CLRM. Once could still estimate $\mathbf{y} = \beta_0 \mathbf{i} + \beta_2 \mathbf{x} + \boldsymbol{\epsilon}$, using only the $d_i = 0$ cases.

(b) This could produce a situation where all or almost all elements of $\mathbf{x} * \mathbf{d}$ are 0's or 1's, which would make it impossible / difficult to estimate β_3.

(IV) $H_0 : \beta_1 = 0$
$\mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}$
$q = 0$
$J = 1$

(V) $H_0 : \beta_3 = 0$
$\mathbf{R} = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}$
$q = 0$
$J = 1$

(VI) $H_0 : \beta_1 = 0, \beta_3 = 0$
$\mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
$\mathbf{q} = \begin{bmatrix} 0 & 0 \end{bmatrix}'$
$J = 2$

2. Question II (7 points)

Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be asymptotically normal, consistent estimators for θ_1 and θ_2, respectively, with $\theta_1, \theta_2 > 0$.

(I) *(1 pt.)* Find a consistent estimator for $\gamma = \frac{log(\theta_1)}{log(\theta_2)}$, and state your reasoning, where "log" denotes the natural logarithm. Call the estimator $\hat{\gamma}$.

(II) *(4 pts.)* Let V_{11} and V_{22} be the asymptotic variances of θ_1 and θ_2, respectively, and assume the asymptotic covariance is zero. Show the general form of the asymptotic variance of $\hat{\gamma}$ (call it $V_{\hat{\gamma}}$), as well as the general form of its asymptotic standard error. The asymptotic variance should be shown as a simple sum of two multiplicative terms.

(III) *(1 pt.)* Assume you have some (asymptotic) estimates for these variances and standard errors, i.e. $\hat{V}_{11}$, $\hat{V}_{22}$, $\hat{s.e.}(\hat{\gamma})$. Show how you would set up a hypothesis test of $\hat{\gamma} = 1$ (you can assume asymptotic normality for $\hat{\gamma}$. Show as many details as possible.

(IV) *(1 pt.)* How would you set up a logically equivalent hypothesis test involving $\hat{\theta}_1$ and $\hat{\theta}_2$? Show as many details as possible.

Solutions:

(I) A good candidate might be $\hat{\gamma} = \frac{log(\hat{\theta}_1)}{log(\hat{\theta}_2)}$. By the Slutsky theorem we have:

$$\text{plim}\left(\hat{\gamma}\right) = \text{plim}\left(\frac{log\left(\hat{\theta}_1\right)}{log\left(\hat{\theta}_2\right)}\right) = \frac{\text{plim}\left(log\left(\hat{\theta}_1\right)\right)}{\text{plim}\left(log\left(\hat{\theta}_2\right)\right)} = \frac{log\left(\text{plim}\left(\hat{\theta}_1\right)\right)}{log\left(\text{plim}\left(\hat{\theta}_2\right)\right)} = \frac{log\left(\theta_1\right)}{log\left(\theta_2\right)} = \gamma$$

(II) Using the Delta method we have:

$$V_{\hat{\gamma}} = \begin{bmatrix} \frac{\partial\gamma}{\partial\theta_1} & \frac{\partial\gamma}{\partial\theta_2} \end{bmatrix} * \begin{bmatrix} V_{11} & 0 \\ 0 & V_{22} \end{bmatrix} * \begin{bmatrix} \frac{\partial\gamma}{\partial\theta_1} \\ \frac{\partial\gamma}{\partial\theta_2} \end{bmatrix} =$$

$$\begin{bmatrix} \left(\theta_1 * log\left(\theta_2\right)\right)^{-1} & -log\left(\theta_1\right)\left(log\left(\theta_2\right)\right)^{-2}\theta_2^{-1} \end{bmatrix} * \begin{bmatrix} V_{11} & 0 \\ 0 & V_{22} \end{bmatrix} * \begin{bmatrix} \left(\theta_1 * log\left(\theta_2\right)\right)^{-1} \\ -log\left(\theta_1\right)\left(log\left(\theta_2\right)\right)^{-2}\theta_2^{-1} \end{bmatrix} =$$

$$\left(\theta_1 * log\left(\theta_2\right)\right)^{-2} * V_{11} + \left(-log\left(\theta_1\right)\right)^2\left(log\left(\theta_2\right)\right)^{-4}\theta_2^{-2} * V_{22}$$

$$s.e.\left(\hat{\gamma}\right) = \sqrt{V_{\hat{\gamma}}}$$

(III) H0: $\gamma = 1$

Determine level of significance (e.g. $\alpha = 0.05$). Compute z-value:

$$z = \frac{\hat{\gamma} - 1}{\hat{s.e.}\left(\hat{\gamma}\right)}$$

Reject if $p(z) < \alpha/2$ (or if $|z| > z_{\alpha/2}$), else fail to reject

(IV) H0: $\theta_1 - \theta_2 = 0$
Compute:

$$\hat{d} = \left(\hat{\theta}_1 - \hat{\theta}_2\right), \quad \text{and}$$
$$s.\hat{e}.\left(\hat{d}\right) = \sqrt{\hat{V}_d} = \sqrt{\hat{V}_{11} + \hat{V}_{22}}$$

Determine level of significance (e.g. $\alpha = 0.05$). Compute z-value:

$$z = \frac{\hat{d}}{s.\hat{e}.\left(\hat{d}\right)}$$

Reject if $p(z) < \alpha/2$ (or if $|z| > z_{\alpha/2}$), else fail to reject

3. Question III (8 points)

Consider the Rayleigh distribution with scale parameter γ:

$$
\begin{aligned}
&f(y) = \gamma^{-2} * y * exp\left(-\left(2\gamma^2\right)^{-1} y^2\right) \quad \gamma > 0,\ 0 < y < \infty \\
&E(y) = \gamma (\pi/2)^{1/2} \\
&V(y) = \gamma^2 (2 - (\pi/2))
\end{aligned}
\tag{2}
$$

(I) *(1 pt.)* Compute $E\left(y^2\right)$ - you will need this for part (IV) below. (Hint: Use a well-known relationship between $V(y)$, $E(y)$, and $E\left(y^2\right)$)

(II) *(2 pts.)* Consider a sample of n draws of $y_i, i = 1 \ldots n$ from this distribution. Derive the sample log-likelihood function $lnL(\gamma)$ and the sample gradient $g(\gamma)$. Simplify as much as possible.

(III) *(1 pt.)* Solve for the MLE estimator $\hat{\gamma}$.

(IV) *(2 pts.)* Using the sample gradient, show that the score identiy holds.

(V) *(2 pts.)* Derive the sample Hessian $H(\gamma)$ and show that your result in (II) is indeed a maximum.

Solutions:

(I)

$$E\left(y^2\right) = V(y) + (E(y))^2 = \gamma^2 (2 - (\pi/2)) + \gamma^2 (\pi/2) = 2\gamma^2$$

(II)

$$
\begin{aligned}
&lnl(\gamma) = -2ln\gamma + lny_i - \left(2\gamma^2\right)^{-1} y_i^2 \\
&lnL(\gamma) = -2nln\gamma + \sum_{i=1}^{n} lny_i - \left(2\gamma^2\right)^{-1} \sum_{i=1}^{n} y_i^2 \\
&g(\gamma) = -2n\gamma^{-1} + 4\gamma\left(2\gamma^2\right)^{-2} \sum_{i=1}^{n} y_i^2 = -2n\gamma^{-1} + \gamma^{-3} \sum_{i=1}^{n} y_i^2
\end{aligned}
$$

(III)

$$\hat{\gamma} = \sqrt{\frac{\sum_{i=1}^{n} y_i^2}{2n}}$$

(IV)

$$E_y(g(\gamma)) = -2n\gamma^{-1} + \gamma^{-3} E\left(\sum_{i=1}^{n} y_i^2\right) = -2n\gamma^{-1} + \gamma^{-3} 2n\gamma^2 = 0$$

(V)

$$H\left(\gamma\right) = 2n\gamma^{-2} - 3\gamma^{-4}\sum_{i=1}^{n} y_i^2$$

$$H\left(\hat{\gamma}\right) = H\left((2n)^{-1/2}\left(\sum_{i=1}^{n} y_i^2\right)^{1/2}\right) =$$

$$2n\left((2n)\left(\sum_{i=1}^{n} y_i^2\right)^{-1}\right) - 3\left(\sum_{i=1}^{n} y_i^2\right)\left((2n)^2\left(\sum_{i=1}^{n} y_i^2\right)^{-2}\right) =$$

$$-8n^2\left(\sum_{i=1}^{n} y_i^2\right)^{-1} < 0$$