AAEC/ECON 5126 FINAL EXAM: SOLUTIONS

SPRING 2016 / INSTRUCTOR: KLAUS MOELTNER

This exam is open-book, open-notes, but please work strictly on your own. Please make sure your name is on every sheet you're handing in. You have 120 minutes to complete this exam. You can collect a maximum of 50 points. Each question is scored as indicated below. Vectors are given in lower-case boldface. Matrices are written in upper-case boldface.

Question I (20 points): De-meaned regression

Consider the CLRM:
$$\mathbf{y} = \mathbf{i}\beta_0 + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim n\left(\mathbf{0}, \sigma^2\mathbf{I}\right) \tag{1}$$

Part (a) *3 points*

Write down the solution for OLS estimator for $\boldsymbol{\beta}$ (call it $\mathbf{b}$) in partitioned regression form.
(You do NOT need to derive the solution mathematically, just show it)

Solution:
$$\begin{aligned} &\mathbf{b} = \left(\mathbf{X}'\mathbf{M}_0\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{M}_0\mathbf{y}, \quad \text{where} \\ &\mathbf{M}_0 = \mathbf{I} - \mathbf{i}\left(\mathbf{i}'\mathbf{i}\right)^{-1}\mathbf{i}' \end{aligned} \tag{2}$$

Part (b) *5 points*

Now consider a version of the model in (1) without an intercept, and with a de-meaned $\mathbf{X}$ matrix, that is an $\mathbf{X}$ matrix with the mean of each column subtracted from each observation in that column, for all columns in $\mathbf{X}$. Call the de-meaned matrix $\tilde{\mathbf{X}}$. Your model is now:

$$\mathbf{y} = \tilde{\mathbf{X}}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim n\left(\mathbf{0}, \sigma^2\mathbf{I}\right) \tag{3}$$

Show the solution for the OLS estimator for this model (call it $\tilde{\mathbf{b}}$) - how does it compare to the solution in part (a)?

Solution:
Note that de-meaning $\mathbf{X}$ amounts to pre-multiplying $\mathbf{X}$ by $\mathbf{M}_0$, i.e. $\tilde{\mathbf{X}} = \mathbf{M}_0\mathbf{X}$ - recognizing this

Date: May 7, 2016.

is key for this entire question. Then:

$$\begin{aligned}
&\tilde{\mathbf{b}} = \left(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\right)^{-1}\tilde{\mathbf{X}}'\mathbf{y} = \\
&\left(\mathbf{X}'\mathbf{M}_0\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{M}_0\mathbf{y} = \mathbf{b}
\end{aligned} \tag{4}$$

Part (c) *5 points*

Now consider a version of the model in (1) without an intercept, and with both a de-meaned $\mathbf{X}$ matrix and $\mathbf{y}$ vector. Call the de-meaned $\mathbf{X}$ matrix $\tilde{\mathbf{X}}$, and the de-meaned $\mathbf{y}$ vector $\tilde{\mathbf{y}}$. Your model is now:

$$\tilde{\mathbf{y}} = \tilde{\mathbf{X}}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim n\left(\mathbf{0}, \sigma^2\mathbf{I}\right) \tag{5}$$

Show the solution for the OLS estimator for this model (call it $\hat{\mathbf{b}}$) - how does it compare to the solutions in parts (a) and (b)?

Solution:

$$\begin{aligned}
&\hat{\mathbf{b}} = \left(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\right)^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{y}} = \\
&\left(\mathbf{X}'\mathbf{M}_0\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{M}_0\mathbf{M}_0\mathbf{y} = \tilde{\mathbf{b}} = \mathbf{b} \quad \text{since} \\
&\mathbf{M}_0' = \mathbf{M}_0, \quad \text{and} \quad \mathbf{M}_0\mathbf{M}_0 = \mathbf{M}_0
\end{aligned} \tag{6}$$

Part (d) *7 points*

Now consider a linear regression model without a constant term, a de-meaned $\mathbf{X}$, and a non-zero mean error term, i.e.:

$$\mathbf{y} = \tilde{\mathbf{X}}\boldsymbol{\gamma} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim n\left(\mathbf{i}\mu, \sigma^2\mathbf{I}\right) \tag{7}$$

Show that the OLS estimator for $\boldsymbol{\gamma}$ (call it $\mathbf{g}$) is unbiased.

Solution:

$$
\begin{aligned}
&E\left(\mathbf{g}|\mathbf{X}\right) = \\
&E\left(\left(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\right)^{-1}\tilde{\mathbf{X}}'\mathbf{y}\right) = \\
&E\left(\left(\mathbf{X}'\mathbf{M}_0\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{M}_0\mathbf{y}\right) = \\
&E\left(\left(\mathbf{X}'\mathbf{M}_0\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{M}_0\mathbf{M}_0\mathbf{X}\boldsymbol{\gamma}\right) + \\
&E\left(\left(\mathbf{X}'\mathbf{M}_0\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{M}_0\boldsymbol{\epsilon}\right) = \\
&\boldsymbol{\gamma} + \left(\mathbf{X}'\mathbf{M}_0\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{M}_0\mathbf{i}\mu = \boldsymbol{\gamma}
\end{aligned}
\tag{8}
$$

since $\mathbf{M}_0\mathbf{i}\mu = \mathbf{0}$.

Question I (10 points): Heteroskedasticity

Consider a linear regression model with dependent variable $\mathbf{y}$, data matrix $\mathbf{X}$ (including a column of ones), coefficient vector $\boldsymbol{\beta}$, and error vector $\boldsymbol{\epsilon}$. The sample size is n. Assume the model satisfies all CLRM assumptions, except for homoskedasticity. Specifically, a fraction of observations αn with $0 < \alpha < 1$ is associated with error terms that have variance σ_1^2, while the remaining $(1-\alpha)\,n$ observations have error variance σ_2^2. For convenience, assume your sample is sorted starting with the σ_1^2 cases, followed by the σ_2^2 cases.

Part (a) *2 points*
Show the explicit contents of the n-by-n variance-covariance matrix $\boldsymbol{\Omega}$ of the sorted error vector in terms of σ_1^2, σ_2^2, and identity and zero matrices of appropriate dimensions.

Solution:

$$\boldsymbol{\Omega} = \begin{bmatrix} \sigma_1^2 \mathbf{I}_{\alpha n} & \mathbf{0}_{\alpha n\, x\, (1-\alpha)n} \\ \mathbf{0}_{(1-\alpha)n\, x\, \alpha n} & \sigma_2^2 \mathbf{I}_{(1-\alpha)n} \end{bmatrix} \tag{9}$$

Part (b) *4 points*
Show the form of the OLS estimator and *derive* its variance (call it $V(\mathbf{b})$). In light of your finding, discuss the implications of ignoring the heteroskedasticity problem and using the conventional expression for $V(\mathbf{b})$ to derive standard errors and t-values.

Solution:

$$\begin{aligned} &\mathbf{b} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y} \\ &V(\mathbf{b}) = E\left(\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\epsilon}\boldsymbol{\epsilon}'\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\right) = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\Omega}\boldsymbol{\epsilon}'\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1} \end{aligned} \tag{10}$$

The naive variance estimator $s^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}$ will be misleading for two reasons: (1) it has the wrong structural form, and (2) s^2 will be a biased estimator for $\sigma_j^2, j = 1,2$.

Part (c) *4 points*
Now assume that there is strong indication that the group-wise heteroskedasticity is driven by an observed indicator variable $\mathbf{D}$, which takes the value of "0" for all σ_1^2 cases, and "1" for all σ_2^2 cases. Outline (in words) how you would derive a feasible GLS estimator (call it $\mathbf{b}_{FGLS}$) for this case. Make sure to show the explicit skedastic function you would use.

Solution:
Skedastic function: $\sigma_i^2 = exp\left(\gamma_0 + \gamma_1 D_i\right)$.

(1) Run OLS and capture the residuals.
(2) Regress the log of squared residuals against a column of ones and $\mathbf{D}$.
(3) Using the estimates for γ_0 and γ_1 generate predicted values for $\sigma_j^2, j = 1,2$.

(4) Insert these predicted values into $\mathbf{\Omega}$ and use this estimator of $\mathbf{\Omega}$ in the GLS formula.

Alternative answer:

(1) Run OLS and capture the residuals.

(2) Compute $\hat{\sigma_j^2} = \frac{\mathbf{e}_j'\mathbf{e}_j}{n_j}, j = 1, 2.$

(3) Insert these predicted values into $\mathbf{\Omega}$ and use this estimator of $\mathbf{\Omega}$ in the GLS formula.

Question III (20 points): Serial correlation

Consider the regression-on-intercept model

$$y_i = \mu + \epsilon_i \tag{11}$$

where the error term ϵ_i has mean zero, constant variance σ^2 and equal correlation ρ with all other errors, with $|\rho| < 1$.

Part (a) *4 points*
Derive the covariance between any two errors and show explicitly the full variance-covariance matrix $\mathbf{\Omega}$.

Solution:

$$\begin{aligned}
&corr\left(\epsilon_i, \epsilon_j\right) = \rho = \frac{cov\left(\epsilon_i, \epsilon_j\right)}{\sigma^2} \rightarrow \\
&cov\left(\epsilon_i, \epsilon_j\right) = \rho\sigma^2 \\
&\mathbf{\Omega} = \sigma^2 \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{bmatrix}
\end{aligned} \tag{12}$$

Part (b) *10 points*
Derive the OLS estimator b and its variance. Prove that b is *not* a consistent estimator for μ by showing that the limit of its variance is not zero, as would be required for consistency.

Solution:

$$
\begin{aligned}
& b = \left(\mathbf{i}'\mathbf{i}\right)^{-1}\mathbf{i}'\mathbf{y} \\
& V(b) = \left(\mathbf{i}'\mathbf{i}\right)^{-1}\mathbf{i}'\mathbf{\Omega}\mathbf{i}\left(\mathbf{i}'\mathbf{i}\right)^{-1} = \\
& \frac{\sigma^2}{n^2}\begin{bmatrix}1 & 1 & \dots & 1\end{bmatrix}\begin{bmatrix}1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1\end{bmatrix}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = \\
& \frac{\sigma^2}{n^2} = \begin{bmatrix}1+(n-1)\rho & 1+(n-1)\rho & \dots & 1+(n-1)\rho\end{bmatrix}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = \\
& \frac{\sigma^2}{n^2}\left(n\left(1+(n-1)\rho\right)\right) = \frac{\sigma^2}{n} + \sigma^2\rho - \frac{\sigma^2\rho}{n} \\
& \lim_{n\to\infty}\left(V(b)\right) = \lim_{n\to\infty}\left(\frac{\sigma^2}{n}\right) + \lim_{n\to\infty}\left(\sigma^2\rho\right) - \lim_{n\to\infty}\left(\frac{\sigma^2\rho}{n}\right) = \sigma^2\rho \neq 0
\end{aligned}
\tag{13}
$$

Part (c) *6 points*
Reconsider the original model, but now define $\mathbf{\Omega}$ as a diagonal matrix with heteroskedastic variances, i.e. $V(\epsilon_i) = \sigma_i^2$, with $\sigma_i^2 < n, \forall i$. Using the same approach as in part (b), show that for this model the OLS estimator is consistent.

Solution:

$$
\begin{aligned}
& V(b) = \tfrac{1}{n^2}\mathbf{i}'\mathbf{\Omega}\mathbf{i} = \\
& \tfrac{1}{n^2}\begin{bmatrix}1 & 1 & \dots & 1\end{bmatrix}\begin{bmatrix}\sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2\end{bmatrix}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = \\
& \tfrac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2 \\
& \lim_{n\to\infty}\left(V(b)\right) = \lim_{n\to\infty}\left(\tfrac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2\right) = 0
\end{aligned}
\tag{14}
$$