Exercise Sheet 2: Specification Tests for Panel Data Models

– Draft Solution –

Question 1

Consider the error-components model

$$\mathbf{y}_i = \mathbf{x}_i\boldsymbol{\beta} + \mathbf{v}_i, \qquad \mathbf{v}_i = \boldsymbol{\iota}_T c_i + \mathbf{u}_i. \tag{1}$$

Suppose that, conditional on $\mathbf{x}_i$, $\mathbf{v}_i \sim \text{Normal}(\mathbf{0}, \boldsymbol{\Omega})$ where

$$\boldsymbol{\Omega} = \sigma_c^2 \boldsymbol{\iota}_T \boldsymbol{\iota}_T' + \sigma_u^2 \mathbf{I}_T.$$

In the following, derive step by step the Breusch-Pagan LM test of the null hypothesis $\sigma_c^2 = 0$. Use (without proof) the fact that the information matrix is block diagonal between $\boldsymbol{\beta}$ and the variance parameters $\boldsymbol{\theta} = (\sigma_u^2, \sigma_c^2)'$ which allows you to compute score and Hessian solely for $\boldsymbol{\theta}$.

(a) Show that for random sampling of $\mathbf{y}_i$ and $\mathbf{x}_i$, $i = 1, \ldots, N$, the conditional log-likelihood function is

$$\ell_i = \frac{1-T}{2}\log\sigma_u^2 - \frac{1}{2}\log(T\sigma_c^2 + \sigma_u^2) - \frac{(T\sigma_c^2 + \sigma_u^2)^{-1}}{2}\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i - \frac{(\sigma_u^2)^{-1}}{2}\mathbf{v}_i'\mathbf{Q}_T\mathbf{v}_i.$$

Note that the conditional log density of the multivariate normal distribution is, up to an irrelevant constant, $\log f(\mathbf{v}_i) = -\frac{1}{2}\log|\boldsymbol{\Omega}| - \frac{1}{2}\mathbf{v}_i'\boldsymbol{\Omega}^{-1}\mathbf{v}_i$ and recall that $\boldsymbol{\Omega}^{-1} = \frac{\phi^2}{\sigma_u^2}\mathbf{J}_T + \frac{1}{\sigma_u^2}\mathbf{Q}_T$. Hint: use the rule $|\mathbf{A}'\mathbf{A} + \mathbf{I}_n| = |\mathbf{A}\mathbf{A}' + I_m|$, where $\mathbf{A}$ is a $m \times n$ matrix, and the rule $|c\mathbf{M}| = c^T|\mathbf{M}|$, where c is a scalar and $\mathbf{M}$ is a $T \times T$ matrix.

<u>Answer:</u> Based on the log density, the conditional log likelihood is, up to an irrelevant constant,

$$\ell_i = -\frac{1}{2}\log|\boldsymbol{\Omega}| - \frac{1}{2}\mathbf{v}_i'\boldsymbol{\Omega}^{-1}\mathbf{v}_i, \qquad \mathbf{v}_i = \mathbf{y}_i - \mathbf{x}_i\boldsymbol{\beta}.$$

To simplify it, first consider $|\boldsymbol{\Omega}|$, substitute its definition and apply the rules stated above:

$$|\boldsymbol{\Omega}| = |\sigma_c^2 \boldsymbol{\iota}_T \boldsymbol{\iota}_T' + \sigma_u^2 \mathbf{I}_T| = \left| \sigma_u^2 \left(\frac{\sigma_c^2}{\sigma_u^2} \boldsymbol{\iota}_T \boldsymbol{\iota}_T' + \mathbf{I}_T \right) \right| = (\sigma_u^2)^T \left| \frac{\sigma_c^2}{\sigma_u^2} \boldsymbol{\iota}_T \boldsymbol{\iota}_T' + \mathbf{I}_T \right|$$
$$= (\sigma_u^2)^T \left| \frac{\sigma_c^2}{\sigma_u^2} \boldsymbol{\iota}_T' \boldsymbol{\iota}_T + \mathbf{I}_1 \right| = (\sigma_u^2)^T \left(\frac{\sigma_c^2}{\sigma_u^2} T + 1 \right) = (\sigma_u^2)^{T-1} (T\sigma_c^2 + \sigma_u^2). \quad (2)$$

Next consider $\boldsymbol{\Omega}^{-1}$ and substitute the definition of ϕ:

$$\boldsymbol{\Omega}^{-1} = \frac{\phi^2}{\sigma_u^2} \mathbf{J}_T + \frac{1}{\sigma_u^2} \mathbf{Q}_T = \frac{1}{T\sigma_c^2 + \sigma_u^2} \mathbf{J}_T + \frac{1}{\sigma_u^2} \mathbf{Q}_T. \quad (3)$$

Now substitute (2) and (3) into the likelihood function and simplify:

$$\ell_i = -\frac{1}{2} \log[(\sigma_u^2)^{T-1} (T\sigma_c^2 + \sigma_u^2)] - \frac{1}{2} \mathbf{v}_i' \left(\frac{1}{T\sigma_c^2 + \sigma_u^2} \mathbf{J}_T + \frac{1}{\sigma_u^2} \mathbf{Q}_T \right) \mathbf{v}_i$$
$$= \frac{1-T}{2} \log \sigma_u^2 - \frac{1}{2} \log(T\sigma_c^2 + \sigma_u^2) - \frac{(T\sigma_c^2 + \sigma_u^2)^{-1}}{2} \mathbf{v}_i' \mathbf{J}_T \mathbf{v}_i - \frac{(\sigma_u^2)^{-1}}{2} \mathbf{v}_i' \mathbf{Q}_T \mathbf{v}_i \quad (4)$$

(b) Find the score $\mathbf{s}_i(\boldsymbol{\theta})$ and Hessian $\mathbf{H}_i(\boldsymbol{\theta})$ of ℓ_i with respect to the parameter vector $\boldsymbol{\theta} = (\sigma_u^2, \sigma_c^2)'$. Evaluate Score and Hessian under the null hypothesis $\sigma_c^2 = 0$.

Answer: To find the score,

$$\mathbf{s}_i(\boldsymbol{\theta}) = [s_{i,1}(\boldsymbol{\theta}), s_{i,2}(\boldsymbol{\theta})]' = \nabla_{\boldsymbol{\theta}} \ell_i(\boldsymbol{\theta})' = \left[\frac{\partial \ell_i}{\partial \sigma_u^2}, \frac{\partial \ell_i}{\partial \sigma_c^2} \right]',$$

we need the first derivatives of ℓ_i with respect to σ_u^2 and σ_c^2:

$$s_{i,1}(\boldsymbol{\theta}) = \frac{\partial \ell_i}{\partial \sigma_u^2} = \frac{1-T}{2} \sigma_u^{-2} - \frac{(T\sigma_c^2 + \sigma_u^2)^{-1}}{2} + \frac{(T\sigma_c^2 + \sigma_u^2)^{-2}}{2} \mathbf{v}_i' \mathbf{J}_T \mathbf{v}_i + \frac{\sigma_u^{-4}}{2} \mathbf{v}_i' \mathbf{Q}_T \mathbf{v}_i \quad (5)$$

$$s_{i,2}(\boldsymbol{\theta}) = \frac{\partial \ell_i}{\partial \sigma_c^2} = -\frac{T}{2} (T\sigma_c^2 + \sigma_u^2)^{-1} + \frac{T}{2} (T\sigma_c^2 + \sigma_u^2)^{-2} \mathbf{v}_i' \mathbf{J}_T \mathbf{v}_i. \quad (6)$$

For the Hessian,

$$\mathbf{H}_i(\boldsymbol{\theta}) = \begin{pmatrix} H_{i,11}(\boldsymbol{\theta}) & H_{i,12}(\boldsymbol{\theta}) \\ H_{i,21}(\boldsymbol{\theta}) & H_{i,22}(\boldsymbol{\theta}) \end{pmatrix} = \nabla_{\boldsymbol{\theta}} \mathbf{s}_i(\boldsymbol{\theta}) = \begin{pmatrix} \frac{\partial^2 \ell_i}{\partial \sigma_u^2 \partial \sigma_u^2} & \frac{\partial^2 \ell_i}{\partial \sigma_u^2 \partial \sigma_c^2} \\ \frac{\partial^2 \ell_i}{\partial \sigma_c^2 \partial \sigma_u^2} & \frac{\partial^2 \ell_i}{\partial \sigma_c^2 \partial \sigma_c^2} \end{pmatrix},$$

we have to find the second derivatives:

$$H_{i,11}(\boldsymbol{\theta}) = \frac{\partial^2 \ell_i}{\partial(\sigma_u^2)^2} = \frac{T-1}{2}\sigma_u^{-4} + \frac{(T\sigma_c^2+\sigma_u^2)^{-2}}{2} - \frac{\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i}{(T\sigma_c^2+\sigma_u^2)^3} - \frac{\mathbf{v}_i'\mathbf{Q}_T\mathbf{v}_i}{\sigma_u^6} \tag{7}$$

$$H_{i,22}(\boldsymbol{\theta}) = \frac{\partial^2 \ell_i}{\partial(\sigma_c^2)^2} = \frac{T^2}{2}(T\sigma_c^2+\sigma_u^2)^{-2} - T^2(T\sigma_c^2+\sigma_u^2)^{-3}\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i \tag{8}$$

$$H_{i,21}(\boldsymbol{\theta}) = \frac{\partial^2 \ell_i}{\partial\sigma_c^2\partial\sigma_u^2} = \frac{T}{2}(T\sigma_c^2+\sigma_u^2)^{-2} - T(T\sigma_c^2+\sigma_u^2)^{-3}\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i. \tag{9}$$

For later use (estimation under the null) we evaluate score and Hessian when setting $\sigma_c^2 = 0$ and thus using the restricted parameter vector $\boldsymbol{\theta}_{H_0} = (\sigma_u^2, 0)'$. This yields score

$$\begin{aligned}
s_{i,1}(\boldsymbol{\theta}_{H_0}) &= \frac{1-T}{2}\sigma_u^{-2} - \frac{1}{2}\sigma_u^{-2} + \frac{\sigma_u^{-4}}{2}\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i + \frac{\sigma_u^{-4}}{2}\mathbf{v}_i'\mathbf{Q}_T\mathbf{v}_i \\
&= -\frac{T}{2}\sigma_u^{-2} + \frac{\sigma_u^{-4}}{2}(\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i + \mathbf{v}_i'\mathbf{Q}_T\mathbf{v}_i) \\
&= -\frac{T}{2}\sigma_u^{-2} + \frac{\sigma_u^{-4}}{2}\mathbf{v}_i'\mathbf{v}_i \\
&= -\frac{T}{2}\sigma_u^{-4}(\sigma_u^2 - T^{-1}\mathbf{v}_i'\mathbf{v}_i)
\end{aligned} \tag{10}$$

where we use $\mathbf{J}_T + \mathbf{Q}_T = \mathbf{I}_T$, and

$$s_{i,2}(\boldsymbol{\theta}_{H_0}) = -\frac{T}{2}\sigma_u^{-2} + \frac{T}{2}\sigma_u^{-4}\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i = -\frac{T}{2}\sigma_u^{-4}(\sigma_u^2 - \mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i). \tag{11}$$

The Hessian becomes

$$H_{i,11}(\boldsymbol{\theta}_{H_0}) = \frac{T}{2}\sigma_u^{-4} - \sigma_u^{-6}\mathbf{v}_i'\mathbf{v}_i \tag{12}$$

$$H_{i,22}(\boldsymbol{\theta}_{H_0}) = \frac{T^2}{2}\sigma_u^{-4} - T^2\sigma_u^{-6}\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i \tag{13}$$

$$H_{i,21}(\boldsymbol{\theta}_{H_0}) = \frac{T}{2}\sigma_u^{-4} - T\sigma_u^{-6}\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i. \tag{14}$$

where we use again $\mathbf{J}_T + \mathbf{Q}_T = \mathbf{I}_T$ in the derivation of $H_{i,11}(\boldsymbol{\theta}_{H_0})$.

(c) Find the conditional expectation $\mathbf{A}(\mathbf{x}_i, \boldsymbol{\theta}) \equiv -\,\mathrm{E}[\mathbf{H}_i(\boldsymbol{\theta})|\mathbf{x}_i]$ under the null hypothesis $\sigma_c^2 = 0$.

Answer: First note that, by the RE assumptions,

$$\mathrm{E}[\mathbf{v}_i'\mathbf{v}_i|\mathbf{x}_i] = \mathrm{E}[v_{i1}^2 + \ldots + v_{iT}^2|\mathbf{x}_i] = T\sigma_u^2.$$

Next use the definition $\mathbf{J}_T = \mathbf{I}_T - \mathbf{Q}_T$ which yields

$$\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i = \mathbf{v}_i'(\mathbf{I}_T - \mathbf{Q}_T)\mathbf{v}_i = \mathbf{v}_i'\mathbf{v}_i - \mathbf{v}_i'\mathbf{Q}_T\mathbf{v}_i = \mathbf{v}_i'\mathbf{v}_i - (\mathbf{Q}_T\mathbf{v}_i)'(\mathbf{Q}_T\mathbf{v}_i).$$

Since $\mathbf{Q}_T$ is the within transformation matrix, we can write

$$(\mathbf{Q}_T\mathbf{v}_i)'(\mathbf{Q}_T\mathbf{v}_i) = \ddot{\mathbf{v}}_i'\ddot{\mathbf{v}}_i = \ddot{v}_{i1}^2 + \ldots + \ddot{v}_{iT}^2,$$

where $\ddot{v}_{it} = v_{it} - \bar{v}_i$. Recall from the discussion of the FE estimator that (under the null hypothesis $\sigma_c^2 = 0$)

$$\mathrm{E}[\ddot{v}_{it}^2|\mathbf{x}_i] = \frac{T-1}{T}\sigma_u^2.$$

Hence,

$$\begin{aligned}\mathrm{E}[\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i|\mathbf{x}_i] &= \mathrm{E}[\mathbf{v}_i'\mathbf{v}_i|\mathbf{x}_i] - \mathrm{E}[\mathbf{v}_i'\mathbf{Q}_T\mathbf{v}_i|\mathbf{x}_i] = T\sigma_u^2 - \mathrm{E}[\ddot{v}_{i1}^2 + \ldots + \ddot{v}_{iT}^2|\mathbf{x}_i]\\ &= T\sigma_u^2 - T\frac{T-1}{T}\sigma_u^2 = \sigma_u^2.\end{aligned}$$

Using these results, we can apply the conditional expectation elementwise to the Hessian, i.e., to (12), (13), and (14). This yields

$$\mathrm{E}[H_{i,11}(\boldsymbol{\theta}_{H_0})|\mathbf{x}_i] = \frac{T}{2}\sigma_u^{-4} - \sigma_u^{-6}T\sigma_u^2 = -\frac{T}{2}\sigma_u^{-4} \tag{15}$$

$$\mathrm{E}[H_{i,22}(\boldsymbol{\theta}_{H_0})|\mathbf{x}_i] = \frac{T^2}{2}\sigma_u^{-4} - T^2\sigma_u^{-6}\sigma_u^2 = -\frac{T^2}{2}\sigma_u^{-4} \tag{16}$$

$$\mathrm{E}[H_{i,21}(\boldsymbol{\theta}_{H_0})|\mathbf{x}_i] = \frac{T}{2}\sigma_u^{-4} - T\sigma_u^{-6}\sigma_u^2 = -\frac{T}{2}\sigma_u^{-4} \tag{17}$$

and thus

$$\mathbf{A}(\mathbf{x}_i, \boldsymbol{\theta}_{H_0}) = \begin{bmatrix} \frac{T}{2}\sigma_u^{-4} & \frac{T}{2}\sigma_u^{-4} \\ \frac{T}{2}\sigma_u^{-4} & \frac{T^2}{2}\sigma_u^{-4} \end{bmatrix} = \frac{T}{2}\sigma_u^{-4}\begin{bmatrix} 1 & 1 \\ 1 & T \end{bmatrix}. \tag{18}$$

(d) Find $\frac{1}{N}\sum_{i=1}^{N}\tilde{\mathbf{s}}_i = \frac{1}{N}\sum_{i=1}^{N}\mathbf{s}_i(\tilde{\boldsymbol{\theta}})$ and $\frac{1}{N}\sum_{i=1}^{N}\tilde{\mathbf{A}}_i = \frac{1}{N}\sum_{i=1}^{N}\mathbf{A}(\mathbf{x}_i,\tilde{\boldsymbol{\theta}})$, where $\tilde{\boldsymbol{\theta}}$ is the CML estimator under the null hypothesis $\sigma_c^2 = 0$. Without proof use $\tilde{\sigma}_u^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{v}_{it}^2$, where $\tilde{v}_{it}$ are the pooled OLS residuals.

Answer: Let us start with the score. Put (10) and (11) into a vector, substitute $\tilde{\boldsymbol{\theta}}$ and average over all N observations:

$$\frac{1}{N}\sum_{i=1}^{N}\tilde{\mathbf{s}}_i = -\frac{T}{2}\tilde{\sigma}_u^{-4}\begin{bmatrix}\frac{1}{N}\sum_{i=1}^{N}(\tilde{\sigma}_u^2 - T^{-1}\mathbf{v}_i'\mathbf{v}_i)\\ \frac{1}{N}\sum_{i=1}^{N}(\tilde{\sigma}_u^2 - \mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i)\end{bmatrix} = -\frac{T}{2}\tilde{\sigma}_u^{-4}\begin{bmatrix}\tilde{\sigma}_u^2 - \frac{1}{NT}\sum_{i=1}^{N}\mathbf{v}_i'\mathbf{v}_i\\ \tilde{\sigma}_u^2 - \frac{1}{N}\sum_{i=1}^{N}\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i\end{bmatrix}$$

$$= -\frac{T}{2}\tilde{\sigma}_u^{-4}\begin{bmatrix}\tilde{\sigma}_u^2 - \tilde{\sigma}_u^2\\ \tilde{\sigma}_u^2 - \frac{1}{N}\sum_{i=1}^{N}\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i\end{bmatrix} = -\frac{T}{2}\tilde{\sigma}_u^{-4}\begin{bmatrix}0\\ \tilde{\sigma}_u^2 - \frac{1}{N}\sum_{i=1}^{N}\mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i\end{bmatrix} \quad (19)$$

Similarly, for the conditional expectation of the Hessian take (18), substitute $\tilde{\boldsymbol{\theta}}$ and average over all N observations:

$$\frac{1}{N}\sum_{i=1}^{N}\tilde{\mathbf{A}}_i = \frac{T}{2}\tilde{\sigma}_u^{-4}\begin{bmatrix}1 & 1\\ 1 & T\end{bmatrix}. \quad (20)$$

(e) Show that the LM statistic

$$LM = N\left(\frac{1}{N}\sum_{i=1}^{N}\tilde{\mathbf{s}}_i\right)'\left(\frac{1}{N}\sum_{i=1}^{N}\tilde{\mathbf{A}}_i\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\tilde{\mathbf{s}}_i\right)$$

can be simplified to

$$LM = \frac{NT}{2(T-1)}\left(\frac{\frac{1}{N}\sum_{i=1}^{N}\tilde{\mathbf{v}}_i'\mathbf{J}_T\tilde{\mathbf{v}}_i}{\tilde{\sigma}_u^2} - 1\right)^2.$$

Answer: First invert the conditional expectation of the Hessian:

$$\left(\frac{1}{N}\sum_{i=1}^{N}\tilde{\mathbf{A}}_i\right)^{-1} = \frac{2\tilde{\sigma}_u^4}{T}\begin{bmatrix}1 & 1\\ 1 & T\end{bmatrix}^{-1} = \frac{2\tilde{\sigma}_u^4}{T(T-1)}\begin{bmatrix}T & -1\\ -1 & 1\end{bmatrix}.$$

Next, substitute it together with (19) into the LM statistic:

$$
\begin{aligned}
LM &= N\frac{T^2\tilde{\sigma}_u^{-8}}{4}\frac{2\tilde{\sigma}_u^4}{T(T-1)}\left[0,\quad \tilde{\sigma}_u^2 - \tfrac{1}{N}\textstyle\sum_{i=1}^N \mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i\right]\begin{bmatrix} T & -1 \\ -1 & 1\end{bmatrix}\begin{bmatrix} 0 \\ \tilde{\sigma}_u^2 - \tfrac{1}{N}\sum_{i=1}^N \mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i\end{bmatrix} \\
&= \frac{NT}{2(T-1)}\tilde{\sigma}_u^{-4}\left(\frac{1}{N}\sum_{i=1}^N \mathbf{v}_i'\mathbf{J}_T\mathbf{v}_i - \tilde{\sigma}_u^2\right)^2 \\
&= \frac{NT}{2(T-1)}\left(\frac{\frac{1}{N}\sum_{i=1}^N \tilde{\mathbf{v}}_i'\mathbf{J}_T\tilde{\mathbf{v}}_i}{\tilde{\sigma}_u^2} - 1\right)^2 \qquad (21)
\end{aligned}
$$

(f) Show that the LM statistic can be reformulated to

$$
LM = \frac{NT(T-1)}{2}\left(\frac{\ddot{\sigma}_u^2}{\tilde{\sigma}_u^2} - 1\right)^2,
$$

where $\ddot{\sigma}_u^2 \equiv \frac{1}{N(T-1)}\sum_{i=1}^N\sum_{t=1}^T(\tilde{v}_{it} - \bar{\tilde{v}}_i)^2$. Interpret the statistic.

Answer: Note that

$$
\tilde{\mathbf{v}}_i'\mathbf{J}_T\tilde{\mathbf{v}}_i = \tilde{\mathbf{v}}_i'(\mathbf{I}_T - \mathbf{Q}_T)\tilde{\mathbf{v}}_i = \tilde{\mathbf{v}}_i'\tilde{\mathbf{v}}_i - \tilde{\mathbf{v}}_i'\mathbf{Q}_T\tilde{\mathbf{v}}_i.
$$

Hence,

$$
\frac{1}{N}\sum_{i=1}^N \tilde{\mathbf{v}}_i'\mathbf{J}_T\tilde{\mathbf{v}}_i = \frac{1}{N}\sum_{i=1}^N \tilde{\mathbf{v}}_i'\tilde{\mathbf{v}}_i - \frac{1}{N}\sum_{i=1}^N \tilde{\mathbf{v}}_i'\mathbf{Q}_T\tilde{\mathbf{v}}_i = T\tilde{\sigma}_u^2 - \frac{T-1}{N(T-1)}\sum_{i=1}^N(\mathbf{Q}_T\tilde{\mathbf{v}}_i)'(\mathbf{Q}_T\tilde{\mathbf{v}}_i).
$$

Since $\mathbf{Q}_T$ is the within transformation matrix, we can write

$$
(\mathbf{Q}_T\tilde{\mathbf{v}}_i)'(\mathbf{Q}_T\tilde{\mathbf{v}}_i) = (\tilde{v}_{i1} - \bar{\tilde{v}}_i)^2 + \ldots + (\tilde{v}_{iT} - \bar{\tilde{v}}_i)^2 = \sum_{t=1}^T(\tilde{v}_{it} - \bar{\tilde{v}}_i)^2,
$$

where $\bar{\tilde{v}}_i = \frac{1}{T}\sum_{t=1}^T \tilde{v}_{it}$ is the mean of the POLS residuals of individual i. Substituting yields

$$
\frac{1}{N}\sum_{i=1}^N \tilde{\mathbf{v}}_i'\mathbf{J}_T\tilde{\mathbf{v}}_i = T\tilde{\sigma}_u^2 - \frac{T-1}{N(T-1)}\sum_{i=1}^N\sum_{t=1}^T(\tilde{v}_{it} - \bar{\tilde{v}}_i)^2.
$$

Now recall from the discussion of the FE estimator that

$$
\mathrm{E}\left(\frac{1}{N(T-1)}\sum_{i=1}^N\sum_{t=1}^T(v_{it} - \bar{v}_i)^2\right) = \sigma_u^2.
$$

This suggests to define the within-type estimator

$$\ddot{\sigma}_u^2 \equiv \frac{1}{N(T-1)} \sum_{i=1}^{N} \sum_{t=1}^{T} (\tilde{v}_{it} - \bar{\tilde{v}}_i)^2.$$

Hence,

$$\frac{1}{N} \sum_{i=1}^{N} \tilde{\mathbf{v}}_i' \mathbf{J}_T \tilde{\mathbf{v}}_i = T\tilde{\sigma}_u^2 - (T-1)\ddot{\sigma}_u^2.$$

Substituting this into the expression for the LM statistic (21) yields

$$LM = \frac{NT}{2(T-1)} \left(\frac{T\tilde{\sigma}_u^2 - (T-1)\ddot{\sigma}_u^2}{\tilde{\sigma}_u^2} - 1 \right)^2 = \frac{NT(T-1)}{2} \left(\frac{\ddot{\sigma}_u^2}{\tilde{\sigma}_u^2} - 1 \right)^2.$$

This statistic lends to a nice interpretation. Under the null $v_{it} = u_{it}$, both estimators of σ_u^2 are consistent and thus

$$\frac{\ddot{\sigma}_u^2}{\tilde{\sigma}_u^2} \xrightarrow{\text{p}} 1.$$

Under the alternative, $\ddot{\sigma}_u^2$ is still consistent because it wipes out the individual effect but $\tilde{\sigma}_u^2$ is not. Hence, the LM statistic effectively checks whether these two estimators are similar (do not reject the null) or rather different (reject the null).

Question 2

Consider the error-components model

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + v_{it}, \qquad v_{it} = c_i + u_{it}$$

with $\mathrm{E}(c_i) = 0$, $\mathrm{E}(u_{it}) = 0$, and $\mathrm{E}(c_i u_{it}) = 0$.

(a) Show without using an expectation operator that for $T = 2$, the within-transformed disturbances $\ddot{u}_{it}$ have first-order autocorrelation $\mathrm{Corr}(\ddot{u}_{i2}, \ddot{u}_{i1}) = -1$.

<u>Answer:</u> Recall that $\ddot{u}_{it} = u_{it} - T^{-1} \sum_{t=1}^{T} u_{it}$. For $T = 2$ we thus obtain

$$\ddot{u}_{1t} = u_{1t} - \frac{1}{2}(u_{i1} + u_{i2}) = \frac{u_{i1} - u_{i2}}{2}$$

and

$$\ddot{u}_{2t} = u_{2t} - \frac{1}{2}(u_{i1} + u_{i2}) = \frac{u_{i2} - u_{i1}}{2} = -\frac{u_{i1} - u_{i2}}{2}.$$

Clearly, $\ddot{u}_{1t} = -\ddot{u}_{2t}$. Hence, they are perfectly negatively correlated with correlation coefficient -1.

(b) Show that in general the within-transformed disturbances $\ddot{u}_{it}$ have first-order autocovariance $\mathrm{E}(\ddot{u}_{it}\ddot{u}_{it-1}) = -\sigma_u^2/T$ and first-order autocorrelation $\mathrm{Corr}(\ddot{u}_{it}, \ddot{u}_{it-1}) = (1-T)^{-1}$ if u_{it} is white noise with variance σ_u^2.

<u>Answer:</u> Let us first find the autocovariance:

$$\begin{aligned}\mathrm{E}(\ddot{u}_{it}\ddot{u}_{it-1}) &= \mathrm{E}[(u_{it} - \bar{u}_i)(u_{it-1} - \bar{u}_i)] \\ &= \mathrm{E}[u_{it}u_{it-1}] - \mathrm{E}[u_{it}\bar{u}_i] - \mathrm{E}[\bar{u}_i u_{it-1}] + \mathrm{E}[\bar{u}_i^2].\end{aligned}$$

Lets go through these four terms one by one.

- $\mathrm{E}[u_{it}u_{it-1}] = 0$ because u_{it} is white noise.
- $\mathrm{E}[u_{it}\bar{u}_i] = \mathrm{E}[u_{it}T^{-1}\sum_{s=1}^{T} u_{is}] = T^{-1}\sum_{s=1}^{T}\mathrm{E}[u_{it}u_{is}] = T^{-1}\sigma_u^2$ because $\mathrm{E}[u_{it}u_{is}] = \sigma_u^2$ for $s = t$ and zero otherwise.
- $\mathrm{E}[u_{it-1}\bar{u}_i] = T^{-1}\sigma_u^2$ by the same reasoning.
- $\mathrm{E}[\bar{u}_i^2] = T^{-1}\sigma_u^2$ because it is the variance of the mean.

Substituting these results yields

$$\mathrm{E}(\ddot{u}_{it}\ddot{u}_{it-1}) = 0 - T^{-1}\sigma_u^2 - T^{-1}\sigma_u^2 + T^{-1}\sigma_u^2 = -T^{-1}\sigma_u^2.$$

To find the autocorrelation, we need the variances of $\ddot{u}_{it}$ and $\ddot{u}_{it-1}$. Under the assumption that u_{it} is white noise with constant variance, we found these variance in Lecture 1 as $\mathrm{Var}(\ddot{u}_{it}) = \mathrm{Var}(\ddot{u}_{it-1}) = \frac{T-1}{T}\sigma_u^2$. Thus the autocorrelation coefficient is

$$\mathrm{Corr}(\ddot{u}_{it}, \ddot{u}_{it-1}) = \frac{\mathrm{E}(\ddot{u}_{it}\ddot{u}_{it-1})}{\sqrt{\mathrm{Var}(\ddot{u}_{it})}\sqrt{\mathrm{Var}(\ddot{u}_{it-1})}} = \frac{-T^{-1}\sigma_u^2}{\frac{T-1}{T}\sigma_u^2} = \frac{1}{1-T}.$$

Question 3

Consider the simple model

$$y_{it} = \alpha + x_t\beta + v_{it},$$

where x_t is a scalar regressor that varies solely with t (e.g., a time dummy or an aggregate control variable). Show that the FE and RE estimators of β are numerically identical. Hint: apply POLS to the within-transformed equation (which yields the FE estimator) and to the quasi-demeaned equation (which yields the RE estimator). In the latter case, use an arbitrary value $\lambda = 1 - \phi$ for the quasi demeaning and show that the two estimators are identical for any λ.

Answer:

Let us start with the FE estimator. The within transformation yields

$$y_{it} - \bar{y}_i = (x_t - \bar{x})\beta + v_{it} - \bar{v}_i, \tag{22}$$

where $\bar{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}$, $\bar{x} = T^{-1}\sum_{t=1}^{T} x_t$, and $\bar{v}_i = T^{-1}\sum_{t=1}^{T} v_{it}$ are scalars. The FE estimator is pooled OLS applied to (22):

$$\hat{\beta}_{FE} = \left(\sum_{i=1}^{N}\sum_{t=1}^{T}(x_t - \bar{x})^2\right)^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}(x_t - \bar{x})(y_{it} - \bar{y}_i)\right) \tag{23}$$

Simplify the second part:

$$\begin{aligned}
\sum_{i=1}^{N}\sum_{t=1}^{T}(x_t - \bar{x})(y_{it} - \bar{y}_i) &= \sum_{t=1}^{T}(x_t - \bar{x})\sum_{i=1}^{N}(y_{it} - \bar{y}_i) \\
&= \sum_{t=1}^{T}(x_t - \bar{x})\sum_{i=1}^{N} y_{it} - \sum_{t=1}^{T}(x_t - \bar{x})\sum_{i=1}^{N}\bar{y}_i \\
&= \sum_{t=1}^{T}(x_t - \bar{x})\sum_{i=1}^{N} y_{it} - \left[\sum_{t=1}^{T}(x_t - \bar{x})\right]\left[\sum_{i=1}^{N}\bar{y}_i\right]
\end{aligned}$$

where the last step uses the fact that $\sum_{i=1}^{N}\bar{y}_i$ does not depend on t. Now recall from basic statistics that $\sum_{t=1}^{T}(x_t - \bar{x}) = T\bar{x} - T\bar{x} = 0$ and substitute back:

$$\sum_{i=1}^{N}\sum_{t=1}^{T}(x_t - \bar{x})(y_{it} - \bar{y}_i) = \sum_{t=1}^{T}(x_t - \bar{x})\sum_{i=1}^{N} y_{it} = \sum_{t=1}^{T}\sum_{i=1}^{N}(x_t - \bar{x})y_{it}.$$

Hence, the FE estimator (23) simplifies to

$$\hat{\beta}_{FE} = \left(\sum_{i=1}^{N} \sum_{t=1}^{T} (x_t - \bar{x})^2 \right)^{-1} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} (x_t - \bar{x}) y_{it} \right) \tag{24}$$

Now let us turn to the RE estimator. Quasi-demeaning with some arbitrary $\lambda = 1 - \phi$ yields

$$y_{it} - \lambda \bar{y}_i = \alpha(1 - \lambda) + (x_t - \lambda \bar{x})\beta + v_{it} - \lambda \bar{v}_i. \tag{25}$$

This can be transformed into an equation that has regressor $x_t - \bar{x}$ and thus the same one as the within-transformed equation:

$$\begin{aligned} y_{it} - \lambda \bar{y}_i &= \alpha(1 - \lambda) + (x_t - \bar{x} + \bar{x} - \lambda \bar{x})\beta + v_{it} - \lambda \bar{v}_i \\ &= \underbrace{\alpha(1 - \lambda) + \beta \bar{x}(1 - \lambda)}_{\mu} + (x_t - \bar{x})\beta + v_{it} - \lambda \bar{v}_i \end{aligned}$$

Now there are two ways to proceed. The first is to recall from basic econometrics that OLS applied to an equation with intercept is identical to OLS applied to the mean-adjusted data without intercept, where the mean is taken from the whole sample. Hence, the lhs becomes

$$y_{it} - \lambda \bar{y}_i - \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (y_{it} - \lambda \bar{y}_i) = y_{it} - \lambda \bar{y}_i - \bar{y} + \lambda \bar{y} = y_{it} - \lambda \bar{y}_i - (1 - \lambda)\bar{y}$$

where $\bar{y} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it} = \frac{1}{N} \sum_{i=1}^{N} \bar{y}_i$, and the rhs remains unchanged

$$x_t - \bar{x} - \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_t - \bar{x}) = x_t - \bar{x} - \bar{x} + \bar{x} = x_t - \bar{x}.$$

Hence, it is equivalent to estimate equation (25) by POLS or this one:

$$y_{it} - \lambda \bar{y}_i - (1 - \lambda)\bar{y} = \mu + (x_t - \bar{x})\beta + v_{it} - \lambda \bar{v}_i - (1 - \lambda)\bar{v}. \tag{26}$$

POLS applied to this equation is

$$\hat{\beta}_{RE} = \left(\sum_{i=1}^{N} \sum_{t=1}^{T} (x_t - \bar{x})^2 \right)^{-1} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} (x_t - \bar{x})(y_{it} - \lambda \bar{y}_i - (1 - \lambda)\bar{y}) \right) \tag{27}$$

The same simplification as used for the FE estimator above yields

$$\sum_{i=1}^{N}\sum_{t=1}^{T}(x_t-\bar{x})(y_{it}-\lambda\bar{y}_i-(1-\lambda)\bar{y})=\sum_{i=1}^{N}\sum_{t=1}^{T}(x_t-\bar{x})y_{it}.$$

Hence, the RE estimator becomes

$$\hat{\beta}_{RE}=\left(\sum_{i=1}^{N}\sum_{t=1}^{T}(x_t-\bar{x})^2\right)^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}(x_t-\bar{x})y_{it}\right). \tag{28}$$

Clearly, (24) and (28) are identical.

The second way to analyze the RE estimator is to apply POLS directly to

$$y_{it}-\lambda\bar{y}_i=\mu+(x_t-\bar{x})\beta+v_{it}-\lambda\bar{v}_i, \tag{29}$$

which implies we have a regression with intercept:

$$\begin{aligned}\begin{pmatrix}\hat{\mu}_{RE}\\ \hat{\beta}_{RE}\end{pmatrix}&=\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\begin{bmatrix}1\\ x_t-\bar{x}\end{bmatrix}\begin{bmatrix}1, & x_t-\bar{x}\end{bmatrix}\right)^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\begin{bmatrix}1\\ x_t-\bar{x}\end{bmatrix}(y_{it}-\lambda\bar{y}_i)\right)\\ &=\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\begin{bmatrix}1 & x_t-\bar{x}\\ x_t-\bar{x} & (x_t-\bar{x})^2\end{bmatrix}\right)^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\begin{bmatrix}y_{it}-\lambda\bar{y}_i\\ (x_t-\bar{x})(y_{it}-\lambda\bar{y}_i)\end{bmatrix}\right)\end{aligned}$$

Due to $\sum_{t=1}^{T}(x_t-\bar{x})=0$ simplify the off-diagonal elements of the first matrix to zero which makes inversion easy:

$$\begin{aligned}\begin{pmatrix}\hat{\mu}_{RE}\\ \hat{\beta}_{RE}\end{pmatrix}&=\begin{bmatrix}NT & 0\\ 0 & \sum_{i=1}^{N}\sum_{t=1}^{T}(x_t-\bar{x})^2\end{bmatrix}^{-1}\begin{bmatrix}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it}-\lambda\bar{y}_i)\\ \sum_{i=1}^{N}\sum_{t=1}^{T}(x_t-\bar{x})(y_{it}-\lambda\bar{y}_i)\end{bmatrix}\\ &=\begin{bmatrix}(NT)^{-1} & 0\\ 0 & \left(\sum_{i=1}^{N}\sum_{t=1}^{T}(x_t-\bar{x})^2\right)^{-1}\end{bmatrix}\begin{bmatrix}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it}-\lambda\bar{y}_i)\\ \sum_{i=1}^{N}\sum_{t=1}^{T}(x_t-\bar{x})(y_{it}-\lambda\bar{y}_i)\end{bmatrix}\\ &=\begin{bmatrix}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it}-\lambda\bar{y}_i)\\ \left(\sum_{i=1}^{N}\sum_{t=1}^{T}(x_t-\bar{x})^2\right)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(x_t-\bar{x})(y_{it}-\lambda\bar{y}_i)\end{bmatrix}\end{aligned}$$

For $\hat{\beta}_{RE}$ this yields

$$\hat{\beta}_{RE}=\left(\sum_{i=1}^{N}\sum_{t=1}^{T}(x_t-\bar{x})^2\right)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(x_t-\bar{x})(y_{it}-\lambda\bar{y}_i).$$

By the same reasoning as above, this is identical to

$$\hat{\beta}_{RE} = \left(\sum_{i=1}^{N} \sum_{t=1}^{T} (x_t - \bar{x})^2 \right)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_t - \bar{x}) y_{it}$$

and thus to the FE estimator.

Question 4

Consider the structural equation

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + v_{it}, \qquad v_{it} = c_i + u_{it}, \tag{30}$$

where $\mathbf{x}_{it}$ is a $1 \times K$ vector which includes only regressors that vary across i and t, and the augmented equation

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \bar{\mathbf{x}}_i\boldsymbol{\delta} + r_{it}, \tag{31}$$

where $\bar{\mathbf{x}}_i$ is a $1 \times K$ vector of time averages. For what follows assume that RE.3 holds and the variance components of the RE estimator are estimated using the Swamy-Arora approach.

(a) Denote the RE estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ in (31) by $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\delta}}$. Show that $\tilde{\boldsymbol{\beta}}$ is identical to the FE estimator of $\boldsymbol{\beta}$ in (30), and $\tilde{\boldsymbol{\delta}}$ is identical to the difference of the between and FE estimators of $\boldsymbol{\beta}$ in (30): $\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{RE}$ and $\tilde{\boldsymbol{\delta}} = \hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_{FE}$. Hint: inversion of the partitioned matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

yields

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{D} & -\mathbf{D}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{D} & \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{D}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \end{bmatrix},$$

where $\mathbf{D} = (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}$.

<u>Answer:</u> To avoid all the sums, write the equations in stacked matrix form as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{V} \tag{32}$$

and

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \bar{\mathbf{X}}\boldsymbol{\delta} + \mathbf{E} = [\mathbf{X}, \bar{\mathbf{X}}] \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\delta} \end{pmatrix} + \mathbf{E} = \mathbf{Z}\boldsymbol{\eta} + \mathbf{E}. \tag{33}$$

The FE estimator applied to (32) yields the usual

$$\hat{\boldsymbol{\beta}}_{FE} = (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\ddot{\mathbf{X}}'\ddot{\mathbf{y}}, \tag{34}$$

where $\ddot{\mathbf{X}} = \mathbf{QX}$ and $\ddot{\mathbf{y}} = \mathbf{Qy}$.

The between estimator applied to (32) yields the usual

$$\hat{\boldsymbol{\beta}}_{B} = (\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\bar{\mathbf{X}}'\bar{\mathbf{y}}, \tag{35}$$

where $\bar{\mathbf{X}} = \mathbf{JX}$ and $\bar{\mathbf{y}} = \mathbf{Jy}$.

The RE estimator applied to (33) equals the OLS estimator applied to the quasi-demeaned equation

$$\mathbf{Ry} = \mathbf{RX}\boldsymbol{\beta} + \mathbf{R}\bar{\mathbf{X}}\boldsymbol{\delta} + \mathbf{RE} = \mathbf{RX}\boldsymbol{\beta} + \hat{\phi}\mathbf{JX}\boldsymbol{\delta} + \mathbf{RE} = \underbrace{[\mathbf{RX}, \hat{\phi}\mathbf{JX}]}_{\mathbf{RZ}}\begin{pmatrix}\boldsymbol{\beta}\\ \boldsymbol{\delta}\end{pmatrix} + \mathbf{RE}$$

because $\mathbf{RJ} = (\mathbf{Q} + \hat{\phi}\mathbf{J})\mathbf{J} = \hat{\phi}\mathbf{J}$ and thus $\mathbf{R}\bar{\mathbf{X}} = \mathbf{RJX} = \hat{\phi}\mathbf{JX}$. Hence,

$$\begin{aligned}
\begin{pmatrix}\tilde{\boldsymbol{\beta}}\\ \tilde{\boldsymbol{\delta}}\end{pmatrix} &= (\mathbf{Z}'\mathbf{R}'\mathbf{RZ})^{-1}\mathbf{Z}'\mathbf{R}'\mathbf{Ry} = \left(\begin{bmatrix}\mathbf{X}'\mathbf{R}'\\ \hat{\phi}\mathbf{X}'\mathbf{J}'\end{bmatrix}[\mathbf{RX}, \quad \hat{\phi}\mathbf{JX}]\right)^{-1}\begin{bmatrix}\mathbf{X}'\mathbf{R}'\\ \hat{\phi}\mathbf{X}'\mathbf{J}'\end{bmatrix}\mathbf{Ry}\\
&= \begin{bmatrix}\mathbf{X}'\mathbf{R}'\mathbf{RX} & \hat{\phi}\mathbf{X}'\mathbf{R}'\mathbf{JX}\\ \hat{\phi}\mathbf{X}'\mathbf{J}'\mathbf{RX} & \hat{\phi}^2\mathbf{X}'\mathbf{J}'\mathbf{JX}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}'\mathbf{R}'\mathbf{Ry}\\ \hat{\phi}\mathbf{X}'\mathbf{J}'\mathbf{Ry}\end{bmatrix}\\
&= \begin{bmatrix}\mathbf{X}'\mathbf{R}'\mathbf{RX} & \hat{\phi}^2\mathbf{X}'\mathbf{J}'\mathbf{JX}\\ \hat{\phi}^2\mathbf{X}'\mathbf{J}'\mathbf{JX} & \hat{\phi}^2\mathbf{X}'\mathbf{J}'\mathbf{JX}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}'\mathbf{R}'\mathbf{Ry}\\ \hat{\phi}^2\mathbf{X}'\mathbf{J}'\mathbf{Jy}\end{bmatrix}\\
&= \underbrace{\begin{bmatrix}\tilde{\mathbf{X}}'\tilde{\mathbf{X}} & \hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}}\\ \hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}} & \hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}}\end{bmatrix}^{-1}}_{\mathbf{A}^{-1}}\begin{bmatrix}\tilde{\mathbf{X}}'\tilde{\mathbf{y}}\\ \hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{y}}\end{bmatrix},
\end{aligned}$$

where we used $\mathbf{RJ} = \hat{\phi}\mathbf{J} = \hat{\phi}\mathbf{J}'\mathbf{J}$. To invert $\mathbf{A}$, we use the rule given above. It turns out that

$$\begin{aligned}
\mathbf{D} &= (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} = (\tilde{\mathbf{X}}'\tilde{\mathbf{X}} - \hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}}(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\\
&= (\tilde{\mathbf{X}}'\tilde{\mathbf{X}} - \hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} = (\mathbf{X}'\mathbf{R}'\mathbf{RX} - \hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} = (\mathbf{X}'\hat{\sigma}_u^2\boldsymbol{\Omega}^{-1}\mathbf{X} - \hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\\
&= (\mathbf{X}'[\mathbf{Q} + \hat{\phi}^2\mathbf{J}]\mathbf{X} - \hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} = (\mathbf{XQX})^{-1} = (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}.
\end{aligned}$$

Hence,

$$\mathbf{A}^{-1} = \begin{bmatrix} (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} & -(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}}(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} \\ -(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}}(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} & (\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} + (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} \end{bmatrix}$$
$$= \begin{bmatrix} (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} & -(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} \\ -(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} & (\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} + (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} \end{bmatrix}.$$

Substituting this into the estimator yields

$$\begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\delta}} \end{pmatrix} = \begin{bmatrix} (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} & -(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} \\ -(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} & (\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} + (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{X}}'\tilde{\mathbf{y}} \\ \hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{y}} \end{bmatrix}$$
$$= \begin{bmatrix} (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{y}} - (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{y}} \\ -(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{y}} + (\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{y}} + (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{y}} \end{bmatrix}.$$

Note that $\tilde{\mathbf{X}}'\tilde{\mathbf{y}} = \mathbf{X}'[\mathbf{Q} + \hat{\phi}^2\mathbf{J}]\mathbf{y} = \ddot{\mathbf{X}}'\ddot{\mathbf{y}} + \hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{y}}$ and thus

$$\begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\delta}} \end{pmatrix} = \begin{bmatrix} (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\ddot{\mathbf{X}}'\ddot{\mathbf{y}} + (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{y}} - (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{y}} \\ -(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\ddot{\mathbf{X}}'\ddot{\mathbf{y}} - (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{y}} + (\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{y}} + (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{y}} \end{bmatrix}$$
$$= \begin{bmatrix} (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\ddot{\mathbf{X}}'\ddot{\mathbf{y}} \\ -(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\ddot{\mathbf{X}}'\ddot{\mathbf{y}} + (\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\bar{\mathbf{X}}'\bar{\mathbf{y}} \end{bmatrix} = \begin{pmatrix} \hat{\boldsymbol{\beta}}_{FE} \\ \hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_{FE} \end{pmatrix}.$$

(b) Find the Wald statistic to test the null hypothesis $\boldsymbol{\delta} = \mathbf{0}$ based on the classical RE estimator of the augmented equation (31).

Answer: Let us start to find the population variance matrix of $\tilde{\boldsymbol{\delta}}$. Recall that under homoscedasticity, the variance matrix of the full parameter vector is

$$\text{Var}\left[\begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\delta}} \end{pmatrix}\right] = (\mathbf{Z}'[\mathbf{I}_N \otimes \boldsymbol{\Omega}^{-1}]\mathbf{Z})^{-1} = \sigma_u^2(\mathbf{Z}'\mathbf{R}'\mathbf{R}\mathbf{Z})^{-1}.$$

The inverse has been found above:

$$\text{Var}\left[\begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\delta}} \end{pmatrix}\right] = \sigma_u^2 \begin{bmatrix} (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} & -(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} \\ -(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} & (\phi^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} + (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} \end{bmatrix}.$$

Hence, the variance matrix of $\tilde{\boldsymbol{\delta}}$ is the lower right part:

$$\text{Var}(\tilde{\boldsymbol{\delta}}) = \sigma_u^2(\phi^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} + \sigma_u^2(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}.$$

The Wald statistic thus is

$$W = \tilde{\boldsymbol{\delta}}' \left[\widehat{\text{Var}}(\tilde{\boldsymbol{\delta}})\right]^{-1} \tilde{\boldsymbol{\delta}} = (\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_{FE})' \left[\hat{\sigma}_u^2(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} + \hat{\sigma}_u^2(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\right]^{-1} (\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_{FE}), \quad (36)$$

where we replace unknown quantities by consistent estimators.

(c) Show that the classical Hausman statistic applied to equation (30) is identical to the Wald statistic computed above.

Answer: The classical Hausman statistic is

$$\begin{aligned} H &= (\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE})'[\widehat{\text{Avar}}(\hat{\boldsymbol{\beta}}_{FE}) - \widehat{\text{Avar}}(\hat{\boldsymbol{\beta}}_{RE})]^{-1}(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE}) \\ &= (\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE})'[\hat{\sigma}_u^2(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} - (\mathbf{X}'[\mathbf{I}_N \otimes \boldsymbol{\Omega}^{-1}]\mathbf{X})^{-1}]^{-1}(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE}) \\ &= (\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE})'[\hat{\sigma}_u^2(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} - (\mathbf{X}'\mathbf{R}'\mathbf{R}\mathbf{X}/\hat{\sigma}_u^2)^{-1}]^{-1}(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE}) \\ &= (\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE})'[\hat{\sigma}_u^2(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} - \hat{\sigma}_u^2(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}]^{-1}(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE}) \end{aligned}$$

In order to transform it into the Wald statistic (36), use the equality

$$\hat{\boldsymbol{\beta}}_{RE} = \mathbf{W}_1\hat{\boldsymbol{\beta}}_B + (\mathbf{I} - \mathbf{W}_1)\hat{\boldsymbol{\beta}}_{FE}$$

where

$$\begin{aligned} \mathbf{W}_1 &= \left(\hat{\phi}^2 \sum_{i=1}^N \mathbf{X}_i'\mathbf{J}_T\mathbf{X}_i + \sum_{i=1}^N \mathbf{X}_i'\mathbf{Q}_T\mathbf{X}_i\right)^{-1} \hat{\phi}^2 \sum_{i=1}^N \mathbf{X}_i'\mathbf{J}_T\mathbf{X}_i \\ &= \left(\hat{\phi}^2\mathbf{X}'\mathbf{J}\mathbf{X} + \mathbf{X}'\mathbf{Q}\mathbf{X}\right)^{-1} \hat{\phi}^2\mathbf{X}'\mathbf{J}\mathbf{X} \\ &= \left(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\right)^{-1} \hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}}. \end{aligned}$$

Substituting the equality yields

$$\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE} = \mathbf{W}_1\hat{\boldsymbol{\beta}}_{FE} - \mathbf{W}_1\hat{\boldsymbol{\beta}}_B = -\mathbf{W}_1(\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_{FE}).$$

Now substitute this into the Hausman statistic:

$$\begin{aligned} H &= (\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_{FE})'\mathbf{W}_1'[\hat{\sigma}_u^2(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} - \hat{\sigma}_u^2(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}]^{-1}\mathbf{W}_1(\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_{FE}) \\ &= (\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_{FE})' \left[\underbrace{(\mathbf{W}_1)^{-1}\left(\hat{\sigma}_u^2(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} - \hat{\sigma}_u^2(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}\right)(\mathbf{W}_1')^{-1}}_{\mathbf{V}}\right]^{-1} (\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_{FE}). \end{aligned}$$

The term in squared brackets simplifies to

$$\begin{aligned}
\mathbf{V} &= (\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\left(\hat{\sigma}_u^2(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} - \hat{\sigma}_u^2(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}\right)\tilde{\mathbf{X}}'\tilde{\mathbf{X}}(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} \\
&= (\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\hat{\sigma}_u^2(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} - (\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\hat{\sigma}_u^2(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} \\
&= \hat{\sigma}_u^2(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}(\ddot{\mathbf{X}}'\ddot{\mathbf{X}} + \hat{\phi}^2\bar{\mathbf{X}}\bar{\mathbf{X}})(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} - \hat{\sigma}_u^2(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} \\
&= \hat{\sigma}_u^2(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} + \hat{\sigma}_u^2(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\hat{\phi}^2\bar{\mathbf{X}}\bar{\mathbf{X}}(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} \\
&\quad - \hat{\sigma}_u^2(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} \\
&= \hat{\sigma}_u^2(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1}(\ddot{\mathbf{X}}'\ddot{\mathbf{X}} + \hat{\phi}^2\bar{\mathbf{X}}\bar{\mathbf{X}})(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} \\
&= \hat{\sigma}_u^2(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} + \hat{\sigma}_u^2(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}.
\end{aligned}$$

Substituting back into the Hausman statistic yields

$$H = (\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_{FE})'\left[\hat{\sigma}_u^2(\hat{\phi}^2\bar{\mathbf{X}}'\bar{\mathbf{X}})^{-1} + \hat{\sigma}_u^2(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\right]^{-1}(\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_{FE}).$$

This is exactly the Wald statistic (36).

It remains to verify that the estimates of the error variances are identical, no matter whether the RE estimator is applied to (30) which yields the Hausman statistic or to (31) which yields the Wald statistic. To this end, recall that the Swamy-Arora approach uses the within and between transformed equations. Applying the within transformation to (30) and (31) yields

$$\ddot{y}_{it} = \ddot{\mathbf{x}}_{it}\boldsymbol{\beta} + \ddot{v}_{it}$$

and

$$\ddot{y}_{it} = \ddot{\mathbf{x}}_{it}\boldsymbol{\beta} + \ddot{\varepsilon}_{it}$$

because the within transformation wipes out the time invariant regressors $\bar{\mathbf{x}}_i$. Hence, the two are identical in terms of their lhs variable and regressors and so is the variance estimator $\hat{\sigma}_u^2$ based on this equation. Applying the between transformation to (30) and (31) yields

$$\bar{y}_i = \bar{\mathbf{x}}_i\boldsymbol{\beta} + \bar{v}_i$$

and

$$\bar{y}_i = \bar{\mathbf{x}}_i\boldsymbol{\beta} + \bar{\mathbf{x}}_i\boldsymbol{\delta} + \bar{\varepsilon}_i = \bar{\mathbf{x}}_i(\boldsymbol{\beta} + \boldsymbol{\delta}) + \bar{\varepsilon}_i.$$

Again, these equations are identical in terms of their lhs variable and regressors. Hence, they produce the same variance estimator $\hat{\sigma}_c^2 + \hat{\sigma}_u^2/T$. Consequently, both approaches also lead to the same estimator $\hat{\phi}$.