Set-up

Consider the model for $i = 1, \ldots, N, t = 1, \ldots, T$,

$$y_{it} = \beta x_{it} + \underbrace{\alpha_i + \epsilon_{it}}_{\equiv v_{it}}$$

$$\begin{aligned} \epsilon_{it} \text{ i.i.d. over } i, t, &\quad var(\epsilon_{it}) = \sigma_\epsilon^2 \\ \alpha_i \text{ i.i.d. over } i, &\quad var(\alpha_i) = \sigma_\alpha^2 \\ cov(\alpha_i, \epsilon_{it}) = 0 & \end{aligned}$$

For simplicity, assume everything is mean zero, and $\beta \in \mathbb{R}$, then

$$E[y_{it}] = \beta E[x_{it}] + E[\alpha_i] + E[\epsilon_{it}] = 0.$$

Let $\Omega = var([v_{i1}, v_{i2}, \ldots, v_{iT}]') = var(\mathbf{v}_i) = E(\mathbf{v_i}\mathbf{v_i}')$. Then we get

$$\Omega = \begin{bmatrix} \sigma_\alpha^2 + \sigma_\epsilon^2 & \sigma_\alpha^2 & \ldots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 + \sigma_\epsilon^2 & \ldots & \sigma_\alpha^2 \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \ldots & \sigma_\alpha^2 + \sigma_\epsilon^2 \end{bmatrix} = \sigma_\epsilon^2 I_T + \sigma_\alpha^2 \mathbf{ee}'$$

where $\mathbf{e} = [1, 1, \ldots, 1]'$. Also, we get

$$\Omega^{-1} = \frac{1}{\sigma_\epsilon^2}\left(I_T - \frac{\sigma_\alpha^2}{(T\sigma_\alpha^2 + \sigma_\epsilon^2)}\mathbf{ee}'\right)$$

Within-Estimator and GLS Estimator with Local Misspecification

With fixed effect model, within-estimator is efficient. On the other hand, with random effect model where $cov(x_{it}, \alpha_i) = 0$ for all i, t, GLS estimator is efficient. Define

$$
\mathbf{y}_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{bmatrix}, \quad \mathbf{x}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iT} \end{bmatrix}, \quad \mathbf{v}_i = \begin{bmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{iT} \end{bmatrix}, \quad \epsilon_i = \begin{bmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \vdots \\ \epsilon_{iT} \end{bmatrix}, \quad \mathbf{a}_i = \begin{bmatrix} \alpha_i \\ \alpha_i \\ \vdots \\ \alpha_i \end{bmatrix}
$$

Also, let's consider local misspecification in the form of

$$
\sum_{t=1}^{T} cov(x_{it}, \alpha_i) = \sum_{t=1}^{T} E[x_{it}\alpha_i] = \frac{\delta}{\sqrt{N}}, \quad \delta \neq 0
$$

(a) Within-Estimator

Define $Q = I_T - \frac{1}{T}\mathbf{e}\mathbf{e}'$, which is demeaning matrix and idempotent. The within-estimator $\widehat{\beta}_{FE}$ is

$$
\begin{aligned}
\widehat{\beta}_{FE} &= \left(\sum_{i=1}^{N} \mathbf{x}_i' Q \mathbf{x}_i\right)^{-1} \left(\sum_{i=1}^{N} \mathbf{x}_i' Q \mathbf{y}_i\right) \\
&= \left(\sum_{i=1}^{N} \mathbf{x}_i' Q \mathbf{x}_i\right)^{-1} \left(\sum_{i=1}^{N} \mathbf{x}_i' Q (\mathbf{x}_i \beta + \mathbf{v}_i)\right) \\
&= \beta + \left(\sum_{i=1}^{N} \mathbf{x}_i' Q \mathbf{x}_i\right)^{-1} \left(\sum_{i=1}^{N} \mathbf{x}_i' Q \mathbf{v}_i\right) \\
&= \beta + \left(\sum_{i=1}^{N} \mathbf{x}_i' Q \mathbf{x}_i\right)^{-1} \left(\sum_{i=1}^{N} \mathbf{x}_i' Q \epsilon_i\right)
\end{aligned}
$$

Assume $E[\epsilon_i \epsilon_i' \mid \mathbf{x}_i, \alpha_i] = \sigma_\epsilon^2 I$ and $E[\mathbf{x}_i' Q \epsilon_i] = 0$. Then the asymptotic distribution of $\widehat{\beta}_{FE}$ is derived as

$$
\sqrt{N}(\widehat{\beta}_{FE} - \beta) \to_d N\big(0,\ \sigma_\epsilon^2 E[\mathbf{x}_i' Q \mathbf{x}_i]^{-1}\big)
$$

(b) GLS Estimator

Using Ω matrix defined before, the GLS estimator $\widehat{\beta}_{GLS}$ is

$$
\begin{aligned}
\widehat{\beta}_{GLS} &= \left(\sum_{i=1}^{N} \mathbf{x}_i' \Omega^{-1} \mathbf{x}_i\right)^{-1} \left(\sum_{i=1}^{N} \mathbf{x}_i' \Omega^{-1} \mathbf{y}_i\right) \\
&= \left(\sum_{i=1}^{N} \mathbf{x}_i' \Omega^{-1} \mathbf{x}_i\right)^{-1} \left(\sum_{i=1}^{N} \mathbf{x}_i' \Omega^{-1} (\mathbf{x}_i \beta + \mathbf{v}_i)\right) \\
&= \beta + \left(\sum_{i=1}^{N} \mathbf{x}_i' \Omega^{-1} \mathbf{x}_i\right)^{-1} \left(\sum_{i=1}^{N} \mathbf{x}_i' \Omega^{-1} \mathbf{v}_i\right)
\end{aligned}
$$

Compute $E[\mathbf{x_i}'\Omega^{-1}\mathbf{v}_i]$. Assume $E[\mathbf{x_i}'\Omega^{-1}\epsilon_i] = 0$. Then,

$$
\begin{aligned}
E[\mathbf{x_i}'\Omega^{-1}\mathbf{v}_i] &= E[\mathbf{x_i}'\Omega^{-1}(\mathbf{a}_i + \epsilon_i)] \\
&= E[\mathbf{x_i}'\Omega^{-1}\mathbf{a}_i] \\
&= E\left[\mathbf{x}_i'\left(\frac{1}{\sigma_\epsilon^2}\big(I_T - \frac{\sigma_\alpha^2}{(T\sigma_\alpha^2 + \sigma_\epsilon^2)}\mathbf{e}\mathbf{e}'\big)\right)\mathbf{a}_i\right] \\
&= \frac{1}{\sigma_\epsilon^2} E[\mathbf{x}_i'\mathbf{a}_i] - \frac{\sigma_\alpha^2}{\sigma_\epsilon^2(T\sigma_\alpha^2 + \sigma_\epsilon^2)} E[\mathbf{x}_i'\mathbf{e}\mathbf{e}'\mathbf{a}_i] \\
&= \frac{1}{\sigma_\epsilon^2} E[\sum_{t=1}^{T} x_{it}\alpha_i] - \frac{\sigma_\alpha^2}{\sigma_\epsilon^2(T\sigma_\alpha^2 + \sigma_\epsilon^2)} E[T\sum_{t=1}^{T} x_{it}\alpha_i] \\
&= \frac{1}{\sigma_\epsilon^2}\left(1 - \frac{T\sigma_\alpha^2}{T\sigma_\alpha^2 + \sigma_\epsilon^2}\right) E[\sum_{t=1}^{T} x_{it}\alpha_i] \\
&= \frac{1}{T\sigma_\alpha^2 + \sigma_\epsilon^2} E[\sum_{t=1}^{T} x_{it}\alpha_i]
\end{aligned}
$$

Note that the last term is relevant to local misspecification and this results in asymptotic bias. Asymptotic variance of $\widehat{\beta}_{GLS}$ is same as standard case, i.e., $Var(\widehat{\beta}_{GLS}) = E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i]^{-1}$. To conclude, the asymptotic distribution of $\widehat{\beta}_{GLS}$ under local misspecification is

$$
\sqrt{N}(\widehat{\beta}_{GLS} - \beta) \to_d N\big(\frac{\delta}{T\sigma_\alpha^2 + \sigma_\epsilon^2} E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i]^{-1},\ E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i]^{-1}\big)
$$

(c) Joint Distribution of $\widehat{\beta}_{FE}$ and $\widehat{\beta}_{GLS}$

$$\sqrt{N}\begin{bmatrix}\widehat{\beta}_{FE}-\beta\\ \widehat{\beta}_{GLS}-\beta\end{bmatrix}\to_d N\Bigg(\begin{bmatrix}0\\ \frac{\delta}{T\sigma_\alpha^2+\sigma_\epsilon^2}E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i]^{-1}\end{bmatrix},\quad \mathbf{V}\Bigg)$$

where

$$\mathbf{V}=\begin{bmatrix}Var(\widehat{\beta}_{FE}) & Cov(\widehat{\beta}_{FE},\widehat{\beta}_{GLS})\\ Cov(\widehat{\beta}_{FE},\widehat{\beta}_{GLS}) & Var(\widehat{\beta}_{GLS})\end{bmatrix}\equiv\begin{bmatrix}\sigma_{FE}^2 & \sigma_{FG}\\ \sigma_{FG} & \sigma_{GLS}^2\end{bmatrix}$$

$$\sigma_{FE}^2=\sigma_\epsilon^2E[\mathbf{x}_i'Q\mathbf{x}_i]^{-1},\qquad \sigma_{GLS}^2=E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i]^{-1}$$

$$\sigma_{FG}=\sigma_\epsilon^2E[\mathbf{x}_i'Q\mathbf{x}_i]^{-1}E[\mathbf{x}_i'Q\Omega^{-1}\mathbf{x}_i]E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i]^{-1}$$

Comparison to Hausman Test

$$\begin{aligned}\sqrt{N}(\widehat{\beta}_{GLS}-\widehat{\beta}_{FE}) &= \sqrt{N}(1\quad -1)\begin{bmatrix}\widehat{\beta}_{GLS}-\beta\\ \widehat{\beta}_{FE}-\beta\end{bmatrix}\\ &\to_d N\Bigg(\frac{\delta}{T\sigma_\alpha^2+\sigma_\epsilon^2}E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i]^{-1},\quad \underbrace{(1\quad -1)\begin{bmatrix}\sigma_{GLS}^2 & \sigma_{FG}\\ \sigma_{FG} & \sigma_{FE}^2\end{bmatrix}\begin{bmatrix}1\\ -1\end{bmatrix}}_{\equiv\Sigma}\Bigg)\end{aligned}$$

We can see that

$$\Sigma=Var(\widehat{\beta}_{FE}-\widehat{\beta}_{GLS})=Var(\widehat{\beta}_{FE})-Var(\widehat{\beta}_{GLS})$$

Last equality holds since GLS is efficient under the null of $cov(x_{it},\alpha_i)=0$ (Hausman). Therefore, under the null (i.e., $\delta=0$)

$$N(\widehat{\beta}_{GLS}-\widehat{\beta}_{FE})\Sigma^{-1}(\widehat{\beta}_{GLS}-\widehat{\beta}_{FE})\to\chi^2(1)$$

AMSE-Optimal Weight

We can compute the weight combining $\widehat{\beta}_{FE}$ and $\widehat{\beta}_{GLS}$ to minimize AMSE, i.e.,

$$\omega^* = argmin_{\omega \in [0,1]} \ AMSE(\omega\widehat{\beta}_{GLS} + (1-\omega)\widehat{\beta}_{FE}).$$

Let $\widehat{\beta}(\omega) = \omega\widehat{\beta}_{GLS} + (1-\omega)\widehat{\beta}_{FE}$. Then,

$$\begin{aligned} Bias(\widehat{\beta}(\omega)) &= \omega Bias(\widehat{\beta}_{GLS}) + (1-\omega) Bias(\widehat{\beta}_{FE}) \\ &= \omega \frac{\delta}{T\sigma_\alpha^2 + \sigma_\epsilon^2} E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i]^{-1} = \omega \frac{\delta}{T\sigma_\alpha^2 + \sigma_\epsilon^2}\sigma_{GLS}^2 \end{aligned}$$

$$\begin{aligned} Var(\widehat{\beta}(\omega)) &= (\omega \quad 1-\omega) \begin{bmatrix} \sigma_{GLS}^2 & \sigma_{FG} \\ \sigma_{FG} & \sigma_{FE}^2 \end{bmatrix} \begin{bmatrix} \omega \\ 1-\omega \end{bmatrix} \\ &= (\omega\sigma_{GLS}^2 + (1-\omega)\sigma_{FG} \quad \omega\sigma_{FG} + (1-\omega)\sigma_{FG}^2) \begin{bmatrix} \omega \\ 1-\omega \end{bmatrix} \\ &= \omega^2\sigma_{GLS}^2 + 2\omega(1-\omega)\sigma_{FG} + (1-\omega)^2\sigma_{FE}^2 \end{aligned}$$

Then, we can get

$$AMSE(\widehat{\beta}(\omega)) = \omega^2 \frac{\delta^2}{(T\sigma_\alpha^2 + \sigma_\epsilon^2)^2}(\sigma_{GLS}^2)^2 + \omega^2\sigma_{GLS}^2 + 2\omega(1-\omega)\sigma_{FG} + (1-\omega)^2\sigma_{FE}^2$$

Taking F.O.C. to find minimizer of $AMSE$,

$$\begin{aligned} 2\omega\left(\frac{\delta^2}{(T\sigma_\alpha^2 + \sigma_\epsilon^2)^2}(\sigma_{GLS}^2)^2 + \sigma_{GLS}^2 - 2\sigma_{FG} + \sigma_{FE}^2\right) + 2\sigma_{FG} - 2\sigma_{FE}^2 = 0 \\ \therefore \quad \omega^* = \frac{\sigma_{FE}^2 - \sigma_{FG}}{\frac{\delta^2}{(T\sigma_\alpha^2+\sigma_\epsilon^2)^2}(\sigma_{GLS}^2)^2 + \sigma_{GLS}^2 - 2\sigma_{FG} + \sigma_{FE}^2} \end{aligned}$$

Note that all the terms in ω^* except for δ can be consistently estimated. What we can do is to think of asymptotically unbiased estimator of δ and δ^2 accordingly.

Asymptotically Unbiased Estimator for δ^2

Consider the moment condition related to the GLS estimator. Plug in $\widehat{\beta}_{FE}$ to the moment condition. Denote the consistent estimators for $\sigma_\alpha^2, \sigma_\epsilon^2$ as $\widehat{\sigma}_\alpha^2, \widehat{\sigma}_\epsilon^2$ respectively. Define

$$\widehat{K}_N = \frac{1}{N}\sum_{i=1}^N \mathbf{x}_i' Q \mathbf{x}_i, \qquad \widetilde{K}_N = \left(\frac{1}{N}\sum_{i=1}^N \mathbf{x}_i'\Omega^{-1}\mathbf{x}_i\right)^{-1}$$

Consider

$$\begin{aligned}
\widehat{\delta} &= \frac{1}{\sqrt{N}}\sum_{i=1}^N \mathbf{x}_i'\Omega^{-1}(\mathbf{y}_i - \mathbf{x}_i\widehat{\beta}_{FE}) \cdot (T\widehat{\sigma}_\alpha^2 + \widehat{\sigma}_\epsilon^2) \\
&= (T\widehat{\sigma}_\alpha^2 + \widehat{\sigma}_\epsilon^2) \cdot \left(\frac{1}{\sqrt{N}}\sum_{i=1}^N \mathbf{x}_i'\Omega^{-1}\mathbf{v}_i - \underbrace{\frac{1}{N}\sum_{i=1}^N \mathbf{x}_i'\Omega^{-1}\mathbf{x}_i}_{=\widetilde{K}_N^{-1}} \cdot \underbrace{\sqrt{N}(\widehat{\beta}_{FE} - \beta)}_{=\widehat{K}_N^{-1}\frac{1}{\sqrt{N}}\sum \mathbf{x}_i' Q \mathbf{v}_i}\right) \\
&= \begin{bmatrix} 1 & \dfrac{-(T\widehat{\sigma}_\alpha^2 + \widehat{\sigma}_\epsilon^2)}{\widehat{K}_N\widetilde{K}_N} \end{bmatrix} \begin{bmatrix} \frac{(T\widehat{\sigma}_\alpha^2 + \widehat{\sigma}_\epsilon^2)}{\sqrt{N}}\sum_{i=1}^N \mathbf{x}_i'\Omega^{-1}\mathbf{v}_i \\ \frac{1}{\sqrt{N}}\sum_{i=1}^N \mathbf{x}_i' Q \mathbf{v}_i \end{bmatrix} \qquad \ldots (***)
\end{aligned}$$

We can get

$$\begin{bmatrix} \frac{(T\widehat{\sigma}_\alpha^2 + \widehat{\sigma}_\epsilon^2)}{\sqrt{N}}\sum_{i=1}^N \mathbf{x}_i'\Omega^{-1}\mathbf{v}_i \\ \frac{1}{\sqrt{N}}\sum_{i=1}^N \mathbf{x}_i' Q \mathbf{v}_i \end{bmatrix} \to_d N\left(\begin{bmatrix}\delta \\ 0\end{bmatrix}, \quad \widetilde{V}\right)$$

where

$$\widetilde{V} = \begin{bmatrix} (T\sigma_\alpha^2 + \sigma_\epsilon^2)^2 E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i] & (T\sigma_\alpha^2 + \sigma_\epsilon^2)E[\mathbf{x}_i' Q\mathbf{x}_i] \\ (T\sigma_\alpha^2 + \sigma_\epsilon^2)E[\mathbf{x}_i' Q\mathbf{x}_i] & \sigma_\epsilon^2 E[\mathbf{x}_i' Q \mathbf{x}_i] \end{bmatrix}$$

The covariance term is from

$$E[\mathbf{x}_i'\Omega^{-1}\mathbf{v}_i\mathbf{v}_i' Q\mathbf{x}_i] = E[\mathbf{x}_i'\Omega^{-1}\Omega Q\mathbf{x}_i] = E[\mathbf{x}_i' Q\mathbf{x}_i].$$

Going back to $(***)$, we get

$$\widehat{\delta} \;\to_d\; N\left(\delta,\quad \underbrace{\left[1 \qquad \frac{-(T\sigma_\alpha^2+\sigma_\epsilon^2)}{\widehat{K}\widetilde{K}}\right]\widetilde{V}\left[1 \qquad \frac{-(T\sigma_\alpha^2+\sigma_\epsilon^2)}{\widehat{K}\widetilde{K}}\right]'}_{\equiv\widehat{\sigma}_\delta^2}\right)$$

where $\widehat{K} = E[\mathbf{x}_i' Q\mathbf{x}_i]$ and $\widetilde{K} = E[\mathbf{x_i}'\Omega^{-1}\mathbf{x}_i]^{-1}$. Lastly, the asymptotically unbiased estimator of δ^2 can be obtained as $\widehat{\delta}^2 - \widehat{\sigma}_\delta^2$.

Comparison between Asymptotic Variances of $\widehat{\beta}_{FE}$ and $\widehat{\beta}_{GLS}$

First, think about the way to simplify the asymptotic variance of GLS estimator (with RE model).

$$\begin{aligned}
\Omega &= \sigma_\epsilon^2 I_T + \sigma_\alpha^2 \mathbf{e}\mathbf{e}' \\
&= \sigma_\epsilon^2 I_T + T\sigma_\alpha^2 \underbrace{\mathbf{e}(\mathbf{e}'\mathbf{e})^{-1}\mathbf{e}'}_{\equiv P_T} \\
&= \sigma_\epsilon^2 I_T + T\sigma_\alpha^2 P_T \\
&= (T\sigma_\alpha^2+\sigma_\epsilon^2)\left(P_T + \underbrace{\frac{\sigma_\epsilon^2}{T\sigma_\alpha^2+\sigma_\epsilon^2}}_{\equiv\eta}\underbrace{(I_T - \mathbf{e}(\mathbf{e}'\mathbf{e})^{-1}\mathbf{e}')}_{=I_T-P_T\equiv Q_T}\right) \\
&= (T\sigma_\alpha^2+\sigma_\epsilon^2)(\underbrace{P_T + \eta Q_T}_{\equiv S_T})
\end{aligned}$$

We can see that

$$S_T^{-1} = P_T + \frac{1}{\eta}Q_T$$

$$S_T^{-\frac{1}{2}} = P_T + \frac{1}{\sqrt{\eta}}Q_T \qquad (\because P_T Q_T = 0)$$

Let's denote $S_T^{-\frac{1}{2}}$ as follows:

$$S_T^{-\frac{1}{2}} = (1-\lambda)^{-1}(I_T - \lambda P_T), \qquad \text{where } \lambda \equiv 1-\sqrt{\eta} = 1-\sqrt{\frac{\sigma_\epsilon^2}{T\sigma_\alpha^2+\sigma_\epsilon^2}}$$

This implies that

$$\Omega^{-\frac{1}{2}} = (T\sigma_\alpha^2 + \sigma_\epsilon^2)^{-\frac{1}{2}}(1-\lambda)^{-1}(I_T - \lambda P_T) = \frac{1}{\sigma_\epsilon}\underbrace{(I_T - \lambda P_T)}_{\equiv C_T}$$

Therefore, we can see the GLS estimator (under RE model) is equivalent to OLS estimation with $\widetilde{\mathbf{y}}_i \equiv C_T \mathbf{y}_i$ and $\widetilde{\mathbf{x}}_i \equiv C_T \mathbf{x}_i$, i.e.,

$$C_T \mathbf{y}_i = C_T \mathbf{x}_i \beta + C_T \mathbf{v}_i \Leftrightarrow \widetilde{\mathbf{y}}_i = \widetilde{\mathbf{x}}_i \beta + \widetilde{\mathbf{v}}_i$$

Note that $E[\widetilde{\mathbf{v}}_i \widetilde{\mathbf{v}}_i'] = E[C_T \mathbf{v}_i \mathbf{v}_i' C_T'] = \sigma_\epsilon^2 I_T$. Therefore, OLS estimator from the above model is BLUE from Gauss-Markov Theorem. Also, we can see that $\widetilde{\mathbf{y}}_i$ is $\mathbf{y}_i - \lambda \bar{\mathbf{y}}_i$. (Wooldridge calls this quasi time-demeaning.) Fixed effect estimator is the same as with $\lambda = 1$.

With finite T and $\sigma_\alpha^2 > 0$, we have $\lambda < 1$. By comparing the asymptotic variances of fixed effect estimator and random effect estimator,

$$Var(\widehat{\beta}_{GLS}) = \sigma_\epsilon^2 E[\mathbf{x}_i' C_T' C_T \mathbf{x}_i]^{-1} = \sigma_\epsilon^2 E[(\mathbf{x}_i - \lambda P_T \mathbf{x}_i)'(\mathbf{x}_i - \lambda P_T \mathbf{x}_i)]^{-1}$$

$$Var(\widehat{\beta}_{FE}) = \sigma_\epsilon^2 E[\mathbf{x}_i' Q' Q \mathbf{x}_i]^{-1} = \sigma_\epsilon^2 E[(\mathbf{x}_i - P_T \mathbf{x}_i)'(\mathbf{x}_i - P_T \mathbf{x}_i)]^{-1}$$

To show that $A - B \geq 0$ ($p.s.d.$), it is equivalent to show $B^{-1} - A^{-1} \geq 0$. Now, consider

$$\begin{aligned} Var(\widehat{\beta}_{GLS})^{-1} - Var(\widehat{\beta}_{FE})^{-1} &= \sigma_\epsilon^2 \big(E(\mathbf{x}_i - \lambda P_T \mathbf{x}_i)'(\mathbf{x}_i - \lambda P_T \mathbf{x}_i) - E(\mathbf{x}_i - P_T \mathbf{x}_i)'(\mathbf{x}_i - P_T \mathbf{x}_i)\big) \\ &= \sigma_\epsilon^2 E\big(2(1-\lambda)\mathbf{x}_i' P_T \mathbf{x}_i - (1-\lambda^2)\mathbf{x}_i' P_T \mathbf{x}_i\big) \\ &= \sigma_\epsilon^2 (1-\lambda)^2 E[\mathbf{x}_i' P_T \mathbf{x}_i] \geq 0 \quad (\because P_T^2 = P_T,\ P_T' = P_T) \end{aligned}$$

Therefore, we have $Var(\widehat{\beta}_{FE}) - Var(\widehat{\beta}_{RE}) \geq 0$.

Lastly, covariance of two estimator can be written as

$$
\begin{aligned}
Cov(\widehat{\beta}_{GLS}, \widehat{\beta}_{FE}) &= E[\mathbf{x}_i' Q \mathbf{x}_i]^{-1} E[\mathbf{x}_i' Q \mathbf{v}_i \mathbf{v}_i' C_T' C_T \mathbf{x}_i] E[\mathbf{x}_i' C_T' C_T \mathbf{x}_i]^{-1} \\
&= E[\mathbf{x}_i' Q \mathbf{x}_i]^{-1} E[\mathbf{x}_i' Q (\mathbf{a}_i + \epsilon_i) \mathbf{v}_i' C_T' C_T \mathbf{x}_i] E[\mathbf{x}_i' C_T' C_T \mathbf{x}_i]^{-1} \\
&= E[\mathbf{x}_i' Q \mathbf{x}_i]^{-1} E[\mathbf{x}_i' Q \epsilon_i (\mathbf{a}_i + \epsilon_i)' C_T' C_T \mathbf{x}_i] E[\mathbf{x}_i' C_T' C_T \mathbf{x}_i]^{-1} \\
&= \sigma_\epsilon^2 E[\mathbf{x}_i' Q \mathbf{x}_i]^{-1} E[\mathbf{x}_i' Q C_T' C_T \mathbf{x}_i] E[\mathbf{x}_i' C_T' C_T \mathbf{x}_i]^{-1}
\end{aligned}
$$

Joint Distribution of $\widehat{\beta}_{RE}, \widehat{\beta}_{FE}$ and $\widehat{\delta}$

Using the results derived upto this point, we can get the joint distribution of $\widehat{\beta}_{RE}, \widehat{\beta}_{FE}$ and $\widehat{\delta}$. From the section of $\widehat{\delta}$, we know

$$
\widehat{\delta} \to_d \begin{bmatrix} 1 & \dfrac{-(T\sigma_\alpha^2 + \sigma_\epsilon^2)}{\widehat{K}\widetilde{K}} \end{bmatrix} \left(\begin{bmatrix} \delta \\ 0 \end{bmatrix} + M \right)
$$

where $\widehat{K} = E[\mathbf{x}_i' Q \mathbf{x}_i], \widetilde{K} = E[\mathbf{x_i}' \Omega^{-1} \mathbf{x}_i]^{-1}$, and $M \sim N(0, \widetilde{V})$, with

$$
\widetilde{V} = \begin{bmatrix} (T\sigma_\alpha^2 + \sigma_\epsilon^2)^2 E[\mathbf{x}_i' \Omega^{-1} \mathbf{x}_i] & (T\sigma_\alpha^2 + \sigma_\epsilon^2) E[\mathbf{x}_i' Q \mathbf{x}_i] \\ (T\sigma_\alpha^2 + \sigma_\epsilon^2) E[\mathbf{x}_i' Q \mathbf{x}_i] & \sigma_\epsilon^2 E[\mathbf{x}_i' Q \mathbf{x}_i] \end{bmatrix}
$$

Also, we know that

$$
\sqrt{N}(\widehat{\beta}_{FE} - \beta) \to_d N\big(0,\ \sigma_\epsilon^2 E[\mathbf{x}_i' Q \mathbf{x}_i]^{-1}\big)
$$

We can find A which satisfies, $[0 \quad A][\delta \quad 0]' = 0$ and

$$
[0 \quad A]\widetilde{V}[0 \quad A]' = \sigma_\epsilon^2 E[\mathbf{x}_i' Q \mathbf{x}_i]^{-1}
$$

That is,

$$
\begin{aligned}
&[0 \quad A] \begin{bmatrix} (T\sigma_\alpha^2+\sigma_\epsilon^2)^2 E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i] & (T\sigma_\alpha^2+\sigma_\epsilon^2)E[\mathbf{x}_i'Q\mathbf{x}_i] \\ (T\sigma_\alpha^2+\sigma_\epsilon^2)E[\mathbf{x}_i'Q\mathbf{x}_i] & \sigma_\epsilon^2 E[\mathbf{x}_i'Q\mathbf{x}_i] \end{bmatrix} \begin{bmatrix} 0 \\ A \end{bmatrix} \\
&= A\sigma_\epsilon^2 E[\mathbf{x}_i'Q\mathbf{x}_i]A \\
&= \sigma_\epsilon^2 E[\mathbf{x}_i'Q\mathbf{x}_i]^{-1} \qquad\qquad \therefore A = E[\mathbf{x}_i'Q\mathbf{x}_i]^{-1}
\end{aligned}
$$

Likewise, we know that

$$
\sqrt{N}(\widehat{\beta}_{GLS}-\beta) \to_d N\big(\frac{\delta}{T\sigma_\alpha^2+\sigma_\epsilon^2}E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i]^{-1},\ E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i]^{-1}\big)
$$

We can fine B such that $[B \quad 0][\delta \quad 0]' = \frac{\delta}{T\sigma_\alpha^2+\sigma_\epsilon^2}E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i]^{-1}$ and $[B \quad 0]\widetilde{V}[B \quad 0]' = E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i]^{-1}$. We can easily get that

$$
B = \frac{E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i]^{-1}}{T\sigma_\alpha^2+\sigma_\epsilon^2}.
$$

To conclude, the joint distribution of $\widehat{\beta}_{RE}, \widehat{\beta}_{FE}$ and $\widehat{\delta}$ is

$$
\begin{bmatrix} \sqrt{N}(\widehat{\beta}_{RE}-\beta) \\ \sqrt{N}(\widehat{\beta}_{FE}-\beta) \\ \widehat{\delta} \end{bmatrix} \to_d \begin{bmatrix} \frac{\widetilde{K}}{T\sigma_\alpha^2+\sigma_\epsilon^2} & 0 \\ 0 & \widehat{K}^{-1} \\ 1 & \frac{-(T\sigma_\alpha^2+\sigma_\epsilon^2)}{\widehat{K}\widetilde{K}} \end{bmatrix} \Big(\begin{bmatrix} \delta \\ 0 \end{bmatrix} + M \Big)
$$

where $M \sim N(0, \widetilde{V})$, with

$$
\widetilde{V} = \begin{bmatrix} (T\sigma_\alpha^2+\sigma_\epsilon^2)^2 E[\mathbf{x}_i'\Omega^{-1}\mathbf{x}_i] & (T\sigma_\alpha^2+\sigma_\epsilon^2)E[\mathbf{x}_i'Q\mathbf{x}_i] \\ (T\sigma_\alpha^2+\sigma_\epsilon^2)E[\mathbf{x}_i'Q\mathbf{x}_i] & \sigma_\epsilon^2 E[\mathbf{x}_i'Q\mathbf{x}_i] \end{bmatrix}
$$

We can rewrite it as follows: (I will denote $\widehat{\delta}$ as $\widehat{\tau}$ from now on, as we discussed before.)

$$\begin{bmatrix} \sqrt{N}(\widehat{\beta}_{RE}-\beta) \\ \sqrt{N}(\widehat{\beta}_{FE}-\beta) \\ \widehat{\tau} \end{bmatrix} \quad \rightarrow_d \quad N\left(\quad \begin{bmatrix} \frac{\widetilde{K}}{T\sigma_\alpha^2+\sigma_\epsilon^2}\tau \\ 0 \\ \tau \end{bmatrix}, \quad \mathcal{V} \quad \right)$$

where

$$\mathcal{V} = \begin{bmatrix} \widetilde{K} & \widetilde{K} & 0 \\ \widetilde{K} & \sigma_\epsilon^2 \widehat{K}^{-1} & (T\sigma_\alpha^2+\sigma_\epsilon^2)\left(1-\frac{\sigma_\epsilon^2}{\widehat{K}\widetilde{K}}\right) \\ 0 & (T\sigma_\alpha^2+\sigma_\epsilon^2)\left(1-\frac{\sigma_\epsilon^2}{\widehat{K}\widetilde{K}}\right) & -\frac{(T\sigma_\alpha^2+\sigma_\epsilon^2)^2}{\widetilde{K}}\left(1-\frac{\sigma_\epsilon^2}{\widehat{K}\widetilde{K}}\right) \end{bmatrix}$$

Define

$$\begin{aligned} c &= \frac{\widetilde{K}}{T\sigma_\alpha^2+\sigma_\epsilon^2} \\ \sigma^2 &= -\frac{(T\sigma_\alpha^2+\sigma_\epsilon^2)^2}{\widetilde{K}}\left(1-\frac{\sigma_\epsilon^2}{\widehat{K}\widetilde{K}}\right) \\ \eta^2 &= \widetilde{K} \end{aligned}$$

We can easily verify that $\sigma_\epsilon^2 \widehat{K}^{-1} = c^2\sigma^2 + \widetilde{K}$:

$$\begin{aligned} c^2\sigma^2 + \widetilde{K} &= \frac{\widetilde{K}^2}{(T\sigma_\alpha^2+\sigma_\epsilon^2)^2}\left(-\frac{(T\sigma_\alpha^2+\sigma_\epsilon^2)^2}{\widetilde{K}}\left(1-\frac{\sigma_\epsilon^2}{\widehat{K}\widetilde{K}}\right)\right) + \widetilde{K} \\ &= -\widetilde{K}\left(1-\frac{\sigma_\epsilon^2}{\widehat{K}\widetilde{K}}\right) + \widetilde{K} = \sigma_\epsilon^2 \widehat{K}^{-1} \end{aligned}$$

Then the previous joint distribution can be simplified as

$$\begin{bmatrix} \sqrt{N}(\widehat{\beta}_{RE}-\beta) \\ \sqrt{N}(\widehat{\beta}_{FE}-\beta) \\ \widehat{\tau} \end{bmatrix} \quad \rightarrow_d \quad N\left(\quad \begin{bmatrix} c\tau \\ 0 \\ \tau \end{bmatrix}, \quad \begin{bmatrix} \eta^2 & \eta^2 & 0 \\ \eta^2 & c^2\sigma^2+\eta^2 & -c\sigma^2 \\ 0 & -c\sigma^2 & \sigma^2 \end{bmatrix} \quad \right)$$

Using the expression for conditional distribution of jointly normally distributed random variables, we get

$$\begin{bmatrix} \sqrt{N}(\widehat{\beta}_{RE} - \beta) \\ \sqrt{N}(\widehat{\beta}_{FE} - \beta) \end{bmatrix} | \widehat{\tau} \quad \to_d \quad N \left(\quad \begin{bmatrix} c\tau \\ c\tau - c\widehat{\tau} \end{bmatrix}, \quad \eta^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \right)$$

Now we can think of AMSE of fixed/random effect estimators:

$$AMSE(\widehat{\beta}_{RE}) = (c\tau)^2 + \widetilde{K}$$
$$AMSE(\widehat{\beta}_{FE}) = 0^2 + c^2\sigma^2 + \widetilde{K}$$

By plugging in $\widehat{\tau}^2 - \widehat{\sigma}_\tau^2$ instead of τ^2, we get FMSC: Use $\widehat{\beta}_{RE}$ if

$$\begin{aligned} c^2(\widehat{\tau}^2 - \widehat{\sigma}_\tau^2) + \widetilde{K} &\leq c^2\sigma^2 + \widetilde{K} \\ &\Leftrightarrow \widehat{\tau}^2 \leq \sigma^2 + \sigma^2 \qquad (\because \widehat{\sigma}_\tau^2 = \sigma^2) \\ &\Leftrightarrow \mid \widehat{\tau} \mid \leq \sqrt{2}\sigma \end{aligned}$$

Then,

$$\sqrt{N}(\widehat{\beta}_{FMSC} - \beta) = I(\mid \widehat{\tau} \mid \leq \sqrt{2}\sigma)\sqrt{N}(\widehat{\beta}_{RE} - \beta) + I(\mid \widehat{\tau} \mid > \sqrt{2}\sigma)\sqrt{N}(\widehat{\beta}_{FE} - \beta)$$