

Set-up

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

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

$$\epsilon_{it} \text{ i.i.d. over } i, t, \quad \text{var}(\epsilon_{it}) = \sigma_\epsilon^2$$

$$\alpha_i \text{ i.i.d. over } i, \quad \text{var}(\alpha_i) = \sigma_\alpha^2$$

$$\text{cov}(\alpha_i, \epsilon_{it}) = 0$$

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

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

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

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

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 $\text{cov}(x_{it}, \alpha_i) = 0$ for all i, t , GLS estimator is efficient. Define

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

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 $\hat{\beta}_{FE}$ is

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

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

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

(b) GLS Estimator

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

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

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

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

Note that the last term is relevant to local misspecification and this results in asymptotic bias.

Asymptotic variance of $\hat{\beta}_{GLS}$ is same as standard case, i.e., $Var(\hat{\beta}_{GLS}) = E[\mathbf{x}_i' \Omega^{-1} \mathbf{x}_i]^{-1}$. To conclude, the asymptotic distribution of $\hat{\beta}_{GLS}$ under local misspecification is

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

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

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

where

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

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

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

Comparison to Hausman Test

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

We can see that

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

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

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

AMSE-Optimal Weight

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

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

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

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

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

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 $AMSE$,

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

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 $\hat{\beta}_{FE}$ to the moment condition.

Denote the consistent estimators for $\sigma_\alpha^2, \sigma_\epsilon^2$ as $\hat{\sigma}_\alpha^2, \hat{\sigma}_\epsilon^2$ respectively. Define

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

Consider

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

We can get

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

where

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

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

$$\hat{\delta} \rightarrow_d N\left(\delta, \underbrace{\begin{bmatrix} 1 & \frac{-(T\sigma_\alpha^2 + \sigma_\epsilon^2)}{\hat{K}\tilde{K}} \end{bmatrix} \tilde{V} \begin{bmatrix} 1 & \frac{-(T\sigma_\alpha^2 + \sigma_\epsilon^2)}{\hat{K}\tilde{K}} \end{bmatrix}'}_{\equiv \hat{\sigma}_\delta^2}\right)$$

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

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

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

$$\begin{aligned} S_T^{-1} &= P_T + \frac{1}{\eta} Q_T \\ S_T^{-\frac{1}{2}} &= P_T + \frac{1}{\sqrt{\eta}} Q_T \quad (\because P_T Q_T = 0) \end{aligned}$$

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

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

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 $\tilde{\mathbf{y}}_i \equiv C_T \mathbf{y}_i$ and $\tilde{\mathbf{x}}_i \equiv C_T \mathbf{x}_i$, i.e.,

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

Note that $E[\tilde{\mathbf{v}}_i \tilde{\mathbf{v}}_i'] = E[C_T \mathbf{v}_i \mathbf{v}_i' C_T'] = \sigma_\epsilon^2 I_T$. Therefore, OLS estimator from the above model is BLUE from Gauss-Markov Theorem. Also, we can see that $\tilde{\mathbf{y}}_i$ 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,

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

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

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

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

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

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

$$\widehat{\delta} \rightarrow_d \left[1 \quad \frac{-(T\sigma_\alpha^2 + \sigma_\epsilon^2)}{\widehat{K} \widetilde{K}} \right] \left(\begin{bmatrix} \delta \\ 0 \end{bmatrix} + M \right)$$

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

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

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,

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

Likewise, we know that

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

We can find 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 $[B \quad 0]\tilde{V}[B \quad 0]' = E[\mathbf{x}_i' \Omega^{-1} \mathbf{x}_i]^{-1}$.

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 $\hat{\beta}_{RE}$, $\hat{\beta}_{FE}$ and $\hat{\delta}$ is

$$\begin{bmatrix} \sqrt{N}(\hat{\beta}_{RE} - \beta) \\ \sqrt{N}(\hat{\beta}_{FE} - \beta) \\ \hat{\delta} \end{bmatrix} \rightarrow_d \begin{bmatrix} \frac{\tilde{K}}{T\sigma_\alpha^2 + \sigma_\epsilon^2} & 0 \\ 0 & \hat{K}^{-1} \\ 1 & \frac{-(T\sigma_\alpha^2 + \sigma_\epsilon^2)}{\hat{K}\tilde{K}} \end{bmatrix} \left(\begin{bmatrix} \delta \\ 0 \end{bmatrix} + M \right)$$

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

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

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

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

where

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

Define

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

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

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

Then the previous joint distribution can be simplified as

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

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

we get

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

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

$$AMSE(\hat{\beta}_{RE}) = (c\tau)^2 + \tilde{K}$$

$$AMSE(\hat{\beta}_{FE}) = 0^2 + c^2\sigma^2 + \tilde{K}$$

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

$$c^2(\hat{\tau}^2 - \hat{\sigma}_\tau^2) + \tilde{K} \leq c^2\sigma^2 + \tilde{K}$$

$$\Leftrightarrow \hat{\tau}^2 \leq \sigma^2 + \sigma^2 \quad (\because \hat{\sigma}_\tau^2 = \sigma^2)$$

$$\Leftrightarrow |\hat{\tau}| \leq \sqrt{2}\sigma$$

Then,

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