# Final

Due by 5/21 @ 11:59PM

**Generate a `Python` or `R` script for your answers. You will submit your script only. Read the questions carefully. All of the terms that you will need to use in your functions are already defined in the questions.*

(1) (50 pts) Consider the following two-way error component model:

$$y_{it} = \alpha + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it}, \quad \varepsilon_{it} = \mu_i + \lambda_t + \nu_{it}$$

for $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$. If the individual specific effects ($\mu_i$) and time specific effects ($\lambda_t$) are correlated with the time-varying regressors ($\boldsymbol{x}_{it}$) in an arbitrary manner, they have to be treated as fixed parameters. We can write the model for the entire sample by stacking up the observations:

$$\boldsymbol{y} = \alpha \boldsymbol{l}_{NT} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}_{\mu}\boldsymbol{\mu} + \boldsymbol{Z}_{\lambda}\boldsymbol{\lambda} + \boldsymbol{\nu} \tag{1}$$

where $\boldsymbol{Z}_{\mu} = \mathbf{I}_N \otimes \boldsymbol{l}_T$, $\boldsymbol{Z}_{\lambda} = \boldsymbol{l}_N \otimes \mathbf{I}_T$, $\mathbf{I}_N$ is the identity matrix of size $N$, $\mathbf{I}_T$ is the identity matrix of size $T$, $\boldsymbol{l}_T$ is a $T \times 1$ vector of ones, $\boldsymbol{l}_N$ is a $N \times 1$ vector of ones, and $\boldsymbol{l}_{NT}$ is a $NT \times 1$ vector of ones.

Assume for the identification that $\alpha = 0$. Let $\mathbb{P}_{\mu} = \boldsymbol{Z}_{\mu}(\boldsymbol{Z}_{\mu}'\boldsymbol{Z}_{\mu})^{-1}\boldsymbol{Z}_{\mu}' = \mathbf{I}_N \otimes (\frac{1}{T}\boldsymbol{l}_T\boldsymbol{l}_T')$, and similarly let $\mathbb{P}_{\lambda} = \boldsymbol{Z}_{\lambda}(\boldsymbol{Z}_{\lambda}'\boldsymbol{Z}_{\lambda})^{-1}\boldsymbol{Z}_{\lambda}' = (\frac{1}{N}\boldsymbol{l}_N\boldsymbol{l}_N') \otimes \mathbf{I}_T$. Also, let $\mathbb{P}_{\mu\lambda} = \mathbb{P}_{\mu}\mathbb{P}_{\lambda} = (\frac{1}{N}\boldsymbol{l}_N\boldsymbol{l}_N') \otimes (\frac{1}{T}\boldsymbol{l}_T\boldsymbol{l}_T')$. Now define

$$\mathbb{Q} = \mathbf{I}_{NT} - \mathbb{P}_{\mu} - \mathbb{P}_{\lambda} + \mathbb{P}_{\mu\lambda}.$$

Notice that $\mathbb{P}_{\mu}\boldsymbol{y}$ produces a vector with typical elements $\bar{y}_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T} y_{it}$; and $\mathbb{P}_{\mu}\boldsymbol{X}$ produces $\bar{\boldsymbol{x}}_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}_{it}$. Similarly, $\mathbb{P}_{\lambda}\boldsymbol{y}$ produces a vector with typical elements $\bar{y}_{\cdot t} = \frac{1}{N}\sum_{i=1}^{N} y_{it}$ ; $\mathbb{P}_{\lambda}\boldsymbol{X}$ produces $\bar{\boldsymbol{x}}_{\cdot t} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_{it}$. Also, $\mathbb{P}_{\mu\lambda}\boldsymbol{y}$ produces a vector with typical elements, $\bar{y}_{\cdot\cdot} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}$, and $\mathbb{P}_{\mu\lambda}\boldsymbol{X}$ produces $\bar{\boldsymbol{x}}_{\cdot\cdot} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\boldsymbol{x}_{it}$. Hence, $\mathbb{Q}\boldsymbol{y}$ produces a vector with typical elements, $y_{it} - \bar{y}_{i\cdot} - \bar{y}_{\cdot t} + \bar{y}_{\cdot\cdot}$, and $\mathbb{Q}\boldsymbol{X}$ produces a vector with typical elements, $\boldsymbol{x}_{it} - \bar{\boldsymbol{x}}_{i\cdot} - \bar{\boldsymbol{x}}_{\cdot t} + \bar{\boldsymbol{x}}_{\cdot\cdot}$. Notice also that $\mathbb{Q}$ is idempotent and symmetric.

The within estimator can be defined by premultiplying both sides of the model (1) with $\mathbb{Q}$ and then applying the least squares methodology. In other words,

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\mathbb{Q}\boldsymbol{X})^{-1}\boldsymbol{X}'\mathbb{Q}\boldsymbol{y}.$$

Assume that the following assumption holds.

**Assumption 1.** *(i) The idiosyncratic errors $\nu_{it} \sim iid(0, \sigma_{\nu}^2)$ for all $i$ and $t$; and (ii) $(\boldsymbol{x}_{it}', \mu_i, \lambda_t)'$ are uncorrelated with $\nu_{it}$.*

Under Assumption 1, an unbiased estimator of $\sigma_{\nu}^2$ is given as

$$\widehat{\sigma}_{\nu}^2 = \widehat{\boldsymbol{\nu}}'\widehat{\boldsymbol{\nu}}/\big((N-1)(T-1) - K\big)$$

where $\widehat{\nu}_{it} = y_{it} - \boldsymbol{x}_{it}^{'}\widehat{\boldsymbol{\beta}} - \widehat{\mu}_i - \widehat{\lambda}_t$, and $K$ is the number of columns in $\boldsymbol{X}$. Moreover, $\widehat{\mu}_i = \bar{y}_{i\cdot} - \bar{\boldsymbol{x}}_{i\cdot}^{'}\widehat{\boldsymbol{\beta}}$, and $\widehat{\lambda}_t = \bar{y}_{\cdot t} - \bar{\boldsymbol{x}}_{\cdot t}^{'}\widehat{\boldsymbol{\beta}}$.

(a) Write a function that will take $\boldsymbol{y}$ and $\boldsymbol{X}$ as inputs, fit a two-way fixed effects model using the within methodology, and return $\widehat{\boldsymbol{\beta}}$ and $\widehat{\sigma}_{\nu}^2$.

(b) Using the Munnell data set from Assignment 7, estimate the following two-way fixed effects model using your function:

$$\begin{aligned}\ln \mathrm{GSP}_{it} = \alpha &+ \beta_1 \ln(\mathrm{P\_CAP}_{it} + \mathrm{HWY}_{it} + \mathrm{WATER}_{it} + \mathrm{UTIL}_{it}) \\ &+ \beta_2 \ln \mathrm{PC}_{it} + \beta_3 \ln \mathrm{EMP}_{it} + \beta_4 \mathrm{UNEMP}_{it} + \mu_i + \lambda_t + \nu_{it}.\end{aligned}$$

(c) Estimate the model above with the dummy variable approach, i.e., by including state and year dummies in the regression. Compare the results with the ones from (b). Are they the same?

(2) (50 points) Let the population model be given by

$$\boldsymbol{y} = \mathbb{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} = \boldsymbol{Z}_{\mu}\boldsymbol{\mu} + \boldsymbol{\nu} \tag{2}$$

where $\mathbb{X} = (\boldsymbol{l}_{NT}, \boldsymbol{X})$, $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^{'})^{'}$, $\boldsymbol{Z}_{\mu} = \mathbf{I}_N \otimes \boldsymbol{l}_T$, $\mathbf{I}_N$ is the identity matrix of size $N$, $\boldsymbol{l}_T$ is a $T \times 1$ vector of ones, and $\boldsymbol{l}_{NT}$ is a $NT \times 1$ vector of ones. Assume that the following assumption holds.

**Assumption 2.** *(i) The idiosyncratic errors $\nu_{it}$s are iid normal with $(0, \sigma_{\nu}^2)$; (ii) $\mu_i$s are iid normal random variables with $(0, \sigma_{\mu}^2)$, and they are uncorrelated with $\boldsymbol{x}_{it}$s, (iii) $\mu_i$s are independent of $\nu_{it}$s.*

Let $\boldsymbol{\Omega}$ denote $\mathrm{Var}(\boldsymbol{\varepsilon})$. Then, one can show that

$$\begin{aligned}\boldsymbol{\Omega} &= \mathbb{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{'}|\mathbb{X}) = \boldsymbol{Z}_{\mu}\mathbb{E}(\boldsymbol{\mu}\boldsymbol{\mu}^{'}|\mathbb{X})\boldsymbol{Z}_{\mu}^{'} + \mathbb{E}(\boldsymbol{\nu}\boldsymbol{\nu}^{'}|\mathbb{X}) \\ &= \sigma_{\mu}^2(\mathbf{I}_N \otimes \mathbf{J}_T) + \sigma_{\nu}^2(\mathbf{I}_N \otimes \mathbf{I}_T)\end{aligned}$$

where $\mathbf{J}_T = \boldsymbol{l}_T\boldsymbol{l}_T^{'}$, a $T \times T$ matrix of ones. Using the spectral decomposition theorem, it can be shown that

$$\begin{aligned}\boldsymbol{\Omega} &= (T\sigma_{\mu}^2 + \sigma_{\nu}^2)(\mathbf{I}_T \otimes \bar{\mathbf{J}}_T) + \sigma_{\nu}^2(\mathbf{I}_N \otimes \mathbb{Q}_T) \\ &= \sigma_1^2\mathbb{P}_{\mu} + \sigma_{\nu}^2\mathbb{Q}_{\mu}\end{aligned}$$

where $\sigma_1^2 = T\sigma_{\mu}^2 + \sigma_{\nu}^2$, $\bar{\mathbf{J}}_T = \frac{1}{T}\mathbf{J}_T$ and $\mathbb{Q}_T = \mathbf{I}_T - \bar{\mathbf{J}}_T$.Then,

$$\boldsymbol{\Omega}^{-1} = \sigma_1^{-2}\mathbb{P}_{\mu} + \sigma_{\nu}^{-2}\mathbb{Q}_{\mu}$$

and

$$|\boldsymbol{\Omega}| = (\sigma_{\nu}^2)^{N(T-1)}(\sigma_1^2)^N$$

where $|\cdot|$ denotes the determinant of a square matrix. Then, the log-likelihood of the model is given by

$$\ln f(\boldsymbol{y}; \mathbb{X}\boldsymbol{\theta}, \boldsymbol{\Omega}) = -\frac{NT}{2}\ln(2\pi) - \frac{1}{2}\ln|\boldsymbol{\Omega}| - \frac{1}{2}(\boldsymbol{y} - \mathbb{X}\boldsymbol{\theta})'\boldsymbol{\Omega}^{-1}(\boldsymbol{y} - \mathbb{X}\boldsymbol{\theta})$$

and the maximum likelihood estimators of $(\boldsymbol{\theta}', \sigma_\nu^2, \sigma_\mu^2)'$ are given by

$$(\widehat{\boldsymbol{\theta}}, \widehat{\sigma}_\nu^2, \widehat{\sigma}_\mu^2)' = \underset{\{\boldsymbol{\theta}, \sigma_\nu^2, \sigma_\mu^2\}}{\arg\max} \ln f(\boldsymbol{y}; \mathbb{X}\boldsymbol{\theta}, \boldsymbol{\Omega}).$$

(a) Write a function that will take $\boldsymbol{y}$ and $\boldsymbol{X}$ as inputs, fit a one-way random individual effects model using the maximum likelihood methodology, and return $\widehat{\boldsymbol{\theta}}$, $\widehat{\sigma}_\mu^2$, and $\widehat{\sigma}_\nu^2$. To this end, you can use `minimize()` from `scipy.optimize` in Python, and `optim()` or `nlm()` in R. Note that these are minimizers and the objective function needs to be stated as a minimization problem. Also, choose `BFGS` as the method.

(b) Using the Munnell data set from Assignment 7, estimate the following one-way random individual effects model using your function:

$$\begin{aligned} \ln \mathrm{GSP}_{it} = \alpha &+ \beta_1 \ln(\mathrm{P\_CAP}_{it} + \mathrm{HWY}_{it} + \mathrm{WATER}_{it} + \mathrm{UTIL}_{it}) \\ &+ \beta_2 \ln \mathrm{PC}_{it} + \beta_3 \ln \mathrm{EMP}_{it} + \beta_4 \mathrm{UNEMP}_{it} + \mu_i + \nu_{it}. \end{aligned}$$