FWL Theorem, econometric application: FE estimator

Consider the panel LRM

$$y_{it} = \boldsymbol{x}'_{it}\boldsymbol{\beta} + \iota\alpha_i + \varepsilon_{it}.$$

y_{it}: an observation for individual i at time t.

$\boldsymbol{x}'_{it}$: K observations for K regressors for individual i at time t. $1 \times K$.

$\boldsymbol{\beta}$: vector of true coefficients. $K \times 1$.

ι: scalar with a value of 1. Greek letter 'iota'.

α_i: time invariant constant term specific to individual i in the panel. Potentially correlated with $\boldsymbol{x}'_{it}$. It captures individual heterogeneity.

ε_{it}: error term. It meets the OLS assumptions.

FWL Theorem, econometric application: FE estimator

There are T observations available for each i. If we stack the T observations, we obtain

$$\mathbf{y}_i = \mathbf{X}_i'\boldsymbol{\beta} + \boldsymbol{\iota}\alpha_i + \boldsymbol{\varepsilon}_i.$$

$\mathbf{y}_i$: $T \times 1$.

$\mathbf{X}_i'$: T observations for i for K independent variables. $T \times K$.

$\mathbf{x}_{it}'$: row vector in row t of $\mathbf{X}_i'$. $1 \times K$. It contains k observations for k regressors for individual i at time t.

$\boldsymbol{\iota}$: column vector containing 1 in every row. $T \times 1$.

$\boldsymbol{\varepsilon}_i$: $T \times 1$.

FWL Theorem, econometric application: FE estimator

There are N individuals. If we stack the N individuals, we obtain

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{D}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}.$$

$\mathbf{y}$: $NT \times 1$.

$\mathbf{X}$: $NT \times K$.

$\mathbf{D}$: has N diagonal elements. Each element of the diagonal is a vector, is the same, and is given by the column vector $\boldsymbol{\iota}$. All of the off-diagonal elements are $\mathbf{0}$ column vectors of size $T \times 1$. Hence, $\mathbf{D}$ is $NT \times N$.

$\boldsymbol{\alpha}$: $N \times 1$ since there are N different α_is.

This is the Least Squares Dummy Variable (LSDV) model.

FWL Theorem, econometric application: FE estimator

For individual i, $T = 3$, and $K = 3$,

$$\begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \end{bmatrix} = \begin{bmatrix} k_{i1} & l_{i1} & m_{i1} \\ k_{i2} & l_{i2} & m_{i2} \\ k_{i3} & l_{i3} & m_{i3} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \alpha_i + \begin{bmatrix} \varepsilon_i \\ \varepsilon_i \\ \varepsilon_i \end{bmatrix}$$

where k, l, m represent three different regressors. Putting them into row vector $\mathbf{x}'_{it}$,

$$\underbrace{\begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \end{bmatrix}}_{\mathbf{y}_i} = \underbrace{\begin{bmatrix} \mathbf{x}'_{i1} \\ \mathbf{x}'_{i2} \\ \mathbf{x}'_{i3} \end{bmatrix}}_{\mathbf{X}'_i} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{\boldsymbol{\iota}} \alpha_i + \underbrace{\begin{bmatrix} \varepsilon_i \\ \varepsilon_i \\ \varepsilon_i \end{bmatrix}}_{\boldsymbol{\varepsilon}_i}$$

FWL Theorem, econometric application: FE estimator

Assume $N = 3$. Stack N individuals to obtain

$$\underset{NT\times 1}{\underbrace{\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\\ \mathbf{y}_3\end{bmatrix}}_{\mathbf{y}}} = \underset{NT\times K}{\underbrace{\begin{bmatrix}\mathbf{X}_1'\\ \mathbf{X}_2'\\ \mathbf{X}_3'\end{bmatrix}}_{\mathbf{X}}} \underset{K\times 1}{\underbrace{\begin{bmatrix}\beta_1\\ \beta_2\\ \beta_3\end{bmatrix}}_{\boldsymbol{\beta}}} + \underset{NT\times N}{\underbrace{\begin{bmatrix}\boldsymbol{\iota} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\iota} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \boldsymbol{\iota}\end{bmatrix}}_{\boldsymbol{D}}} \underset{N\times 1}{\underbrace{\begin{bmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{bmatrix}}_{\boldsymbol{\alpha}}} + \underset{N\times 1}{\underbrace{\begin{bmatrix}\boldsymbol{\varepsilon}_1\\ \boldsymbol{\varepsilon}_2\\ \boldsymbol{\varepsilon}_3\end{bmatrix}}_{\boldsymbol{\varepsilon}}}$$

where

$$\boldsymbol{D} = \begin{bmatrix}1 & 0 & 0\\ 1 & 0 & 0\\ 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 1\\ 0 & 0 & 1\end{bmatrix}$$

FWL Theorem, econometric application: FE estimator

The LSDV model has two problems. First, it requires the inversion of a very large matrix due to $\boldsymbol{D}$. Second, it requires estimation of the large number of intercept terms contained in α. Could we avoid these problems?

FWL Theorem, econometric application: FE estimator

Consider again the panel model for individual i at time t

$$y_{it} = \boldsymbol{x}'_{it}\boldsymbol{\beta} + \iota\alpha_i + \varepsilon_{it}.$$

Take the average over all t for individual i to obtain

$$\bar{y}_i = \bar{\boldsymbol{x}}'_i\boldsymbol{\beta} + \iota\alpha_i + \bar{\varepsilon}_i,$$

where

$$\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}.$$

Subtract the second equation from the first to obtain

$$y_{it} - \bar{y}_i = (\boldsymbol{x}'_{it} - \bar{\boldsymbol{x}}'_i)\boldsymbol{\beta} + \varepsilon_{it} - \bar{\varepsilon}_i.$$

The is the fixed effects transformation. The time invariant individual specific constant term α_i disappeared!

FWL Theorem, econometric application: FE estimator

We have carried out the fixed effects transformation for individual i using his T observations. We need to take account of the fact that we have n individuals in the panel data.

FWL Theorem, econometric application: FE estimator

The panel model for N individuals described above is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{D}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}.$$

Consider the model

$$\mathbf{y} = \mathbf{M}_D\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\upsilon}.$$

where

$$\mathbf{M}_D = \mathbf{I} - \mathbf{P}_D.$$

and

$$\mathbf{P}_D = \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'.$$

FWL theorem: OLS estimator of $\boldsymbol{\beta}$ in the stated two models is the same and given by

$$\begin{aligned}\mathbf{b}_{OLS} &= ((\mathbf{M}_D\mathbf{X})'(\mathbf{M}_D\mathbf{X}))^{-1}(\mathbf{M}_D\mathbf{X})'\mathbf{y} = (\mathbf{X}'\mathbf{M}_D\mathbf{X})^{-1}\mathbf{X}'\mathbf{M}_D\mathbf{y}\\ &= \mathbf{b}_{FE}.\end{aligned}$$

This is the fixed effects estimator.

FWL Theorem, econometric application: FE estimator

Why does $\mathbf{M}_D$ demeans the data? Take a closer look at

$$\mathbf{M}_D = \mathbf{I} - \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'.$$

Realise that

$$\mathbf{D}'\mathbf{D} = T.$$

and hence

$$\mathbf{M}_D = \mathbf{I} - \frac{1}{T}\mathbf{D}\mathbf{D}'$$

which is called the centering matrix.

FWL Theorem, econometric application: FE estimator

If $T = 3$ and $N = 2$,

$$\boldsymbol{M}_D = \begin{bmatrix} 1-\frac{1}{T} & -\frac{1}{T} & -\frac{1}{T} & 0 & 0 & 0 \\ -\frac{1}{T} & 1-\frac{1}{T} & -\frac{1}{T} & 0 & 0 & 0 \\ -\frac{1}{T} & -\frac{1}{T} & 1-\frac{1}{T} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-\frac{1}{T} & -\frac{1}{T} & -\frac{1}{T} \\ 0 & 0 & 0 & -\frac{1}{T} & 1-\frac{1}{T} & -\frac{1}{T} \\ 0 & 0 & 0 & -\frac{1}{T} & -\frac{1}{T} & 1-\frac{1}{T} \end{bmatrix}$$

FWL Theorem, econometric application: FE estimator

Then, assuming values for $\boldsymbol{X}$, $\boldsymbol{X}$ in deviation form is

$$
\boldsymbol{M}_D\boldsymbol{X} = \begin{bmatrix} 1-\frac{1}{T} & -\frac{1}{T} & -\frac{1}{T} & 0 & 0 & 0 \\ -\frac{1}{T} & 1-\frac{1}{T} & -\frac{1}{T} & 0 & 0 & 0 \\ -\frac{1}{T} & -\frac{1}{T} & 1-\frac{1}{T} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-\frac{1}{T} & -\frac{1}{T} & -\frac{1}{T} \\ 0 & 0 & 0 & -\frac{1}{T} & 1-\frac{1}{T} & -\frac{1}{T} \\ 0 & 0 & 0 & -\frac{1}{T} & -\frac{1}{T} & 1-\frac{1}{T} \end{bmatrix} \begin{bmatrix} 1 & 12 \\ 1 & 13 \\ 1 & 11 \\ 3 & 43 \\ 4 & 46 \\ 4 & 41 \end{bmatrix}
$$

$$
= \begin{bmatrix} 1-\frac{1+1+1}{3} & 12-\frac{12+13+11}{3} \\ 1-\frac{1+1+1}{3} & 13-\frac{12+13+11}{3} \\ 1-\frac{1+1+1}{3} & 11-\frac{12+13+11}{3} \\ 3-\frac{3+4+4}{3} & 43-\frac{43+46+41}{3} \\ 4-\frac{3+4+4}{3} & 46-\frac{43+46+41}{3} \\ 4-\frac{3+4+4}{3} & 41-\frac{43+46+41}{3} \end{bmatrix}
$$