Dynamic Panel Data Models

Exercises

1. Consider the simple dynamic model,

$$\begin{aligned} y_{it} &= \gamma y_{i,t-1} + \alpha_i^* + \varepsilon_{it},\ i = 1, ..., N,\ t = 1, ..., T \\ \alpha_i^* &= \alpha + \alpha_i \end{aligned}$$

where $|\gamma| < 1$, y_{i0} is observable, $E(\varepsilon_{it}) = 0$, $E(\varepsilon_{it}\varepsilon_{jt}) = \sigma_\varepsilon^2$, $if\ i = j$ and $t = s$, and $E(\varepsilon_{it}\varepsilon_{jt}) = 0$ otherwise.

a) Define the LSDV estimators of α_i and γ.

b) Define the LSDV bias considering T fixed and $N \to \infty$.

c) Indicate two alternative estimators which overcome the bias problem discussed in b).

Proposed Solutions:

1. **a)** Define the LSDV estimators of α_i and γ.

Solution: Note that

$$\begin{aligned} \widehat{\alpha}_i &= \overline{y}_i - \widehat{\gamma}_{LSDV}\overline{y}_{i,-1} \\ \widehat{\gamma}_{LSDV} &= \left[\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{i,t-1} - \overline{y}_{i,-1})^2\right]^{-1}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{i,t-1} - \overline{y}_{i,-1})(y_{i,t} - \overline{y}_i)\right] \end{aligned}$$

where $\overline{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}$ and $\overline{y}_{i,-1} = \frac{1}{T}\sum_{t=1}^{T} y_{i,t-1}$.

1. **b)** LSDV bias.

Solution: Note that

$$\widehat{\gamma}_{LSDV} - \gamma = \left[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{i,t-1} - \overline{y}_{i,-1})^2\right]^{-1}\left[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{i,t-1} - \overline{y}_{i,-1})(\varepsilon_{i,t} - \overline{\varepsilon}_i)\right]$$

Consider first the numerator

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{i,t-1} - \overline{y}_{i,-1})(\varepsilon_{i,t} - \overline{\varepsilon}_i)$$

and assume that T is fixed. Thus,

$$\underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{i,t-1}-\overline{y}_{i,-1})(\varepsilon_{i,t}-\overline{\varepsilon}_i) \tag{1}$$

$$\begin{aligned} = \quad & \underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}y_{i,t-1}\varepsilon_{i,t}-\underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\overline{y}_{i,-1}\varepsilon_{i,t} \\ & -\underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}y_{i,t-1}\overline{\varepsilon}_i+\underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\overline{y}_{i,-1}\overline{\varepsilon}_i \end{aligned}$$

By definition,

$$\underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}y_{i,t-1}\varepsilon_{i,t}=E\left(y_{i,t-1}\varepsilon_{i,t}\right)=0$$

$$\begin{aligned} \underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}y_{i,t-1}\overline{\varepsilon}_i &= \underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\overline{\varepsilon}_i\sum_{t=1}^{T}y_{i,t-1} \\ &= \underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\overline{\varepsilon}_iT\overline{y}_{i,-1} \\ &= \underset{N\to\infty}{p\lim}\frac{1}{N}\sum_{i=1}^{N}\overline{\varepsilon}_i\overline{y}_{i,-1} \end{aligned}$$

Similarly,

$$\begin{aligned} \underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\overline{y}_{i,-1}\varepsilon_{i,t} &= \underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\overline{y}_{i,-1}\sum_{t=1}^{T}\varepsilon_{i,t} \\ &= \underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\overline{y}_{i,-1}T\overline{\varepsilon}_i \\ &= \underset{N\to\infty}{p\lim}\frac{1}{N}\sum_{i=1}^{N}\overline{y}_{i,-1}\overline{\varepsilon}_i \end{aligned}$$

$$\underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\overline{y}_{i,-1}\overline{\varepsilon}_i=\underset{N\to\infty}{p\lim}\frac{1}{N}\sum_{i=1}^{N}\overline{y}_{i,-1}\overline{\varepsilon}_i$$

Thus, substituting these results into (1) we observe for the numerator of the LSDV that,

$$\underset{N\to\infty}{p\lim}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{i,t-1}-\overline{y}_{i,-1})(\varepsilon_{i,t}-\overline{\varepsilon}_i)=-\underset{N\to\infty}{p\lim}\frac{1}{N}\sum_{i=1}^{N}\overline{y}_{i,-1}\overline{\varepsilon}_i.$$

Moreover, note that by recursive substitution y_{it} can be written as,

$$\begin{aligned} y_{it} &= \varepsilon_{it} + \gamma\varepsilon_{i,t-1} + \gamma^2\varepsilon_{i,t-2} + ... + \gamma^{t-1}\varepsilon_{i1} \\ &+ \frac{1-\gamma^t}{1-\gamma}\alpha_i^* + \gamma^t y_{i0} \end{aligned}$$

Similarly, for $y_{i,t-1}$

$$\begin{aligned} y_{i,t-1} &= \varepsilon_{i,t-1} + \gamma\varepsilon_{i,t-2} + \gamma^2\varepsilon_{i,t-3} + ... + \gamma^{t-2}\varepsilon_{i1} \\ &+ \frac{1-\gamma^{t-1}}{1-\gamma}\alpha_i^* + \gamma^{t-1} y_{i0} \end{aligned}$$

and so on for the other y_{ik}, $k = t-2, ..., 1$.

Summing $y_{i,t-1}$ over t we establish that,

$$\begin{aligned} \sum_{t=1}^{T} y_{i,t-1} &= \varepsilon_{i,T-1} + \frac{1-\gamma^2}{1-\gamma}\varepsilon_{i,T-2} + ... + \frac{1-\gamma^{T-1}}{1-\gamma}\varepsilon_{i1} \\ &+ \frac{(T-1) - \gamma T + \gamma^T}{(1-\gamma)^2}\alpha_i^* + \frac{1-\gamma^T}{1-\gamma} y_{i0} \end{aligned}$$

Note that $\sum_{t=1}^{T} y_{i,t-1} = T\overline{y}_{i,-1}$.

Hence,

$$\begin{aligned} \underset{N\to\infty}{p\lim}\frac{1}{N}\sum_{i=1}^{N}\overline{y}_{i,-1}\overline{\varepsilon}_i &= \underset{N\to\infty}{p\lim}\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\left\{\varepsilon_{i,T-1} + \frac{1-\gamma^2}{1-\gamma}\varepsilon_{i,T-2} + ... + \frac{1-\gamma^{T-1}}{1-\gamma}\varepsilon_{i1}\right. \\ &\left. + \frac{(T-1) - \gamma T + \gamma^T}{(1-\gamma)^2}\alpha_i^* + \frac{1-\gamma^T}{1-\gamma} y_{i0}\right\} \times \\ &\times \left(\frac{1}{T}(\varepsilon_{i1} + ... + \varepsilon_{iT})\right) \end{aligned}$$

2. For $T = 2$ consider a standard unobserved effects model

$$y_{it} = \mathbf{x}_{it}\beta + c_i + u_{it}, \ t = 1, 2$$

Let $\widehat{\beta}_{FE}$ and $\widehat{\beta}_{FD}$ denote the fixed effects and first difference estimators, respectively.

a) Show that the FE and FD estimates are numerically identical.

b) Show that the error variance estimates from the FE and FD methods are numerically identical.

2. **Solution: a)** Let

$$\begin{aligned}\overline{x}_i &= \frac{(x_{i1}+x_{i2})}{2} \\ \overline{y}_i &= \frac{(y_{i1}+y_{i2})}{2}\end{aligned}$$

and

$$\begin{aligned}\widetilde{x}_{i1} &= x_{i1}-\overline{x}_i \qquad \widetilde{x}_{i2}=x_{i2}-\overline{x}_i \\ \widetilde{y}_{i1} &= y_{i1}-\overline{y}_i \qquad \widetilde{y}_{i2}=y_{i2}-\overline{y}_i\end{aligned}$$

For T=2 the fixed estimator is,

$$\widehat{\beta}_{FE}=\frac{\sum_{i=1}^{N}(\widetilde{x}'_{i1}\widetilde{y}_{i1}+\widetilde{x}'_{i2}\widetilde{y}_{i2})}{\sum_{i=1}^{N}(\widetilde{x}'_{i1}\widetilde{x}_{i1}+\widetilde{x}'_{i2}\widetilde{x}_{i2})}$$

Now by simple algebra

$$\begin{array}{c}\widetilde{x}_{i1}=\frac{(x_{i1}-x_{i2})}{2}=-\frac{\Delta x_{i2}}{2} \\ \widetilde{x}_{i2}=\frac{(x_{i2}-x_{i1})}{2}=\frac{\Delta x_{i2}}{2} \\ \widetilde{y}_{i1}=\frac{(y_{i1}-y_{i2})}{2}=-\frac{\Delta y_{i2}}{2} \\ \widetilde{y}_{i2}=\frac{(y_{i2}-y_{i1})}{2}=\frac{\Delta y_{i2}}{2}\end{array}$$

Therefore,

$$\begin{aligned}(\widetilde{x}'_{i1}\widetilde{x}_{i1}+\widetilde{x}'_{i2}\widetilde{x}_{i2}) &= \frac{\Delta x'_{i2}\Delta x_{i2}}{4}+\frac{\Delta x'_{i2}\Delta x_{i2}}{4} \\ &= \frac{\Delta x'_{i2}\Delta x_{i2}}{2}\end{aligned}$$

and

$$\begin{aligned}(\widetilde{x}'_{i1}\widetilde{y}_{i1}+\widetilde{x}'_{i2}\widetilde{y}_{i2}) &= \frac{\Delta x'_{i2}\Delta y_{i2}}{4}+\frac{\Delta x'_{i2}\Delta y_{i2}}{4} \\ &= \frac{\Delta x'_{i2}\Delta y_{i2}}{2}\end{aligned}$$

Thus,

$$
\begin{aligned}
\widehat{\beta}_{FE} &= \frac{\sum_{i=1}^{N} \left(\widetilde{x}'_{i1}\widetilde{y}_{i1} + \widetilde{x}'_{i2}\widetilde{y}_{i2}\right)}{\sum_{i=1}^{N} \left(\widetilde{x}'_{i1}\widetilde{x}_{i1} + \widetilde{x}'_{i2}\widetilde{x}_{i2}\right)} \\
&= \frac{\sum_{i=1}^{N} \frac{\Delta x'_{i2}\Delta y_{i2}}{2}}{\sum_{i=1}^{N} \frac{\Delta x'_{i2}\Delta x_{i2}}{2}} \\
&= \frac{\sum_{i=1}^{N} \Delta x'_{i2}\Delta y_{i2}}{\sum_{i=1}^{N} \Delta x'_{i2}\Delta x_{i2}} \\
&= \widehat{\beta}_{FD}
\end{aligned}
$$

2. **Solution: b)** Let

$$
\begin{aligned}
\widehat{u}_{i1} &= \widetilde{y}_{i1} - \widetilde{x}_{i1}\widehat{\beta}_{FE} \\
\widehat{u}_{i2} &= \widetilde{y}_{i2} - \widetilde{x}_{i2}\widehat{\beta}_{FE}
\end{aligned}
$$

be the fixed effects residuals for two time periods for cross section observation i.

Since $\widehat{\beta}_{FE} = \widehat{\beta}_{FD}$,

$$
\begin{aligned}
\widehat{u}_{i1} &= -\frac{\Delta y_{i2}}{2} - \left(-\frac{\Delta x_{i2}}{2}\right)\widehat{\beta}_{FD} \\
&= -\left(\frac{\Delta y_{i2}}{2} - \frac{\Delta x_{i2}}{2}\widehat{\beta}_{FD}\right) \\
&= -\frac{\widehat{e}_i}{2}
\end{aligned}
$$

$$
\begin{aligned}
\widehat{u}_{i2} &= \frac{\Delta y_{i2}}{2} - \frac{\Delta x_{i2}}{2}\widehat{\beta}_{FD} \\
&= \frac{\widehat{e}_i}{2}
\end{aligned}
$$

where $\widehat{e}_i$ are the first difference residuals.

Therefore,

$$\sum_{i=1}^{N} (\widehat{u}_{i1} + \widehat{u}_{i2}) = \frac{1}{2} \sum_{i=1}^{N} \widehat{e}_i$$

Since $T = 2$:

The FE variance estimate is $\frac{\sum_{i=1}^{N} (\widehat{u}_{i1}+\widehat{u}_{i2})}{N-k}$ and the FD variance estimate is $\frac{1}{2} \frac{\sum_{i=1}^{N} \widehat{e}_i}{N-k}$, thus, $\widehat{\sigma}_u^2 = \frac{\widehat{\sigma}_e^2}{2}$.

However,

$$\begin{aligned} Var(\widehat{\beta}_{FE}) &= \widehat{\sigma}_u^2 \left[\sum_{i=1}^{N} (\widetilde{x}_{i1}' \widetilde{x}_{i1} + \widetilde{x}_{i2}' \widetilde{x}_{i2}) \right]^{-1} \\ &= \frac{\widehat{\sigma}_e^2}{2} \left[\sum_{i=1}^{N} \frac{\Delta x_{i2}' \Delta x_{i2}}{2} \right]^{-1} \\ &= \widehat{\sigma}_e^2 \left[\sum_{i=1}^{N} \Delta x_{i2}' \Delta x_{i2} \right]^{-1} \\ &= Var(\widehat{\beta}_{FD}). \end{aligned}$$