

HW1

Wonjun Choi

September 19, 2022

1. (a) First, we assume $\{x_i, y_i\}$ is i.i.d.

Second, we assume $E[V_{it}|\eta_i, X_i] = 0$. For η_i , $E[\eta_i|X_i] = 0$ by construction:

$$E[\eta_i|\bar{X}_i, X_{it}] = E[\eta_i|X_{i1}, X_{i2}, \dots, X_{iT}] = E[\eta_i|X_i] = 0.$$

Now we show the consistency. Define $\theta \equiv (\delta_0, \gamma'_0, \beta'_0)'$ and $\tilde{X}_{it} \equiv (1, \bar{X}'_i, X'_{it})'$. Denote $(1, \dots, 1)' \in \mathbb{R}^k$ as ι_k . Then,

$$\begin{aligned} \hat{\theta}_{n, POLS} &= \left(\sum_i \sum_t \tilde{X}_{it} \tilde{X}'_{it} \right)^{-1} \left(\sum_i \sum_t \tilde{X}_{it} Y_{it} \right) \\ &= \left(\sum_i \sum_t \tilde{X}_{it} \tilde{X}'_{it} \right)^{-1} \left(\sum_i \sum_t \tilde{X}_{it} (\delta_0 + \bar{X}'_i \gamma_0 + X'_{it} \beta_0 + \eta_i + V_{it}) \right) \\ &= \theta_0 + \left(\sum_i \sum_t \tilde{X}_{it} \tilde{X}'_{it} \right)^{-1} \left(\sum_i \sum_t \tilde{X}_{it} (\eta_i + V_{it}) \right) \\ &= \theta_0 + \left(\sum_i \tilde{X}'_i \tilde{X}_i \right)^{-1} \left(\sum_i \tilde{X}'_i (\eta_i \iota_T + V_i) \right) \end{aligned}$$

To apply WLLN, we further assume $E[\tilde{X}'_i \tilde{X}_i]$ is finite and invertible and $E[\tilde{X}'_i \eta_i \tilde{X}_i] < \infty$, $E[\tilde{X}'_i V_i V'_i \tilde{X}_i] < \infty$. Note

$$\begin{aligned} E(\tilde{X}'_i \eta_i \iota_T) &= E[E(\tilde{X}'_i \eta_i \iota_T | X_i)] = 0, \\ E(\tilde{X}'_i V_i) &= E[E(\tilde{X}'_i V_i | X_i)] = 0. \end{aligned}$$

Applying WLLN, CMT, and Slutsky's lemma shows the second term converges in probability to 0. \square

- (b) Apply the FWL theorem to (1.46). First, observe that $\bar{X} = (\bar{X}'_1, \dots, \bar{X}'_n)'$ can be constructed by projecting X onto D where $D = (D'_1, \dots, D'_n)'$,

$D_i = (D_{it})_{t=1}^T = (D_{i1}, \dots, D_{iT})'$, $D'_{it} = (\delta_{j,it})_{j=1}^T$, $\delta_{j,it} = 1_{[j=i],it}$, $1_{[\cdot]}$ is an indicator function. Then

$$\bar{X} = D(D'D)^{-1}DX := P_DX$$

Let $M_X = I - X(X'X)^{-1}X'$. Then

$$\begin{aligned}\hat{\beta}_{FWL} &= (X'M_{(1,\bar{X})}X)^{-1}(X'M_{(1,\bar{X})}Y) \\ &= (X'M_{(1,P_DX)}X)^{-1}(X'M_{(1,P_DX)}Y) \\ &= (X'M_{(1,D)}X)^{-1}(X'M_{(1,D)}Y) \\ &= (X'M_DX)^{-1}(X'M_DY) \\ &= (\dot{X}'\dot{X})^{-1}(\dot{X}'\dot{Y}) \\ &= \hat{\beta}_{n,FE}\end{aligned}$$

where \dot{X} is a within-transform of X .

- (c) Check if CRE satisfies the assumptions of FE. Since CRE contains a time-invariant regressors 1 and \bar{X}_i , $E(\dot{X}'_i\dot{X}_i)$ is not invertible. Thus, CRE does not belong to a family of FE (X_{it} does not contain time-invariant regressors in CRE for invertibility, too). On the other hand, FE does not assume any specific functional form of the correlation between α_i and x_i so FE does not belong to a family of CRE. Therefore, we cannot say the statement is true.

(d)

$$\frac{\partial E[\alpha_i|X_i]}{\partial X'_i} = \begin{pmatrix} \gamma_1/T & \cdots & \gamma_k/T \\ \vdots & \vdots & \vdots \\ \gamma_1/T & \cdots & \gamma_k/T \end{pmatrix} = 0 \Leftrightarrow \gamma = 0$$

Therefore, we can simply test $H_0 : \gamma = 0$ from the OLS of (1.46).

2. (a) (Symmetric)

$$Q' = (I_n \otimes (I_T - J_T/T))' = I'_n \otimes (I'_T - J'_T/T) = Q$$

(Idempotent)

$$\begin{aligned}Q^2 &= (I_n \otimes (I_T - J_T/T))(I_n \otimes (I_T - J_T/T)) \\ &= I_n I_n \otimes (I_T - J_T/T)(I_T - J_T/T) \\ &= I_n \otimes [I_T - 2J_T/T + J_T^2/T^2] \\ &= I_n \otimes [I_T - 2J_T/T + T J_T/T^2] \\ &= I_n \otimes [I_T - J_T/T] = Q\end{aligned}$$

(Q is not invertible) Since Q is idempotent,

$$\text{rank}(Q) = \text{tr}(Q^2) = \text{tr}(Q) = n(T-1) < nT.$$

(b) Consider QY .

$$\begin{aligned} QY &= [I_n \otimes (I_T - J_T/T)]Y \\ &= \begin{pmatrix} I_T - J_T/T & O & \cdots & O \\ O & I_T - J_T/T & \cdots & O \\ \vdots & \vdots & \vdots & \vdots \\ O & \cdots & O & I_T - J_T/T \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \\ &= \dot{Y} \end{aligned}$$

Therefore,

$$\begin{aligned} \hat{\beta}_{n,FE} &= (\dot{X}'\dot{X})^{-1}(\dot{X}'\dot{Y}) \\ &= (XQ'QX)^{-1}(XQ'QY) \\ &= (X'QX)^{-1}(X'QY) \end{aligned}$$

(c) Denote $x_i = X'_i$, $y_i = Y_i$, $Q_i = I_T - J_T/T$. Consider the variance of \dot{v} .

$$\begin{aligned} \Omega_i &= E(\dot{v}_i\dot{v}'_i) \\ &= E(Q_i v_i v'_i Q_i) \\ &= \sigma_V^2 Q_i \end{aligned}$$

so that $\Omega \equiv E(\dot{v}\dot{v}') = I_n \otimes \sigma_V^2 Q_i = \sigma_V^2 Q$.

Hence, the GLS of (1.24) with the weighting matrix $1/\sigma_V^2 Q^-$ is

$$\begin{aligned} \hat{\beta}_{GLS} &= (\dot{X}' \frac{1/\sigma_V^2 Q^-}{\sigma_V^2} \dot{X})^{-1} (\dot{X}' \frac{1/\sigma_V^2 Q^-}{\sigma_V^2} \dot{Y}) \\ &= (X' Q Q^- Q X)^{-1} (X' Q Q^- Q Y) \\ &= (X' Q X)^{-1} (X' Q Y) \\ &= \hat{\beta}_{FE} \end{aligned}$$

Finally, we show the above GLS (in fact WLS) is efficient.

(d) Observe

$$\begin{aligned}
\sum_{i=1}^n \sum_{t=1}^2 (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' &= \sum_{i=1}^n \sum_{t=1}^2 (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \\
&= \sum_{i=1}^n \left[\left(x_{i2} - \frac{x_{i1} + x_{i2}}{2} \right) \left(x_{i2} - \frac{x_{i1} + x_{i2}}{2} \right)' \right. \\
&\quad \left. + \left(x_{i1} - \frac{x_{i1} + x_{i2}}{2} \right) \left(x_{i1} - \frac{x_{i1} + x_{i2}}{2} \right)' \right] \\
&= \sum_{i=1}^n \left(\frac{x_{i2}}{2} - \frac{x_{i1}}{2} \right) \left(\frac{x_{i2}}{2} - \frac{x_{i1}}{2} \right)' + \left(\frac{x_{i1}}{2} - \frac{x_{i2}}{2} \right) \left(\frac{x_{i1}}{2} - \frac{x_{i2}}{2} \right)' \\
&= \frac{1}{2} \sum_i \Delta x_{i2} \Delta x'_{i2}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\hat{\beta}_{n,FE} &= \left(\sum_i \sum_{t=1}^2 (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right)^{-1} \left(\sum_i \sum_{t=1}^2 (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i) \right) \\
&= \left(\frac{1}{2} \sum_i \Delta x_{i2} \Delta x'_{i2} \right)^{-1} \left(\frac{1}{2} \sum_i \Delta x_{i2} \Delta y_{i2} \right) = \hat{\beta}_{n,FD}.
\end{aligned}$$

3. (a) Plugging $y_{it} = \frac{1}{1-\beta_0} \alpha_i + \sum_{s=0}^{\infty} \beta_0^s v_{i,t-s}$ into (1.47),

$$\begin{aligned}
RHS &= \beta_0 \left(\frac{1}{1-\beta_0} \alpha_i + \sum_{s=0}^{\infty} \beta_0^s v_{i,t-1-s} \right) + \alpha_i + v_{it} \\
&= \frac{1}{1-\beta_0} \alpha_i + \sum_{s=0}^{\infty} \beta_0^{s+1} v_{i,t-1-s} + v_{it} \\
&= \frac{1}{1-\beta_0} \alpha_i + \sum_{s=1}^{\infty} \beta_0^s v_{i,t-s} + v_{it} \\
&= LHS
\end{aligned}$$

(b) Note

$$\begin{aligned}
\Delta y_{it} &= y_{it} - y_{i,t-1} = \sum_{s=0}^{\infty} \beta_0^s (v_{i,t-s} - v_{i,t-1-s}) \\
&= (v_{it} - v_{i,t-1}) + \beta_0(v_{i,t-1} - v_{i,t-2}) + \cdots \\
&= v_{it} - (1 - \beta_0)v_{i,t-1} - \beta_0(1 - \beta_0)v_{i,t-2} + \cdots \\
&= v_{it} - (1 - \beta_0) \sum_{s=0}^{\infty} \beta_0^s v_{i,t-1-s}
\end{aligned}$$

Then, denoting $\text{var}(v_{it}) = \sigma^2$,

$$\begin{aligned}
\Delta y_{it} \Delta y_{it} &= (v_{it}^2 + (1 - \beta_0)^2 v_{i,t-1}^2 + (1 - \beta_0)^2 \beta_0^2 v_{i,t-2}^2 + \cdots) \\
&\quad + (\text{cross products}) \\
E(\Delta y_{it} \Delta y_{it}) &= \sigma^2 + \frac{(1 - \beta_0)^2 \sigma^2}{1 - \beta_0^2} = \frac{2\sigma^2}{1 + \beta_0}
\end{aligned}$$

*Integration is done over the joint distribution of $(v_{it}, v_{i,t-1}, \dots)$.

Similarly,

$$\begin{aligned}
\Delta y_{it} \Delta y_{i,t-1} &= -(1 - \beta_0)v_{i,t-1}^2 + (1 - \beta_0)^2 \beta_0 v_{i,t-2}^2 \\
&\quad + (1 - \beta_0)^2 \beta_0^3 v_{i,t-3}^2 + \cdots \\
\Rightarrow E(\Delta y_{it} \Delta y_{i,t-1}) &= \frac{-1 + \beta_0}{1 + \beta_0} \sigma^2
\end{aligned}$$

Then, applying LLN over (i, t) ,

$$\begin{aligned}
\hat{\beta}_{n,FD} &= \left(\frac{1}{nt} \sum_{(i,t)} \Delta y_{i,t-1} \Delta y_{i,t-1} \right)^{-1} \left(\frac{1}{nt} \sum_{(i,t)} \Delta y_{i,t-1} \Delta y_{it} \right) \\
&\xrightarrow{p} \left(\frac{2\sigma^2}{1 + \beta_0} \right)^{-1} \frac{-1 + \beta_0}{1 + \beta_0} \sigma_V^2 = \frac{\beta_0 - 1}{2}
\end{aligned}$$

Thus,

$$\hat{\beta}_{n,FD} - \beta_0 \xrightarrow{p} \frac{-1 - \beta_0}{2}$$

so it does not depend on T ¹.

¹if finite period T is observed, just apply LLN over i and sum it over T

(c) Similarly to (b),

$$E(y_{i0})^2 = \frac{\sigma_\alpha^2}{(1 - \beta_0)^2} + \frac{\sigma^2}{1 - \beta_0^2}.$$

Meanwhile,

$$\begin{aligned} E[\Delta y_{i1} y_{i0}] &= E[-(1 - \beta_0) v_{i0}^2 - (1 - \beta_0) \beta_0^2 v_{i,-1}^2 - \dots + (\text{C.P.})] \\ &= -\frac{1 - \beta_0}{1 - \beta_0^2} \sigma^2 = \frac{-\sigma^2}{1 + \beta_0} \end{aligned}$$

Then,

$$\hat{\beta}_{OLS} = \left(\frac{1}{n} \sum_i \Delta y_{i1} \Delta y_{i1} \right)^{-1} \left(\frac{1}{n} \sum_i \Delta y_{i1} y_{i0} \right) \xrightarrow{p} \frac{(1 - \beta_0)^2 \sigma^2}{(1 + \beta_0) \sigma_\alpha^2 + (1 - \beta_0) \sigma^2}$$

Thus, if $\sigma_\alpha^2 \gg \sigma^2$ or $\beta_0 \approx 1$, the coefficient is close to zero, which implies that Y_{i0} is a weak instrument.

4. (a) In our simulation settings ($\sigma_\alpha^2 = \sigma^2 = 1$), we can speculate the followings

- i. $\hat{\beta}_{n,POLS} - \beta_0 \xrightarrow{p} \frac{(1-\beta_0)(1+\beta_0)\sigma_\alpha^2}{(1+\beta_0)\sigma_\alpha^2 + (1-\beta_0)\sigma^2} = \frac{1-\beta_0^2}{2}$ (omitted variable bias)
- ii. $\hat{\beta}_{n,FE} - \beta_0 \xrightarrow{p} \frac{1 + \beta_0}{\frac{2\beta_0}{1-\beta_0} - \frac{T-1}{1-\beta_0^{T-1}}}$ (Nickell, 1981)
- iii. $\hat{\beta}_{n,FD} - \beta_0 \xrightarrow{p} -\frac{1 + \beta_0}{2}$ (Q3)
- iv. $\hat{\beta}_{n,AH} - \beta_0 \xrightarrow{p} 0$ (Anderson and Hsiao, 1982), but this will suffer from weak instruments (Blundell and Bond, 1998). The first stage estimation coefficient (if we do 2SLS) is $\frac{(1-\beta_0)^2}{2}$.

The simulation results are presented in Figure 1. First, the bias of $\hat{\beta}_{n,POLS}$ decreased as β_0 approaches 1 as expected. The bias of $\hat{\beta}_{n,FD}$ decreased until $\beta_0 = 0.75$. $\hat{\beta}_{n,FDIV}$ gave an unbiased result, but gave a large standard error as $\beta_0 \rightarrow 1$.

(b) The simulation results are presented in Figure 2 and 3. When T was short ($T = 3$), $\hat{\beta}_{n,FDIV}$ gave unstable results. It also gave a large variance regardlessly of T if β_0 is close to 1.

5.

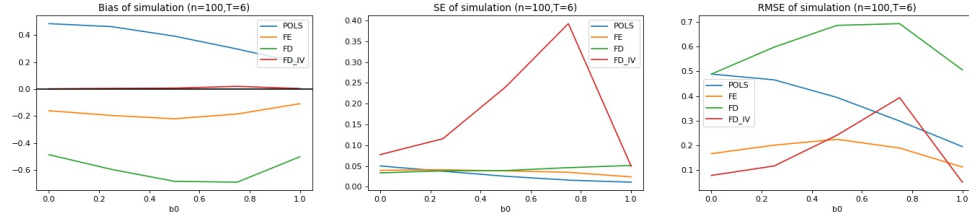


Figure 1: The simulation results of $n=100$, $T=6$

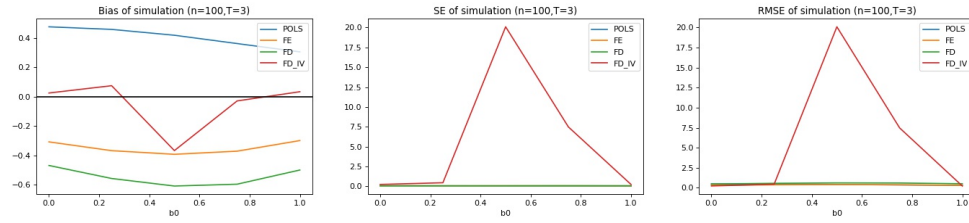


Figure 2: The simulation results of $n=100$, $T=3$

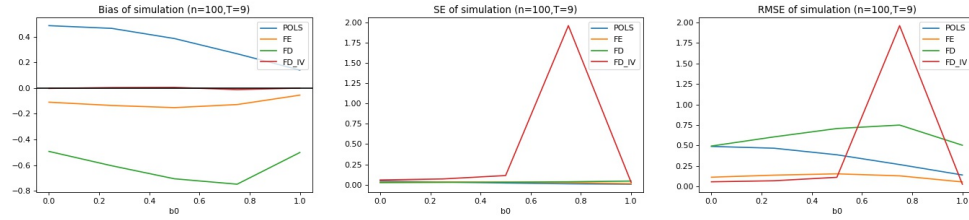


Figure 3: The simulation results of $n=100$, $T=9$