Introduction to Dynamic Panel Data: Autoregressive Models with Fixed Effects

Eric Zivot

Winter, 2013

Dynamic Panel Data (without covariates)

$$\begin{aligned} y_{im} &= \rho y_{i,m-1} + \alpha_i + \eta_{im} \\ i &= 1, \ldots, n \text{ (individuals)} \\ m &= 1, \ldots, M \text{ (time periods)} \end{aligned}$$

Typical assumptions

1. Stationarity: $|\rho| < 1$

2. $E[\eta_{im} | y_{i0}, \ldots, y_{i,m-1}, \alpha_i] = 0$

3a. No serial correlation: $\eta_{im} \sim iid\ (0, \sigma_\eta^2)$

3b. Homoskedasticity: $\alpha_i \sim iid\ (0, \sigma_\alpha^2)$

Stationary Model Representation

$$y_{im} = \rho y_{i,m-1} + \alpha_i + \eta_{im}$$

By recursive substitution

$$\begin{aligned}
y_{i1} &= \rho y_{i0} + \alpha_i + \eta_{i1} \\
y_{i2} &= \rho y_{i1} + \alpha_i + \eta_{i2} \\
&= \rho[\rho y_{i0} + \alpha_i + \eta_{i1}] + \alpha_i + \eta_{i2} \\
&= \rho^2 y_{i0} + (1+\rho)\alpha_i + \rho\eta_{i1} + \eta_{i2} \\
&\vdots \\
y_{im} &= \rho^m y_{i0} + \alpha_i \left(\sum_{s=0}^{m-1} \rho^s \right) + \sum_{s=0}^{m-1} \rho^s \eta_{i,m-s}
\end{aligned}$$

Now

$$
\begin{aligned}
E[y_{im}|\alpha_i] &= \rho^m E[y_{i0}|\alpha_i] + \alpha_i \left(\sum_{s=0}^{m-1} \rho^s \right) \\
&\quad + \sum_{s=0}^{m-1} \rho^s E[\eta_{i,m-s}|\alpha_i] \\
&= \rho^m E[y_{i0}|\alpha_i] + \alpha_i \left(\sum_{s=0}^{m-1} \rho^s \right)
\end{aligned}
$$

For large m

$$
\begin{aligned}
\rho^m &\approx 0 \\
\sum_{s=0}^{m-1} \rho^s &\approx \frac{1}{1-\rho}
\end{aligned}
$$

so that (for large m)

$$
\begin{aligned}
E[y_{im}|\alpha_i] &\approx \frac{\alpha_i}{1-\rho} \\
\text{var}[y_{im}|\alpha_i] &\approx \sum_{s=0}^{\infty} \text{var}[\rho^s \eta_{i,m-s}|\alpha_i] \\
&= \frac{\sigma_\eta^2}{1-\rho^2}
\end{aligned}
$$

Estimation

$$\begin{aligned} y_{im} &= \rho y_{i,m-1} + \alpha_i + \eta_{im} \\ &= \rho^m y_{i0} + \alpha_i \left(\sum_{s=0}^{m-1} \rho^s \right) + \sum_{s=0}^{m-1} \rho^s \eta_{i,m-s} \end{aligned}$$

Can't use RE estimation because

$$\begin{aligned} & E[y_{i,m-1} \cdot \alpha_i] \\ = \; & E\left[\left(\rho^{m-1} y_{i0} + \alpha_i \left(\sum_{s=0}^{m-2} \rho^s \right) + \sum_{s=0}^{m-2} \rho^s \eta_{i,m-s} \right) \alpha_i \right] \neq 0 \end{aligned}$$

What about FE estimation?

Write the model in FE matrix notation as

$$\begin{aligned}
\mathbf{y}_i &= \rho \mathbf{y}_{i,-1} + \alpha_i \mathbf{1}_M + \boldsymbol{\eta}_i \\
\underset{M\times 1}{\mathbf{y}_i} &= \begin{bmatrix} y_{i1} \\ \vdots \\ y_{iM} \end{bmatrix}, \ \underset{M\times 1}{\mathbf{y}_{i,-1}} = \begin{bmatrix} y_{i0} \\ \vdots \\ y_{i,M-1} \end{bmatrix}, \ \boldsymbol{\eta}_i = \begin{bmatrix} \eta_{i1} \\ \vdots \\ \eta_{iM} \end{bmatrix}
\end{aligned}$$

Define

$$\mathbf{Q}_M = \mathbf{I}_M - \mathbf{P}_{\mathbf{1}_M}$$

so that

$$\mathbf{Q}_M \mathbf{y}_i = \tilde{\mathbf{y}}_i = \begin{pmatrix} y_{i1} - \bar{y} \\ \vdots \\ y_{iM} - \bar{y} \end{pmatrix}, \ \bar{y} = \frac{1}{M} \sum_{m=1}^{M} y_{im}$$

Then, the transformed model is

$$\begin{aligned}
\mathbf{Q}_M \mathbf{y}_i &= \rho \mathbf{Q}_M \mathbf{y}_{i,-1} + \alpha_i \mathbf{Q}_M \mathbf{1}_M + \mathbf{Q}_M \boldsymbol{\eta}_i \Rightarrow \\
\tilde{\mathbf{y}}_i &= \rho \tilde{\mathbf{y}}_{i,-1} + \tilde{\boldsymbol{\eta}}_i
\end{aligned}$$

The FE estimator of ρ is the pooled OLS estimator on the transformed model

$$\begin{aligned}
\hat{\rho}_{FE} &= \left(\sum_{i=1}^{n} \tilde{\mathbf{y}}'_{i,-1} \tilde{\mathbf{y}}_{i,-1} \right)^{-1} \sum_{i=1}^{n} \tilde{\mathbf{y}}'_{i,-1} \tilde{\mathbf{y}}_i \\
&= \left(\sum_{i=1}^{n} \mathbf{y}'_{i,-1} \mathbf{Q}_M \mathbf{y}_{i,-1} \right)^{-1} \sum_{i=1}^{n} \mathbf{y}'_{i,-1} \mathbf{Q}_M \mathbf{y}_i
\end{aligned}$$

For consistency, consider

$$\hat{\rho}_{FE} - \rho = \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{y}'_{i,-1} \mathbf{Q}_M \mathbf{y}_{i,-1} \right)^{-1} \frac{1}{n} \sum_{i=1}^{n} \mathbf{y}'_{i,-1} \mathbf{Q}_M \boldsymbol{\eta}_i$$

$\hat{\rho}_{FE}$ will be consistent if

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathbf{y}'_{i,-1} \mathbf{Q}_M \boldsymbol{\eta}_i = E[\mathbf{y}'_{i,-1} \mathbf{Q}_M \boldsymbol{\eta}_i] = 0$$

Note: In the asymptotic analysis, the cross section dimension $n \to \infty$, but the time series dimension M is held fixed!

Now

$$
\begin{aligned}
\mathbf{y}_{i,-1}'\mathbf{Q}_M\boldsymbol{\eta}_i &= \left[y_{i0},\ldots,y_{i,M-1}\right]\begin{bmatrix}\eta_{i1}-\bar{\eta}_i\\ \vdots \\ \eta_{iM}-\bar{\eta}_i\end{bmatrix}\\
&= \sum_{m=1}^{M} y_{i,m-1}(\eta_{im}-\bar{\eta}_i)\\
\bar{\eta}_i &= \frac{1}{M}\sum_{m=1}^{M}\eta_{im}\\
y_{i,m-1} &= \rho^{m-1}y_{i0}+\alpha_i\left(\sum_{s=0}^{m-2}\rho^s\right)+\sum_{s=0}^{m-2}\rho^s\eta_{i,m-s}
\end{aligned}
$$

Using

$$y_{i,m-1} = \rho^{m-1} y_{i0} + \alpha_i \left(\sum_{s=0}^{m-2} \rho^s \right) + \sum_{s=0}^{m-2} \rho^s \eta_{i,m-s}$$

it follows that

$$E[\mathbf{y}'_{i,-1} \mathbf{Q}_M \boldsymbol{\eta}_i] = \sum_{m=1}^{M} E\left[y_{i,m-1} (\eta_{im} - \bar{\eta}_i) \right] \neq 0$$

because

$$E[y_{i,m-1} \bar{\eta}_i] = E\left[y_{i,m-1} \left(\frac{1}{M} \sum_{m=1}^{M} \eta_{im} \right) \right] \neq 0$$

Result:

$$\hat{\rho}_{FE} \overset{p}{\nrightarrow} \rho \text{ as } n \longrightarrow \infty \text{ for fixed } M$$

due to correlation between $y_{i,m-1}$ and $\bar{\eta}_i$.

Remark:

If both $n \longrightarrow \infty$ and $M \longrightarrow \infty$ then it can be shown that

$$\hat{\rho}_{FE} \overset{p}{\rightarrow} \rho$$

because

$$\bar{\eta}_i = \frac{1}{M} \sum_{m=1}^{M} \eta_{im} \overset{p}{\rightarrow} 0 \Rightarrow E[y_{i,m-1}\bar{\eta}_i] = 0$$

This is an example of "double asymptotic analysis" that is common in the analysis of panel data models.

Bias of Fixed Effects Estimator

In the stationary model with $\eta_{im} \sim$ iid $(0, \sigma_\eta^2)$, Nickell (1981, Ecta) showed that for fixed n and M

$$\begin{aligned} E[\mathbf{y}'_{i,-1}\mathbf{Q}_M \boldsymbol{\eta}_i] &= -\sigma_\eta^2 h_M(\rho) \\ h_M(\rho) &= \frac{1}{1-\rho}\left[1 - \frac{1}{M}\left(\frac{1-\rho^M}{1-\rho}\right)\right] \\ &\Rightarrow \hat{\rho}_{FE} \text{ is downward biased} \end{aligned}$$

Further, Nickell showed that for fixed M as $n \to \infty$

$$\hat{\rho}_{FE} - \rho \xrightarrow{p} -\frac{(1-\rho^2)h_M}{M-1}\left(1 - \frac{2\rho h_M(\rho)}{M-1}\right)$$

Notice that as $n \to \infty$ and $M \to \infty$ sequentially

$$\hat{\rho}_{FE} - \rho \xrightarrow{p} 0$$

Example of Bias in FE estimation of dynamic panel model

FE bias: $\hat{\rho}_{FE} - \rho$			
M/ρ	.05	.5	.95
2	-0.52	-0.75	-0.97
3	-0.35	-0.54	-0.73
10	-0.11	-0.16	-0.26
15	-0.07	-0.11	-0.17

Remarks

1. If $\rho > 0$ the bias is always negative, and massive for very small values of M

2. As M increases, the bias decreases but even with $M = 15$ the bias is still substantial for large ρ.

IV Estimation

Anderson and Hsiao (1981) suggested the following approach

1. First eliminate the fixed effect by taking first differences

$$\begin{aligned} y_{im} - y_{i,m-1} &= \rho(y_{i,m-1} - y_{i,m-2}) + \eta_{im} - \eta_{i,m-1} \\ \Delta y_{im} &= \rho \Delta y_{i,m-1} + \Delta \eta_{im}, \ m = 2, \ldots, M \end{aligned}$$

Note

$$E[\Delta y_{i,m-1} \Delta \eta_{im}] \neq 0$$

due to correlation between $y_{i,m-1}$ and $\eta_{i,m-1}$ terms.

Stacking across time gives

$$\underset{(M-1)\times 1}{\Delta \mathbf{y}_i} = \rho \Delta \mathbf{y}_{i,-1} + \Delta \boldsymbol{\eta}_i$$

2. Do IV estimation using $y_{i,m-2}$ as an instrument for $\Delta y_{i,m-1}$

$$\begin{aligned} E[y_{i,m-2}\Delta y_{i,m-1}] &= E[y_{i,m-2}(y_{i,m-1} - y_{i,m-2})] \neq 0 \\ E[y_{i,m-2}\Delta \eta_{i,m}] &= E[y_{i,m-2}\left(\eta_{im} - \eta_{i,m-1}\right)] = 0 \end{aligned}$$

since η_{im} is iid.

The IV estimator is then

$$\begin{aligned}
\hat{\rho}_{IV} &= \left(\sum_{i=1}^{n} \mathbf{y}'_{i,-2}\boldsymbol{\Delta}\mathbf{y}_{i,-1}\right)^{-1} \sum_{i=1}^{n} \mathbf{y}'_{i,-2}\boldsymbol{\Delta}\mathbf{y}_{i} \\
&= \frac{\sum_{i=1}^{n}\sum_{m=2}^{M} y_{i,m-2}\Delta y_{i,m}}{\sum_{i=1}^{n}\sum_{m=2}^{M} y_{i,m-2}\Delta y_{i,m-1}}
\end{aligned}$$

which is consistent by construction.

Remark:

The Anderson-Hsiao estimator does not exploit all the relevant moment conditions so it is not the most efficient GMM estimator.

Arellano-Bond GMM Estimator

"Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, 58, 1991

Arellano and Bond (AB) derived all of the relevant moment conditions from the dynamic panel data model to be used in GMM estimation.

The moment condtions are based on the first differenced model

$$\Delta y_{im} = \rho \Delta y_{i,m-1} + \Delta \eta_{im},\ m = 2, \ldots, M$$

They showed that the number of moment conditions depends on M (number of time periods)

Example: $M = 4 \Rightarrow 6$ moment condtions

$$
\begin{aligned}
m &= 2: E[\Delta\eta_{i2}y_{i0}] = 0 \\
m &= 3: \begin{array}{c} E[\Delta\eta_{i3}y_{i0}] = 0 \\ E[\Delta\eta_{i3}y_{i1}] = 0 \end{array} \\
m &= 4: \begin{array}{c} E[\Delta\eta_{i4}y_{i0}] = 0 \\ E[\Delta\eta_{i4}y_{i1}] = 0 \\ E[\Delta\eta_{i4}y_{i2}] = 0 \end{array}
\end{aligned}
$$

For GMM estimation, define

$$
\mathbf{g}_i^4(\rho) = \begin{pmatrix} \Delta\eta_{i2}y_{i0} \\ \Delta\eta_{i3}y_{i0} \\ \Delta\eta_{i3}y_{i1} \\ \Delta\eta_{i4}y_{i0} \\ \Delta\eta_{i4}y_{i1} \\ \Delta\eta_{i4}y_{i2} \end{pmatrix} = \begin{pmatrix} (\Delta y_{i2} - \rho\Delta y_{i1})\, y_{i0} \\ (\Delta y_{i3} - \rho\Delta y_{i3})\, y_{i0} \\ (\Delta y_{i3} - \rho\Delta y_{i2})\, y_{i1} \\ (\Delta y_{i4} - \rho\Delta y_{i3})\, y_{i0} \\ (\Delta y_{i4} - \rho\Delta y_{i3})\, y_{i1} \\ (\Delta y_{i4} - \rho\Delta y_{i3})\, y_{i2} \end{pmatrix}
$$

Notice that $\mathbf{g}_i^4(\rho)$ is a linear function of ρ. It may be re-expressed in matrix form as

$$\begin{aligned}
\mathbf{g}_i^4(\rho) &= \begin{bmatrix} y_{i0} & 0 & 0 \\ 0 & y_{i0} & 0 \\ 0 & y_{i1} & 0 \\ 0 & 0 & y_{i0} \\ 0 & 0 & y_{i1} \\ 0 & 0 & y_{i2} \end{bmatrix} \begin{bmatrix} \Delta y_{i2} - \rho \Delta y_{i1} \\ \Delta y_{i3} - \rho \Delta y_{i3} \\ \Delta y_{i4} - \rho \Delta y_{i3} \end{bmatrix} \\
&= \mathbf{X}_i^{4\prime}[\Delta \mathbf{y}_i^4 - \rho \Delta \mathbf{y}_{i,-1}^4]
\end{aligned}$$

where

$$\Delta \mathbf{y}_i^4 = \begin{bmatrix} \Delta y_{i2} \\ \Delta y_{i3} \\ \Delta y_{i4} \end{bmatrix}, \ \Delta \mathbf{y}_{i,-1}^4 = \begin{bmatrix} \Delta y_{i1} \\ \Delta y_{i2} \\ \Delta y_{i3} \end{bmatrix}$$

$$\mathbf{X}_i^4 = \begin{bmatrix} y_{i0} & 0 & 0 & 0 & 0 & 0 \\ 0 & y_{i0} & y_{i1} & 0 & 0 & 0 \\ 0 & 0 & 0 & y_{i0} & y_{i1} & y_{i2} \end{bmatrix}$$

The sample moments used for GMM estimation are then

$$\begin{aligned}
\mathbf{g}_n^4(\rho) &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i^{4\prime}[\Delta\mathbf{y}_i^4 - \rho\Delta\mathbf{y}_{i,-1}^4] \\
&= \mathbf{S}_{x\Delta y}^4 - \mathbf{S}_{x\Delta y_{-1}}^4\rho, \\
\mathbf{S}_{x\Delta y}^4 &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i^{4\prime}\Delta\mathbf{y}_i^4, \\
\mathbf{S}_{x\Delta y_{-1}}^4 &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i^{4\prime}\Delta\mathbf{y}_{i,-1}^4
\end{aligned}$$

Because η_{im} is randomly sampled over i it follows that

$$\underset{6\times 6}{\mathbf{S}} = E[\mathbf{g}_i^4(\rho)\mathbf{g}_i^4(\rho)'] = E[\mathbf{X}_i^{4\prime}\Delta\boldsymbol{\eta}_i\Delta\boldsymbol{\eta}_i'\mathbf{X}_i^4]$$

Under conditional heteroskedasticity, a consistent estimate is

$$\begin{aligned}\hat{\mathbf{S}} &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i^{4\prime}\Delta\hat{\boldsymbol{\eta}}_i\Delta\hat{\boldsymbol{\eta}}_i'\mathbf{X}_i^4 \\ \Delta\hat{\boldsymbol{\eta}}_i &= [\Delta\mathbf{y}_i^4 - \hat{\rho}\Delta\mathbf{y}_{i,-1}^4] \\ & \hat{\rho} \xrightarrow{p} \rho\end{aligned}$$

For example, can use the Anderson and Hsiao IV estimate of ρ.

Under conditional homoskedasticity

$$\begin{aligned}\mathbf{S} &= E[\mathbf{X}_i^{4\prime}\boldsymbol{\Delta\eta}_i\boldsymbol{\Delta\eta}_i'\mathbf{X}_i^4] = E[E[\mathbf{X}_i^{4\prime}\mathbf{C}'\boldsymbol{\eta}_i\boldsymbol{\eta}_i'\mathbf{C}\mathbf{X}_i^4|\mathbf{X}_i^4]] = \\ &= E[\mathbf{X}_i^{4\prime}\mathbf{C}'E[\boldsymbol{\eta}_i\boldsymbol{\eta}_i'|\mathbf{X}_i^4]\mathbf{C}\mathbf{X}_i^4] = \sigma_\eta^2 E[\mathbf{X}_i^{4\prime}\mathbf{C}'\mathbf{C}\mathbf{X}_i^4]\end{aligned}$$

where

$$\underset{(M-1)\times M}{\mathbf{C}'} = \begin{pmatrix} -1 & 1 & 0 & & 0 \\ 0 & -1 & 1 & & 0 \\ & & & & \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

So a consistent estimate of $\mathbf{S}$ is

$$\hat{\mathbf{S}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i^{4\prime}\mathbf{C}'\mathbf{C}\mathbf{X}_i^4$$

Estimation of σ_η^2 is not required for GMM estimation because it cancels out in the resulting estimator.

The efficient GMM estimator solves

$$\min_{\rho} J(\rho, \hat{\mathbf{S}}^{-1}) = n\mathbf{g}_n^4(\rho)\hat{\mathbf{S}}^{-1}\mathbf{g}_n^4(\rho)$$
$$= n[\mathbf{S}^4_{x\Delta y} - \mathbf{S}^4_{x\Delta y_{-1}}\rho]'\hat{\mathbf{S}}^{-1}[\mathbf{S}^4_{x\Delta y} - \mathbf{S}^4_{x\Delta y_{-1}}\rho]$$

Since $J(\rho, \hat{\mathbf{S}}^{-1})$ is linear in ρ, the analytic solution is

$$\hat{\rho}_{AB}(\hat{\mathbf{S}}^{-1}) = (\mathbf{S}^{4\prime}_{x\Delta y_{-1}}\hat{\mathbf{S}}^{-1}\mathbf{S}^4_{x\Delta y_{-1}})\mathbf{S}^{4\prime}_{x\Delta y_{-1}}\hat{\mathbf{S}}^{-1}\mathbf{S}^4_{x\Delta y}$$

This estimator is known as the Arellano-Bond GMM estimator.

Example: International Difference in Output Growth Rates (Hayashi, Section 5.4)

Q: Do poor countries grow faster than rich countries? If so how much faster?

Neoclassical growth theory foundations

$$\begin{aligned} q(t) &= \text{output per effective labor at time } t \text{ for a country} \\ &\rightarrow\ q^* = \text{steady state} \end{aligned}$$

The log-linear approximation around q^* gives the adjustment equation

$$\begin{aligned} \frac{d\,\ln(q(t))}{dt} &= \lambda \cdot [\ln(q^*) - \ln(q(t)] \\ \lambda &= \text{speed of convergence } > 0 \end{aligned}$$

Between any two time periods t_{m-1} and t_m, the log-linear adjustment equation implies

$$\begin{aligned} \ln(q(t_m)) &= (1-\rho)\cdot\ln(q^*) + \rho\ln(q(t_{m-1})) \\ \rho &= \exp\left[-\lambda\cdot(t_m - t_{m-1})\right] \end{aligned}$$

Define

$$\begin{aligned} q(t) &= \frac{Y(t)}{A(t)L(t)} \\ Y(t) &= \text{aggregate output} \\ L(t) &= \text{aggregate hours worked} \\ A(t) &= \text{level of labor augmenting technical progress} \end{aligned}$$

Assume $A(t)$ grows at a constant rate g. Then $A(t) = A(0)\exp(g \cdot t)$ and

$$\ln q(t) = \ln\left(\frac{Y(t)}{L(t)}\right) - \ln(A(0)) - g \cdot t$$

Substituting into the output equation gives

$$\begin{aligned} \ln\left(\frac{Y(t_m)}{L(t_m)}\right) &= \rho \ln\left(\frac{Y(t_{m-1})}{L(t_{m-1})}\right) + (1-\rho)\left[\ln(q^*) - \ln(A(0))\right] + \phi_m \\ \phi_m &= g \cdot (t_m - \rho \cdot t_{m-1}) \end{aligned}$$

Subtracting $\ln\left(\frac{Y(t_{m-1})}{L(t_{m-1})}\right)$ from both sides gives the growth equation

$$\Delta \ln\left(\frac{Y(t_m)}{L(t_m)}\right) = (\rho - 1)\ln\left(\frac{Y(t_{m-1})}{L(t_{m-1})}\right) + (1-\rho)\left[\ln(q^*) - \ln(A(0)\right] + \phi_m$$

Because $\rho < 1$, the level of per capita output has a negative effect on growth. Hence, poor countries should grow faster than rich countries.

Turning the output equation into a dynamic panel data model

Assume

$$\begin{aligned} \ln\left(\frac{Y(t_m)}{L(t_m)}\right) &= \rho \ln\left(\frac{Y(t_{m-1})}{L(t_{m-1})}\right) + (1-\rho)\left[\ln(q^*) - \ln(A(0)\right] + \phi_m \\ \phi_m &= g \cdot (t_m - \rho \cdot t_{m-1}) \end{aligned}$$

holds for every country i where ρ and g are the same for every country but that q^* and A might differ. Then we can write

$$\begin{aligned} y_{im} &= \phi_m + \rho y_{i,m-1} + \alpha_i + \eta_{im} \\ y_{im} &= \ln\left(Y(t_m)/L(t_m)\right) = \text{log per capita output at time } t_m \\ \alpha_i &= (1-\rho)\{\ln(q_i^*) - \ln(A_i(0))\} \text{ for country } i \\ \eta_{im} &= \text{country and time specific shock (e.g. business cycle)} \end{aligned}$$