Dynamic Heterogeneous Panels

Contents

1 Asymptotic property of LS and IV estimator

1.1 Brief the source of bias

Consider the dynamic heterogeneous panels data model:

$$y_{i,t} = \alpha_i + \phi_i y_{i,t-1} + \beta_i x_{i,t} + u_{i,t},\ \ for\ i = 1, \ldots N;\ t = 1, \ldots, T\,, \tag{1}$$

In above model, we assume the number of regressor is one and the model can rewritten as

$$\Delta y_{i,t} = \alpha_i + -(1-\phi_i)(y_{i,t} - \pi_i x_{i,t}) + \varepsilon_{i,t}, \tag{2}$$

where $\pi_i = \frac{\beta_i}{1-\phi_i}$. And we defined $\theta_i = (1-\phi_i)$. Then we suppose

$$\begin{aligned} \theta_i &= \theta + \eta_{i1}, \\ \pi_i &= \pi + \eta_{i3}. \end{aligned} \tag{3}$$

Therefore, we know

$$\beta_i = \pi_i\theta_i = (\pi + \eta_{i3})(\theta + \eta_{i1}) = \pi\theta + \pi\eta_{i,1} + \theta\eta_{i3} + \eta_{i1}\eta_{i3} \tag{4}$$

And we defined $\eta_{i2} = \pi\eta_{i,1} + \theta\eta_{i3} + \eta_{i1}\eta_{i3}$. Then,we know $\beta_i = \pi\theta + \eta_{i2}$. Therefore, from equation (1), we have

$$\begin{aligned} y_{i,t} &= \alpha_i + \phi_i y_{i,t-1} + (\pi\theta + \pi\eta_{i,1} + \theta\eta_{i3} + \eta_{i1}\eta_{i3})\, x_{i,t} + u_{i,t} \\ &= \alpha_i + (1-\theta_i)\, y_{i,t-1} + \beta x_{i,t} + \eta_{i,2}x_{i,t} + u_{i,t} \\ &= \alpha_i + (1-\theta)y_{i,t-1} - \eta_{i1}y_{i,t-1} + \beta x_{i,t} + \eta_{i2}x_{i,t} + u_{i,t} \\ &= \alpha_i + \phi y_{i,t-1} + \beta x_{i,t} + (u_{i,t} - \eta_{i1}y_{i,t-1} + \eta_{i2}x_{i,t}) \\ &= \alpha_i + \phi y_{i,t-1} + \beta x_{i,t} + v_{i,t}, \end{aligned} \tag{5}$$

where $v_{i,t} = (\varepsilon_{i,t} - \eta_{1i}y_{i,t-1} + \eta_{2i}x_{i,t})$. Then,we can see that $y_{i,t-1}$ and $x_{i,t}$ are correlated with $v_{i,t}$.

1.2 Asymptotic property of LS estimator

We define our interested parameter as

$$(\phi_i, \beta_i)' = \boldsymbol{\theta}_i = \boldsymbol{\theta} + \boldsymbol{\lambda}_i, \tag{6}$$

where $\boldsymbol{\lambda}_i \overset{i.i.d.}{\sim} (\mathbf{0}, \boldsymbol{\Sigma}_\lambda)$.

Based on heterogenous dynamic panel data model (1), we can obtain fixed effect estimator as

$$\hat{\boldsymbol{\theta}}_{LS,i} = \begin{pmatrix} \hat{\phi}_i \\ \hat{\beta}_i \end{pmatrix} = \begin{pmatrix} \frac{\sum_{t=1}^T \tilde{y}_{i,t-1}^2}{T} & \frac{\sum_{t=1}^T \tilde{y}_{i,t-1}\tilde{x}_{i,t}}{T} \\ \frac{\sum_{t=1}^T \tilde{x}_{i,t}\tilde{y}_{i,t-1}}{T} & \frac{\sum_{t=1}^T \tilde{x}_{i,t}^2}{T} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\sum_{t=1}^T \tilde{y}_{i,t-1}\tilde{y}_{i,t}}{T} \\ \frac{\sum_{t=1}^T \tilde{x}_{i,t}\tilde{y}_{i,t}}{T} \end{pmatrix}, \tag{7}$$

where $\tilde{y}_{i,t} = y_{i,t} - \bar{y}_i$, $\tilde{y}_{i,t-1} = y_{i,t-1} - \bar{y}_{i,-1}\iota_T$ and $\tilde{x}_{i,t} = x_{i,t} - \bar{x}_i$ with $\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{i,t}$, $\bar{y}_{i,-1} = \frac{1}{T}\sum_{t=1}^{T} y_{i,t-1}$, $\bar{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{i,t}$. Under equation (1), we have

$$\begin{pmatrix} \hat{\phi}_i - \phi_i \\ \hat{\beta}_i - \beta_i \end{pmatrix} = \begin{pmatrix} \frac{\sum_{t=1}^{T} \tilde{y}_{i,t-1}^2}{T} & \frac{\sum_{t=1}^{T} \tilde{y}_{i,t-1}\tilde{x}_{i,t}}{T} \\ \frac{\sum_{t=1}^{T} \tilde{x}_{i,t}\tilde{y}_{i,t-1}}{T} & \frac{\sum_{t=1}^{T} \tilde{x}_{i,t}^2}{T} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\sum_{t=1}^{T} \tilde{y}_{i,t-1}\tilde{u}_{i,t}}{T} \\ \frac{\sum_{t=1}^{T} \tilde{x}_{i,t}\tilde{u}_{i,t}}{T} \end{pmatrix}, \tag{8}$$

Now, we can investigate asymptotic bias by taking the probability limit as

$$A_{\phi i}^1 = \operatorname*{plim}_{T\to\infty} \left(\frac{\sum_{t=1}^{T} \tilde{y}_{i,t-1}\tilde{u}_{i,t}}{T} \right). \tag{9}$$

Then A_i can be taken expectations as

$$\begin{aligned} A_{\phi i}^1 &= E\left(y_{i,t-1} - \bar{y}_{i,-1}\right)\left(u_{i,t} - \bar{u}_i\right) \\ &= E\left(y_{i,t-1}u_{i,t}\right) - E\left(y_{i,t-1}\bar{u}_i\right) - E\left(\bar{y}_{i,-1}u_{i,t}\right) + E\left(\bar{y}_{i,-1}\bar{u}_i\right), \end{aligned} \tag{10}$$

where $E\left(y_{i,t-1}u_{i,t}\right) = 0$.

And we assume $y_{i,t}$ has started from a long time period in the past, so we have

$$y_{i,t} = \frac{\alpha_i}{(1-\phi_i)} + \sum_{s=0}^{\infty} \beta_i \phi_i^s x_{i,t-s} + \sum_{s=0}^{\infty} \phi_i^s u_{i,t-s}, \tag{11}$$

Then, we have

$$\begin{aligned} A_{\phi i}^1 = -E\left(\left(\sum_{s=0}^{\infty} \phi^s u_{i,t-s-1}\right)\left(\frac{1}{T}\sum_{t=1}^{T} u_{i,t}\right)\right) - E\left(\frac{u_{i,t}}{T}\sum_{t=1}^{T}\sum_{s=0}^{\infty} \phi^s u_{i,t-s-1}\right) + \\ \left(\frac{1}{T}\sum_{t=1}^{T}\sum_{s=0}^{\infty} \phi^s u_{i,t-s-1}\right)\left(\frac{1}{T}\sum_{t=1}^{T} u_{i,t}\right). \end{aligned} \tag{12}$$

Hence, from above equation, we have

$$\begin{aligned} A_{\phi i}^1 =& -\frac{1}{T}E\left(\left(u_{i,t-1} + u_{i,t-2}\phi^1 + u_{i,t-3}\phi^2 + \ldots\right)\left(u_{i,1} + \ldots + u_{i,t-1} + u_{i,t} + \ldots + u_{i,T}\right)\right) - \\ & \frac{1}{T}E\left(u_{i,t}\sum_{s=1}^{T}\left(u_{i,s-1}\phi^0 + u_{i,s-2}\phi^1 + \ldots + u_{i,s-t-1}\phi^t + \ldots + u_{i,s-T-1}\phi^T + \ldots\right)\right) + \\ & + \frac{1}{T}E(\left(\sum_{s=1}^{T} u_{i,s-1}\phi^0 + \sum_{s=1}^{T} u_{i,s-2}\phi^1 + \ldots + \sum_{s=1}^{T} u_{i,s-t-1}\phi^t + \ldots + \sum_{s=1}^{T} u_{is-T-1}\phi^T + \ldots\right) \\ & \left(\frac{1}{T}\sum_{s=1}^{T} u_{i,s}\right)) \\ =& -\frac{\sigma_u^2}{T}\frac{(1-\phi^{t-1})}{1-\phi} - \frac{\sigma_u^2}{T}\frac{(1-\phi^{T-t})}{(1-\phi)} + \frac{\sigma_u^2}{T}\left(\frac{1}{1-\phi} - \frac{1}{T}\frac{(1-\phi^T)}{(1-\phi)^2}\right) \\ =& -\frac{\sigma_u^2}{T(1-\phi)}\left(1 - \phi^{t-1} - \phi^{T-t} + \frac{1}{T}\frac{(1-\phi^T)}{(1-\phi)}\right). \end{aligned} \tag{13}$$

Therefore, we can see the bias of $\hat{\phi}_i$ is $O\left(T^{-1}\right)$.

Now, we try to derive the asymptotic variance of $\hat{\boldsymbol{\theta}}_{LS,i}$.

$$\boldsymbol{\Sigma}_{LS,i} = \operatorname*{plim}_{T\to\infty} \boldsymbol{u}_i \boldsymbol{u}_i^{'} \tag{14}$$

Then, we have

$$\operatorname*{plim}_{T\to\infty} Var\left(\hat{\boldsymbol{\theta}}_{LS,i}\right) = \left(\tilde{\boldsymbol{W}}_i^{'}\tilde{\boldsymbol{W}}_i\right)^{-1} \tilde{\boldsymbol{W}}_i^{'}\boldsymbol{\Sigma}_{LS,i}\tilde{\boldsymbol{W}}_i \left(\tilde{\boldsymbol{W}}_i^{'}\tilde{\boldsymbol{W}}_i\right)^{-1}, \tag{15}$$

where $\tilde{\boldsymbol{W}}_i = \left(\tilde{\boldsymbol{w}}_{i,1}^{'}, \ldots, \tilde{\boldsymbol{w}}_{i,T}^{'}\right)^{'}$ is $T \times 2$ matrix and $\boldsymbol{P}_i = (p_{i,1}, \ldots, p_{i,T})\, \boldsymbol{I}_T$ is $T \times T$ matrix with $\boldsymbol{I}_T$ is $T \times T$ identity matrix, $\tilde{\boldsymbol{w}}_{i,t} = (y_{i,t-1} - \bar{y}_{i,-1}, x_{i,t} - \bar{x}_i)$, for $t = 1, \ldots, T$ and $i = 1, \ldots, N$..

Now, we define the mean group estimator of $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{\theta}}_{LSMG} = \frac{1}{N}\sum_{i=1}^{N} \hat{\boldsymbol{\theta}}_{LSi}. \tag{16}$$

And we can show that the asymptotic property of $\hat{\boldsymbol{\theta}}_{IVMG}$, as

$$\sqrt{N}\left(\hat{\boldsymbol{\theta}}_{LSMG} - \boldsymbol{\theta}\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} \boldsymbol{\lambda}_i + O\left(T^{-1}\right) \tag{17}$$

And we can see that

$$\sqrt{N}\left(\hat{\boldsymbol{\theta}}_{LSMG} - \boldsymbol{\theta}\right) \xrightarrow{d} N\left(\mathbf{0}, \boldsymbol{\Sigma}_{LS,\lambda}\right) \tag{18}$$

1.3 Asymptotic property of IV estimator

We use current and lagged values of $\boldsymbol{x}_i$ as instruments, as

$$\boldsymbol{z}_{i,t} = \left(x_{i,t}, x_{i,t-1}\right), \tag{19}$$

where $\boldsymbol{z}_{i,t}$ is 1×2 vector. And we define $\boldsymbol{w}_{i,t} = (\tilde{y}_{i,t-1}, \tilde{x}_{i,t})$. Then the IV estimator can be expressed as

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{IVi} = \begin{pmatrix} \phi_i \\ \beta_i \end{pmatrix} &= \left(\left(\frac{\sum_{t=1}^{T} \boldsymbol{z}_{i,t}^{'}\boldsymbol{w}_{i,t}}{T}\right)^{'} \left(\frac{\sum_{t=1}^{T} \boldsymbol{z}_{i,t}^{'}\boldsymbol{z}_{i,t}}{T}\right)^{-1} \left(\frac{\sum_{t=1}^{T} \boldsymbol{z}_{i,t}^{'}\boldsymbol{w}_{i,t}}{T}\right)\right)^{-1} \times \\ &\quad \left(\left(\frac{\sum_{t=1}^{T} \boldsymbol{z}_{i,t}^{'}\boldsymbol{w}_{i,t}}{T}\right)^{'} \left(\frac{\sum_{t=1}^{T} \boldsymbol{z}_{i,t}^{'}\boldsymbol{z}_{i,t}}{T}\right)^{-1} \left(\frac{\sum_{t=1}^{T} \boldsymbol{z}_{i,t}^{'}\boldsymbol{y}_{i,t}}{T}\right)\right) \\ &= \left(\frac{\sum_{t=1}^{T} \boldsymbol{w}_{i,t}^{'}p_{i,t}\boldsymbol{w}_{i,t}}{T}\right)^{-1} \left(\frac{\sum_{t=1}^{T} \boldsymbol{w}_{i,t}^{'}p_{i,t}\boldsymbol{y}_{i,t}}{T}\right), \end{aligned} \tag{20}$$

where $p_{i,t} = \boldsymbol{z}_{i,t}\left(\boldsymbol{z}_{i,t}^{'}\boldsymbol{z}_{i,t}\right)^{-1}\boldsymbol{z}_{i,t}^{'}$.

Under (5), we have

$$\begin{pmatrix} \hat{\phi}_i - \phi_i \\ \hat{\beta}_i - \beta_i \end{pmatrix} = \begin{pmatrix} \frac{\sum_{t=1}^{T} \tilde{y}_{i,t-1} p_{i,t} \tilde{y}_{i,t-1}}{T} & \frac{\sum_{t=1}^{T} \tilde{y}_{i,t-1} p_{i,t} \tilde{x}_{i,t}}{T} \\ \frac{\sum_{t=1}^{T} \tilde{x}_{i,t} p_{i,t} \tilde{y}_{i,t-1}}{T} & \frac{\sum_{t=1}^{T} \tilde{x}_{i,t} p_{i,t} \tilde{x}_{i,t}}{T} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\sum_{t=1}^{T} \tilde{y}_{i,t-1} p_{i,t} \tilde{u}_{i,t}}{T} \\ \frac{\sum_{t=1}^{T} \tilde{x}_{i,t} p_{i,t} \tilde{u}_{i,t}}{T} \end{pmatrix}, \tag{21}$$

We can see the asymptotic bias by taking the probability limit as

$$A_{\phi i}^2 = \operatorname*{plim}_{T\to\infty} \left(\frac{\sum_{t=1}^{T} \left(y_{i,t-1} - \bar{y}_i\right) p_{i,t} \left(u_{i,t-1} - \bar{u}_i\right)}{T} \right) \tag{22}$$

Again, we have

$$\begin{aligned} A_{\phi i}^2 &= E\left(y_{i,t-1} - \bar{y}_{i,-1}\right) p_{i,t} \left(u_{i,t} - \bar{u}_i\right) \\ &= E\left(y_{i,t-1} p_{i,t} u_{i,t}\right) - E\left(y_{i,t-1} p_{i,t} \bar{u}_i\right) - E\left(\bar{y}_{i,-1} p_{i,t} u_{i,t}\right) + E\left(\bar{y}_{i,-1} p_{i,t} \bar{u}_i\right), \end{aligned} \tag{23}$$

Also, we can show

$$E\left(y_{i,t-1} p_{i,t} u_{i,t}\right) = 0 \tag{24}$$

$$\begin{aligned} &E\left(y_{i,t-1} p_{i,t} \bar{u}_i\right) = E\left(\left(\sum_{s=0}^{\infty} \beta_i \phi_i^s x_{i,t-s-1} + \sum_{s=0}^{\infty} \phi_i^s u_{i,t-s-1}\right) p_{i,t} \left(\frac{1}{T} \sum_{t=1}^{T} u_{i,t}\right)\right) \\ &= E\left(\left(\sum_{s=0}^{\infty} \beta_i \phi_i^s x_{i,t-s-1} + \sum_{s=0}^{\infty} \phi_i^s u_{i,t-s-1}\right) \boldsymbol{z}_{i,t} \left(\boldsymbol{z}_{i,t}^{'} \boldsymbol{z}_{i,t}\right)^{-1} \boldsymbol{z}_{i,t}^{'} \left(\frac{1}{T} \sum_{t=1}^{T} u_{i,t}\right)\right) \\ &= 0 \end{aligned} \tag{25}$$

$$\begin{aligned} &E\left(\bar{y}_{i,-1} p_{i,t} u_{i,t}\right) = \frac{1}{T} E\left(\sum_{t=1}^{T} \left(\sum_{s=0}^{\infty} \beta_i \phi_i^s x_{i,t-s-1} + \sum_{s=0}^{\infty} \phi_i^s u_{i,t-s-1}\right) p_{i,t} u_{i,t}\right) \\ &= \frac{1}{T} E\left(\sum_{t=1}^{T} \left(\sum_{s=0}^{\infty} \beta_i \phi_i^s x_{i,t-s-1} + \sum_{s=0}^{\infty} \phi_i^s u_{i,t-s-1}\right) \boldsymbol{z}_{i,t} \left(\boldsymbol{z}_{i,t}^{'} \boldsymbol{z}_{i,t}\right)^{-1} \boldsymbol{z}_{i,t}^{'} u_{i,t}\right) \\ &= 0 \end{aligned} \tag{26}$$

$$\begin{aligned} &E\left(\bar{y}_{i,-1} p_{i,t} \bar{u}_i\right) = \frac{1}{T} E\left(\sum_{t=1}^{T} \left(\sum_{s=0}^{\infty} \beta_i \phi_i^s x_{i,t-s-1} + \sum_{s=0}^{\infty} \phi_i^s u_{i,t-s-1}\right) p_{i,t} \left(\frac{1}{T} \sum_{t=1}^{T} u_{i,t}\right)\right) \\ &= \frac{1}{T} E\left(\sum_{t=1}^{T} \left(\sum_{s=0}^{\infty} \beta_i \phi_i^s x_{i,t-s-1} + \sum_{s=0}^{\infty} \phi_i^s u_{i,t-s-1}\right) \boldsymbol{z}_{i,t} \left(\boldsymbol{z}_{i,t}^{'} \boldsymbol{z}_{i,t}\right)^{-1} \boldsymbol{z}_{i,t}^{'} \left(\frac{1}{T} \sum_{t=1}^{T} u_{i,t}\right)\right) \\ &= 0 \end{aligned} \tag{27}$$

Now, we try to derive the asymptotic variance of $\hat{\boldsymbol{\theta}}_{IV,i}$. We define

$$\boldsymbol{\Sigma}_{IV,i} = \operatorname*{plim}_{T\to\infty} \boldsymbol{u}_i \boldsymbol{u}_i^{'} \tag{28}$$

Then, we have

$$\operatorname*{plim}_{T\to\infty} Var\left(\hat{\boldsymbol{\theta}}_i\right) = \left(\boldsymbol{W}_i^{'}\boldsymbol{P}_i\boldsymbol{W}_i\right)^{-1}\boldsymbol{W}_i^{'}\boldsymbol{P}_i\boldsymbol{\Sigma}_{IV,i}\boldsymbol{P}_i\boldsymbol{W}_i\left(\boldsymbol{W}_i^{'}\boldsymbol{P}_i\boldsymbol{W}_i\right)^{-1}, \tag{29}$$

where $\boldsymbol{W}_i = \left(w_{i,1}^{'}, \ldots, w_{i,T}^{'}\right)^{'}$ is $T \times 2$ matrix and $\boldsymbol{P}_i = (p_{i,1}, \ldots, p_{i,T})\,\boldsymbol{I}_T$.

Now, we define the mean group estimator of $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{\theta}}_{IVMG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\theta}}_{IVi}. \tag{30}$$

And we can show that the asymptotic property of $\hat{\boldsymbol{\theta}}_{IVMG}$, as

$$\sqrt{N}\left(\hat{\boldsymbol{\theta}}_{IVMG} - \boldsymbol{\theta}\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\lambda}_i + o\,(1) \tag{31}$$

And we can see that

$$\sqrt{N}\left(\hat{\boldsymbol{\theta}}_{IVMG} - \boldsymbol{\theta}\right) \xrightarrow{d} N\left(\mathbf{0}, \boldsymbol{\Sigma}_{IV,\lambda}\right) \tag{32}$$

2 Estimation method on dynamic heterogeneous panel data model

For convenient, we assume the number of regressor is 1 and we express the model as

$$y_{i,t} = \phi_i y_{i,t-1} + \sum_{\ell=1}^{k}\beta_{\ell i}x_{\ell i,t} + u_{i,t},\;\; for\ i = 1,\ldots N;\ t = 1,\ldots,T,\ \ell = 1,\ldots,k. \tag{33}$$

We stack the T observations for each i yield

$$\boldsymbol{y}_i = \boldsymbol{y}_{i,-1}\phi_i + \sum_{\ell=1}^{k}\boldsymbol{x}_{\ell i}\boldsymbol{\beta}_{\ell i} + \boldsymbol{u}_i, \tag{34}$$

where $\boldsymbol{y}_i = (y_{i,1}, \ldots, y_{i,T})^{'}$, $\boldsymbol{y}_{i,-1} = (y_{i,0}, \ldots, y_{i,T-1})^{'}$, $\boldsymbol{x}_{\ell i} = (x_{\ell i,1}, \ldots, x_{\ell i,T})^{'}$ and $\boldsymbol{u}_i = (u_{i,1}, \ldots, u_{i,T})$. To be more compressive, the model can be expressed as

$$\boldsymbol{y}_i = \boldsymbol{W}_i\boldsymbol{\varphi}_i + \boldsymbol{u}_i, \tag{35}$$

where $\boldsymbol{W}_i = \left(\boldsymbol{y}_{i,-1}, \boldsymbol{X}_i\right)$ and $\boldsymbol{\varphi}_i = \left(\phi_i, \boldsymbol{\beta}_i^{'}\right)^{'}$

2.1 LSMG estimator

The LS (least square) estimator is defined as

$$\hat{\boldsymbol{\varphi}}_{LSi} = \left(\frac{\boldsymbol{W}_i^{'}\boldsymbol{W}_i}{T}\right)^{-1}\left(\frac{\boldsymbol{W}_i^{'}\boldsymbol{y}_i}{T}\right) \tag{36}$$

Follow Pesaran and Smith (1995), we define the LSMG (least square mean group) estimator as

$$\hat{\boldsymbol{\varphi}}_{LSMG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\varphi}}_{LSi}. \tag{37}$$

2.2 IVMG estimator

We use current and lagged values of $\boldsymbol{x}_i$ as instruments, as

$$\boldsymbol{Z}_i = (\boldsymbol{X}_i, \boldsymbol{X}_{i,-1}), \tag{38}$$

where $\boldsymbol{Z}_i$ is $T \times 2k$ matrix.

The IV (instrument variable) estimator is defined as

$$\hat{\boldsymbol{\varphi}}_{IVi} = \left(\left(\frac{\boldsymbol{Z}_i'\boldsymbol{W}_i}{T}\right)'\left(\frac{\boldsymbol{Z}_i'\boldsymbol{Z}_i}{T}\right)^{-1}\left(\frac{\boldsymbol{Z}_i'\boldsymbol{W}_i}{T}\right)\right)^{-1}\left(\left(\frac{\boldsymbol{Z}_i'\boldsymbol{W}_i}{T}\right)'\left(\frac{\boldsymbol{Z}_i'\boldsymbol{Z}_i}{T}\right)^{-1}\left(\frac{\boldsymbol{Z}_i'\boldsymbol{y}_i}{T}\right)\right) \tag{39}$$

We also define the IVMG (instrument variable mean group) estimator as

$$\hat{\boldsymbol{\varphi}}_{IVMG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\varphi}}_{IVi}. \tag{40}$$

3 Estimation method on dynamic heterogeneous panel data model with multifactor error structure

Consider the model (33), we drawn $x_{\ell i,t}$ as

$$x_{\ell i,t} = \boldsymbol{\gamma}_{xi}^{0\prime}\boldsymbol{f}_{xt}^{0} + \varepsilon_{xi,t} \tag{41}$$

and the idiosyncratic errors of the process for $y_{i,t}$ as

$$u_{i,t} = \boldsymbol{\gamma}_{yi}^{0\prime}\boldsymbol{f}_{yt}^{0} + \varepsilon_{yi,t}, \tag{42}$$

where $\boldsymbol{\gamma}_{yi}^0$ and $\boldsymbol{\gamma}_{xi}^0$ are $m_y \times 1$ and $m_x \times 1$ true factor loading respectively, $\boldsymbol{f}_{yt}^0$ and $\boldsymbol{f}_{xt}^0$ are $m_y \times 1$ and $m_x \times 1$ true vector of unobservable factors respectively.

3.1 Norkutes' (2019) IVMG estimator

We asymptotically eliminate the common factor in $\boldsymbol{x}_i$ by projecting matrix, $\boldsymbol{M}_{F_x^0}$.

$$\boldsymbol{M}_{F_x^0} = \boldsymbol{I}_T - \boldsymbol{F}_x^0\left(\boldsymbol{F}_x^{0\prime}\boldsymbol{F}_x^0\right)^{-1}\boldsymbol{F}_x^{0\prime}; \boldsymbol{M}_{F_{x,-1}^0} = \boldsymbol{I}_T - \boldsymbol{F}_{x,-1}^0\left(\boldsymbol{F}_{x,-1}^{0\prime}\boldsymbol{F}_{x,-1}^0\right)^{-1}\boldsymbol{F}_{x,-1}^{0\prime} \tag{43}$$

And using the defactored covariates as instruments, as

$$\boldsymbol{Z}_{IVi} = \left(\boldsymbol{M}_{F_X^0}\boldsymbol{x}_i, \boldsymbol{M}_{F_{x,-1}^0}\boldsymbol{X}_{i,-1}\right) \tag{44}$$

The first step IV estimator can be expressed as

$$\begin{aligned}\hat{\boldsymbol{\varphi}}_{IVi} = &\left(\left(\frac{\boldsymbol{Z}_i'\boldsymbol{M}_{F_X^0}\boldsymbol{W}_i}{T}\right)'\left(\frac{\boldsymbol{Z}_i'\boldsymbol{M}_{F_X^0}\boldsymbol{Z}_i}{T}\right)^{-1}\left(\frac{\boldsymbol{Z}_i'\boldsymbol{M}_{F_X^0}\boldsymbol{W}_i}{T}\right)\right)^{-1}\\ &\left(\left(\frac{\boldsymbol{Z}_i'\boldsymbol{M}_{F_X^0}\boldsymbol{W}_i}{T}\right)'\left(\frac{\boldsymbol{Z}_i'\boldsymbol{M}_{F_X^0}\boldsymbol{Z}_i}{T}\right)^{-1}\left(\frac{\boldsymbol{Z}_i'\boldsymbol{M}_{F_X^0}\boldsymbol{y}_i}{T}\right)\right).\end{aligned} \tag{45}$$

4 Monte Carlo simulation design

4.1 dynamic heterogeneous panels data model without error factor structure

The data generating process:

$$
\begin{aligned}
y_{i,t} &= \phi_i y_{i,t-1} + \sum_{\ell=1}^{k} \beta_{\ell i} x_{\ell i,t} + u_{i,t}, \;\; for \;\; i = 1, \ldots N; \; t = -49, \ldots, T\,, \\
x_{\ell i,t} &= \sum_{\ell=1}^{k} \phi_{\ell i} x_{\ell i,t-1} + v_{\ell i,t},
\end{aligned}
\tag{46}
$$

where $u_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}(0,\, 1)$, and $v_{\ell i,t} = \rho_{v,\ell} v_{\ell i,t-1} + \left(1 - \rho_{v,\ell}^2\right)^{\frac{1}{2}} \varpi_{\ell i,t}, \varpi_{\ell i,t} \overset{i.i.d.}{\sim} U(0.5, 1.5)\,, \rho_{v,\ell} =$ 0.5.

The slope coefficients are generated as

$$
\phi_i = \phi + \eta_{\phi i}, \;\; \beta_{1,i} = \beta_1 + \eta_{\beta_1 i} \, and \, \beta_{2,i} = \beta_2 + \eta_{\beta_2 i}. \tag{47}
$$

Here we consider $\phi \in \{0.5\}$, $\beta_1 = 3$ and $\beta_2 = 1$. For the design of heterogenous slopes, $\eta_{\phi i} \overset{i.i.d.}{\sim} U\,(-c, c)$, and

$$
\eta_{\beta_\ell i} = \left(1 - \rho_\beta^2\right)^{1/2} \eta_{\phi i}. \tag{48}
$$

Here, we set $c = 0.2$, $\rho_\beta = 0.4$ for $\ell = 1, 2$.

4.2 Dynamic heterogeneous panels data model with multi-factor error structure

This Monte Carlo simulation design same as Norkute et al. (2019). For convenience, we rewrite the data generating process as bellow

$$
y_{i,t} = \alpha_i + \phi_i y_{i,t-1} + \sum_{\ell=1}^{k} \beta_{\ell i} x_{\ell i,t} + u_{i,t}, \;\; for \;\; i = 1, \ldots N; \; t = -49, \ldots, T\,. \tag{49}
$$

$$
\tag{50}
$$

We allow error factor structure in the model as

$$
u_{i,t} = \sum_{s=1}^{m_y} \gamma_{si}^0 f_{s,t}^0 + \varepsilon_{i,t}, \tag{51}
$$

where

$$
f_{s,t}^0 = \rho_{s,t}^0 f_{s,t-1}^0 + \left(1 - \rho_{fs}^2\right)^{1/2} \zeta_{s,t}, \tag{52}
$$

with $\zeta_{s,t} \overset{i.i.d.}{\sim} N(0,1)$ for $s = 1, \ldots m_y$. We assume $k = 2$ and $m_y = 1 + k = 3$ and set $\rho_{s,t}^0 = 0.5$ for all s. The error term, $\varepsilon_{i,t}$, setting as

$$
\varepsilon_{i,t} = \varsigma_\varepsilon \sigma_{it} \left(\epsilon_{it} - 1\right) / \sqrt{2}, \tag{53}
$$

where $\epsilon_{it} \overset{i.i.d.}{\sim} \chi^2_1$, $\sigma^2_{it} = \eta_i \varphi_t$, $\eta_i \overset{i.i.d.}{\sim} \chi^2_2/2$, and $\varphi_t = t/T \; for \; t = 0, \ldots, T$. And we set

$$\varsigma_\varepsilon = \frac{\pi_\mu}{1-\pi_\mu} m_y. \tag{54}$$

we set $\pi_\mu \in \{3/4\}$.

The process of regressors is

$$x_{\ell it} = \mu_{\ell i} + + \sum_{\ell=1}^{k} \phi_{\ell i} x_{\ell i,t-1} + \sum_{s=1}^{m_x} \gamma^0_{\ell si} f^0_{s,t} + v_{\ell it}, \; for \; i = 1, \ldots N; \, t = -49, \ldots, T \, ; \, \ell = 1, 2. \tag{55}$$

We set number of factor, m_x, is 2. Therefore, $\boldsymbol{f}^0_{y,t} = \left(f^0_{1t}, f^0_{2t}, f^0_{3t}\right)'$ and $\boldsymbol{f}^0_{x,t} = \left(f^0_{1t}, f^0_{2t}\right)'$. We set

$$v_{\ell i,t} = \rho_{v,\ell} v_{\ell i,t-1} + \left(1 - \rho^2_{v,\ell}\right)^{\frac{1}{2}} \varpi_{\ell i,t}, for \; \ell = 1, 2, \tag{56}$$

where $\rho_{v,\ell} = 0.5$ for all ℓ. The individual effect is

$$\alpha^*_i \overset{i.i.d.}{\sim} N\left(0, (1-\rho_i)^2\right), \; \mu^*_{\ell i} = \rho_{\mu,\ell} \alpha^*_i + \left(1 - \rho^2_{\mu,\ell}\right)^{1/2} \omega_{\ell i}, \tag{57}$$

where $\omega \overset{i.i.d.}{\sim} N\left(0, (1-\rho_i)^2\right)$ and $\rho_{\mu,\ell} = 0.5$.

Now, we define the factor loading in $u_{i,t}$ are generated as $\gamma^{0*}_{si} \overset{i.i.d.}{\sim} N(0,1)$, for $s = 1, \ldots, m_y = 3$, and the factor loading in x_{1it} and x_{2it} are drawn as

$$\begin{aligned} \gamma^{0*}_{1si} &= \rho_{\gamma,1s} \gamma^{0*}_{3i} + \left(1 - \rho^2_{\gamma,1s}\right)^{1/2} \xi_{1si}; \; \xi_{1si} \overset{i.i.d.}{\sim} N(0,1); \\ \gamma^{0*}_{2si} &= \rho_{\gamma,2s} \gamma^{0*}_{si} + \left(1 - \rho^2_{\gamma,2s}\right)^{1/2} \xi_{2si}; \; \xi_{2si} \overset{i.i.d.}{\sim} N(0,1); \end{aligned} \tag{58}$$

for $s = 1, \ldots, m_x = 2$. We set $\rho_{\gamma,11} = \rho_{\gamma,12} \in \{0.5\}$ and $\rho_{\gamma,21} = \rho_{\gamma,22} = 0.5$. The factor loading are generated as

$$\boldsymbol{\Gamma} = \boldsymbol{\Gamma}^0 + \boldsymbol{\Gamma}^{0*}_i \tag{59}$$

where

$$\boldsymbol{\Gamma}^0_i = \begin{bmatrix} \gamma^0_{1i} & \gamma^0_{11i} & \gamma^0_{21i} \\ \gamma^0_{2i} & \gamma^0_{12i} & \gamma^0_{22i} \\ \gamma^0_{3i} & 0 & 0 \end{bmatrix} \tag{60}$$

and

$$\boldsymbol{\Gamma}^{0*}_i = \begin{bmatrix} \gamma^{0*}_{1i} & \gamma^{0*}_{11i} & \gamma^{0*}_{21i} \\ \gamma^{0*}_{2i} & \gamma^{0*}_{12i} & \gamma^{0*}_{22i} \\ \gamma^{0*}_{3i} & 0 & 0 \end{bmatrix}. \tag{61}$$

We set

$$\boldsymbol{\Gamma}^0 = \begin{bmatrix} 1/4 & 1/4 & -1 \\ 1/2 & -1 & 1/4 \\ 1/2 & 0 & 0 \end{bmatrix}. \tag{62}$$

And

$$\alpha_i = \alpha + \alpha_i^*,\ \mu_{\ell i} = \mu_\ell + \mu_{\ell i}^*, \tag{63}$$

where $\alpha = 1/2$, $\mu_1 = 1$, $\mu_2 = -1/2$.

The slope coefficients are generated as

$$\phi_i = \phi + \eta_{\phi i},\ \beta_{1,i} = \beta_1 + \eta_{\beta_1 i}\, and\, \beta_{2,i} = \beta_2 + \eta_{\beta_2 i}. \tag{64}$$

Here we consider $\phi \in \{0.5\}$, $\beta_1 = 3$ and $\beta_2 = 1$. For the design of heterogenous slopes, $\eta_{\phi i} \overset{i.i.d.}{\sim} U\left(-c, c\right)$, and

$$\eta_{\beta_\ell i} = \left[(2c)^2/12\right] \rho_\beta \xi_{\beta\ell i} + \left(1 - \rho_\beta^2\right)^{1/2} \eta_{\phi i}, \tag{65}$$

where

$$\xi_{\beta\ell i} = \frac{\bar{v_{\ell i}^2} - \bar{v_\ell^2}}{\left[N^{-1} \sum_{i=1}^N \left(\bar{v_{\ell i}^2} - \bar{v_\ell^2}\right)^2\right]^{1/2}}, \tag{66}$$

with $\bar{v_{elli}^2} = T^{-1} \sum_{t=1}^T v_{\ell it}^2$, $\bar{v_\ell^2} = N^{-1} \sum_{i=1}^N \bar{v_{\ell i}^2}$, for $\ell = 1, 2$. Here, we set $c = 0.2$, $\rho_\beta = 0.4$ for $\ell = 1, 2$. And

$$\varsigma_v^2 = \varsigma_\varepsilon^2 \left[SNR - \frac{\rho_v^2}{1 - \rho_v^2}\right] \left(\frac{\beta_1^2 + \beta_2^2}{1 - \rho_v^2}\right)^{-1}, \tag{67}$$

where $SNR = 4$. For the (T, N), we consider $T \in \{25, 50, 100, 200\}$ and $N \in \{25, 50, 100, 200\}$.

5 Monte Carlo simulation results

5.1 Dynamic Heterogeneous Panels without multifactor error structure

We consider ARDL(1,0) model.
$\phi \in \{0.5\}$.
$\beta_1 = 3$.
$\beta_2 = 1$.
$u_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}(0,\ 1)$.
$\varpi_{\ell i,t} \overset{i.i.d.}{\sim} U(0.5, 1.5)$.
$\rho_{v,\ell} = 0.5$.
$c = 0.2$.
$\rho_\beta = 0.4$.
$T \in \{25, 50, 100, 200\}$.
$N \in \{25, 50, 100, 200\}$.

LSMG estimator is provided in sheet 1 of MC.xlsx file.
IVMG estimator is provided in sheet 2 of MC.xlsx file.

5.2 Dynamic Heterogeneous Panels with multifactor error structure

We consider ARDL(1,0) model.
$\phi \in \{0.5\}$.
$\beta_1 = 3.$
$\beta_2 = 1.$
$k = 2.$
$m_y = 1 + k = 3.$
$m_x = k = 2.$
$\zeta_{s,t} \overset{i.i.d.}{\sim} N(0, 1)$
$\pi_\mu \in \{3/4\}$.
$\rho^0_{s,t} = 0.5.$
$\rho_{v,\ell} = 0.5.$
$\rho_{\mu,\ell} = 0.5.$
$\gamma^{0*}_{si} \overset{i.i.d.}{\sim} N(0, 1)$.
$\xi_{1si} \overset{i.i.d.}{\sim} N(0, 1)$.
$\xi_{2si} \overset{i.i.d.}{\sim} N(0, 1)$.
$\rho_{\gamma,11} = \rho_{\gamma,12} \in \{0.5\}$.
$\rho_{\gamma,21} = \rho_{\gamma,22} = 0.5.$
$\mathbf{\Gamma}^0 = \begin{bmatrix} 1/4 & 1/4 & -1 \\ 1/2 & -1 & 1/4 \\ 1/2 & 0 & 0 \end{bmatrix}.$
$\alpha = 1/2.$
$\mu_1 = 1.$
$\mu_2 = -1/2.$
$c = 0.2.$
$\rho_\beta = 0.4.$
$SNR = 4.$
$T \in \{25, 50, 100, 200\}$.
$N \in \{25, 50, 100, 200\}$.

IVMG estimator is provided in sheet 3 of MC.xlsx file.

6 Short summary

6.1 Dynamic Heterogeneous Panels without multifactor error structure

1. The performance of IVMG estimator is better than LSMG estimator in bias and RMSE.

6.2 Dynamic Heterogeneous Panels with multifactor error structure

1. When N and T increase, the performance of IVMG estimator is good in bias and RMSE.
Related literature to dynamic Heterogeneous Panels with multifactor error structure: Chudik and Pesaran (2015) and Norkute et al. (2019).
Related literature to choosing number of instruments: Donald and Newey (2001), Swanson (2005), Carrasco (2012), Bai and Ng (2010) and Kang (2019).

References

Bai, J. and S. Ng (2010). Instrumental variable estimation in a data rich enviroment. *Econometric Theory 26*, 1577–1606.

Carrasco, M. (2012). A regularization approach to the many instruments problem. *Journal of Econometrics 170*, 383–398.

Chudik, A. and M. H. Pesaran (2015). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics 188*, 393–420.

Donald, S. G. and W. K. Newey (2001). Choosing the number of instruments. *Econometrica 69*, 1161–1191.

Kang, B. (2019). Higher order approximation of iv estimators with invalid instruments.

Norkute, M., V. Sarafidis, T. Yamagata, and G. Cui (2019). Instrumental variable estimation of dynamic linear panel data models with defactored regressors and a multifactor error structure. ISER Discussion Paper No. 1019.

Pesaran, M. and R. Smith (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics 68*, 79 – 113.

Swanson, J. C. C. N. R. (2005). Consistent estimation with a large number of weak instruments. *Econometrica 73*, 1673–1692.