

---

# Dynamic Heterogeneous Panels

---

# Contents

<b>1</b>	<b>Asymptotic property of LS and IV estimator</b>	<b>2</b>
1.1	Asymptotic property of LS estimator . . . . .	2
1.1.1	Mean group LS estimator . . . . .	3
1.2	Asymptotic property of IV estimator . . . . .	6
1.2.1	Mean group IV estimator . . . . .	7
<b>2</b>	<b>Estimation method on dynamic heterogeneous panel data model with multifactor error structure</b>	<b>8</b>
2.1	Norkutes' (2019) IVMG estimator . . . . .	8
<b>3</b>	<b>Monte Carlo simulation design</b>	<b>8</b>
3.1	dynamic heterogeneous panels data model without error factor structure	8
3.2	Dynamic heterogeneous panels data model with multi-factor error structure . . . . .	9
<b>4</b>	<b>Monte Carlo simulation results</b>	<b>11</b>
4.1	Dynamic Heterogeneous Panels without multifactor error structure .	11
4.2	Dynamic Heterogeneous Panels with multifactor error structure . . .	11
<b>5</b>	<b>Short summary</b>	<b>12</b>
5.1	Dynamic Heterogeneous Panels without multifactor error structure .	12
5.2	Dynamic Heterogeneous Panels with multifactor error structure . . .	12
	<b>Reference</b>	<b>13</b>

# 1 Asymptotic property of LS and IV estimator

## 1.1 Asymptotic property of LS estimator

Consider the dynamic heterogeneous panels data model:

$$y_{i,t} = \alpha_i + \phi_i y_{i,t-1} + \beta_i x_{i,t} + u_{i,t}, \text{ for } i = 1, \dots, N; t = 1, \dots, T, \quad (1)$$

### Assumption 1

$x_{i,t}$  and  $u_{i,t}$  are independently distributed for all  $t$  and  $s$ .

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

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

where  $\tilde{y}_{i,t} = y_{i,t} - \bar{y}_i$ ,  $\tilde{y}_{i,t-1} = y_{i,t-1} - \bar{y}_{i,-1}$  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

$$\begin{pmatrix} \hat{\phi}_i - \phi_i \\ \hat{\beta}_i - \beta_i \end{pmatrix} = \begin{pmatrix} \frac{\sum_{t=1}^T \tilde{y}_{i,t-1}^2}{T} & \frac{\sum_{t=1}^T \tilde{y}_{i,t-1} \tilde{x}_{i,t}}{T} \\ \frac{\sum_{t=1}^T \tilde{x}_{i,t} \tilde{y}_{i,t-1}}{T} & \frac{\sum_{t=1}^T \tilde{x}_{i,t}^2}{T} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\sum_{t=1}^T \tilde{y}_{i,t-1} \tilde{u}_{i,t}}{T} \\ \frac{\sum_{t=1}^T \tilde{x}_{i,t} \tilde{u}_{i,t}}{T} \end{pmatrix}, \quad (3)$$

where  $\tilde{u}_{i,t}$  is  $u_{i,t} - \bar{u}_i$  with  $\bar{u}_i = \frac{1}{T} \sum_{t=1}^T u_{i,t}$ .

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

$$A_{\phi i}^{(1)} = \text{plim}_{T \rightarrow \infty} \left( \frac{\sum_{t=1}^T \tilde{y}_{i,t-1} \tilde{u}_{i,t}}{T} \right). \quad (4)$$

Then  $A_i^{(1)}$  can be taken expectations as

$$\begin{aligned} A_{\phi i}^{(1)} &= E(y_{i,t-1} - \bar{y}_{i,-1})(u_{i,t} - \bar{u}_i) \\ &= E(y_{i,t-1} u_{i,t}) - E(y_{i,t-1} \bar{u}_i) - E(\bar{y}_{i,-1} u_{i,t}) + E(\bar{y}_{i,-1} \bar{u}_i), \end{aligned} \quad (5)$$

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

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

$$y_{i,t} = \frac{\alpha_i}{(1 - \phi_i)} + \sum_{s=0}^{\infty} \beta_i \phi_i^s x_{i,t-s} + \sum_{s=0}^{\infty} \phi_i^s u_{i,t-s}, \quad (6)$$

Then, we have

$$\begin{aligned} A_{\phi i}^{(1)} &= -E \left( \left( \sum_{s=0}^{\infty} \phi_i^s u_{i,t-s-1} \right) \left( \frac{1}{T} \sum_{t=1}^T u_{i,t} \right) \right) - E \left( \frac{u_{i,t}}{T} \sum_{t=1}^T \sum_{s=0}^{\infty} \phi_i^s u_{i,t-s-1} \right) + \\ &\quad \left( \frac{1}{T} \sum_{t=1}^T \sum_{s=0}^{\infty} \phi_i^s u_{i,t-s-1} \right) \left( \frac{1}{T} \sum_{t=1}^T u_{i,t} \right). \end{aligned} \quad (7)$$

Hence, from above equation, we have

$$\begin{aligned}
A_{\phi_i}^{(1)} &= -\frac{1}{T}E\left\{(u_{i,t-1} + u_{i,t-2}\phi_i^1 + u_{i,t-3}\phi_i^2 + \dots)(u_{i,1} + \dots + u_{i,t-1} + u_{i,t} + \dots + u_{i,T})\right\} - \\
&\quad \frac{1}{T}E\left\{u_{i,t}\sum_{s=1}^T(u_{i,s-1}\phi_i^0 + u_{i,s-2}\phi_i^1 + \dots + u_{i,s-t-1}\phi_i^t + \dots + u_{i,s-T-1}\phi_i^T + \dots)\right\} + \\
&\quad +\frac{1}{T}E\left\{\left(\sum_{s=1}^T u_{i,s-1}\phi_i^0 + \sum_{s=1}^T u_{i,s-2}\phi_i^1 + \dots + \sum_{s=1}^T u_{i,s-t-1}\phi_i^t + \dots + \sum_{s=1}^T u_{i,s-T-1}\phi_i^T + \dots\right)\right. \\
&\quad \left.\left(\frac{1}{T}\sum_{s=1}^T u_{i,s}\right)\right\} \\
&= -\frac{\sigma_u^2(1-\phi_i^{t-1})}{T(1-\phi_i)} - \frac{\sigma_u^2(1-\phi_i^{T-t})}{T(1-\phi_i)} + \frac{\sigma_u^2}{T}\left(\frac{1}{1-\phi_i} - \frac{1}{T}\frac{(1-\phi_i^T)}{(1-\phi_i)^2}\right) \\
&= -\frac{\sigma_u^2}{T(1-\phi_i)}\left(1 - \phi_i^{t-1} - \phi_i^{T-t} + \frac{1}{T}\frac{(1-\phi_i^T)}{(1-\phi_i)}\right).
\end{aligned} \tag{8}$$

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

To be more compact, we can rewrite the model as,

$$\tilde{\mathbf{y}}_i = \tilde{\mathbf{W}}_i \boldsymbol{\theta}_i + \tilde{\mathbf{u}}_i, \tag{9}$$

where  $\tilde{\mathbf{y}}_i = (\tilde{y}_{i,1}, \dots, \tilde{y}_{i,T})'$  is  $T \times 1$  vector,  $\tilde{\mathbf{W}}_i = (\tilde{\mathbf{w}}_{i,1}, \dots, \tilde{\mathbf{w}}_{i,T})'$  is  $T \times 2$  matrix and  $\tilde{\mathbf{u}}_i = (\tilde{u}_{i,1}, \dots, \tilde{u}_{i,T})'$  is  $T \times 1$  vector with  $\tilde{\mathbf{w}}_{i,t} = (y_{i,t-1} - \bar{y}_{i,-1}, x_{i,t} - \bar{x}_i)'$ , for  $t = 1, \dots, T$  and  $i = 1, \dots, N$ .

Also, we define our interested parameter as

$$(\phi_i, \beta_i)' = \boldsymbol{\theta}_i = \boldsymbol{\theta} + \boldsymbol{\lambda}_i, \tag{10}$$

where  $\boldsymbol{\lambda}_i \stackrel{i.i.d.}{\sim} (\mathbf{0}, \boldsymbol{\Sigma}_\lambda)$ . The least square estimator,  $\hat{\boldsymbol{\theta}}_{LS,i}$ , can be expressed as

$$\hat{\boldsymbol{\theta}}_{LS,i} = \left(\frac{\tilde{\mathbf{W}}_i' \tilde{\mathbf{W}}_i}{T}\right)^{-1} \frac{\tilde{\mathbf{W}}_i' \tilde{\mathbf{y}}_i}{T}. \tag{11}$$

From above assumptions, we have following theorem

**Theorem 1**

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i) \xrightarrow{d} N\left(\mathbf{0}, \left(\mathbf{W}_i' \mathbf{W}_i\right)^{-1} \boldsymbol{\Sigma}_{LS,i} \left(\mathbf{W}_i' \mathbf{W}_i\right)^{-1}\right), \tag{12}$$

where  $\boldsymbol{\Sigma}_{LS,i} = \text{plim}_{T \rightarrow \infty} T^{-1} \mathbf{W}_i' \mathbf{u}_i \mathbf{u}_i' \mathbf{W}_i$ .

### 1.1.1 Mean group LS estimator

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

$$\hat{\boldsymbol{\theta}}_{LSMG} = \frac{1}{N} \sum_{i=1}^N \hat{\boldsymbol{\theta}}_{LSi}. \tag{13}$$

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

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{LSMG} &= N^{-1} \sum_{i=1}^N \left( \frac{\tilde{\mathbf{W}}_i' \tilde{\mathbf{W}}_i}{T} \right)^{-1} \frac{\tilde{\mathbf{W}}_i' \tilde{\mathbf{y}}_i}{T} \\ &= \bar{\boldsymbol{\theta}} + N^{-1} \sum_{i=1}^N \left( \frac{\tilde{\mathbf{W}}_i' \tilde{\mathbf{W}}_i}{T} \right)^{-1} \frac{\tilde{\mathbf{W}}_i' \tilde{\mathbf{u}}_i}{T},\end{aligned}\tag{14}$$

where  $\bar{\boldsymbol{\theta}} = N^{-1} \sum_{i=1}^N \boldsymbol{\theta}_i$ . For fixed  $N$  and large  $T$ , we have

$$\text{plim}_{T \rightarrow \infty} \hat{\boldsymbol{\theta}}_{LSMG} = \bar{\boldsymbol{\theta}} + N^{-1} \sum_{i=1}^N \text{plim}_{T \rightarrow \infty} \left( \frac{\tilde{\mathbf{W}}_i' \tilde{\mathbf{W}}_i}{T} \right)^{-1} \text{plim}_{T \rightarrow \infty} \left( \frac{\tilde{\mathbf{W}}_i' \tilde{\mathbf{u}}_i}{T} \right)\tag{15}$$

Then, under the regularity conditions, we can obtain

$$\text{plim}_{T \rightarrow \infty} \hat{\boldsymbol{\theta}}_{LSMG} = \bar{\boldsymbol{\theta}}.\tag{16}$$

When  $N \rightarrow \infty$  and  $T \rightarrow \infty$  and by the law of large numbers, we can see that

$$\text{plim}_{T \rightarrow \infty, N \rightarrow \infty} \hat{\boldsymbol{\theta}}_{LSMG} = \boldsymbol{\theta}.\tag{17}$$

And the variance estimator of  $\hat{\boldsymbol{\theta}}_{LSMG}$  is given by

$$\hat{\boldsymbol{\Sigma}}_{LS,\lambda} = \frac{1}{N(N-1)} \sum_{i=1}^N \left( \hat{\boldsymbol{\theta}}_{LS,i} - \hat{\boldsymbol{\theta}}_{LSMG} \right) \left( \hat{\boldsymbol{\theta}}_{LS,i} - \hat{\boldsymbol{\theta}}_{LSMG} \right)'. \tag{18}$$

We show that  $\hat{\boldsymbol{\Sigma}}_{LS,\lambda}$  is consistent when  $N \rightarrow \infty$  and  $T \rightarrow \infty$ .

$$\begin{aligned}\sum_{i=1}^N \left( \hat{\boldsymbol{\theta}}_{LS,i} - \hat{\boldsymbol{\theta}}_{LSMG} \right) \left( \hat{\boldsymbol{\theta}}_{LS,i} - \hat{\boldsymbol{\theta}}_{LSMG} \right)' &= \sum_{i=1}^N \left( \hat{\boldsymbol{\theta}}_{LS,i} - E \left( \hat{\boldsymbol{\theta}}_{LS,i} \right) \right) \left( \hat{\boldsymbol{\theta}}_{LS,i} - E \left( \hat{\boldsymbol{\theta}}_{LS,i} \right) \right)' \\ &\quad - \sum_{i=1}^N \left( \hat{\boldsymbol{\theta}}_{LSMG} - E \left( \hat{\boldsymbol{\theta}}_{LS,i} \right) \right) \left( \hat{\boldsymbol{\theta}}_{LSMG} - E \left( \hat{\boldsymbol{\theta}}_{LS,i} \right) \right)' \\ &\quad + \sum_{i=1}^N \left( \hat{\boldsymbol{\theta}}_{LSMG} - E \left( \hat{\boldsymbol{\theta}}_{LS,i} \right) \right) \left( \hat{\boldsymbol{\theta}}_{LSMG} - E \left( \hat{\boldsymbol{\theta}}_{LS,i} \right) \right)' \\ &\quad - \sum_{i=1}^N \left( \hat{\boldsymbol{\theta}}_{LS,i} - E \left( \hat{\boldsymbol{\theta}}_{LS,i} \right) \right) \left( \hat{\boldsymbol{\theta}}_{LS,i} - E \left( \hat{\boldsymbol{\theta}}_{LS,i} \right) \right)'.\end{aligned}\tag{19}$$

Taking expectation on equation (19), we have

$$\begin{aligned}
E \left( \sum_{i=1}^N \left( \hat{\theta}_{LS,i} - \hat{\theta}_{LSMG} \right) \left( \hat{\theta}_{LS,i} - \hat{\theta}_{LSMG} \right)' \right) &= \sum_{i=1}^N Var \left( \hat{\theta}_{LS,i} \right) + \sum_{i=1}^N \hat{\theta}_{LSMG} \hat{\theta}_{LSMG}' + \\
&\sum_{i=1}^N E \left( \hat{\theta}_{LS,i} \right) E \left( \hat{\theta}_{LS,i} \right)' - \sum_{i=1}^N \hat{\theta}_{LSMG} \hat{\theta}_{LSMG}' + \sum_{i=1}^N E \left( \hat{\theta}_{LS,i} \right) \hat{\theta}_{LSMG}' - \sum_{i=1}^N \hat{\theta}_{LSMG} E \left( \hat{\theta}_{LS,i} \right)' - \\
&\sum_{i=1}^N \hat{\theta}_{LSMG} \hat{\theta}_{LSMG}' + \sum_{i=1}^N \hat{\theta}_{LSMG} E \left( \hat{\theta}_{LS,i} \right)' - \sum_{i=1}^N E \left( \hat{\theta}_{LS,i} \right) E \left( \hat{\theta}_{LS,i} \right)' = \\
&\sum_{i=1}^N Var \left( \hat{\theta}_{LS,i} \right) - NE \left( \hat{\theta}_{LSMG} \hat{\theta}_{LSMG}' \right) + \sum_{i=1}^N E \left( \hat{\theta}_{LS,i} \right) E \left( \hat{\theta}_{LS,i} \right)'
\end{aligned} \tag{20}$$

and

$$E \left( \hat{\theta}_{LSMG} \hat{\theta}_{LSMG}' \right) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N E \left( \hat{\theta}_{LS,i} \hat{\theta}_{LS,i}' \right). \tag{21}$$

And, we also have

$$E \left( \hat{\theta}_{LSMG} \hat{\theta}_{LSMG}' \right) = \frac{1}{N^2} \left( \sum_{i=1}^N Var \left( \hat{\theta}_{LS,i} \right) + \sum_{i=1}^N \sum_{j=1}^N E \left( \hat{\theta}_{LS,i} \hat{\theta}_{LS,i}' \right) \right). \tag{22}$$

Then,

$$\begin{aligned}
E \left( \sum_{i=1}^N \left( \hat{\theta}_{LS,i} - \hat{\theta}_{LSMG} \right) \left( \hat{\theta}_{LS,i} - \hat{\theta}_{LSMG} \right)' \right) &= \\
\left( 1 - \frac{1}{N} \right) \sum_{i=1}^N Var \left( \hat{\theta}_{LS,i} \right) + \sum_{i=1}^N E \left( \hat{\theta}_{LS,i} \hat{\theta}_{LS,i}' \right) - \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N E \left( \hat{\theta}_{LS,i} \hat{\theta}_{LS,i}' \right).
\end{aligned} \tag{23}$$

From above equation (23), we can observe the bias term as,

$$\aleph = \sum_{i=1}^N E \left( \hat{\theta}_{LS,i} \hat{\theta}_{LS,i}' \right) - \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N E \left( \hat{\theta}_{LS,i} \hat{\theta}_{LS,i}' \right). \tag{24}$$

Taking expectation on equation (26), we have

$$E \left( \hat{\theta}_{LS,i} \right) = \theta + E \left( b_i \right), \tag{25}$$

From equation (24), we know that

$$\aleph = \sum_{i=1}^N E \left( b_i \right) E \left( b_i' \right) - \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N E \left( b_i \right) E \left( b_j' \right). \tag{26}$$

When  $T \rightarrow \infty$ ,  $\aleph = 0$ . Therefore, we have following theorem

### Theorem 3

When  $(T, N) \xrightarrow{d} \infty$  such that  $N/T \rightarrow c$  with  $0 < c < \infty$ ,

1.

$$\sqrt{N} \left( \hat{\boldsymbol{\theta}}_{LSMG} - \boldsymbol{\theta} \right) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_{LS,\lambda}). \quad (27)$$

2.

$$\hat{\boldsymbol{\Sigma}}_{LS,\lambda} \xrightarrow{p} \boldsymbol{\Sigma}_{LS,\lambda} \quad (28)$$

where

$$\hat{\boldsymbol{\Sigma}}_{LS,\lambda} = \frac{1}{N(N-1)} \sum_{i=1}^N \left( \hat{\boldsymbol{\theta}}_{LS,i} - \hat{\boldsymbol{\theta}}_{LSMG} \right) \left( \hat{\boldsymbol{\theta}}_{LS,i} - \hat{\boldsymbol{\theta}}_{LSMG} \right)' . \quad (29)$$

## 1.2 Asymptotic property of IV estimator

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

$$\mathbf{Z}_{i,t} = (\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_{i,-1})', \quad (30)$$

where  $\mathbf{Z}_i$  is  $T \times 2$  vector.

### Assumption 3

$\mathbf{A}_i = \text{plim}_{T \rightarrow \infty} \tilde{\mathbf{A}}_{i,T}$  has full column rank,  $\mathbf{B}_i = \text{plim}_{T \rightarrow \infty} \tilde{\mathbf{B}}_{i,T}$  and  $\boldsymbol{\Sigma}_i = \text{plim}_{T \rightarrow \infty} T^{-1} \mathbf{Z}_i' \mathbf{u}_i \mathbf{u}_i' \mathbf{Z}_i$  has positive definite, uniformly.

Then, the IV estimator can be expressed as

$$\hat{\boldsymbol{\theta}}_{IV,i} = \left( \tilde{\mathbf{A}}_{i,T}' \tilde{\mathbf{B}}_{i,T}^{-1} \tilde{\mathbf{A}}_{i,T} \right) \tilde{\mathbf{A}}_{i,T}' \tilde{\mathbf{B}}_{i,T}^{-1} \tilde{\mathbf{g}}_{i,T}, \quad (31)$$

where

$$\tilde{\mathbf{A}}_{i,T} = \frac{1}{T} \mathbf{Z}_i' \mathbf{W}_i, \quad \tilde{\mathbf{B}}_{i,T} = \frac{1}{T} \mathbf{Z}_i' \mathbf{Z}_i, \quad \tilde{\mathbf{g}}_{i,T} = \frac{1}{T} \mathbf{Z}_i' \tilde{\mathbf{y}}_i, \quad (32)$$

and  $\mathbf{W}_i = (\tilde{w}_{i,1}', \dots, \tilde{w}_{i,T}')'$  is  $T \times 2$  matrix

From above equation, we have

$$\sqrt{T} \left( \hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i \right) = \left( \tilde{\mathbf{A}}_{i,T}' \tilde{\mathbf{B}}_{i,T}^{-1} \tilde{\mathbf{A}}_{i,T} \right) \tilde{\mathbf{A}}_{i,T}' \tilde{\mathbf{B}}_{i,T}^{-1} \left( T^{-1/2} \mathbf{Z}_i' \mathbf{u}_i \right) \quad (33)$$

Then, the property of  $T^{-1/2} \mathbf{Z}_i' \mathbf{u}_i$  is given by following proposition.

### Proposition 1

Under above assumptions, as  $(N, T) \xrightarrow{j} \infty$  such that  $N/T \rightarrow c$  with  $0 < c < \infty$ , for each  $i$ , we have

$$\begin{aligned} E \left( \mathbf{Z}_i' \mathbf{u}_i \right) &= 0 \\ \text{and} \\ T^{-1/2} \mathbf{Z}_i' \mathbf{u}_i &\xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_i). \end{aligned} \quad (34)$$

Thus, IV estimator,  $\hat{\boldsymbol{\theta}}_{IV,i}$  is  $\sqrt{T}$  consistent to  $\boldsymbol{\theta}_i$  and this estimator does not have Nickell's bias. Then, we have following theorem

### Theorem 1

As  $(N, T) \rightarrow \infty$  such that  $N/T \rightarrow c$  with  $0 < c < \infty$ . for each  $i$ ,

$$\sqrt{T} \left( \hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i \right) \xrightarrow{d} N \left( \mathbf{0}, \left( \mathbf{A}_i' \mathbf{B}_i^{-1} \mathbf{A}_i \right)^{-1} \mathbf{A}_i' \mathbf{B}_i^{-1} \boldsymbol{\Sigma}_i \mathbf{B}_i^{-1} \mathbf{A}_i \left( \mathbf{A}_i' \mathbf{B}_i^{-1} \mathbf{A}_i \right) \right). \quad (35)$$

### 1.2.1 Mean group IV estimator

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

$$\hat{\boldsymbol{\theta}}_{IVMG} = \frac{1}{N} \sum_{i=1}^N \hat{\boldsymbol{\theta}}_{IVi}. \quad (36)$$

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

$$\hat{\boldsymbol{\theta}}_{IVMG} - \boldsymbol{\theta} = \frac{1}{N} \sum_{i=1}^N (\hat{\boldsymbol{\theta}}_{IVMG} - \boldsymbol{\theta}) = \frac{1}{N} \sum_{i=1}^N (\hat{\boldsymbol{\theta}}_{IVMG} - \boldsymbol{\theta}_i) + \frac{1}{N} \sum_{i=1}^N \boldsymbol{\lambda}_i. \quad (37)$$

$$\sqrt{N} (\hat{\boldsymbol{\theta}}_{IVMG} - \boldsymbol{\theta}) = \frac{1}{\sqrt{N}} \sum_{i=1}^N \boldsymbol{\lambda}_i + o_p(1). \quad (38)$$

And the variance estimator of  $\hat{\boldsymbol{\theta}}_{IVMG}$  is given by

$$\hat{\boldsymbol{\Sigma}}_{IV,\lambda} = \frac{1}{N-1} \sum_{i=1}^N (\hat{\boldsymbol{\theta}}_{IV,i} - \hat{\boldsymbol{\theta}}_{IVMG}) (\hat{\boldsymbol{\theta}}_{IV,i} - \hat{\boldsymbol{\theta}}_{IVMG})'. \quad (39)$$

Follow [Norkute et al. \(2019\)](#), we can show that  $\hat{\boldsymbol{\Sigma}}_{IV,\lambda}$  is consistent and it does not have small  $T$  bias. Firstly, we decompose (39) as

$$\begin{aligned} & \sum_{i=1}^N (\hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta} + \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{IVMG}) (\hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta} + \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{IVMG})' = \\ & \sum_{i=1}^N \boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' + \sum_{i=1}^N (\hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i) (\hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i)' + \sum_{i=1}^N (\hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i) \boldsymbol{\lambda}_i + \sum_{i=1}^N \boldsymbol{\lambda}_i (\hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i) - \\ & N (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{IVMG})' (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{IVMG}). \end{aligned} \quad (40)$$

Then we can show consistent of  $\hat{\boldsymbol{\Sigma}}_{IV,\lambda}$  as

$$\begin{aligned} \hat{\boldsymbol{\Sigma}}_{IV,\lambda} - \boldsymbol{\Sigma}_{IV,\lambda} &= \frac{1}{N-1} \sum_{i=1}^N (\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{IV,\lambda}) + \frac{1}{N-1} \sum_{i=1}^N (\hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i) (\hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i)' \\ &+ \frac{1}{N-1} \sum_{i=1}^N (\hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i) \boldsymbol{\lambda}_i + \frac{1}{N-1} \sum_{i=1}^N \boldsymbol{\lambda}_i (\hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i) - \\ &\frac{N}{N-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{IVMG})' (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{IVMG}) = o_p(1). \end{aligned} \quad (41)$$

*(The proof is similar as )*

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

$$\sqrt{N} (\hat{\boldsymbol{\theta}}_{IVMG} - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_{IV,\lambda}). \quad (42)$$



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

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

$$y_{i,t} = \phi_i y_{i,t-1} + \sum_{\ell=1}^k \beta_{\ell i} x_{\ell i,t} + u_{i,t}, \text{ for } i = 1, \dots, N; t = 1, \dots, T, \ell = 1, \dots, k. \quad (43)$$

Consider the model (40), we drawn  $x_{\ell i,t}$  as

$$x_{\ell i,t} = \gamma_{xi}^{0'} \mathbf{f}_{xt}^0 + \varepsilon_{xi,t} \quad (44)$$

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

$$u_{i,t} = \gamma_{yi}^{0'} \mathbf{f}_{yt}^0 + \varepsilon_{yi,t}, \quad (45)$$

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

### 2.1 Norkutes' (2019) IVMG estimator

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

$$\mathbf{M}_{F_x^0} = \mathbf{I}_T - \mathbf{F}_x^0 \left( \mathbf{F}_x^{0'} \mathbf{F}_x^0 \right)^{-1} \mathbf{F}_x^{0'}; \mathbf{M}_{F_{x,-1}^0} = \mathbf{I}_T - \mathbf{F}_{x,-1}^0 \left( \mathbf{F}_{x,-1}^{0'} \mathbf{F}_{x,-1}^0 \right)^{-1} \mathbf{F}_{x,-1}^{0'} \quad (46)$$

And using the defactored covariates as instruments, as

$$\mathbf{Z}_{IVi} = \left( \mathbf{M}_{F_X^0} \mathbf{x}_i, \mathbf{M}_{F_{x,-1}^0} \mathbf{X}_{i,-1} \right) \quad (47)$$

The first step IV estimator can be expressed as

$$\begin{aligned} \hat{\varphi}_{IVi} = & \left( \left( \frac{\mathbf{Z}_i' \mathbf{M}_{F_X^0} \mathbf{W}_i}{T} \right)' \left( \frac{\mathbf{Z}_i' \mathbf{M}_{F_X^0} \mathbf{Z}_i}{T} \right)^{-1} \left( \frac{\mathbf{Z}_i' \mathbf{M}_{F_X^0} \mathbf{W}_i}{T} \right) \right)^{-1} \\ & \left( \left( \frac{\mathbf{Z}_i' \mathbf{M}_{F_X^0} \mathbf{W}_i}{T} \right)' \left( \frac{\mathbf{Z}_i' \mathbf{M}_{F_X^0} \mathbf{Z}_i}{T} \right)^{-1} \left( \frac{\mathbf{Z}_i' \mathbf{M}_{F_X^0} \mathbf{y}_i}{T} \right) \right). \end{aligned} \quad (48)$$

## 3 Monte Carlo simulation design

### 3.1 dynamic heterogeneous panels data model without error factor structure

The data generating process:

$$\begin{aligned} y_{i,t} &= \phi_i y_{i,t-1} + \sum_{\ell=1}^k \beta_{\ell i} x_{\ell i,t} + u_{i,t}, \text{ for } i = 1, \dots, N; t = -49, \dots, T, \\ x_{\ell i,t} &= \sum_{\ell=1}^k \phi_{\ell i} x_{\ell i,t-1} + v_{\ell i,t}, \end{aligned} \quad (49)$$

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

The slope coefficients are generated as

$$\phi_i = \phi + \eta_{\phi i}, \beta_{1,i} = \beta_1 + \eta_{\beta_1 i} \text{ and } \beta_{2,i} = \beta_2 + \eta_{\beta_2 i}. \quad (50)$$

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

$$\eta_{\beta_{\ell} i} = (1 - \rho_{\beta}^2)^{1/2} \eta_{\phi i}. \quad (51)$$

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

### 3.2 Dynamic heterogeneous panels data model with multifactor error structure

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

$$y_{i,t} = \alpha_i + \phi_i y_{i,t-1} + \sum_{\ell=1}^k \beta_{\ell i} x_{\ell i,t} + u_{i,t}, \text{ for } i = 1, \dots, N; t = -49, \dots, T. \quad (52)$$

$$(53)$$

We allow error factor structure in the model as

$$u_{i,t} = \sum_{s=1}^{m_y} \gamma_{si}^0 f_{s,t}^0 + \varepsilon_{i,t}, \quad (54)$$

where

$$f_{s,t}^0 = \rho_{s,t}^0 f_{s,t-1}^0 + (1 - \rho_{s,t}^2)^{1/2} \zeta_{s,t}, \quad (55)$$

with  $\zeta_{s,t} \stackrel{i.i.d.}{\sim} N(0, 1)$  for  $s = 1, \dots, m_y$ . 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

$$\varepsilon_{i,t} = \varsigma_{\varepsilon} \sigma_{it} (\epsilon_{it} - 1) / \sqrt{2}, \quad (56)$$

where  $\epsilon_{it} \stackrel{i.i.d.}{\sim} \chi_1^2$ ,  $\sigma_{it}^2 = \eta_i \varphi_t$ ,  $\eta_i \stackrel{i.i.d.}{\sim} \chi_2^2/2$ , and  $\varphi_t = t/T$  for  $t = 0, \dots, T$ . And we set

$$\varsigma_{\varepsilon} = \frac{\pi_{\mu}}{1 - \pi_{\mu}} m_y. \quad (57)$$

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

The process of regressors is

$$x_{\ell i,t} = \mu_{\ell i} + \sum_{\ell=1}^k \phi_{\ell i} x_{\ell i,t-1} + \sum_{s=1}^{m_x} \gamma_{\ell si}^0 f_{s,t}^0 + v_{\ell i,t}, \text{ for } i = 1, \dots, N; t = -49, \dots, T; \ell = 1, 2. \quad (58)$$

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

$$v_{\ell i,t} = \rho_{v,\ell} v_{\ell i,t-1} + (1 - \rho_{v,\ell}^2)^{\frac{1}{2}} \varpi_{\ell i,t}, \text{ for } \ell = 1, 2, \quad (59)$$

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

$$\alpha_i^* \stackrel{i.i.d.}{\sim} N(0, (1 - \rho_i)^2), \mu_{\ell i}^* = \rho_{\mu,\ell} \alpha_i^* + (1 - \rho_{\mu,\ell}^2)^{1/2} \omega_{\ell i}, \quad (60)$$

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

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

$$\begin{aligned} \gamma_{1si}^{0*} &= \rho_{\gamma,1s} \gamma_{3i}^{0*} + (1 - \rho_{\gamma,1s}^2)^{1/2} \xi_{1si}; \xi_{1si} \stackrel{i.i.d.}{\sim} N(0, 1); \\ \gamma_{2si}^{0*} &= \rho_{\gamma,2s} \gamma_{3i}^{0*} + (1 - \rho_{\gamma,2s}^2)^{1/2} \xi_{2si}; \xi_{2si} \stackrel{i.i.d.}{\sim} N(0, 1); \end{aligned} \quad (61)$$

for  $s = 1, \dots, m_x = 2$ . 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

$$\mathbf{\Gamma} = \mathbf{\Gamma}^0 + \mathbf{\Gamma}_i^{0*} \quad (62)$$

where

$$\mathbf{\Gamma}_i^0 = \begin{bmatrix} \gamma_{1i}^0 & \gamma_{11i}^0 & \gamma_{21i}^0 \\ \gamma_{2i}^0 & \gamma_{12i}^0 & \gamma_{22i}^0 \\ \gamma_{3i}^0 & 0 & 0 \end{bmatrix} \quad (63)$$

and

$$\mathbf{\Gamma}_i^{0*} = \begin{bmatrix} \gamma_{1i}^{0*} & \gamma_{11i}^{0*} & \gamma_{21i}^{0*} \\ \gamma_{2i}^{0*} & \gamma_{12i}^{0*} & \gamma_{22i}^{0*} \\ \gamma_{3i}^{0*} & 0 & 0 \end{bmatrix}. \quad (64)$$

We set

$$\mathbf{\Gamma}^0 = \begin{bmatrix} 1/4 & 1/4 & -1 \\ 1/2 & -1 & 1/4 \\ 1/2 & 0 & 0 \end{bmatrix}. \quad (65)$$

And

$$\alpha_i = \alpha + \alpha_i^*, \mu_{\ell i} = \mu_{\ell} + \mu_{\ell i}^*, \quad (66)$$

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

The slope coefficients are generated as

$$\phi_i = \phi + \eta_{\phi i}, \beta_{1,i} = \beta_1 + \eta_{\beta 1i} \text{ and } \beta_{2,i} = \beta_2 + \eta_{\beta 2i}. \quad (67)$$

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

$$\eta_{\beta \ell i} = [(2c)^2/12] \rho_{\beta} \xi_{\beta \ell i} + (1 - \rho_{\beta}^2)^{1/2} \eta_{\phi i}, \quad (68)$$

where

$$\xi_{\beta\ell i} = \frac{v_{\ell i}^2 - \bar{v}_\ell^2}{\left[ N^{-1} \sum_{i=1}^N (v_{\ell i}^2 - \bar{v}_\ell^2)^2 \right]^{1/2}}, \quad (69)$$

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

$$\varsigma_v^2 = \varsigma_\varepsilon^2 \left[ SNR - \frac{\rho_v^2}{1 - \rho_v^2} \right] \left( \frac{\beta_1^2 + \beta_2^2}{1 - \rho_v^2} \right)^{-1}, \quad (70)$$

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

## 4 Monte Carlo simulation results

### 4.1 Dynamic Heterogeneous Panels without multifactor error structure

We consider ARDL(1,0) model.

$\phi \in \{0.5\}$ .

$\beta_1 = 3$ .

$\beta_2 = 1$ .

$u_{i,t} \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$ .

$\varpi_{\ell i,t} \stackrel{i.i.d.}{\sim} U(0.5, 1.5)$ .

$\rho_{v,\ell} = 0.5$ .

$c = 0.2$ .

$\rho_\beta = 0.4$ .

$T \in \{25, 50, 100, 200\}$ .

$N \in \{25, 50, 100, 200\}$ .

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

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

### 4.2 Dynamic Heterogeneous Panels with multifactor error structure

We consider ARDL(1,0) model.

$\phi \in \{0.5\}$ .

$\beta_1 = 3$ .

$\beta_2 = 1$ .

$k = 2$ .

$m_y = 1 + k = 3$ .

$m_x = k = 2$ .

$\zeta_{s,t} \stackrel{i.i.d.}{\sim} N(0, 1)$

$$\begin{aligned}
\pi_\mu &\in \{3/4\}. \\
\rho_{s,t}^0 &= 0.5. \\
\rho_{v,\ell} &= 0.5. \\
\rho_{\mu,\ell} &= 0.5. \\
\gamma_{si}^{0*} &\overset{i.i.d.}{\sim} N(0, 1). \\
\xi_{1si} &\overset{i.i.d.}{\sim} N(0, 1). \\
\xi_{2si} &\overset{i.i.d.}{\sim} N(0, 1). \\
\rho_{\gamma,11} = \rho_{\gamma,12} &\in \{0.5\}. \\
\rho_{\gamma,21} = \rho_{\gamma,22} &= 0.5. \\
\mathbf{\Gamma}^0 &= \begin{bmatrix} 1/4 & 1/4 & -1 \\ 1/2 & -1 & 1/4 \\ 1/2 & 0 & 0 \end{bmatrix}. \\
\alpha &= 1/2. \\
\mu_1 &= 1. \\
\mu_2 &= -1/2. \\
c &= 0.2. \\
\rho_\beta &= 0.4. \\
SNR &= 4. \\
T &\in \{25, 50, 100, 200\}. \\
N &\in \{25, 50, 100, 200\}.
\end{aligned}$$

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

## 5 Short summary

### 5.1 Dynamic Heterogeneous Panels without multifactor error structure

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

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