# Dynamic Heterogeneous Panels

# Contents

# 1 Asymptotic bias of LS estimator

## 1.1 Brief the source of bias

Consider the dynamic heterogeneous panels data model:

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

Form above model, the model can rewritten as

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

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

$$\begin{aligned} \theta_i &= \theta + \eta_{i1}, \\ \pi_i &= \pi + \eta_{i2}. \end{aligned} \tag{3}$$

Therefore, we know

$$\beta_i = \pi_i \theta_i = (\pi + \eta_{i2})(\theta + \eta_{i1}) = \pi\theta + \pi\eta_{i,1} + \theta\eta_{i2} + \eta_{i1}\eta_{i2} \tag{4}$$

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

$$\begin{aligned} y_{i,t} &= \phi_i y_{i,t-1} + (\pi\theta + \pi\eta_{i,1} + \theta\eta_{i2} + \eta_{i1}\eta_{i2}) x_{i,t} + u_{i,t} \\ &= (1 - \theta_i) y_{i,t-1} + \beta x_{i,t} + \eta_{i,3} x_{i,t} + u_{i,t} \\ &= (1 - \theta) y_{i,t-1} - \eta_{i1} y_{i,t-1} + \beta x_{i,t} + \eta_{i3} x_{i,t} + u_{i,t} \\ &= \phi y_{i,t-1} + \beta x_{i,t} + (u_{i,t} - \eta_{i1} y_{i,t-1} + \eta_{i3} x_{i,t}) \\ &= \phi y_{i,t-1} + \beta x_{i,t} + v_{i,t}, \end{aligned} \tag{5}$$

where $v_{i,t} = (u_{i,t} - \eta_{i1} y_{i,t-1} + \eta_{i3} x_{i,t})$. Then,we can see that $y_{i,t-1}$ and $x_{i,t}$ are correlated with $v_{i,t}$.

## 1.2 Asymptotic bias of LS estimator

To be continue.

# 2 estimation methods

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

$$y_{i,t} = \phi_i y_{i,t-1} + \beta_{1i} x_{i,t} + u_{i,t}, \ for \ i = 1, \ldots N; \ t = 1, \ldots, T. \tag{6}$$

We stack the $T$ observations for each $i$ yield

$$\boldsymbol{y}_i = \phi \boldsymbol{y}_{i,-1} + \boldsymbol{x}_i \beta + \boldsymbol{u}_i, \tag{7}$$

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

$$\boldsymbol{y}_i = \boldsymbol{W}_i \boldsymbol{\varphi} + \boldsymbol{u}_i, \tag{8}$$

where $\boldsymbol{W}_i = (\boldsymbol{y}_{i,-1}, \boldsymbol{x}_i)$ and $\boldsymbol{\varphi} = (\phi, \beta)'$

## 2.1 LSMG estimator

The LS (least square) estimator is defined as

$$\hat{\varphi}_{LSi} = \left( \boldsymbol{W}_i' \boldsymbol{W}_i \right)^{-1} \left( \boldsymbol{W}_i' \boldsymbol{y}_i \right) \tag{9}$$

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

$$\hat{\varphi}_{LSMG} = \frac{1}{N} \sum_{i=1}^{N} \hat{\varphi}_{LSi}. \tag{10}$$

## 2.2 IVMG estimator

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

$$\boldsymbol{Z}_i = \left( \boldsymbol{x}_i, \boldsymbol{x}_{i,-1} \right), \tag{11}$$

where $\boldsymbol{x}_{i,-1} = \left( x_{i,0}, \ldots, x_{i,T-1} \right)'$.

The IV (instrument variable) estimator is defined as

$$\hat{\varphi}_{IVi} = \left( \boldsymbol{Z}_i' \boldsymbol{W}_i \right)^{-1} \left( \boldsymbol{Z}_i' \boldsymbol{y}_i \right) \tag{12}$$

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

$$\hat{\varphi}_{IVMG} = \frac{1}{N} \sum_{i=1}^{N} \hat{\varphi}_{IVi}. \tag{13}$$

## 2.3 Norkutes' (2019) IVMG estimator

Consider the dynamic heterogeneous panel data model with multi-factor error structure. The error term $u_{i,t}$ and $\xi_{i,t}$ can be extend as

$$\begin{aligned} u_{i,t} &= \boldsymbol{\gamma}_{yi}^{0\prime} \boldsymbol{f}_{yt}^0 + \varepsilon_{yi,t}, \\ \xi_{i,t} &= \boldsymbol{\gamma}_{xi}^{0\prime} \boldsymbol{f}_{xt}^0 + \varepsilon_{xi,t}, \end{aligned} \tag{14}$$

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. And, we set $\varepsilon_{yi,t} \sim \mathcal{N}(0, \sigma_u^2)$, , $\varepsilon_{xi,t} \sim \mathcal{N}(0, \sigma_\xi^2)$. In the first step, we asymptotically eliminate the common factor in $\boldsymbol{x}_i$ by projecting matrix, $\boldsymbol{M}_{F_x^0}$.

$$\boldsymbol{M}_{F_x^0} = \boldsymbol{I}_T - \boldsymbol{F}_x^0 \left( \boldsymbol{F}_x^{0\prime} \boldsymbol{F}_x^0 \right)^{-1} \boldsymbol{F}_x^{0\prime}; \boldsymbol{M}_{F_{x,-1}^0} = \boldsymbol{I}_T - \boldsymbol{F}_{x,-1}^0 \left( \boldsymbol{F}_{x,-1}^{0\prime} \boldsymbol{F}_{x,-1}^0 \right)^{-1} \boldsymbol{F}_{x,-1}^{0\prime} \tag{15}$$

And using the defactored covariates as instruments, as

$$\boldsymbol{Z}_{IVi} = \left( \boldsymbol{M}_{F_x^0} \boldsymbol{x}_i, \boldsymbol{M}_{F_{x,-1}^0} \boldsymbol{x}_{i,-1} \right) \tag{16}$$

We use $\mathbf{Z}_{IVi}$ as instruments to obtain first step estimator. Then, we can use CCE (common correlated effects) method or PC (principal component) methods to estimate $\mathbf{F}_y^0$ from the residuals in first step IV regression. In second step, we asymptotically eliminate the common factor, $\mathbf{F}_y^0$ by projecting matrix, $\mathbf{M}_{F_y^0}$. Therefore, we premultiply $\mathbf{M}_{F_y^0}$ in both side of equation (8) as

$$\mathbf{M}_{F_y^0}\mathbf{y}_i = \mathbf{M}_{F_y^0}\mathbf{W}_i\boldsymbol{\varphi} + \mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i, \tag{17}$$

where $\boldsymbol{\varepsilon}_i = (\varepsilon_{i,1}, \ldots, \varepsilon_{i,T})$. The second step IV estimator can be expressed as

$$\hat{\hat{\boldsymbol{\varphi}}}_{IV} = \left( \left(\mathbf{W}_i'\mathbf{M}_{F_y^0}\mathbf{Z}_{IVi}\right) \left(\mathbf{Z}_{IVi}'\mathbf{M}_{F_y^0}\mathbf{Z}_{IVi}\right)^{-1} \left(\mathbf{Z}_{IVi}'\mathbf{M}_{F_y^0}\mathbf{W}\right) \right)^{-1} \times \tag{18}$$
$$\left(\mathbf{W}_i'\mathbf{M}_{F_y^0}\mathbf{Z}_{IVi}\right) \left(\mathbf{Z}_{IVi}'\mathbf{M}_{F_y^0}\mathbf{Z}_{IVi}\right)^{-1} \left(\mathbf{Z}_{IVi}'\mathbf{M}_{F_y^0}\mathbf{y}_i\right)$$

# 3 MC setting

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

Consider the dynamic heterogeneous panels data model:

$$y_{i,t} = \phi_i y_{i,t-1} + \beta_{1i}x_{i,t} + u_{i,t}, \ \ for \ i = 1, \ldots N; \ t = 1, \ldots, T, \tag{19}$$
$$x_{i,t} = \rho x_{i,t-1} + \xi_{i,t},$$

where $u_{i,t} \sim \mathcal{N}(0, \sigma_u^2)$, and $\xi_{i,t} \sim \mathcal{N}(0, \sigma_\xi^2)$.

## 3.2 dynamic heterogeneous panels data model with multi-factor error structure

The generating process of $y$ and $x$ follow (19), but we assume

$$u_{i,t} = \boldsymbol{\gamma}_{yi}'\mathbf{f}_{yt} + \varepsilon_{yi,t}, \tag{20}$$
$$\xi_{i,t} = \boldsymbol{\gamma}_{xi}'\mathbf{f}_{xt} + \varepsilon_{xi,t},$$

where $\boldsymbol{\gamma}_{yi}$ and $\boldsymbol{\gamma}_{xi}$ are $m_y \times 1$ and $m_x \times 1$ factor loading respectively, $\mathbf{f}_{yt}$ and $\mathbf{f}_{xt}$ are $m_y \times 1$ and $m_x \times 1$ unobservable factors respectively. And, we set $\varepsilon_{yi,t} \sim \mathcal{N}(0, \sigma_u^2)$, , $\varepsilon_{xi,t} \sim \mathcal{N}(0, \sigma_\xi^2)$.

$$\phi_i = \phi + \eta_{1i} \tag{21}$$
$$\beta_{1i} = \beta_1 + \eta_{2i},$$

where $\eta_{1i} \sim \mathcal{IIDU}(0.5, 1)$, and $\eta_{2i} \sim \mathcal{IIDU}(0, 0.8)$ We also try $\phi_i$ is fixed across group, which means $\phi_i = \phi$.

We chose $\phi = \{0.25\}$ and $\beta_1 = 1 - \phi$ and $\rho = \{0, 0.5\}$. Also, we selected $\sigma_u^2 = 1$ and $\sigma_\xi^2 = 1$. The replication is 1000. We define LSMG (lest square mean group) estimator and instrument variable mean group (IVMG) estimator.

In the simulation results, we provide the bias and RMSE (root mean square errors) for LSMG estimator and IVMG estimator.

# 4 MC results

Report in excel file.
In Excel file:

## 4.1 Model without factor structure

Sheet 1: ARDL(1,0); $\phi_i$ ; $\beta_i$ .
Sheet 2: ARDL(1,0); $\phi$ ; $\beta_i$ .
Sheet 3: ARDL(1,1); $\phi_i$ ; $\beta_i$ .
Sheet 4: ARDL(1,1); $\phi$ ; $\beta_i$ .

## 4.2 Model have factor structure 1

We estimate the number of factor is 0, so we use traditional OLS estimation method
and IV estimation method.
Sheet 1: ARDL(1,0); $\phi_i$ ; $\beta_i$.
Sheet 2: ARDL(1,0); $\phi$ ; $\beta_i$
Sheet 3: ARDL(1,1); $\phi_i$ ; $\beta_i$.
Sheet 4: ARDL(1,1); $\phi$ ; $\beta_i$ .

## 4.3 Model have factor structure 2

We provide MC simulation results for IV estimator that is been provided by Norkute
et al. (2019).
Sheet 1: ARDL(1,0); $\phi_i$ ; $\beta_i$.
Sheet 2: ARDL(1,0); $\phi$ ; $\beta_i$ .
Sheet 3: ARDL(1,1); $\phi_i$ ; $\beta_i$ .
Sheet 4: ARDL(1,1); $\phi$ ; $\beta_i$.

# References

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.