Seemingly Unrelated Regression Estimation for VAR Models with Explosive Roots*

Ye Chen⋆, Jian Li⋆⋆, Qiyuan Li⋆⋆⋆
⋆Capital University of Economics and Business
⋆⋆Huazhong Agricultural University
⋆⋆⋆Singapore Management University

March 16, 2023

This Online Supplement provides proofs of the main results and subsidiary Lemmas in Chen et al. (2022). In particular, Section A provide the proof for Preliminary Lemmas, Theorem 3.1, Corollary 2.2 and Corollary 3.2. Section B presents additional Tables for simulations. Section C introduces the method to test for explosiveness in the data. Section D states the asymptotics of IV estimator for VAR models with distinct explosive roots.

A Preliminary Lemmas and Some Proof

In this section, we introduce some basic limit results, which assist in deriving the limit distribution in this paper. These asymptotics are direct extensions of Phillips & Magdalinos (2007), Phillips & Magdalinos (2008), Magdalinos & Phillips (2009), and Phillips & Magdalinos (2013).

Lemma A.1 *Assuming the integer-valued sequence κ_n satisfying*

$$\sum_{i=1}^{\infty} n\rho_i^{-2\kappa_n} < \infty \text{ and } \sum_{n=1}^{\infty} n\rho_i^{-2n+2\kappa_n} < \infty \text{ as } n \to \infty,$$

under Assumptions 2.1 and 2.2, given the following data generating process for $i = 1, \ldots, k$ as $x_{i,t} = \rho_i x_{i,t-1} + u_{i,t}, \quad \rho_i > 1,$

*We thank the Editor Anindya Banerjee and two anonymous referees for very helpful comments on earlier versions of the paper, which improved the quality of the paper significantly. Chen acknowledges support from National Natural Science Foundation of China (No. 71803138), the Project of Construction and Support for high-level Innovative Teams of Beijing Municipal Institutions (BPHR20220119), and the Project of Cultivation for Young Top-notch Talents of Beijing Municipal Institutions (BPHR202203171). Li acknowledges support from National Natural Science Foundation of China (No.72173052 & No.71803058). Ye Chen, International School of Economics and Management, Capital University of Economics and Business; Email: chenye@cueb.edu.cn. Jian Li (Corresponding author), College of Economics and Management, Huazhong Agricultural University; E-mail: hzaulj@126.com. Qiyuan Li, School of Economics, Singapore Management University, Email: qyli.2019@phdecons.smu.edu.sg.

(i) Let $Q_n(\boldsymbol{\rho}) = [Q_{1,n}(\rho_1), \ldots, Q_{k,n}(\rho_k)]^\top$ with $Q_{i,n} := \sum_{i=1}^n \rho_i^{-t} \frac{u_{i,t}}{\sqrt{\sigma_{i,i}}}$ for $i = 1, \ldots, k$. We have $Q_n(\boldsymbol{\rho}) \Rightarrow Q(\boldsymbol{\rho}) = [Q_1(\rho_1), \ldots, Q_k(\rho_k)]^\top$, with $Q_i(\rho_i) = \sum_{i=1}^\infty \rho_i^{-t} \frac{u_{i,t}}{\sqrt{\sigma_{i,i}}}$. The subscript i of $x_i(\cdot)$ corresponds to $u_{i,t}$.

(ii) Let $\widetilde{Q}_n(\boldsymbol{\rho}) = [\widetilde{Q}_{1,n}(\rho_1), \ldots, \widetilde{Q}_{k,n}(\rho_k)]^\top$ with $\widetilde{Q}_{j,n}(\rho_i) = \sum_{t=1}^n \rho_i^{-(n-t)-1} \frac{u_{j,t}}{\sqrt{\sigma_{j,j}}}$ for $i, j = 1, \ldots, k$. We have $\widetilde{Q}_n(\boldsymbol{\rho}) \Rightarrow \widetilde{Q}(\boldsymbol{\rho}) = [\widetilde{Q}_1(\rho_1), \ldots, \widetilde{Q}_k(\rho_k)]^\top$, with $\widetilde{Q}_j(\rho_i) = \sum_{t=1}^\infty \rho_i^{-(n-t)-1} \frac{u_{j,t}}{\sqrt{\sigma_{j,j}}}$. The subscript j of $\widetilde{Q}_j(\cdot)$ corresponds to $u_{j,t}$.

(iii) $Q(\boldsymbol{\rho})$ and $\widetilde{Q}(\boldsymbol{\rho})$ are asymptotically independent.

Proof. (i) Following the argument in Lemma 1 of Phillips & Magdalinos (2013), by virtue of the martingale convergence theorem, we have the following almost sure convergence results:

$$
\begin{aligned}
Q_{i,n}(\rho_i) &= \sum_{i=1}^{\kappa_n} \rho_i^{-t} \frac{u_{i,t}}{\sqrt{\sigma_{ii}}} + \sum_{i=\kappa_n+1}^{n} \rho_i^{-t} \frac{u_{i,t}}{\sqrt{\sigma_{ii}}} \\
&= \sum_{i=1}^{\kappa_n} \rho_i^{-t} \frac{u_{i,t}}{\sqrt{\sigma_{ii}}} + o_{a.s}\left(\frac{1}{\sqrt{n}}\right) \to^{a.s.} \sum_{i=1}^{\infty} \rho_i^{-t} \frac{u_{i,t}}{\sqrt{\sigma_{ii}}} =: Q_i(\rho_i),
\end{aligned} \tag{A.1}
$$

since

$$
\max_{\kappa_n+1 \le t \le n} \left| \sum_{i=\kappa_n+1}^{n} \rho_i^{-t} \frac{u_{i,t}}{\sqrt{\sigma_{ii}}} \right| = o_{a.s}\left(\frac{1}{\sqrt{n}}\right),
$$

as $\sum_{i=1}^\infty n\rho_i^{-2\kappa_n} < \infty$ from Lemma 4.1 in Phillips & Magdalinos (2008). Hence, $Q_n(\boldsymbol{\rho}) \Rightarrow Q(\boldsymbol{\rho})$.

(ii) We first show $\max_{1\le t\le \kappa_n} \left|\sum_{t=1}^{\kappa_n} \rho_i^{-(n-t)-1} \frac{u_{j,t}}{\sqrt{\sigma_{j,j}}}\right| = o_{a.s}\left(\frac{1}{\sqrt{n}}\right)$ in the following. Follows the argument in Lemma 4.1 of Phillips & Magdalinos (2008), using Doob's inequality for martingales, we obtain

$$
\begin{aligned}
&\sum_{n=1}^{\infty} \mathbb{P}\left(\max_{1\le \kappa_n \le \infty} \left| \sum_{t=1}^{\kappa_n} \rho_i^{-(n-t)-1} \frac{u_{j,t}}{\sqrt{\sigma_{j,j}}} \right| > \frac{M}{\sqrt{n}} \right) \\
\le\ & \frac{1}{M^2} \sum_{n=1}^{\infty} n \mathbb{E}\left[\left| \sum_{t=1}^{\kappa_n} \rho_i^{-(n-t)-1} \frac{u_{j,t}}{\sqrt{\sigma_{j,j}}} \right|^2 \right] = \frac{\mathbb{E}\left[\left| \frac{u_{j,t}}{\sqrt{\sigma_{j,j}}} \right|^2 \right]}{M^2} \sum_{n=1}^{\infty} n \sum_{t=1}^{\kappa_n} \rho_i^{-2(n-t)-2} \\
=\ & M \sum_{n=1}^{\infty} n\rho_i^{-2n+2\kappa_n} < \infty
\end{aligned}
$$

from Assumption 2.1 and 2.2.

Hence, we obtain,

$$
\begin{aligned}
\widetilde{Q}_{j,n}(\rho_i) &= \sum_{t=1}^{\kappa_n} \rho_i^{-(n-t)-1} \frac{u_{j,t}}{\sqrt{\sigma_{j,j}}} + \sum_{i=\kappa_n+1}^{n} \rho_i^{-(n-t)-1} \frac{u_{j,t}}{\sqrt{\sigma_{j,j}}} \\
&= \sum_{i=\kappa_n+1}^{n} \rho_i^{-(n-t)-1} \frac{u_{j,t}}{\sqrt{\sigma_{j,j}}} + o_{a.s}\left(\frac{1}{\sqrt{n}}\right) \to^{a.s.} \sum_{i=1}^{\infty} \rho_i^{-(n-t)-1} \frac{u_{j,t}}{\sqrt{\sigma_{j,j}}} =: \widetilde{x}_j(\rho_i).
\end{aligned} \tag{A.2}
$$

Moreover, we have $Q_n(\boldsymbol{\rho}) \Rightarrow Q(\boldsymbol{\rho})$.

(iii) The asymptotical independence follows from the sample splitting argument, as stated in (A.1) and (A.2). ■

Lemma A.2 *Under Assumptions 2.1 and 2.2, given the following data generating process for* $i = 1,\ldots,k$ *as* $x_{i,t} = \rho_i x_{i,t-1} + u_{i,t}, \quad \rho_i > 1$, *we have,*

(i)

$$\frac{1}{\rho_i^n}\sum_{t=1}^{n} x_{i,t-1} = \frac{1}{\rho_i - 1}\left(\sum_{j=1}^{n}\rho_i^{-j}u_{i,t}\right) + o_p(1) \Rightarrow \frac{\sqrt{\sigma_{i,i}}Q_i(\rho_i)}{\rho_i - 1}, \tag{A.3}$$

(ii)

$$\begin{aligned}\frac{1}{\rho_i^n}\sum_{t=1}^{n} x_{i,t-1}u_{j,t} &= \left(\sum_{t=1}^{n}\rho_i^{-(n-t)-1}u_{j,t}\right)\left(\sum_{s=1}^{n}\rho_i^{-s}u_{i,s}\right) + o_p(1) \\ &\Rightarrow \sqrt{\sigma_{i,i}\sigma_{j,j}}Q_i(\rho_i)\widetilde{Q}_j(\rho_i),\end{aligned} \tag{A.4}$$

(iii)

$$\frac{1}{\rho_i^{2n}}\sum_{t=1}^{n} x_{i,t-1}^2 = \frac{\rho_i^{-2n}}{\rho_i^2 - 1}x_{i,n}^2 + O_p\left(\rho_i^{-2n}n\right) \Rightarrow \frac{\sigma_{i,i}Q_i(\rho_i)^2}{\rho_i^2 - 1}, \tag{A.5}$$

(iv)

$$\begin{aligned}\frac{1}{\rho_i^n\rho_j^n}\sum_{t=1}^{n} x_{i,t-1}x_{j,t-1} &= \frac{1}{\rho_i\rho_j - 1}\left(\sum_{t=1}^{n}\rho_i^{-t}u_{i,t}\right)\left(\sum_{t=1}^{n}\rho_j^{-t}u_{j,t}\right) \\ &\Rightarrow \frac{\sqrt{\sigma_{i,i}\sigma_{j,j}}}{\rho_i\rho_j - 1}Q_i(\rho_i)Q_j(\rho_j),\end{aligned} \tag{A.6}$$

(v) Consider the martingale array

$$U_n(s) = \left[\frac{1}{\rho_1^n}\sum_{t=1}^{n} x_{1,t-1}u_{j,t}, \ldots, \frac{1}{\rho_p^n}\sum_{t=1}^{n} x_{p,t-1}u_{j,t},\right].$$

Then the following joint convergence applies:

$$\begin{bmatrix} U_n(s) \\ \frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor ns\rfloor} u_t \end{bmatrix} \Rightarrow \begin{bmatrix} U(s) \\ B(s) \end{bmatrix}$$

for any $p \in (1,\ldots,k)$, *on the Skorokhod space* $\mathbb{D}_{\mathbb{R}^{p+k}}[0,1]$ *where* U *and* B *are independent Brownian motions with variance* $\sigma_{j,j}\sum_{t=1}^{\infty} R^{-(n-t)-1}\Sigma_u^\top \widetilde{Q}(\boldsymbol{\rho})\widetilde{Q}(\boldsymbol{\rho})^\top \Sigma_u R^{-(n-t)-1}$ *and* Σ_u *respectively.*

Proof. (i) For $\frac{1}{\rho_i^n}\sum_{t=1}^{n} x_{i,t-1}$. Since $\sum_{j=1}^{n}\rho_i^{-j}u_{i,t} \to^{a.s.} \sqrt{\sigma_{i,i}}Q_i(\rho_i)$ from Lemma A.1, $\frac{1}{\rho_i-1}\frac{x_{i,0}}{\rho_i^n} = o_p\left(\rho_i^{-n}\right)$ from Assumption 2.2, and $\frac{1}{\rho_i-1}\frac{\sum_{t=1}^{n}u_{i,t}}{\rho_i^n} = O_p\left(n^{1/2}\rho_i^{-n}\right)$ by central limit theorem, we obtain,

$$\begin{aligned}\frac{1}{\rho_i^n}\sum_{t=1}^{n} x_{i,t-1} &= \frac{1}{\rho_i - 1}\frac{x_{i,n}}{\rho_i^n} + \frac{1}{\rho_i - 1}\frac{x_{i,0}}{\rho_i^n} + \frac{1}{\rho_i - 1}\frac{\sum_{t=1}^{n}u_{i,t}}{\rho_i^n} \\ &= \frac{1}{\rho_i - 1}\left(\sum_{j=1}^{n}\rho_i^{-j}u_{i,t}\right) + o_p(1) \Rightarrow \frac{\sqrt{\sigma_{i,i}}Q_i(\rho_i)}{\rho_i - 1}.\end{aligned}$$

(ii) For $\frac{1}{\rho_i^n}\sum_{t=1}^n x_{i,t-1}u_{j,t}$. We have

$$\begin{aligned}
\frac{1}{\rho_i^n}\sum_{t=1}^n x_{i,t-1}u_{j,t} &= \frac{1}{\rho_i^n}\sum_{t=1}^n\sum_{s=1}^{t-1}\rho_i^{t-s-1}u_{i,s}u_{j,t} + \frac{1}{\rho_i^n}\sum_{t=1}^n \rho_i^t x_{i,0}u_{j,t} \\
&= \frac{1}{\rho_i^n}\sum_{t=1}^n\sum_{s=1}^{n}\rho_i^{t-s-1}u_{i,s}u_{j,t} - \frac{1}{\rho_i^n}\sum_{t=1}^n\sum_{s=t}^{n}\rho_i^{t-s-1}u_{i,s}u_{j,t} + \frac{1}{\rho_i^n}\sum_{t=1}^n \rho_i^t x_{i,0}u_{j,t} \\
&= \left(\sum_{s=1}^n \rho_i^{-s}u_{i,s}\right)\left(\sum_{t=1}^n \rho_i^{-(n-t)-1}u_{j,t}\right) + o_p(1) \\
&\Rightarrow \sqrt{\sigma_{i,i}\sigma_{j,j}}Q_i(\rho_i)\widetilde{Q}_j(\rho_i),
\end{aligned}$$

since (a)

$$\left|\frac{1}{\rho_i^n}\sum_{t=1}^n \rho_i^t x_{i,0}u_{j,t}\right| \le \left|\frac{1}{\rho_i^n}\sum_{t=1}^n \rho_i^t\right| |x_{i,0}| \max_t |u_{j,t}| = o_p(1)$$

from Assumption 2.1 and 2.2; (b)

$$\mathbb{E}\left[\frac{1}{\rho_i^n}\sum_{t=1}^n\sum_{j=t}^n \rho_i^{t-j-1}u_{i,t}u_{j,t}\right]^2 = \frac{1}{\rho_i^{2n}}\sum_{t=1}^n\sum_{j=t}^n \rho_i^{2(t-j-1)}\sigma_{i,i}\sigma_{j,j} = O_p\left(n\rho_i^{-2n}\right),$$

from Assumption 2.1.

(iii) For $\frac{1}{\rho_i^{2n}}\sum_{t=1}^n x_{i,t-1}^2$. We have

$$\begin{aligned}
\frac{1}{\rho_i^{2n}}\sum_{t=1}^n x_{i,t-1}^2 &= \frac{1}{\rho_i^2-1}\left[\rho_i^{-2n}\left(x_{i,n}^2 - x_{i,0}^2\right) - 2\rho_i^{-2n+1}\sum_{t=1}^n x_{i,t-1}u_{i,t} - \rho_i^{-2n}\sum_{t=1}^n u_{i,t}^2\right] \\
&= \frac{\rho_i^{-2n}}{\rho_i^2-1}x_{i,n}^2 + o_p\left(\rho_i^{-2n}\right) + O_p\left(\rho_i^{-n}\right) + O_p\left(\rho_i^{-2n}n\right) \\
&\Rightarrow \frac{\sigma_{i,i}Q_i(\rho_i)^2}{\rho_i^2-1},
\end{aligned}$$

since $x_{i,0} = o_p(1)$ from Assumption 2.2, $\sum_{t=1}^n x_{i,t-1}u_{j,t} = O_p(\rho_i^n)$ from (ii), and $\sum_{t=1}^n u_{i,t}^2 = O_p(n)$ by WLLN.

(iv) First, we have

$$\sum_{t=1}^n x_{i,t}x_{j,t} = \rho_i\rho_j\sum_{t=1}^n x_{i,t-1}x_{j,t-1} + \rho_j\sum_{t=1}^n x_{j,t-1}u_{i,t} + \rho_i\sum_{t=1}^n x_{i,t-1}u_{j,t} + \sum_{t=1}^n u_{i,t}u_{j,t},$$

which leads to

$$(1-\rho_i\rho_j)\sum_{t=1}^n x_{i,t-1}x_{j,t-1} = -x_{i,0}x_{j,0} + x_{i,n}x_{j,n} - \rho_j\sum_{t=1}^n x_{j,t-1}u_{i,t} - \rho_i\sum_{t=1}^n x_{i,t-1}u_{j,t} - \sum_{t=1}^n u_{i,t}u_{j,t}.$$

We show in the following that

$$\frac{(1-\rho_i\rho_j)}{\rho_i^n\rho_j^n}\sum_{t=1}^n x_{i,t-1}x_{j,t-1} = \frac{x_{i,n}x_{j,n}}{\rho_i^n\rho_j^n} + o_p(1) \Rightarrow \sqrt{\sigma_{i,i}\sigma_{j,j}}Q_i(\rho_i)Q_j(\rho_j). \tag{A.7}$$

since $\frac{x_{i,0}x_{j,0}}{\rho_i^n\rho_j^n} = o_p(\rho_{i,n}^{-n}\rho_j^{-n}) = o_p(1)$ from Assumption 2.2, $\frac{\sum_{t=1}^n x_{j,t-1}u_{i,t}}{\rho_i^n\rho_j^n} = O_p(\rho_i^{-n}) = o_p(1)$ and $\frac{\sum_{t=1}^n x_{i,t-1}u_{j,t}}{\rho_i^n\rho_j^n} = O_p(\rho_j^{-n}) = o_p(1)$ from (ii); and $\frac{\sum_{t=1}^n u_{i,t}u_{j,t}}{\rho_i^n\rho_j^n} = O_p(n^{1/2}\rho_i^{-n}\rho_j^{-n}) = o_p(1)$. The proof is therefore completed by noting that

$$\frac{x_{i,n}x_{j,n}}{\rho_i^n\rho_j^n} = \frac{x_{i,n}}{\rho_i^n}\frac{x_{j,n}}{\rho_j^n} \Rightarrow \sqrt{\sigma_{i,i}\sigma_{j,j}}Q_i(\rho_i)Q_j(\rho_j).$$

(v) Let $\mathcal{F}_{n,i} = \sigma(u_i, u_{i-1}, \ldots)$. Then we define

$$\begin{aligned}
\xi_{n,t} &= \left(\frac{1}{\rho_1^n}x_{1,t-1}u_{j,t}, \ldots, \frac{1}{\rho_p^n}x_{p,t-1}u_{j,t}, , \frac{1}{\sqrt{n}}u_t\right) \\
&= \left(\frac{1}{n^{\alpha/2}}\left(\rho_1^{-(n-t)-1}u_{j,t}\right)\left(\sum_{s=1}^n \rho_1^{-s}u_{1,s}\right), \ldots, \frac{1}{n^{\alpha/2}}\left(\rho_p^{-(n-t)-1}u_{j,t}\right)\left(\sum_{s=1}^n \rho_p^{-s}u_{p,s}\right), \frac{1}{\sqrt{n}}u_t\right) \\
&= \left(\frac{1}{n^{\alpha/2}}\left(\rho_1^{-(n-t)-1}u_{j,t}\right)Y_1(\rho_{1n}), \ldots, \frac{1}{n^{\alpha/2}}\left(\rho_k^{-(n-t)-1}u_{j,t}\right)Y_p(\rho_{pn}), \frac{1}{\sqrt{n}}u_t\right) + o_p(1),
\end{aligned}$$

which is a $\mathbb{R}^{p+k}$ -valued martingale difference array with respect to $\mathcal{F}_{n,t}$. We further denote $M_{n,j} := \sum_{t=1}^j \xi_{n,t}$, whose conditional variance is given by

$$\langle M_n\rangle_j := \sum_{t=1}^n \mathbb{E}[\xi_{n,t}\xi_{n,t}^\top|\mathcal{F}_{n,t-1}] \to^p \begin{pmatrix} \sigma_{j,j}\sum_{t=1}^\infty R^{-(n-t)-1}\Sigma_u^\top\widetilde{Q}(\boldsymbol{\rho})\widetilde{Q}(\boldsymbol{\rho})^\top\Sigma_u R^{-(n-t)-1} & 0 \\ 0 & \Sigma_u \end{pmatrix},$$

with $R = diag(\rho_1, \ldots, \rho_p)$, which is a $\mathbb{R}^{2k}$ -valued martingale difference array with respect to $\mathcal{F}_{n,t}$. The conditional variance matrix implies that U is independent from B. ■

Lemma A.3 *Under Assumptions 2.1 and 2.2, given the following data generating process for $i = 1, \ldots, k$ as*

$$x_{i,t} = \mu_i + \rho_i x_{i,t-1} + u_{i,t}, \quad \rho_i > 1, \tag{A.8}$$

where $\mu_i = \widetilde{\mu}_i n^{-\eta_i}$ with $\eta_i \geq 0$. We also have result (i)-(iv) in Lemma A.2.

Proof. (i) From the equation (A.8), we rewrite it as

$$\begin{aligned}
x_{i,t} &= \mu_i\sum_{j=0}^t \rho_i^j + \widetilde{x}_{i,t}, \\
\widetilde{x}_{i,t} &= \rho_i\widetilde{x}_{i,t-1} + u_{i,t},.
\end{aligned}$$

Therefore, we rewrite $\frac{1}{\rho_i^n}\sum_{t=1}^n x_{i,t-1}$ as

$$\begin{aligned}
\frac{1}{\rho_i^n}\sum_{t=1}^n x_{i,t-1} &= \frac{1}{\rho_i^n}\sum_{t=1}^n\left(\mu_i\sum_{j=0}^{t-1}\rho_i^j + \widetilde{x}_{i,t-1}\right) = \frac{\mu_i}{\rho_i^n}\sum_{t=1}^n\sum_{j=0}^{t-1}\rho_i^j + \frac{1}{\rho_i^n}\sum_{t=1}^n\widetilde{x}_{i,t-1} \\
&= O_p\left(n^{-\eta_i}\rho_i^{-n}\right) + \frac{1}{\rho_i^n}\sum_{t=1}^n\widetilde{x}_{i,t-1}. \qquad (A.9)
\end{aligned}$$

The equation (A.9) converges to $\frac{\sqrt{\sigma_{i,i}}Q_i(\rho_i)}{\rho_i - 1}$ from the result of (A.3).

(ii) Since

$$\frac{1}{\rho_i^n}\sum_{t=1}^{n}\left(\mu_i\left(\sum_{l=0}^{t-1}\rho_i^l\right)u_{j,t}\right)=\frac{\mu_i\rho_i}{(1-\rho_i)}\sum_{t=1}^{n}\left(\frac{(1-\rho_i^t)}{\rho_i^n}\mu_i\right)=O_p\left(n^{\frac{1}{2}-\eta_i}\rho_i^{-n}\right),$$

We robtain

$$\begin{aligned}\frac{1}{\rho_i^n}\sum_{t=1}^{n}x_{i,t-1}u_{j,t} &= \frac{1}{\rho_i^n}\sum_{t=1}^{n}\left(\mu_i\left(\sum_{l=0}^{t-1}\rho_i^l\right)u_{j,t}\right)+\frac{1}{\rho^n}\sum_{t=1}^{n}\widetilde{x}_{i,t-1}u_{j,t}\\ &\Rightarrow \sqrt{\sigma_{i,i}\sigma_{j,j}}Q_j(\rho_i)\widetilde{Q}_i(\rho_i)\end{aligned}$$

(iii) We rewrite $\frac{1}{\rho_i^{2n}}\sum_{t=1}^{n}x_{i,t-1}^2$ as

$$\begin{aligned}&\frac{1}{\rho_i^{2n}}\sum_{t=1}^{n}x_{i,t-1}^2\\ =\ &\frac{1}{\rho_i^{2n}}\sum_{t=1}^{n}\left\{\mu_i^2\left(\sum_{j=0}^{t-1}\rho_i^j\right)^2+2\mu\left(\sum_{j=0}^{t-1}\rho_i^j\right)\widetilde{x}_{i,t-1}+\widetilde{x}_{i,t-1}^2\right\}\\ =\ &\frac{\mu_i^2\sum_{t=1}^{n}\left(1-2\rho_i^t+\rho_i^{2t}\right)}{(1-\rho_i)^2\rho_i^{2n}}+\frac{2\mu_i\sum_{t=1}^{n}(1-\rho_i^t)\,\widetilde{x}_{i,t-1}}{(1-\rho_i)\,\rho_i^{2n}}+\frac{\sum_{t=1}^{n}\widetilde{x}_{i,t-1}^2}{\rho_i^{2n}}\end{aligned}\tag{A.10}$$

The first term on the right hand side of equation (A.10) is $o_p(1)$ since

$$\frac{\mu_i^2\sum_{t=1}^{n}\left(1-2\rho_i^t+\rho_i^{2t}\right)}{(1-\rho_i)^2\rho_i^{2n}}=\frac{\mu_i^2\left(n-2\frac{\rho_i(1-\rho_i^n)}{1-\rho_i}+\frac{\rho_i^2\left(1-\rho_i^{2n}\right)}{(1-\rho_i)^2}\right)}{(1-\rho_i)^2\rho_i^{2n}}=O_p\left(n^{-2\eta_i}\right)=o_p(1).$$

Second, from (A.9), we have

$$\frac{\mu_i\sum_{t=1}^{n}\widetilde{x}_{i,t-1}}{\rho_i^{2n}}=O_p\left(n^{-\eta_i}\rho_i^{-n}\right)=o_p(1).\tag{A.11}$$

Further, we have

$$\mu_i\frac{\sum_{t=1}^{n}\rho_i^t\widetilde{x}_{i,t-1}}{\rho_i^{2n}}\le\mu_i\max_t|\widetilde{x}_{i,t-1}|\frac{\sum_{t=1}^{n}\rho_i^t}{\rho_i^{2n}}=O_p\left(n^{-\eta_i}\right)=O_p(1).\tag{A.12}$$

Hence, combined the equation (A.11) and (A.12), we can show that the second term on the right hand side of equation (A.10) is $o_p(1)$.

Third, using the result from Lemma A.2 (iii), we complete the proof.

(iv) The result can be proved by the same argument from (iii). ∎

Lemma A.4 *Under Assumptions 3.1 and 3.2, given the following data generating process for* $\alpha\in(0,1)$, *and* $c_i>0$ *for* $i=1,\ldots,k$ *as* $x_{i,t}=\rho_{in}x_{i,t-1}+u_{i,t},\rho_{in}=1+\frac{c_i}{n^\alpha}$, *we have:*

(i) For all $i,j=1,\ldots,k$, *the following joint convergence applies:*

$$\left(\frac{1}{n^{\alpha/2}}\sum_{t=1}^{n}\rho_{in}^{-t}\frac{u_{i,t}}{\sqrt{\sigma_{j,j}}},\frac{1}{n^{\alpha/2}}\sum_{t=1}^{n}\rho_{in}^{-(n-t)-1}\frac{u_{j,t}}{\sqrt{\sigma_{j,j}}}\right)\Rightarrow\left(\sqrt{\frac{\sigma_{i,i}}{2c_i}}Y_i\left(c_i\right),\sqrt{\frac{\sigma_{j,j}}{2c_i}}\widetilde{Y}_j\left(c_i\right)\right),$$

where $Y_i(c_i) \equiv \mathcal{N}(0,1)$ and $\widetilde{Y}_j(c_i) \equiv \mathcal{N}(0,1)$, and they are independent.

(ii)

$$\frac{1}{n^{3\alpha/2}\rho_{in}^{n}}\sum_{t=1}^{n} x_{i,t-1} = \frac{n^{\alpha}}{c_i}\frac{1}{n^{\alpha}}\frac{1}{\sqrt{n^{\alpha}}}\sum_{t=1}^{n}\rho_{in}^{-j}u_{i,t} + o_p(1) \Rightarrow \frac{1}{c_i}\sqrt{\frac{\sigma_{i,i}}{2c_i}}Y_i(c_i) \tag{A.13}$$

$$\begin{aligned}\frac{1}{n^{\alpha}\rho_{in}^{n}}\sum_{t=1}^{n} x_{i,t-1}u_{j,t} &= \left(\sum_{t=1}^{n}\rho_i^{-(n-t)-1}u_{j,t}\right)\left(\sum_{s=1}^{n}\rho_i^{-s}u_{i,s}\right) + o_p(1) \\ &\Rightarrow \frac{\sqrt{\sigma_{i,i}\sigma_{j,j}}}{2c_i}Y_i(c_i)\widetilde{Y}_j(c_i)\end{aligned} \tag{A.14}$$

$$\frac{1}{n^{2\alpha}\rho_{in}^{2n}}\sum_{t=1}^{n} x_{i,t-1}^{2} = \frac{1}{2c_i}\left(\frac{1}{\sqrt{n^{\alpha}}}\sum_{t=1}^{n}\rho_{in}^{-j}u_{i,t}\right)^2 + o_p(1) \Rightarrow \frac{\sigma_{i,i}}{4c_i^2}\left[Y_i(c_i)\right]^2 \tag{A.15}$$

$$\begin{aligned}\frac{1}{n^{2\alpha}\rho_{in}^{n}\rho_{jn}^{n}}\sum_{t=1}^{n} x_{i,t-1}x_{j,t-1} &= \frac{1}{c_i+c_j}\left(\frac{1}{n^{\alpha/2}}\sum_{t=1}^{n}\rho_{in}^{-t}u_{i,t}\right)\left(\frac{1}{n^{\alpha/2}}\sum_{t=1}^{n}\rho_{jn}^{-t}u_{j,t}\right) + o_p(1) \\ &\Rightarrow \frac{1}{2(c_i+c_j)}\sqrt{\frac{\sigma_{i,i}\sigma_{j,j}}{c_ic_j}}Y_i(c_i)Y_j(c_j)\end{aligned} \tag{A.16}$$

(iii) Let $\widetilde{Y}(\boldsymbol{\rho}) = [\frac{1}{\sqrt{2c_1}}\widetilde{Y}_j(\rho_1), \ldots, \frac{1}{\sqrt{2c_p}}\widetilde{Y}_j(\rho_p)]^{\top}$. Consider the martingale array $U_n(s) = \left[\frac{1}{\rho_{1n}^{n}}\sum_{t=1}^{n} x_{1,t-1}u_{j,t}, \ldots, \frac{1}{\rho_{kn}^{n}}\sum_{t=1}^{n} x_{k,t-1}u_{j,t},\right]$, then the following joint convergence applies:

$$\begin{bmatrix} U_n(s) \\ \frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor ns\rfloor} u_t \end{bmatrix} \Rightarrow \begin{bmatrix} U(s) \\ B(s) \end{bmatrix}$$

for any $p \in (1, \ldots, k)$, on the Skorokhod space $\mathbb{D}_{\mathbb{R}^{p+k}}[0,1]$ where U and B are independent Brownian motions with variance $\sigma_{j,j}\int_0^{\infty} e^{-sC}\Sigma_u^{\top}\widetilde{Y}(\boldsymbol{\rho})\widetilde{Y}(\boldsymbol{\rho})^{\top}\Sigma_u e^{-sC}ds$ and Σ_u respectively.

Proof. (i) Follows the argument from Lemma 4.1 of Magdalinos & Phillips (2009), we obtain

$$\frac{1}{n^{\alpha/2}}\sum_{t=1}^{n}\rho_{in}^{-t}u_{i,t} \Rightarrow Y_i(c_i) = \frac{\sigma_{i,i}}{\sqrt{2c_i}}\mathcal{N}(0,1),$$

and

$$\frac{1}{n^{\alpha/2}}\sum_{t=1}^{n}\rho_{in}^{-(n-t)-1}u_{j,t} \Rightarrow \widetilde{Y}_j(c_i) = \frac{\sigma_{j,j}}{\sqrt{2c_i}}\mathcal{N}(0,1)$$

$Y_i(c_i)$ and $\widetilde{Y}_j(c_i)$ are asymptotically independent since

$$\mathbb{E}\left[\left(\frac{1}{n^{\alpha/2}}\sum_{t=1}^{n}\rho_{in}^{-t}u_{i,t}\right)\left(\frac{1}{n^{\alpha/2}}\sum_{t=1}^{n}\rho_{in}^{-(n-t)-1}u_{j,t}\right)\right] = \frac{\rho_n^{-n+1}}{n^{\alpha}}\sum_{j=1}^{n}\mathbb{E}\left[u_{i,t}u_{j,t}\right] \to 0.$$

(ii) The results follow from similar arguments as in the proof of Lemma A.4.

(iii) Let $\widetilde{Y}_n(\boldsymbol{\rho}) = \left(\frac{1}{\rho_1^{n}}\sum_{i=1}^{\infty}\rho_1^{-(n-t)-1}\frac{u_{j,t}}{\sqrt{\sigma_{j,j}}}, \ldots, \frac{1}{\rho_p^{n}}\sum_{i=1}^{\infty}\rho_p^{-(n-t)-1}\frac{u_{j,t}}{\sqrt{\sigma_{j,j}}}\right)^{\top}$. From (ii) and (iii), we have

$$\widetilde{Y}_n(\boldsymbol{\rho}) \Rightarrow \widetilde{Y}(\boldsymbol{\rho}).$$

The remaining proof follows the same argument from (Magdalinos & Phillips (2009), Proposition A1). Let $\mathcal{F}_{n,i} = \sigma(u_i, u_{i-1}, \ldots)$. Define

$$\begin{aligned}
\xi_{n,t} &= \left(\frac{1}{n^{\alpha/2}\rho_{1n}^n} x_{1,t-1}u_{j,t}, \ldots, \frac{1}{n^{\alpha/2}\rho_{pn}^n} x_{p,t-1}u_{j,t}, , \frac{1}{\sqrt{n}} u_t\right) \\
&= \left(\frac{1}{n^{\alpha/2}}\left(\rho_{1n}^{-(n-t)-1}u_{j,t}\right)\left(\sum_{s=1}^n \rho_{1n}^{-s}u_{1,s}\right), \ldots, \frac{1}{n^{\alpha/2}}\left(\rho_{pn}^{-(n-t)-1}u_{j,t}\right)\left(\sum_{s=1}^n \rho_{pn}^{-s}u_{p,s}\right), \frac{1}{\sqrt{n}}u_t\right) \\
&= \left(\frac{1}{n^{\alpha/2}}\left(\rho_{1n}^{-(n-t)-1}u_{j,t}\right)Y_1(\rho_{1n}), \ldots, \frac{1}{n^{\alpha/2}}\left(\rho_{pn}^{-(n-t)-1}u_{j,t}\right)Y_p(\rho_{pn}), \frac{1}{\sqrt{n}}u_t\right) + o_p(1),
\end{aligned}$$

which is a $\mathbb{R}^{2k}$-valued martingale difference array with respect to $\mathcal{F}_{n,t}$. We further denote $M_{n,j} := \sum_{t=1}^j \xi_{n,t}$, whose conditional variance is given by

$$\langle M_n\rangle_j := \sum_{t=1}^n \mathbb{E}[\xi_{n,t}\xi_{n,t}^\top|\mathcal{F}_{n,t-1}] \to^p \begin{pmatrix} \sigma_{j,j}\int_0^\infty e^{-sC}\Sigma_u^\top \widetilde{Y}(\boldsymbol{\rho})\widetilde{Y}(\boldsymbol{\rho})^\top \Sigma_u e^{-sC}ds & 0 \\ 0 & \Sigma_u \end{pmatrix},$$

The conditional variance matrix implies that U is independent from B. The conditional Lindeberg condition can be verified using the argument from (Magdalinos & Phillips (2009), Proposition A1). ■

Lemma A.5 *Under Assumption 3.1 and 3.2, given the following data generating process for $\alpha \in (0,1)$, and $c_i > 0$ for $i = 1, \ldots, k$ as*

$$x_{i,t} = \mu + \rho_{in}x_{i,t-1} + u_{i,t}, \rho_{in} = 1 + \frac{c_i}{n^\alpha} \tag{A.17}$$

where $\mu_i = \widetilde{\mu}_i n^{-\eta}$ with $\eta \geq 0$, we also have result (i)-(iv) in Lemma A.4.

Proof. The proof using the similar argument as in Lemma A.3. ■

A.1 Proof of Theorem 3.1

Proof. The proof is analogous to the proof of Theorem 2.1. Except that, now we are using Lemma A.4 instead of A.1. Normalized by $\frac{1}{n^{2\alpha}\rho_{in}^n\rho_{jn}^n}$, B_{ij} has the following asymptotics,

$$\frac{1}{n^{2\alpha}\rho_{in}^n\rho_{jn}^n}B_{ij} = \frac{1}{n^{2\alpha}\rho_{in}^n\rho_{jn}^n}\sigma^{ij}\sum_{t=1}^n x_{i,t-1}x_{jt-1} \Rightarrow \frac{\sigma^{i,j}}{2(c_i+c_j)}\sqrt{\frac{\sigma_{i,i}\sigma_{j,j}}{c_ic_j}}Y_i(\rho_{in})Y_j(\rho_{jn}).$$

Therefore the denominator has the following asymptotics,

$$\begin{aligned}
\left(\frac{1}{n^{2\alpha}}\right)^k\left(\prod_{j=1}^k \rho_{jn}^n\right)^2 |B| &\Rightarrow \left|\begin{bmatrix} \frac{\sigma^{1,1}}{2(c_1+c_1)}\sqrt{\frac{\sigma_{1,1}\sigma_{1,1}}{c_1c_1}}Y_1(c_1)Y_1(c_1) & \cdots & \frac{\sigma^{1,k}}{2(c_1+c_k)}\sqrt{\frac{\sigma_{1,1}\sigma_{k,k}}{c_1c_k}}Y_1(c_1)Y_k(c_k) \\ \vdots & \ddots & \vdots \\ \frac{\sigma^{k,1}}{2(c_k+c_1)}\sqrt{\frac{\sigma_{k,k}\sigma_{1,1}}{c_kc_1}}Y_k(c_k)Y_1(c_1) & \cdots & \frac{\sigma^{k,k}}{2(c_k+c_k)}\sqrt{\frac{\sigma_{k,k}\sigma_{k,k}}{c_kc_k}}Y_k(c_k)Y_k(c_k) \end{bmatrix}\right| \\
&= \left(\prod_{j=1}^k \frac{\sigma_{j,j}}{2c_j}Y_j(\rho_{jn})^2\right)\zeta_0(\boldsymbol{c}),
\end{aligned} \tag{A.18}$$

where

$$\zeta_0(\boldsymbol{c}) := \left| \begin{bmatrix} \frac{\sigma^{1,1}}{c_1+c_2} & \cdots & \frac{\sigma^{1,k}}{c_1+c_k} \\ \vdots & \ddots & \vdots \\ \frac{\sigma^{k,1}}{c_k+c_1} & \cdots & \frac{\sigma^{k,k}}{c_k+c_k} \end{bmatrix} \right|. \tag{A.19}$$

The j-th summand in the numerator of (A.9) has the following asymptotics

$$\begin{aligned} &\left(\prod_{\ell \neq i} \frac{1}{n^\alpha \rho_{\ell n}^n}\right)^2 \frac{1}{n^\alpha \rho_{jn}^n} \left| \begin{bmatrix} B_{1,1} & \cdots & \sigma^{1,j} \sum_{t=1}^n x_{1,t-1} u_{j,t} & \cdots & B_{1,k} \\ \vdots & & \vdots & & \vdots \\ B_{k,1} & \cdots & \sigma^{k,j} \sum_{t=1}^n x_{k,t-1} u_{j,t} & \cdots & B_{k,k} \end{bmatrix} \right| \\ \Rightarrow & \left| \begin{bmatrix} \frac{\sigma^{1,1}}{2(c_1+c_1)} \sqrt{\frac{\sigma_{1,1}\sigma_{1,1}}{c_1 c_1}} Y_1(c_1) Y_1(c_1) & \cdots & \frac{\sqrt{\sigma_{1,1}\sigma_{j,j}}}{2c_1} \sigma^{1,j} Y_1(c_1) \widetilde{Y}_j(c_1) & \cdots & \frac{\sigma^{1,k}}{2(c_1+c_k)} \sqrt{\frac{\sigma_{1,1}\sigma_{k,k}}{c_1 c_k}} Y_1(\rho_{1n}) Y_k(c_k) \\ \vdots & & \vdots & & \vdots \\ \frac{\sigma^{k,1}}{2(c_k+c_1)} \sqrt{\frac{\sigma_{k,k}\sigma_{1,1}}{c_k c_1}} Y_k(c_k) Y_1(c_1) & \cdots & \frac{\sqrt{\sigma_{k,k}\sigma_{j,j}}}{2c_k} \sigma^{k,j} Y_k(c_k) \widetilde{Y}_j(c_k) & \cdots & \frac{\sigma^{k,k}}{2(c_k+c_k)} \sqrt{\frac{\sigma_{k,k}\sigma_{k,k}}{c_k c_k}} Y_k(c_k) Y_k(c_k) \end{bmatrix} \right| \\ = & \left(\prod_{j=1}^k \frac{\sigma_{j,j}}{2c_j} Y_j(\rho_{jn})^2\right) \sqrt{\frac{\sigma_{j,j}}{\sigma_{i,i}}} \frac{1}{Y_i(\rho_{in})} \left| \begin{bmatrix} \frac{\sigma^{1,1}}{c_1+c_1} & \cdots & \frac{\sigma^{1,j}}{c_1} \widetilde{Y}_j(c_i) & \cdots & \frac{\sigma^{1,k}}{c_1+c_k} \\ \vdots & & \vdots & & \vdots \\ \frac{\sigma^{k,1}}{c_k+c_1} & \cdots & \frac{\sigma^{k,j}}{c_k} \widetilde{Y}_j(c_k) & \cdots & \frac{\sigma^{k,k}}{c_k+c_k} \end{bmatrix} \right|. \end{aligned} \tag{A.20}$$

Combine the the result from (A.18) and (A.20), we obtain

$$n^\alpha \rho_{in}^n \left(\widehat{\rho}_{i,SUR} - \rho_{in}\right) \Rightarrow \frac{\zeta_i(\boldsymbol{c})}{\zeta_0(\boldsymbol{c})},$$

where

$$\zeta_0(\boldsymbol{c}) := \sum_{j=1}^k \sqrt{\frac{\sigma_{j,j}}{\sigma_{i,i}}} \frac{1}{Y_i(\rho_{in})} \left| \begin{bmatrix} \frac{\sigma^{1,1}}{c_1+c_1} & \cdots & \frac{\sigma^{1,j}}{c_1} \widetilde{Y}_j(c_i) & \cdots & \frac{\sigma^{1,k}}{c_1+c_k} \\ \vdots & & \vdots & & \vdots \\ \frac{\sigma^{k,1}}{c_k+c_1} & \cdots & \frac{\sigma^{k,j}}{c_k} \widetilde{Y}_j(c_k) & \cdots & \frac{\sigma^{k,k}}{c_k+c_k} \end{bmatrix} \right|. \tag{A.21}$$

■

A.2 Proof of Corollary 2.2

Proof. Let

$$B := \begin{bmatrix} n\sigma^{1,1} & \sigma^{1,1} \sum_{t=1}^n x_{1,t-1} & \cdots & n\sigma^{1,k} & \sigma^{1,k} \sum_{t=1}^n x_{k,t-1} \\ \sigma^{1,1} \sum_{t=1}^n x_{1,t-1} & \sigma^{1,1} \sum_{t=1}^n x_{1,t-1}^2 & \cdots & \sigma^{1,k} \sum_{t=1}^n x_{1,t-1} & \sigma^{1,k} \sum_{t=1}^n x_{1,t-1} x_{k,t-1} \\ n\sigma^{2,1} & \sigma^{2,1} \sum_{t=1}^n x_{1,t-1} & \cdots & n\sigma^{2,k} & \sigma^{2,k} \sum_{t=1}^n x_{k,t-1} \\ \sigma^{2,1} \sum_{t=1}^n x_{2,t-1} & \sigma^{2,1} \sum_{t=1}^n x_{2,t-1} x_{1,t-1} & \cdots & \sigma^{2,k} \sum_{t=1}^n x_{2,t-1} & \sigma^{2,k} \sum_{t=1}^n x_{2,t-1} x_{k,t-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ n\sigma^{k,1} & \sigma^{k,1} \sum_{t=1}^n x_{1,t-1} & \cdots & n\sigma^{k,k} & \sigma^{k,k} \sum_{t=1}^n x_{k,t-1} \\ \sigma^{k,1} \sum_{t=1}^n x_{k,t-1} & \sigma^{k,1} \sum_{t=1}^n x_{k,t-1} x_{1,t-1} & \cdots & \sigma^{k,k} \sum_{t=1}^n x_{k,t-1} & \sigma^{k,k} \sum_{t=1}^n x_{k,t-1}^2 \end{bmatrix},$$

and

$$D := \begin{bmatrix} \sigma^{1,1}\sum_{t=1}^n x_{1,t} + \sigma^{1,2}\sum_{t=1}^n x_{2,t} + \cdots + \sigma^{1,k}\sum_{t=1}^n x_{k,t} \\ \sigma^{1,1}\sum_{t=1}^n x_{1,t-1}x_{1,t} + \sigma^{1,2}\sum_{t=1}^n x_{1,t-1}x_{2,t} + \cdots + \sigma^{1,k}\sum_{t=1}^n x_{1,t-1}x_{k,t} \\ \sigma^{2,1}\sum_{t=1}^n x_{1,t} + \sigma^{2,2}\sum_{t=1}^n x_{2,t} + \cdots + \sigma^{2,k}\sum_{t=1}^n x_{k,t} \\ \sigma^{2,1}\sum_{t=1}^n x_{2,t-1}x_{1,t} + \sigma^{2,2}\sum_{t=1}^n x_{2,t-1}x_{2,t} + \cdots + \sigma^{2,k}\sum_{t=1}^n x_{2,t-1}x_{k,t} \\ \vdots \\ \sigma^{k,1}\sum_{t=1}^n x_{1,t} + \sigma^{k,2}\sum_{t=1}^n x_{2,t} + \cdots + \sigma^{k,k}\sum_{t=1}^n x_{k,t} \\ \sigma^{k,1}\sum_{t=1}^n x_{k,t-1}x_{1,t} + \sigma^{k,2}\sum_{t=1}^n x_{k,t-1}x_{2,t} + \cdots + \sigma^{k,k}\sum_{t=1}^n x_{k,t-1}x_{k,t} \end{bmatrix}_{2k\times 1}.$$

Then the SUR estimator of A is

$$\begin{aligned} \widehat{A}_{SUR} &= [X_-^\top(\Sigma_u \otimes I_n)^{-1}X_-]^{-1}[X_-^\top(\Sigma_u \otimes I_n)^{-1}X] \\ &= B_{2k\times 2k}^{-1}D_{2k\times 1} = \frac{1}{|B|}B_{2k\times 2k}^{*}D_{2k\times 1}, \end{aligned}$$

where B^* denotes adjoint matrix of B.

From Lemma A.3 (i) and (iii), we have the limiting result for $|B|$ such that,

$$\begin{aligned} &\left(\frac{1}{\prod_{j=1}^k \rho_j^n}\right)^2 \left(\frac{1}{n}\right)^k |B| \\ \Rightarrow\ &\left|\begin{bmatrix} \sigma^{1,1} & O_p\left(n^{-1}\right) & \cdots & \sigma^{1,k} & O_p\left(n^{-1}\right) \\ \sigma^{1,1}\frac{\sqrt{\sigma_{1,1}}Q_1(\rho_1)}{\rho_1-1} & \frac{\sigma^{1,1}\sigma_{1,1}Q_1(\rho_1)^2}{\rho_1^2-1} & \cdots & \sigma^{1,k}\frac{\sqrt{\sigma_{1,1}}Q_1(\rho_1)}{\rho_1-1} & \frac{\sigma^{1,k}\sqrt{\sigma_{1,1}\sigma_{k,k}}Q_1(\rho_1)Q_k(\rho_k)}{\rho_1\rho_k-1} \\ \sigma^{2,1} & O_p\left(n^{-1}\right) & \cdots & \sigma^{2,k} & O_p\left(n^{-1}\right) \\ \sigma^{2,1}\frac{\sqrt{\sigma_{2,2}}Q_2(\rho_2)}{\rho_2-1} & \frac{\sigma^{2,1}\sqrt{\sigma_{2,2}\sigma_{1,1}}Q_2(\rho_2)Q_1(\rho_1)}{\rho_2\rho_1-1} & \cdots & \sigma^{2,k}\frac{\sqrt{\sigma_{2,2}}Q_2(\rho_2)}{\rho_2-1} & \frac{\sigma^{2,k}\sqrt{\sigma_{2,2}\sigma_{k,k}}Q_2(\rho_2)Q_k(\rho_k)}{\rho_2\rho_k-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma^{k,1} & O_p\left(n^{-1}\right) & \cdots & \sigma^{k,k} & O_p\left(n^{-1}\right) \\ \sigma^{k,1}\frac{Q_k(\rho_k)\sqrt{\sigma_{k,k}}}{\rho_k-1} & \frac{\sigma^{k,1}\sqrt{\sigma_{k,k}\sigma_{1,1}}Q_k(\rho_k)Q_1(\rho_1)}{\rho_k\rho_1-1} & \cdots & \sigma^{k,k}\frac{\sqrt{\sigma_{k,k}}Q_k(\rho_k)}{\rho_k-1} & \frac{\sigma^{k,k}\sigma_{k,k}Q_k(\rho_k)^2}{\rho_k^2-1} \end{bmatrix}\right| \\ =\ &\left(\prod_{j=1}^k \sigma_{j,j}Q_j\left(\rho_j\right)^2\right)\frac{\xi_0(\boldsymbol{\rho})}{|\Sigma_u|}. \end{aligned} \tag{A.22}$$

Note that, the $(2i-1)$ -th term of $\widehat{A}_{SUR}$ is

$$\begin{aligned} \hat{\mu}_{i,SUR} &= \frac{1}{|B|}\begin{bmatrix}0 & 0 & \cdots & \underset{\substack{\uparrow \\ 2i-1}}{1} & \cdots 0\end{bmatrix} B_{2k\times 2k}^{*}D_{2k\times 1} \\ &= \frac{1}{|B|}\times\left|\begin{bmatrix} B_{1,1} & B_{1,2} & \cdots & B_{1,2i-2} & D_1 & B_{1,2i} & \cdots & B_{1,2k} \\ B_{2,1} & B_{2,2} & \cdots & B_{2,2i-2} & D_2 & B_{2,2i} & \cdots & B_{2,2k} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ B_{2k,1} & B_{2k,2} & \cdots & B_{2k,2i-2} & D_{2k} & B_{2k,2i} & \cdots & B_{2k,2k} \end{bmatrix}\right|. \end{aligned}$$

We rewrite $\sum_{t=1}^{n} x_{jt}$ as

$$\sum_{t=1}^{n} x_{j,t} = \sum_{t=1}^{n} (\mu_j + \rho_j x_{j,t-1} + u_{j,t}) = n\mu_j + \rho_j \sum_{t=1}^{n} x_{j,t-1} + \sum_{t=1}^{n} u_{j,t},$$

and then we are able to expand the vector $D = [D_1, \ldots, D_k]$ as

$$\begin{aligned} & D_{2k\times 1} \\ = \; & \mu_1 \begin{bmatrix} B_{1,1} \\ B_{2,1} \\ \vdots \\ B_{2k,1} \end{bmatrix} + \rho_1 \begin{bmatrix} B_{1,2} \\ B_{2,2} \\ \vdots \\ B_{2k,2} \end{bmatrix} + \cdots + \mu_k \begin{bmatrix} B_{1,2k-1} \\ B_{2,2k-1} \\ \vdots \\ B_{2k,2k-1} \end{bmatrix} + \rho_k \begin{bmatrix} B_{1,2k} \\ B_{2,2k} \\ \vdots \\ B_{2k,2k} \end{bmatrix} \\ & + \begin{bmatrix} \sigma^{1,1} \sum_{t=1}^{n} u_{1,t} + \sigma^{1,2} \sum_{t=1}^{n} u_{2,t} + \cdots + \sigma^{1,k} \sum_{t=1}^{n} u_{k,t} \\ \sigma^{1,1} \sum_{t=1}^{n} x_{1,t-1} u_{1,t} + \sigma^{1,2} \sum_{t=1}^{n} x_{1,t-1} u_{2,t} + \cdots + \sigma^{1,k} \sum_{t=1}^{n} x_{1,t-1} u_{k,t} \\ \sigma^{2,1} \sum_{t=1}^{n} u_{1,t} + \sigma^{22} \sum_{t=1}^{n} u_{2,t} + \cdots + \sigma^{2,k} \sum_{t=1}^{n} u_{k,t} \\ \sigma^{2,1} \sum_{t=1}^{n} x_{2,t-1} u_{1,t} + \sigma^{2,2} \sum_{t=1}^{n} x_{2,t-1} u_{2,t} + \cdots + \sigma^{2,k} \sum_{t=1}^{n} x_{2,t-1} u_{k,t} \\ \vdots \\ \sigma^{k,1} \sum_{t=1}^{n} u_{1,t} + \sigma^{k,2} \sum_{t=1}^{n} u_{2,t} + \cdots + \sigma^{k,k} \sum_{t=1}^{n} u_{k,t} \\ \sigma^{k,1} \sum_{t=1}^{n} x_{k,t-1} u_{1,t} + \sigma^{k,2} \sum_{t=1}^{n} x_{k,t-1} u_{2,t} + \cdots + \sigma^{k,k} \sum_{t=1}^{n} x_{k,t-1} u_{k,t} \end{bmatrix}. \end{aligned}$$

Hence,

$$\begin{aligned} & \hat{\mu}_{i,SUR} \\ = \; & \frac{1}{|B|} \times \left(\mu_i |B| + \sum_{p=1}^{k} \left| \begin{bmatrix} B_{1,1} & \cdots & B_{1,2i-2} & \sigma^{1,p} \sum_{t=1}^{n} u_{p,t} & B_{1,2i} & \cdots & B_{1,2k} \\ B_{2,1} & \cdots & B_{2,2i-2} & \sigma^{1,p} \sum_{t=1}^{n} x_{1,t-1} u_{p,t} & B_{2,2i} & \cdots & B_{2,2k} \\ B_{3,1} & \cdots & B_{3,2i-2} & \sigma^{2,p} \sum_{t=1}^{n} u_{p,t} & B_{3,2i} & \cdots & B_{3,2k} \\ B_{4,1} & \cdots & B_{4,2i-2} & \sigma^{2,p} \sum_{t=1}^{n} x_{2,t-1} u_{p,t} & B_{4,2i} & \cdots & B_{4,2k} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ B_{2k-1,1} & \cdots & B_{2k-1,2i-2} & \sigma^{k,p} \sum_{t=1}^{n} u_{p,t} & B_{2k-1,2i} & \cdots & B_{2k-1,2k} \\ B_{2k,1} & \cdots & B_{2k,2i-2} & \sigma^{k,p} \sum_{t=1}^{n} x_{k,t-1} u_{p,t} & B_{2k,2i} & \cdots & B_{2k,2k} \end{bmatrix} \right| \right) \\ = \; & \mu_i + \frac{1}{|B|} \sum_{p=1}^{k} |(B)_{\mu_i,p}|. \end{aligned}$$

Note that, we have the joint convergence result from Lemma A.1 (iii) such that

$$\begin{aligned} \frac{1}{n^{\frac{1}{2}}} \sum_{t=1}^{n} u_{i,t} \quad & \Rightarrow \quad B_i(1), \\ \frac{1}{\rho_i^n} \sum_{t=1}^{n} x_{i,t-1}^2 u_{j,t} \quad & \Rightarrow \quad \frac{\sqrt{\sigma_{i,i}\sigma_{j,j}}}{\rho_i^2 - 1} Q_i(\rho_i) \widetilde{Q}_j(\rho_i), \end{aligned}$$

where $B_i(1)$ is independent from any $Q_i(\rho_i)$ and $\widetilde{Q}_j(\rho_i)$ for any $i, j = 1, 2, 3, \cdots, k$.

Therefore the numerator $|(B)_{\mu_i,p}|$ has the asymptotics

$$\left(\frac{1}{\prod_{j=1}^{k} \rho_j^n} \right)^2 \left(\frac{1}{n} \right)^k \frac{1}{n^{\frac{1}{2}}} |(B)_{\mu_i,p}|$$

$$
\Rightarrow \left|\begin{bmatrix}
\sigma^{1,1} & \sigma^{1,1}\frac{\sqrt{\sigma_{1,1}}Q_1(\rho_1)}{\rho_1-1} & \cdots & \sigma^{1,p}B_p(1) & \sigma^{1,i}\frac{\sqrt{\sigma_{i,i}}}{\rho_i-1}Q_i(\rho_i) & \cdots & \sigma^{1,k} & \sigma^{1,k}\frac{\sqrt{\sigma_{k,k}}Q_k(\rho_k)}{\rho_k-1} \\
O_p\left(n^{-1/2}\right) & \frac{\sigma^{1,1}\sigma_{1,1}Q_1(\rho_1)^2}{\rho_1^2-1} & \cdots & O_p\left(n^{-1/2}\right) & \frac{\sigma^{1,i}\sqrt{\sigma_{1,1}\sigma_{i,i}}Q_1(\rho_1)Q_i(\rho_i)}{\rho_1\rho_i-1} & \cdots & O_p\left(n^{-1/2}\right) & \frac{\sigma^{1,k}\sqrt{\sigma_{1,1}\sigma_{k,k}}Q_1(\rho_1)Q_k(\rho_k)}{\rho_1\rho_k-1} \\
\sigma^{2,1} & \sigma^{2,1}\frac{\sqrt{\sigma_{1,1}}Q_1(\rho_1)}{\rho_1-1} & \cdots & \sigma^{2,p}B_p(1) & \sigma^{2,i}\frac{\sqrt{\sigma_{i,i}}}{\rho_i-1}Q_i(\rho_i) & \cdots & \sigma^{2,k} & \sigma^{2,k}\frac{\sqrt{\sigma_{k,k}}Q_k(\rho_k)}{\rho_k-1} \\
O_p\left(n^{-1/2}\right) & \frac{\sigma^{2,1}\sqrt{\sigma_{2,2}\sigma_{1,1}}Q_2(\rho_2)Q_1(\rho_1)}{\rho^2-1} & \cdots & O_p\left(n^{-1/2}\right) & \frac{\sigma^{2,i}\sqrt{\sigma_{2,2}\sigma_{i,i}}Q_2(\rho_2)Q_i(\rho_i)}{\rho_2\rho_i-1} & \cdots & O_p\left(n^{-1/2}\right) & \frac{\sigma^{2,k}\sqrt{\sigma_{2,2}\sigma_{k,k}}Q_2(\rho_2)Q_k(\rho_k)}{\rho_2\rho_k-1} \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\
\sigma^{k,1} & \sigma^{k,1}\frac{\sqrt{\sigma_{1,1}}Q_1(\rho_1)}{\rho_1-1} & \cdots & \sigma^{k,p}B_p(1) & \sigma^{k,i}\frac{\sqrt{\sigma_{i,i}}}{\rho_i-1}Q_i(\rho_i) & \cdots & \sigma^{k,k} & \sigma^{k,k}\frac{\sqrt{\sigma_{k,k}}Q_k(\rho_k)}{\rho_k-1} \\
O_p\left(n^{-1/2}\right) & \frac{\sigma^{k,1}\sqrt{\sigma_{k,k}\sigma_{11}}Q_k(\rho_k)Q_1(\rho_1)}{\rho_k\rho_1-1} & \cdots & O_p\left(n^{-1/2}\right) & \frac{\sigma^{ki}\sqrt{\sigma_{k,k}\sigma_{i,i}}Q_k(\rho_k)Q_i(\rho_i)}{\rho_k\rho_i-1} & \cdots & O_p\left(n^{-1/2}\right) & \frac{\sigma^{kk}\sigma_{k,k}Q_k(\rho_k)^2}{\rho^2-1}
\end{bmatrix}\right|
$$

$$
= \begin{cases} \left(\prod_{j=1}^k \sigma_{jj}Q_j\left(\rho_j\right)^2\right)\frac{\xi_0(\boldsymbol{\rho})}{|\Sigma_u|}B_p(1) & \text{if } p=i, \\ 0 & \text{otherwise.} \end{cases} \tag{A.23}
$$

Combine the limiting results from (A.22) and (A.23), we obtain

$$
n^{\frac{1}{2}}\left(\hat{\mu}_{i,SUR}-\mu_i\right) = \frac{\left(\frac{1}{\prod_{j=1}^k\rho_j^n}\right)^2\left(\frac{1}{n}\right)^{k-1}\frac{1}{n^{\frac{1}{2}}}\sum_{p=1}^k|(B)_{\mu_i,p}|}{\left(\frac{1}{\prod_{j=1}^k\rho_j^n}\right)^2\left(\frac{1}{n}\right)^k|B|}
$$

$$
\Rightarrow \frac{\left(\prod_{j=1}^k\sigma_{j,j}Q_j\left(\rho_j\right)^2\right)\frac{\xi_0(\boldsymbol{\rho})}{|\Sigma_u|}B_p(1)}{\left(\prod_{j=1}^k\sigma_{j,j}Q_j\left(\rho_j\right)^2\right)\frac{\xi_0(\boldsymbol{\rho})}{|\Sigma_u|}} = B_i\left(1\right).
$$

By the same argument, the $2i$-th term of $\widehat{A}_{SUR}$ is

$$
\hat{\rho}_{i,SUR} = \rho_i + \frac{1}{|B|}\times\left|\begin{bmatrix}
B_{1,1} & \cdots & B_{1,2i-1} & \sigma^{1,p}\sum_{t=1}^n u_{p,t} & B_{1,2i+1} & \cdots & B_{1,2k} \\
B_{2,1} & \cdots & B_{2,2i-1} & \sigma^{1,p}\sum_{t=1}^n x_{1,t-1}u_{p,t} & B_{2,2i+1} & \cdots & B_{2,2k} \\
B_{3,1} & \cdots & B_{3,2i-1} & \sigma^{2,p}\sum_{t=1}^n u_{p,t} & B_{3,2i+1} & \cdots & B_{3,2k} \\
B_{4,1} & \cdots & B_{4,2i-1} & \sigma^{2,p}\sum_{t=1}^n x_{2,t-1}u_{p,t} & B_{4,2i+1} & \cdots & B_{4,2k} \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
B_{2k-1,1} & \cdots & B_{2k-1,2i-1} & \sigma^{k,p}\sum_{t=1}^n u_{p,t} & B_{2k-1,2i+1} & \cdots & B_{2k-1,2k} \\
B_{2k,1} & \cdots & B_{2k,2i-1} & \sigma^{k,p}\sum_{t=1}^n x_{k,t-1}u_{p,t} & B_{2k,2i+1} & \cdots & B_{2k,2k}
\end{bmatrix}\right|
$$

$$
= \rho_i + \frac{1}{|B|}\sum_{p=1}^k|(B)_{\rho_i,p}|.
$$

The numerator $|(B)_{\rho_i,p}|$ has the asymptotics

$$
\left(\frac{1}{\prod_{j\neq i}\rho_j^n}\right)^2\frac{1}{\rho_i^n}\left(\frac{1}{n}\right)^k|(B)_{\rho_i,p}|
$$

$$
\Rightarrow \left|\begin{bmatrix}
\sigma^{1,1} & O_p\left(n^{-1/2}\right) & \cdots & \sigma^{1,i} & \sigma^{1p}B_p(1) & \cdots & \sigma^{1,k} & O_p\left(n^{-1/2}\right) \\
O_p\left(n^{-1/2}\right) & \frac{\sigma^{1,1}}{\rho_1^2-1}\sigma_{11}Q_1(\rho_1)^2 & \cdots & O_p\left(n^{-1/2}\right) & \frac{\sigma^{1,p}\sqrt{\sigma_{1,1}\sigma_{p,p}}Q_1(\rho_1)\tilde{Q}_p(\rho_1)}{\rho_1^2-1} & \cdots & O_p\left(n^{-1/2}\right) & \frac{\sigma^{1k}\sqrt{\sigma_{1,1}\sigma_{k,k}}Q_1(\rho_1)Q_k(\rho_k)}{\rho_1\rho_k-1} \\
\sigma^{2,1} & O_p\left(n^{-1/2}\right) & \cdots & \sigma^{2,i} & \sigma^{2,p}B_p(1) & \cdots & \sigma^{2,k} & O_p\left(n^{-1/2}\right) \\
O_p\left(n^{-1/2}\right) & \frac{\sigma^{2,1}\sqrt{\sigma_{2,2}\sigma_{1,1}}Q_2(\rho_2)Q_1(\rho_1)}{\rho_2\rho_1-1} & \cdots & O_p\left(n^{-1/2}\right) & \frac{\sigma^{2,p}\sqrt{\sigma_{2,2}\sigma_{p,p}}Q_2(\rho_2)\tilde{Q}_p(\rho_2)}{\rho_2^2-1} & \cdots & O_p\left(n^{-1/2}\right) & \frac{\sigma^{2,k}\sqrt{\sigma_{2,2}\sigma_{k,k}}Q_2(\rho_2)Q_k(\rho_k)}{\rho_2\rho_k-1} \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\
\sigma^{k,1} & O_p\left(n^{-1/2}\right) & \cdots & \sigma^{k,i} & \sigma^{k,p}B_p(1) & \cdots & \sigma^{k,k} & O_p\left(n^{-1/2}\right) \\
O_p\left(n^{-1/2}\right) & \frac{\sigma^{k,1}\sqrt{\sigma_{k,k}\sigma_{1,1}}Q_k(\rho_k)Q_1(\rho_1)}{\rho_k\rho_1-1} & \cdots & O_p\left(n^{-1/2}\right) & \frac{\sigma^{k,p}\sqrt{\sigma_{k,k}\sigma_{p,p}}Q_k(\rho_k)\tilde{Q}_p(\rho_k)}{\rho_k^2-1} & \cdots & O_p\left(n^{-1/2}\right) & \frac{\sigma^{k,k}\sigma_{k,k}Q_k(\rho_k)^2}{\rho_k^2-1}
\end{bmatrix}\right|
$$

$$= \left(\prod_{j=1}^{k}\sigma_{j,j}Q_j(\rho_j)^2\right)\sqrt{\frac{\sigma_{p,p}}{\sigma_{i,i}}}\frac{1}{|\Sigma_u|Q_i(\rho_i)}\left|\begin{bmatrix} \frac{\sigma^{1,1}}{\rho_1^2-1} & \cdots & \sigma^{1,p}\widetilde{Q}_p(\rho_1) & \cdots & \frac{\sigma^{1,k}}{\rho_1\rho_k-1} \\ \vdots & & \vdots & & \vdots \\ \frac{\sigma^{k,1}}{\rho_k\rho_1-1} & \cdots & \sigma^{k,p}\widetilde{Q}_p(\rho_k) & \cdots & \frac{\sigma^{k,k}}{\rho_k^2-1}\end{bmatrix}\right|.$$

Hence

$$\rho_i^n\left(\hat{\rho}_{i,SUR}-\rho_i\right) = \frac{\left(\frac{1}{\prod_{j\neq i}\rho_j^n}\right)^2\frac{1}{\rho_i^n}\left(\frac{1}{n}\right)^k\sum_{p=1}^{k}|(B)_{\rho_i,p}|}{\left(\frac{1}{\prod_{j=1}^{k}\rho_j^n}\right)^2\left(\frac{1}{n}\right)^k|B|} \Rightarrow \frac{\xi_i(\boldsymbol{\rho})}{\xi_0(\boldsymbol{\rho})}.$$

Given the result of joint convergence from Lemma A.1, the SUR estimator of A have the following asymptotics:

$$\operatorname{diag}\left(n^{\frac{1}{2}},\rho_1^n,n^{\frac{1}{2}},\rho_2^n,\ldots,n^{\frac{1}{2}},\rho_k^n\right)\left(\widehat{A}_{SUR}-A\right)\Rightarrow\left[B_1(1),\frac{\xi_1(\boldsymbol{\rho})}{\xi_0(\boldsymbol{\rho})},B_2(1),\frac{\xi_2(\boldsymbol{\rho})}{\xi_0(\boldsymbol{\rho})},\ldots,B_p(1),\frac{\xi_k(\boldsymbol{\rho})}{\xi_0(\boldsymbol{\rho})}\right]_{2k\times 1}^{\top}$$

■

A.3 Proof of Corollary 3.2

Proof. The proof is analogous to Corollary 2.2, except replacing Lemma A.5 with A.3. ■

B Pre-test method

We test for the explosiveness using the method introduced in Chen et al. (2017). Suppose the data is recorded over a time interval $[0,T]$ with sampling interval h. Hence, the number of observation is $N = T/h$. $h = 1/12$ is for monthly data.

For each of the price series, we fit it with the univariate continuous explosive model

$$dx(t) = \kappa(\mu - x(t))\,dt + \sigma_{xx}^{1/2}dB_x(t),\ x(0) = O_p(1),\ \kappa < 0, \tag{B.1}$$

where B_x is a standard Brownian motion. The discrete time counterpart of the (B.1) is

$$x_{th} = a_h(\kappa)\,x_{(t-1)h} + g_h + u_{x,th},\ \ x_{0h} = x_0 = O_p(1),$$

with

$$\begin{aligned} a_h(\kappa) &= \exp(-\kappa h), \\ g_h &= \kappa^{-1}\left(I_K - e^{-\kappa h}\right)\kappa\mu, \\ u_{x,th} &= \int_{(t-1)h}^{th} e_{xx}^{-\kappa(th-s)}\sigma_{xx}dB_x(s) \overset{d}{=} \mathcal{N}(0,\sigma_{xx}h). \end{aligned}$$

Chen et al. (2017) shows that

$$t_\kappa = \frac{\widehat{\kappa}-\kappa^0}{s_\kappa} = \frac{a_h^N\left(\widehat{\kappa}-\kappa^0\right)}{\left\{s_x^2\left(\frac{1}{a_h^{2N}/h^2}\sum_{t=1}^N x_{(t-1)h}^2 - \frac{1}{N}\left(\frac{1}{a_h^N/h}\sum_{t=1}^N x_{(t-1)h}\right)^2\right)^{-1}\right\}^{1/2}} \Rightarrow \mathcal{N}(0,1).$$

With the result available above, we test for the explosiveness for the commodity closing prices considered in the empirical section.

Reported in Table S1 are the point estimates of the persistence parameter κ, the autoregressive coefficient $a_h(\kappa)$, the estimated standard errors, and the t-statistic for each commodity prices. Take the closing price of soybeans for example, the LS estimate of κ is -0.023 (whose autoregressive coefficient is 1.034), with the estimated standard error of 0.0001 and the t statistic of -1.514. The result confirm explosive behavior in soybean closing price over this period. Moreover, upon comparison of the test reported in Table S1, it is apparent that soybean, rice, wheat, corn, and sugar exhibiting explosive behavior under 5% significant level.

Table S1: The estimated persistence parameter in the CBOT agricultural commodity market.

Commodity	κ	a_h	$S.E.(\kappa)$	$T(\kappa)$
Soybean	-0.6239	1.0121	0.0634	-2.4778
Rice	-0.6851	1.0133	0.0837	-2.3675
Wheat	-0.8134	1.0158	0.2309	-1.6927
Corn	-0.9483	1.0184	0.1344	-2.5869
Sugar	-0.9177	1.0178	0.2151	-1.9785

C Additional Tables and Figures for Simulation

C.1 Regressors with two variates

This section includes Tables and Figures for regressors with Two variates. In particular, Figure C.1 and C.2 present the finite-sample distribution of intercepts estimate for DGP (2.6). Figure C.3 and C.4 present the finite-sample distribution of intercepts estimate for DGP (3.2). Table S2 reports the finite-sample performance comparison between OLS, SUR and IV estimators for explosive VAR models with no intercept as in DGP (2.1). Table S3 reports the finite-sample performance comparison between OLS, SUR and IV estimators for explosive VAR models with no intercept as in DGP (3.1).

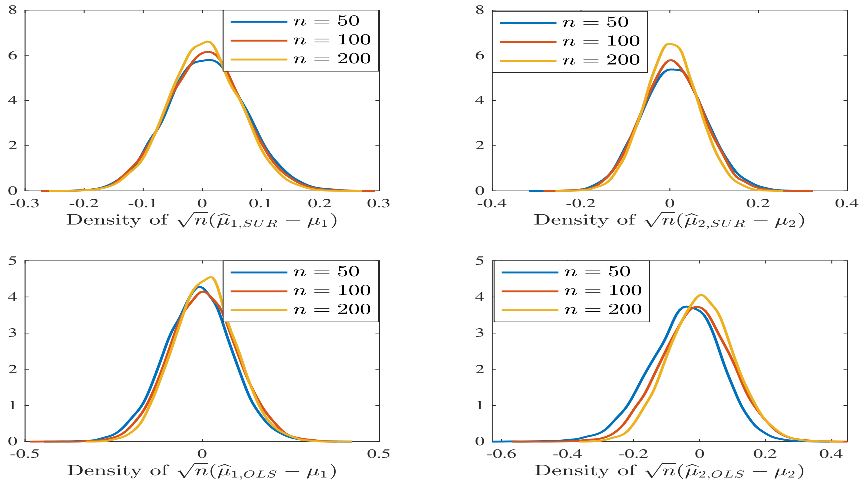

Fig. C.1: Finite-sample distribution of SUR and OLS estimators for DGP (2.6) with distinct explosive roots.

C.2 Regressors with three variates

First we consider the explosive process of x_t. We compare the finite-sample performance between OLS estimator, SUR estimator and IV estimator proposed in PM, in terms of bias, variance and mean squared error, for various values of $n = 50, 100, 200$. The data are generated from model (2.1) for $k = 3$ with $\rho = 1.02$, and

$$\Sigma_u = \begin{bmatrix} 1 & 0.8 & 0.6 \\ 0.8 & 1 & 0.7 \\ 0.6 & 0.7 & 1 \end{bmatrix}.$$

The results are reported in Table S4. We also consider the DGP (2.1) with intercept

$$\mu = [\mu_1, \mu_2, \mu_3]^\top = \left[1 \times n^{-1/2}, 0.8 \times n^{-1/2}, 6 \times n^{-1/2}\right]^\top.$$

The result are reported in Table S5.

Second, we consider the mildly explosive process of x_t. The data are generated from model (3.1) for $k = 3$ with $c = 4$ and $\alpha = 0.95$. The configuration of remaining parameters is the same as that in Table S4. The results are reported in Table S6. We also consider the DGP with intercept

$$\mu = [\mu_1, \mu_2, \mu_3]^\top = \left[1 \times n^{-1/2}, 0.8 \times n^{-1/2}, 6 \times n^{-1/2}\right]^\top.$$

The result are reported in Table S7.

C.3 Robustness check with non-Gaussian noise

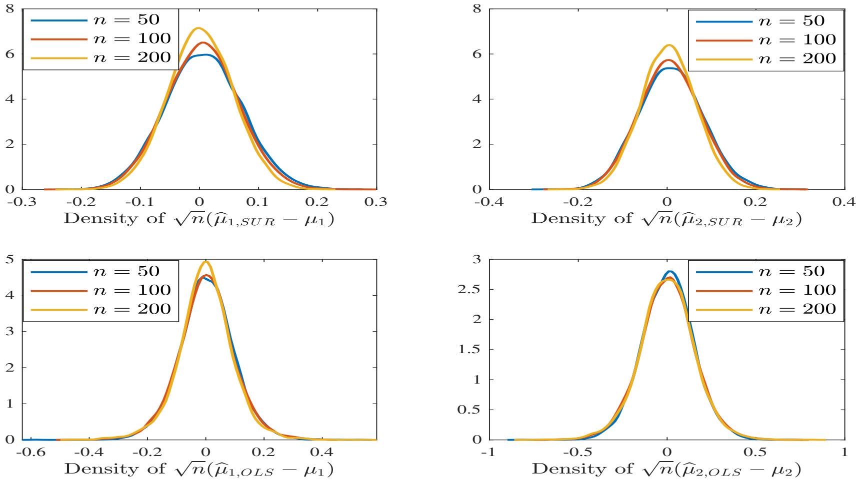

Fig. C.2: Finite-sample distribution of SUR and OLS estimators for DGP (2.6) with common explosive roots.

We redo table 1 where the noise now follows a $t(5)$ distribution. Table S8 shows that our SUR estimate can achieve good finite-sample performance even with non-Gaussian noise.

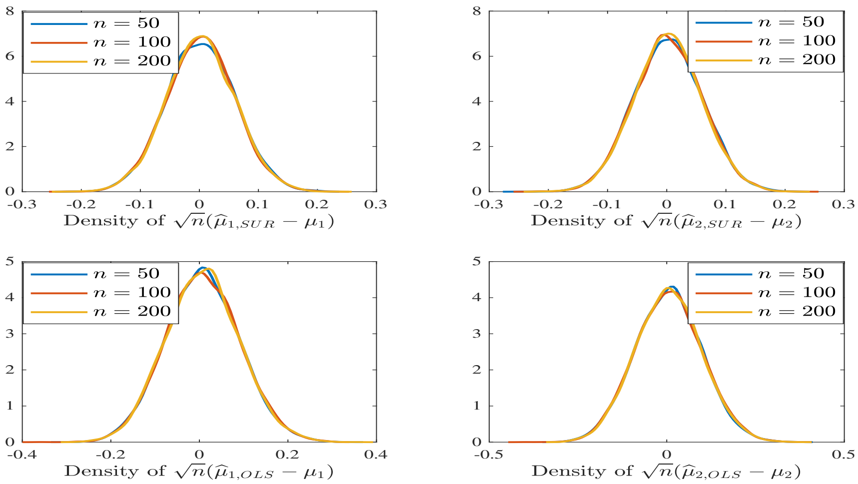

Fig. C.3: Finite-sample distribution of $n^{\frac{1}{2}}\left(\hat{\mu}_{i,SUR} - \mu_i\right)$ for $i = 1, 2$ for DGP (3.2) with distinct explosive roots.

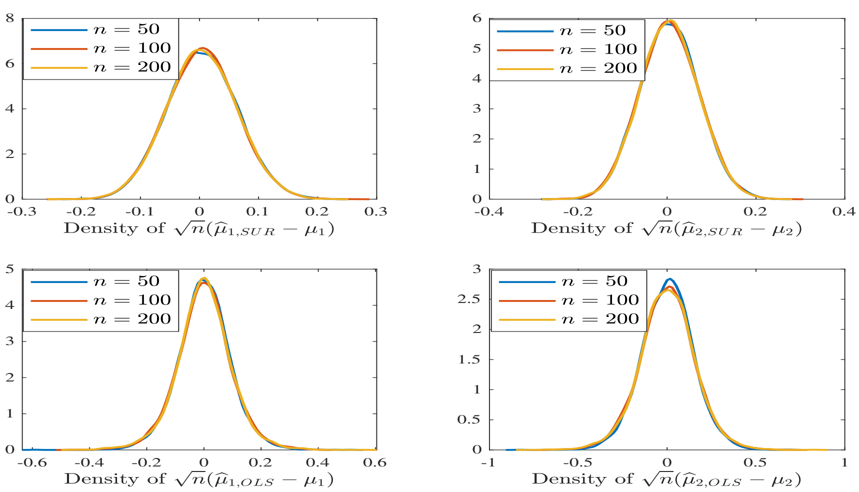

Fig. C.4: Finite-sample distribution of $n^{\frac{1}{2}}\left(\hat{\mu}_{i,SUR} - \mu_i\right)$ for $i = 1, 2$ for DGP (3.2) with common explosive roots.

Table S2: Finite-sample performance comparison between OLS, SUR and IV estimators for explosive VAR with no intercept as in DGP (2.1).

n	True		OLS			SUR			IV		
			bias	var	mse	bias	var	mse	bias	var	mse
Panel A. Explosive regressor for distinct explosive roots											
50	ρ_1	1.0050	−0.0736	0.0362	0.0416	−0.0252	0.0016	0.0022	0.2097	0.6632	0.7063
	ρ_2	1.0100	−0.0698	0.0332	0.0380	−0.0241	0.0016	0.0022	0.0538	6.8704	6.8733
100	ρ_1	1.0050	−0.0379	0.0083	0.0097	−0.0118	0.0004	0.0005	0.0952	0.0136	0.0227
	ρ_2	1.0100	−0.0323	0.0070	0.0080	−0.0106	0.0004	0.0005	0.0100	0.0115	0.0116
200	ρ_1	1.0050	−0.0176	0.0015	0.0018	−0.0048	0.0001	0.0001	0.0403	0.0015	0.0031
	ρ_2	1.0100	−0.0110	0.0011	0.0012	−0.0038	0.0001	0.0001	−0.0037	0.0010	0.0010
Panel B. Explosive regressor for common explosive roots											
50	ρ	1.0100	−0.0757	0.0365	0.0422	−0.0245	0.0016	0.0022	0.1588	0.6712	0.6964
		1.0100	−0.0724	0.0358	0.0410	−0.0244	0.0016	0.0022	0.1096	3.5936	3.6056
100	ρ	1.0100	−0.0419	0.0092	0.0110	−0.0110	0.0003	0.0005	0.0539	0.0149	0.0178
		1.0100	−0.0382	0.0092	0.0106	−0.0109	0.0003	0.0005	0.0510	0.0145	0.0171
200	ρ	1.0100	−0.0237	0.0027	0.0033	−0.0040	0.0001	0.0001	0.0207	0.0028	0.0032
		1.0100	−0.0222	0.0026	0.0031	−0.0041	0.0001	0.0001	0.0197	0.0027	0.0031

Table S3: Finite-sample performance comparison between OLS, SUR and IV estimators for mildly explosive VAR with no intercept as in DGP (3.1).

n	True		OLS			SUR			IV		
			bias	var	mse	bias	var	mse	bias	var	mse
Panel A. Mildly explosive regressor for distinct explosive roots											
50	ρ_1	1.0243	−0.0646	0.0192	0.0234	−0.0173	0.0010	0.0013	0.1814	0.1742	0.2070
	ρ_2	1.0486	−0.0360	0.0143	0.0156	−0.0125	0.0009	0.0010	−0.0171	0.0493	0.0496
100	ρ_1	1.0126	−0.0336	0.0046	0.0058	−0.0083	0.0002	0.0003	0.0744	0.0064	0.0120
	ρ_2	1.0252	−0.0166	0.0032	0.0034	−0.0058	0.0002	0.0002	−0.0076	0.0037	0.0120
200	ρ_1	1.0065	−0.0168	0.0011	0.0014	−0.0040	0.0001	0.0001	0.0340	0.0011	0.0022
	ρ_2	1.0130	−0.0081	0.0008	0.0008	−0.0027	0.0001	0.0001	−0.0038	0.0006	0.0006
Panel B. Mildly explosive regressor for common explosive roots											
50	ρ	1.0243	−0.0811	0.0377	0.0442	−0.0161	0.0013	0.0018	0.1234	1.7980	1.8132
		1.0243	−0.0781	0.0369	0.0430	−0.0151	0.0013	0.0018	0.1404	5.9024	5.9221
100	ρ	1.0126	−0.0431	0.0094	0.0112	−0.0103	0.0003	0.0004	0.0513	0.0147	0.0173
		1.0126	−0.0393	0.0093	0.0108	−0.0102	0.0003	0.0004	0.0482	0.0143	0.0166
200	ρ	1.0065	−0.0218	0.0025	0.0030	−0.0050	0.0001	0.0001	0.0235	0.0029	0.0035
		1.0065	−0.0205	0.0025	0.0029	−0.0051	0.0001	0.0001	0.0226	0.0028	0.0033

Table S4: Finite-sample performance comparison between OLS, SUR and IV estimators for the explosive regressors with no intercept as in DGP (2.1) for $k = 3$.

n	True		OLS			SUR			IV		
			bias	var	mse	bias	var	mse	bias	var	mse
50	ρ	1.0200	−0.1179	0.0282	0.0421	−0.0257	0.0017	0.0023	0.2395	1.1553	1.2126
		1.0200	−0.1236	0.0356	0.0509	−0.0237	0.0012	0.0018	0.2202	0.9964	1.0449
		1.0200	−0.1145	0.0221	0.0352	−0.0226	0.0016	0.0021	0.2081	0.6833	0.7266
100	ρ	1.0200	−0.0725	0.0080	0.0133	−0.0089	0.0003	0.0003	0.0787	0.0185	0.0247
		1.0200	−0.0675	0.0105	0.0151	−0.0078	0.0002	0.0003	0.0656	0.0196	0.0239
		1.0200	−0.0698	0.0056	0.0104	−0.0096	0.0004	0.0005	0.0759	0.0111	0.0169
200	ρ	1.0200	−0.0462	0.0026	0.0048	−0.0016	0.0000	0.0000	0.0260	0.0028	0.0035
		1.0200	−0.0471	0.0031	0.0053	−0.0015	0.0000	0.0000	0.0302	0.0035	0.0044
		1.0200	−0.0429	0.0018	0.0036	−0.0018	0.0000	0.0000	0.0291	0.0022	0.0031

Table S5: Finite-sample performance comparison between OLS, SUR and IV estimators for the explosive regressors with intercept as in DGP (2.1) for $k = 3$.

n	True		OLS			SUR			IV		
			bias	var	mse	bias	var	mse	bias	var	mse
50	μ_1	0.1414	−0.0026	0.4864	0.4864	0.1656	0.1458	0.1733	−0.1183	17.5425	17.5565
	μ_2	0.1131	0.0071	0.4753	0.4754	0.1534	0.1454	0.1689	0.0665	10.3898	10.3942
	μ_3	0.8485	0.0315	0.2928	0.2937	0.1461	0.0809	0.1023	0.0322	6.3710	6.3721
	ρ	1.0200	−0.2391	0.0415	0.0986	−0.0568	0.0023	0.0055	0.5994	10.8350	11.1943
		1.0200	−0.2467	0.0422	0.1031	−0.0564	0.0019	0.0051	0.4072	14.8992	15.0650
		1.0200	0.0115	0.0017	0.0019	−0.0041	0.0000	0.0001	−0.0532	0.2667	0.2695
100	μ_1	0.1000	−0.0168	0.2756	0.2759	0.0634	0.0583	0.0623	0.0042	0.1520	0.1520
	μ_2	0.0800	0.0070	0.2980	0.2980	0.0621	0.0602	0.0640	−0.0013	0.1646	0.1646
	μ_3	0.6000	−0.0173	0.1492	0.1495	0.0654	0.0307	0.0350	0.0072	0.1263	0.1264
	ρ	1.0200	−0.1323	0.0109	0.0284	−0.0188	0.0005	0.0009	0.1310	0.0394	0.0565
		1.0200	−0.1473	0.0113	0.0330	−0.0183	0.0004	0.0007	0.1434	0.0427	0.0633
		1.0200	0.0064	0.0005	0.0006	−0.0008	0.0000	0.0000	−0.0070	0.0011	0.0012
200	μ_1	0.0707	−0.0115	0.1430	0.1431	−0.0080	0.0119	0.0120	−0.0037	0.0497	0.0497
	μ_2	0.0566	−0.0069	0.1379	0.1380	0.0086	0.0115	0.0116	−0.0032	0.0479	0.0480
	μ_3	0.4243	−0.0515	0.0737	0.0763	0.0101	0.0101	0.0102	0.0223	0.0385	0.0390
	ρ	1.0200	−0.0822	0.0029	0.0097	−0.0023	0.0000	0.0000	0.0500	0.0052	0.0077
		1.0200	−0.0896	0.0031	0.0111	−0.0024	0.0000	0.0000	0.0586	0.0055	0.0089
		1.0200	0.0026	0.0003	0.0003	0.0000	0.0000	0.0000	−0.0014	0.0002	0.0002

Table S6: Finite-sample performance comparison between OLS, SUR and IV estimators for the mildly explosive regressors with no intercept as in DGP (3.1) for $k = 3$.

n	True		OLS			SUR			IV		
			bias	var	mse	bias	var	mse	bias	var	mse
50	ρ_1	1.0973	−0.1855	0.0374	0.0718	−0.0032	0.0002	0.0002	0.2434	4.2067	4.2659
	ρ_2	1.0973	−0.1970	0.0490	0.0878	−0.0038	0.0002	0.0002	0.1820	3.2926	3.3258
	ρ_3	1.0973	−0.1512	0.0306	0.0633	−0.0042	0.0003	0.0003	0.3079	18.8933	18.9881
100	ρ_1	1.0504	−0.1031	0.0106	0.0212	−0.0020	0.0001	0.0001	0.0687	0.0210	0.0258
	ρ_2	1.0504	−0.0957	0.0140	0.0231	−0.0011	0.0000	0.0000	0.0532	0.0208	0.0236
	ρ_3	1.0504	−0.1016	0.0075	0.0178	−0.0021	0.0001	0.0001	0.0652	0.0123	0.0166
200	ρ_1	1.0261	−0.0528	0.0030	0.0058	−0.0005	0.0000	0.0000	0.0240	0.0029	0.0035
	ρ_2	1.0261	−0.0542	0.0035	0.0065	−0.0007	0.0000	0.0000	0.0291	0.0037	0.0045
	ρ_3	1.0261	−0.0493	0.0021	0.0045	−0.0010	0.0000	0.0000	0.0276	0.0023	0.0030

Table S7: Finite-sample performance comparison between OLS, SUR and IV estimators for the mildly explosive regressors with intercept as in DGP (3.1) for $k = 3$.

n		True	OLS			SUR			IV		
			bias	var	mse	bias	var	mse	bias	var	mse
50	μ_1	0.1414	−0.0343	0.4746	0.4758	0.0052	0.0395	0.0396	−0.1033	31.2874	31.2981
	μ_2	0.1131	−0.0183	0.4682	0.4685	0.0064	0.0370	0.0371	0.0749	29.8923	29.8979
	μ_3	0.8485	−0.1037	0.2840	0.2948	0.0141	0.0362	0.0364	0.0033	19.2076	19.2076
	ρ_1	1.0973	−0.3344	0.0403	0.1521	−0.0053	0.0003	0.0004	0.5703	20.3038	20.6290
	ρ_2	1.0973	−0.3464	0.0423	0.1623	−0.0055	0.0004	0.0004	0.4673	18.6579	18.8762
	ρ_3	1.0973	0.0058	0.0046	0.0046	0.0000	0.0000	0.0000	−0.0514	0.4971	0.4997
100	μ_1	0.1000	−0.0271	0.2739	0.2747	−0.0008	0.0180	0.0180	0.0150	0.1402	0.1404
	μ_2	0.0800	−0.0005	0.2901	0.2901	0.0006	0.0192	0.0192	0.0069	0.1517	0.1517
	μ_3	0.6000	−0.0891	0.1480	0.1560	0.0043	0.0166	0.0167	0.0438	0.1153	0.1173
	ρ_1	1.0504	−0.1734	0.0112	0.0412	−0.0026	0.0001	0.0001	0.1196	0.0448	0.0591
	ρ_2	1.0504	−0.1885	0.0121	0.0476	−0.0020	0.0001	0.0001	0.1274	0.0475	0.0637
	ρ_3	1.0504	0.0021	0.0012	0.0012	0.0000	0.0000	0.0000	−0.0032	0.0016	0.0016
200	μ_1	0.0707	−0.0145	0.1421	0.1423	0.0031	0.0087	0.0087	−0.0025	0.0468	0.0468
	μ_2	0.0566	−0.0095	0.1368	0.1369	0.0027	0.0084	0.0084	−0.0022	0.0454	0.0454
	μ_3	0.4243	−0.0708	0.0745	0.0795	0.0030	0.0084	0.0084	0.0270	0.0360	0.0367
	ρ_1	1.0261	−0.0904	0.0030	0.0112	−0.0008	0.0000	0.0000	0.0474	0.0054	0.0076
	ρ_2	1.0261	−0.0989	0.0032	0.0130	−0.0009	0.0000	0.0000	0.0563	0.0057	0.0089
	ρ_3	1.0261	0.0012	0.0004	0.0004	0.0000	0.0000	0.0000	−0.0007	0.0003	0.0003

Table S8: Finite-sample performance comparison between OLS, SUR and IV estimators for explosive VAR with intercept as in DGP (2.6) for t-distribution with degree of freedom 5.

n	True		OLS		SUR		IV		OLS-IV with Pre-test	
			100*bias	100*var	100*bias	100*var	100*bias	100*var	100*bias	100*var
Panel A. Explosive regressor for distinct explosive roots										
50	μ_1	0.1414	−0.1643	0.0181	0.1072	0.0087	0.0364	0.0193	−0.1541	0.0182
	μ_2	0.1131	−0.6496	0.0235	0.1197	0.0103	0.3270	0.0246	−0.6008	0.0235
	ρ_1	1.0050	1.3469	0.4683	−0.0244	0.0004	−0.8991	0.6883	1.2346	0.4793
	ρ_2	1.0100	−4.8615	0.7411	−0.0292	0.0007	4.6449	1.2814	−4.3862	0.7681
100	μ_1	0.1000	0.0408	0.0091	0.0620	0.0042	−0.0498	0.0091	0.0363	0.0091
	μ_2	0.0800	−0.1078	0.0116	0.0656	0.0047	0.0269	0.0114	−0.1011	0.0116
	ρ_1	1.0050	−0.0213	0.0229	−0.0086	0.0001	−0.0811	0.0241	−0.0162	0.0229
	ρ_2	1.0100	−0.2431	0.0257	−0.0083	0.0001	0.1508	0.0281	−0.2234	0.0258
200	μ_1	0.0707	0.0589	0.0038	0.0256	0.0018	−0.0412	0.0038	0.0539	0.0038
	μ_2	0.0566	0.0307	0.0048	0.0232	0.0018	−0.0260	0.0048	0.0279	0.0048
	ρ_1	1.0050	−0.0225	0.0009	−0.0016	0.0000	0.0306	0.0009	−0.0198	0.0009
	ρ_2	1.0100	0.0028	0.0005	−0.0008	0.0000	−0.0125	0.0005	0.0020	0.0005
Panel B. Explosive regressor for common explosive roots										
50	μ_1	0.1414	0.0528	0.0188	0.0995	0.0083	−0.0921	0.3707	−0.0849	0.3531
	μ_2	0.1131	0.0572	0.0442	0.1162	0.0103	−0.1512	1.0047	−0.1408	0.9567
	ρ	1.0100	11.156	2.1009	−0.0201	0.0004	−20.344	346.20	−18.769	329.00
		1.0100	−30.964	4.4297	−0.0275	0.0007	55.253	2105.8	50.932	2000.7
100	μ_1	0.1000	0.0204	0.0094	0.0540	0.0037	−0.0076	0.0080	−0.0062	0.0081
	μ_2	0.0800	0.0205	0.0240	0.0639	0.0048	−0.0084	0.0124	−0.0070	0.0130
	ρ	1.0100	6.0717	0.5361	−0.0060	0.0001	−6.4571	1.1154	−5.8307	1.0864
		1.0100	−16.975	1.2103	−0.0078	0.0001	17.942	2.8943	16.196	2.8101
200	μ_1	0.0707	0.0034	0.0043	0.0177	0.0015	−0.0040	0.0031	−0.0040	0.0032
	μ_2	0.0566	0.0027	0.0119	0.0206	0.0019	−0.0086	0.0046	−0.0080	0.0050
	ρ	1.0100	3.4329	0.1360	0.0006	0.0000	−2.7558	0.1906	−2.4464	0.1879
		1.0100	−0.9378	0.3173	0.0007	0.0000	7.5992	0.4841	6.7373	0.4758

D The limiting distribution of IV method for distinct explosive roots

We stated the limiting distribution of IV method under DGP with explosive regressors and mildly explosive regressors. The details are as follows:

Explosive process with distinct explosive roots in Model (2.1):

$$X_t = RX_{t-1} + u_t, \quad t = 1, \ldots, n,$$

where X_t is a k-dimensional vector with $X_t = [x_{1,t}, \ldots, x_{k,t}]^\top$. The initial value is set as $x_{i,0} = 0$ for $i = 1, \ldots, k$ for the ease purpose. The residual $u_t = [u_{1,t}, \ldots, u_{k,t}]^\top$ is assumed to be a martingale difference sequence with respect to $\mathcal{F}_t = \sigma(u_t, u_{t-1}, \ldots)$ satisfying $\mathbb{E}[u_t u_t^\top | \mathcal{F}_{t-1}] = \Sigma_u$ with $\mathrm{Cov}(u_{i,t}, u_{j,t}) = \sigma_{i,j}$ for $i, j = 1, \ldots, k$. The autoregressive matrix is defined as $R = \mathrm{diag}(\rho_1, \ldots, \rho_k)$ with $\rho_i > 1$ for $i = 1, \ldots, k$ and $\rho_i \neq \rho_j$ for $i, j = 1, \ldots, k$.

The limiting distribution of the IV method under model (**2.1**) with distinct explosive roots is as follows:

Since $X_t = RX_{t-1} + u_t$, and $X_{t+k} = R^{k+1}X_{t-1} + \sum_{j=0}^{k} R^j u_{t+j}$, and from the definition of IV in Phillips and Magdalinos (2013) such that

$$\widehat{R}_{n,p} = \left(\sum_{t=1}^{n-p} X_t X_{t+p}^\top\right)\left(\sum_{t=1}^{n-p} X_{t-1} X_{t+p}^\top\right)^{-1}, \text{ for } p \in (0, 1, 2, \ldots),$$

we have

$$\begin{aligned}
& \widehat{R}_{n,p} - R \\
= \; & \sum_{t=1}^{n-p} u_t X_{t+p}' \left(\sum_{t=1}^{n-p} X_{t-1} X_{t+p}'\right)^{-1} \\
= \; & \left[\sum_{t=1}^{n-p} u_t X_{t-1}' R^{p+1} + \sum_{t=1}^{n-p} u_t \left(\sum_{j=0}^{p} R^j u_{t+j}\right)'\right]\left[\sum_{t=1}^{n-p} X_{t-1} X_{t-1}' R^{p+1} + \sum_{t=1}^{n-p} X_{t-1} \left(\sum_{j=0}^{p} R^j u_{t+j}\right)'\right]^{-1}
\end{aligned} \tag{D.1}$$

Since $\sum_{t=1}^{n-p} u_t u_{t+j}' = O_p(n)$ by WLLN and R^j is finite for $j = 0, 1, \ldots, p$ and $p \in (0, 1, 2, \ldots)$, the numerator of (D.1) equals to $\sum_{t=1}^{n-p} u_t x_{t-1}' R^{p+1} + O_p(n) = O_p(R^n)$ and the denominator of (D.1) equals to $\sum_{t=1}^{n-p} X_{t-1} X_{t-1}' R^{p+1} + O_p(n) = O_p\left(R^{2n}\right)$ using the limiting result from Magdalinos and Phillips (2008). Hence, the limting theory of Theorem 2.2 in Magdalinos and Phillips (2008) applies. In particular, we have

$$vec\left(\widehat{R}_{n,p} - R\right) \Rightarrow MN\left(0, \left(\sum_{j=0}^{\infty} R^{-j} X_R X_R' R^{-j}\right)^{-1} \otimes \Sigma_u\right)$$

Mildly explosive process with distinct explosive roots in Model (3.1):

$$X_t = R_n X_{t-1} + u_t. \tag{D.2}$$

In particular, $R_n = \text{diag}\left(\rho_{1n}, \cdots, \rho_{kn}\right)$, and the autoregressive root ρ_{in} is defined as $\rho_{in} = \rho_n = 1 + \frac{c_i}{n^\alpha}$ with $\alpha \in (0,1)$. The root is moderately deviated from unity, hence covered a larger neighborhood around unity than the local-to-unity processes. Other assumptions are the same as that in Section 2. Let $\{u_t, \mathcal{F}_t\}$ be an independent and identically distributed random variables with

$$\mathbb{E}[u_t|\mathcal{F}_{t-1}] = 0, \text{ and } \mathbb{E}[u_t u_t^\top|\mathcal{F}_{t-1}] = \Sigma_u < \infty, \text{ for all } t \leq n.$$

We have the numerator of (D.1) equals to $\sum_{t=1}^{n-p} u_t x_{t-1}' R^{p+1} + O_p(n) = O_p(n^\alpha R^n)$ and the denominator of (D.1) equals to $\sum_{t=1}^{n-p} x_{t-1} x_{t-1}' R^{p+1} + O_p(n) = O_p\left(n^{2\alpha} R^{2n}\right)$ using the limiting result from Phillips and Magdalinos (2009). Hence, the limiting theory of Theorem 4.1 in Phillips and Magdalinos (2009) applies. In particular, we have

$$n^\alpha \left(\widehat{R}_{n,p} - R_n\right) R_n^n \Rightarrow MN\left(0, \left(\int_0^\infty e^{-mC} Y_C Y_C' e^{-mC} dm\right)^{-1} \otimes \Sigma_u\right),$$

where $Y_C =_d N\left(0, \int_0^\infty e^{-mC} \Omega_u e^{-mC} dm\right)$ and $C = diag\left(c_1, \cdots, c_k\right)$.

References

Chen, Y., Phillips, P. C. B. & Yu, J. (2017), 'Inference in continuous systems with mildly explosive regressors', *Journal of Econometrics* **201**(2), 400–416.

Chen, Y. Z., Li, J. & Li, Q. (2022), 'Online supplement to "seemingly unrelated regression estimation for var models with explosive roots"', *Working paper* .

Magdalinos, T. & Phillips, P. C. B. (2009), 'Limit theory for cointegrated systems with moderately integrated and moderately explosive regressors', *Econometric Theory* **25**(2), 482–526.

Phillips, P. C. B. & Magdalinos, T. (2007), 'Limit theory for moderate deviations from a unit root', *Journal of Econometrics* **136**(1), 115–130.

Phillips, P. C. B. & Magdalinos, T. (2008), 'Limit theory for explosively cointegrated systems', *Econometric Theory* **24**(4), 865–887.

Phillips, P. C. B. & Magdalinos, T. (2013), 'Inconsistent VAR regression with common explosive roots', *Econometric Theory* **29**(4), 808–837.