OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 0305-9049
doi: 10.1111/obes.12551

Seemingly Unrelated Regression Estimation for VAR Models with Explosive Roots*

YE CHEN,† JIAN LI‡ and QIYUAN LI§

†*International School of Economics and Management, Capital University of Economics and Business, Beijing, China (e-mail: zoeyechen_cueb@163.com)*
‡*College of Economics and Management, Huazhong Agricultural University, Wuhan, China (e-mail: jli@mail.hzau.edu.cn)*
§*School of Economics, Singapore Management University, Singapore (e-mail: qyli.2019@phdecons.smu.edu.sg)*

Abstract

For VAR models with common explosive root, the OLS estimator of the autoregressive coefficient matrix is inconsistent (refer to Nielsen, 2009 and Phillips and Magdalinos, 2013). Although Phillips & Magdalinos (2013) proposed using the future observations as the instrumental variable for removing the endogeneity from VAR models, type I error occurs when testing for a common explosive root from the distinct explosive roots before the implementation of IV estimation. Such error creates bias and variance in the estimate and further causes incorrect inference in the structural analysis such as forecast error decomposition (FEVD). Hence, we propose using of seemingly unrelated regression (SUR) estimation for VAR models with explosive roots. Our SUR estimator is consistent in the case of both distinct explosive roots and common explosive root. We also consider models with drift in the system for generalization. Simulations show that the SUR estimate performs better than OLS and IV estimate in the case of both a common explosive root and distinct explosive roots case. In structural FEVD analysis, simulations show that SUR yields a different result from OLS and IV. We demonstrate the use of SUR in FEVD for agricultural commodity markets between 3 July 2010, and 29 January 2011.

I. Introduction

The explosive process is able to capture bubbles in asset prices (Diba and Grossman, 1988). Hence, it has been used extensively in recent studies of asset price bubbles. The algorithms provided by Phillips, Wu, and Yu (2011), Phillips, Shi, and Yu (2015a, 2015b) are able

JEL Classification numbers: C12, C13, C58.
*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).

to detect bubble behaviour and date-stamp its origination and collapse. These methods are widely used in the empirical study of asset price bubbles in various markets, including the stock market (see Basse *et al.*, 2021; Horváth, Li, and Liu, 2021; Li, Wang, and Zhao, 2021), housing market (See Phillips and Yu, 2013; Greenaway-McGrevy and Phillips, 2016; Shi *et al.*, 2016), cryptocurrency market (see Cheung, Roca, and Su, 2015; Corbet, Lucey, and Yarovaya, 2018; Bouri, Shahzad, and Roubaud, 2019), oil market (see Fantazzini, 2016; Caspi, Katzke, and Gupta, 2018; Gharib, Mefteh-Wali, and Jabeur, 2021), precious metals market (see Figuerola-Ferretti, Gilbert, and McCrorie, 2015; Pan, 2018; and Ma and Xiong, 2021), exchange rate market and others (see Etienne, Irwin, and Garcia, 2014; Kräussl, Lehnert, and Martelin, 2016; Shi, Hurn, and Phillips, 2020). These models are built in a non-stationary framework. In a stationary framework, Gouriéroux and Zakoı ̈an (2017) propose a non-causal autoregressive process with heavy-tailed errors to capture the local explosive behaviour in the financial time series.

In addition to the study of univariate explosive process, there is increased interest in multivariate explosive processes. For example, Nielsen (2010) studied a vector autoregressive model with one unit root process and one explosive process. From Nielsen (2010), Engsted and Nielsen (2012) proposed a bubble detection mechanism for asset prices in VAR regression. Phillips and Lee (2015) analyse a VAR system with mixed explosive roots. This model allows for a local to unit root from the explosive side and a mildly explosive root. Moreover, they study the Wald test and model selection criterion for testing for common roots. Magdalinos and Phillips (2009) developed limit theory for multivariate co-explosive processes. In particular, they consider the cases of both the distinct explosive roots and common explosive root. Different from the distinct explosive roots case, the common explosive root case yields the singular matrix for the sample variance matrix, hence it requires coordinates rotation in developing the asymptotics. When the regressors are endogenous, Phillips and Lee (2016) consider self-generated instruments in a method called IVX in the co-explosive system. The IVX procedure enables a robust Wald test for regressors with different levels of persistence. The continuous-time counterpart of Magdalinos and Phillips (2009) in discrete time is developed in Chen, Phillips, and Yu (2017). In a stationary framework, Gourieroux and Jasiak (2017) consider a VAR(p) model with mixed causal and non-causal components. They introduce a consistent semi-parametric estimator for model estimation. Cubadda, Hecq, and Telg (2019) studies co-movement features in the non-causal time-series models.

VAR models with explosive roots are of particular interest in this paper. VAR models are the fundamental statistical tool for studying the relationship between multiple time series over time. It provides a framework for structural analysis such as forecast error variance decomposition, which are useful tools for analysing the effect of shocks to the variables in the system. However, for VAR models with a common explosive root, the OLS estimator of the autoregressive coefficient matrix is inconsistent (see Nielsen, 2009; Phillips and Magdalinos, 2013). Phillips and Magdalinos (2013) explained the inconsistency problem in terms of endogeneity induced by co-explosive behaviour. In particular, co-explosive behaviour results in the singularity of the sample variance matrix in the limit. To address this asymptotic singularity, they rotate coordinates by using orthogonal transformation. The transformation yields a term that is correlated with the future residuals; hence, the endogeneity arises from the system. Phillips and Magdalinos (2013) have proposed using future observations as the instrumental variable (IV) to remove the endogeneity from the VAR model. Although the IV estimator is also consistent for explosive regressors with distinct explosive roots[1], the IV estimator causes larger variance than OLS estimator in both distinct explosive roots and common explosive root cases. In addition, we could use IV estimators (Phillips and Magdalinos, 2013) for VAR models with common explosive roots and use OLS estimators for VAR models with distinct explosive roots. However, type I error occurs when performing statistical inference to distinguish the case with distinct explosive roots from that with a common explosive root. Such error creates bias and variance in the estimate and further causes incorrect inference in the structural analysis such as forecast error decomposition. We demonstrate this point in the section on simulation. Hence, new estimators that account for these issues are crucially needed.

To solve this problem, we propose using an SUR estimator for VAR models with a common explosive root and with distinct explosive roots. The use of SUR estimators has a long history in econometric research. Since Zellner (1962) used SUR to study a multivariate system with correlated errors, there have been many studies employing SUR in the literature. For example, in the context of cointegration regressions, Mark, Ogaki, and Sul (2005) propose the dynamic SUR estimation strategy. Moon and Perron (2006) study SUR estimation for a triangular system with integrated regressors and presents an empirical study for testing purchasing power parity among the G-7 countries. In the semiparametric context, Henderson *et al.* (2015) consider the smooth coefficient of the SUR model. Smith and Kohn (2000) investigate the SUR in the non-parametric context. In the spatial modelling context, we have Anselin (1988) and Baltagi and Pirotte (2011). The recent work by Chen *et al.* (2017) analyse the multivariate Ornstein–Uhlenbeck processes with common persistence and shows that SUR estimator performs well in the bias reduction. Ultimately, the SUR estimator is used for either gaining estimation efficiency or imposing restrictions on the difference equations in the system; refer to Moon and Perron (2006).

In VAR models with both a common explosive root and distinct explosive roots, we show that the SUR estimator is consistent. In addition, our SUR estimator follows a mixture normal distribution in the limit. Therefore, it can be used for inference. Moreover, we generalize our analytical study to mildly explosive processes, whose autoregressive roots moderately deviate from unity (Phillips and Magdalinos, 2007). Such roots are closely related to the real data and are useful for practical implementation. For moderately explosive system with distinct explosive roots, our SUR estimator performs well in finite samples. For a moderately explosive system with a common explosive root, the SUR estimator is consistent and has better finite sample performance than the OLS and IV estimator proposed in Magdalinos and Phillips (2009).

The paper is organized as follows. Section II introduces prototypical VAR models with explosive processes, provides the SUR estimation procedure and presents the

[1]A brief discussion regarding the limiting distribution of the IV method for explosive regressors with distinct explosive roots is provided in the Data S1.

limit distribution of normalized estimators. Section III generalizes the results of the mildly explosive VAR models. Section IV reports simulations studying the finite-sample performance of the SUR estimator, in comparison with the OLS estimator and IV estimator from Phillips and Magdalinos (2013). The application to Chicago Board of Trade (CBOT) agricultural commodities is conducted in Section V. Section VI concludes the paper.

Throughout this paper, we denote the Euclidean norm of matrix A by $\|A\| = (A^\top A)^{1/2}$. The notation M is used to denote a finite positive constant, := and =: represent definitional equality, $\to^{a.s.}$ denotes almost sure convergence, and $\Rightarrow$ signifies weak convergence on the relevant probability space. We assume that n goes to infinity.

II. Seemingly unrelated regression of VAR models with explosive regressors

We analyse the asymptotic behaviour of the SUR estimator, and find that it is able to produce consistent estimates for a VAR regression model with both a common explosive root and distinct explosive roots. For generalization, we also consider the case where there are the intercept terms in the VAR system.

The model and the inconsistency problem

Following Phillips and Magdalinos (2013), we consider the following model:

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

where X_t is a k-dimensional vector with $X_t = [x_{1,t}, \dots, x_{k,t}]^\top$. The initial value is set to $x_{i,0} = 0$ for $i = 1, \dots, k$ for simplicity. The residual $u_t = [u_{1,t}, \dots, u_{k,t}]^\top$ is assumed to be a martingale difference sequence with respect to $\mathcal{F}_t = \sigma(u_t, u_{t-1}, \dots)$ 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 $i,j = 1, \dots, k$. The autoregressive coefficient matrix is defined as $R = \mathrm{diag}(\rho_1, \dots, \rho_k)$.[2,3] We consider two cases:

(1) distinct explosive roots: $\rho_i > 1$ for $i = 1, \dots, k$ and $\rho_i \neq \rho_j$ for $i,j = 1, \dots, k$.

(2) common explosive root: $\rho_i = \rho > 1$ for $i = 1, \dots, k$.

The OLS estimator for model (1) with a common explosive root is inconsistent (Phillips and Magdalinos, 2013). Take $k = 2$ for example. First, the standardized sample variance matrix $\sum_{t=1}^n X_t X_t^\top$ is asymptotically singular. To treat the singularity in the limit, Phillips and Magdalinos (2013) use the orthogonal transformation for X_t by $H_n^\top$ such that

$$H_n = \frac{1}{\|X_n\|} \begin{bmatrix} x_{1,n} & -x_{2,n} \\ x_{2,n} & x_{1,n} \end{bmatrix} = \frac{1}{\|X_n\|} \left[X_n, \mathcal{R}_{\frac{\pi}{2}} X_n \right],$$

with $\mathcal{R}_{\frac{\pi}{2}} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Let $Z_t = H_n^\top X_t$. The limiting OLS estimator is

[2] The diagonal structure of R is found in predictive regression with multiple predictors (Amihud and Hurvich, 2004). In addition, the variables can be correlated through the covariance matrix of errors Σ_u. We can test for the diagonal structure of R using the SUR method. Refer to Judge (1982).

[3] We could generalize our model to consider the inclusion of lagged variables of $x_{i,t-p}$ for some $p \geq 1$. Such generalizations would complicate in the development of asymptotics and hence are left for future work.

$$\widehat{R} - R = \left(\frac{1}{n}\sum_{t=1}^{n} u_t Z_{t-1}^{\top}\right)\left(\frac{1}{n}\sum_{t=1}^{n} Z_t Z_t^{\top}\right)^{-1} H_n^{\top}.$$

The transformed variable Z_{t-1} can be rewritten as $Z_{t-1} = \frac{1}{\|X_n\|}\begin{bmatrix} X_n^{\top} X_{t-1} \\ -\mathcal{R}_{\frac{\pi}{2}} X_n \sum_{j=t}^{n} \rho^{-(j-t+1)} u_j \end{bmatrix}$. In the limit of $\frac{\mathcal{R}_{\frac{\pi}{2}} X_n}{\|X_n\|}$, there is a term $\sum_{j=1}^{\infty} \rho^{-j} u_j$, that is correlated with the error term u_t. Hence, we have endogeneity in the regressors, which further causes inconsistency of the OLS estimator in the limit.

To remove the endogeneity from VAR models with common explosive roots, Phillips and Magdalinos (2013) propose an IV estimator. However, there is a type I error associated with this pre-test, since we have to test for the common explosive root before the implementation of IV estimation. Although the IV estimate is consistent for VAR models with distinct explosive roots, the use of IV estimation produces larger variance than OLS estimation and further entails incorrect inference in structural analysis such as forecast error decomposition as reported in our simulation section. Therefore, we propose using the SUR estimator. The SUR estimator is proved to be consistent in the cases of both a common explosive root and distinct explosive roots. Moreover, we demonstrate in the simulation section that the SUR estimator has better finite sample performance in terms of bias and variance than OLS, IV, IV-OLS method with a pre-test procedure[4].

Seemingly unrelated regression estimate and asymptotics

Let the ith regression model of model (1) be $x_{i,t} = \rho_i x_{i,t-1} + u_{i,t}$. Define $X_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,n}]^{\top}$, $X_{i\_} = [x_{i,0}, x_{i,1}, \ldots, x_{i,n-1}]^{\top}$, and $U_i = [u_{i,1}, u_{i,2}, \ldots, u_{i,n}]^{\top}$, which are an $n \times 1$ vector. Let $A = [\rho_1, \ldots, \rho_k]^{\top}$, $X = [X_1^{\top}, \ldots, X_k^{\top}]^{\top}$, $U = [U_1, \ldots, U_k]^{\top}$, and

$$X\_ = \begin{bmatrix} X_{1\_} & 0_{n\times 1} & \cdots & 0_{n\times 1} \\ 0_{n\times 1} & X_{2\_} & \cdots & 0_{n\times 1} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{n\times 1} & 0_{n\times 1} & \cdots & X_{k\_} \end{bmatrix},$$

which are of dimension $k \times 1$, $nk \times 1$, $nk \times 1$, and $nk \times k$ respectively. Hence, the system of the SUR is $X = X\_A + U$. Furthermore, note that $\mathrm{Var}(U) = \Sigma_u \otimes I_n$ is a $nk \times nk$ matrix. The SUR estimator of A for model (1) is defined as follows:

$$\widehat{A}_{SUR} = [X\_^{\top}(\Sigma_u \otimes I_n)^{-1} X\_]^{-1} [X\_^{\top}(\Sigma_u \otimes I_n)^{-1} X]. \tag{2}$$

We start by stating assumptions on the variables and errors that facilitate the development of the asymptotic theory.

[4]The implementation of the IV-OLS method with a pre-test procedure is as follows: for a given DGP, we first test for the common explosive root, and then implement the IV method if the test fails to reject the null of a common explosive root. If the test rejects the null of a common explosive root, we use OLS instead. The test is from Chen, Phillips and Yu (2017) and is described in the Data S1.

Assumption 1. Define the filtration $\mathcal{F}_t = \sigma\{u_t, u_{t-1}, \dots\}$. Let $\{u_t, \mathcal{F}_t\}$ be an independent martingale difference sequence (*mds*) with

$$\mathbb{E}[u_t u_t^\top | \mathcal{F}_{t-1}] = \Sigma_u, \text{ and } \mathbb{E}[\|u_t\|] \geq M \;\; a.s. \;\; \text{for all} \;\; t \leq n, \tag{3}$$

for some $M > 0$ and positive definite matrix Σ_u. Moreover, we have $u_t \sim i.i.d.\mathcal{N}(0, \Sigma_u)$.

Assumption 2. The initial value of the explosive series $\{x_{i,t}\}$ is $x_{i,0} = o_p(1)$ for $i = 1, \dots, k$.

Assumption 1 discusses the property of error terms. In contrast to Phillips and Magdalinos (2008, 2013), we require normal distribution for the error terms. This assumption facilitates the proof of joint convergence, which is important in Lemma 2 (v). Under (3) of Assumption 1, $Q_i(\rho) = \lim_{n\to\infty} \sum_{i=1}^n \rho^{-t} \frac{u_{i,t}}{\sqrt{\sigma_{i,i}}}$ is a.s. not zero. This result is proved in Lai and Wei (1983), and stated in Phillips and Magdalinos (2013). Assumption 2 presents the asymptotic order of the initial condition.

Define $Q_i(\rho) = \lim_{n\to\infty} \sum_{i=1}^n \rho^{-t} \frac{u_{i,t}}{\sqrt{\sigma_{i,i}}}$ and $\widetilde{Q}_i(\rho) = \lim_{n\to\infty} \sum_{i=1}^n \rho^{-(n-t)-1} \frac{u_{i,t}}{\sqrt{\sigma_{i,i}}}$. Both $Q_i(\rho)$ and $\widetilde{Q}_i(\rho)$ are of central importance in the limits of the SUR estimator. As shown in the proof of Theorem 1, the sample variance $\frac{1}{\rho_i^{2n}} \sum_{t=1}^n x_{i,t-1}^2$ is approximated by $\left(\sum_{j=1}^n \rho_i^{-j} u_{i,t}\right)^2$ where $\sum_{j=1}^n \rho_i^{-j} u_{i,t}$ is dominated by the sum of the first κ_n shocks, that is,

$$\begin{aligned}
\frac{1}{\rho_i^{2n}} \sum_{t=1}^n x_{i,t-1}^2 &= \frac{\rho_i^{-2n}}{\rho_i^2 - 1} x_{i,n}^2 + O_p\left(\rho_i^{-2n} n\right) \\
&= \frac{\sigma_{i,i}}{\rho_i^2 - 1} \left[\sum_{i=1}^{\kappa_n} \rho_i^{-t} \frac{u_{i,t}}{\sqrt{\sigma_{i,i}}} + o_{a.s.}\left(n^{-1/2}\right) \right]^2 + O_p\left(\rho_i^{-2n} n\right) \\
&\Rightarrow \frac{\sigma_{i,i} Q_i(\rho_i)^2}{\rho_i^2 - 1},
\end{aligned}$$

The sample covariance $\frac{1}{\rho_i^n} \sum_{t=1}^n x_{i,t-1} u_{j,t}$ is approximated by $\left(\sum_{t=1}^n \rho_i^{-(n-t)-1} u_{j,t}\right)\left(\sum_{j=1}^n \rho_i^{-j} u_{j,t}\right)$ where $\sum_{t=1}^n \rho_i^{-(n-t)-1} u_{j,t}$ is dominated by the sum of last κ_n shocks, that is,

$$\begin{aligned}
\frac{1}{\rho_i^n} \sum_{t=1}^n x_{i,t-1} u_{j,t} = &\left(\sum_{i=\kappa_n+1}^n \rho_i^{-(n-t)-1} u_{j,t} + o_{a.s.}\left(n^{-1/2}\right) \right) \left(\sum_{s=1}^{\kappa_n} \rho_i^{-s} u_{i,s} + o_{a.s.}\left(n^{-1/2}\right) \right) \\
&+ o_p(1) \Rightarrow \sqrt{\sigma_{i,i}\sigma_{j,j}} Q_i(\rho_i) \widetilde{Q}_j(\rho_i).
\end{aligned}$$

Both $\sum_{i=\kappa_n+1}^n \rho_i^{-t} \frac{u_{i,t}}{\sqrt{\sigma_{i,i}}}$ and $\sum_{t=1}^{\kappa_n} \rho_i^{-(n-t)-1} \frac{u_{j,t}}{\sqrt{\sigma_{j,j}}}$ are $o_{a.s.}\left(n^{-1/2}\right)$ as the integer-valued sequence κ_n satisfying $\sum_{i=1}^\infty n\rho_i^{-2\kappa_n} < \infty$ and $\sum_{n=1}^\infty n\rho_i^{-2n+2\kappa_n}$ as $n \to \infty$. This result is stated in Lemma 1. Hence, by employing the sample splitting argument as used in

Magdalinos and Phillips (2009), we have the the limiting theory for $\widehat{A}_{SUR} - A$ in the following Theorem.

Theorem 1. Under Assumptions 1 and 2, the SUR estimator of model (1) has the following limit as $n \to \infty$:

$$\begin{cases} \mathrm{diag}(\rho_1^n, \ldots, \rho_k^n)(\widehat{A}_{SUR} - A) \Rightarrow \frac{1}{\xi_0(\boldsymbol{\rho})}[\xi_1(\boldsymbol{\rho}), \ldots, \xi_k(\boldsymbol{\rho})]^\top & \text{if } \rho_i \neq \rho_j \text{ for } i,j = 1, \ldots, k \\ \rho^n(\widehat{A}_{SUR} - A) \Rightarrow (\rho^2 - 1)\left(\frac{\widetilde{Q}_1(\rho)}{Q_1(\rho)}, \ldots, \frac{\widetilde{Q}_k(\rho)}{Q_k(\rho)}\right)^\top & \text{if } \rho_i = \rho > 1 \text{ for } i = 1, \ldots, k \end{cases},$$

where $\boldsymbol{\rho} = [\rho_1, \ldots, \rho_k]^\top$, $\xi_i(\boldsymbol{\rho}) := \sum_{j=1}^k \sqrt{\frac{\sigma_{j,j}}{\sigma_{i,i}}} \frac{1}{Q_i(\rho_i)} \left\| \begin{bmatrix} \frac{\sigma^{1,1}}{\rho_1^2 - 1} & \cdots & \sigma^{1,j}\widetilde{x}_j(\rho_1) & \cdots & \frac{\sigma^{1,k}}{\rho_1\rho_k - 1} \\ \vdots & & \vdots & & \vdots \\ \frac{\sigma^{k,1}}{\rho_k\rho_1 - 1} & \cdots & \sigma^{k,j}\widetilde{x}_j(\rho_k) & \cdots & \frac{\sigma^{k,k}}{\rho_k^2 - 1} \end{bmatrix} \right\|$

for $i = 1, \ldots, k$, and $\xi_0(\boldsymbol{\rho}) := \left\| \begin{bmatrix} \frac{\sigma^{1,1}}{\rho_1^2 - 1} & \cdots & \frac{\sigma^{1,k}}{\rho_1\rho_k - 1} \\ \vdots & & \vdots \\ \frac{\sigma^{k,1}}{\rho_k\rho_1 - 1} & \cdots & \frac{\sigma^{k,k}}{\rho_k^2 - 1} \end{bmatrix} \right\|$ with $\sigma^{i,j} := [\Sigma_u^{-1}]_{i,j}$.

Remark 1. Theorem 1 implies that $\rho_i^n(\widehat{\rho}_{i,SUR} - \rho_i) \Rightarrow \frac{\xi_i(\boldsymbol{\rho})}{\xi_0(\boldsymbol{\rho})}$ in $i = 1, \ldots, k$ for the case of distinct explosive roots and $\rho_i^n(\widehat{\rho}_{i,\mathrm{SUR}} - \rho_i) \Rightarrow (\rho^2 - 1)\frac{\widetilde{Q}_i(\rho)}{Q_i(\rho)}$ in the case of a common explosive root. It indicates that the SUR estimator is consistent since $\widehat{\rho}_{i,SUR} = \rho_i + o_p(1)$ in the cases of both distinct explosive roots and a common explosive root. Unlike the SUR estimator, the OLS estimator is inconsistent in the case of a common explosive root from Phillips and Magdalinos (2013) such that $\widehat{\rho}_{i,OLS} = \rho_i + O_p(1)$. Such inconsistency is explained by the endogeneity induced by the coordinate rotation when deriving the asymptotics for $\hat{\rho}_{i,\mathrm{OLS}}$.

Remark 2. For example, in the case of a bivariate common explosive root, different from the OLS estimator, we do not use coordinate rotation to derive the sample moments limits, since the sample variance limit for the SUR estimator is invertible. In particular, the proof of Theorem 1 shows that the sample variance of the SUR estimator in the limit is

$$\rho^{-n}\begin{bmatrix} \sigma^{1,1}X_{1\_}^\top X_{1\_} & \sigma^{1,2}X_{1\_}^\top X_{2\_} \\ \sigma^{2,1}X_{2\_}^\top X_{1\_} & \sigma^{2,2}X_{2\_}^\top X_{2\_} \end{bmatrix} \to_{a.s.} \begin{bmatrix} \sigma^{1,1}\frac{\sigma_{1,1}}{\rho^2-1}Q_1(\rho)^2 & \sqrt{\sigma^{1,2}\sigma^{2,1}}\frac{\sqrt{\sigma_{1,1}\sigma_{2,2}}}{\rho^2-1}Q_1(\rho)Q_2(\rho) \\ \sqrt{\sigma^{1,2}\sigma^{2,1}}\frac{\sqrt{\sigma_{1,1}\sigma_{2,2}}}{\rho^2-1}Q_1(\rho)Q_2(\rho) & \sigma^{2,2}\frac{\sigma_{2,2}}{\rho^2-1}Q_2(\rho)^2 \end{bmatrix}. \tag{4}$$

where $\Sigma_u^{-1} = \begin{bmatrix} \sigma^{1,1} & \sigma^{1,2} \\ \sigma^{2,1} & \sigma^{2,2} \end{bmatrix}$. In contrast, the sample covariance limit (5) for the OLS estimator is singular since

$$\rho^{-n}\begin{bmatrix} X_{1\_}^\top X_{1\_} & X_{1\_}^\top X_{2\_} \\ X_{2\_}^\top X_{1\_} & X_{2\_}^\top X_{2\_} \end{bmatrix} \to_{a.s.} \begin{bmatrix} \frac{\sigma_{1,1}}{\rho^2-1}Q_1(\rho)^2 & \frac{\sqrt{\sigma_{1,1}\sigma_{2,2}}}{\rho^2-1}Q_1(\rho)Q_2(\rho) \\ \frac{\sqrt{\sigma_{1,1}\sigma_{2,2}}}{\rho^2-1}Q_1(\rho)Q_2(\rho) & \frac{\sigma_{2,2}}{\rho^2-1}Q_2(\rho)^2 \end{bmatrix}. \tag{5}$$

Clearly, by incorporating the information from the error covariance, the sample covariance limit (4) is invertible.

Remark 3. The SUR estimator is not feasible since Σ_u is unknown. Hence, we replace it with its OLS estimator in application. This replacement yields good finite sample performance, which is presented in the simulation section.

For the purpose of generalization, we also consider the VAR system with intercepts. Following Phillips, Shi, and Yu (2014), we introduce the localizing drift term in the system. In particular, the data are generated as follows

$$X_t = \mu + RX_{t-1} + u_t, \tag{6}$$

where $\mu = [\mu_1, \ldots, \mu_k]^\top$ with $\mu_i = \widetilde{\mu}_i n^{-\eta}$ for $\eta \geq 0$. The drift term depends on the sample size, since the constant term is sample size or frequency dependent for financial time series, see the discussion in Phillips *et al.* (2014).

The system of the SUR for model (6) is written as $X = X\_A + U$, where $X\_$ is now defined as

$$X\_ = \begin{bmatrix} 1_{n\times 1} & X_{1\_} & 0_{n\times 1} & 0_{n\times 1} & \cdots & 0_{n\times 1} & 0_{n\times 1} \\ 0_{n\times 1} & 0_{n\times 1} & 1_{n\times 1} & X_{2\_} & \cdots & 0_{n\times 1} & 0_{n\times 1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{n\times 1} & 0_{n\times 1} & 0_{n\times 1} & 0_{n\times 1} & \cdots & 1_{n\times 1} & X_{k\_} \end{bmatrix}.$$

and $A = [\mu_1, \rho_1, \ldots, \mu_k, \rho_k]^\top$. The SUR estimator of A for model (6) is

$$\widehat{A}_{\text{SUR}} = [X_\_^\top (\Sigma_u \otimes I_n)^{-1} X\_]^{-1} [X_\_^\top (\Sigma_u \otimes I_n)^{-1} X]. \tag{7}$$

We present the limiting behaviour of SUR estimator $\widehat{A}_{\text{SUR}}$ in the following theorem.

Corollary 1. The SUR estimator for model (6) as $n \to \infty$ is

$$\begin{cases} \text{diag}\left(n^{\frac{1}{2}}, \rho_1^n, n^{\frac{1}{2}}, \rho_2^n, \ldots, n^{\frac{1}{2}}, \rho_k^n\right)\left(\widehat{A}_{SUR} - A\right) \\ \Rightarrow \left[B_1(1), \frac{\xi_1(\boldsymbol{\rho})}{\xi_0(\boldsymbol{\rho})}, \ldots, B_p(1), \frac{\xi_k(\boldsymbol{\rho})}{\xi_0(\boldsymbol{\rho})}\right]_{2k\times 1}^\top & \text{if } \rho_i \neq \rho_j \text{ for } i,j = 1, \ldots, k \\ \qquad \text{diag}\left(n^{\frac{1}{2}}, \rho^n, n^{\frac{1}{2}}, \rho^n, \ldots, n^{\frac{1}{2}}, \rho^n\right)\left(\widehat{A}_{SUR} - A\right) \\ \Rightarrow \left[B_1(1), (\rho^2 - 1)\frac{\widetilde{Q}_1(\rho)}{Q_1(\rho)}, \ldots, B_p(1), (\rho^2 - 1)\frac{\widetilde{Q}_p(\rho)}{Q_p(\rho)}\right]_{2k\times 1}^\top & \text{if } \rho_i = \rho > 1 \text{ for } i = 1, \ldots, k \end{cases},$$

where $B_i(1)$, for $i = 1, \ldots, k$ are Brownian motions with variance $\sigma_{i,i}$. $\xi_0(\boldsymbol{\rho})$ and $\xi_i(\boldsymbol{\rho})$ are the same as in Theorem 1.

Remark 4. The asymptotics of $\hat{\rho}_{i,\text{SUR}}$ are the same as those in Theorem 1. This is because the magnitude of the localizing drift term is $O_p(n^{-\eta})$. We assume that the drift term is asymptotically negligible, hence it will not be a dominant component in the limit.

Figure 1 presents the finite-sample distribution of the SUR estimator and the OLS estimator for model (6) with distinct explosive roots. Apparently, the OLS estimator

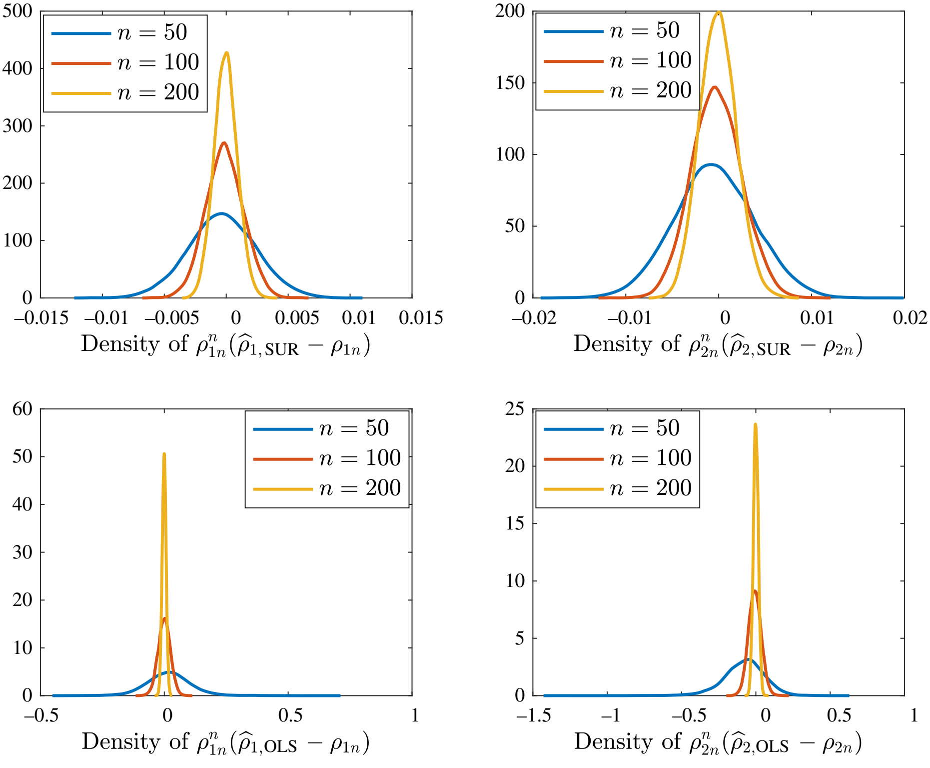

Figure 1. Finite-sample distribution of SUR and OLS estimators for DGP (6) with distinct explosive roots, where $\rho_1 = 1.005$, $\rho_2 = 1.01$, and $n = 50,100,200$. The configuration of the remaining parameters is the same as that in Table 1 [Colour figure can be viewed at *wileyonlinelibrary.com*]

clearly produces larger variance than the SUR estimator as shown in Table 1. As n increases from 50 to 200, the variance of both OLS and SUR estimators are decreasing. Figure 2 presents finite-sample distribution of SUR estimator and the OLS estimator for model (6) with a common explosive root. Similar to the distinct explosive root case in 1, the OLS estimate produces larger bias and variance than the SUR estimate in Figure 2. In particular, since the OLS estimator of ρ_1 is left skewed and the OLS estimator of ρ_2 is right skewed, the OLS estimator of ρ_1 is upward biased and the OLS estimator of ρ_2 is downward biased as shown in Table 1. In the Data S1, Figures C1 and C2 present the finite-sample distribution of the intercept estimate for the case of distinct explosive roots and a common explosive root respectively. The densities of the intercept SUR estimate are symmetric, which indicates that the intercept estimates are unbiased. The OLS estimate produces larger variance than the SUR estimate in the cases of both distinct explosive roots and a common explosive root.

III. Seemingly unrelated regression of VAR with multiple mildly explosive regressors

The OLS inconsistency problem also occurs in VAR models with common mildly explosive regressors. The mildly explosive process is an important tool in characterizing

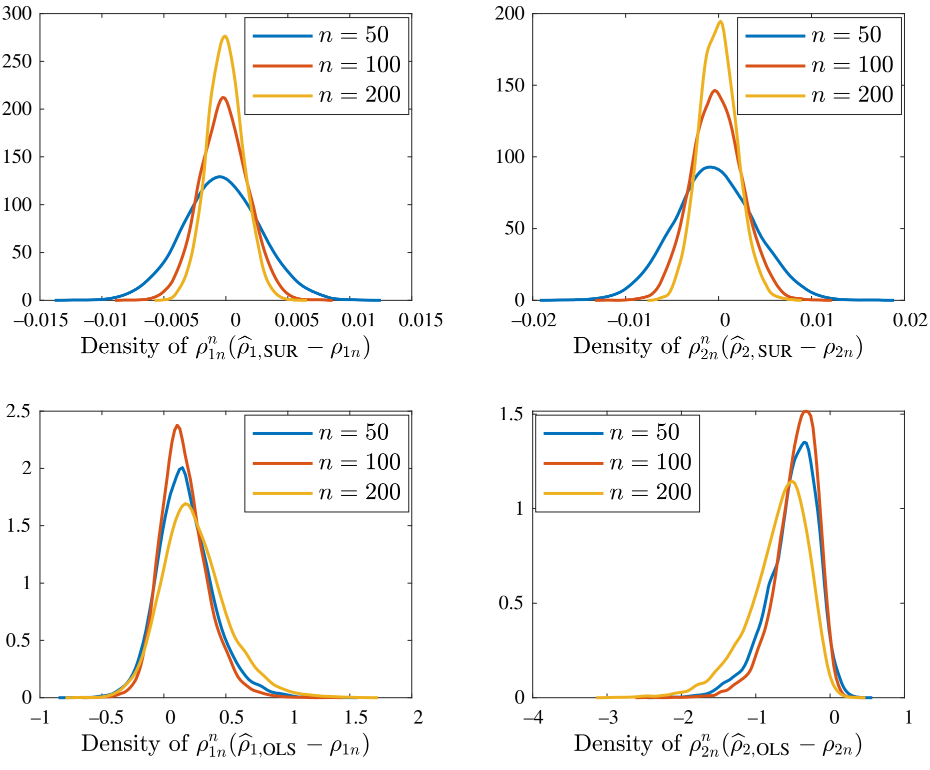

Figure 2. Finite-sample distribution of SUR and OLS estimators for DGP (6) with common explosive roots, where $\rho_1 = \rho_2 = 1.01$, and $n = 50,100,200$. The configuration of the remaining parameters is the same as that in Table 1 [Colour figure can be viewed at *wileyonlinelibrary.com*]

asset price bubbles and plays a fundamental role in the bubble detection literature (see Phillips *et al.*, 2011, 2015a, 2015b; Shi and Phillips, 2021). In this section, we extend the results in Section II to VAR systems with mildly explosive regressors.

Replacing R with R_n in DGP (1), we further assume the data are generated from the following model:

$$X_t = R_n X_{t-1} + u_t. \tag{8}$$

In particular, $R_n = \mathrm{diag}(\rho_{1n}, \ldots, \rho_{kn})$, and the autoregressive root ρ_{in} is defined as $\rho_{in} = 1 + \frac{c_i}{n^{\alpha}}$ with $\alpha \in (0,1)$. The root is moderately deviated from unity and hence covered a larger neighborhood around unity than local-to-unity processes.

We start by stating assumptions on the variables and errors which assist in the development of the asymptotic theory.

Assumption 3. Define the filtration $\mathcal{F}_t = \sigma\{u_t, u_{t-1}, \ldots\}$. Let $\{u_t, \mathcal{F}_t\}$ be independent and identically distributed random variables with

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

TABLE 1

Finite sample performance comparison between OLS, SUR and IV estimators for explosive VAR models with intercept as in DGP (6)

			OLS		*SUR*		*IV*		*OLS-IV with pre-test*	
n		*True*	*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	0.1617	0.0181	0.1071	0.0088	0.0381	0.0195	0.1555	0.0182
	μ_2	0.1131	−0.6488	0.0231	0.1191	0.0103	0.3352	0.0246	−0.5996	0.0232
	ρ_1	1.0050	1.3431	0.4672	−0.0244	0.0005	−0.8968	0.6885	1.2311	0.4783
	ρ_2	1.0100	−4.8674	0.7372	−0.0287	0.0007	4.6688	1.2266	−4.8433	0.7387
100	μ_1	0.1000	0.0393	0.0090	0.0596	0.0041	−0.0514	0.0090	0.0348	0.0090
	μ_2	0.0800	−0.1074	0.0115	0.0630	0.0047	0.0269	0.0114	−0.1007	0.0115
	ρ_1	1.0050	−0.0212	0.0227	−0.0082	0.0001	0.0819	0.0239	−0.0161	0.0227
	ρ_2	1.0100	−0.2366	0.0256	−0.0078	0.0001	0.1554	0.0280	−0.2170	0.0257
200	μ_1	0.0707	0.0584	0.0038	0.0255	0.0018	−0.0417	0.0038	0.0534	0.0038
	μ_2	0.0566	0.0310	0.0048	0.0236	0.0018	−0.0261	0.0048	0.0281	0.0048
	ρ_1	1.0050	−0.0222	0.0009	−0.0016	0.0000	0.0309	0.0009	−0.0196	0.0009
	ρ_2	1.0100	0.0027	0.0005	−0.0008	0.0000	−0.0127	0.0005	0.0019	0.0005
Panel B. Explosive regressor for common explosive roots										
50	μ_1	0.1414	0.0540	0.0177	0.0999	0.0083	−0.0546	0.1659	−0.0492	0.1585
	μ_2	0.1131	0.0604	0.0425	0.1174	0.0104	−0.0776	0.3632	−0.0707	0.4732
	ρ	1.0100	10.779	1.8901	−0.0206	0.0004	−16.544	151.80	−15.178	144.30
		1.0100	−30.603	4.1941	−0.0281	0.0007	47.263	683.07	43.370	649.12
100	μ_1	0.1000	0.0175	0.0093	0.0501	0.0038	−0.0082	0.0079	−0.0069	0.0080
	μ_2	0.0800	0.0136	0.0238	0.0601	0.0048	−0.0074	0.0118	−0.0064	0.0124
	ρ	1.0100	5.9175	0.4985	−0.0055	0.0001	−6.3239	1.0461	−5.7119	1.0187
		1.0100	−16.809	1.1517	−0.0072	0.0001	17.797	2.8374	16.067	2.7531
200	μ_1	0.0707	−0.0054	0.0044	0.0181	0.0015	−0.0020	0.0031	−0.0022	0.0032
	μ_2	0.0566	−0.0007	0.0120	0.0218	0.0019	−0.0041	0.0045	−0.0039	0.0049
	ρ	1.0100	3.3873	0.1316	0.0006	0.0000	−2.7089	0.1837	−2.4041	0.1811
		1.0100	−9.6007	0.3138	0.0007	0.0000	7.5490	0.4678	−2.4041	0.4601

Assumption 4. The initial value of the explosive series $\{x_{i,t}\}$ is $x_{i,0} = o_p(1)$ for $i = 1, \ldots, k$.

Assumption 3 provides the conditions for the error terms. In particular, there is no requirement for the distribution condition. This assumption is the same as Phillips and Magdalinos (2007). We require $o_p(1)$ as an initial condition in Assumption 4. However, we can relax the initial condition to be $o_p\left(n^{\alpha/2}\right)$ as stated in (Phillips and Magdalinos, 2007).

Define $Y_i(c_i) \sim \mathcal{N}(0,1)$ and $\widetilde{Y}_j(c_i) \sim \mathcal{N}(0,1)$. As $Q_i(\rho)$ and $\widetilde{Q}_i(\rho)$ are of central importance in the limits of the SUR estimator for VAR models with explosive regressors, both $Y_i(c_i)$ and $\widetilde{Y}_j(c_i)$ play an important role in the asympototics for VAR models with mildly explosive regressors. We show in the proof of Theorem 2 that the sample variance $\frac{1}{\rho_i^{2n}}\sum_{t=1}^n x_{i,t-1}^2$ is approximated by $\left(\sum_{t=1}^n \rho_{in}^{-j} u_{i,t}\right)^2$ that is,

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

The sample covariance $\frac{1}{\rho_i^n}\sum_{t=1}^n x_{i,t-1}u_{j,t}$ is approximated by $\left(\sum_{t=1}^n \rho_i^{-(n-t)-1}u_{j,t}\right)\left(\sum_{s=1}^n \rho_i^{-s}u_{i,s}\right)$ where

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

Unlike VAR models with explosive regressors, we do not use the sample splitting argument in the proof, as $Y_i(c_i)$ and $\widetilde{Y}_j(c_i)$ are asymptotically independent since

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

We present the limit result of the SUR estimator for the DGP (8).

Theorem 2. Under Assumptions 1 and 2, the SUR estimator for model (8) has the following limit as $n \to \infty$:

$$\begin{cases} \operatorname{diag}(n^\alpha\rho_{1n}^n, \ldots, n^\alpha\rho_{kn}^n)(\widehat{A}_{\mathrm{SUR}} - A) \Rightarrow \frac{1}{\zeta_0(\boldsymbol{c})}[\zeta_1(\boldsymbol{c}), \ldots, \zeta_k(\boldsymbol{c})]^\top & \text{if } c_i \neq c_j \text{ for } i,j = 1, \ldots, k \\ n^\alpha\rho_n^n(\widehat{A}_{\mathrm{SUR}} - A) \Rightarrow 2c\left(\frac{\widetilde{Y}_1(c)}{Y_1(c)}, \ldots, \frac{\widetilde{Y}_k(c)}{Y_k(c)}\right)^\top & \text{if } c_i = c \text{ for } i = 1, \ldots, k \end{cases},$$

where $\boldsymbol{c} = [c_1, \ldots, c_k]^\top$, $\zeta_i(\boldsymbol{c}) := \sum_{j=1}^{k} \sqrt{\frac{\sigma_{j,j}}{\sigma_{i,i}}} \frac{1}{Y_i(c_i)} \left\| \begin{bmatrix} \frac{\sigma^{1,1}}{c_1+c_1} & \cdots & \frac{\sigma^{1,j}}{c_1}\widetilde{Y}_j(c_i) & \cdots & \frac{\sigma^{1,k}}{c_1+c_k} \\ \vdots & & \vdots & & \vdots \\ \frac{\sigma^{k,1}}{c_k+c_1} & \cdots & \frac{\sigma^{k,j}}{c_k}\widetilde{Y}_j(c_k) & \cdots & \frac{\sigma^{k,k}}{c_k+c_k} \end{bmatrix} \right\|$,

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

Remark 5. The limiting distribution of $n^\alpha \rho_{in}^n \left(\hat{\rho}_{in,\mathrm{SUR}} - \rho_{in}\right)$ is asymptotically Cauchy. There is no Gaussian assumption for the error terms. This is different from the explosive regressor case in the previous section, where the limiting distribution depends on the error distributions. Hence, the invariance principle (refer to Phillips and Magdalinos, 2007) applies in our mildly explosive VAR models.

Remark 6. Unlike the result for explosive VAR models (1) in previous section, where the converge rate is ρ_i^n for $\hat{\rho}_{i,SUR}$, the convergence rate for $\hat{\rho}_{in,SUR}$ is $n^\alpha \rho_{in}^n$ for mildly explosive VAR models (8). The differences in the convergence rate depend on the configuration of the autoregressive coefficient.

Remark 7. Moreover, SUR estimator $\hat{\rho}_{in,\mathrm{SUR}}$ is asymptotically mixture normal, and its associated Wald test is chi-squared distributed. Hence, we can use Wald test statistics for inference in the applied work.

Next, we include the intercept term in the multivariate system. The model is generated as follows

$$X_t = \mu + R_n X_{t-1} + u_t, \tag{9}$$

where $\mu = [\mu_1, \ldots, \mu_k]^\top$. Following Phillips *et al.* (2014), we set $\mu_i = \widetilde{\mu}_i n^{-\eta}$ for $\eta \geq 0$. The associated asymptotics are as follows:

Corollary 2. The SUR estimator for model (9) as $n \to \infty$ is

$$\begin{cases} \mathrm{diag}\left(n^{\frac{1}{2}}, n^\alpha \rho_{1n}^n, n^{\frac{1}{2}}, n^\alpha \rho_{2n}^n, \ldots, n^{\frac{1}{2}}, n^\alpha \rho_{kn}^n\right)\left(\widehat{A}_{\mathrm{SUR}} - A\right) \\ \Rightarrow \left[B_1(1), \zeta_1(\boldsymbol{c})\zeta_0(\boldsymbol{c}), B_2(1), \frac{\zeta_2(\boldsymbol{c})}{\zeta_0(\boldsymbol{c})}, \ldots, B_k(1), \frac{\zeta_k(\boldsymbol{c})}{\zeta_0(\boldsymbol{c})}\right]_{2k\times 1}^\top & \text{if } c_i \neq c_j \text{ for } i,j = 1, \ldots, k \\ n^\alpha \rho_n^n(\widehat{A}_{\mathrm{SUR}} - A) \Rightarrow 2c\left(\frac{\widetilde{Y}_1(c)}{Y_1(c)}, \ldots, \frac{\widetilde{Y}_k(c)}{Y_k(c)}\right)^\top & \text{if } c_i = c \text{ for } i = 1, \ldots, k \end{cases},$$

where $B_i(1)$, for $i = 1, \ldots, k$ are defined in the Appendix. $\zeta_0(\boldsymbol{c})$ and $\zeta_i(\boldsymbol{c})$ are the same as in Theorem 2.

Remark 8. Similar to the explosive VAR models (1), the SUR estimator $\hat{\rho}_{in,\mathrm{SUR}}$ with drift shares the same limit as that without drift. This is because the drift term is asymptotically negligible in the system.

We compare the finite-sample distribution of the proposed SUR estimator and the OLS estimator. Figure 3 presents the finite-sample distribution of SUR estimator and the OLS estimator for DGP (9) with distinct explosive roots. Similar to the explosive VAR case in previous section, the OLS estimator produces larger variance than the SUR

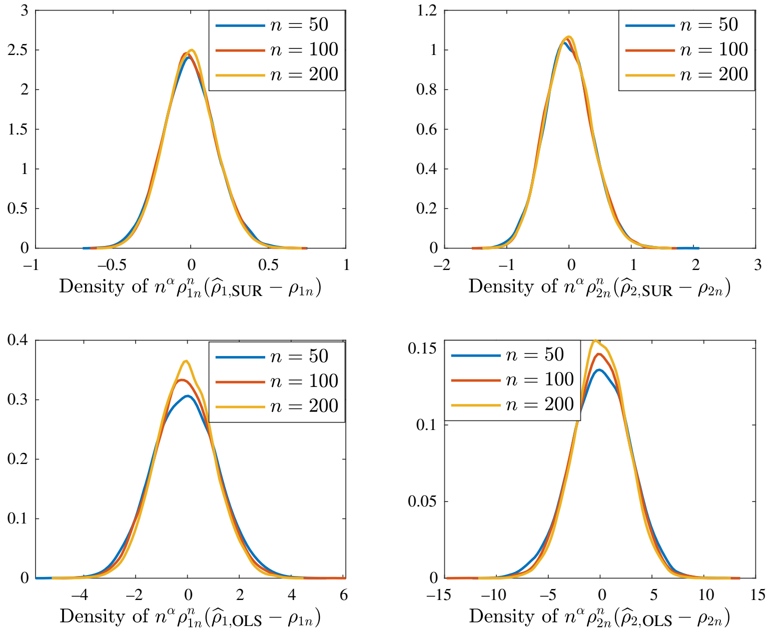

Figure 3. Finite-sample distribution of SUR and OLS estimators for DGP (9) with distinct explosive roots, where $c_1 = 1$, $c_2 = 2$, $\alpha = 0.95$, and $n = 50, 100, 200$. The configuration of the remaining parameters is the same as that in Table 2 [Colour figure can be viewed at *wileyonlinelibrary.com*]

estimator as shown in Table 2. As n increases from 50 to 200, the variance of the both OLS and SUR estimator increase. Figure 4 presents the finite sample distribution of SUR estimator and the OLS estimator for model (6) with a common explosive root. Similar to the VAR with an explosive regressor case in the previous Section, the OLS estimator produces larger bias and variance than the SUR estimator as shown in Figure 4. Since the OLS estimate of ρ_1 is left skewed and the OLS estimate of ρ_2 is right skewed, the OLS estimate of ρ_1 has upward bias and the OLS estimate of ρ_2 has downward bias, as shown in Table 2. In Data S1, Figures C3 and C4 present the finite-sample distribution of the intercept estimate for the cases of distinct explosive roots and common explosive root, respectively. The densities of the intercept estimate are symmetric, which indicates that the intercept estimates are unbiased. The OLS estimator produces larger variance than the SUR estimator in the cases of both distinct explosive roots case and a common explosive root.

IV. Simulations

First we compare the finite-sample performance between OLS estimator, SUR estimator, and IV estimator[5] in terms of 100 × bias, 100 × variance and 100 × mean squared error, for various values of $n = 50, 100, 200$. In particular, the IV estimator from

[5]The IV-OLS estimator with a pre-test procedure is considered in Table 1. Although it improves the finite-sample performance of OLS and IV to some extent, it is not considered in other tables and figures since the improvement is insignificant.

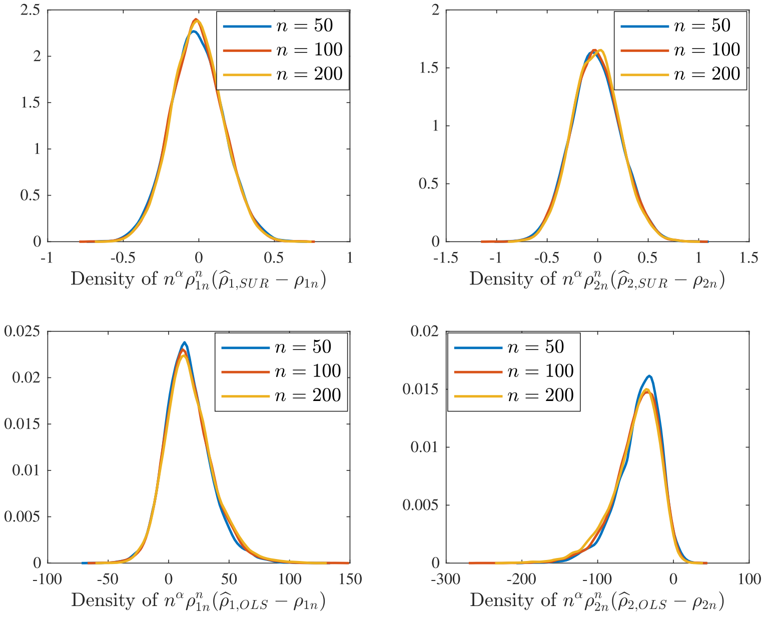

Figure 4. Finite-sample distribution of SUR and OLS estimators for DGP (9) with common explosive roots, where $c_1 = c_2 = 1, \alpha = 0.95, n = 50{,}100, 200$. The configuration of the remaining parameters is the same as that in Table 2 [Colour figure can be viewed at *wileyonlinelibrary.com*]

Phillips and Magdalinos (2013) is defined as $\widehat{R}_{\mathrm{IV}} = (\sum_{t=1}^{n-p} X_t X_{t+p}^{\top})(\sum_{t=1}^{n-p} X_{t-1} X_{t+p}^{\top})^{-1}$, for $p \in \{0, 1, 2, \ldots\}$. In the section, we set $p = 1$ for the IV estimator. Since the variance-covariance matrix Σ_u is not available when using SUR estimation, we propose using the feasible SUR estimator for the implementation by the following procedures.

Step 1: Estimate Σ_u by OLS instead. Then, compute the estimated error terms u_t.

Step 2: Calculate the SUR estimate of the autoregressive coefficient matrix using the estimate of Σ_u from step 1. Then, compute the estimated error terms u_t using the SUR estimate.

Step 3: Repeat steps 1 and 2 until the SUR estimate of R converges.

First, we consider DGP (6) for $k = 2$. We choose $\rho_1 = 1.005$ and $\rho_2 = 1.01$ for $n = 50{,}100, 200$. In addition, we set $\Sigma_u = \begin{bmatrix} 0.0012 & 0.0012 \\ 0.0012 & 0.0015 \end{bmatrix}$ and $\mu = [\mu_1, \mu_2]^{\top} = \left[1 \times n^{-1/2}, 0.8 \times n^{-1/2}\right]^{\top}$. The number of replications is 10,000. The parameters are calibrated to our CBOT agricultural commodity application. In particular, we calibrate the x_t process to the sample period between 3 July 2010 and 29 January 2011. The selection of the sample period is guided by the literature as stated in the empirical section. We estimate the autoregressive coefficient for each of the commodity price series, and forms the autoregressive coefficient interval. The chosen $\rho_1 = 1.005$, and $\rho_2 = 1.01$ lie in the interval. The initial values of the explosive processes are set to $x_0 = [0, 0]^{\top}$.

The results are reported in Table 1. For the case of distinct explosive roots case, first, as the sample size increases, all autoregressive coefficient estimates have smaller bias and variance. Second, among these estimates, the SUR estimate has the smallest bias and

TABLE 2

Finite sample performance comparison between OLS, SUR and IV estimators for explosive VAR models with intercept as in DGP (9)

			OLS			*SUR*			*IV*		
n		*True*	*100 × bias*	*100 × var*	*100 × mse*	*100 × bias*	*100× var*	*100 × mse*	*100 × bias*	*100 × var*	*100 × mse*
Panel A. Mildly explosive regressor for distinct explosive roots											
50	μ_1	0.1414	0.1023	0.0139	0.0140	0.0413	0.0068	0.0068	−0.0625	0.0142	0.0142
	μ_2	0.1131	0.0689	0.0175	0.0175	0.0345	0.0067	0.0067	−0.0445	0.0178	0.0178
	ρ_1	1.0243	−0.0618	0.0085	0.0086	−0.0052	0.0002	0.0002	0.0906	0.0096	0.0097
	ρ_2	1.0486	0.0119	0.0042	0.0042	−0.0021	0.0001	0.0001	−0.0428	0.0049	0.0049
100	μ_1	0.1000	0.0783	0.0070	0.0071	0.0265	0.0033	0.0033	−0.0434	0.0071	0.0071
	μ_2	0.0800	0.0557	0.0089	0.0089	0.0227	0.0033	0.0033	−0.0328	0.0090	0.0090
	ρ_1	1.0126	−0.0332	0.0018	0.0018	−0.0021	0.0000	0.0000	0.0385	0.0019	0.0019
	ρ_2	1.0252	0.0078	0.0008	0.0008	−0.0007	0.0000	0.0000	−0.0188	0.0009	0.0009
200	μ_1	0.0707	0.0540	0.0035	0.0035	0.0181	0.0017	0.0017	−0.0335	0.0035	0.0035
	μ_2	0.0566	0.0407	0.0043	0.0044	0.0154	0.0016	0.0016	−0.0250	0.0043	0.0043
	ρ_1	1.0065	−0.0147	0.0004	0.0004	−0.0008	0.0000	0.0000	0.0195	0.0004	0.0004
	ρ_2	1.0130	0.0036	0.0002	0.0002	−0.0002	0.0000	0.0000	−0.0090	0.0002	0.0002
Panel B. Mildly explosive regressor for common explosive roots											
50	μ_1	0.1414	0.0300	0.0166	0.0166	0.0655	0.0071	0.0071	−0.0283	0.1632	0.1633
	μ_2	0.1131	0.0413	0.0418	0.0419	0.0763	0.0088	0.0089	−0.0456	0.4173	0.4174
	ρ	1.0243	11.663	1.9311	3.2913	−0.0095	0.0002	0.0002	−16.995	252.20	255.09
		1.0243	−33.092	4.2389	15.190	−0.0122	0.0003	0.0003	47.875	1238.2	1261.1
100	μ_1	0.1000	0.0113	0.0091	0.0091	0.0423	0.0035	0.0035	−0.0041	0.0077	0.0077
	μ_2	0.0800	0.0087	0.0237	0.0237	0.0507	0.0045	0.0045	−0.0042	0.0116	0.0116
	ρ	1.0126	6.0856	0.5012	0.8716	−0.0040	0.0000	0.0000	−6.2172	1.0551	1.4416
		1.0126	−17.283	1.1549	4.1418	−0.0051	0.0001	0.0001	17.493	2.8551	5.9151
200	μ_1	0.0707	0.0053	0.0046	0.0046	0.0293	0.0018	0.0018	−0.0074	0.0034	0.0034
	μ_2	0.0566	0.0079	0.0121	0.0121	0.0352	0.0022	0.0022	−0.0083	0.0048	0.0048
	ρ	1.0065	3.1426	0.1290	0.2278	−0.0018	0.0000	0.0000	−2.8441	0.1796	0.2605
		1.0065	−8.9136	0.3103	1.1048	−0.0023	0.0000	0.0000	7.9256	0.4601	1.0883

variance. For the case of common explosive root, first, the OLS estimate has upward bias for ρ_1 and downward bias for ρ_2. In contrast, the IV estimate has downward bias for ρ_1 and upward bias for ρ_2. Second, the SUR estimator is able to deliver the best finite-sample performance in terms of bias and variance. Third, both bias and variance decrease as the sample size increases for all four estimates.

Second, we consider the mildly explosive process of x_t. The data are generated from model (9) for $k = 2$ with $c_1 = 1$, $c_2 = 2$, $\alpha = 0.95$ for the case of distinct explosive roots case and $c = 1$, $\alpha = 0.95$ for the case of a common explosive root. The configuration of the remaining parameters is the same as that in Table 1. The results are reported in Table 2. For the case of the distinct explosive roots, SUR estimator has the smallest variance and bias. For the case of a common explosive root, Table 2 shows that the SUR estimator is consistent and outperforms the other estimators. Similar to the case of an explosive regressor, for ρ_2, the OLS estimate is inconsistent and has downward bias with a larger variance, and the IV estimate is upward biased.

We conclude that in the cases of both distinct explosive roots and a common explosive root, the SUR estimator delivers the estimate with a smaller bias, variance and mean squared error than the OLS and IV estimators in nearly all cases.

Next, we perform a structural analysis of the forecast error variance decomposition for our explosive VAR models. The forecast error variance decomposition (FEVD) is intended to decompose the variance of the forecast error and to compute the contributions of specific exogenous shocks. In particular, FEVD computes the percentage of variation in each x_t at time $t + k$ that is due to shocks in each x_t (including itself) at time t. In Figure 5, we compare the performance of the forecast error variance decomposition between the OLS estimator, SUR estimator and IV estimator in the case of a common explosive root. Figure 5 reports FEVD at 20-month horizons ahead. The data generating process is the same as in Table 1. On the one hand, when using the SUR or IV estimator, we can attribute most of the forecast error variance of variable 1 to a shock to variable 1. For the OLS estimator, after several periods, the contribution of variable 1 to variable 1 decreases and the contribution of variable 2 to variable 1 increases. On the other hand, for the OLS and SUR estimators, we can attribute most of the forecast error variance of variable 2 to a shock to variable 2. While for the IV estimator, the contribution of shock 1 to variable 2 is increasing while the contribution of shock 2 to variable 2 is decreasing over the sample period.

V. Empirical application

Agricultural commodities are a popular investment for optimizing portfolio return and diversifying portfolio risk. Except for market fundamentals (supply, demand shifters), these commodity prices tend to respond to financial and economic events and hence move together in the market (Etienne *et al.*, 2014; Chavas and Li, 2020; Li and Chavas, 2023). Moreover, the financialization of commodity markets has increased the integration in agricultural commodity markets and across different commodity markets (Tang and Xiong, 2012). Therefore, it is interesting to explore and analyse the impact of agricultural commodity shocks in the market. There are some studies applying forecast error variance decomposition to agricultural commodities and other markets (Yang,

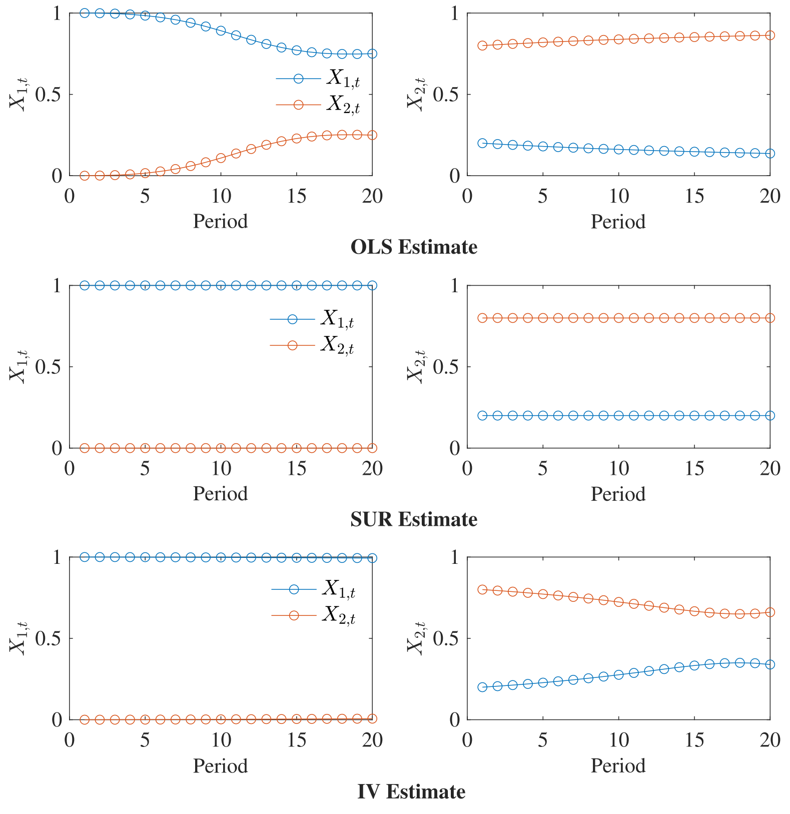

Figure 5. FEVD performance by using different estimators [Colour figure can be viewed at *wileyonlinelibrary.com*]

Balyeat, and Leatham, 2005; Wang and McPhail, 2014; Antonakakis and Kizys, 2015; Luo and Ji, 2018; Balli *et al.*, 2019; Xiao *et al.*, 2020). Since impacts may be intense during financial bubbles, in this empirical section, we investigate the interrelationships among the price shocks and study the variation between the essential CBOT agricultural commodities using FEVD during such period.

As we demonstrated in the simulation section, the use of an OLS estimator in FEVD yields a possibly incorrect result when computing the decomposition of forecast error variance. In this section, we demonstrate the difference in FEVD when using different estimators. We apply OLS, SUR and IV estimators to commodity price data using the closing prices of soybeans, rice, wheat, corn and sugar futures contracts. Weekly data for these prices were downloaded from the Wind database.

The sample period is between 3 July 2010, and 29 January 2011. The selection of sample period is guided by Li *et al.* (2017a), who studied the relationship between commodity price bubbles and macroeconomic factors. Etienne *et al.* (2014) and Li, Li, and Chavas (2017b) found that bubbles primarily occurred over the period 2010–11. Hence, our focus is on the period between 3 July 2010, and 29 January 2011. Figure 6 plots the closing prices of the selected commodities. We observe price run-ups in the sample period. In addition, we use the method in Chen *et al.* (2017) to explore the presence of

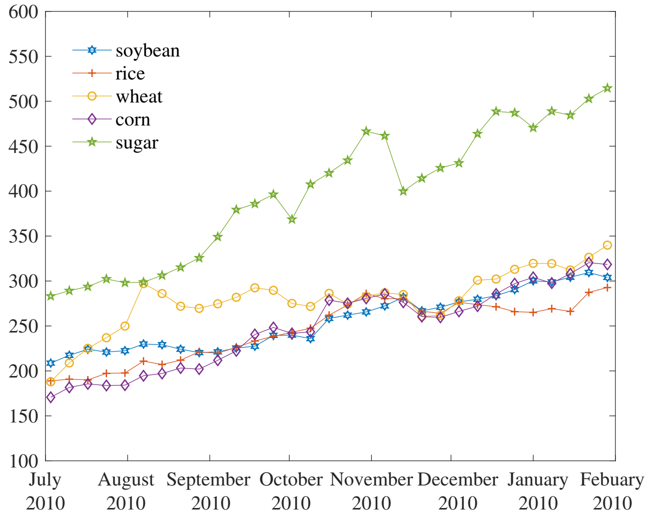

Figure 6. Closing prices of the CBOT agricultural commodity futures between July 3, 2010 and January 29, 2011 [Colour figure can be viewed at *wileyonlinelibrary.com*]

explosive behaviour in the data. In particular, we find that soybeans, rice, wheat, corn, and sugar exhibit explosive behaviour[6], hence, we further study the structural behaviour of these data.

Figure 7 reports the results of FEVD using the OLS, SUR and IV estimators. Consider soybean prices for example. When using the OLS estimator, the forecast error variance is mainly explained by the soybean price shock itself in the first few months. In particular, at the horizon of 1 month, the variance is almost entirely explained by the price itself. The contribution of soybean price shocks to its forecast error variance declines of the period studied, while the contribution of wheat prices increases. Rice, corn, and sugar contribute very little to soybean prices, with maximum contribution less than 10%. When using the SUR estimator, the variance of soybean oil is almost entirely explained by its own shock. The rice shock, wheat shock, corn shock and sugar shock contribute very little to the soybean price. When using the IV estimator, at the horizon of 1 month, the result is the same as that using the SUR estimator. However, at horizons other than 1 month, most of the variance is explained by corn and sugar. The wheat price contributes little to soybean price variance.

For rice, when using the OLS estimator, at the 1-month horizon, the variance is almost entirely explained by the rice shock itself. Thereafter, the contribution of corn and sugar price shocks increases, and the contribution of rice shocks declines to approximately 30% during the period of study. The contribution of soybeans and wheat is approximately zero for all of the considered period. When using the SUR estimator for rice, over 90% of the variance is explained by its own shock; 10% of the variance is explained by soybeans. The wheat, corn and sugar shocks contribute very little to the rice price. When using the IV estimator for rice, at a 1-month horizon, the result is the same as when using the OLS

[6]The results of the analysis are reported in the Appendix.

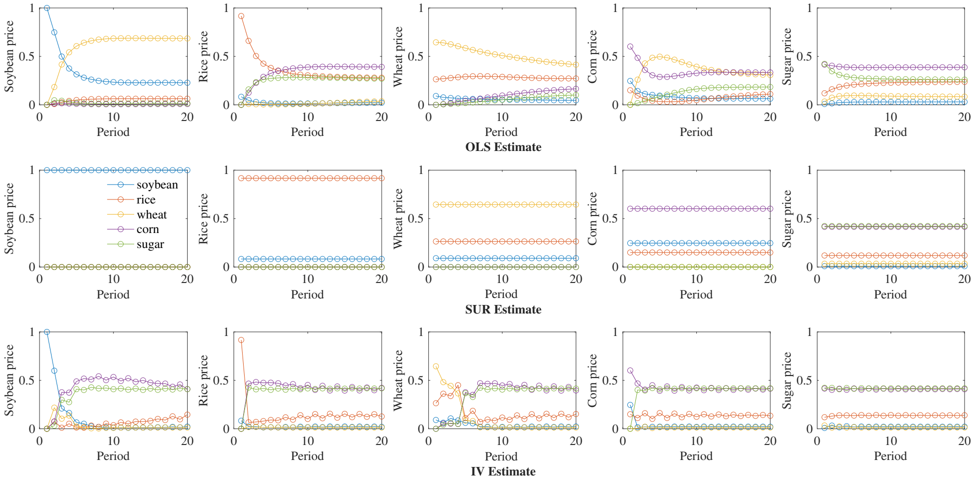

Figure 7. FEVD for price of soybeans, rice, wheat, corn, and sugar in CBOT agricultural commodity market. The top panels are estimation using OLS, the middle using SUR, and the bottom panels are using IV [Colour figure can be viewed at *wileyonlinelibrary.com*]

and SUR estimators. However, at horizons other than 1 month, most of the variance is explained by corn and sugar prices, approximately 80%. Rice contributes approximately 20% to its own variance. Soybeans and wheat contribute little to rice's variance.

Regarding the OLS estimates for wheat, the variance is almost entirely explained by the wheat shock itself and the rice shock. Soybeans and sugar contribute very little to sugar's variance. Over the next few periods considered, the contribution of the wheat price itself decreases while that of corn increases. When applying the SUR estimator to wheat, over all the periods considered, the results are similar to the first month of the OLS estimates. For sugar when using the IV estimator, at a horizon of 1 month, the result is the same as that using the OLS and SUR estimators. However, at horizons other than 1 month, most of the variance is explained by corn and sugar, approximately 80%. Rice accounts for approximately 15% of wheat's variance. Soybeans and sugar contribute little to wheat's variance.

Regarding the OLS estimates for corn, at a horizon of first month, 60% of the variance is explained by the corn shock itself; 30% of the variance is explained by soybean and rice shocks. The wheat shock contribute little to corn's variance. Thereafter, the contribution of corn decreases, and that of wheat increases. Regarding the SUR estimates for corn, over all the periods considered, the result is similar to the first month of the OLS estimates. When using the IV estimator for corn, at a horizon of 1 month, the result is the same as that using the OLS and SUR estimators. However, at horizons other than 1 month, most of the variance is explained by corn and sugar. Rice accounts for approximately 15% of corn's variance. Soybeans and wheat contribute little to corn's variance.

Regarding the OLS estimates for sugar, at a one-month horizon, 80% of the variance is explained by sugar and corn; 20% of the variance is explained by the rice shock. The soybean shock contributes little to corn's variance. Thereafter, the contribution of sugar decreases and that of rice increases. When applying the SUR and IV estimators to sugar, over all periods considered, the result is similar to the first month of the OLS estimates.

Ultimately, there is a significant difference in the FEVD when using different estimators for explosive VAR models. Hence, we suggest the use of the SUR estimator when the focus is on bubble periods.

VI. Conclusion

The VAR models with explosive regressors is of particular interest in studying the structural relationship during the bubble period. When the VAR models with common explosive root is estimated with OLS, the inconsistency problem occurs. The OLS estimator will produce a confusing result and it will further produce a misleading inferential result when analyzing the systems with explosive regressors. Phillips and Magdalinos (2013) discuss and explains the problem, and further propose removing the inconsistency from such system by using the IV estimator. However, we have to test for the common explosive root before applying the IV estimator. In order to avoid the type I error before implementing the IV estimator, we propose a SUR estimator for the VAR models with distinct explosive roots and common explosive root. We show that the SUR estimator is able to deliver the consistent estimator with good finite sample performance. The asymptotic result is provided for both the explosive regressors and the mildly explosive regressors, and for the system with or without intercept terms. Moreover, we demonstrate in simulation that our SUR estimator gives a different result in FEVD, comparing to OLS and IV estimator.

As to the future research work, we expect to see the application of SUR in financial connectedness, particularly in the studying of financial connectedness in the bubble period. Financial connectedness measures the shares of forecast error variation in various locations, due to shocks arising elsewhere (Diebold and Yilmaz, 2009). In the Diebold–Yilmaz framework, we apply the FEVD in studying the connectedness between returns, defaults, contracts, and systems, see Diebold and Yilmaz (2009, 2012, 2014) and Demirer *et al.* (2018). The financial connectedness plays an important role in understanding the financial risk management, hence, we expect to see the performance of our SUR estimator in measuring the financial connectedness.

Appendix

Useful lemmas

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 and Magdalinos (2007, 2008, 2013) and Magdalinos and Phillips (2009).

Lemma 1. Assuming the integer-valued sequence κ_n satisfying $\sum_{i=1}^{\infty} n\rho_i^{-2\kappa_n} < \infty$ and $\sum_{n=1}^{\infty} n\rho_i^{-2n+2\kappa_n} < \infty$ as $n \to \infty$, under Assumptions 1 and 2, given the following data generating process for $i = 1, \ldots, k$ as $x_{i,t} = \rho_i x_{i,t-1} + u_{i,t}$, $\rho_i > 1$,

(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.

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

(i)
$$\frac{1}{\rho_i^n} \sum_{t=1}^{n} x_{i,t-1} \Rightarrow \frac{\sqrt{\sigma_{i,i}} Q_i(\rho_i)}{\rho_i - 1}, \tag{A1}$$

(ii)
$$\frac{1}{\rho_i^n} \sum_{t=1}^{n} x_{i,t-1} u_{j,t} \Rightarrow \sqrt{\sigma_{i,i}\sigma_{j,j}} Q_i(\rho_i) \widetilde{Q}_j(\rho_i), \tag{A2}$$

(iii)
$$\frac{1}{\rho_i^{2n}} \sum_{t=1}^{n} x_{i,t-1}^2 \Rightarrow \frac{\sigma_{i,i} Q_i(\rho_i)^2}{\rho_i^2 - 1}, \tag{A3}$$

(iv)
$$\frac{1}{\rho_i^n \rho_j^n} \sum_{t=1}^{n} x_{i,t-1} x_{j,t-1} \Rightarrow \frac{\sqrt{\sigma_{i,i}\sigma_{j,j}}}{\rho_i \rho_j - 1} Q_i(\rho_i) Q_j(\rho_j), \tag{A4}$$

(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.

Lemma 3. Under Assumptions 1 and 2, given the following data generating process for $i = 1, \ldots, k$ as $x_{i,t} = \mu_i + \rho_i x_{i,t-1} + u_{i,t}$, $\rho_i > 1$, where $\mu_i = \widetilde{\mu}_i n^{-\eta_i}$ with $\eta_i \geq 0$. We also have results (i)–(iv) in Lemma 2.

Lemma 4. Under Assumptions 1 and 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:

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

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

(ii)

$$\frac{1}{n^{3\alpha/2}\rho_{in}^{n}}\sum_{t=1}^{n} x_{i,t-1} \Rightarrow \frac{1}{c_i}\sqrt{\frac{\sigma_{i,i}}{2c_i}}Y_i(c_i), \tag{A5}$$

$$\frac{1}{n^{\alpha}\rho_{in}^{n}}\sum_{t=1}^{n} x_{i,t-1}u_{j,t} \Rightarrow \frac{\sqrt{\sigma_{i,i}\sigma_{j,j}}}{2c_i}Y_i(c_i)\,\widetilde{Y}_j(c_i), \tag{A6}$$

$$\frac{1}{n^{2\alpha}\rho_{in}^{2n}}\sum_{t=1}^{n} x_{i,t-1}^{2} \Rightarrow \frac{\sigma_{i,i}}{4c_i^2}[Y_i(c_i)]^2, \tag{A7}$$

$$\frac{1}{n^{2\alpha}\rho_{in}^{n}\rho_{jn}^{n}}\sum_{t=1}^{n} x_{i,t-1}x_{j,t-1} \Rightarrow \frac{1}{2\left(c_i+c_j\right)}\sqrt{\frac{\sigma_{i,i}\sigma_{j,j}}{c_ic_j}}Y_i(c_i)\,Y_j\left(c_j\right). \tag{A8}$$

(iii) Let $\widetilde{Y}(\boldsymbol{\rho}) = [\frac{1}{\sqrt{2c_1}}\widetilde{Y}_j(\rho_1), \ldots, \frac{1}{\sqrt{2c_k}}\widetilde{Y}_j(\rho_k)]^\top$. 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.

Lemma 5. Under Assumptions 1 and 2, given the following data generating process for $\alpha \in (0,1)$, and $c_i > 0$ for $i = 1, \ldots, k$ as $x_{i,t} = \mu_i + \rho_{in}x_{i,t-1} + u_{i,t}, \rho_{in} = 1 + \frac{c_i}{n^\alpha}$, where $\mu_i = \widetilde{\mu}_i n^{-\eta}$ with $\eta \geq 0$, we also have results (i)-(iii) in Lemma 4.

Proof of Theorem 1

Proof. Let

$$\Sigma_u^{-1} := \begin{bmatrix} \sigma^{1,1} & \cdots & \sigma^{1,k} \\ \vdots & & \vdots \\ \sigma^{k,1} & \cdots & \sigma^{k,k} \end{bmatrix}_{k\times k} \quad \text{and} \quad B := \begin{bmatrix} \sigma^{1,1}X_{1\_}^\top X_{1\_} & \cdots & \sigma^{1,k}X_{1\_}^\top X_{k\_} \\ \vdots & & \vdots \\ \sigma^{k,1}X_{k\_}^\top X_{1\_} & \cdots & \sigma^{k,k}X_{k\_}^\top X_{k\_} \end{bmatrix}_{k\times k}.$$

Hence, the SUR estimator of A can be rewritten as

$$\widehat{A}_{SUR} = [X_\_^\top(\Sigma_u \otimes I_n)^{-1}X_\_]^{-1}[X_\_^\top(\Sigma_u \otimes I_n)^{-1}X] = \frac{1}{|B|}B^* \begin{bmatrix} \sum_{j=1}^{k}\sigma^{1,j}X_{1\_}^\top X_j \\ \vdots \\ \sum_{j=1}^{k}\sigma^{k,j}X_{k\_}^\top X_j \end{bmatrix}.$$

where B^* denotes adjoint matrix of B. Further, the i-th term of $\widehat{A}_{SUR}$ is

$$\widehat{\rho}_{i,SUR} = \frac{1}{|B|} \times \left\| \begin{bmatrix} B_{1,1} & \cdots & B_{1,i-1} & \sum_{j=1}^{k} \sigma^{1,j} X_{1\_}^{\top} X_j & B_{1,i+1} & \cdots & B_{1,k} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ B_{k,1} & \cdots & B_{k,i-1} & \sum_{j=1}^{k} \sigma^{k,j} X_{k\_}^{\top} X_j & B_{k,i+1} & \cdots & B_{k,k} \end{bmatrix} \right\|$$
$$= \rho_i + \frac{1}{|B|} \times \sum_{j=1}^{k} \left\| \begin{bmatrix} B_{1,1} & \cdots & \sigma^{1,j} \sum_{t=1}^{n} x_{1,t-1} u_{j,t} & \cdots & B_{1,k} \\ \vdots & & \vdots & & \vdots \\ B_{k,1} & \cdots & \sigma^{kj} \sum_{t=1}^{n} x_{k,t-1} u_{j,t} & \cdots & B_{k,k} \end{bmatrix} \right\|,$$

since

$$\begin{bmatrix} \sum_{j=1}^{k} \sigma^{1,j} X_{1\_}^{\top} X_j \\ \vdots \\ \sum_{j=1}^{k} \sigma^{k,j} X_{k\_}^{\top} X_j \end{bmatrix} = \sum_{j=1}^{k} \rho_j \begin{bmatrix} B_{1,j} \\ \vdots \\ B_{k,j} \end{bmatrix} + \begin{bmatrix} \sum_{j=1}^{k} \sigma^{1,j} (\sum_{t=1}^{n} x_{1,t-1} u_{j,t}) \\ \vdots \\ \sum_{j=1}^{k} \sigma^{k,j} (\sum_{t=1}^{n} x_{k,t-1} u_{j,t}) \end{bmatrix}.$$

Therefore, we have

$$\widehat{\rho}_{i,SUR} - \rho_i = \frac{1}{|B|} \times \sum_{j=1}^{k} \left\| \begin{bmatrix} B_{1,1} & \cdots & \sigma^{1,j} \sum_{t=1}^{n} x_{1,t-1} u_{j,t} & \cdots & B_{1,k} \\ \vdots & & \vdots & & \vdots \\ B_{k,1} & \cdots & \sigma^{k,j} \sum_{t=1}^{n} x_{k,t-1} u_{j,t} & \cdots & B_{k,k} \end{bmatrix} \right\|. \tag{A9}$$

Using the result in Lemma 2, we obtain the following asymptotics:

$$\frac{1}{\rho_i^n \rho_j^n} B_{i,j} = \frac{1}{\rho_i^n \rho_j^n} \sigma^{i,j} \sum_{t=1}^{n} x_{i,t-1} x_{j,t-1} \Rightarrow \frac{\sqrt{\sigma_{i,i}\sigma_{j,j}}}{\rho_i \rho_j - 1} \sigma^{i,j} Q_i(\rho_i) Q_j(\rho_j). \tag{A10}$$

Therefore, $|B|$, the denominator of equation (A9), has the following asymptotics:

$$\left(\frac{1}{\prod_{j=1}^{k} \rho_j^n} \right)^2 |B| \Rightarrow \left(\prod_{j=1}^{k} \sigma_{j,j} Q_j(\rho_j)^2 \right) \xi_0(\boldsymbol{\rho}), \tag{A11}$$

where

$$\xi_0(\boldsymbol{\rho}) := \left\| \begin{bmatrix} \frac{\sigma^{1,1}}{\rho_1^2 - 1} & \cdots & \frac{\sigma^{1,k}}{\rho_1 \rho_k - 1} \\ \vdots & & \vdots \\ \frac{\sigma^{k,1}}{\rho_k \rho_1 - 1} & \cdots & \frac{\sigma^{k,k}}{\rho_k^2 - 1} \end{bmatrix} \right\|. \tag{A12}$$

Further, using the limiting result from Lemma 2 such that $\frac{1}{\rho_i^n} \sigma^{i,j} \sum_{t=1}^{n} x_{i,t-1} u_{j,t-1} \Rightarrow \sigma^{i,j} \sqrt{\sigma_{i,i}\sigma_{j,j}} Q_i(\rho_i) \widetilde{Q}_j(\rho_i)$, the jth summand in numerator of equation (A9) has the following asymptotics:

$$\left(\frac{1}{\prod_{\ell\neq i}\rho_\ell^n}\right)^2 \frac{1}{\rho_i^n}\left\|\begin{bmatrix} B_{1,1} & \cdots & \sigma^{1,p}\sum_{t=1}^n x_{1,t-1}u_{j,t} & \cdots & B_{1,k} \\ \vdots & & \vdots & & \vdots \\ B_{k,1} & \cdots & \sigma^{k,p}\sum_{t=1}^n x_{k,t-1}u_{j,t} & \cdots & B_{k,k}\end{bmatrix}\right\|$$
$$\Rightarrow \left(\prod_{\ell=1}^k \sigma_{\ell,\ell}x_\ell(\rho_\ell)^2\right)\sqrt{\frac{\sigma_{j,j}}{\sigma_{i,i}}}\frac{1}{Q_i(\rho_i)}\left\|\begin{bmatrix} \frac{\sigma^{1,1}}{\rho_1^2-1} & \cdots & \sigma^{1,j}\widetilde{x}_j(\rho_1) & \cdots & \frac{\sigma^{1,k}}{\rho_1\rho_k-1} \\ \vdots & & \vdots & & \vdots \\ \frac{\sigma^{k,1}}{\rho_k\rho_1-1} & \cdots & \sigma^{k,j}\widetilde{x}_j(\rho_k) & \cdots & \frac{\sigma^{k,k}}{\rho_k^2-1}\end{bmatrix}\right\|. \tag{A13}$$

Combine the result from equations (A11) and (A13), we obtain $\rho_i^n(\widehat{\rho}_{i,SUR}-\rho_i) \Rightarrow \frac{\xi_i(\boldsymbol{\rho})}{\xi_0(\boldsymbol{\rho})}$, where

$$\xi_i(\boldsymbol{\rho}) := \sum_{j=1}^k \sqrt{\frac{\sigma_{j,j}}{\sigma_{i,i}}}\frac{1}{Q_i(\rho_i)}\left\|\begin{bmatrix} \frac{\sigma^{1,1}}{\rho_1^2-1} & \cdots & \sigma^{1,j}\widetilde{Q}_j(\rho_1) & \cdots & \frac{\sigma^{1,k}}{\rho_1\rho_k-1} \\ \vdots & & \vdots & & \vdots \\ \frac{\sigma^{k,1}}{\rho_k\rho_1-1} & \cdots & \sigma^{k,j}\widetilde{Q}_j(\rho_k) & \cdots & \frac{\sigma^{k,k}}{\rho_k^2-1}\end{bmatrix}\right\|. \tag{A14}$$

This completes the proof of Theorem 1. □

Proof of Theorem 2

Proof . The proof is analogous to the proof of Theorem 1. The detailed proof is provided in the Data S1. □

Final Manuscript Received: May 2022

References

Amihud, Y. and Hurvich, C. M. (2004). 'Predictive regressions: a reduced-bias estimation method', *Journal of Financial and Quantitative Analysis*, Vol. 39, pp. 813–841.

Anselin, L. (1988). 'A test for spatial autocorrelation in seemingly unrelated regressions', *Economics Letters*, Vol. 28, pp. 335–341.

Antonakakis, N. and Kizys, R. (2015). 'Dynamic spillovers between commodity and currency markets', *International Review of Financial Analysis*, Vol. 41, pp. 303–319.

Balli, F., Naeem, M. A., Shahzad, S. J. H. and de Bruin, A. (2019). 'Spillover network of commodity uncertainties', *Energy Economics*, Vol. 81, pp. 914–927.

Baltagi, B. H. and Pirotte, A. (2011). 'Seemingly unrelated regressions with spatial error components', *Empirical Economics*, Vol. 40, pp. 5–49.

Basse, T., Klein, T., Vigne, S. A. and Wegener, C. (2021). 'Us stock prices and the dot. com-bubble: can dividend policy rescue the efficient market hypothesis?', *Journal of Corporate Finance*, Vol. 67, 101892.

Bouri, E., Shahzad, S. J. H. and Roubaud, D. (2019). 'Co-explosivity in the cryptocurrency market', *Finance Research Letters*, Vol. 29, pp. 178–183.

Caspi, I., Katzke, N. and Gupta, R. (2018). 'Date stamping historical periods of oil price explosivity: 1876–2014', *Energy Economics*, Vol. 70, pp. 582–587.
Chavas, J.-P. and Li, J. (2020). 'A quantile autoregression analysis of price volatility in agricultural markets', *Agricultural Economics*, Vol. 51, pp. 273–289.
Chen, Y., Phillips, P. C. B. and Yu, J. (2017). 'Inference in continuous systems with mildly explosive regressors', *Journal of Econometrics*, Vol. 201, pp. 400–416.
Cheung, A., Roca, E. and Su, J.-J. (2015). 'Crypto-currency bubbles: an application of the phillips–shi–yu (2013) methodology on mt. gox bitcoin prices', *Applied Economics*, Vol. 47, pp. 2348–2358.
Corbet, S., Lucey, B. and Yarovaya, L. (2018). 'Datestamping the bitcoin and ethereum bubbles', *Finance Research Letters*, Vol. 26, pp. 81–88.
Cubadda, G., Hecq, A. and Telg, S. (2019). 'Detecting co-movements in non-causal time series', *Oxford Bulletin of Economics and Statistics*, Vol. 81, pp. 697–715.
Demirer, M., Diebold, F. X., Liu, L. and Yilmaz, K. (2018). 'Estimating global bank network connectedness', *Journal of Applied Econometrics*, Vol. 33, pp. 1–15.
Diba, B. T. and Grossman, H. I. (1988). 'Explosive rational bubbles in stock prices?', *The American Economic Review*, Vol. 78, pp. 520–530.
Diebold, F. X. and Yilmaz, K. (2009). 'Measuring financial asset return and volatility spillovers, with application to global equity markets', *The Economic Journal*, Vol. 119, pp. 158–171.
Diebold, F. X. and Yilmaz, K. (2012). 'Better to give than to receive: predictive directional measurement of volatility spillovers', *International Journal of Forecasting*, Vol. 28, pp. 57–66.
Diebold, F. X. and Yılmaz, K. (2014). 'On the network topology of variance decompositions: Measuring the connectedness of financial firms', *Journal of Econometrics*, Vol. 182, pp. 119–134.
Engsted, T. and Nielsen, B. (2012). 'Testing for rational bubbles in a coexplosive vector autoregression', *The Econometrics Journal*, Vol. 15, pp. 226–254.
Etienne, X. L., Irwin, S. H. and Garcia, P. (2014). 'Bubbles in food commodity markets: four decades of evidence', *Journal of International Money and Finance*, Vol. 42, pp. 129–155.
Fantazzini, D. (2016). 'The oil price crash in 2014/15: was there a (negative) financial bubble?', *Energy Policy*, Vol. 96, pp. 383–396.
Figuerola-Ferretti, I., Gilbert, C. L. and McCrorie, J. R. (2015). 'Testing for mild explosivity and bubbles in lme non-ferrous metals prices', *Journal of Time Series Analysis*, Vol. 36, pp. 763–782.
Gharib, C., Mefteh-Wali, S. and Jabeur, S. B. (2021). 'The bubble contagion effect of covid-19 outbreak: evidence from crude oil and gold markets', *Finance research letters*, Vol. 38, 101703.
Gourieroux, C. and Jasiak, J. (2017). 'Noncausal vector autoregressive process: representation, identification and semi-parametric estimation', *Journal of Econometrics*, Vol. 200, pp. 118–134.
Gouriéroux, C. and Zakoı¨an, J.-M. (2017). 'Local explosion modelling by non-causal process', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 79, pp. 737–756.
Greenaway-McGrevy, R. and Phillips, P. C. B. (2016). 'Hot property in New Zealand: empirical evidence of housing bubbles in the metropolitan centres', *New Zealand Economic Papers*, Vol. 50, pp. 88–113.
Henderson, D. J., Kumbhakar, S. C., Li, Q. and Parmeter, C. F. (2015). 'Smooth coefficient estimation of a seemingly unrelated regression', *Journal of Econometrics*, Vol. 189, pp. 148–162.
Horváth, L., Li, H. and Liu, Z. (2021). 'How to identify the different phases of stock market bubbles statistically?', *Finance Research Letters*, Vol. 46, 102366.
Judge, G. G. (1982). *Introduction to the Theory and Practice of Econometrics*, New York: Wiley.
Kräussl, R., Lehnert, T. and Martelin, N. (2016). 'Is there a bubble in the art market?', *Journal of Empirical Finance*, Vol. 35, pp. 99–109.
Lai, T. and Wei, C. (1983). 'Asymptotic properties of general autoregressive models and strong consistency of least-squares estimates of their parameters', *Journal of Multivariate Analysis*, Vol. 13, pp. 1–23.
Li, J. and Chavas, J.-P. (2023). 'A dynamic analysis of the distribution of commodity futures and spot prices', *American Journal of Agricultural Economics*, Vol. 105, pp. 122–143.
Li, J., Chavas, J.-P., Etienne, X. L. and Li, C. (2017a). 'Commodity price bubbles and macroeconomics: evidence from the chinese agricultural markets', *Agricultural Economics*, Vol. 48, pp. 755–768.
Li, J., Li, C. and Chavas, J.-P. (2017b). 'Food price bubbles and government intervention: is China different?', *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, Vol. 65, pp. 135–157.

Li, Y., Wang, S. and Zhao, Q. (2021). 'When does the stock market recover from a crisis?', *Finance Research Letters*, Vol. 39, 101642.

Luo, J. and Ji, Q. (2018). 'High-frequency volatility connectedness between the us crude oil market and China's agricultural commodity markets', *Energy Economics*, Vol. 76, pp. 424–438.

Ma, R. R. and Xiong, T. (2021). 'Price explosiveness in nonferrous metal futures markets', *Economic Modelling*, Vol. 94, pp. 75–90.

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

Mark, N. C., Ogaki, M. and Sul, D. (2005). 'Dynamic seemingly unrelated cointegrating regressions', *The Review of Economic Studies*, Vol. 72, pp. 797–820.

Moon, H. R. and Perron, B. (2006). *Seemingly Unrelated Regressions*, New York: Palgrave.

Nielsen, B. (2009). *Singular Vector Autoregressions with Deterministic Terms: Strong Consistency and Lag Order Determination*, University of Oxford.

Nielsen, B. (2010). 'Analysis of coexplosive processes', *Econometric Theory*, Vol. 26, pp. 882–915.

Pan, W.-F. (2018). 'Sentiment and asset price bubble in the precious metals markets', *Finance Research Letters*, Vol. 26, pp. 106–111.

Phillips, P. C. B. and Lee, J. H. (2015). 'Limit theory for vars with mixed roots near unity', *Econometric Reviews*, Vol. 34, pp. 1035–1056.

Phillips, P. C. B. and Lee, J. H. (2016). 'Robust econometric inference with mixed integrated and mildly explosive regressors', *Journal of Econometrics*, Vol. 192, pp. 433–450.

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

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

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

Phillips, P. C. B. and Yu, J. (2013). 'Bubble or roller coaster in world stock markets', *The Business Times*.

Phillips, P. C. B., Wu, Y. and Yu, J. (2011). 'Explosive behavior in the 1990s nasdaq: when did exuberance escalate asset values?', *International economic review*, Vol. 52, pp. 201–226.

Phillips, P. C. B., Shi, S. and Yu, J. (2014). 'Specification sensitivity in right-tailed unit root testing for explosive behaviour', *Oxford Bulletin of Economics and Statistics*, Vol. 76, pp. 315–333.

Phillips, P. C. B., Shi, S. and Yu, J. (2015a). 'Testing for multiple bubbles: historical episodes of exuberance and collapse in the S&P 500', *International Economic Review*, Vol. 56, pp. 1043–1078.

Phillips, P. C. B., Shi, S. and Yu, J. (2015b). 'Testing for multiple bubbles: limit theory of real-time detectors', *International Economic Review*, Vol. 56, pp. 1079–1134.

Shi, S. and Phillips, P. C. B. (2021). 'Diagnosing housing fever with an econometric thermometer', *Journal of Economic Surveys*, Vol. 37, pp. 159–186. https://doi.org/10.1111/joes.12430.

Shi, S., Valadkhani, A., Smyth, R. and Vahid, F. (2016). 'Dating the timeline of house price bubbles in Australian capital cities', *Economic Record*, Vol. 92, pp. 590–605.

Shi, S., Hurn, S. and Phillips, P. C. B. (2020). 'Causal change detection in possibly integrated systems: revisiting the money–income relationship', *Journal of Financial Econometrics*, Vol. 18, pp. 158–180.

Smith, M. and Kohn, R. (2000). 'Nonparametric seemingly unrelated regression', *Journal of Econometrics*, Vol. 98, pp. 257–281.

Tang, K. and Xiong, W. (2012). 'Index investment and the financialization of commodities', *Financial Analysts Journal*, Vol. 68, pp. 54–74.

Wang, S. L. and McPhail, L. (2014). 'Impacts of energy shocks on us agricultural productivity growth and commodity prices—a structural var analysis', *Energy Economics*, Vol. 46, pp. 435–444.

Xiao, B., Yu, H., Fang, L. and Ding, S. (2020). 'Estimating the connectedness of commodity futures using a network approach', *Journal of Futures Markets*, Vol. 40, pp. 598–616.

Yang, J., Balyeat, R. B. and Leatham, D. J. (2005). 'Futures trading activity and commodity cash price volatility', *Journal of Business Finance & Accounting*, Vol. 32, pp. 297–323.

Zellner, A. (1962). 'An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias', *Journal of the American Statistical Association*, Vol. 57, pp. 348–368.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1. Supporting information