INCONSISTENT VAR REGRESSION WITH COMMON EXPLOSIVE ROOTS

BY

Peter C.B. Phillips and Tassos Magdalinos

COWLES FOUNDATION PAPER NO. 1383

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

2013

http://cowles.econ.yale.edu/

Econometric Theory, **29**, 2013, 808–837.
doi:10.1017/S0266466612000709

INCONSISTENT VAR REGRESSION WITH COMMON EXPLOSIVE ROOTS

Peter C.B. Phillips
Yale University, University of Auckland, University of Southampton, and Singapore Management University

Tassos Magdalinos
University of Southampton

Nielsen (Working paper, University of Oxford, 2009) shows that vector autoregression is inconsistent when there are common explosive roots with geometric multiplicity greater than unity. This paper discusses that result, provides a coexplosive system extension and an illustrative example that helps to explain the finding, gives a consistent instrumental variable procedure, and reports some simulations. Some exact limit distribution theory is derived and a useful new reverse martingale central limit theorem is proved.

1. BACKGROUND AND MOTIVATION

Financial exuberance and market bubbles have led to a new interest among empirical researchers in autoregressive time series with explosive roots. Recent research has focused on the detection of bubble activity by means of right-sided recursive unit root tests (Phillips, Wu, and Yu, 2011) and date stamping the origination and termination of this type of phenomenon in the data (Phillips and Yu, 2011). These methods have attracted the attention of empirical researchers interested in bubbles. Explosive roots are also known to arise in certain present value models that link decision variables and explanatory variables in the absence of Granger causality (Fanelli, 2007) or when nonstationary data are bootstrapped (Swensen, 2006; Cavaliere, Rahbek, and Taylor, 2010, 2012).

Theoretical econometric research has also attracted interest and increased relevance for practical work by developing new concepts and associated limit theory for mildly explosive processes (Phillips and Magdalinos, 2007) and by extending the notion of comovement to include coexplosive processes (Phillips and Magdalinos, 2008 (hereafter PM); Magdalinos and Phillips, 2009). These

This paper was inspired by reading Nielsen (2009). We thank two referees and the co-editor for helpful comments on earlier versions. Magdalinos thanks the Cowles Foundation, Yale University, for hospitality in Fall 2009 when the paper was first drafted and acknowledges research support from the Economic and Social Research Council under grant no. RES-061-25-0517. Phillips acknowledges partial research support from the National Science Foundation under grant no. SES 09-56687. Address correspondence to Peter Phillips, Cowles Foundation for Research in Economics, Yale University, Box 208281, New Haven, CT 06520-8281, USA; e-mail: peter.phillips@yale.edu.

processes are relevant in practical work with data where contagion effects are suspected. Coexplosive processes arise when there are common explosive roots and these lead to an asymptotic singularity in the signal matrix, which produces complications in limit theory.

In related work, Nielsen (2009) (hereafter NN) considers a vector autoregression (VAR) with common explosive roots and shows that least squares regression (and Gaussian maximum likelihood) is inconsistent. This result is intriguing because the model is correctly specified in terms of its lag and error structure and falls within a framework where ordinary least squares (OLS) is well known to be generally consistent with good asymptotic properties. The model is unremarkable except for the occurrence of common explosive roots with geometric multiplicity exceeding unity. The simplest case is a VAR(1) with scalar coefficient matrix ρI and $\rho > 1$. The common explosive roots produce coexplosive behavior and lead to an asymptotic singularity in the signal matrix, analogous to that studied in PM in structural models. The singularity has fatal consequences in the VAR case. Importantly, Nielsen's result provides a new context where (unrestricted) maximum likelihood is inconsistent.

The present work explores the result by considering an example that helps to explain the inconsistency in terms of the endogeneity that is induced by coexplosive behavior. In an explosive autoregression the variables behave like exponential trends (with random coefficients) that are informative about the future trajectory. Coexplosive behavior in a VAR produces common exponential trends that are close to the future in the sense that certain linear combinations of the variables depend explicitly on future residuals, thereby producing an endogeneity in the regressors. While least squares regression is inconsistent, simple instrumental variable (IV) estimation with contemporaneous or future values of the variables as instruments is shown to be consistent and to provide a basis for econometric testing. The OLS regression inconsistency phenomenon can also occur in triangular systems, such as those studied in PM, and a similar IV remedy may be implemented in that context.

The inconsistency of OLS regression is to a random limit involving a matrix quotient of random variables. The exact marginal limit distributions are obtained for the case where the VAR innovations are Gaussian. The limit random variables are bounded and the distributions have asymptotes at the boundaries. Simulations reveal a corresponding bimodality in the finite sample distributions.

2. MAIN RESULTS

2.1. A Prototypical Model

For simplicity of exposition of the main ideas, we consider the bivariate VAR(1) model

$$x_t = R x_{t-1} + u_t, \qquad t = 1, \ldots, n, \tag{1}$$

with $R = \rho I_2$, $\rho > 1$, $x_0 = 0$, and martingale difference innovations u_t satisfying Assumption 1 below.

The bivariate system in (1) can be written in component form as

$$\begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}, \tag{2}$$

with the same explosive autoregressive coefficient $\rho > 1$, so the algebraic and geometric multiplicity of this system is two. The results below extend in a straightforward way to more complex multivariate VAR systems with common explosive roots.

As pointed out by Anderson (1959) and discussed in PM and NN, equality of the autoregressive coefficients in (2) induces coexplosive behavior in the series x_{1t} and x_{2t} that results in a singular limit for the standardized sample moment matrix:

$$\rho^{-2n} X'X = \rho^{-2n} \sum_{t=1}^{n} x_{t-1} x_{t-1}' \to_{a.s.} \frac{1}{\rho^2 - 1} X(\rho) X(\rho)', \tag{3}$$

where

$$X(\rho) = \sum_{j=1}^{\infty} \rho^{-j} u_j = \lim_{n\to\infty} \frac{x_n}{\rho^n} \quad \text{a.s.} \tag{4}$$

When u_t is a zero mean uncorrelated sequence with bounded second moments, the infinite series in (4) can easily be shown to converge almost surely and in L_2. The more restrictive assumption of u_t being a martingale difference sequence satisfying a local Marcinkiewicz-Zygmund condition ensures that $X(\rho) \neq 0$ $a.s.$; see Lai and Wei (1983). Our asymptotic development also requires a constant conditional variance assumption and a uniform integrability type of condition on the sequence $\left\{ \|u_t\|^2 : 1 \le t \le n \right\}$. The dependence and moment structure of the innovation sequence is explicitly stated below. We denote by $\|x\| = \left(x'x\right)^{1/2}$ the Euclidian norm and by $\|x\|_{L_p} = \left[\mathbb{E}\left(\|x\|^p\right)\right]^{1/p}$ the L_p norm of a random vector x. When x is a random matrix we define $\|x\| = \|\text{vec}(x)\|$.

Assumption 1. Let u_t be a martingale difference sequence with respect to $\mathcal{F}_t = \sigma(u_t, u_{t-1}, \ldots)$ satisfying

$$\mathbb{E}_{\mathcal{F}_{t-1}}\left(u_t u_t'\right) = \Sigma_u \quad \text{and} \quad \mathbb{E}_{\mathcal{F}_{t-1}} \|u_t\| \ge \delta \quad \text{a.s. for all } t \tag{5}$$

for some $\delta > 0$ and positive definite matrix Σ_u and

$$\max_{1\le t\le n} \mathbb{E}\left(\|u_t\|^2 \mathbf{1}\{\|u_t\| > \lambda_n\}\right) \to 0 \quad \text{as } n \to \infty \tag{6}$$

for any sequence $(\lambda_n)_{n\in\mathbb{N}}$ such that $\lambda_n \to \infty$

To treat the limiting singularity (3) induced by coexplosive behavior we perform a coordinate rotation as developed in PM. Here it is convenient to use the (sample size dependent) orthogonal transformation

$$z_t = H_n' x_t, \tag{7}$$

where

$$H_n = \frac{1}{\|x_n\|}\begin{bmatrix} x_{1n} & -x_{2n} \\ x_{2n} & x_{1n} \end{bmatrix} = \frac{1}{\|x_n\|}\left[x_n, \mathcal{R}_{\frac{\pi}{2}} x_n\right], \tag{8}$$

in which the orthogonal matrix

$$\mathcal{R}_{\frac{\pi}{2}} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

rotates vectors in the plane by $\pi/2$ radians in the positive direction. In view of (7), the transformed variate z_t forms an array, but for notational simplicity the additional subscript is not employed. The large sample behavior of the random rotation matrix in (8) is characterized by the following lemma, proved in Section 5.

LEMMA 1. *Under Assumption 1,* $\rho^{-n} x_n \to_{a.s.} X(\rho)$, $X(\rho) \neq 0$ *a.s. and*

$$H_n \to_{a.s.} \frac{1}{\|X(\rho)\|}\left[X(\rho), \mathcal{R}_{\frac{\pi}{2}} X(\rho)\right] \quad \text{as } n \to \infty. \tag{9}$$

The transformed regressor variate in (7) may be analyzed by combining the identity

$$x_{t-1} = \rho^{-(n-t+1)} x_n - \sum_{j=t}^{n} \rho^{-(j-t+1)} u_j \tag{10}$$

and the orthogonality condition $(\mathcal{R}_{\frac{\pi}{2}} x_n)' x_n = 0$ as

$$\begin{aligned} z_{t-1} = H_n' x_{t-1} &= \frac{1}{\|x_n\|}\begin{bmatrix} x_n' x_{t-1} \\ \left(\mathcal{R}_{\frac{\pi}{2}} x_n\right)' x_{t-1} \end{bmatrix} \\ &= \frac{1}{\|x_n\|}\begin{bmatrix} x_n' x_{t-1} \\ -\left(\mathcal{R}_{\frac{\pi}{2}} x_n\right)' \zeta_{n,t} \end{bmatrix} =: \begin{bmatrix} z_{1t-1} \\ z_{2t-1} \end{bmatrix}, \end{aligned} \tag{11}$$

which is conformably partitioned with x_t, where

$$\zeta_{n,t} = \sum_{j=t}^{n} \rho^{-(j-t+1)} u_j \tag{12}$$

is a (forward filtered) linear process with l_1 summable coefficients.

The transformed variate z_{t-1} has an explosive component (z_{1t-1}) and a nonexplosive component (z_{2t-1}). However, unlike similar transformations in models with trend-induced degeneracies (such as models with some deterministic trends and some stochastic trends; see Park and Phillips, 1988, 1989), the nonexplosive component z_{2t-1} involves linear combinations that are data dependent and random, even asymptotically, as is apparent from the limit of $\mathcal{R}_{\frac{\pi}{2}} x_n / \|x_n\|$ in (9). It follows from the form of $X(\rho) = \sum_{j=1}^{\infty} \rho^{-j} u_j$ that the random linear combination present in z_{2t-1} introduces an endogeneity into the regressor that leads to the inconsistency of least squares. In particular, the component $\rho^{-t} u_t$ of $X(\rho)$ is correlated with the regression error u_t, as is the component $\zeta_{n,t}$ of the transformed regressor z_{2t-1}.

Intriguingly, under a martingale difference assumption on the innovation sequence u_t, the regressor x_{t-1} in the original system (1) satisfies $\mathbb{E}(u_t | x_{t-1}) = 0$ $a.s.$, thereby fulfilling one of the usual conditions for consistent least squares estimation. However, the limiting singularity in the sample moment matrix involves the data dependent vector $X(\rho)$ and induces an endogeneity in the (transformed) system that takes into account the coexplosive behavior present in x_t. To see the reason for the endogeneity more clearly, note that $\rho^{-(t-1)} x_{t-1} = X(\rho) - \sum_{k=t}^{\infty} \rho^{-k} u_k$, so the difference $\rho^{-(t-1)} x_{t-1} - X(\rho)$ contains information about future disturbances and, in particular, is correlated with u_t. When the system is unidimensional this almost sure (a.s.) limit behavior is not enough to induce endogeneity. But in a multidimensional system with common explosive roots (and geometric multiplicity greater than unity) information is sourced from more than one component of x_{t-1}, and the resulting singularity in the signal matrix reveals information about $X(\rho)$ and the null space of the (asymptotic) signal matrix. It is this information that leads to the residual process $\zeta_{n,t}$ that is correlated with u_t. When geometric multiplicity is unity, there are cross effects in the coefficient matrix R (which is no longer diagonal) that complicate the signal matrix and eliminate the endogeneity in the regressor.

Given the form of z_{2t-1} and (12), it is apparent that dynamic timing also plays a role in the endogeneity that is manifest in $\mathbb{E}\left(u_t | \zeta_{n,t}\right) \neq 0$, since $\zeta_{n,t}$ itself depends on u_t. As we shall see, this type of endogeneity can arise even in the triangular (coexplosive) system considered in PM. Like most forms of endogeneity, it can be dealt with by suitable instrumentation that adjusts the dynamic timing, as discussed in Section 2.3.

2.2. Least Squares Limit Theory

Coexplosive behavior induces a singularity of the form (3) in the limiting sample moment matrix. The degeneracy occurs along the direction vector $[-X_2(\rho), X_1(\rho)]$. The inverse sample moment matrix sustains a similar singularity, which can be conveniently expressed in terms of the transformed system. More generally, Lemma 2 below describes the asymptotic behavior of the inverse of sample moment matrices involving the transformed variates z_{t-1} and z_{t+k} for some fixed value of $k \geq 0$. The lemma also characterizes the condition number limit behavior of the least squares regression matrix $X'X$. The lemma is useful in developing a limit theory for both least squares and instrumental variable estimates.

LEMMA 2. *Under Assumption 1, the following hold as $n \to \infty$ for any fixed k:*

(i) $n^{-1}\sum_{t=1}^{n} u_t u_t' \to_{L_1} \Sigma_u$,

(ii) $\max_{1\le j\le n}\left\| n^{-1}\sum_{t=1}^{n} u_t u_{t+j}'\right\|_{L_1} \to 0$,

(iii) $n^{-1}\sum_{t=1}^{n-k} u_t \zeta_{n,t+k}' \to_{L_1} 0,\ k \geq 1$,

(iv) $n^{-1}\sum_{t=1}^{n-k} \zeta_{n,t}\zeta_{n,t+k}' \to_{L_1} \rho^{-k}\left(\rho^2-1\right)^{-1}\Sigma_u,\ k \geq 0$,

(v) $\left(\frac{1}{n}\sum_{t=1}^{n-k} z_{t-1}z_{t+k}'\right)^{-1} \to_p diag\left(0, \left\{\frac{\rho^{-k-1}}{\rho^2-1}\frac{X(\rho)'\mathcal{R}_{\frac{\pi}{2}}'\Sigma_u\mathcal{R}_{\frac{\pi}{2}}X(\rho)}{\|X(\rho)\|^2}\right\}^{-1}\right)$, $k \geq -1$,

(vi) *Letting $\lambda_{\max}(X'X)$ and $\lambda_{\min}(X'X)$ denote the largest and smallest eigenvalues of the matrix $X'X = \sum_{t=1}^{n} x_{t-1}x_{t-1}'$,*

$$\frac{\log\left\{\lambda_{\max}\left(X'X\right)\right\}}{\lambda_{\min}\left(X'X\right)} \to_p 2\left(\log\rho\right)\left(\rho^2-1\right)\frac{\|X(\rho)\|^2}{X(\rho)'\mathcal{R}_{\frac{\pi}{2}}'\Sigma_u\mathcal{R}_{\frac{\pi}{2}}X(\rho)} > 0 \quad a.s. \tag{13}$$

We proceed to establish a central limit theorem (CLT) for the sample covariances of the form $\sum_{t=1}^{n}\zeta_{n,t+k}u_t'$. In view of the forward filtered nature of $\zeta_{n,t}$ in (12), sample covariances of this process and u_t have a type of reverse martingale structure, which has to be transformed to a (forward) martingale in order for standard martingale central limit theory to apply. This transformation is carried out in (15) below.

LEMMA 3. *Consider the $\mathcal{F}_j$-martingale difference array*

$$\eta_{n,j} = n^{-1/2}u_j \otimes \left(\sum_{i=k+1}^{j-1}\rho^{-i}u_{j-i}\right). \tag{14}$$

Under Assumption 1 and $\mathbb{E}(\|u_j\|^2\|u_{j+i}\|^2) < \infty$ for all $i, j \geq 1$ we have

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n-k-1}\left(\zeta_{n,t+k+1}\otimes u_t\right) = \rho^k\sum_{j=k+2}^{n}\eta_{n,j} \Rightarrow N\left(0, \frac{1}{\rho^2-1}\Sigma_u\otimes\Sigma_u\right) \tag{15}$$

as $n \to \infty$ for any fixed $k \geq 0$.

Remark. When u_t is an independent and identically distributed (i.i.d.) sequence, an alternative approach to Lemma 3 is to work directly with the reverse martingale difference $\xi_{n,t} = n^{-1/2}\zeta_{n,t+k+1}\otimes u_t$. For each fixed $k \geq 0$,

$S_{n,\tau} = \sum_{t=\tau}^{n} \xi_{n,t}$ is a reverse martingale array with respect to the reverse filtration $\mathcal{F}^{\tau} = \sigma(u_{\tau}, u_{\tau+1}, ...)$ that can be further reversed into a martingale array $(M_{n,\tau}, \mathcal{G}_{n,\tau}, 1 \leq \tau \leq n)$ by letting $M_{n,\tau} = S_{n,n-\tau}$ and $\mathcal{G}_{n,\tau} = \mathcal{F}^{n-\tau}$. The identities

$$\sum_{\tau=0}^{n-1} \Delta M_{n,\tau} = \sum_{\tau=0}^{n-1} \xi_{n,n-\tau} = \sum_{t=1}^{n} \xi_{n,t}$$

then imply that the limit distribution of $\sum_{t=1}^{n} \xi_{n,t}$ can be derived by a standard martingale CLT (MGCLT) on $\sum_{\tau=0}^{n-1} \Delta M_{n,\tau}$ (e.g., Cor. 3.1 of Hall and Heyde, 1980). One application of this result is to the CLT stated in equation (26) of PM.[1]

Define $X' = [x_1, ..., x_n]$ and $X'_{-1} = [x_0, ..., x_{n-1}]$, and the least squares regression matrix $\hat{R}_n = X'X_{-1}(X'_{-1}X_{-1})^{-1}$. The following result characterizes the limit of $\hat{R}_n$.

THEOREM 1. *Under Assumption 1, the OLS estimator in (1) has the following limit as $n \to \infty$:*

$$\hat{R}_n - R \to_p \quad -\frac{\rho^2-1}{\rho} \frac{\Sigma_u \mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}'_{\frac{\pi}{2}}}{X(\rho)' \mathcal{R}'_{\frac{\pi}{2}} \Sigma_u \mathcal{R}_{\frac{\pi}{2}} X(\rho)}$$

$$= -\frac{\rho^2-1}{\rho} \frac{\Sigma_u \begin{bmatrix} X_2(\rho)^2 & -X_1(\rho) X_2(\rho) \\ -X_1(\rho) X_2(\rho) & X_1(\rho)^2 \end{bmatrix}}{\sigma_1^2 X_2(\rho)^2 - 2\sigma_{12} X_1(\rho) X_2(\rho) + \sigma_2^2 X_1(\rho)^2}. \tag{16}$$

Remark 1. The inconsistency of $\hat{R}_n$ is explained by the endogeneity of the regressors discussed earlier. Lai and Wei (1982) showed consistency of least squares in time series regression models with martingale difference errors under second-moment conditions on the errors, an excitation condition on the smallest eigenvalue of the regression matrix $X'X$ and a condition number requirement for which the ratio

$$\frac{\log\left\{\lambda_{\max}\left(X'X\right)\right\}}{\lambda_{\min}(X'X)} \to_{a.s.} 0 \quad \text{as } n \to \infty. \tag{17}$$

As demonstrated in Lemma 2(vi), the ratio in (17) converges in probability to an almost surely positive random variable, thereby invalidating the condition number requirement. Thus, the sufficient conditions for consistency given in Lai and Wei (1982) fail in the present case. Interestingly, the asymptotic bias of $\hat{R}_n$ can be written is terms of the probability limit of the eigenvalue ratio in (17). In particular, (13) and (16) imply that

$$\hat{R}_n - R \to_p -\left(\operatorname*{plim}_{n\to\infty} \frac{\log \lambda_{\max}\left(X'X\right)}{\lambda_{\min}(X'X)}\right) \frac{\Sigma_u}{2(\log\rho)\rho} \frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}'_{\frac{\pi}{2}}}{X(\rho)' X(\rho)}.$$

Remark 2. All elements of the regression matrix $\hat{R}_n$ converge to random variates that depend on $X(\rho) = (X_1(\rho), X_2(\rho))'$, the error covariance matrix Σ_u, and the common explosive coefficient ρ. The limit distribution (16) is singular and is of rank unity, corresponding to $X(\rho)$. Defining $\xi = \Sigma_u^{1/2} \mathcal{R}_{\frac{\pi}{2}} X(\rho)$ and $h = \xi(\xi'\xi)^{-1/2}$, the limit (16) may be written more simply as

$$-\frac{\rho^2 - 1}{\rho} \Sigma_u^{1/2} h h' \Sigma_u^{-1/2}$$

in terms of the vector h, which is distributed on the unit sphere.

Remark 3. Figures 1 and 2 show the results of simulations of the fitted regression coefficients in the least squares regression

$$x_{1t} = \hat{\rho} x_{1t-1} + \hat{\beta} x_{2t-1} + \hat{u}_{1t} \quad \textbf{(18)}$$

for various values of n $(= 200, 400, 800)$ against the limit distribution (16) (see also (20) and (21) in Theorem 2) when the data are generated according to (1) with $\rho = 1.04$ and

$$u_t \sim iid\ N(0, I_2).$$

The finite sample and limit distributions are bimodal in both cases, although the limit distributions have compact support and the densities asymptote at the boundaries. The limit distributions are obtained explicitly in Theorem 2 and discussed in

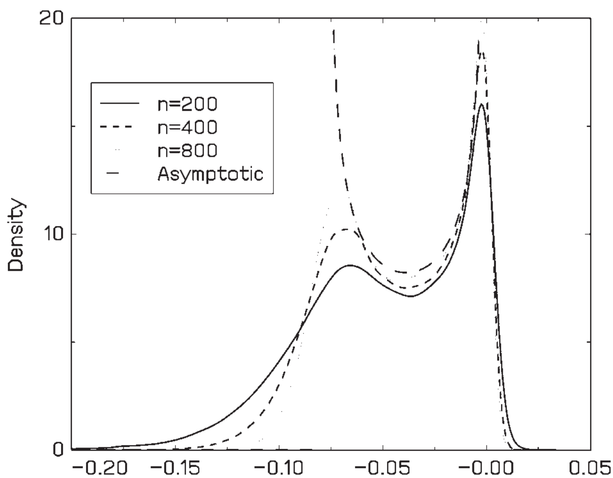

FIGURE 1. Finite sample densities of $\hat{\rho} - \rho$ from $R = 80,000$ replications in the fitted model $X_{1t} = \hat{\rho} X_{1t-1} + \hat{\beta} X_{2t-1} + \hat{u}_{1t}$ with $\rho = 1.04$ and $\sigma_{12} = 0$. The limit density has bounded support and is computed from the exact formula (20).

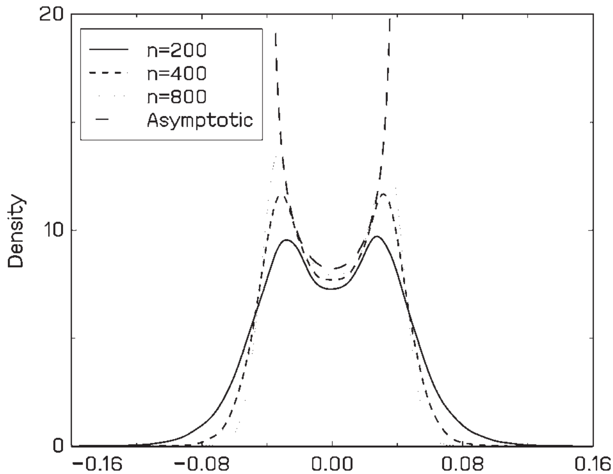

FIGURE 2. Finite sample densities of $\hat{\beta}$ from $R = 80,000$ replications in the fitted model $X_{1t} = \hat{\rho} X_{1t-1} + \hat{\beta} X_{2t-1} + \hat{u}_{1t}$ with $\rho = 1.04$ and $\sigma_{12} = 0$. The limit density has bounded support and is computed from the exact formula (21).

the remarks below. The distribution of $\hat{\beta}$ appears symmetric about the origin. The finite sample distribution of $\hat{\rho} - \rho$ is asymmetric, shows downward bias, and the convergence to the limit distribution appears to be a little slower. Similar findings were obtained for covariance structures with $\sigma_{12} = \mathbb{E}(u_{1t} u_{2t}) \neq 0$.

Remark 4. The limit random variables corresponding to $\hat{\rho}$ and $\hat{\beta}$ in (18) are given in (16). When $u_t \sim iid\ N(0, \sigma^2 I_2)$, these limits become

$$\begin{aligned} \hat{\rho} - \rho &\to_p -\frac{\rho^2 - 1}{\rho} \frac{X_2(\rho)^2}{X_2(\rho)^2 + X_1(\rho)^2}, \\ \hat{\beta} - \beta &\to_p \frac{\rho^2 - 1}{\rho} \frac{X_1(\rho) X_2(\rho)}{X_2(\rho)^2 + X_1(\rho)^2}, \end{aligned} \tag{19}$$

and since $X(\rho) =_d N(0, \sigma^2 (\rho^2 - 1)^{-1} I_2)$, we have

$$\hat{\rho} - \rho \to_p -\frac{\rho^2 - 1}{\rho} \frac{\xi_1^2}{\xi_1^2 + \xi_2^2}, \qquad \hat{\beta} - \beta \to_p \frac{\rho^2 - 1}{\rho} \frac{\xi_1 \xi_2}{\xi_1^2 + \xi_2^2},$$

where $\xi = (\xi_1, \xi_2)' =_d N(0, I_2)$ and $\beta = 0$. The exact marginal densities are given in the following result.

THEOREM 2. *If $u_t \sim iid\ N\left(0, \sigma^2 I_2\right)$ then the marginal densities of the limit distributions of $\hat{\rho} - \rho$ and $\hat{\beta} - \beta = \hat{\beta}$ are*

$$pdf_{\hat{\rho}}(y) = \frac{1}{\pi\left\{(-y)\left(a_\rho + y\right)\right\}^{1/2}}, \quad \textit{for } y \in \left(-a_\rho, 0\right), \tag{20}$$

$$pdf_{\hat{\beta}}(y) = \frac{2}{\pi\left\{a_\rho^2 - 4y^2\right\}^{1/2}}, \quad \textit{for } |y| < a_\rho/2, \tag{21}$$

where $a_\rho = \left(\rho^2 - 1\right)/\rho$.

Remark 5. The supports of the limit distributions (20) and (21) are finite and are determined by a_ρ. As $\rho \to 1$, $a_\rho \to 0$ and the supports shrink to the origin, which corresponds to the (well-known) consistent estimation of ρ and β when $\rho = 1$.

Remark 6. Figures 1 and 2 also show the limit densities $pdf_{\hat{\rho}}(y)$ and $pdf_{\hat{\beta}}(y)$ for $\rho = 1.04$ and $a_\rho = 0.078$. The density $pdf_{\hat{\rho}}(y)$ is that of a (translated) arc sine law. Each of the densities has bounded support and asymptotes at the limits of the domain of definition. These limit distributions are derived simply from the distribution of $h = \xi/(\xi'\xi)^{1/2}$, which is uniform on the unit circle, and have been studied in earlier work on structural estimation (Phillips, 1984). In a related form, the distribution (20) also appears in Luati and Paruolo (2002).

Remark 7. Importantly, the support of $pdf_{\hat{\rho}}(y)$ is negative, whereas the support of $pdf_{\hat{\beta}}(y)$ is symmetric about the origin. The implied downward bias in the limit distribution of $\hat{\rho}$ is explained by the presence of the coexplosive time series x_{2t-1} in regression (18). The regressor x_{2t-1} is asymptotically collinear to x_{1t-1} when $\rho > 1$. The explosive signal is then shared between these two regressors, reducing the impact of the own lagged dependent variable x_{1t-1} and, in this case, producing an inconsistency and resulting in the downward bias for $\hat{\rho}$ in the limit that is apparent in (19) and Figure 1. As discussed earlier, the inconsistency arises from the endogeneity induced by the comovement of the regressors and the random nature of the directional vector $X(\rho)$ of the comovement, which depends on the regression error u_t.

Remark 8. The bimodality in the finite sample distributions shown in Figures 1 and 2 is also a consequence of the common explosive signal that is shared between the regressors x_{1t-1} and x_{2t-1}. The distributions of the corresponding regression coefficients interact by way of the linear combination $\hat{\rho} + \hat{\beta} X_2(\rho)/X_1(\rho)$, which serves as the "effective" own lag coefficient in the regression (18). This interaction either attenuates or accentuates the downward bias in $\hat{\rho}$, producing a compensating bimodality in the two distributions and compensating asymptotes in the two limit distributions.

2.3. Consistent Estimation by Instrumental Variables

As indicated above, dynamic timing plays a role in the inconsistency of least squares regression because of the dependence of the forward filtered process $\zeta_{n,t}$

and hence the (transformed) regressor z_{2t-1} on the contemporaneous error u_t. This dependence can be avoided by the use of a suitable instrumental variable. In particular, future values of the system variables remove this dependency, and we may use x_{t+k} for any integer $k \geq 0$ as an instrument for x_{t-1}. The corresponding IV estimators of R have the simple form

$$\hat{R}_{n,k} = \sum_{t=1}^{n-k} x_t x_{t+k}' \left(\sum_{t=1}^{n-k} x_{t-1} x_{t+k}' \right)^{-1}, \qquad k \in \{0, 1, 2, ...\}.$$

The estimator $\hat{R}_{n,k}$ is consistent and has the following limit distribution.

THEOREM 3.

(i) Under the assumptions of Lemma 3,

$$\sqrt{n} vec\left(\hat{R}_{n,k} - R\right) \Rightarrow -\rho^{k+1}\left(\rho^2 - 1\right) \frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}_{\frac{\pi}{2}}' \otimes I_2}{X(\rho)' \mathcal{R}_{\frac{\pi}{2}}' \Sigma_u \mathcal{R}_{\frac{\pi}{2}} X(\rho)} U \tag{22}$$

as $n \to \infty$, for each fixed $k \geq 0$, where U is an $N(0, (\rho^2 - 1)^{-1} \Sigma_u \otimes \Sigma_u)$ random vector. If (6) is replaced by uniform integrability of the sequence $(\|u_t\|^2)_{t \in \mathbb{N}}$, then U and $X(\rho)$ are uncorrelated random vectors.

(ii) If, in addition, u_t is an m-dependent sequence[2] for some $m \in \mathbb{N}$ such that $m \to \infty$ and $m/n \to 0$, U and $X(\rho)$ are independent random vectors and the limit distribution is mixed normal (MN); i.e.,

$$\sqrt{n} vec\left(\hat{R}_{n,k} - R\right) \Rightarrow \rho^{k+1} \sqrt{\rho^2 - 1}\, MN\left(0, \frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}_{\frac{\pi}{2}}' \otimes \Sigma_u}{X(\rho)' \mathcal{R}_{\frac{\pi}{2}}' \Sigma_u \mathcal{R}_{\frac{\pi}{2}} X(\rho)}\right). \tag{23}$$

Remark 9. The limit theory (22) relies on the central limit theorem for sample covariance matrices of Lemma 3 and shows that the IV estimator $\hat{R}_{n,k}$ is $\sqrt{n}$-consistent. However, as $X(\rho)$ is not necessarily Gaussian, lack of correlation between the Gaussian random vector U and $X(\rho)$ does not guarantee independence. In other words, a martingale difference assumption on a (non-Gaussian) sequence of innovations u_t is not sufficient for asymptotic mixed normality of the IV estimator $\hat{R}_{n,k}$. However, the limit random vector in (22) will have a mixed normal distribution if asymptotic independence is imposed on the sequence u_t.

Remark 10. Observe that the limit distribution of $\sqrt{n}(\hat{R}_{n,k} - R)$ is degenerate in the direction $X(\rho)$ in view of the singularity of the limit random matrix

$\mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}'_{\frac{\pi}{2}}$. In particular, as shown in the proof of Theorem 3, we have the representation

$$\sqrt{n}\left(\hat{R}_{n,k} - R\right) = \frac{\frac{1}{\sqrt{n}} \sum_{t=1}^{n-k-1} u_t \zeta'_{n,t+k+1} \left(\frac{\mathcal{R}_{\frac{\pi}{2}} x_n}{\|x_n\|}\right) \left(\frac{\mathcal{R}_{\frac{\pi}{2}} x_n}{\|x_n\|}\right)'}{\left(\frac{\mathcal{R}_{\frac{\pi}{2}} x_n}{\|x_n\|}\right)' \frac{1}{n} \sum_{t=1}^{n-k-1} \zeta_{n,t} \zeta'_{n,t+k+1} \left(\frac{\mathcal{R}_{\frac{\pi}{2}} x_n}{\|x_n\|}\right)} + O_p\left(\frac{1}{\sqrt{n}}\right),$$

so that $\sqrt{n}\left(\hat{R}_{n,k} - R\right) x_n / \|x_n\| = o_p(1)$.

Remark 11. When the sequence u_t is (asymptotically) independent, the mixed normal limit (23) facilitates inference, which may be conducted in the usual manner in view of the following arguments. First, from Lemmas 1 and 2(iii) we obtain

$$\begin{aligned}\left(\frac{1}{n}\sum_{t=1}^{n} x_{t-1}x'_{t+k}\right)^{-1} &= H_n \left(\frac{1}{n}\sum_{t=1}^{n-k} z_{t-1}z'_{t+k}\right)^{-1} H'_n \\ &\to_p \rho^{k+1}\left(\rho^2 - 1\right) \frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}_{\frac{\pi}{2}}}{X(\rho)' \mathcal{R}'_{\frac{\pi}{2}} \Sigma_u \mathcal{R}_{\frac{\pi}{2}} X(\rho)}.\end{aligned}$$

Next, define the residual moment matrix $\hat{\Sigma}_{uk} = n^{-1} \sum_{t=1}^{n} \hat{u}_{tk} \hat{u}'_{tk}$, where the residuals are constructed using the IV estimator: $\hat{u}_{tk} = x_t - \hat{R}_{n,k} x_{t-1}$. As shown in Section 5,

$$\hat{\Sigma}_{uk} \to_p \Sigma_u, \tag{24}$$

and then

$$\left(\frac{1}{n}\sum_{t=1}^{n} x_{t-1}x'_{t+k}\right)^{-1} \otimes \hat{\Sigma}_{uk} \to_p \rho^{k+1}\left(\rho^2 - 1\right) \left(\frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}_{\frac{\pi}{2}}}{X(\rho)' \mathcal{R}'_{\frac{\pi}{2}} \Sigma_u \mathcal{R}_{\frac{\pi}{2}} X(\rho)}\right) \otimes \Sigma_u,$$

giving a consistent estimator of the covariance matrix in (23). Thus, inference about R may be conducted using the standard formula for the variance matrix of $\hat{R}_k$, that is $(\sum_{t=1}^{n} x_{t-1} x'_{t+k})^{-1} \otimes \hat{\Sigma}_{uk}$.

Remark 12. The variance of the limit distribution (23) increases with k and is minimized for $k = 0$. This is explained by the fact that the instrument x_{t+k} is most effective for x_{t-1} when $k = 0$, and the relevance of the instrument deteriorates as k increases.

3. COEXPLOSIVE COINTEGRATED SYSTEMS

PM studied a triangular system with possibly coexplosive regressors. A simpler version of this system, which will be sufficient to demonstrate our findings, is given by

$$y_t = A w_t + \varepsilon_t, \tag{25}$$

$$w_t = x_t, \qquad x_t = R x_{t-1} + u_t, \tag{26}$$

$$R = \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix}, \qquad \rho > 1, \tag{27}$$

where A is an $m \times 2$ matrix of cointegrating coefficients, x_t is a bivariate vector of coexplosive autoregressions initialized at $x_0 = 0$, and $v_t = (\varepsilon_t', u_t')'$ is a sequence of independent, identically distributed $(0, \Sigma)$ random vectors with absolutely continuous density, where

$$\Sigma = \begin{bmatrix} \Sigma_\varepsilon & 0 \\ 0 & \Sigma_u \end{bmatrix} \tag{28}$$

is a positive definite matrix partitioned conformably with v_t. The regressor x_t is therefore uncorrelated with the system shocks ε_t.

PM noted that asymptotic behavior of the least squares estimator

$$\hat{A}_n = \left(\sum_{t=1}^{n} y_t x_t' \right) \left(\sum_{t=1}^{n} x_t x_t' \right)^{-1}$$

depends on the relationship between the regressors in (26), i.e., on the precise form of the autoregressive matrix R. When R has the form (27), so the regressors are coexplosive, $\hat{A}_n$ is consistent for A but has a degenerate mixed normal limiting distribution with convergence rate $n^{1/2}$. In particular, Theorem 2.3 of PM shows that

$$\begin{aligned} \sqrt{\frac{n}{\rho^2 - 1}} \left(\hat{A}_n - A \right) &\Rightarrow MN \left(0, H_\perp \left(H_\perp' \Sigma_{uu} H_\perp \right)^{-1} H_\perp' \otimes \Sigma_{\varepsilon\varepsilon} \right) \\ &= MN \left(0, \frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}_{\frac{\pi}{2}}'}{X(\rho)' \mathcal{R}_{\frac{\pi}{2}}' \Sigma_{uu} \mathcal{R}_{\frac{\pi}{2}} X(\rho)} \otimes \Sigma_{\varepsilon\varepsilon} \right), \end{aligned} \tag{29}$$

where $H_\perp = \mathcal{R}_{\frac{\pi}{2}} X(\rho) / \|X(\rho)\|$ in the notation of the limiting rotation matrix (9) given earlier. In proving (29), PM assumed that Σ has the block diagonal structure (28), so that x_t is uncorrelated with ε_t. However, as shown in Section 5, (29) continues to hold when the covariance structure is given by

$$\Sigma = \begin{bmatrix} \Sigma_\varepsilon & \Sigma_{\varepsilon u} \\ \Sigma_{u\varepsilon} & \Sigma_u \end{bmatrix}, \quad \text{with } \Sigma_{u\varepsilon} \neq 0. \tag{30}$$

From this result, it would seem that coexplosive behavior in the regressors does not cause an inconsistency, contrary to the VAR result (16) in Theorem 3. However, suppose that $w_t = x_{t-1}$ in (26), so that there is a simple time lag in the long-run structural relation. Such a lag has no effect on conventional cointegration limit theory. However, as we now demonstrate, in the context of coexplosive time series, the impact of dynamic timing is considerable. Let the corresponding least squares estimator of A, when $w_t = x_{t-1}$, be $\tilde{A}_n = (\sum_{t=1}^n y_t x_{t-1}')(\sum_{t=1}^n x_{t-1}x_{t-1}')^{-1}$.

THEOREM 4. *In the model (25)–(27) with $w_t = x_{t-1}$ and Σ is given by (30),*

$$\tilde{A}_n - A \to_p -\frac{\rho^2-1}{X(\rho)'\mathcal{R}_{\frac{\pi}{2}}'\Sigma_u\mathcal{R}_{\frac{\pi}{2}}X(\rho)}\Sigma_{\varepsilon u}\mathcal{R}_{\frac{\pi}{2}}X(\rho)X(\rho)'\mathcal{R}_{\frac{\pi}{2}}' \quad \text{as } n\to\infty.$$

Evidently, when there are coexplosive regressors, the critical factor in determining consistency of least squares regression is the dynamic timing of the regression system rather than independence (exogeneity) of the regressor in the system. As in the case of vector autoregression, consistency in estimation can be accomplished by using x_t as an instrument for x_{t-1} in the regression. This finding shows that weak exogeneity (Engle, Hendry, and Richard, 1983) in regression with explosive regressors can depend subtly on dynamic timing and in a manner quite different from stationary systems.[3] Under the condition $(\varepsilon_t, u_t) \sim iid(0, \Sigma)$, convention would dictate that x_{t-1} is predetermined (and exogenous) for A in the system $y_t = Ax_{t-1} + \varepsilon_t$, but jointly dependent and correlated with ε_t in the system $y_t = Ax_t + \varepsilon_t$. Curiously, however, in the presence of coexplosive regressors, least squares is consistent in the system $y_t = Ax_t + \varepsilon_t$ but inconsistent in the system $y_t = Ax_{t-1} + \varepsilon_t$. The explanation is the same as that for a VAR. In particular, the limiting singularity in the sample moment matrix that is caused by coexplosive behavior induces an endogeneity in the regressor x_{t-1}. As before, dynamic timing plays a role in the resulting endogeneity because upon transformation to resolve the effects of coexplosive behavior, the stationary component of the transformed regressor, which is forward looking and depends on u_t, is correlated with ε_t when $\Sigma_{u\varepsilon} \neq 0$.

4. CONCLUSIONS

Besides the intriguing nature of the inconsistency of least squares and the implicit endogeneity that arises in coexplosive VARs and structural systems, the limit distributions themselves have some interesting features. The supports of the limit distributions are bounded, and the densities have asymptotes at the boundary. In the VAR case, the limit distribution of the centered (own) autoregressive estimator $\hat{\rho} - \rho$ is an arc sine law, and its support is on the negative part of the real line. The finite sample distributions are bimodal with modes that are close to the boundary asymptotes in the limit distributions. When the explosive parameter $\rho \to 1$, the support of the limit distribution shrinks to the origin, and the least squares estimates are again consistent.

5. PROOFS

Proof of Lemma 2 (i) and (ii). We employ uniform integrability (U.I.) of the sequence $\|u_t\|^2$, as appears in (6), combined with a truncation argument. Choose a sequence $(k_n)_{n\in\mathbb{N}}$ such that $k_n \to \infty$ and $k_n/n^{1/2} \to 0$ as $n \to \infty$.

For part (i), define the random matrices

$$v_t = u_t u_t' - \Sigma_u \quad \text{and} \quad v_t^{(n)} = u_t u_t' \mathbf{1}\{\|u_t\| \le k_n\} - \mathbb{E}_{\mathcal{F}_{t-1}}\left(u_t u_t' \mathbf{1}\{\|u_t\| \le k_n\}\right).$$

It is easy to see that, for each n, $\{v_t^{(n)} : 1 \le t \le n\}$ is a $\mathcal{F}_t$ martingale difference array with finite second-moment. The identity $1 - \mathbf{1}\{\|u_t\| \le k_n\} = \mathbf{1}\{\|u_t\| > k_n\}$ yields

$$\begin{aligned}\left\|\frac{1}{n}\sum_{t=1}^{n}\left(v_t - v_t^{(n)}\right)\right\|_{L_1} &\le \frac{1}{n}\sum_{t=1}^{n}\left\|u_t u_t' \mathbf{1}\{\|u_t\| > k_n\} - \mathbb{E}_{\mathcal{F}_{t-1}}\left(u_t u_t' \mathbf{1}\{\|u_t\| > k_n\}\right)\right\|_{L_1} \\ &\le 2\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}\left(\|u_t\|^2 \mathbf{1}\{\|u_t\| > k_n\}\right) \\ &\le 2\max_{1\le t\le n}\mathbb{E}\left(\|u_t\|^2 \mathbf{1}\{\|u_t\| > k_n\}\right) \to 0 \qquad (31)\end{aligned}$$

as $n \to \infty$ by U.I. of $\|u_t\|^2$. By the Jensen inequality for conditional expectations,

$$\begin{aligned}\left\|v_t^{(n)}\right\|^2 &\le \left[\|u_t\|^2 \mathbf{1}\{\|u_t\| \le k_n\} + \mathbb{E}_{\mathcal{F}_{t-1}}\left(\|u_t\|^2 \mathbf{1}\{\|u_t\| \le k_n\}\right)\right]^2 \\ &\le 2\left[\|u_t\|^4 \mathbf{1}\{\|u_t\| \le k_n\} + \mathbb{E}_{\mathcal{F}_{t-1}}\left(\|u_t\|^4 \mathbf{1}\{\|u_t\| \le k_n\}\right)\right] \\ &\le 2k_n^2\left[\|u_t\|^2 + \mathbb{E}_{\mathcal{F}_{t-1}}\left(\|u_t\|^2\right)\right].\end{aligned}$$

Orthogonality of the martingale difference $v_t^{(n)}$ and the choice $k_n = o(n^{1/2})$ yield

$$\mathbb{E}\left\|\frac{1}{n}\sum_{t=1}^{n} v_t^{(n)}\right\|^2 = \frac{1}{n^2}\sum_{t=1}^{n}\mathbb{E}\left\|v_t^{(n)}\right\|^2 \le \frac{4k_n^2}{n^2}\sum_{t=1}^{n}\mathbb{E}\|u_t\|^2 = 4tr\left(\Sigma_u\right)\frac{k_n^2}{n} \to 0.$$

Hence, $n^{-1}\sum_{t=1}^{n} v_t^{(n)} \to 0$ in L_2 and, in view of (31), $n^{-1}\sum_{t=1}^{n} v_t \to 0$ in L_1.

For part (ii), define $\tilde{u}_t^{(n)} = u_t \mathbf{1}\{\|u_t\| \le k_n\}$ and note that

$$\frac{1}{n}\sum_{t=1}^{n} u_t u_{t+j}' = \frac{1}{n}\sum_{t=1}^{n}\tilde{u}_t^{(n)} u_{t+j}' + \frac{1}{n}\sum_{t=1}^{n} u_t \mathbf{1}\{\|u_t\| > k_n\} u_{t+j}'. \qquad (32)$$

Since $\tilde{u}_t^{(n)}$ is $\mathcal{F}_t$-adapted, $\{\tilde{u}_t^{(n)} \otimes u_{t+j} : 1 \le t \le n\}$ is a $\mathcal{F}_{t+j}$-martingale difference array for all $j \ge 1$ satisfying

$$\left\|\tilde{u}_t^{(n)}\right\|^2 \left\|u_{t+j}\right\|^2 \le k_n^2 \left\|u_{t+j}\right\|^2$$

for all $j, n \geq 1$. Orthogonality of $\{\tilde{u}_t^{(n)} \otimes u_{t+j}\}$ and the choice $k_n = o(n^{1/2})$ yield

$$\mathbb{E}\left\|\frac{1}{n}\sum_{t=1}^{n}\left(\tilde{u}_t^{(n)} \otimes u_{t+j}\right)\right\|^2 = \frac{1}{n^2}\sum_{t=1}^{n}\mathbb{E}\left(\left\|\tilde{u}_t^{(n)} \otimes u_{t+j}\right\|^2\right) \leq \frac{k_n^2}{n}\operatorname{tr}(\Sigma_u) \to 0,$$

so the first term on the right of (32) converges to zero in L_2. For the second term of (32), the Cauchy Schwarz inequality yields

$$\begin{aligned}\left\|\frac{1}{n}\sum_{t=1}^{n} u_t \mathbf{1}\{\|u_t\| > k_n\} u_{t+j}'\right\|_{L_1} &\leq \max_{1\leq t\leq n}\mathbb{E}\left(\|u_t\|\mathbf{1}\{\|u_t\| > k_n\}\left\|u_{t+j}\right\|\right)\\ &\leq \max_{1\leq t\leq n}\left[\mathbb{E}\left\|u_{t+j}\right\|^2\right]^{1/2}\left[\mathbb{E}\|u_t\|^2\mathbf{1}\{\|u_t\| > k_n\}\right]^{1/2}\\ &= \operatorname{tr}(\Sigma_u)^{1/2}\left[\max_{1\leq t\leq n}\mathbb{E}\left(\|u_t\|^2\mathbf{1}\{\|u_t\| > k_n\}\right)\right]^{1/2} \to 0\end{aligned}$$

by U.I. of $\|u_t\|^2$ in (6).

LEMMA 0. *Consider the $\mathcal{F}_{j-k-1}$-adapted sequence*

$$\psi_j = \sum_{i=k+1}^{j-1}\rho^{-i}u_{j-i}. \tag{33}$$

Under Assumption 1, the following apply for each fixed $k \geq 0$:

(i) the sequence $\{\left\|\psi_j\right\|^2, 1 \leq j \leq n\}$ is uniformly integrable:

$$\max_{1\leq j\leq n}\mathbb{E}\left(\left\|\psi_j\right\|^2\mathbf{1}\left\{\left\|\psi_j\right\| > \lambda_n\right\}\right) \to 0 \quad \text{as } n \to \infty \tag{34}$$

for any sequence $\lambda_n \to \infty$.

(ii) $n^{-1}\sum_{j=k+2}^{n}\psi_j\psi_j' \to_{L_1} \rho^{-2k}\Sigma_u/(\rho^2-1)$ as $n \to \infty$.

Proof. Denote $A_{n,j} = \{\|\psi_j\| > \lambda_n\}$. Since $\|\psi_j\| \leq \sum_{i=k+1}^{j-1}\rho^{-i}\|u_{j-i}\|$, we obtain, for each $j \in \{1, \ldots, n\}$,

$$\begin{aligned}\mathbb{E}\left(\left\|\psi_j\right\|^2\mathbf{1}_{A_{n,j}}\right) &\leq \sum_{i,l=k+1}^{j-1}\rho^{-i-l}\mathbb{E}\left[\left(\left\|u_{j-i}\right\|\mathbf{1}_{A_{n,j}}\right)\left(\left\|u_{j-l}\right\|\mathbf{1}_{A_{n,j}}\right)\right]\\ &\leq \sum_{i,l=k+1}^{j-1}\rho^{-i-l}\left(\mathbb{E}\left\|u_{j-i}\right\|^2\mathbf{1}_{A_{n,j}}\right)^{1/2}\left(\mathbb{E}\left\|u_{j-l}\right\|^2\mathbf{1}_{A_{n,j}}\right)^{1/2}\\ &\leq \max_{1\leq s\leq n}\mathbb{E}\left(\|u_s\|^2\mathbf{1}_{A_{n,j}}\right)\left(\sum_{i=0}^{\infty}\rho^{-i}\right)^2.\end{aligned}$$

Since the series in the final expression is convergent, the above bound shows that

$$\max_{1\le j\le n}\max_{1\le s\le n}\mathbb{E}\left(\|u_s\|^2\mathbf{1}_{A_{n,j}}\right)\to 0 \qquad \text{as } n\to\infty$$

for any sequence $\lambda_n \to \infty$ is sufficient for (34). Now

$$\begin{aligned}\max_{1\le j,s\le n}\mathbb{E}\left(\|u_s\|^2\mathbf{1}_{A_{n,j}}\right) &\le \max_{1\le j,s\le n}\mathbb{E}\left(\|u_s\|^2\mathbf{1}\left\{\|u_s\|>\lambda_n^{1/2}\right\}\mathbf{1}_{A_{n,j}}\right)\\ &\quad+\max_{1\le j,s\le n}\mathbb{E}\left(\|u_s\|^2\mathbf{1}\left\{\|u_s\|\le\lambda_n^{1/2}\right\}\mathbf{1}_{A_{n,j}}\right)\\ &\le \max_{1\le s\le n}\mathbb{E}\left(\|u_s\|^2\mathbf{1}\left\{\|u_s\|>\lambda_n^{1/2}\right\}\right)+\lambda_n\max_{1\le j\le n}P\left(A_{n,j}\right)\\ &\le o(1)+\frac{1}{\lambda_n}\max_{1\le j\le n}\mathbb{E}\left\|\psi_j\right\|^2=o(1)\end{aligned}$$

for any sequence $\lambda_n\to\infty$ as $n\to\infty$, by U.I. of the sequence $\|u_s\|^2$ and the Chebyshev inequality. This establishes (34). For part (ii), we can write

$$\begin{aligned}\sum_{j=k+2}^{n}\psi_j\psi_j' &= \sum_{j=k+2}^{n}\sum_{i=k+1}^{j-1}\sum_{l=k+1}^{j-1}\rho^{-i-l}u_{j-i}u_{j-l}' = \sum_{i=k+1}^{n-1}\sum_{j=i+1}^{n}\sum_{l=k+1}^{j-1}\rho^{-i-l}u_{j-i}u_{j-l}'\\ &=\sum_{i=k+1}^{n-1}\rho^{-i}\sum_{j=i+1}^{n}u_{j-i}\left(\sum_{l=k+1}^{i-1}\rho^{-l}u_{j-l}'+\sum_{l=i}^{j-1}\rho^{-l}u_{j-l}'\right)\\ &=\sum_{i=k+1}^{n-1}\rho^{-i}\sum_{l=k+1}^{i-1}\rho^{-l}\sum_{j=i+1}^{n}u_{j-i}u_{j-l}'+\sum_{i=k+1}^{n-1}\rho^{-i}\sum_{l=i+1}^{n-1}\rho^{-l}\sum_{j=l+1}^{n}u_{j-i}u_{j-l}'\\ &\quad+\sum_{i=k+1}^{n-1}\rho^{-2i}\sum_{j=i+1}^{n}u_{j-i}u_{j-i}'.\end{aligned}\tag{35}$$

For the first term of (35), since $\sum_{j=i+1}^{n}u_{j-i}u_{j-l}'=\sum_{j=1}^{n-i}u_ju_{j+i-l}'$ and $\sum_{i=1}^{\infty}i\rho^{-i}<\infty$,

$$\left\|\sum_{i=k+1}^{n-1}\rho^{-i}\sum_{l=k+1}^{i-1}\rho^{-l}\sum_{j=i+1}^{n}u_{j-i}u_{j-l}'\right\|_{L_1}\le\max_{1\le j\le n}\left\|\sum_{t=1}^{n}u_tu_{t+j}'\right\|_{L_1}\left(\sum_{i=1}^{\infty}\rho^{-i}\right)^2+O(1).$$

The same bound applies for the second term of (35) so, upon normalization by n^{-1}, the first two terms of (35) converge to 0 in L_1. Finally, since $\|\sum_{i=k+1}^{n-1}\rho^{-2i}\sum_{j=n-i+1}^{n}u_ju_j'\|_{L_1}$ is bounded by $\operatorname{tr}(\Sigma_u)\sum_{i=1}^{\infty}i\rho^{-2i}$, (35) yields

$$\begin{aligned}\frac{1}{n}\sum_{j=k+2}^{n}\psi_j\psi_j' &= \frac{1}{n}\sum_{i=k+1}^{n-1}\rho^{-2i}\sum_{j=1}^{n-i}u_ju_j' + o_{L_1}(1)\\ &= \left(\sum_{i=k+1}^{n-1}\rho^{-2i}\right)\frac{1}{n}\sum_{j=1}^{n}u_ju_j' + o_{L_1}(1)\\ &= \left(\sum_{i=k+1}^{\infty}\rho^{-2i}\right)\Sigma_u + o_{L_1}(1)\end{aligned}$$

by Lemma 2(i), as required, where $o_{L_1}(1)$ denotes convergence to 0 in L_1.

Proof of Lemma 1. Almost sure convergence of the infinite series in (4) can be proved by applying the Rademacher-Menchoff convergence theorem for orthogonal random variables (Thm. 2.3.2 of Stout, 1974) when $(u_t)_{t\geq 1}$ is an uncorrelated sequence or by the L_2 martingale convergence theorem applied to $\rho^{-n}x_n = \sum_{t=1}^{n}\rho^{-t}u_t$ when $(u_t)_{t\geq 1}$ is a martingale difference sequence. Lai and Wei (1983) show that, under (5) of Assumption 1, $X(\rho)\neq 0\ a.s.$ Almost sure convergence of H_n follows by applying $\rho^{-n}x_n \to_{a.s.} X(\rho)$ to (8) and using continuity of norms.

Proof of Lemma 2 (iii)–(vi). For part (iii), (12) and part (ii) yield, for any fixed $k\geq 1$,

$$\begin{aligned}\left\|\frac{1}{n}\sum_{t=1}^{n-k}u_t\zeta_{n,t+k}'\right\|_{L_1} &= \rho^{k-1}\left\|\frac{1}{n}\sum_{t=1}^{n-k}u_t\sum_{j=k}^{n-t}\rho^{-j}u_{t+j}'\right\|_{L_1}\\ &= \rho^{k-1}\left\|\sum_{j=k}^{n}\rho^{-j}\frac{1}{n}\sum_{t=1}^{n}u_tu_{t+j}'\right\|_{L_1} + O\left(\frac{1}{n}\right)\\ &\leq \max_{1\leq j\leq n}\left\|\frac{1}{n}\sum_{t=1}^{n}u_tu_{t+j}'\right\|_{L_1}\rho^{k-1}\sum_{j=1}^{\infty}\rho^{-j} = o(1),\end{aligned}$$

by part (ii), since both $\|n^{-1}\sum_{t=1}^{n}u_t\sum_{j=n-t+1}^{n}\rho^{-j}u_{t+j}'\|_{L_1}$ and $\|n^{-1}\sum_{t=n-k+1}^{n} u_tu_{t+j}'\|_{L_1}$ are of order $O(n^{-1})$.

For part (iv), the definition of $\zeta_{n,t}$ in (12) yields the identities

$$\zeta_{n,t} = \rho^{-k}\zeta_{n,t+k} + \sum_{j=0}^{k-1}\rho^{-(j+1)}u_{t+j} \quad \text{and} \quad \zeta_{n,t+k} = \rho^{-1}\zeta_{n,t+k+1} + \rho^{-1}u_{t+k} \tag{36}$$

for any $t\leq n$ and any fixed $k\geq 0$. Since, by part (iii),

$$\left\|\frac{1}{n}\sum_{t=1}^{n-k}\left(\sum_{j=0}^{k-1}\rho^{-(j+1)}u_{t+j}\right)\zeta_{n,t+k}'\right\|_{L_1} \leq \sum_{j=0}^{k-1}\rho^{-(j+1)}\left\|\frac{1}{n}\sum_{t=1}^{n-k}u_{t+j}\zeta_{n,t+k}'\right\|_{L_1} \to 0,$$

the first identity in (36) implies that

$$\left\| \frac{1}{n}\sum_{t=1}^{n-k} \zeta_{n,t}\zeta_{n,t+k}' - \rho^{-k}\frac{1}{n}\sum_{t=1}^{n-k} \zeta_{n,t+k}\zeta_{n,t+k}' \right\|_{L_1} \to 0 \quad \text{as } n \to \infty \tag{37}$$

for all fixed $k \geq 0$. The second identity in (36) yields

$$\zeta_{n,t+k}\zeta_{n,t+k}' = \rho^{-2}\left(\zeta_{n,t+k+1}\zeta_{n,t+k+1}' + u_{t+k}\zeta_{n,t+k+1}' + \zeta_{n,t+k+1}u_{t+k}' + u_{t+k}u_{t+k}'\right).$$

Summing over $t \in \{1, ..., n-k-1\}$ yields

$$\left(1-\rho^{-2}\right)\frac{1}{n}\sum_{t=1}^{n-k-1} \zeta_{n,t+k}\zeta_{n,t+k}' = \rho^{-2}\frac{1}{n}\sum_{t=1}^{n-k-1} u_{t+k}u_{t+k}' + R_n,$$

where the remainder term

$$R_n = \frac{\rho^{-2}}{n}\left(\zeta_{n,n}\zeta_{n,n}' - \zeta_{n,k+1}\zeta_{n,k+1}' + \sum_{t=k+1}^{n-1} u_t\zeta_{n,t+1}' + \sum_{t=k+1}^{n-1} \zeta_{n,t+1}u_t'\right)$$

tends to 0 in L_1 by part (iii) with $k = 1$. Therefore,

$$\left\| \frac{1}{n}\sum_{t=1}^{n-k-1} \zeta_{n,t+k}\zeta_{n,t+k}' - \frac{1}{\rho^2 - 1}\frac{1}{n}\sum_{t=k+1}^{n-1} u_t u_t' \right\|_{L_1} \to 0 \quad \text{as } n \to \infty \tag{38}$$

for all fixed $k \geq 0$. Part (iv) of the lemma follows by combining (37), (38), and the law of large numbers for $n^{-1}\sum_{t=1}^{n} u_t u_t'$ of part (i), which applies in view of the fact that k is a fixed integer independent of n.

For part (v), using (11) to expand z_{1t-1} and z_{2t-1}, we obtain the following rates of convergence for the elements of the matrix $\sum_{t=1}^{n-k} z_{t-1}z_{t+k}'$ for each fixed $k \geq -1$:

$$\sum_{t=1}^{n-k} z_{1t-1}z_{1t+k} = O_p\left(\rho^{2n}\right), \tag{39}$$

$$\max\left\{\sum_{t=1}^{n-k} z_{1t-1}z_{2t+k}, \sum_{t=1}^{n-k} z_{2t-1}z_{1t+k}\right\} = O_p\left(\rho^{n}\right). \tag{40}$$

The first part of (40) can be deduced by (11) and the bound

$$\begin{aligned}\left\| \sum_{t=1}^{n-k-1} x_{t-1}\zeta_{n,t+k+1}' \right\|_{L_1} &\leq \sum_{t=1}^{n-k-1} \left(\mathbb{E}\left\|x_{t-1}\right\|^2 \mathbb{E}\left\|\zeta_{n,t+k+1}\right\|^2\right)^{1/2} \\ &\leq C\sum_{t=1}^{n}\left(\sum_{j=0}^{t-2}\rho^{2j}\right)^{1/2} = O\left(\rho^{n}\right),\end{aligned} \tag{41}$$

where $C = \mathrm{tr}\Sigma_u/\sqrt{\rho^2-1}$. Since k is fixed, the second part of (40) can be deduced by an identical argument. For (39), the identity $z_{t+k} = \rho^{k+1} z_{t-1} + H_n' \sum_{j=0}^{k} \rho^{k-j} u_{t+j}$ for all $k \geq -1$ (when $k=-1$ the empty sum is equal to 0) yields

$$\begin{aligned}\sum_{t=1}^{n-k} z_{1t-1} z_{1t+k} &= \rho^{k+1} \sum_{t=1}^{n-k} z_{1t-1}^2 + \frac{x_n'}{\|x_n\|} \sum_{t=1}^{n-k} x_{t-1} \left(\sum_{j=0}^{k} \rho^{k-j} u_{t+j} \right)' \frac{x_n}{\|x_n\|} \\ &= \rho^{k+1} \sum_{t=1}^{n-k} z_{1t-1}^2 + O_p\left(\rho^n\right),\end{aligned}$$

where the last order of magnitude is obtained by the same argument used to prove (40). Now (39) follows by direct computation of $\sum_{t=1}^{n-k} \mathbb{E}\|x_{t-1}\|^2$. For the remaining element of $\sum_{t=1}^{n-k} z_{t-1} z_{t+k}'$, recalling that $z_{2n} = 0$ and using (11) and part (iv), we obtain as $n \to \infty$,

$$\begin{aligned}\frac{1}{n} \sum_{t=1}^{n-k-1} z_{2t-1} z_{2t+k} &= \left(\frac{\mathcal{R}_{\frac{\pi}{2}} x_n}{\|x_n\|} \right)' \frac{1}{n} \sum_{t=1}^{n-k-1} \zeta_{n,t} \zeta_{n,t+k+1}' \left(\frac{\mathcal{R}_{\frac{\pi}{2}} x_n}{\|x_n\|} \right) \\ &\to_p \frac{\rho^{-k-1}}{\rho^2-1} \left(\mathcal{R}_{\frac{\pi}{2}} \frac{X(\rho)}{\|X(\rho)\|} \right)' \Sigma_u \left(\mathcal{R}_{\frac{\pi}{2}} \frac{X(\rho)}{\|X(\rho)\|} \right). \end{aligned} \tag{42}$$

The determinant of the matrix $\sum_{t=1}^{n-k} z_{t-1} z_{t+k}'$ is given by

$$\begin{aligned}D_n &= \left(\sum_{t=1}^{n-k} z_{1t-1} z_{1t+k} \right) \left(\sum_{t=1}^{n-k-1} z_{2t-1} z_{2t+k} \right) \\ &\quad - \left(\sum_{t=1}^{n-k} z_{2t-1} z_{1t+k} \right) \left(\sum_{t=1}^{n-k-1} z_{1t-1} z_{2t+k} \right) \\ &= \left(\sum_{t=1}^{n-k} z_{1t-1} z_{1t+k} \right) \left(\sum_{t=1}^{n-k-1} z_{2t-1} z_{2t+k} \right) \left\{ 1 + O_p\left(n^{-1}\right) \right\}. \end{aligned} \tag{43}$$

Using (39)–(43), the inverse of the signal matrix is given by

$$\begin{aligned}\left(\frac{1}{n} \sum_{t=1}^{n-k} z_{t-1} z_{t+k}' \right)^{-1} &= \frac{1}{n^{-1} D_n} \begin{bmatrix} \sum_{t=1}^{n-k-1} z_{2t-1} z_{2t+k} & -\sum_{t=1}^{n-k-1} z_{1t-1} z_{2t+k} \\ -\sum_{t=1}^{n-k} z_{2t-1} z_{1t+k} & \sum_{t=1}^{n-k} z_{1t-1} z_{1t+k} \end{bmatrix} \\ &= \begin{bmatrix} n \left(\sum_{t=1}^{n-k} z_{1t-1} z_{1t+k} \right)^{-1} & -O_p\left(\rho^{-2n} \sum_{t=1}^{n-k-1} \times z_{1t-1} z_{2t+k} \right) \\ -O_p\left(\rho^{-2n} \sum_{t=1}^{n-k} \times z_{2t-1} z_{1t+k} \right) & n \left(\sum_{t=1}^{n-k-1} z_{2t-1} z_{2t+k} \right)^{-1} \end{bmatrix},\end{aligned}$$

where the last equality holds up to $1+O_p(n^{-1})$. Thus, for each fixed $k \geq -1$,

$$\left(\frac{1}{n}\sum_{t=1}^{n-k} z_{t-1}z_{t+k}'\right)^{-1} = \left[1+O_p\left(\frac{1}{n}\right)\right]\begin{bmatrix} O_p\left(n\rho^{-2n}\right) & O_p\left(\rho^{-n}\right) \\ O_p\left(\rho^{-n}\right) & \left(\frac{1}{n}\sum_{t=1}^{n-k-1} z_{2t-1}z_{2t+k}\right)^{-1}\end{bmatrix}, \quad \textbf{(44)}$$

and the result follows from (42).

For part (vi), after a long but elementary calculation, the identities

$$\sum_{t=1}^{n} x_{it-1}^2 = \frac{1}{\rho^2-1}\left(x_{in}^2 - 2\rho\sum_{t=1}^{n} x_{it-1}u_{it} - \sum_{t=1}^{n} u_{it}^2\right),$$

$$\sum_{t=1}^{n} x_{1t-1}x_{2t-1} = \frac{1}{\rho^2-1}\left(x_{1n}x_{2n} - \rho\sum_{t=1}^{n} x_{1t-1}u_{2t} - \rho\sum_{t=1}^{n} x_{2t-1}u_{1t} - \sum_{t=1}^{n} u_{1t}u_{2t}\right),$$

$$\sum_{t=1}^{n} x_{it-1}u_{jt} = x_{in}\sum_{t=1}^{n}\rho^{-(n-t+1)}u_{jt} - \rho^{-1}\sum_{t=1}^{n} u_{it}u_{jt} - \sum_{t=1}^{n}\left(\sum_{s=t+1}^{n}\rho^{t-1-s}u_{is}\right)u_{jt},$$

for each $i, j \in \{1,2\}$ yield the following expressions for the determinant and trace of the sample moment matrix:

$$\begin{aligned}\frac{\rho^{-2n}}{n}\det\left(X'X\right) &= \frac{1}{\left(\rho^2-1\right)^2}\frac{\rho^{-2n}}{n}\left(x_{1n}^2\sum_{t=1}^{n} u_{2t}^2 + x_{2n}^2\sum_{t=1}^{n} u_{1t}^2 - 2x_{1n}x_{2n}\sum_{t=1}^{n} u_{1t}u_{2t}\right) \\ &\quad + O_p\left(\frac{1}{\sqrt{n}}\right) \\ &\to_p \frac{1}{\left(\rho^2-1\right)^2}\left[X_1(\rho)^2\sigma_2^2 + X_2(\rho)^2\sigma_1^2 - 2X_1(\rho)X_2(\rho)\sigma_{12}\right] \quad \textbf{(45)}\end{aligned}$$

$$\rho^{-2n}tr\left(X'X\right) = \frac{\rho^{-2n}}{\rho^2-1}\left(x_{1n}^2 + x_{2n}^2\right) + O_p\left(\rho^{-n}\right) \to_p \frac{1}{\rho^2-1}\left[X_1(\rho)^2 + X_2(\rho)^2\right]. \quad \textbf{(46)}$$

The asymptotic behavior of the eigenvalues of $X'X$ can be obtained from (45) and (46):

$$\begin{aligned}\lambda_{\max}\left(X'X\right) &= \frac{1}{2}\left[\mathrm{tr}\left(X'X\right) + \sqrt{(\mathrm{tr}X'X)^2 - 4\det(X'X)}\right] \\ &= \mathrm{tr}\left(X'X\right) + O_p\left[\frac{\det\left(X'X\right)}{(\mathrm{tr}X'X)}\right],\end{aligned}$$

$$\begin{aligned}\lambda_{\min}\left(X'X\right) &= \frac{1}{2}\left[\mathrm{tr}\left(X'X\right) - \sqrt{(\mathrm{tr}X'X)^2 - 4\det(X'X)}\right] \\ &= \frac{\det\left(X'X\right)}{\mathrm{tr}(X'X)} + O_p\left[\frac{\det\left(X'X\right)^2}{(\mathrm{tr}(X'X))^2}\right].\end{aligned}$$

Combining the above expressions and noting that $\mathrm{tr}(\rho^{-2n}X'X) \to_p \|X(\rho)\|^2 > 0$ a.s., we obtain

$$\begin{aligned}
\frac{\log \lambda_{\max}\left(X'X\right)}{\lambda_{\min}\left(X'X\right)} &= \frac{\mathrm{tr}\left(X'X\right)\log\left[\rho^{2n}\mathrm{tr}\left(\rho^{-2n}X'X\right)\right]}{\det\left(X'X\right)} + o_p(1) \\
&= \frac{2(\log\rho)\,\mathrm{tr}\left(\rho^{-2n}X'X\right)}{n^{-1}\rho^{-2n}\det\left(X'X\right)} + O_p\left(\frac{\mathrm{tr}\left(X'X\right)}{\det\left(X'X\right)}\right) \\
&\to_p \frac{2(\log\rho)\left(\rho^2-1\right)\left[X_1(\rho)^2 + X_2(\rho)^2\right]}{\sigma_2^2 X_1(\rho)^2 + \sigma_1^2 X_2(\rho)^2 - 2\sigma_{12}X_1(\rho)X_2(\rho)} \\
&= 2(\log\rho)\left(\rho^2-1\right)\frac{\|X(\rho)\|^2}{X(\rho)'\mathcal{R}_{\frac{\pi}{2}}'\Sigma_u\mathcal{R}_{\frac{\pi}{2}}X(\rho)}
\end{aligned}$$

by (45) and (46), as required. Almost sure positivity of the above probability limit is ensured, since $\rho > 1$, $\|X(\rho)\| > 0$ *a.s.* by Assumption 1 and $\mathcal{R}_{\frac{\pi}{2}}'\Sigma_u\mathcal{R}_{\frac{\pi}{2}}$ is a positive definite matrix.

Proof of Lemma 3. Using (12) and changing the order of summation, we obtain

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{t=1}^{n-k-1}\left(\zeta_{n,t+k+1}\otimes u_t\right) &= \frac{\rho^k}{\sqrt{n}}\sum_{t=1}^{n-k-1}\sum_{j=t+k+1}^{n}\rho^{-(j-t)}\left(u_j\otimes u_t\right) \\
&= \frac{\rho^k}{\sqrt{n}}\sum_{j=k+2}^{n}\sum_{t=1}^{j-k-1}\rho^{-(j-t)}\left(u_j\otimes u_t\right) \\
&= \frac{\rho^k}{\sqrt{n}}\sum_{j=k+2}^{n}\left(u_j\otimes\psi_j\right) = \rho^k\sum_{j=k+2}^{n}\eta_{n,j}, \qquad \textbf{(47)}
\end{aligned}$$

recalling the definitions (14) and (33): $\eta_{n,j} = n^{-1/2}u_j\otimes\psi_j$ with $\psi_j = \sum_{i=k+1}^{j-1}\rho^{-i}u_{j-i}$. Since ψ_j is $\mathcal{F}_{j-k-1}$-adapted and $\mathbb{E}(\|u_j\|^2\|u_{j+i}\|^2) < \infty$ for all $i, j \geq 1$, $\eta_{n,j}$ is a square integrable $\mathcal{F}_j$-martingale difference array and the limit distribution of $\sum_{j=k+2}^{n}\eta_{n,j}$ can be derived by a standard martingale CLT; e.g., Cor. 3.1 of Hall and Heyde (1980). Using part (ii) of Lemma 0, the conditional variance of $\sum_{j=k+2}^{n}\eta_{n,j}$ is given by

$$\begin{aligned}
\sum_{j=k+2}^{n}\mathbb{E}_{\mathcal{F}_{j-1}}\left(\eta_{n,j}\eta_{n,j}'\right) &= \frac{1}{n}\sum_{j=k+2}^{n}\left[\mathbb{E}_{\mathcal{F}_{j-1}}\left(u_ju_j'\right)\otimes\psi_j\psi_j'\right] \\
&= \Sigma_u\otimes\left(\frac{1}{n}\sum_{j=k+2}^{n}\psi_j\psi_j'\right) \to_p \frac{\rho^{-2k}\left(\Sigma_u\otimes\Sigma_u\right)}{\rho^2-1},
\end{aligned}$$

as $n \to \infty$ for each fixed $k \geq 0$. Combined with (47), the above limit yields the required asymptotic variance. For the Lindeberg condition, $\mathcal{F}_{j-1}$-measurability of ψ_j and the inequality $\mathbf{1}\{\|u_j\|\|\psi_j\| > \sqrt{n}\delta\} \leq \mathbf{1}\{\|\psi_j\| > n^{1/4}\sqrt{\delta}\} + \mathbf{1}\{\|u_j\| > n^{1/4}\sqrt{\delta}\}$ yield the bound

$$\sum_{j=k+2}^{n} \mathbb{E}_{\mathcal{F}_{j-1}}\left(\left\|\eta_{n,j}\right\|^2 \mathbf{1}\left\{\left\|\eta_{n,j}\right\| > \delta\right\}\right) \leq I_1(n) + I_2(n), \tag{48}$$

where

$$I_1(n) = \frac{1}{n}\sum_{j=k+2}^{n} \left\|\psi_j\right\|^2 \mathbb{E}_{\mathcal{F}_{j-1}}\left(\left\|u_j\right\|^2 \mathbf{1}\left\{\left\|u_j\right\| > n^{1/4}\sqrt{\delta}\right\}\right),$$
$$I_2(n) = \frac{1}{n}\sum_{j=k+2}^{n} \left\|\psi_j\right\|^2 \mathbf{1}\left\{\left\|\psi_j\right\| > n^{1/4}\sqrt{\delta}\right\} \mathbb{E}_{\mathcal{F}_{j-1}}\left(\left\|u_j\right\|^2\right).$$

Since $\mathbb{E}_{\mathcal{F}_{j-1}}(\|u_j\|^2) = \mathrm{tr}(\Sigma_u)$, $I_2(n) \to 0$ in L_1 by U.I. of the sequence $\{\|\psi_j\|^2 : j \leq n\}$, established in part (i) of Lemma 0. To show that $I_1(n) \to_{L_1} 0$, let

$$\kappa_{n,\delta} = \max_{1\leq j\leq n} \mathbb{E}\left(\left\|u_j\right\|^2 \mathbf{1}\left\{\left\|u_j\right\| > n^{1/4}\sqrt{\delta}\right\}\right).$$

By (6) we know that $\kappa_{n,\delta} \to 0$ as $n \to \infty$ for any $\delta > 0$. We may obtain a bound for $I_1(n)$ by using the identity $1 = \mathbf{1}\{\|\psi_j\|^2 > \kappa_{n,\delta}^{-1/2}\} + \mathbf{1}\{\|\psi_j\|^2 \leq \kappa_{n,\delta}^{-1/2}\}$ as follows:

$$\begin{aligned} I_1(n) \leq{}& \frac{1}{n}\sum_{j=k+2}^{n} \left\|\psi_j\right\|^2 \mathbf{1}\left\{\left\|\psi_j\right\|^2 > \kappa_{n,\delta}^{-1/2}\right\} \mathbb{E}_{\mathcal{F}_{j-1}}\left(\left\|u_j\right\|^2\right) \\ &+ \kappa_{n,\delta}^{-1/2}\frac{1}{n}\sum_{j=k+2}^{n} \mathbb{E}_{\mathcal{F}_{j-1}}\left(\left\|u_j\right\|^2 \mathbf{1}\left\{\left\|u_j\right\| > n^{1/4}\sqrt{\delta}\right\}\right). \end{aligned}$$

Since, for any $\delta > 0$, $\kappa_{n,\delta}^{-1/2} \to \infty$ as $n \to \infty$ and $\mathbb{E}_{\mathcal{F}_{j-1}}(\|u_j\|^2)$ is constant, taking expectations and using (34) and (6) we obtain

$$\begin{aligned} \|I_1(n)\|_{L_1} &\leq \kappa_{n,\delta}^{-1/2}\frac{1}{n}\sum_{j=k+2}^{n} \mathbb{E}\left(\left\|u_j\right\|^2 \mathbf{1}\left\{\left\|u_j\right\| > n^{1/4}\sqrt{\delta}\right\}\right) + o(1) \\ &\leq \kappa_{n,\delta}^{-1/2} \max_{1\leq j\leq n} \mathbb{E}\left(\left\|u_j\right\|^2 \mathbf{1}\left\{\left\|u_j\right\| > n^{1/4}\sqrt{\delta}\right\}\right) + o(1) \\ &= \kappa_{n,\delta}^{1/2} + o(1) = o(1). \end{aligned}$$

Since both $I_1(n)$ and $I_2(n)$ tend to 0 in L_1, (48) implies that the Lindeberg condition is satisfied.

Proof of Theorem 1. Using the representation $H_n = \frac{1}{\|x_n\|}[x_n, \mathcal{R}_{\frac{\pi}{2}} x_n]$ and (44) with $k = -1$, we obtain

$$\hat{R}_n - R = \sum_{t=1}^{n} u_t x_{t-1}' \left(\sum_{t=1}^{n} x_{t-1} x_{t-1}' \right)^{-1} = \frac{1}{n} \sum_{t=1}^{n} u_t z_{t-1}' \left(\frac{1}{n} \sum_{t=1}^{n} z_{t-1} z_{t-1}' \right)^{-1} H_n'$$
$$= \left\{ 1 + O_p\left(\frac{1}{n}\right) \right\} \left[O_p\left(\rho^{-2n} \sum_{t=1}^{n} u_t x_{t-1}' \right), \left(n^{-1} \sum_{t=1}^{n} z_{2t-1}^2 \right)^{-1} \frac{1}{n} \sum_{t=1}^{n} u_t z_{2t-1} \right] H_n'$$
$$= \left\{ 1 + O_p\left(\frac{1}{n}\right) \right\} \left[O_p\left(\rho^{-n} \right), \left(n^{-1} \sum_{t=1}^{n} z_{2t-1}^2 \right)^{-1} \frac{1}{n} \sum_{t=1}^{n} u_t z_{2t-1} + O_p\left(\frac{1}{n}\right) \right] H_n'$$
$$= \left\{ 1 + O_p\left(\frac{1}{n}\right) \right\} \left(n^{-1} \sum_{t=1}^{n} z_{2t-1}^2 \right)^{-1} \left(\frac{1}{n} \sum_{t=1}^{n} u_t z_{2t-1} \right) \frac{x_n'}{\|x_n\|} \mathcal{R}_{\frac{\pi}{2}}', \tag{49}$$

since $\|\rho^{-n} \sum_{t=1}^{n} u_t x_{t-1}'\|_{L_1} \leq (\mathrm{tr}\Sigma_u)^{1/2} \rho^{-n} \sum_{t=1}^{n} (\mathbb{E}\|x_{t-1}\|^2)^{1/2} \leq \mathrm{tr}\Sigma_u / (\rho^2 - 1)^{3/2}$. Using (11) and the second identity in (36) we can write

$$\frac{1}{n} \sum_{t=1}^{n} u_t z_{2t-1} = -\frac{1}{n} \sum_{t=1}^{n-1} u_t \zeta_{n,t}' \mathcal{R}_{\frac{\pi}{2}} \frac{x_n}{\|x_n\|} + O_p\left(\frac{1}{n}\right)$$
$$= -\rho^{-1} \left(\frac{1}{n} \sum_{t=1}^{n-1} u_t u_t' + \frac{1}{n} \sum_{t=1}^{n-1} u_t \zeta_{n,t+1}' \right) \mathcal{R}_{\frac{\pi}{2}} \frac{x_n}{\|x_n\|}$$
$$= -\rho^{-1} \Sigma_u \mathcal{R}_{\frac{\pi}{2}} \frac{x_n}{\|x_n\|} + o_p(1) \tag{50}$$

by parts (i) and (iii) of Lemma 2. Combining (49) and (50) we obtain

$$\hat{R}_n - R = -\rho^{-1} \left[\left(\frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho)}{\|X(\rho)\|} \right)' \frac{\Sigma_u}{\rho^2 - 1} \frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho)}{\|X(\rho)\|} \right]^{-1} \Sigma_u \frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho)}{\|X(\rho)\|} \left(\frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho)}{\|X(\rho)\|} \right)'$$
$$+ o_p(1)$$
$$= -\frac{\rho^2 - 1}{\rho} \frac{\Sigma_u \mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}_{\frac{\pi}{2}}'}{X(\rho)' \mathcal{R}_{\frac{\pi}{2}}' \Sigma_u \mathcal{R}_{\frac{\pi}{2}} X(\rho)} + o_p(1)$$

by Lemma 1 and (42) with $k = -1$.

Proof of Theorem 2. For the limit density of $\hat{\rho}$, define $Y = -a_\rho \frac{\xi_1^2}{\xi_1^2 + \xi_2^2}$ with $a_\rho = \frac{\rho^2 - 1}{\rho}$, and observe that $Y = -a_\rho h_1^2$, where

$$h = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} := \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \left(\xi_1^2 + \xi_2^2 \right)^{-1/2} \tag{51}$$

is uniformly distributed on the sphere $h'h = 1$ (cf. Phillips, 1984). Using the representation $(h_1, h_2) = (\cos\theta, -\sin\theta)$, we have $Y = -a_\rho \cos^2\theta$ and so

$$dY/d\theta = 2a_\rho \cos\theta \sin\theta = 2a_\rho \left(\frac{-Y}{a_\rho} \right)^{1/2} \left\{ 1 - \frac{-Y}{a_\rho} \right\}^{1/2}$$
$$= 2(-Y)^{1/2} \left(a_\rho + Y \right)^{1/2}.$$

A full range of values of h_1^2 is accommodated by restricting the domain of θ to the subinterval $[0, \pi/2]$. Over this domain, θ is uniformly distributed with density $\frac{2}{\pi}$. We deduce that

$$pdf_Y(y) = \frac{2}{\pi}\left|\frac{d\theta}{dY}\right| = \frac{1}{\pi}\frac{1}{(-y)^{1/2}\left(a_\rho + y\right)^{1/2}}, \quad \text{for } y \in \left(-a_\rho, 0\right).$$

This density is that of an arc sine law and is shown in Figure 1 for $\rho = 1.04$ and $a_\rho = 0.078$.

Next, for the limit density of $\hat{\beta}$, define $Z = a_\rho \frac{\xi_1 \xi_2}{\xi_1^2 + \xi_2^2}$. Using (51), we have

$$Z = a_\rho h_1 h_2 = -a_\rho \cos\theta \sin\theta = -a_\rho \sin(2\theta)/2,$$

so that the Jacobian is

$$dZ/d\theta = a_\rho \cos(2\theta) = a_\rho \left\{1 - \frac{4}{a_\rho^2} Z^2\right\}^{1/2} = \left\{a_\rho^2 - 4Z^2\right\}^{1/2}.$$

Again, we can restrict the domain of θ to the subinterval $[0, \pi/2]$ with density $\frac{2}{\pi}$, as the Jacobian involves only $Z^2 = a_\rho^2 h_1^2 h_2^2$ and is therefore invariant to the sign of $h_1 h_2$. It follows that the density of Z is

$$pdf_Z(z) = \frac{2}{\pi}\left|\frac{d\theta}{dZ}\right| = \frac{2}{\pi}\frac{1}{\left\{a_\rho^2 - 4Z^2\right\}^{1/2}}, \quad \text{for } z \in \left(-\frac{a_\rho}{2}, \frac{a_\rho}{2}\right),$$

as stated.

Proof of Theorem 3. For any fixed $k \geq 0$ orthogonality of the matrix H_n and (44) yield, up to $1 + O_p(n^{-1})$,

$$\begin{aligned}
\hat{R}_{n,k} - R &= \frac{1}{n}\sum_{t=1}^{n-k} u_t z'_{t+k}\left(\frac{1}{n}\sum_{t=1}^{n-k} z_{t-1} z'_{t+k}\right)^{-1} H'_n \\
&= \frac{1}{n}\sum_{t=1}^{n-k} u_t \left[z_{1t+k}, z_{2t+k}\right]\begin{bmatrix} O_p\left(n\rho^{-2n}\right) & O_p\left(\rho^{-n}\right) \\ O_p\left(\rho^{-n}\right) & \left(n^{-1}\sum_{t=1}^{n-k-1} z_{2t-1} z_{2t+k}\right)^{-1}\end{bmatrix} H'_n.
\end{aligned}$$

Recalling the definition of z_{1t} and z_{2t} in (11) and (12), $\sum_{t=1}^{n-k} u_t z_{1t+k} = O_p(\rho^n)$ because $\sum_{t=1}^{n-k} u_t x'_{t+k} = O_p(\rho^n)$ by using an identical method to the bound in (41), and $\sum_{t=1}^{n-k} u_t z_{2t+k} = O_p(n^{1/2})$ because $z_{2n} = 0\ a.s.$ and $\sum_{t=1}^{n-k-1} u_t \zeta'_{n,t+k+1} = O_p(n^{1/2})$ by Lemma 3. The above asymptotic orders yield

$$\hat{R}_{n,k} - R = \frac{1}{n}\left[O_p\left(n\rho^{-n}\right) 1_2, \frac{\sum_{t=1}^{n-k-1} z_{2t+k} u_t}{n^{-1}\sum_{t=1}^{n-k-1} z_{2t-1} z_{2t+k}} + O_p(1) 1_2\right] H'_n, \qquad \textbf{(52)}$$

where $1_2 = (1,1)'$. Since $H_n = O_{a.s.}(1)$ by Lemma 1, the IV estimator becomes

$$\sqrt{n}\left(\hat{R}_{n,k} - R\right) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n-k-1} u_t z_{2t+k} \left(n^{-1} \sum_{t=1}^{n-k-1} z_{2t-1} z_{2t+k}\right)^{-1} \frac{x_n'}{\|x_n\|} \mathcal{R}_{\frac{\pi}{2}}' + O_p\left(\frac{1}{\sqrt{n}}\right)$$

$$= -\frac{\frac{1}{\sqrt{n}} \sum_{t=1}^{n-k-1} u_t \zeta_{n,t+k+1}' \left(\frac{\mathcal{R}_{\frac{\pi}{2}} x_n}{\|x_n\|}\right) \left(\frac{\mathcal{R}_{\frac{\pi}{2}} x_n}{\|x_n\|}\right)'}{\left(\frac{\mathcal{R}_{\frac{\pi}{2}} x_n}{\|x_n\|}\right)' \frac{1}{n} \sum_{t=1}^{n-k-1} \zeta_{n,t} \zeta_{n,t+k+1}' \left(\frac{\mathcal{R}_{\frac{\pi}{2}} x_n}{\|x_n\|}\right)} + O_p\left(\frac{1}{\sqrt{n}}\right). \quad \textbf{(53)}$$

Given an integer-valued sequence $(m_n)_{n\in\mathbb{N}}$ satisfying $m_n \to \infty$ and $m_n/n \to 0$ as $n \to \infty$, define the sequences

$$X_{m_n} = \sum_{j=1}^{m_n} \rho^{-j} u_j \quad \text{and} \quad U_{n,k} = \frac{1}{\sqrt{n}} \sum_{t=2m_n+1}^{n-k-1} \left(\zeta_{n,t+k+1} \otimes u_t\right). \quad \textbf{(54)}$$

By Lemma 1 and direct calculation,

$$X_n - X_{m_n} \to_{a.s.} 0 \quad \text{and} \quad \left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n-k-1} \left(\zeta_{n,t+k+1} \otimes u_t\right) - U_{n,k} \right\|_{L_2}^2 = O\left(\frac{m_n}{n}\right). \quad \textbf{(55)}$$

Lemma 3 then implies that $U_{n,k} \Rightarrow U$ as $n \to \infty$, where U is a $N(0, (\rho^2 - 1)^{-1} \Sigma_u \otimes \Sigma_u)$ random vector. Therefore, using (55) and applying Lemma 2(iv) to the denominator of (53) yields

$$\sqrt{n}\text{vec}\left(\hat{R}_{n,k} - R\right) = -\rho^{k+1}\left(\rho^2 - 1\right) \frac{\left(\frac{\mathcal{R}_{\frac{\pi}{2}} X_{m_n}}{\|X_{m_n}\|}\right) \left(\frac{\mathcal{R}_{\frac{\pi}{2}} X_{m_n}}{\|X_{m_n}\|}\right)' \otimes I_2}{\left(\frac{\mathcal{R}_{\frac{\pi}{2}} X_{m_n}}{\|X_{m_n}\|}\right)' \Sigma_u \left(\frac{\mathcal{R}_{\frac{\pi}{2}} X_{m_n}}{\|X_{m_n}\|}\right)} U_{n,k} + o_p(1)$$

$$\Rightarrow -\rho^{k+1}\left(\rho^2 - 1\right) \frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}_{\frac{\pi}{2}}' \otimes I_2}{X(\rho)' \mathcal{R}_{\frac{\pi}{2}}' \Sigma_u \mathcal{R}_{\frac{\pi}{2}} X(\rho)} U. \quad \textbf{(56)}$$

If $(\|X_{m_n}\| \|U_{n,k}\|)_{n\geq 1}$ is a U.I. sequence, $\mathbb{E}(X(\rho)U') = \lim_{n\to\infty} \mathbb{E}(X_{m_n} U_{n,k}') = 0$, so $X(\rho)$ and U are uncorrelated. It remains to show the required uniform integrability. If (6) is replaced by the stronger condition that $(\|u_t\|^2)_{t\in\mathbb{N}}$ is a U.I. sequence, uniform integrability of the sequence $(\|X_{m_n}\|^2)_{n\geq 1}$ can be established by using an identical argument to the proof of Lemma 0 (i). Since $\mathbb{E}\|U_{n,k}\|^2 \leq c = (\text{tr}\Sigma_u)^2/(\rho - 1)$ for all n, letting $G_{n,\lambda} = \{\|X_{m_n}\| \|U_{n,k}\| > \lambda\}$, $C_{n,\lambda} = \{\|X_{m_n}\| > \lambda^{1/2}\}$ and $D_{n,\lambda} = \{\|U_{n,k}\| > \lambda^{1/2}\}$ the Cauchy Schwarz inequality yields, as $\lambda \to \infty$,

$$\begin{aligned}\sup_{n\geq 1}\left[\mathbb{E}\left(\|X_{m_n}\|\,\|U_{n,k}\|\,\mathbf{1}_{G_{n,\lambda}}\right)\right]^2 &\leq c\sup_{n\geq 1}\mathbb{E}\left(\|X_{m_n}\|^2\mathbf{1}_{G_{n,\lambda}}\right)\\ &\leq c\sup_{n\geq 1}\mathbb{E}\left(\|X_{m_n}\|^2\mathbf{1}_{D_{n,\lambda}}\right)+c\sup_{n\geq 1}\mathbb{E}\left(\|X_{m_n}\|^2\mathbf{1}_{C_{n,\lambda}}\right)\\ &= c\sup_{n\geq 1}\mathbb{E}\left[\|X_{m_n}\|^2\mathbf{1}_{D_{n,\lambda}}\left(\mathbf{1}_{C_{n,\sqrt{\lambda}}}+\mathbf{1}_{C^c_{n,\sqrt{\lambda}}}\right)\right]+o(1)\\ &\leq c\sup_{n\geq 1}\left[\mathbb{E}\left(\|X_{m_n}\|^2\mathbf{1}_{C_{n,\sqrt{\lambda}}}\mathbf{1}_{D_{n,\lambda}}\right)+\lambda^{1/2}P\left(D_{n,\lambda}\right)\right]\\ &\quad +o(1)\\ &\leq c\sup_{n\geq 1}\mathbb{E}\left(\|X_{m_n}\|^2\mathbf{1}_{C_{n,\sqrt{\lambda}}}\right)+\frac{c^2}{\lambda^{1/2}}+o(1)=o(1),\end{aligned}$$

where the bound $\lambda^{1/2}P(D_{n,\lambda})$ applies by noting that $C^c_{n,\sqrt{\lambda}}=\{\|X_{m_n}\|\leq\lambda^{1/4}\}$ and the last line by the Chebyshev inequality and uniform integrability of $(\|X_{m_n}\|^2)_{n\geq 1}$.

If, in addition, (u_t) is an m-dependent sequence, $U_{n,k}$ and X_n in (54) are independent, so U and $X(\rho)$ are independent random vectors, and the limit in (56) has the mixed normal distribution given in (23).

Proof of (24). Using the identity $\hat{u}_{tk}=u_t-(\hat{R}_{n,k}-R)x_{t-1}$ for the residuals, the estimator of Σ_u can be written as

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{n}\hat{u}_{tk}\hat{u}'_{tk}&=\frac{1}{n}\sum_{t=1}^{n}u_tu'_t-\frac{1}{n}\sum_{t=1}^{n}u_tx'_{t-1}\left(\hat{R}_{n,k}-R\right)'-\frac{1}{n}\left(\hat{R}_{n,k}-R\right)\sum_{t=1}^{n}x_{t-1}u'_t\\ &\quad+\frac{1}{n}\left(\hat{R}_{n,k}-R\right)\sum_{t=1}^{n}x_{t-1}x'_{t-1}\left(\hat{R}_{n,k}-R\right)'.\end{aligned}\tag{57}$$

The consistency result in (24) will follow from Lemma 2(i) and

$$\frac{1}{n}\sum_{t=1}^{n}\hat{u}_{tk}\hat{u}'_{tk}=\frac{1}{n}\sum_{t=1}^{n}u_tu'_t+O_p\left(n^{-1/2}\right).\tag{58}$$

To show (58), using the orthogonality of H_n and (52) we obtain

$$\begin{aligned}\left(\hat{R}_{n,k}-R\right)\sum_{t=1}^{n}x_{t-1}u'_t&=\left(\hat{R}_{n,k}-R\right)H_n\sum_{t=1}^{n}z_{t-1}u'_t\\ &=\frac{1}{n}\left[O_p\left(n\rho^{-n}\right)1_2,\ \frac{\sum_{t=1}^{n-k}z_{2t+k}u_t+O_p\left(1_2\right)}{n^{-1}\sum_{t=1}^{n-k}z_{2t-1}z_{2t+k}}\right]\\ &\quad\times\begin{bmatrix}\sum_{t=1}^{n}z_{1t-1}u'_t\\ \sum_{t=1}^{n}z_{2t-1}u'_t\end{bmatrix}\\ &=O_p(1)I_2+\frac{1}{n}\frac{\left(\sum_{t=1}^{n-k}z_{2t+k}u_t\right)\left(\sum_{t=1}^{n}z_{2t-1}u'_t\right)}{n^{-1}\sum_{t=1}^{n-k}z_{2t-1}z_{2t+k}}\\ &=O_p\left(n^{1/2}\right)I_2,\end{aligned}\tag{59}$$

since $\sum_{t=1}^{n} z_{1t-1}u_t' = O_p(\rho^n)$ by PM, $\sum_{t=1}^{n} z_{2t-1}u_t' = O_p(n)$ by (50), and $\sum_{t=1}^{n-k} z_{2t+k}u_t = O_p(n^{1/2})$ by Lemma 3. This shows that the second and third terms of (57) have order $O_p(n^{-1/2})$. For the last term of (57), the identity

$$\sum_{t=1}^{n} x_{t-1}x_{t-1}' = \frac{1}{\rho^2-1}\left\{x_n x_n' - \rho\sum_{t=1}^{n} x_{t-1}u_t' - \rho\sum_{t=1}^{n} u_t x_{t-1}' - \sum_{t=1}^{n} u_t u_t'\right\},$$

the fact that $\hat{R}_{n,k} - R = O_p(n^{-1/2})$, and (59) imply that

$$\left(\hat{R}_{n,k}-R\right)\sum_{t=1}^{n} x_{t-1}x_{t-1}'\left(\hat{R}_{n,k}-R\right)' = \frac{1}{\rho^2-1}\left(\hat{R}_{n,k}-R\right)H_n z_n z_n' H_n' \times\left(\hat{R}_{n,k}-R\right)' + O_p\left(I_2\right).$$

Since $z_n = (\|x_n\|, 0)' = O_{a.s.}(\rho^n)$, (52) yields

$$\left(\hat{R}_{n,k}-R\right)H_n z_n = \frac{1}{n}O_p\left(n\rho^{-n}\right)\|x_n\| 1_2 = O_p(1)\, 1_2.$$

This shows that the last term of (57) has order $O_p(n^{-1})$ and establishes (58).

Proof of (29) when $\Sigma_{u\varepsilon} \neq 0$. Letting $v_t = (u_t, \varepsilon_t')'$, $\mathcal{F}_{v,t} = \sigma(v_t, v_{t-1}, \ldots)$ and replacing u_t by v_t in Assumption 1, we can show that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}\left(\zeta_{n,t+1}\otimes\varepsilon_t\right) \Rightarrow N\left(0, \frac{1}{\rho^2-1}\Sigma_u\otimes\Sigma_\varepsilon\right), \tag{60}$$

where Σ_u and Σ_ε are defined in (30). The proof of (60) is identical to the proof of Lemma 3. Using (44) and proceeding as in the proof of Theorem 3, we obtain

$$\begin{aligned}\sqrt{n}\left(\hat{A}_n - A\right) &= \left(\frac{1}{n}\sum_{t=1}^{n-1} z_{2t}^2\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}\varepsilon_t z_{2t}\right)\frac{x_n'}{\|x_n\|}\mathcal{R}_{\frac{\pi}{2}}' + o_p(1)\\ &= -\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}\varepsilon_t\zeta_{n,t+1}'\right)\frac{\mathcal{R}_{\frac{\pi}{2}}x_n x_n'\mathcal{R}_{\frac{\pi}{2}}'}{x_n'\mathcal{R}_{\frac{\pi}{2}}'\frac{1}{n}\sum_{t=1}^{n-1}\zeta_{n,t+1}\zeta_{n,t+1}'\mathcal{R}_{\frac{\pi}{2}}x_n}\\ &= -(\rho^2-1)\left(\frac{1}{\sqrt{n}}\sum_{t=2m_n+1}^{n-1}\varepsilon_t\zeta_{n,t+1}'\right)\frac{\mathcal{R}_{\frac{\pi}{2}}X_{m_n}X_{m_n}'\mathcal{R}_{\frac{\pi}{2}}'}{X_{m_n}'\mathcal{R}_{\frac{\pi}{2}}'\Sigma_u\mathcal{R}_{\frac{\pi}{2}}X_{m_n}} + o_p(1),\end{aligned}$$

where m_n and X_{m_n} are defined as in the proof of Theorem 3. The proof then follows by (60), and asymptotic mixed normality is ensured by the independence assumption on the sequence v_t.

Proof of Theorem 4. Proceeding as in the proof of Theorem 3,

$$\begin{aligned}\tilde{A}_n - A &= \frac{1}{n}\sum_{t=1}^{n}\varepsilon_t x_{t-1}'\left(\frac{1}{n}\sum_{t=1}^{n}x_{t-1}x_{t-1}'\right)^{-1}\\ &= \left(n^{-1}\sum_{t=1}^{n}z_{2t-1}^2\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t z_{2t-1}\right)\frac{x_n'}{\|x_n\|}\mathcal{R}_{\frac{\pi}{2}}' + o_p(1)\\ &= -\left[I_m + o_p(1)\right]\left(\rho^2-1\right)\left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t\zeta_{n,t}'\right)\frac{\mathcal{R}_{\frac{\pi}{2}}X(\rho)X(\rho)'\mathcal{R}_{\frac{\pi}{2}}'}{X(\rho)'\mathcal{R}_{\frac{\pi}{2}}'\Sigma_u\mathcal{R}_{\frac{\pi}{2}}X(\rho)},\end{aligned} \quad \textbf{(61)}$$

by Lemma 2(ii). Using the identity $\zeta_{n,t} = \rho^{-1}(u_t + \zeta_{n,t+1})$ we obtain

$$\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t\zeta_{n,t}' = \rho^{-1}\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t u_t' + \rho^{-1}\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t\zeta_{n,t+1}' \to_p \Sigma_{\varepsilon u}, \quad \textbf{(62)}$$

since $n^{-1}\sum_{t=1}^{n}\varepsilon_t\zeta_{n,t+1}' \to 0$ in L_1 as in Lemma 2(iii). The result follows by combining (61) and (62).

NOTES

1. The argument given in the proof of equation (26) of PM is incorrect because the sum is not a martingale. However, upon reversion, as discussed in the Remark above, an MGCLT applies and the stated result in PM holds.
2. We refer to Stout (1974, p. 207) for the definition of m-dependence.
3. It is already known (e.g., Fischer, 1993) that dynamic stability conditions can violate weak exogeneity in cointegrated VARs. The present results show that dynamic timing and dynamic instability can induce endogeneity in the regressors of a multivariate comoving system.

REFERENCES

Anderson, T.W. (1959) On asymptotic distributions of estimates of parameters of stochastic difference equations. *Annals of Mathematical Statistics* 30, 676–687.

Cavaliere, G., A. Rahbek, & A.M.R. Taylor (2010) Co-integration rank testing under conditional heteroskedasticity. *Econometric Theory* 26, 1719–1760.

Cavaliere, G., A. Rahbek, & A.M.R. Taylor (2012) Bootstrap determination of the co-integration rank in VAR models. *Econometrica* 80, 1721–1740.

Engle, R.F., D.F. Hendry, & J-F Richard (1983) Exogeneity. *Econometrica* 51, 277–304.

Fanelli, L. (2007) Present value relations, Granger noncausality, and VAR stability. *Econometric Theory* 23, 1254–1260.

Fischer, A.M. (1993) Weak exogeneity and dynamic stability in cointegrated VARs. *Economics Letters* 43, 167–170.

Hall, P. & C.C. Heyde (1980) *Martingale Limit Theory and Its Application*. Academic Press.

Lai, T.L. & C.Z. Wei (1982) Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. *Annals of Statistics* 10, 154–166.

Lai, T.L. & C.Z. Wei (1983) A note on martingale difference sequences satisfying the local Marcinkiewich-Zygmund condition. *Bulletin of the Institute of Mathematics Academia Sinica* 11, 1–13.

Luati, A. & P. Paruolo (2002) Sulla distribuzione di una base di norma unitaria del complemento ortogonale di un vettore gaussiano: Il caso bidimensionale. *Statistica* LXII, 33–38.

Magdalinos, T. & P.C.B Phillips (2009) Limit theory for cointegrated systems with moderately integrated and moderately explosive regressors. *Econometric Theory* 25, 482–526.

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

Park, J.Y. & P.C.B. Phillips (1988) Statistical inference in regressions with integrated processes: Part 1. *Econometric Theory* 4, 468–497.

Park, J.Y. & P.C.B. Phillips (1989) Statistical inference in regressions with integrated processes: Part 2. *Econometric Theory* 5, 95–131.

Phillips, P.C.B. (1984) The exact distribution of LIML: I. *International Economic Review* 25, 249–261.

Phillips, P.C.B. & T. Magdalinos (2007) Limit theory for moderate deviations from a unit root. *Journal of Econometrics* 136, 115–130.

Phillips, P.C.B. & T. Magdalinos (2008) Limit theory for explosively cointegrated systems. *Econometric Theory* 24, 865–887.

Phillips P.C.B., Y. Wu, & J. Yu (2011) Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? *International Economic Review* 52, 201–226.

Phillips P.C.B. & J. Yu (2011) Dating the timeline of financial bubbles during the subprime crisis. *Quantitative Economics* 2, 455–491.

Stout, W.F. (1974) *Almost Sure Convergence*. Academic Press.

Swensen, A.R. (2006) Bootstrap algorithms for testing and determining the cointegration rank in VAR models. *Econometrica* 74, 1699–1714.