# INCONSISTENT VAR REGRESSION WITH COMMON EXPLOSIVE ROOTS

**By**

**Peter C. B. Phillips and Tassos Magdalinos**

**January 2011**

**COWLES FOUNDATION DISCUSSION PAPER NO. 1777**

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

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

# Inconsistent VAR Regression with Common Explosive Roots[1]

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

Tassos Magdalinos
*Granger Centre*
*University of Nottingham*

November 27, 2010

[1]This paper was inspired by reading Nielsen (2009). We thank two referees and the CoEditor for helpful comments on the original version. Magdalinos thanks the Cowles Foundation, Yale University, for hospitality in the Fall of 2009 when the paper was first drafted. Phillips acknowledges partial research support from the NSF under Grant No. SES 09-56687.

**Abstract**

Nielsen (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 co-explosive 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.

*Keywords:* Co-explosive behavior, Common roots, Endogeneity, Forward instrumentation, Geometric multiplicity, Reverse martingale.

*JEL classification:* C22

# 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 focussed on the detection of bubble activity by means of right sided recursive unit root tests (Phillips, Yu and Wu, 2010) and date stamping the origination and termination of this type of phenomenon in the data (Phillips and Yu, 2010). These methods have attracted the attention of empirical researchers interested in bubbles. 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 co-movement to include co-explosive processes (Phillips and Magdalinos [PM], 2008; Magdalinos and Phillips, 2009). These processes are relevant in practical work with data where contagion effects are suspected. Co-explosive processes arise when there are common explosive roots and these lead to an asymptotic singularity in the signal matrix, which produces complications in the limit theory.

In related work, Nielsen (2009, [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 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 $\rho I$ and $\rho > 1$. The common explosive roots produce co-explosive behavior and lead to an asymptotic singularity in the signal matrix, analogous to that studied in Phillips and Magdalinos (2008, [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 co-explosive behavior. In an explosive autoregression the variables behave like exponential trends (with random coefficients) that are informative about the future trajectory. Co-explosive 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.

To establish the limit theory here, a new reverse martingale central limit theorem is proved that is of some independent interest. 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 (2008), 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

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

with $R = \rho I_2$, $\rho > 1$, $x_0 = 0$, and uncorrelated innovations $u_t$ with $E(u_t) = 0$ for all $t$. Note that no martingale difference structure has been imposed on the innovations at this point. Inconsistency of the OLS estimator in (1) applies for more general classes of uncorrelated innovation processes, see Assumption 2 below, and martingale theory will only be used for the derivation of the limit distribution of an IV correction in Section 2.3.

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 co-explosive behavior in the series $x_{1t}$ and $x_{2t}$ that results to 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

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

When $u_t$ is a zero mean uncorrelated sequence with bounded second moments, the infinite series in (4) is shown to converge almost surely in Lemma 1 below. The following assumption ensures that $X(\rho) \neq 0$ almost surely.

**Assumption 1.** *Each random variable in the sequence $(u_t)_{t\in\mathbb{N}}$ admits an absolutely continuous density with respect to Lebesgue measure.*

To treat the limiting singularity that is induced by this co-explosive behavior, we perform a coordinate rotation as developed in PM. Here it is convenient to use the (sample size dependent) orthogonal transformation

$$z_t = H_n' x_t, \tag{5}$$

where

$$H_n = \frac{1}{\|x_n\|}\begin{bmatrix} x_{1n} & -x_{2n} \\ x_{2n} & x_{1n} \end{bmatrix} = \frac{1}{\|x_n\|}\left[x_n, \mathcal{R}_{\frac{\pi}{2}} x_n\right], \tag{6}$$

in which the orthogonal matrix

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

rotates vectors in the plane by an angle $\pi/2$ radians in the positive direction. In view of (5), the transformed variate $z_t$ forms an array, but for notational simplicity the additional subscript is not employed. The large sample behaviour of the random rotation matrix in (6) is characterised by the following lemma, proved in the Appendix.

**Lemma 1.** *Let $u_t$ be a zero mean uncorrelated sequence with $\sup_{t\geq 0} E\|u_t\|^2 < \infty$. Then $\rho^{-n}x_n \to_{a.s.} X(\rho)$. Moreover, under Assumption 1, $X(\rho) \neq 0$ a.s. and*

$$H_n \to_{a.s.} \frac{1}{\|X(\rho)\|}\left[X(\rho), \mathcal{R}_{\frac{\pi}{2}} X(\rho)\right] \quad \text{as } n \to \infty. \tag{7}$$

The transformed regressor variate in (5) may be analysed by combining the identity

$$x_{t-1} = \rho^{-(n-t+1)} x_n - \sum_{j=t}^{n} \rho^{-(j-t+1)} u_j \tag{8}$$

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

$$\begin{aligned} z_{t-1} &= H_n' x_{t-1} = \frac{1}{\|x_n\|}\begin{bmatrix} x_n' x_{t-1} \\ \left(\mathcal{R}_{\frac{\pi}{2}} x_n\right)' x_{t-1} \end{bmatrix} \\ &= \frac{1}{\|x_n\|}\begin{bmatrix} x_n' x_{t-1} \\ -\left(\mathcal{R}_{\frac{\pi}{2}} x_n\right)' \zeta_{n,t} \end{bmatrix} =: \begin{bmatrix} z_{1t-1} \\ z_{2t-1} \end{bmatrix}, \end{aligned} \tag{9}$$

which is conformably partitioned with $x_t$, where

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

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 non-explosive 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 non-explosive 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 (7). 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 $E(u_t | x_{t-1}) = 0$ *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 which takes into account the co-explosive 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 *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 (10), it is apparent that dynamic timing also plays a role in the endogeneity that is manifest in $E\left(u_t | \zeta_{n,t}\right) \neq 0$ since $\zeta_{n,t}$ itself depends on $u_t$. As we shall see, this type of endogeneity can arise even in the triangular (co-explosive) 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.

We now proceed with the asymptotic development, starting with the following assumptions on the innovation sequence $u_t$. We denote by

$$\mathcal{F}_t = \sigma(u_t, u_{t-1}, ...) \quad \text{and} \quad \mathcal{F}^t = \sigma(u_t, u_{t+1}, ...) \tag{11}$$

the natural filtration and the reverse filtration of the innovation sequence and let

$$U_{t,j} = u_t E_{\mathcal{F}_{t+j-1}}\left(u_{t+j}'\right) \quad \text{and} \quad V_{t,j} = E_{\mathcal{F}^{t+1}}(u_t)\, u_{t+j}'.$$

**Assumption 2.** *Let $u_t$ be a zero mean, uncorrelated sequence satisfying*

$$\max_{1\leq t\leq n} E\left(\|u_t\|^2 \mathbf{1}\left\{\|u_t\| > \lambda_n\right\}\right) \to 0 \quad as\ n \to \infty. \tag{12}$$

*for any sequence $(\lambda_n)_{n\in\mathbb{N}}$ such that $\lambda_n \to \infty$, and one of the following conditions: For each $t \geq 1$ and some positive definite matrix $\Sigma_u$*

$$E_{\mathcal{F}_{t-1}}\left(u_t u_t'\right) = \Sigma_u \ a.s. \quad and \quad \max_{j\leq L, t\leq n} \sum_{r=2L}^{n} \left\|E\left(U_{t,j} U_{t+r,j}'\right)\right\| = o\left(n\right) \tag{13}$$

*or*

$$E_{\mathcal{F}^{t+1}}\left(u_t u_t'\right) = \Sigma_u \ a.s. \quad and \quad \max_{j\leq L, t\leq n} \sum_{r=2L}^{n} \left\|E\left(V_{t,j} V_{t+r,j}'\right)\right\| = o\left(n\right) \tag{14}$$

*for any $L \in \mathbb{N}$ such that $L \to \infty$ and $L/n \to 0$.*

Condition (12) is a uniform integrability type of assumption on the sequence $\left\{\|u_t\|^2 : 1 \leq t \leq n\right\}$. Since both $E\|U_{t,j}\|^2$ $E\|V_{t,j}\|^2$ are bounded by $(\mathrm{tr}\Sigma_u)^2$, see (38) in the Appendix, both expectations in (13) and (14) exist. Assumptions (13) and (14) impose a standard constant conditional variance condition on the sequence $u_t$ and an asymptotic weak dependence condition on the sequences $U_{t,j}$ and $V_{t,j}$, respectively. The latter is trivially satisfied when $u_t$ is a martingale difference sequence in (13) or a reverse martingale difference sequence in (14). It also holds for uncorrelated processes $u_t$ that are not martingale differences but satisfy certain asymptotic independence conditions, such as $m$-dependence. Recall that $u_t$ is an $m$-dependent sequence if and only if the sequences of $\sigma$-algebras $\mathcal{F}_t$ and $\mathcal{F}^{t+m}$ are independent for all $t$ and $m \geq 1$. If $u_t$ is an uncorrelated $L$-dependent sequence (i.e. $m$-dependent with $m = L$), then for all $r \geq 2L$ and $j \in \{1, ..., L\}$:

$$\begin{aligned} E\left(U_{t,j} U_{t+r,j}'\right) &= E\left[U_{t,j} E_{\mathcal{F}_{t+j-1}}\left(U_{t+r,j}'\right)\right] = E\left[U_{t,j} E_{\mathcal{F}_{t+j-1}} E_{\mathcal{F}_{t+r+j-1}}\left(u_{t+r+j} u_{t+r}'\right)\right] \\ &= E\left[U_{t,j} E_{\mathcal{F}_{t+j-1}}\left(u_{t+r+j} u_{t+r}'\right)\right] = E\left[U_{t,j} E\left(u_{t+r+j} u_{t+r}'\right)\right] = 0, \end{aligned}$$

by independence for $r-j \geq L$ and uniform boundedness of $\left\|E_{\mathcal{F}_{t+j-1}} E_{\mathcal{F}_{t+r+j-1}}\left(u_{t+r+j} u_{t+r}'\right)\right\|$ in view of (38). Similarly,

$$E\left(V_{t,j} V_{t+r,j}'\right) = E\left[E_{\mathcal{F}^{t+r+1}}\left(u_t u_{t+j}'\right) V_{t+r,j}'\right] = E\left[U_{t,j} E\left(u_{t+r+j} u_{t+r}'\right)\right] = 0,$$

showing that the second part of (13) and (14) apply for any $L$-dependent sequence $u_t$. This asymptotic independence assumption will be employed for the derivation of a mixed normal limit distribution for the IV estimator of Theorem 3.

## 2.2. Least Squares Limit Theory

Co-explosive 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 behaviour 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.** *Let $u_t$ be a zero mean sequence satisfying (12) and either (13) or (14) for some positive definite matrix $\Sigma_u$. 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) $\max_{1\leq j\leq L}\left\| n^{-1}\sum_{t=1}^{n} u_t u_{t+j}'\right\|_{L_1} \to 0$

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

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

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

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

$$\frac{\log\{\lambda_{\max}(X'X)\}}{\lambda_{\min}(X'X)} \to_p 2(\log\rho)\left(\rho^2-1\right)\frac{\|X(\rho)\|^2}{X(\rho)'\mathcal{R}_{\frac{\pi}{2}}'\Sigma_u\mathcal{R}_{\frac{\pi}{2}}X(\rho)} > 0 \;\; a.s. \qquad (15)$$

*where part (v) applies under Assumption 1 and $L$ is defined in Assumption 2.*

In order to obtain a central limit theorem (CLT), the asymptotic orthogonality conditions of (13) and (14) must be replaced by the stronger assumption of martingale and reverse martingale differences with constant conditional variance. In view of the forward filtered nature of $\zeta_{n,t}$, sample covariances of this process and $u_t$, such as $\sum_{t=1}^{n}\zeta_{n,t+1}u_t'$, have a type of reverse martingale structure, which can be exploited to develop a limit theory. The next result gives a new reverse martingale central limit theorem that is useful for such sample covariances. One application of this result is to the CLT stated in equation (26) of PM (2008)[1].

[1] The argument given in the proof of equation (26) of PM (2008) is incorrect because the sum is not a martingale. However, upon reversion, as shown here in the proof of Lemma 3, a MG CLT applies and the stated result holds by Lemma 2.

**Lemma 3.** *Let $u_t$ satisfy (12) and one of the following conditions:*

(i) $E_{\mathcal{F}^{t+1}}(u_t) = 0$ *and* $E_{\mathcal{F}^{t+1}}(u_t u_t') = \Sigma_u$ *a.s. for all* $t$,

(ii) $E_{\mathcal{F}_{t-1}}(u_t) = 0$ *and* $E_{\mathcal{F}_{t-1}}(u_t u_t') = \Sigma_u$ *a.s. for all* $t$

*where* $\mathcal{F}^{t+1}$ *and* $\mathcal{F}_{t-1}$ *are defined in (11). Then, for any fixed* $k \geq 0$, *we have*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left( \zeta_{n,t+k+1} \otimes u_t \right) \Rightarrow N\left(0, \frac{1}{\rho^2 - 1} \Sigma_u \otimes \Sigma_u \right) \quad \text{as } n \to \infty. \tag{16}$$

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

**Theorem 1.** *Under the conditions of Lemma 2 and Assumption 1, the OLS estimator in (1) has the following limit as* $n \to \infty$ :

$$\begin{aligned} \hat{R}_n - R \quad &\to_p \quad -\frac{\rho^2 - 1}{\rho} \frac{\Sigma_u \mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}'_{\frac{\pi}{2}}}{X(\rho)' \mathcal{R}'_{\frac{\pi}{2}} \Sigma_u \mathcal{R}_{\frac{\pi}{2}} X(\rho)} \\ &= \quad -\frac{\rho^2 - 1}{\rho} \frac{\Sigma_u \begin{bmatrix} X_2(\rho)^2 & -X_1(\rho) X_2(\rho) \\ -X_1(\rho) X_2(\rho) & X_1(\rho)^2 \end{bmatrix}}{\sigma_1^2 X_2(\rho)^2 - 2\sigma_{12} X_1(\rho) X_2(\rho) + \sigma_2^2 X_1(\rho)^2}. \end{aligned} \tag{17}$$

**Remarks**

1. The inconsistency of $\hat{R}_n$ is explained by the endogeneity of the regressors discussed earlier. Lai and Wei (1981) 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
$$\frac{\log \{\lambda_{\max}(X'X)\}}{\lambda_{\min}(X'X)} \to_{a.s.} 0 \quad \text{as } n \to \infty. \tag{18}$$
As demonstrated in Lemma 2(vi), the ratio in (18) 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 (1981) 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 on the left side of (18). In particular, (15) and (17) imply that
$$\hat{R}_n - R \to_p -\left( \operatorname*{plim}_{n \to \infty} \frac{\log \lambda_{\max}(X'X)}{\lambda_{\min}(X'X)} \right) \frac{\Sigma_u}{2(\log \rho)\rho} \frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}'_{\frac{\pi}{2}}}{X(\rho)' X(\rho)}. \tag{19}$$

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 $\Sigma_u$ and the common explosive coefficient $\rho$. The limit distribution (17) 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 (17) may be written more simply as

$$-\frac{\rho^2-1}{\rho}\Sigma_u^{1/2}hh'\Sigma_u^{-1/2} \tag{20}$$

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

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

$$x_{1t} = \hat{\rho}x_{1t-1} + \hat{\beta}x_{2t-1} + \hat{u}_{1t}, \tag{21}$$

for various values of $n$ $(= 200, 400, 800)$ against the limit distribution (17) (see also (23) and (24) 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 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 $\sigma_{12} = E(u_{1t}u_{2t}) \neq 0$.

4. The limit random variables corresponding to $\hat{\rho}$ and $\hat{\beta}$ in (21) are given in (17). When $u_t \sim iid\ N(0, \sigma^2 I_2)$, these limits become

$$\begin{aligned} \hat{\rho} - \rho &\to_p -\frac{\rho^2-1}{\rho}\frac{X_2(\rho)^2}{X_2(\rho)^2 + X_1(\rho)^2}, \\ \hat{\beta} - \beta &\to_p \frac{\rho^2-1}{\rho}\frac{X_1(\rho)X_2(\rho)}{X_2(\rho)^2 + X_1(\rho)^2}, \end{aligned} \tag{22}$$

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

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

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

$$pdf_{\hat{\rho}}(y) = \frac{1}{\pi \{(-y)(a_\rho + y)\}^{1/2}}, \quad for\ y \in (-a_\rho, 0) \tag{23}$$

$$pdf_{\hat{\beta}}(y) = \frac{2}{\pi \{a_\rho^2 - 4y^2\}^{1/2}}, \quad for\ |y| < a_\rho/2 \tag{24}$$

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

**Remarks**

1. The supports of the limit distributions (23) and (24) are finite and are determined by $a_\rho$. As $\rho \to 1$, $a_\rho \to 0$ and the supports shrink to the origin, which corresponds to the (well known) consistent estimation of $\rho$ and $\beta$ when $\rho = 1$.

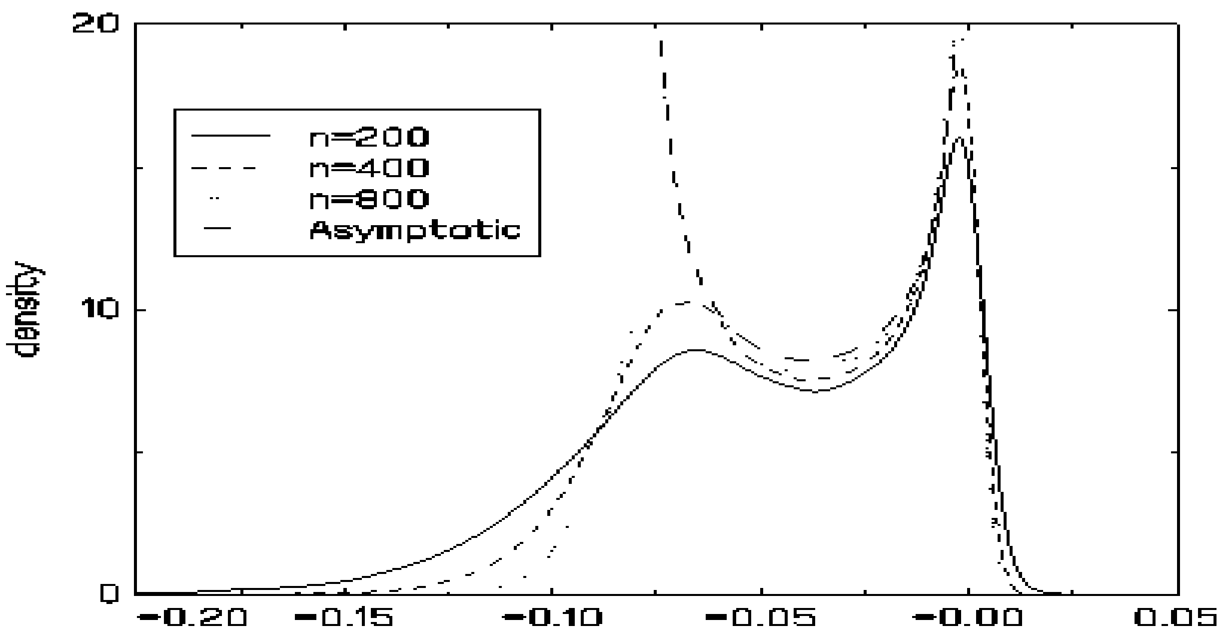


Fig. 1: Finite sample densities of $\hat{\rho} - \rho$ from $R = 80,000$ replications in the fitted model $X_{1t} = \hat{\rho} X_{1t-1} + \hat{\beta} X_{2t-1} + \hat{u}_{1t}$ with $\rho = 1.04$ and $\sigma_{12} = 0$. The limit density has bounded support and is computed from the exact formula (23).

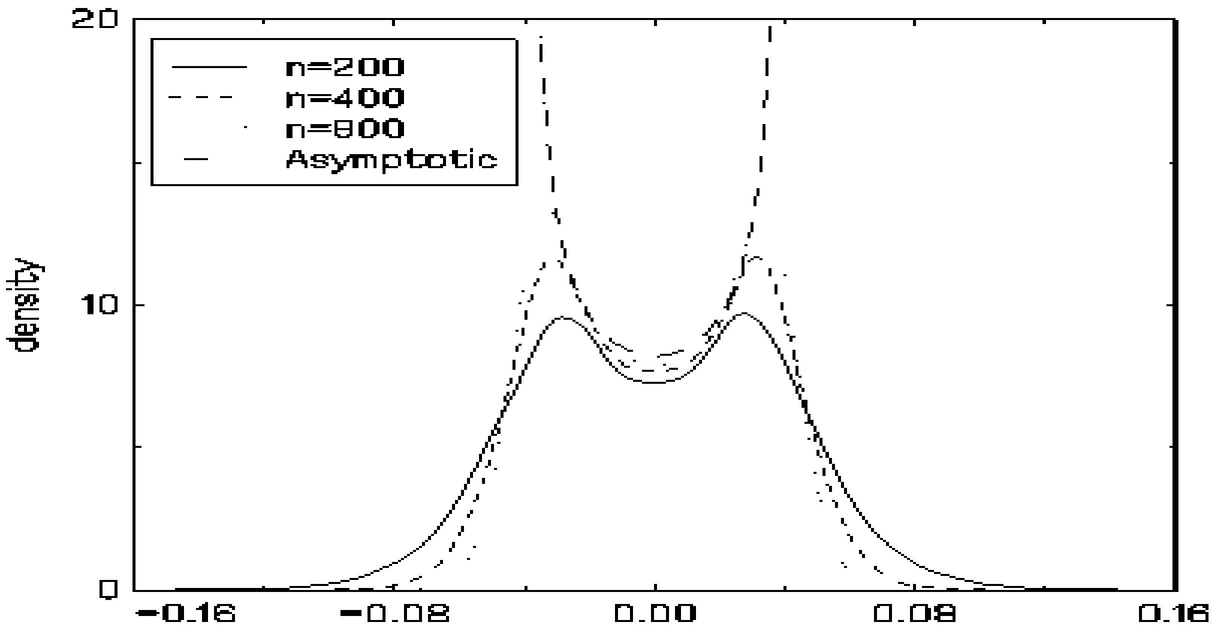


Fig. 2: Finite sample densities of $\hat{\beta}$ from $R = 80,000$ replications in the fitted model $X_{1t} = \hat{\rho}X_{1t-1} + \hat{\beta}X_{2t-1} + \hat{u}_{1t}$ with $\rho = 1.04$ and $\sigma_{12} = 0$. The limit density has bounded support and is computed from the exact formula (24).

2. Figs. 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. 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 co-explosive time series $x_{2t-1}$ in the regression (21). 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 (22) and Fig. 1. As discussed earlier, the inconsistency arises from the endogeneity induced by the co-movement of the regressors and the random nature of the directional vector $X(\rho)$ of the co-movement which depends on the regression error $u_t$.

3. The bimodality in the finite sample distributions shown in Figs. 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 $\left(\hat{\rho} + \hat{\beta}X_2(\rho)/X_1(\rho)\right)$ which serves as the "effective" own lag coefficient in the regression (21). 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

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

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

**Theorem 3.**

(i) *Under the assumptions of Lemma 3 and Assumption 1*

$$\sqrt{n}\mathrm{vec}\left(\hat{R}_{n,k} - R\right) \Rightarrow -\rho^{k+1}\left(\rho^2 - 1\right) \frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}_{\frac{\pi}{2}}' \otimes I_2}{X(\rho)' \mathcal{R}_{\frac{\pi}{2}}' \Sigma_u \mathcal{R}_{\frac{\pi}{2}} X(\rho)} U \tag{25}$$

*as $n \to \infty$, for each fixed $k \geq 0$, where $U$ is a $N\left(0, \left(\rho^2 - 1\right)^{-1} \Sigma_u \otimes \Sigma_u\right)$ random vector. If (12) is replaced by uniform integrability of the sequence $\left(\|u_t\|^2\right)_{t \in \mathbb{N}}$, then $U$ and $X(\rho)$ are uncorrelated random vectors.*

(ii) *If, in addition to the assumptions of Lemma 3, $u_t$ is an $L$-dependent sequence for some $L \in \mathbb{N}$ such that $L \to \infty$ and $L/n \to 0$, $U$ and $X(\rho)$ are independent random vectors and*

$$\sqrt{n}\mathrm{vec}\left(\hat{R}_{n,k} - R\right) \Rightarrow \rho^{k+1}\sqrt{\rho^2 - 1}\, MN\left(0, \frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}_{\frac{\pi}{2}}' \otimes \Sigma_u}{X(\rho)' \mathcal{R}_{\frac{\pi}{2}}' \Sigma_u \mathcal{R}_{\frac{\pi}{2}} X(\rho)}\right). \tag{26}$$

**Remarks**

1. The limit theory (25) relies on the central limit theorem for sample covariance matrices given by (16) 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 (reverse) martingale difference assumption on the innovations $u_t$ is not sufficient for asymptotic mixed normality of the IV estimator $\hat{R}_{n,k}$. However, the limit random vector in (25) will have a mixed normal distribution if asymptotic independence is imposed on the sequence $u_t$.

2. Observe that the limit distribution of $\sqrt{n}\left(\hat{R}_{n,k}-R\right)$ 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

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

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

3. When the sequence $u_t$ is independent, the mixed normal limit (26) facilitates inference, which may be conducted in the usual manner in view of the following arguments. First, from Lemma 1 and Lemma 2(iii) we obtain

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

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 the Appendix,

$$\hat{\Sigma}_{uk}\to_p\Sigma_u, \tag{27}$$

and then

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

giving a consistent estimator of the covariance matrix in (26). Thus, inference about $R$ may be conducted using the standard formula for the variance matrix of $\hat{R}_k$, that is $\left(\sum_{t=1}^{n}x_{t-1}x_{t+k}'\right)^{-1}\otimes\hat{\Sigma}_u$.

4. The variance of the limit distribution (26) 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. Co-explosive Cointegrated Systems

PM (2008) studied a triangular system with possibly co-explosive regressors. A simpler version of this system, which will be sufficient to demonstrate our findings, is given by

$$y_t = Aw_t + \varepsilon_t, \tag{28}$$

$$w_t = x_t, \quad x_t = Rx_{t-1} + u_t \tag{29}$$

$$R = \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix}, \quad \rho > 1, \tag{30}$$

where $A$ is an $m \times 2$ matrix of cointegrating coefficients, $x_t$ is a bivariate vector of co-explosive 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

$$\Sigma = \begin{bmatrix} \Sigma_\varepsilon & 0 \\ 0 & \Sigma_u \end{bmatrix} \tag{31}$$

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

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 (29), i.e. on the precise form of the autoregressive matrix $R$. When $R$ has the form (30), so the regressors are co-explosive, $\hat{A}_n$ 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

$$\begin{aligned} \sqrt{\frac{n}{\rho^2 - 1}} \left( \hat{A}_n - A \right) &\Rightarrow MN\left(0, H_\perp \left(H_\perp' \Sigma_{uu} H_\perp\right)^{-1} H_\perp' \otimes \Sigma_{\varepsilon\varepsilon}\right) \\ &= MN\left(0, \frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}_{\frac{\pi}{2}}'}{X(\rho)' \mathcal{R}_{\frac{\pi}{2}}' \Sigma_{uu} \mathcal{R}_{\frac{\pi}{2}} X(\rho)} \otimes \Sigma_{\varepsilon\varepsilon}\right) \end{aligned} \tag{32}$$

where $H_\perp = \mathcal{R}_{\frac{\pi}{2}} X(\rho) / \|X(\rho)\|$ in the notation of the limiting rotation matrix (7) given earlier. In proving (32), PM assumed that $\Sigma$ has the block diagonal structure (31), so that $x_t$ is uncorrelated with $\varepsilon_t$. However, as shown in the Appendix, (32) continues to hold when the covariance structure is given by

$$\Sigma = \begin{bmatrix} \Sigma_\varepsilon & \Sigma_{\varepsilon u} \\ \Sigma_{u\varepsilon} & \Sigma_u \end{bmatrix}, \quad \text{with } \Sigma_{u\varepsilon} \neq 0. \tag{33}$$

From this result, it would seem that co-explosive behavior in the regressors does not cause an inconsistency, contrary to the VAR regression result (17) in Theorem 3. However, suppose that $w_t = x_{t-1}$ in (29), 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 co-explosive time series, the impact of dynamic timing is considerable. Let the corresponding least squares estimator of $A$, when $w_t = x_{t-1}$, be $\tilde{A}_n = \left(\sum_{t=1}^n y_t x_{t-1}'\right)\left(\sum_{t=1}^n x_{t-1}x_{t-1}'\right)^{-1}$.

**Theorem 4.** *In the model (28)-(30) with $w_t = x_{t-1}$ and $\Sigma$ is given by (33)*

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

Evidently, when there are co-explosive 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 in regression with explosive regressors can depend subtly on dynamic timing and in a manner quite different from stationary systems. Under the condition $(\varepsilon_t, u_t) \sim iid(0, \Sigma)$, convention would dictate that $x_{t-1}$ is weakly exogenous for $A$ in the system $y_t = Ax_{t-1} + \varepsilon_t$, but jointly dependent and correlated with $\varepsilon_t$ in the system $y_t = Ax_t + \varepsilon_t$. Curiously, however, in the presence of co-explosive 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 regression. In particular, the limiting singularity in the sample moment matrix that is caused by co-explosive 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 co-explosive behavior, the stationary component of the transformed regressor, which is forward looking and depends on $u_t$, is correlated with $\varepsilon_t$ when $\Sigma_{u\varepsilon} \neq 0$.

# 4. Conclusions

Besides the intriguing nature of the inconsistency in co-explosive VARs and structural systems, the limit distributions of the least squares estimates 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 centred (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

**Lemma 0.** *Condition (12) implies that the sequence* $\left\{ \left\| \zeta_{n,t+k} \right\|^2, 1 \le t \le n \right\}$ *is uniformly integrable:*

$$\max_{1 \le t \le n} E\left( \left\| \zeta_{n,t+k} \right\|^2 \mathbf{1}\left\{ \left\| \zeta_{n,t+k} \right\| > \lambda_n \right\} \right) \to 0 \quad \textit{as } n \to \infty \tag{34}$$

*for any sequence* $\lambda_n \to \infty$ *and any fixed* $k \ge 0$.

**Proof.** Denote $A_{n,t} = \left\{ \left\| \zeta_{n,t+k} \right\| > \lambda_n \right\}$. Since, $\left\| \zeta_{n,t+k} \right\| \le \sum_{j=0}^{n-t-k} \rho^{-(j+1)} \left\| u_{t+k+j} \right\|$, we obtain, for each $t \in \{1, ..., n\}$

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

Since the last series is convergent, the above bound shows that

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

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

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

for any sequence $\lambda_n \to \infty$ as $n \to \infty$, by U.I. of the sequence $\left\| u_s \right\|^2$ and the Chebyshev inequality. This establishes (34).

**Proof of Lemma 1.** We prove $a.s.$ convergence of the infinite series in (4) by applying the Rademacher-Menchoff convergence theorem for orthogonal random variables

(Theorem 2.3.2 of Stout, 1974). Since $(\rho^{-t}u_t)_{t\geq 1}$ is an orthogonal sequence of random vectors with

$$\sum_{t=1}^{\infty}\left(\log^2 t\right) E\left\|\rho^{-t}u_t\right\|^2 \leq \sup_{j\geq 1} E\left\|u_j\right\|^2 \sum_{t=1}^{\infty} t^2\rho^{-2t} < \infty,$$

$\sum_{t=1}^{n}\rho^{-t}u_t$ converges $a.s.$ as $n\to\infty$. Therefore, the infinite series $X(\rho)$ in (4) exists as the $a.s.$ limit of $\sum_{t=1}^{n}\rho^{-t}u_t = \rho^{-n}x_n$ and is non-zero almost surely by Assumption 1. Almost sure convergence of $H_n$ follows by applying $\rho^{-n}x_n \rightarrow_{a.s.} X(\rho)$ to (6) and using continuity of norms.

**Proof of Lemma 2.** Let $j\geq 0$. Denoting by $\mathcal{H}_{t,j}$ either $\mathcal{F}_{t+j-1}$ or $\mathcal{F}^{t+1}$,

$$\frac{1}{n}\sum_{t=1}^{n} u_t u_{t+j}' = \frac{1}{n}\sum_{t=1}^{n}\left[u_t u_{t+j}' - E_{\mathcal{H}_{t,j}}\left(u_t u_{t+j}'\right)\right] + \frac{1}{n}\sum_{t=1}^{n} E_{\mathcal{H}_{t,j}}\left(u_t u_{t+j}'\right). \tag{35}$$

We will show that uniform integrability (U.I.) of the sequence $\|u_t\|^2$ implies that the first term on the right of (35) converges to 0 in $L_1$ norm. Choose a sequence $(k_n)_{n\in\mathbb{N}}$ such that $k_n\to\infty$ and $k_n/n^{1/2}\to 0$ as $n\to\infty$ and define the random matrices

$$\begin{aligned} v_{t,j} &= u_t u_{t+j}' - E_{\mathcal{H}_{t,j}}\left(u_t u_{t+j}'\right) \\ v_{t,j}^{(n)} &= u_t u_{t+j}'\mathbf{1}\{\|u_t\|\leq k_n\} - E_{\mathcal{H}_{t,j}}\left(u_t u_{t+j}'\mathbf{1}\{\|u_t\|\leq k_n\}\right). \end{aligned}$$

It is easy to see that $\left\{v_{t,j}^{(n)} : 1\leq t\leq n\right\}$ is a zero mean uncorrelated sequence: when $\mathcal{H}_{t,j} = \mathcal{F}_{t+j-1}$, $v_{t,j}^{(n)}$ is a (matrix valued) $\mathcal{F}_{t+j}$ martingale difference array; when $\mathcal{H}_{t,j} = \mathcal{F}^{t+1}$, $v_{t,j}^{(n)}$ is a reverse martingale difference array in the sense that it is $\mathcal{F}^t$-adapted and $E_{\mathcal{F}^{t+1}}v_{t,j}^{(n)} = 0$ for all $t$, so uncorrelatedness follows from the law of iterated expectations as in the forward martingale difference case. The identity $1-\mathbf{1}\{\|u_t\|\leq k_n\} = \mathbf{1}\{\|u_t\|>k_n\}$ yields

$$\begin{aligned} \left\|\frac{1}{n}\sum_{t=1}^{n}\left(v_{t,j}-v_{t,j}^{(n)}\right)\right\|_{L_1} &= \frac{1}{n}\sum_{t=1}^{n}\left\|u_t u_{t+j}'\mathbf{1}\{\|u_t\|>k_n\} - E_{\mathcal{H}_{t,j}}\left(u_t u_{t+j}'\mathbf{1}\{\|u_t\|>k_n\}\right)\right\|_{L_1} \\ &\leq 2\frac{1}{n}\sum_{t=1}^{n} E\left(\|u_t\|\,\|u_{t+j}\|\,\mathbf{1}\{\|u_t\|>k_n\}\right) \\ &\leq 2\frac{1}{n}\sum_{t=1}^{n}\left[E\left(\|u_t\|^2\mathbf{1}\{\|u_t\|>k_n\}\right)E\left(\|u_{t+j}\|^2\right)\right]^{1/2} \\ &\leq 2\left[\operatorname{tr}(\Sigma_u)\max_{1\leq t\leq n} E\left(\|u_t\|^2\mathbf{1}\{\|u_t\|>k_n\}\right)\right]^{1/2}\to 0 \end{aligned} \tag{36}$$

as $n \to \infty$ by U.I. of $\|u_t\|^2$. Also, since $\left\{v_{t,j}^{(n)} : 1 \le t \le n\right\}$ is an uncorrelated sequence, the Jensen inequality for conditional expectations yields

$$\begin{aligned} E\left\|\frac{1}{n}\sum_{t=1}^{n} v_{t,j}^{(n)}\right\|^2 &= \frac{1}{n^2}\sum_{t=1}^{n} E\left\|v_{t,j}^{(n)}\right\|^2 \le \frac{2}{n^2}\sum_{t=1}^{n} E\left[\|u_t\|^2\|u_{t+j}\|^2 \mathbf{1}\{\|u_t\| \le k_n\}\right] \\ &+\frac{2}{n^2}\sum_{t=1}^{n} E\left[E_{\mathcal{H}_{t,j}}\left(\|u_t\|\|u_{t+j}\| \mathbf{1}\{\|u_t\| \le k_n\}\right)\right]^2 \\ &\le \frac{4}{n^2}\sum_{t=1}^{n} E\left[\|u_t\|^2\|u_{t+j}\|^2 \mathbf{1}\{\|u_t\| \le k_n\}\right] \\ &\le 4\mathrm{tr}\left(\Sigma_u\right)\frac{k_n^2}{n} \to 0 \end{aligned}$$

since $k_n = o\left(n^{1/2}\right)$. Hence, $n^{-1}\sum_{t=1}^{n} v_{t,j}^{(n)} \to 0$ in $L_2$ and, in view of (36), $n^{-1}\sum_{t=1}^{n} v_{t,j} \to 0$ in $L_1$. Moreover, this convergence applies uniformly for $j \ge 0$, so (35) yields

$$\sup_{j\ge 0}\left\|\frac{1}{n}\sum_{t=1}^{n} u_t u_{t+j}' - \frac{1}{n}\sum_{t=1}^{n} E_{\mathcal{H}_{t,j}}\left(u_t u_{t+j}'\right)\right\|_{L_1} \to 0 \quad \text{as } n \to \infty. \tag{37}$$

Part (i) of the lemma follows immediately by (37) since $E_{\mathcal{H}_{t,j}}\left(u_t u_t'\right) = \Sigma_u$ $a.s.$ for all $t$ by (13) and (14).

For part (ii), we will show that the second term of (37), which is equal to either $n^{-1}\sum_{t=1}^{n} U_{t,j}$ or $n^{-1}\sum_{t=1}^{n} V_{t,j}$ according to whether $\mathcal{H}_{t,j}$ equals $\mathcal{F}_{t+j-1}$ or $\mathcal{F}^{t+1}$, converges to 0 in $L_2$ uniformly for $j \in \{1, ..., L\}$. First note that

$$\begin{aligned} E\|U_{t,j}\|^2 &\le E\|u_t\|^2 E\left(\left\|E_{\mathcal{F}_{t+j-1}}\left(u_{t+j}\right)\right\|^2\right) \\ &\le E\|u_t\|^2 E\left(E_{\mathcal{F}_{t+j-1}}\|u_{t+j}\|^2\right) = \left(\mathrm{tr}\Sigma_u\right)^2 \end{aligned} \tag{38}$$

by the Cauchy-Schwarz inequality followed by the Jensen inequality for conditional expectations and the first part of (13). An identical argument shows that $E\|V_{t,j}\|^2 \le \left(\mathrm{tr}\Sigma_u\right)^2$. Therefore, for any $L \in \mathbb{N}$ such that $L \to \infty$ $L/n \to 0$,

$$\begin{aligned} E\left\|\frac{1}{n}\sum_{t=1}^{n} U_{t,j}\right\|^2 &= \frac{1}{n^2}\sum_{t=1}^{n} E\|U_{t,j}\|^2 + 2\frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=t+1}^{n} E\left(U_{t,j}U_{s,j}'\right) \\ &= \frac{2}{n^2}\sum_{t=1}^{n}\left[\sum_{s=t+2L+1}^{n} EU_{t,j}U_{s,j}' + \sum_{s=t+1}^{t+2L} EU_{t,j}U_{s,j}'\right] + O\left(n^{-1}\right) \\ &\le \frac{2}{n^2}\sum_{t=1}^{n}\left(\sum_{s=2L+1}^{n-t}\left\|EU_{t,j}U_{t+s,j}'\right\| + 2L\mathrm{tr}\left(\Sigma_u\right)\right) + O\left(n^{-1}\right) \\ &\le \frac{2}{n}\max_{j\le L}\max_{t\le n}\sum_{s=2L+1}^{n}\left\|EU_{t,j}U_{t+s,j}'\right\| + O\left(\frac{L}{n}\right) = o(1) \end{aligned}$$

as $n \to \infty$ by (13). An identical argument shows that (14) implies that $n^{-1} \sum_{t=1}^{n} V_{t,j} \to_{L_2} 0$ uniformly for $j \in \{1, ..., L\}$. This shows part (ii).

For part (iii), (10) and part (ii) yield, for any fixed $k \geq 1$,

$$\begin{aligned}
\left\| \frac{1}{n} \sum_{t=1}^{n} u_t \zeta'_{n,t+k} \right\|_{L_1} &= \left\| \frac{1}{n} \sum_{t=1}^{n} u_t \sum_{j=0}^{n-t-k} \rho^{-(j+1)} u'_{t+k+j} \right\|_{L_1} \\
&= \left\| \frac{1}{n} \sum_{t=1}^{n} u_t \sum_{j=0}^{L-k} \rho^{-(j+1)} u'_{t+k+j} \right\|_{L_1} + O\left(\rho^{-L}\right) \\
&\leq \max_{1 \leq l \leq L} \left\| \frac{1}{n} \sum_{t=1}^{n} u_t u'_{t+l} \right\|_{L_1} \sum_{j=0}^{\infty} \rho^{-(j+1)} = o(1),
\end{aligned}$$

for any integer $L$ for which part (ii) applies. For part (iv), the definition of $\zeta_{n,t}$ in (10) yields the following identities:

$$\zeta_{n,t} = \rho^{-k} \zeta_{n,t+k} + \sum_{j=0}^{k-1} \rho^{-(j+1)} u_{t+j} \quad \text{and} \quad \zeta_{n,t+k} = \rho^{-1} \zeta_{n,t+k+1} + \rho^{-1} u_{t+k} \tag{39}$$

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

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

the first identity in (39) implies that

$$\left\| \frac{1}{n} \sum_{t=1}^{n} \zeta_{n,t} \zeta'_{n,t+k} - \rho^{-k} \frac{1}{n} \sum_{t=1}^{n} \zeta_{n,t+k} \zeta'_{n,t+k} \right\|_{L_1} \to 0 \quad \text{as } n \to \infty \tag{40}$$

for all fixed $k \geq 0$. The second identity in (39) 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 $t \in \{1, ..., n\}$ yields

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

where the remainder term

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

tends to 0 in $L_1$ by part (iii). Therefore,

$$\left\| \frac{1}{n} \sum_{t=1}^{n} \zeta_{n,t+k} \zeta_{n,t+k}' - \frac{1}{\rho^2 - 1} \frac{1}{n} \sum_{t=1}^{n} u_{t+k} u_{t+k}' \right\|_{L_1} \to 0 \quad \text{as } n \to \infty \tag{41}$$

for all fixed $k \geq 0$. Part (iv) of the lemma follows by combining (40), (41) and the law of large numbers for $n^{-1} \sum_{t=1}^{n} u_t u_t'$ of part (i).

For part (v), using (9) to expand $z_{1t-1}$ and $z_{2t-1}$, we obtain the following rates of convergence for the elements of the matrix $\sum_{t=1}^{n} z_{t-1} z_{t+k}'$ for each fixed $k \geq -1$:

$$\sum_{t=1}^{n} z_{1t-1} z_{1t+k} = O_p\left(\rho^{2n}\right) \tag{42}$$

$$\max \left\{ \sum_{t=1}^{n} z_{1t-1} z_{2t+k}, \sum_{t=1}^{n} z_{2t-1} z_{1t+k} \right\} = O_p\left(\rho^{n}\right). \tag{43}$$

The first part of (43) can be deduced by (9) and the following bound:

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

where $C = \text{tr}\Sigma_u / \sqrt{\rho^2 - 1}$. Since $k$ is fixed, the second part of (43) can be deduced by an identical argument. For (42), 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

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

where the last order of magnitude is obtained by the same argument used to prove (43). Now (42) follows since, by direct computation of $\sum_{t=1}^{n} E \left\| x_{t-1} \right\|^2$. For the remaining element, using (9) and part (iv), we obtain as $n \to \infty$

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^{n} z_{2t-1} z_{2t+k} &= \left( \frac{\mathcal{R}_{\frac{\pi}{2}} x_n}{\|x_n\|} \right)' \frac{1}{n} \sum_{t=1}^{n} \zeta_{n,t} \zeta_{n,t+k+1}' \left( \frac{\mathcal{R}_{\frac{\pi}{2}} x_n}{\|x_n\|} \right) \\ &\to_p \frac{\rho^{-k-1}}{\rho^2 - 1} \left( \mathcal{R}_{\frac{\pi}{2}} \frac{X(\rho)}{\|X(\rho)\|} \right)' \Sigma_u \left( \mathcal{R}_{\frac{\pi}{2}} \frac{X(\rho)}{\|X(\rho)\|} \right). \end{aligned} \tag{44}$$

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

$$\begin{aligned} D_n &= \left(\sum_{t=1}^{n} z_{1t-1}z_{1t+k}\right)\left(\sum_{t=1}^{n} z_{2t-1}z_{2t+k}\right) - \left(\sum_{t=1}^{n} z_{2t-1}z_{1t+k}\right)\left(\sum_{t=1}^{n} z_{1t-1}z_{2t+k}\right) \\ &= \left(\sum_{t=1}^{n} z_{1t-1}z_{1t+k}\right)\left(\sum_{t=1}^{n} z_{2t-1}z_{2t+k}\right)\left\{1 + O_p\left(n^{-1}\right)\right\}. \qquad (45) \end{aligned}$$

Using (42)-(45), the inverse of the signal matrix is given by

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

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

$$\left(\frac{1}{n}\sum_{t=1}^{n} z_{t-1}z_{t+k}'\right)^{-1} = \left[1 + O_p\left(\frac{1}{n}\right)\right]\begin{bmatrix} O_p\left(n\rho^{-2n}\right) & O_p\left(\rho^{-n}\right) \\ O_p\left(\rho^{-n}\right) & \left(\frac{1}{n}\sum_{t=1}^{n} z_{2t-1}z_{2t+k}\right)^{-1} \end{bmatrix} \qquad (46)$$

and the result follows from (44).

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

$$\begin{aligned} \sum_{t=1}^{n} x_{it-1}^2 &= \frac{1}{\rho^2 - 1}\left(x_{in}^2 - 2\rho\sum_{t=1}^{n} x_{it-1}u_{it} - \sum_{t=1}^{n} u_{it}^2\right) \\ \sum_{t=1}^{n} x_{1t-1}x_{2t-1} &= \frac{1}{\rho^2 - 1}\left(x_{1n}x_{2n} - \rho\sum_{t=1}^{n} x_{1t-1}u_{2t} - \rho\sum_{t=1}^{n} x_{2t-1}u_{1t} - \sum_{t=1}^{n} u_{1t}u_{2t}\right) \\ \sum_{t=1}^{n} x_{it-1}u_{jt} &= x_{in}\sum_{t=1}^{n}\rho^{-(n-t+1)}u_{jt} - \rho^{-1}\sum_{t=1}^{n} u_{it}u_{jt} - \sum_{t=1}^{n}\left(\sum_{s=t+1}^{n}\rho^{t-1-s}u_{is}\right)u_{jt} \end{aligned}$$

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

$$\begin{aligned} \frac{\rho^{-2n}}{n}\det\left(X'X\right) &= \frac{1}{\left(\rho^2 - 1\right)^2}\frac{\rho^{-2n}}{n}\left(x_{1n}^2\sum_{t=1}^{n} u_{2t}^2 + x_{2n}^2\sum_{t=1}^{n} u_{1t}^2 - 2x_{1n}x_{2n}\sum_{t=1}^{n} u_{1t}u_{2t}\right) \\ &\quad + O_p\left(\frac{1}{\sqrt{n}}\right) \\ &\to_p \frac{1}{\left(\rho^2 - 1\right)^2}\left[X_1(\rho)^2\sigma_2^2 + X_2(\rho)^2\sigma_1^2 - 2X_1(\rho)X_2(\rho)\sigma_{12}\right] \qquad (47) \end{aligned}$$

$$\rho^{-2n} tr\left(X'X\right) = \frac{\rho^{-2n}}{\rho^2-1}\left(x_{1n}^2 + x_{2n}^2\right) + O_p\left(\rho^{-n}\right) \to_p \frac{1}{\rho^2-1}\left[X_1\left(\rho\right)^2 + X_2\left(\rho\right)^2\right]. \quad (48)$$

The asymptotic behaviour of the eigenvalues of $X'X$ can be obtained from (47) and (48):

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

$$\begin{aligned}
\lambda_{\min}\left(X'X\right) &= \frac{1}{2}\left[tr\left(X'X\right) - \sqrt{\left(trX'X\right)^2 - 4\det\left(X'X\right)}\right] \\
&= \frac{\det\left(X'X\right)}{tr(X'X)} + O_p\left[\frac{\left(\det\left(X'X\right)\right)^2}{\left(trX'X\right)^3}\right].
\end{aligned}$$

Combining the above expressions and noting from (48) that $tr\left(\rho^{-2n}X'X\right) \to_p \left\|X\left(\rho\right)\right\|^2 / \left(\rho^2-1\right) > 0$ $a.s.$ we obtain

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

by (47) and (48), as required. Almost sure positivity of the above probability limit is ensured since $\rho > 1$, $\left\|X\left(\rho\right)\right\| > 0$ $a.s.$ 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.** We first give the proof under condition (i) of the lemma. For each fixed $k \geq 0$, let

$$S_{n,\tau} = \sum_{t=\tau}^{n}\xi_{n,t} \qquad \xi_{n,t} = \frac{1}{\sqrt{n}}\zeta_{n,t+k+1}\otimes u_t.$$

and $\mathcal{F}^t$ be the reverse filtration defined in (11). Then $\left(S_{n,\tau}, \mathcal{F}^\tau, 1 \leq \tau \leq n\right)$ is a reverse martingale array that can be reversed into a martingale array $\left(M_{n,\tau}, \mathcal{G}_{n,\tau}, 1 \leq \tau \leq n\right)$ by letting

$$M_{n,\tau} = S_{n,n-\tau} \quad \text{and} \quad \mathcal{G}_{n,\tau} = \mathcal{F}^{n-\tau},$$

since $\mathcal{G}_{n,\tau} = \sigma(u_{n-\tau}, u_{n-\tau+1}, ...)$ is now a (forward) filtration with respect to $\tau$ and $\zeta_{n,n-\tau+k+1}$ is $\mathcal{G}_{n,\tau-1}$-measurable for all $k \geq 0$. The identities

$$\sum_{\tau=0}^{n-1} \Delta M_{n,\tau} = \sum_{\tau=0}^{n-1} \xi_{n,n-\tau} = \sum_{t=1}^{n} \xi_{n,t} = S_{n,1} \tag{49}$$

then imply that the limit distribution of $S_{n,1}$ can be derived by a standard martingale CLT on $\sum_{\tau=0}^{n-1} \Delta M_{n,\tau}$, e.g. Corollary 3.1 of Hall and Heyde (1980). The conditional variance of $M_{n,\tau}$ is given by

$$\begin{aligned}
\sum_{\tau=0}^{n-1} E_{\mathcal{G}_{n,\tau-1}}\left(\xi_{n,n-\tau}\xi_{n,n-\tau}'\right) &= \frac{1}{n}\sum_{\tau=0}^{n-1} \zeta_{n,n-\tau+k+1}\zeta_{n,n-\tau+k+1}' \otimes E_{\mathcal{F}^{n-\tau+1}}\left(u_{n-\tau}u_{n-\tau}'\right) \\
&= \frac{1}{n}\sum_{j=1}^{n} \zeta_{j+k+1}\zeta_{j+k+1}' \otimes \Sigma_u \\
&\to_p \frac{1}{\rho^2-1}\Sigma_u \otimes \Sigma_u,
\end{aligned}$$

by Lemma 2 (ii) since $k$ is fixed.

To establish the Lindeberg condition, we employ Lemma 0. Letting $t = n - \tau$ in the index of summation of (49), and noting that $\mathcal{G}_{n,n-t-1} = \mathcal{F}^{t+1}$, the Lindeberg condition is equivalent to

$$\sum_{t=1}^{n} E_{\mathcal{F}^{t+1}}\left(\left\|\xi_{n,t}\right\|^2 \mathbf{1}_{B_{n,t}(\delta)}\right) = \frac{1}{n}\sum_{t=1}^{n}\left\|\zeta_{n,t+k+1}\right\|^2 E_{\mathcal{F}^{t+1}}\left(\|u_t\|^2 \mathbf{1}_{B_{n,t}(\delta)}\right) \to_p 0 \tag{50}$$

for any $\delta > 0$, where the events $B_{n,t}(\delta)$ are defined by

$$\begin{aligned}
B_{n,t}(\delta) &= \left\{\left\|\xi_{n,t}\right\| > \delta\right\} = \left\{\left\|\zeta_{n,t+k+1}\right\| \|u_t\| > \sqrt{n}\delta\right\} \\
&\subseteq \left\{\left\|\zeta_{n,t+k+1}\right\| > n^{1/4}\sqrt{\delta}\right\} \cup \left\{\|u_t\| > n^{1/4}\sqrt{\delta}\right\}.
\end{aligned}$$

Hence, the right side of (50) is bounded by $I_1(n) + I_2(n)$, where

$$\begin{aligned}
I_1(n) &= \frac{1}{n}\sum_{t=1}^{n}\left\|\zeta_{n,t+k+1}\right\|^2 \mathbf{1}\left\{\left\|\zeta_{n,t+k+1}\right\| > n^{1/4}\sqrt{\delta}\right\} E_{\mathcal{F}^{t+1}}\left(\|u_t\|^2\right) \\
I_2(n) &= \frac{1}{n}\sum_{t=1}^{n}\left\|\zeta_{n,t+k+1}\right\|^2 E_{\mathcal{F}^{t+1}}\left(\|u_t\|^2 \mathbf{1}\left\{\|u_t\| > n^{1/4}\sqrt{\delta}\right\}\right).
\end{aligned}$$

Since $E_{\mathcal{F}^{t+1}}\left(\|u_t\|^2\right) = \mathrm{tr}(\Sigma_u)$ for all $t$, $I_1(n) \to_{L_1} 0$ as $n \to \infty$ by (34). To show that $I_2(n) \to_{L_1} 0$, let

$$\psi_{n,\delta} = \max_{1\leq t\leq n} E\left(\|u_t\|^2 \mathbf{1}\left\{\|u_t\| > n^{1/4}\sqrt{\delta}\right\}\right).$$

By (12) we know that $\psi_{n,\delta} \to 0$ as $n \to \infty$ for any $\delta > 0$. Using the identity $1 = \mathbf{1}\left\{\left\|\zeta_{n,t+k+1}\right\|^2 > \psi_{n,\delta}^{-1/2}\right\} + \mathbf{1}\left\{\left\|\zeta_{n,t+k+1}\right\|^2 \leq \psi_{n,\delta}^{-1/2}\right\}$ we obtain

$$\begin{aligned} I_2(n) \leq\ & \frac{1}{n}\sum_{t=1}^{n}\left\|\zeta_{n,t+k+1}\right\|^2 \mathbf{1}\left\{\left\|\zeta_{n,t+k+1}\right\| > \psi_{n,\delta}^{-1/2}\right\} E_{\mathcal{F}^{t+1}}\left(\|u_t\|^2\right) \\ & + \psi_{n,\delta}^{-1/2}\frac{1}{n}\sum_{t=1}^{n} E_{\mathcal{F}^{t+1}}\left(\|u_t\|^2 \mathbf{1}\left\{\|u_t\| > n^{1/4}\sqrt{\delta}\right\}\right). \end{aligned}$$

Since, for any $\delta > 0$, $\psi_{n,\delta}^{-1/2} \to \infty$ as $n \to \infty$ and $E_{\mathcal{F}^{t+1}}\left(\|u_t\|^2\right)$ is constant, taking expectations and using (34) and (12) we obtain

$$\begin{aligned} \left\|I_2(n)\right\|_{L_1} \leq\ & \psi_{n,\delta}^{-1/2}\frac{1}{n}\sum_{t=1}^{n} E\left(\|u_t\|^2 \mathbf{1}\left\{\|u_t\| > n^{1/4}\sqrt{\delta}\right\}\right) + o(1) \\ \leq\ & \psi_{n,\delta}^{-1/2}\max_{1\leq t\leq n} E\left(\|u_t\|^2 \mathbf{1}\left\{\|u_t\| > n^{1/4}\sqrt{\delta}\right\}\right) + o(1) \\ =\ & \psi_{n,\delta}^{1/2} + o(1) = o(1). \end{aligned}$$

Under condition (ii) of the lemma, $u_t$ is a martingale difference sequence and the above argument does not apply in general. However, $S_{n,1}$ can be approximated by a martingale array as follows: Let $(\kappa_n)_{n\in\mathbb{N}}$ be an integer valued sequence such that $\kappa_n \to \infty$ and $\kappa_n/\sqrt{n} \to 0$ as $n \to \infty$. Then, changing the order of summation,

$$\begin{aligned} S_{n,1} =\ & \frac{\rho^k}{\sqrt{n}}\sum_{t=1}^{n}\sum_{j=t+k+1}^{n}\rho^{-(j-t)}\left(u_j \otimes u_t\right) = \frac{\rho^k}{\sqrt{n}}\sum_{j=k+2}^{n}\sum_{t=1}^{j-k-1}\rho^{-(j-t)}\left(u_j \otimes u_t\right) \\ =\ & \frac{\rho^k}{\sqrt{n}}\sum_{j=k+2}^{n}\sum_{i=k+1}^{j-1}\rho^{-i}\left(u_j \otimes u_{j-i}\right) = \frac{\rho^k}{\sqrt{n}}\sum_{j=\kappa_n}^{n}\sum_{i=k+1}^{j-1}\rho^{-i}\left(u_j \otimes u_{j-i}\right) + O_p\left(\frac{\kappa_n}{\sqrt{n}}\right) \\ =\ & \rho^k\frac{1}{\sqrt{n}}\sum_{j=\kappa_n}^{n}\left(u_j \otimes \sum_{i=k+1}^{\kappa_n}\rho^{-i}u_{j-i}\right) + O_p\left(\rho^{-\kappa_n}n^{1/2}\right), \end{aligned} \tag{51}$$

where all approximations apply in $L_1$ norm. Clearly, $\eta_{n,j} = n^{-1/2}u_j \otimes \sum_{i=k+1}^{\kappa_n}\rho^{-i}u_{j-i}$

is an $\mathcal{F}_j$ martingale difference array with

$$\begin{aligned}
\sum_{j=\kappa_n}^{n} E_{\mathcal{F}_{j-1}}\left(\eta_{n,j}\eta_{n,j}'\right) &= \Sigma_u \otimes \frac{1}{n}\sum_{j=\kappa_n}^{n}\sum_{i,l=k+1}^{\kappa_n}\rho^{-i-l}u_{j-i}u_{j-l}' \\
&= \Sigma_u \otimes \sum_{i,l=k+1}^{\kappa_n}\rho^{-i-l}\left(\frac{1}{n}\sum_{j=\kappa_n}^{n}u_{j-i}u_{j-l}'\right) \\
&= \Sigma_u \otimes \sum_{i=k+1}^{\kappa_n}\rho^{-2i}\left(\frac{1}{n}\sum_{j=\kappa_n}^{n}u_{j-i}u_{j-i}'\right) + o_p(1) \\
&= \frac{\rho^{-2k}}{\rho^2-1}\Sigma_u \otimes \Sigma_u + o_p(1) \qquad (52)
\end{aligned}$$

where the third line follows by Lemma 2(ii) with $L = \kappa_n$ and dominated convergence, and the last line by Lemma 2(i). Combining (51) with (52) yields the required asymptotic variance. The Lindeberg condition can be established by an identical argument to part (i).

**Proof of Theorem 1.** Using the representation $H_n = \frac{1}{\|x_n\|}\left[x_n, \mathcal{R}_{\frac{\pi}{2}}x_n\right]$ and (46) with $k = -1$ we obtain

$$\begin{aligned}
\hat{R}_n - R &= \sum_{t=1}^{n}u_t x_{t-1}'\left(\sum_{t=1}^{n}x_{t-1}x_{t-1}'\right)^{-1} = \frac{1}{n}\sum_{t=1}^{n}u_t z_{t-1}'\left(\frac{1}{n}\sum_{t=1}^{n}z_{t-1}z_{t-1}'\right)^{-1}H_n' \\
&= \left[O_p\left(\rho^{-n}\frac{1}{n}\sum_{t=1}^{n}u_t x_{t-1}'\right), \left(n^{-1}\sum_{t=1}^{n}z_{2t-1}^2\right)^{-1}\frac{1}{n}\sum_{t=1}^{n}u_t z_{2t-1}\right]H_n' \\
&= \left[O_p\left(\rho^{-n}\right), \left(n^{-1}\sum_{t=1}^{n}z_{2t-1}^2\right)^{-1}\frac{1}{n}\sum_{t=1}^{n}u_t z_{2t-1} + O_p\left(\frac{1}{n}\right)\right]H_n' \\
&= \left\{1 + O_p\left(\frac{1}{n}\right)\right\}\left(n^{-1}\sum_{t=1}^{n}z_{2t-1}^2\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n}u_t z_{2t-1}\right)\frac{x_n'}{\|x_n\|}\mathcal{R}_{\frac{\pi}{2}}' \qquad (53)
\end{aligned}$$

since $\left\|\rho^{-n}\sum_{t=1}^{n}u_t x_{t-1}'\right\|_{L_1} \leq (\mathrm{tr}\Sigma_u)^{1/2}\rho^{-n}\sum_{t=1}^{n}\left(E\|x_{t-1}\|^2\right)^{1/2} \leq \mathrm{tr}\Sigma_u/\left(\rho^2-1\right)^{3/2}$. Using (9) and the second identity in (39) we can write

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}u_t z_{2t-1} &= -\frac{1}{n}\sum_{t=1}^{n}u_t\zeta_{n,t}'\mathcal{R}_{\frac{\pi}{2}}\frac{x_n}{\|x_n\|} \\
&= -\rho^{-1}\left(\frac{1}{n}\sum_{t=1}^{n}u_t u_t' + \frac{1}{n}\sum_{t=1}^{n}u_t\zeta_{n,t+1}\right)\mathcal{R}_{\frac{\pi}{2}}\frac{x_n}{\|x_n\|} \\
&= -\rho^{-1}\Sigma_u\mathcal{R}_{\frac{\pi}{2}}\frac{x_n}{\|x_n\|} + o_p(1) \qquad (54)
\end{aligned}$$

by parts (i) and (iii) of Lemma 2. Combining (53) and (54) we obtain

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

by Lemma 1 and (44) 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

$$h = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} := \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}\left(\xi_1^2+\xi_2^2\right)^{-1/2} \tag{55}$$

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

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

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

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

This density is that of an arc sine law and is shown in Fig. 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 (55) 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 $\theta$ 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 (46) yield, up to $1 + O_p(n^{-1})$,

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

By computing the second moment of $\sum_{t=1}^{n} u_t\zeta'_{n,t+k+1}$ and of $\sum_{t=1}^{n} u_t x'_{t+k}$, we obtain that $\sum_{t=1}^{n} u_t z_{2t+k} = O_p\left(n^{1/2}\right)$ and $\sum_{t=1}^{n} u_t z_{1t+k} = O_p\left(\rho^n\right)$. Hence,

$$\hat{R}_{n,k} - R = \frac{1}{n}\left[O_p\left(n\rho^{-n}\right)1_2, \frac{\sum_{t=1}^{n-k} z_{2t+k}u_t}{n^{-1}\sum_{t=1}^{n-k} z_{2t-1}z_{2t+k}} + O_p(1)\,1_2\right]H'_n \tag{56}$$

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

$$\begin{aligned}\sqrt{n}\left(\hat{R}_{n,k} - R\right) &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n-k} u_t z_{2t+k}\left(n^{-1}\sum_{t=1}^{n-k} z_{2t-1}z_{2t+k}\right)^{-1}\frac{x'_n}{\|x_n\|}\mathcal{R}'_{\frac{\pi}{2}} + O_p\left(\frac{1}{\sqrt{n}}\right) \\ &= -\frac{\frac{1}{\sqrt{n}}\sum_{t=1}^{n-k} u_t\zeta'_{n,t+k+1}\left(\frac{\mathcal{R}_{\frac{\pi}{2}}x_n}{\|x_n\|}\right)\left(\frac{\mathcal{R}_{\frac{\pi}{2}}x_n}{\|x_n\|}\right)'}{\left(\frac{\mathcal{R}_{\frac{\pi}{2}}x_n}{\|x_n\|}\right)'\frac{1}{n}\sum_{t=1}^{n-k}\zeta_{n,t}\zeta'_{n,t+k+1}\left(\frac{\mathcal{R}_{\frac{\pi}{2}}x_n}{\|x_n\|}\right)} + O_p\left(\frac{1}{\sqrt{n}}\right)\end{aligned} \tag{57}$$

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

$$X_{L_n} = \sum_{j=1}^{L_n}\rho^{-j}u_j \quad \text{and} \quad U_{n,k} = \frac{1}{\sqrt{n}}\sum_{t=2L_n+1}^{n-k}\left(\zeta_{n,t+k+1}\otimes u_t\right). \tag{58}$$

By Lemma 1 and direct calculation,

$$X_n - X_{L_n} \to_{a.s.} 0 \quad \text{and} \quad \left\|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-k}\left(\zeta_{n,t+k+1}\otimes u_t\right) - U_{n,k}\right\|_{L_2}^2 = O\left(\frac{L_n}{n}\right). \tag{59}$$

Lemma 3 then implies that $U_{n,k} \Rightarrow U$ as $n \to \infty$, where $U$ is a $N\left(0, \left(\rho^2-1\right)^{-1}\Sigma_u \otimes \Sigma_u\right)$ random vector. Therefore, using (59), and applying Lemma 2(iv) to the denominator of (57) yields

$$\begin{aligned}\sqrt{n}\mathrm{vec}\left(\hat{R}_{n,k} - R\right) &= -\rho^{k+1}\left(\rho^2-1\right)\frac{\left(\frac{\mathcal{R}_{\frac{\pi}{2}}X_{L_n}}{\|X_{L_n}\|}\right)\left(\frac{\mathcal{R}_{\frac{\pi}{2}}X_{L_n}}{\|X_{L_n}\|}\right)'\otimes I_2}{\left(\frac{\mathcal{R}_{\frac{\pi}{2}}X_{L_n}}{\|X_{L_n}\|}\right)'\Sigma_u\left(\frac{\mathcal{R}_{\frac{\pi}{2}}X_{L_n}}{\|X_{L_n}\|}\right)}U_{n,k} + o_p(1) \\ &\Rightarrow -\rho^{k+1}\left(\rho^2-1\right)\frac{\mathcal{R}_{\frac{\pi}{2}}X(\rho)X(\rho)'\mathcal{R}'_{\frac{\pi}{2}}\otimes I_2}{X(\rho)'\mathcal{R}'_{\frac{\pi}{2}}\Sigma_u\mathcal{R}_{\frac{\pi}{2}}X(\rho)}U.\end{aligned} \tag{60}$$

If $(\|X_{L_n}\| \|U_{n,k}\|)_{n\geq 1}$ is a U.I. sequence, $E(X(\rho)U') = \lim_{n\to\infty} E\left(X_{L_n}U'_{n,k}\right) = 0$, so $X(\rho)$ and $U$ are uncorrelated. It remains to show the required uniform integrability. If (12) is replaced by the stronger condition that $\left(\|u_t\|^2\right)_{t\in\mathbb{N}}$ is a U.I. sequence, uniform integrability of the sequence $\left(\|X_{L_n}\|^2\right)_{n\geq 1}$ can be established by using an identical argument to the proof of Lemma 0. Thus, since $E\|U_{n,k}\|^2 \leq c = (\mathrm{tr}\Sigma_u)^2/(\rho-1)$ for all $n$, letting $G_{n,\lambda} = \{\|X_{L_n}\| \|U_{n,k}\| > \lambda\}$, $C_{n,\lambda} = \left\{\|X_{L_n}\| > \lambda^{1/2}\right\}$ and $D_{n,\lambda} = \left\{\|U_{n,k}\| > \lambda^{1/2}\right\}$ the Cauchy Schwarz inequality yields, as $\lambda\to\infty$,

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

by the Chebyshev inequality and uniform integrability of $\left(\|X_{L_n}\|^2\right)_{n\geq 1}$.

If, in addition, $(u_t)$ is an $L_n$-dependent sequence, $U_{n,k}$ and $X_n$ in (58) are independent, so $U$ and $X(\rho)$ are independent random vectors and the limit in (60) has the mixed normal distribution given in (26).

**Proof of (27).** Using the identity $\hat{u}_{tk} = u_t - \left(\hat{R}_{n,k} - R\right)x_{t-1}$ for the residuals, the estimator of $\Sigma_u$ can be written as:

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}\hat{u}_{tk}\hat{u}'_{tk} &= \frac{1}{n}\sum_{t=1}^{n}u_t u'_t - \frac{1}{n}\sum_{t=1}^{n}u_t x'_{t-1}\left(\hat{R}_{n,k} - R\right)' - \frac{1}{n}\left(\hat{R}_{n,k} - R\right)\sum_{t=1}^{n}x_{t-1}u'_t \\
&\quad + \frac{1}{n}\left(\hat{R}_{n,k} - R\right)\sum_{t=1}^{n}x_{t-1}x'_{t-1}\left(\hat{R}_{n,k} - R\right)'. \qquad (61)
\end{aligned}$$

The consistency result in (27) will follow from Lemma 2(i) and

$$\frac{1}{n}\sum_{t=1}^{n}\hat{u}_{tk}\hat{u}'_{tk} = \frac{1}{n}\sum_{t=1}^{n}u_t u'_t + O_p\left(n^{-1/2}\right). \qquad (62)$$

To show (62), using the orthogonality of $H_n$ and (56) we obtain

$$\begin{aligned}
\left(\hat{R}_{n,k}-R\right)\sum_{t=1}^{n}x_{t-1}u_t' &= \left(\hat{R}_{n,k}-R\right)H_n\sum_{t=1}^{n}z_{t-1}u_t' \\
&= \frac{1}{n}\left[O_p\left(n\rho^{-n}\right)1_2, \frac{\sum_{t=1}^{n-k}z_{2t+k}u_t+O_p\left(1_2\right)}{n^{-1}\sum_{t=1}^{n-k}z_{2t-1}z_{2t+k}}\right]\left[\begin{array}{c}\sum_{t=1}^{n}z_{1t-1}u_t' \\ \sum_{t=1}^{n}z_{2t-1}u_t'\end{array}\right] \\
&= O_p\left(1\right)I_2+\frac{1}{n}\frac{\left(\sum_{t=1}^{n-k}z_{2t+k}u_t\right)\left(\sum_{t=1}^{n}z_{2t-1}u_t'\right)}{n^{-1}\sum_{t=1}^{n-k}z_{2t-1}z_{2t+k}} \\
&= O_p\left(n^{1/2}\right)I_2, \qquad (63)
\end{aligned}$$

since $\sum_{t=1}^{n}z_{1t-1}u_t'=O_p\left(\rho^n\right)$ by PM, $\sum_{t=1}^{n}z_{2t-1}u_t'=O_p\left(n\right)$ by (54) and $\sum_{t=1}^{n-k}z_{2t+k}u_t=O_p\left(n^{1/2}\right)$ by Lemma 3. This shows that the second and third terms of (61) have order $O_p\left(n^{-1/2}\right)$. For the last term of (61), the identity

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

the fact that $\hat{R}_{n,k}-R=O_p\left(n^{-1/2}\right)$ and (63) imply that

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

Since $z_n=\left(\left\|x_n\right\|,0\right)'=O_{a.s.}\left(\rho^n\right)$, (56) yields

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

This shows that the last term of (61) has order $O_p\left(n^{-1}\right)$ and establishes (62).

**Proof of (32) when** $\Sigma_{u\varepsilon}\neq 0$ **.** Letting $v_t=\left(u_t,\varepsilon_t'\right)'$, $\mathcal{F}_{v,t}=\sigma\left(v_t,v_{t-1},...\right)$ and $\mathcal{F}_v^t=\sigma\left(v_t,v_{t+1},...\right)$ and replacing $u_t$ by $v_t$ in the assumptions of Lemma 3, we can show that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(\zeta_{n,t+1}\otimes\varepsilon_t\right)\Rightarrow N\left(0,\frac{1}{\rho^2-1}\Sigma_u\otimes\Sigma_\varepsilon\right) \qquad (64)$$

where $\Sigma_u$ and $\Sigma_\varepsilon$ are defined in (33). The proof of (64) is identical to the proof of Lemma 3. Using (46) and proceeding as in the proof of Theorem 3, we obtain

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

where $L_n$ and $X_{L_n}$ are defined as in the proof of Theorem 3. The proof then follows by (64) 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,

$$\begin{aligned}
\tilde{A}_n - A &= \frac{1}{n}\sum_{t=1}^{n}\varepsilon_t x_{t-1}'\left(\frac{1}{n}\sum_{t=1}^{n} x_{t-1}x_{t-1}'\right)^{-1} \\
&= \left(n^{-1}\sum_{t=1}^{n} z_{2t-1}^2\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t z_{2t-1}\right)\frac{x_n'}{\|x_n\|}\mathcal{R}_{\frac{\pi}{2}}' + o_p(1) \\
&= -\left[I_m + o_p(1)\right]\left(\rho^2 - 1\right)\left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t\zeta_{n,t}'\right)\frac{\mathcal{R}_{\frac{\pi}{2}} X(\rho) X(\rho)' \mathcal{R}_{\frac{\pi}{2}}'}{X(\rho)'\mathcal{R}_{\frac{\pi}{2}}'\Sigma_u\mathcal{R}_{\frac{\pi}{2}} X(\rho)}, \quad (65)
\end{aligned}$$

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

$$\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t\zeta_{n,t}' = \rho^{-1}\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t u_t' + \rho^{-1}\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t\zeta_{n,t+1}' \to_p \Sigma_{\varepsilon u}, \qquad (66)$$

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 (65) and (66).

# References


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

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

Lai, T.L. and 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.

Magdalinos, T. and 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." University of Oxford working paper.

Park, J. Y. and P. C. B. Phillips (1988). "Statistical Inference in Regressions With Integrated Processes: Part 1," *Econometric Theory* 4, 468–497.

Park, J. Y. and 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. and T. Magdalinos (2008). "Limit theory for explosively cointegrated systems", *Econometric Theory*, 24, 865-887.

Phillips P. C. B., Y. Wu and J. Yu (2010). "Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values?", *International Economic Review* (forthcoming)

Phillips P. C. B. and J. Yu (2010). "Dating the Timeline of Financial Bubbles during the Subprime Crisis" Yale University Working Paper.

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