Estimation and Testing for Partially Nonstationary Vector Autoregressive Models with GARCH: WLS versus QMLE *

Chor-yiu SIN

Department of Economics

Hong Kong Baptist University, Hong Kong

June 4, 2004

Abstract

Macroeconomic or financial data are often modelled with cointegration and GARCH. Noticeable examples include those studies of price discovery, in which stock prices of the same underlying asset are cointegrated and they exhibit multivariate GARCH. Modifying the asymptotic theories developed in Li, Ling and Wong (2001) and Sin and Ling (2004), this paper proposes a WLS (weighted least squares) for the parameters of an ECM (error-correction model). Apart from its computational simplicity, by construction, the consistency of WLS is insensitive to possible mis-specification in conditional variance. Further, asymmetrically distributed deflated error is allowed, at the expense of more deliberate estimation procedures. Efficiency loss relative to QMLE (quasi-maximum likelihood estimator) is discussed within the class of LABF (locally asymptotically Brownian functional) models. The insensitivity and efficiency of WLS in finite samples are examined through Monte Carlo experiments. We also apply the WLS to an empirical example of HSI (Hang Seng Index), HSIF (Hang Seng Index Futures) and TraHK (Hong Kong Tracker Fund).

Key Words: Asymmetric distribution; Cointegration; LABF models; Multivariate GARCH; Price discovery; WLS.

JEL Codes: C32, C51, G14

* *Acknowledgments:* I thank helpful comments from Chi-shing CHAN and Shiqing LING. This research is partially supported by the Hong Kong Research Grant Council competitive earmarked grant HKBU2014/02H. The usual disclaimers apply.

1 Introduction

Throughout this paper, we consider an $m-$dimensional autoregressive (AR) process $\{Y_t\}$, which is generated by

$$Y_t = \Phi_1 Y_{t-1} + \cdots + \Phi_s Y_{t-s} + \varepsilon_t, \tag{1.1}$$

$$E(\varepsilon_t \mid \mathcal{I}_{t-1}) = 0, \tag{1.2}$$

where Φ_j's are constant matrices, and $\mathcal{I}_t = \sigma\{\varepsilon_s, s = t, t-1, \ldots\}$.

Assuming the ε_t's are i.i.d., under further conditions on Φ_j's (See Assumptions 2.1 and 2.2 below), Ahn and Reinsel (1990) (see also Johansen, 1996) show that, although some component series of $\{Y_t\}$ exhibit nonstationary/$I(1)$ behaviour, there are r linear combinations of $\{Y_t\}$ that are stationary/$I(0)$. This phenomenon, which is called cointegration in the literature of economics, was first investigated by Granger (1983) (see also Engle and Granger, 1987). The partially nonstationary multivariate AR model or cointegrating time series models without GARCH have been extensively discussed over the past twenty years. Other noticeable examples include Phillips and Durlauf (1986) and Stock and Watson (1993).

Economic time series related to financial markets often exhibit time-varying variances. As far as we know, Li, Ling and Wong (2001) first investigate multivariate time series that exhibit both cointegration and time-varying variances, where the heteroskedasticity part is the random coefficient AR model and the scope of applications is thus restricted. Extending Li et al. (2001)'s estimation results, Sin and Ling (2004) construct a likelihood ratio (LR) test for reduced rank; and modify their model. They consider a multivariate GARCH with constant correlations, which was first suggested by Bollerslev (1990) and widely used in many papers in the economics and finance literature. More precisely, the conditional variance-covariance matrix, denoted as $\tilde{V}_{t-1}$, is *modelled* as $\tilde{D}_{t-1}\tilde{\Gamma}\tilde{D}_{t-1}$, where $\tilde{D}_{t-1} = diag(\sqrt{\tilde{h}_{1t-1}}, \ldots, \sqrt{\tilde{h}_{mt-1}})$

and:

$$\tilde{h}_{it-1} = \tilde{a}_{i0} + \sum_{j=1}^{q} \tilde{a}_{ij}\varepsilon_{it-j}^2 + \sum_{k=1}^{p} \tilde{b}_{ik}\tilde{h}_{it-1-k}, \tag{1.3}$$

$$\tilde{\Gamma} \equiv (\tilde{\gamma}_{ij})_{m\times m}, \text{ a symmetric positive definite matrix with } \tilde{\gamma}_{ii} = 1. \tag{1.4}$$

This paper assumes the existence of some *pseudo true* parameters of this multivariate GARCH model (1.3)-(1.4), which satisfy Assumption 2.4 below. However, in view of the possible mis-specification in variance (see, for instance, the GJR model first suggested in Glosten, Jagannathan and Runkle, 1993, the extended model first suggested in Jeantheau, 1998, and the time-varying correlation model first suggested in Tse and Tsui, 2002), instead of a QMLE (quasi-maximum likelihood estimator), we consider a WLS (weighted least squares), which is computationally simpler on the one hand, and is insensitive to possible mis-specification of variance on the other hand. See Section 2 of Sin (2003) for a related study in the purely stationary case. Further, we also allow asymmetrically distributed deflated errors, at the expense of (i) a more involved estimation procedure; and/or (ii) a more involved asymptotic distribution. Efficiency loss relative to the QMLE is discussed within the class of LABF (locally asymptotically Brownian functional) models.

This paper proceeds as follows. Section 2 discusses the structure of the DGP (data generating process) or the model (1.1)-(1.4). Assuming a symmetric distribution, Section 3 derives the asymptotic distributions of the full rank estimator, the reduced rank estimator, and a test for reduced rank. Relaxing the symmetry assumption, Section 4 considers a modified weighting matrix (modified upon the original one $\tilde{V}_{t-1}^{-1}$) and thus the efficiency of the estimators may be altered. In Section 5, we maintain the original weighting matrix and consider an alternative estimation. Section 6 contains a brief discussion on the estimation procedures as well as their efficiency of the estimators discussed so far. Monte Carlo experiments and an illustrative empirical example are discussed in the subsequent sections. We conclude in the last section.

Readers who are much interested in the algorithms may jump to (3.7) for the full rank estimation, to (3.14)-(3.15) for the reduced rank estimation, (3.17) (or (3.19) for a robust version) for the test statistics for reduced rank. In Section 4, the counterparts can be found in (4.5), (4.6)-(4.7) and (4.8) (a robust version) respectively. In Section 5, they are found in (5.1), (5.3)-(5.4), and (5.5) (or (5.11) for a robust version) respectively. Throughout, $\longrightarrow_{\mathcal{L}}$ denotes convergence in distribution, $\longrightarrow_p$ denotes convergence in probability, $O_p(1)$ denotes a series of random numbers that are bounded in probability, and $o_p(1)$ denotes a series of random numbers converging to zero in probability. $\tilde{\theta}$ denotes a *generic* version of the *(pseudo) true* parameter θ, while $\hat{\theta}$ denotes an *initial* estimator and $\dot{\theta}$ denotes a *full rank* or *reduced rank* estimator.

2 Basic Properties of the Models

Denote L as the lag operator. Refer to (1.1)-(1.2) and define $\Phi(L) = I_m - \sum_{j=1}^{s} \Phi_j L^j$. We first make the following assumption:

Assumption 2.1. *$\mid \Phi(z) \mid = 0$ implies that either $\mid z \mid > 1$ or $z = 1$.* □

Define $W_t = Y_t - Y_{t-1}$, $\Pi_j = -\sum_{k=j+1}^{s} \Phi_k$ and $C = -\Phi(1) = -(I_m - \sum_{j=1}^{s} \Phi_j)$. By a Taylor's formula, $\Phi(L)$ can be decomposed as:

$$\Phi(z) = (1-z)I_m - Cz - \sum_{j=1}^{s-1} \Pi_j (1-z) z^j. \tag{2.1}$$

Thus, we can reparameterize process (1.1) as:

$$W_t = CY_{t-1} + \sum_{j=1}^{s-1} \Pi_j W_{t-j} + \varepsilon_t. \tag{2.2}$$

Following Ahn and Reinsel (1990) and Johansen (1996), we can decompose $C = AB$, where A and B are respectively $m \times r$ and $r \times m$ matrices of rank r. Define $d = m - r$. Denote $B_\perp$ as a $d \times m$ matrix of full rank such that $BB_\perp' = 0_{r \times d}$, $\bar{B} = (BB')^{-1}B$ and $\bar{B}_\perp = (B_\perp B_\perp')^{-1} B_\perp$, and $A_\perp$ as an $m \times d$ matrix of full rank such that $A'A_\perp = 0_{r\times d}$, $\bar{A} = A(A'A)^{-1}$ and $\bar{A}_\perp = A_\perp (A_\perp' A_\perp)^{-1}$. We impose the following conditions:

Assumption 2.2. $\mid A'_\perp(I_m - \sum_{j=1}^{s-1}\Pi_j)B'_\perp \mid \neq 0.$ $\square$

Assumption 2.3. $\{\varepsilon_t\}$ *is a stationary process,* $E(vec[\varepsilon_t\varepsilon_t']vec[\varepsilon_t\varepsilon_t']') < \infty.$ $\square$

Unlike Ahn and Reinsel (1990), we do not assume the existence of their Jordan canonical form and include some DGPs such as that in Exercise 4.3, pp.62-63 of Johansen (1996), which approach is essentially adopted here. Given Assumptions 2.1-2.2, by the proof of Theorem (4.2) in Johansen (1996),

$$\Psi(L)\begin{bmatrix} (1-L)B_\perp Y_t \\ BY_t \end{bmatrix} = (\bar{A}_\perp, \bar{A})'\varepsilon_t, \tag{2.3}$$

where $\Psi(z) = (\bar{A}_\perp, \bar{A})'\Phi(z)(\bar{B}'_\perp, \bar{B}'(1-z)^{-1})$ is invertible for $\mid z \mid < 1+\rho$ for some $\rho > 0$. In other words, similar to Ahn and Reinsel (1990), we can consider the following transformation:

$$Z_{1t} = B_\perp Y_t = Z_{1t-1} + u_{1t}, \text{ and } Z_{2t} = BY_t = u_{2t}, \tag{2.4}$$

where $u_t = (u'_{1t}, u'_{2t})' = \psi(L)a_t$, $\psi(L) \equiv \Psi^{-1}(L)$ and $a_t \equiv (\bar{A}_\perp, \bar{A})'\varepsilon_t$. Note in Assumption 2.3, the i.i.d assumption in Johansen (1996) is replaced by a stationarity assumption. Given this, by (2.3)-(2.4), Z_{1t} is $I(1)$ while Z_{2t} is $I(0)$.

Further, we make the following assumptions on (1.3)-(1.4).

Assumption 2.4. *For* $i = 1, \ldots, m$, *the pseudo true parameters* $a_{i0} > 0$, $a_{i1}, \ldots, a_{iq}, b_{i1}, \ldots, b_{ip} \geq 0$, $\sum_{j=1}^{q} a_{ij} + \sum_{k=1}^{p} b_{ik} < 1$; *and* $\{\eta_{it} \equiv \varepsilon_{it}/\sqrt{h_{it-1}}\}$ *is a stationary process.* $\square$

Assumption 2.5. $\eta_t \equiv (\eta_{1t}, \ldots, \eta_{mt})'$ *is symmetrically distributed.* $\square$

The stationarity assumption in Assumptions 2.3-2.4 can be weakened to heterogeneous and mixing process assumption, as in Phillips and Durlauf (1986). We keep this for simplicity. (1.2) and Assumption 2.4 imply that although $E(\eta_t \mid \mathcal{I}_{t-1}) = 0$, in general $E(\eta_t \mid \mathcal{I}_{t-1}) \neq \Gamma$. In view of this possible mis-specification, we do not make primitive assumptions that render stationarity and finite fourth moments in Assumption 2.3. We will come back to this point in a subsequent section.

The symmetry assumption in Assumption 2.5 will be used in the next section. It will be relaxed in Sections 4 and 5.

We close this section with a basic lemma, which is found useful in proving the results in the subsequent sections. Let $(W_m'(u),\ W_m^{*\prime}(u))'$ be a $2m-$dimensional Brownian motion (BM) with the covariance matrix:

$$u\Omega \equiv u \left(\begin{array}{cc} V^* & E(\varepsilon_t\varepsilon_t' V_{t-1}^{-1}) \\ E(V_{t-1}^{-1}\varepsilon_t\varepsilon_t') & \Omega_1^* \end{array} \right),$$

where $V^* = E\varepsilon_t\varepsilon_t'$, and $\Omega_1^* = E(V_{t-1}^{-1}\varepsilon_t\varepsilon_t' V_{t-1}^{-1})$. Let $B_d(u) = \Omega_{a_1}^{-1/2}[I_d, 0_{dxr}]\,\Omega_a^{1/2} V^{*-1/2} W_m(u)$, where $\Omega_a = E(a_t a_t')$ and $\Omega_{a_1} = [I_d, 0_{dxr}]\Omega_a[I_d, 0_{dxr}]'$.

Lemma 2.1. *Suppose Assumptions 2.1-2.4 hold. Then*

$$(a)\qquad n^{-2}\sum_{t=1}^{n} Z_{1t-1}Z_{1t-1}' \otimes V_{t-1}^{-1} \longrightarrow_{\mathcal{L}} (\psi_{11}\Omega_{a_1}^{1/2}\int_0^1 B_d(u)B_d(u)'\Omega_{a_1}^{1/2}\psi_{11}') \otimes \Omega_1,$$

$$(b)\qquad n^{-3/2}\sum_{t=1}^{n} Z_{1t-1}U_{t-1}' \otimes V_{t-1}^{-1} \longrightarrow_{\mathcal{L}} (\psi_{11}\Omega_{a_1}^{1/2}\int_0^1 B_d(u)du \otimes I_m)E(U_{t-1}' \otimes V_{t-1}^{-1}),$$

(c) *If in addition, Assumption 2.5 holds,*

$$n^{-3/2}\sum_{t=1}^{n} Z_{1t-1}U_{t-1}' \otimes V_{t-1}^{-1} = o_p(1),$$

$$(d)\qquad n^{-1}\sum_{t=1}^{n} U_{t-1}U_{t-1}' \otimes V_{t-1}^{-1} \longrightarrow_p \Omega_2,$$

$$(e)\qquad n^{-1}\sum_{t=1}^{n} Z_{1t-1} \otimes V_{t-1}^{-1}\varepsilon_t \longrightarrow_{\mathcal{L}} vec[(\int_0^1 B_d(u)dW_m^*(u)')'\Omega_{a_1}^{1/2}\psi_{11}],$$

$$(f)\qquad n^{-1/2}\sum_{t=1}^{n} U_{t-1} \otimes V_{t-1}^{-1}\varepsilon_t \longrightarrow_{\mathcal{L}} \Omega_2^{*1/2}\Phi,$$

where $\Phi \sim N(0, I_{rm+(s-1)m^2})$, $\psi_{11} \equiv [I_d, 0](\sum_{k=1}^{\infty}\psi_k)[I_d, 0]'$, $\Omega_1 \equiv E(V_{t-1}^{-1})$, $\Omega_2 \equiv E(U_{t-1}U_{t-1}' \otimes V_{t-1}^{-1})$, $\Omega_2^* \equiv E(U_{t-1}U_{t-1}' \otimes V_{t-1}^{-1}\varepsilon_t\varepsilon_t' V_{t-1}^{-1})$, *and* $U_{t-1} = [(BY_{t-1})', W_{t-1}', \cdots, W_{t-s+1}']'$

3 Assuming Symmetric Distribution

In this section, we follow the lines in Sin and Ling (2004) and assume a symmetric distribution of the deflated error η_t. See Assumption 2.5 above. The procedures of the full rank estimator, the reduced rank estimator, and a test for reduced rank as well as their asymptotic distributions resemble those in Sections 3-5 in Sin and Ling (2004).

3.1 Full Rank Estimation

Refer to Process (2.2). In this section, we consider the full rank estimator for $\varphi \equiv vec[C, \Pi_1, \ldots, \Pi_{s-1}]$, given some initial estimator for the *pseudo true* parameter of the conditional variance (see Model (1.3)-(1-4) and Assumption (2.4) above). Similar to above, the *generic* "variance parameter" is denoted as $\tilde{\delta}$ while the pseudo true "variance parameter" is simply denoted as δ.

Given $\{Y_t : t = 1, \cdots, n\}$, conditional on the initial values $Y_s = 0$ for $s \leq 0$, the log-likelihood function (LF) (with a constant ignored) can be written as:

$$l(\tilde{\varphi}, \tilde{\delta}) = \sum_{t=1}^{n} l_t(\tilde{\varphi}, \tilde{\delta}) \text{ and } l_t(\tilde{\varphi}, \tilde{\delta}) = -\frac{1}{2}\tilde{\varepsilon}_t' \tilde{V}_{t-1}^{-1} \tilde{\varepsilon}_t - \frac{1}{2} \ln |\tilde{V}_{t-1}|, \tag{3.1}$$

where $\tilde{V}_{t-1} = \tilde{D}_{t-1}\tilde{\Gamma}\tilde{D}_{t-1}$, $\tilde{D}_{t-1} = diag(\sqrt{\tilde{h}_{1t-1}}, \ldots, \sqrt{\tilde{h}_{mt-1}})$. $\tilde{\varepsilon}_t$, $\tilde{V}_{t-1}$, $\tilde{D}_{t-1}$ and $\tilde{h}_{it-1}$'s are functions of the *generic* parameter $(\tilde{\varphi}, \tilde{\delta})$. Further denote $\tilde{h}_{t-1} = (\tilde{h}_{1t-1}, \ldots, \tilde{h}_{mt-1})'$ and $\tilde{H}_{t-1} = (\tilde{h}_{1t-1}^{-1}, \ldots, \tilde{h}_{mt-1}^{-1})'$. Define $X_{t-1} \equiv [Y_{t-1}', W_{t-1}', \ldots, W_{t-s+1}']'$. The score function w.r.t. φ can be written as:

$$\nabla_\varphi \tilde{l}_t = -\frac{1}{2}\nabla_\varphi \tilde{h}_{t-1}(\iota - w(\tilde{\varepsilon}_t\tilde{\varepsilon}_t'\tilde{V}_{t-1}^{-1})) \odot \tilde{H}_{t-1} + (X_{t-1} \otimes I_m)\tilde{V}_{t-1}^{-1}\tilde{\varepsilon}_t, \tag{3.2}$$

where $\iota = (1, 1, \ldots, 1)_{m\times 1}'$ and $w(.)$ is a vector containing the diagonal elements of a square matrix. In Sin and Ling (2004), the score function (3.2) is used. As one can see in that paper, the algorithm for the one-step estimator is quite involved. More importantly, if the multivariate GARCH is mis-specified and for all $(\tilde{\varphi}, \tilde{\delta})$, $Prob\{E[\nabla_\varphi \tilde{h}_{t-1}(\iota - w(\tilde{\varepsilon}_t\tilde{\varepsilon}_t'\tilde{V}_{t-1}^{-1})) \odot \tilde{H}_{t-1} \mid \mathcal{I}_{t-1}] = 0\} < 1$, it is unclear what the asymptotic properties of the one-step estimator carries. In view of that, for our *WLS*, we only consider the second part of the score function:

$$\tilde{f}_t \equiv (X_{t-1} \otimes I_m)\tilde{V}_{t-1}^{-1}\tilde{\varepsilon}_t. \tag{3.3}$$

Denote $Q^* = \text{diag}([B_\perp', B']' \otimes I_m, I_{(s-1)m^2})$ and $D^* = \text{diag}(nI_{dm}, \sqrt{n}I_{rm+(s-1)m^2})$. For any fixed positive constant K, let $\Theta_n \equiv \{(\tilde{\varphi}, \tilde{\delta}) : \|D^* Q^{*'-1}(\tilde{\varphi} - \varphi)\| \leq K$ *and* $\|\sqrt{n}(\tilde{\delta} - \delta)\| \leq K\}$, where (φ, δ) is the *true* parameter. Using Assumptions

2.1-2.5 and a similar method as in Ling and Li (1998), the derivative of $\tilde{f}_t$ on Θ_n can be *simplified* as follows:

$$D^{*-1}Q^*(\sum_{t=1}^{n}\nabla_{\varphi'}\tilde{f}_t)Q^{*'}D^{*-1} = \sum_{t=1}^{n}D^{*-1}Q^*\tilde{F}_tQ^{*'}D^{*-1} + o_p(1), \tag{3.4}$$

where $\tilde{F}_t \equiv -(X_{t-1}X'_{t-1} \otimes \tilde{V}_{t-1}^{-1})$.

Moreover, we can show the following results hold uniformly in Θ_n:

$$\sum_{t=1}^{n}D^{*-1}Q^*(\tilde{F}_t - F_t)Q^{*'}D^{*-1} = o_p(1), \tag{3.5}$$

$$\sum_{t=1}^{n}D^{*-1}Q^*(\tilde{f}_t - f_t) = \sum_{t=1}^{n}D^{*-1}Q^*F_t(\tilde{\varphi} - \varphi) + o_p(1), \tag{3.6}$$

where $F_t = -(X_{t-1}X'_{t-1} \otimes V_{t-1}^{-1})$ and $f_t = (X_{t-1} \otimes I_m)V_{t-1}^{-1}\varepsilon_t$. In view of (3.4)-(3.6), we first find an initial estimator $(\hat{\varphi}, \hat{\delta}) \in \Theta_n$. For instance, $\hat{\varphi}$ can be the least squares (LS) estimator while $\hat{\delta}$ may be the QMLE with the true ε_t replaced by the LS residual $\hat{\varepsilon}_t \equiv W_t - \hat{C}Y_{t-1} - \sum_{j=1}^{s-1}\hat{\Pi}_jW_{t-j}$. See Section 2 of Ling, Li and McAleer (2003). Given $(\hat{\varphi}, \hat{\delta})$, we perform one iteration of the WLS:

$$\dot{\varphi} \;=\; \hat{\varphi} - [\sum_{t=1}^{n}F_t(\hat{\varphi}, \hat{\delta})]^{-1}[\sum_{t=1}^{n}f_t(\hat{\varphi}, \hat{\delta})], \tag{3.7}$$

where $\hat{F}_t = X_{t-1}X'_{t-1} \otimes \hat{V}_{t-1}^{-1}$, $\hat{V}_{t-1}$ is the $\tilde{V}_{t-1}$ in Model (1.3)-(1.4) evaluated at $(\hat{\varphi}, \hat{\delta})$. The following is proved in the Appendix.

Theorem 3.1. *Suppose Assumptions 2.1-2.5 hold. Then*

(*a*) $\quad n(\dot{C} - C)\bar{B}'_{\perp} \longrightarrow_{\mathcal{L}} \Omega_1^{-1}M^*,$

(*b*) $\quad \sqrt{n}vec[(\dot{C} - C)\bar{B}', (\dot{\Pi}_1 - \Pi_1), \ldots, (\dot{\Pi}_{s-1} - \Pi_{s-1})] \longrightarrow_{\mathcal{L}} N(0, \Omega_2^{-1}\Omega_2^*\Omega_2^{-1}),$

where $M^* = (\int_0^1 B_d(u)dW_m^*(u)')'(\int_0^1 B_d(u)B_d(u)'du)^{-1}\Omega_{a_1}^{-1/2}\psi_{11}^{-1}$, *and the remaining variables are defined as in Lemma 2.1.* □

When $E(\varepsilon_t\varepsilon'_t \mid \mathcal{I}_{t-1}) = V_{t-1}$, $\Omega_1^* = \Omega_1$ and $\Omega_2^* = \Omega_2$. On the other hand, the asymptotic distribution of the nonstationary component argument in (a) is independent of that of the stationary component argument in (b). As one can see in the proof, this suffices to have:

$$n^{-3/2}\sum_{t=1}^{n}(Z_{1t-1}\otimes I_m)(U'_{t-1}\otimes V_{t-1}^{-1})=o_p(1),$$

which by Lemma 2.1(c), is implied by $E[U'_{t-1}\otimes V_{t-1}^{-1}]=0$, a result depending on the variance model (1.3)-(1.4) and the symmetry assumption.

3.2 Reduced Rank Estimation

We first rewrite (2.2) in a reduced rank form:

$$W_t = ABY_{t-1}+\sum_{j=1}^{s-1}\Pi_j W_{t-j}+\varepsilon_t, \tag{3.8}$$

where A and B are defined as in Section 2. Denote $\alpha=[\alpha'_1,\alpha'_2]'$ with $\alpha_1\equiv vec[B]$ and $\alpha_2\equiv vec[A,\Pi_1,\ldots,\Pi_{s-1}]$. The LF based on the error-correction form (3.9) is the same as that in (3.1), but now it is a function of the *generic* parameters $\tilde{\alpha}$ and $\tilde{\delta}$. Denote $U^*_{t-1}\equiv[(Y_{t-1}\otimes A')',(U_{t-1}\otimes I_m)']'$, where we recall from Theorem 3.1 that $U_{t-1}=[(BY_{t-1})',W'_{t-1},\ldots,W'_{t-s+1}]'$. Similar to (3.2),

$$\nabla_\alpha\tilde{l}_t=\nabla_\alpha l_t(\tilde{\alpha},\tilde{\delta})=-\frac{1}{2}(\nabla_\alpha\tilde{h}_{t-1})(\iota-w(\tilde{\varepsilon}_t\tilde{\varepsilon}'_t\tilde{V}_{t-1}^{-1}))\odot\tilde{H}_{t-1}+\tilde{U}^*_{t-1}\tilde{V}_{t-1}^{-1}\tilde{\varepsilon}_t. \tag{3.9}$$

Our WLS only considers the second term in (3.9), that is:

$$\tilde{r}_t=r_t(\tilde{\alpha},\tilde{\delta})=(\tilde{r}'_{1t},\tilde{r}'_{2t})', \tag{3.10}$$

where $\tilde{r}_{1t}\equiv(Y_{t-1}\otimes\tilde{A}')\tilde{V}_{t-1}^{-1}\tilde{\varepsilon}_t$ and $\tilde{r}_{2t}\equiv(\tilde{U}_{t-1}\otimes I_m)\tilde{V}_{t-1}^{-1}\tilde{\varepsilon}_t$.

Denote $Q^{**}\equiv diag((B_\perp\otimes I_r),I_{rm+(s-1)m^2})$ and $D^{**}\equiv diag(nI_{rd},\sqrt{n}I_{rm+(s-1)m^2})$. For any fixed positive constant K, let $\Xi_n\equiv\{(\tilde{\alpha},\tilde{\delta}):\|D^{**}Q^{**\prime-1}(\tilde{\alpha}-\alpha)\|\leq K\ and\ \|\sqrt{n}(\tilde{\delta}-\delta)\|\leq K\}$. Given Assumptions (2.1)-(2.5), similar to (3.4), on Ξ_n, the derivative of $\tilde{r}_t$ can be *simplified* as follows:

$$D^{**-1}Q^{**}\sum_{t=1}^{n}\nabla_{\alpha'}\tilde{r}_tQ^{**\prime}D^{**-1}=D^{**-1}Q^{**}\sum_{t=1}^{n}\tilde{R}_tQ^{**\prime}D^{**-1}+o_p(1), \tag{3.11}$$

where $\tilde{R}_t=diag\{\tilde{R}_{1t},\tilde{R}_{2t}\}$, $\tilde{R}_{1t}=-(Y_{t-1}Y'_{t-1}\otimes\tilde{A}'\tilde{V}_{t-1}^{-1}\tilde{A})$, $\tilde{R}_{2t}=-(\tilde{U}_{t-1}\tilde{U}'_{t-1}\otimes\tilde{V}_{t-1}^{-1})$.

Similar to (3.5)-(3.6), the following results hold uniformly in Ξ_n:

$$D^{**-1}Q^{**}\sum_{t=1}^{n}(\tilde{R}_t - R_t)Q^{**\prime}D^{**-1} = o_p(1), \tag{3.12}$$

$$D^{**-1}Q^{**}\sum_{t=1}^{n}(\tilde{r}_t - r_t) = D^{**-1}Q^{**}\sum_{t=1}^{n}R_t(\tilde{\alpha} - \alpha) + o_p(1), \tag{3.13}$$

where R_t and r_t are $\tilde{R}_t$ and $\tilde{r}_t$ evaluated at the *(pseudo) true* parameters α and δ. Given (3.12)-(3.14), we first find an initial estimator $(\hat{\alpha}, \hat{\delta})$ which, after certain *normalization*, belongs to Ξ_n. For instance, $\hat{\alpha} = [\hat{\alpha}_1', \hat{\alpha}_2']'$ may be that from Johansen (1996) (see for instance, Chapter 13 there) while $\hat{\delta}$ may be that from Sub-section 3.1. Given $(\hat{\alpha}, \hat{\delta})$, we first perform one iteration of the WLS for α_1:

$$\dot{\alpha}_1 = \hat{\alpha}_1 - [\sum_{t=1}^{n} R_{1t}(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\delta})]^{-1}[\sum_{t=1}^{n} r_{1t}(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\delta})]. \tag{3.14}$$

Once we obtain the reduced rank estimator for α_1 which incorporates the possible heteroskedasticity, we perform one iteration of the WLS for α_2:

$$\dot{\alpha}_2 = \hat{\alpha}_2 - [\sum_{t=1}^{n} R_{2t}(\dot{\alpha}_1, \hat{\alpha}_2, \hat{\delta})]^{-1}[\sum_{t=1}^{n} r_{2t}(\dot{\alpha}_1, \hat{\alpha}_2, \hat{\delta})]. \tag{3.15}$$

In view of (3.12)-(3.14), the asymptotic distributions of the *normalized* estimators for α_1 and α_2 are given as follows.

Theorem 3.2. *Suppose Assumptions 2.1-2.5 hold. Then*

(*a*) $$n((\dot{B}\bar{B}')^{-1}\dot{B} - B)\bar{B}_\perp' \longrightarrow_{\mathcal{L}} (A'\Omega_1 A)^{-1}A'M^*,$$

(*b*) $$\sqrt{n}vec[(\dot{A}(\dot{B}\bar{B}') - A), (\dot{\Pi}_1 - \Pi_1), \ldots, (\dot{\Pi}_{s-1} - \Pi_{s-1})] \longrightarrow_{\mathcal{L}} N(0, \Omega_2^{-1}\Omega_2^*\Omega_2^{-1}),$$

where the remaining variables are defined as in Lemma 2.1. □

From Theorem 3.2 above, the asymptotic distribution of the nonstationary component argument in (a) is independent of that of the stationary component argument in (b). Similar to the arguments at the end of the last sub-section, this result depends on the variance model (1.3)-(1.4) and the symmetry assumption.

3.3 Testing for Reduced Rank

This section applies the asymptotic distributions in Theorems 3.1 and 3.2 to construct tests for reduced rank. The null and the alternative hypotheses are:

$$H_0 : rank(C) = r < m \text{ vs } H_a : rank(C) = m. \tag{3.16}$$

We first consider the Likelihood Ratio-Type (LRT) test statistic:

$$LRT_G \equiv vec(\dot{C} - \dot{A}\dot{B})'(-\sum_{t=1}^{n} \hat{F}_{11t})vec(\dot{C} - \dot{A}\dot{B}), \tag{3.17}$$

where we recall that $\dot{C}$ is the full rank estimator defined in Sub-section 3.1, $\dot{A}$ and $\dot{B}$ are the reduced rank estimators defined in Sub-section 3.2, while $\hat{F}_{11t} = -(Y_{t-1}Y_{t-1}' \otimes \hat{V}_{t-1}^{-1})$. The following lemma gives the asymptotic distribution of LRT_G.

Lemma 3.3. *Suppose Assumptions 2.1-2.5 hold. Then under H_0 in (3.16), the LRT test statistic for rank,*

$$LRT_G \longrightarrow_{\mathcal{L}} tr[(\int_0^1 B_d(u)dV_d^*(u)')'(\int_0^1 B_d(u)B_d(u)'du)^{-1}(\int_0^1 B_d(u)dV_d^*(u)')],$$

where $V_d^*(u) = \Upsilon B_d(u) + [(A_\perp'\Omega_1^{-1}A_\perp)^{-1/2}A_\perp'\Omega_1^{-1}\Omega_1^*\Omega_1^{-1}A_\perp(A_\perp'\Omega_1^{-1}A_\perp)^{-1/2} - \Upsilon^{**}\Upsilon^{**\prime}]^{1/2}$ $V_d(u)$, $\Upsilon^{**} = (A_\perp'\Omega_1^{-1}A_\perp)^{-1/2}(A_\perp'\Omega_1^{-1}E(V_{t-1}^{-1}\varepsilon_t\varepsilon_t')A_\perp)(A_\perp'V^*A_\perp)^{-1/2}$, and $(B_d'(u), V_d'(u))'$ is a $2d-$dimensional standard Brownian motion. □

When the ε_t's are conditional homoskedastic, $E(V_{t-1}^{-1}\varepsilon_t\varepsilon_t') = I_m$, $\Omega_1^* = \Omega_1 = V^{*-1}$, and hence $\Upsilon = I_d$ and $V_d^*(u) = B_d(u)$. The distribution of LRT_G is exactly the same as that in Reinsel and Ahn (1992) and Johansen (1996). When $E(\varepsilon_t\varepsilon_t' \mid \mathcal{I}_{t-1}) = V_{t-1}$ and V_{t-1} may or may not equal to V^*, the distribution of LRT_G can be simplified as follows.

Theorem 3.3. *Suppose the assumptions in Lemma 3.3 hold. If $E(\varepsilon_t\varepsilon_t' \mid \mathcal{I}_{t-1}) = V_{t-1}$, then*

$$LRT_G \longrightarrow_{\mathcal{L}} tr\{[\zeta(I_d - \Lambda_d)^{1/2} + \Phi\Lambda_d^{1/2}]'[\zeta(I_d - \Lambda_d)^{1/2} + \Phi\Lambda_d^{1/2}]\}, \tag{3.18}$$

where Λ_d is a diagonal matrix containing the d eigenvalues of $(I_d - \Upsilon\Upsilon')$, where $\Upsilon = (A_\perp'\Omega_1^{-1}A_\perp)^{1/2}(A_\perp'V^*A_\perp)^{-1/2}$, $\Phi \sim N(0, I_d)$ and is independent of $\zeta = (\int_0^1 B_d(u)B_d(u)'du)^{-1/2}$ $\cdot (\int_0^1 B_d(u)dB_d(u)')$. □

On the other hand, when $E(\varepsilon_t\varepsilon_t' \mid \mathcal{I}_{t-1}) \neq V_{t-1}$, we may define a modified LRT test statistic:

$$LRT_G^* \equiv [vec(\dot{C}^*) - vec(\dot{A}\dot{B}^*)]'[-\sum_{t=1}^{n} \hat{F}_{11t}^*][vec(\dot{C}^*) - vec(\dot{A}\dot{B}^*)], \tag{3.19}$$

where $vec(\dot{C}^*) = (-\sum_{t=1}^n \hat{F}_{11t}^*)^{-1}(-\sum_{t=1}^n \hat{F}_{11t})vec(\dot{C})$, $\dot{B}^* = (\dot{A}'\dot{\Omega}_1^*\dot{A})^{-1}(\dot{A}'\dot{\Omega}_1\dot{A})\dot{B}$, and $\hat{F}_{11t}^* = -(Y_{t-1}Y_{t-1}' \otimes \hat{V}_{t-1}^{-1}\hat{\varepsilon}_t\hat{\varepsilon}_t'\hat{V}_{t-1}^{-1})$, $\hat{F}_{11t} = -(Y_{t-1}Y_{t-1}' \otimes \hat{V}_{t-1}^{-1})$, where we recall from Sub-section 3.1 that $\hat{\varepsilon}_t \equiv W_t - \hat{C}Y_{t-1} - \sum_{j=1}^{s-1}\hat{\Pi}_j W_{t-j}$, and $\hat{V}_{t-1}$ is defined around (3.7). The following corollary gives the asymptotic distribution of LRT_G^*.

Corollary 3.3. *Suppose the assumptions in Lemma 3.3 hold.*

$$LRT_G^* \longrightarrow_{\mathcal{L}} tr\{[\zeta(I_d - \Lambda_d^*)^{1/2} + \Phi\Lambda_d^{*1/2}]'[\zeta(I_d - \Lambda_d^*)^{1/2} + \Phi\Lambda_d^{*1/2}]\}, \tag{3.20}$$

where Λ_d^* is a diagonal matrix containing the d eigenvalues of $(I_d - \Upsilon^*\Upsilon^{*\prime})$, where $\Upsilon^* = (A_\perp'\Omega_1^{*-1}A_\perp)^{-1/2}(A_\perp'\Omega_1^{*-1}E(V_{t-1}^{-1}\varepsilon_t\varepsilon_t')A_\perp)(A_\perp'V^*A_\perp)^{-1/2}$, and the remaining variables are defined as in Theorem 3.3. □

Some of the critical values for the distributions in (3.18) or (3.20) are tabulated in Section 5 and Appendix B of Sin and Ling (2004). Refer to Theorem 3.3 and Corollary 3.3. In actual empirical applications, one needs to estimate the d eigenvalues of either $I_d - \Upsilon\Upsilon'$ or those of $I_d - \Upsilon^*\Upsilon^*$, which involves V^*, $A_\perp$, $E(V_{t-1}^{-1}\varepsilon_t\varepsilon_t'V_{t-1}^{-1})$, $E(V_{t-1}^{-1}\varepsilon_t\varepsilon_t')$, and $E(V_{t-1}^{-1})$. V^* (see around (3.7) above) can be consistently estimated by $n^{-1}\sum_{t=1}^n \hat{\varepsilon}_t\hat{\varepsilon}_t'$. Similarly, $A_\perp$ (see around (2.2) above) can be consistently estimated by $(I_m - c(\dot{A}'c)^{-1}\dot{A}')c_\perp$, where $c = (I_r, 0_{rxd})'$ and $c_\perp = (0_{dxr}, I_d)'$. See p.48 of Johansen (1996) for details. Further, $E(V_{t-1}^{-1}\varepsilon_t\varepsilon_t'V_{t-1}^{-1})$, $E(V_{t-1}^{-1}\varepsilon_t\varepsilon_t')$ and $E(V_{t-1}^{-1})$ can be consistently estimated by $n^{-1}\sum_{t=1}^n \hat{V}_{t-1}^{-1}\hat{\varepsilon}_t\hat{\varepsilon}_t'\hat{V}_{t-1}^{-1}$, $n^{-1}\sum_{t=1}^n \hat{V}_{t-1}^{-1}\hat{\varepsilon}_t\hat{\varepsilon}_t'$, and $n^{-1}\sum_{t=1}^n \hat{V}_{t-1}^{-1}$ respectively, where .

4 Allowing Asymmetric Distribution: With A Modified Weighting Matrix

In this section, we relax the symmetry assumption and consider the full rank estimation, the reduced rank estimation, as well as the test for reduced rank, parallel to those in Section 3. In Section 3, the only place we need symmetry is $E[U'_{t-1} \otimes V_{t-1}^{-1}] = 0$. In view of this, we consider a *modified* weighting matrix, which is denoted as G_{t-1} when it is evaluated at the *pseudo true* parameters (α, δ). G_{t-1} is defined such that:

$$\nu(G_{t-1}) \equiv \nu(V_{t-1}^{-1}) - E[\nu(V_{t-1}^{-1})U'_{t-1}][EU_{t-1}U'_{t-1}]^{-1}U_{t-1}, \tag{4.1}$$

where $\nu(.)$ is obtained from $vec(.)$ by eliminating all supradiagonal elements of a square matrix (see Magnus, 1988, p.27). It should be noted that by construction, $E[\nu(G_{t-1})U'_{t-1}] = 0$. Further, as $E(U_{t-1}) = 0$, $E(G_{t-1}) = E(V_{t-1}^{-1})$.

Refer to Sub-sections 3.1 and 3.2. Given the initial estimators $(\hat{\varphi}, \hat{\delta}) \in \Theta_n$ and $(\hat{\alpha}, \hat{\delta}) \in \Xi_n$, denote the sample analogue of V_{t-1}^{-1} and U_{t-1} as $\hat{V}_{t-1}^{-1}$ and $\hat{U}_{t-1}$ respectively, where similar to that in (1.3)-(1.4) and Theorem 3.1:

$$\hat{V}_{t-1}^{-1} = V_{t-1}^{-1}(\hat{\varphi}, \hat{\delta}), \tag{4.2}$$

$$\hat{U}_{t-1} = U_{t-1}(\hat{\alpha}, \hat{\delta}) = [(\hat{B}Y_{t-1})', W'_{t-1}, \cdots, W'_{t-s+1}]'. \tag{4.3}$$

In fact, under the null that $rank(C) = r$, $V_{t-1}^{-1}(\hat{\varphi}, \hat{\delta})$ in (4.2) can be replaced by $V_{t-1}^{-1}(\hat{\alpha}, \hat{\delta})$. Given (4.2)-(4.3), we can form a sample analogue of G_{t-1}, denoted as $\hat{G}_{t-1}$, where:

$$\nu(\hat{G}_{t-1}) \equiv \nu(\hat{V}_{t-1}^{-1}) - [\sum_{t=1}^{n} \nu(\hat{V}_{t-1}^{-1})\hat{U}'_{t-1}][\sum_{t=1}^{n} \hat{U}_{t-1}\hat{U}'_{t-1}]^{-1}\hat{U}_{t-1}. \tag{4.4}$$

Given $\hat{\varphi}$ and $\hat{G}_{t-1}$ (see (4.4)), similar to (3.7) and abusing the notation, we perform one iteration of the full rank estimator:

$$\dot{\varphi} = \hat{\varphi} - [\sum_{t=1}^{n} \hat{F}_t]^{-1}[\sum_{t=1}^{n} \hat{f}_t], \tag{4.5}$$

where $\hat{F}_t = -(X_{t-1}X'_{t-1} \otimes \hat{G}_{t-1})$, $\hat{f}_t = (X_{t-1} \otimes I_m)\hat{G}_{t-1}\hat{\varepsilon}_t$, where as in Sub-section 3.1, $\hat{\varepsilon}_t = W_t - \hat{C}Y_{t-1} - \sum_{j=1}^{s-1} \hat{\Pi}_j W_{t-j}$.

In order to derive the asymptotic distribution of $\dot{\varphi}$, once again abusing the notation, we let $(W'_m(u), W^{*\prime}_m(u))'$ be a $2m-$dimensional Brownian motion (BM) with the covariance matrix:

$$u\Omega \equiv u \begin{pmatrix} V^* & E(\varepsilon_t\varepsilon_t' G_{t-1}) \\ E(G_{t-1}\varepsilon_t\varepsilon_t') & \Omega_1^* \end{pmatrix},$$

where $V^* = E\varepsilon_t\varepsilon_t'$, and $\Omega_1^* = E(G_{t-1}\varepsilon_t\varepsilon_t' G_{t-1})$. Let $B_d(u) = \Omega_{a_1}^{-1/2}[I_d, 0_{dxr}]\,\Omega_a^{1/2}V^{*-1/2}W_m(u)$, where $\Omega_a = E(a_t a_t')$ and $\Omega_{a_1} = [I_d, 0_{dxr}]\Omega_a[I_d, 0_{dxr}]'$. The following is proved in the Appendix.

Theorem 4.1. *Suppose Assumptions 2.1-2.4 hold. Then*

$$(a) \qquad n(\dot{C} - C)\bar{B}_\perp' \longrightarrow_{\mathcal{L}} \Omega_1^{-1}M^*,$$

$$(b) \qquad \sqrt{n}vec[(\dot{C} - C)\bar{B}', (\dot{\Pi}_1 - \Pi_1), \ldots, (\dot{\Pi}_{s-1} - \Pi_{s-1})] \longrightarrow_{\mathcal{L}} N(0, \Omega_2^{-1}\Omega_2^*\Omega_2^{-1}),$$

where $\psi_{11} \equiv [I_d, 0](\sum_{k=1}^{\infty}\psi_k)[I_d, 0]'$, $\Omega_1 \equiv E(V_{t-1}^{-1})$, $\Omega_2 \equiv E(U_{t-1}U_{t-1}' \otimes G_{t-1})$, $\Omega_2^* \equiv E(U_{t-1}U_{t-1}' \otimes G_{t-1}\varepsilon_t\varepsilon_t' G_{t-1})$, $U_{t-1} = [(BY_{t-1})', W_{t-1}', \cdots, W_{t-s+1}']'$, *and* $M^* = (\int_0^1 B_d(u)dW_m^*(u)')'(\int_0^1 B_d(u)B_d(u)'du)^{-1}\Omega_{a_1}^{-1/2}\psi_{11}^{-1}$. □

It should be noted that the definitions of Ω_1^*, Ω_2 and Ω_2^* are essentially the same as those in Sub-section 3.1, with V_{t-1}^{-1} replaced by G_{t-1}. The definitions of other variables (in particular that of Ω_1) remain unchanged though.

Next we turn to the reduced rank estimation. Given $\hat{\alpha}$ and $\hat{G}_{t-1}$ (see (4.4)), similar to (3.14)-(3.15) and abusing the notation, we perform one iteration of the reduced rank estimators:

$$\dot{\alpha}_1 = \hat{\alpha}_1 - [\sum_{t=1}^{n}\hat{R}_{1t}]^{-1}[\sum_{t=1}^{n}\hat{r}_{1t}], \tag{4.6}$$

$$\dot{\alpha}_2 = \hat{\alpha}_2 - [\sum_{t=1}^{n}\dot{R}_{2t}]^{-1}[\sum_{t=1}^{n}\dot{r}_{2t}], \tag{4.7}$$

where $\hat{R}_{1t} = -(Y_{t-1}Y_{t-1}' \otimes \hat{A}'\hat{G}_{t-1}\hat{A})$, $\dot{R}_{2t} = -(\dot{U}_{t-1}\dot{U}_{t-1}' \otimes \hat{G}_{t-1})$;
$\hat{r}_{1t} = (Y_{t-1} \otimes \hat{A}')\hat{G}_{t-1}\hat{\varepsilon}_t$, and $\dot{r}_{2t} = (\dot{U}_{t-1} \otimes I_m)\hat{G}_{t-1}\dot{\varepsilon}_t$, where as in Sub-section 3.2, $\dot{U}_{t-1} = [(\dot{B}Y_{t-1})', W_{t-1}', \ldots, W_{t-s+1}']'$ and $\dot{\varepsilon}_t = W_t - \hat{A}\dot{B}Y_{t-1} - \sum_{j=1}^{s-1}\hat{\Pi}_j W_{t-j}$.

Refer to (4.6)-(4.7), the asymptotic distributions of the *normalized* estimators for α_1 and α_2 are given as follows.

Theorem 4.2. *Suppose Assumptions 2.1-2.4 hold. Then*

$$(a) \qquad n((\dot{B}\bar{B}')^{-1}\dot{B} - B)\bar{B}'_{\perp} \longrightarrow_{\mathcal{L}} (A'\Omega_1 A)^{-1}A'M^*,$$

$$(b) \quad \sqrt{n}vec[(\dot{A}(\dot{B}\bar{B}') - A), (\dot{\Pi}_1 - \Pi_1), \ldots, (\dot{\Pi}_{s-1} - \Pi_{s-1})] \longrightarrow_{\mathcal{L}} N(0, \Omega_2^{-1}\Omega_2^*\Omega_2^{-1}),$$

where the remaining variables are defined as in Theorem 4.1. □

Given Theorems 4.1 and 4.2, parallel to Sub-section 3.3, we may define a modified LRT test statistic for reduced rank, where abusing notation:

$$LRT_G^* \equiv [vec(\dot{C}^*) - vec(\dot{A}\dot{B}^*)]'[-\sum_{t=1}^{n} \hat{F}_{11t}^*][vec(\dot{C}^*) - vec(\dot{A}\dot{B}^*)], \tag{4.8}$$

where $vec(\dot{C}^*) = (-\sum_{t=1}^n \hat{F}_{11t}^*)^{-1}(-\sum_{t=1}^n \hat{F}_{11t})vec(\dot{C})$, $\dot{B}^* = (\dot{A}'\dot{\Omega}_1^*\dot{A})^{-1}(\dot{A}'\dot{\Omega}_1\dot{A})\dot{B}$, $\hat{F}_{11t} = -(Y_{t-1}Y_{t-1}' \otimes \hat{G}_{t-1})$, and $\hat{F}_{11t}^* = -(Y_{t-1}Y_{t-1}' \otimes \hat{G}_{t-1}\hat{\varepsilon}_t\hat{\varepsilon}_t\hat{G}_{t-1})$.

The following corollary gives the asymptotic distribution of LRT_G^*.

Corollary 4.3. *Suppose Assumptions 2.1-2.4 hold. Then under H_0 in (3.16), the modified LR test statistic for reduced rank,*

$$LRT_G^* \longrightarrow_{\mathcal{L}} tr\{[\zeta(I_d - \Lambda_d^*)^{1/2} + \Phi\Lambda_d^{*1/2}]'[\zeta(I_d - \Lambda_d^*)^{1/2} + \Phi\Lambda_d^{*1/2}]\}, \tag{4.9}$$

where Λ_d^* is a diagonal matrix containing the d eigenvalues of $(I_d - \Upsilon^*\Upsilon^{*\prime})$, $\Upsilon^* = (A_\perp'\Omega_1^{*-1}A_\perp)^{-1/2}(A_\perp'\Omega_1^{*-1}E(G_{t-1}\varepsilon_t\varepsilon_t')A_\perp)(A_\perp'V^*A_\perp)^{-1/2}$, where we recall that $\Omega_1^* = E(G_{t-1}\varepsilon_t\varepsilon_t'G_{t-1})$ and the remaining variables are defined as in Theorem 4.1. □

5 Allowing Asymmetric Distribution: Without Modifying the Weighting Matrix

Similar to the previous section, this section also relaxes the symmetry assumption and consider the full rank estimation, the reduced rank estimation, as well as the test for reduced rank, parallel to those in Section 3. Nevertheless, in order to maintain the original the weighting matrix, we consider more involved procedures and more elaborate asymptotic distributions.

Consider the full rank estimation and refer to Sub-section 3.1. Given the initial estimators $(\hat{\varphi}, \hat{\delta}) \in \Theta_n$, we perform a one-step iteration of the full rank estimator:

$$\dot{\varphi} = \hat{\varphi} - [\sum_{t=1}^{n} F_t(\hat{\varphi}, \hat{\delta})]^{-1}[\sum_{t=1}^{n} f_t(\hat{\varphi}, \hat{\delta})]. \tag{5.1}$$

It should be noted that the procedure in (5.1) is exactly the same as that in (3.7). Nevertheless, due to the possibly asymmetric distribution, the asymptotic distribution is different, as one can see in the next theorem.

Theorem 5.1. *Suppose Assumptions 2.1-2.4 hold. Then*

$$D^* Q^{*'-1}(\dot{\varphi} - \varphi) \longrightarrow_{\mathcal{L}} \begin{pmatrix} Z \otimes \Omega_1 & (L \otimes I_m)\Sigma \\ \Sigma'(L' \otimes I_m) & \Omega_2 \end{pmatrix}^{-1} \begin{pmatrix} W^* \\ \Omega_2^{*1/2}\Phi \end{pmatrix},$$

where $Z = \psi_{11}\Omega_{a_1}^{1/2}(\int_0^1 B_d(u)B_d(u)'du)\Omega_{a_1}^{1/2}\psi_{11}'$, $L = \psi_{11}\Omega_{a_1}^{1/2}(\int_0^1 B_d(u)du)$, $\Sigma = E(U_{t-1}' \otimes V_{t-1}^{-1})$, $W^* = (\psi_{11}\Omega_{a_1}^{1/2} \otimes I_m)vec[(\int_0^1 B_d dW_m^*(u)')']$, *and the remaining variables are defined as in Lemma 2.1.* □

Next we turn to the reduced rank estimation. Refer to the ECM in (3.8), as $BB_\perp' = 0$,

$$V_{t-1}^{-1/2}W_t = V_{t-1}^{-1/2}\varepsilon_t + V_{t-1}^{-1/2}ABY_{t-1} + \sum_{j=1}^{s-1} V_{t-1}^{-1/2}\Pi_j W_{t-j} + V_{t-1}^{-1/2}AB\bar{B}_\perp' B_\perp Y_{t-1}$$

$$\Longrightarrow \quad V_{t-1}^{-1/2}W_t = V_{t-1}^{-1/2}\varepsilon_t + (B_\perp Y_{t-1} \otimes A'V_{t-1}^{-1/2})vec(B\bar{B}_\perp') + (U_{t-1} \otimes V_{t-1}^{-1/2})\alpha_2,$$

where we recall from Sections 2 and 3 that $\bar{B}_\perp' = B_\perp'(B_\perp B_\perp')^{-1}$ and $U_{t-1} = [(BY_{t-1})', W_{t-1}', \cdots, W_{t-s+1}']'$.

However in deriving the LRT test for reduced rank, we need a $n-$*consistent* estimator for B itself. Note $vec(B\bar{B}_\perp') = [(B_\perp B_\perp')^{-1}B_\perp \otimes I_r]vec(B)$. Define $\hat{B}_\perp \equiv c_\perp'(I_m - \hat{B}'(c'\hat{B}')^{-1}c')$, where $c = (I_r, 0_{rxd})'$ and $c_\perp = (0_{dxr}, I_d)'$. See p.48 of Johansen (1996) for details. $\hat{B}_\perp$ is an $n-$*consistent* estimator for $B_\perp$. Abusing the notation, a natural estimator for α is:

$$diag(\hat{B}_\perp' \otimes I_r, I_{rm+(s-1)m^2})(\sum_{t=1}^{n} \hat{P}_t)^{-1}(\sum_{t=1}^{n} \hat{\phi}_t), \tag{5.2}$$

where

$$\hat{P}_t = \begin{pmatrix} \hat{B}_\perp Y_{t-1}Y_{t-1}'\hat{B}_\perp' \otimes \hat{A}'\hat{V}_{t-1}^{-1}\hat{A} & \hat{B}_\perp Y_{t-1}\hat{U}_{t-1}' \otimes \hat{A}'\hat{V}_{t-1}^{-1} \\ \hat{U}_{t-1}Y_{t-1}'\hat{B}_\perp' \otimes \hat{V}_{t-1}^{-1}\hat{A} & \hat{U}_{t-1}\hat{U}_{t-1}' \otimes \hat{V}_{t-1}^{-1} \end{pmatrix},$$

$$\hat{\phi}_t = \begin{pmatrix} \hat{B}_\perp Y_{t-1} \otimes \hat{A}'\hat{V}_{t-1}^{-1} W_t \\ \hat{U}_{t-1} \otimes \hat{V}_{t-1}^{-1} W_t \end{pmatrix},$$

Following the lines in the previous sections and abusing the notation, as an alternative to (5.2), we can estimate α_1 and α_2 separately, where:

$$\dot{\alpha}_1 = \hat{\alpha}_1 - [\hat{B}_\perp' \otimes I_r, 0_{mrx(rm+(s-1)m^2)}][\sum_{t=1}^n P_t(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\delta})]^{-1}[\sum_{t=1}^n p_t(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\delta})], \quad (5.3)$$

$$\dot{\alpha}_2 = \hat{\alpha}_2 - [0_{(rm+(s-1)m^2)xmd}, I_{rm+(s-1)m^2}][\sum_{t=1}^n P_t(\dot{\alpha}_1, \hat{\alpha}_2, \hat{\delta})]^{-1}[\sum_{t=1}^n p_t(\dot{\alpha}_1, \hat{\alpha}_2, \hat{\delta})], (5.4)$$

where for a *generic* $(\tilde{\alpha}, \tilde{\delta})$,

$$P_t(\tilde{\alpha}, \tilde{\delta}) = \begin{pmatrix} \tilde{B}_\perp Y_{t-1} Y_{t-1}' \tilde{B}_\perp' \otimes \tilde{A}'\tilde{V}_{t-1}^{-1}\tilde{A} & \tilde{B}_\perp Y_{t-1}\tilde{U}_{t-1}' \otimes \tilde{A}'\tilde{V}_{t-1}^{-1} \\ \tilde{U}_{t-1} Y_{t-1}' \tilde{B}_\perp' \otimes \tilde{V}_{t-1}^{-1}\tilde{A} & \tilde{U}_{t-1}\tilde{U}_{t-1}' \otimes \tilde{V}_{t-1}^{-1} \end{pmatrix},$$

$$p_t(\tilde{\alpha}, \tilde{\delta}) = \begin{pmatrix} \tilde{B}_\perp Y_{t-1} \otimes \tilde{A}'\tilde{V}_{t-1}^{-1}\tilde{\varepsilon}_t \\ \tilde{U}_{t-1} \otimes \tilde{V}_{t-1}^{-1}\tilde{\varepsilon}_t \end{pmatrix}.$$

Theorem 5.2. *Suppose Assumptions 2.1-2.4 hold. Then*

$$D^{**}Q^{**'-1}(\ddot{\alpha} - \alpha) \longrightarrow_{\mathcal{L}} \begin{pmatrix} Z \otimes A'\Omega_1 A & (L \otimes A')\Sigma \\ \Sigma'(L' \otimes A) & \Omega_2 \end{pmatrix}^{-1} \begin{pmatrix} (I_d \otimes A')W^* \\ \Omega_2^{*1/2}\Phi \end{pmatrix},$$

where $\ddot{\alpha} = (\ddot{\alpha}_1', \ddot{\alpha}_2')'$, $\ddot{\alpha}_1 = vec((\dot{B}\bar{B}')^{-1}\dot{B})$ *and* $\ddot{\alpha}_2 = vec(\dot{A}(\dot{B}\bar{B}'), \dot{\Pi}_1, \ldots, \dot{\Pi}_{s-1})$, *and the remaining variables are defined as in Lemma 2.1 or Theorem 5.1.* □

Refer to the null and the alternative hypotheses in (3.16). Similar to (3.17) and abusing the notation, we first consider the Likelihood Ratio-Type (LRT) test statistic:

$$LRT_G \equiv [\dot{\varphi} - diag(I_m \otimes \dot{A}, I_{(s-1)m^2})\dot{\chi}]'[-\sum_{t=1}^n \hat{F}_t][\dot{\varphi} - diag(I_m \otimes \dot{A}, I_{(s-1)m^2})\dot{\chi}], \quad (5.5)$$

where $\dot{\chi} \equiv (\dot{\alpha}_1', \dot{\alpha}_{22}')'$, $\dot{\alpha}_{22}$ is the reduced-rank estimator for $\alpha_{22} \equiv vec(\Pi_1, \ldots, \Pi_{s-1})$. Compare (3.17) with (5.5). The inclusion of parameters other than A and B is due to the possible asymmetric distribution. Similar to the proof of Lemma 3.3, it can be shown that:

$$\begin{aligned} & D^* Q^{*'-1}[diag(I_m \otimes \dot{A}, I_{(s-1)m^2})\dot{\chi} - \varphi] \\ = \; & D^* Q^{*'-1}[diag(I_m \otimes \dot{A}(\dot{B}\bar{B}'), I_{(s-1)m^2})(\ddot{\alpha}_1', \dot{\alpha}_{22}') - \varphi] \\ = \; & diag(I_d \otimes A, I_{(s-1)m^2}) D^{**}Q^{**'-1}(\ddot{\alpha} - \alpha) + o_p(1). \end{aligned}$$

As a result,

$$\begin{aligned}
& [D^*Q^{*'-1}(diag(I_m \otimes \dot{A}, I_{(s-1)m^2})\dot{\chi} - \varphi)]'[-D^{*-1}Q^* \sum_{t=1}^{n} \hat{F}_t Q^{*'} D^{*-1}] \\
& \cdot[D^*Q^{*'-1}(diag(I_m \otimes \dot{A}, I_{(s-1)m^2})\dot{\chi} - \varphi)] \\
= \; & [D^{**}Q^{**'-1}(\ddot{\alpha} - \alpha)]'[-D^{**-1}Q^{**} \sum_{t=1}^{n} R_t Q^{**'} D^{**-1}][D^{**}Q^{**'-1}(\ddot{\alpha} - \alpha)] + o_p(1). \quad (5.6)
\end{aligned}$$

In a similar token,

$$\begin{aligned}
& [D^*Q^{*'-1}(diag(I_m \otimes \dot{A}, I_{(s-1)m^2})\dot{\chi} - \varphi)]'[-D^{*-1}Q^* \sum_{t=1}^{n} \hat{F}_t Q^{*'} D^{*-1}][D^*Q^{*'-1}(\dot{\varphi} - \varphi)] \\
= \; & [D^{**}Q^{**'-1}(\ddot{\alpha} - \alpha)]'[-D^{**-1}Q^{**} \sum_{t=1}^{n} R_t Q^{**'} D^{**-1}][D^{**}Q^{**'-1}(\ddot{\alpha} - \alpha)] + o_p(1). \quad (5.7)
\end{aligned}$$

Therefore, by Theorems 5.1 and 5.2,

$$\begin{aligned}
& LRT_G \\
= \; & [D^*Q^{*'-1}(\dot{\varphi} - \varphi)]'[-D^{*-1}Q^* \sum_{t=1}^{n} F_t Q^{*'} D^{*-1}][D^*Q^{*'-1}(\dot{\varphi} - \varphi)] \\
& -[D^{**}Q^{**'-1}(\ddot{\alpha} - \alpha)]'[-D^{**-1}Q^{**} \sum_{t=1}^{n} R_t Q^{**'} D^{**-1}][D^{**}Q^{**'-1}(\ddot{\alpha} - \alpha)] + o_p(1) \\
\longrightarrow_{\mathcal{L}} \; & \begin{pmatrix} w^* \\ \Phi \end{pmatrix}' \begin{pmatrix} z \otimes \Omega_1^{*-1/2}\Omega_1\Omega_1^{*-1/2} & (l \otimes \Omega_1^{*-1/2})\Sigma\Omega_2^{*-1/2} \\ \Omega_2^{*-1/2}\Sigma'(l' \otimes \Omega_1^{*-1/2}) & \Omega_2^{*-1/2}\Omega_2\Omega_2^{*-1/2} \end{pmatrix}^{-1} \begin{pmatrix} w^* \\ \Phi \end{pmatrix} \\
& - \begin{pmatrix} w^* \\ \Phi \end{pmatrix}' \mathcal{A}^*[\mathcal{A}^{*'} \begin{pmatrix} z \otimes \Omega_1^{*-1/2}\Omega_1\Omega_1^{*-1/2} & (l \otimes \Omega_1^{*-1/2})\Sigma\Omega_2^{*-1/2} \\ \Omega_2^{*-1/2}\Sigma'(l' \otimes \Omega_1^{*-1/2}) & \Omega_2^{*-1/2}\Omega_2\Omega_2^{*-1/2} \end{pmatrix} \mathcal{A}^*]^{-1} \\
& \cdot \mathcal{A}^* \begin{pmatrix} w^* \\ \Phi \end{pmatrix}, \qquad (5.8)
\end{aligned}$$

where $w^* = vec[(\int_0^1 B_d dw_m^*(u)')']$, $w_m^*(u) = \Omega_1^{*-1/2} W_m^*(u)$, and $\mathcal{A}^* = diag(I_d \otimes \Omega_1^{*1/2} A, \Omega_1^{*1/2})$, $z = \int_0^1 B_d(u)B_d(u)'du$, and $l = \int_0^1 B_d(u)du$.

Denoting

$$\begin{pmatrix} z \otimes \Omega_1^{*-1/2}\Omega_1\Omega_1^{*-1/2} & (l \otimes \Omega_1^{*-1/2})\Sigma\Omega_2^{*-1/2} \\ \Omega_2^{*-1/2}\Sigma'(l' \otimes \Omega_1^{*-1/2}) & \Omega_2^{*-1/2}\Omega_2\Omega_2^{*-1/2} \end{pmatrix}$$

as Ω^{**}, the following lemma gives the asymptotic distribution of the LRT_G in (5.5).

Lemma 5.3. *Suppose Assumptions 2.1-2.4 hold. Then under H_0 in (3.16), the LRT test statistic for rank in (5.5),*

$$LRT_G \longrightarrow_{\mathcal{L}} \begin{pmatrix} w^* \\ \Phi \end{pmatrix}' \Omega^{**-1}\mathcal{A}_\perp^*(\mathcal{A}_\perp^{*'}\Omega^{**-1}\mathcal{A}_\perp^*)^{-1}\mathcal{A}_\perp^{*'}\Omega^{**-1} \begin{pmatrix} w^* \\ \Phi \end{pmatrix}, \quad (5.9)$$

where $\mathcal{A}_\perp^{*'} = (I_d \otimes A'_\perp \Omega_1^{*-1/2}, 0_{d^2x(rm+(s-1)m^2)})$. $\square$

When the ε_t's are conditional homoskedastic, $E(V_{t-1}^{-1}\varepsilon_t\varepsilon_t') = I_m$, and $\Sigma = 0$, it can be shown that the distribution of LRT_G is exactly the same as that in Reinsel and Ahn (1992) and Johansen (1996). When $E(\varepsilon_t\varepsilon_t' \mid \mathcal{I}_{t-1}) = V_{t-1}$ and V_{t-1} may or may not equal to I_m, the distribution of LRT_G can be simplified as follows.

Theorem 5.3. *Suppose the assumptions in Lemma 5.3 hold. If* $E(\varepsilon_t\varepsilon_t' \mid \mathcal{I}_{t-1}) = V_{t-1}$, *then*

$$LRT_G \longrightarrow_{\mathcal{L}} \begin{pmatrix} w \\ \Phi \end{pmatrix}' \Omega^{-1}\mathcal{A}_\perp(\mathcal{A}'_\perp\Omega^{-1}\mathcal{A}_\perp)^{-1}\mathcal{A}'_\perp\Omega^{-1}\begin{pmatrix} w \\ \Phi \end{pmatrix}, \tag{5.10}$$

where Ω *is defined as*

$$\begin{pmatrix} z \otimes I_m & (l \otimes \Omega_1^{-1/2})\Sigma\Omega_2^{-1/2} \\ \Omega_2^{-1/2}\Sigma'(l' \otimes \Omega_1^{-1/2}) & I_m \end{pmatrix},$$

and $\mathcal{A}'_\perp = (I_d \otimes A'_\perp\Omega_1^{-1/2}, 0_{d^2x(rm+(s-1)m^2)})$, $w = vec[(\int_0^1 B_d dw_m(u)')']$, $w_m(u) = \Omega_1^{-1/2}W_m^*(u)$. $\square$

On the other hand, when $E(\varepsilon_t\varepsilon_t' \mid \mathcal{I}_{t-1}) \neq V_{t-1}$, we may define a modified LRT test statistic:

$$\begin{aligned} LRT_G^* \equiv\ & [\dot{\varphi} - diag(I_m \otimes \dot{A}, I_{(s-1)m^2})\dot{\chi}]'[-\sum_{t=1}^n \hat{F}_t][-\sum_{t=1}^n \hat{F}_t^*]^{-1} \\ & \cdot[-\sum_{t=1}^n \hat{F}_t][\dot{\varphi} - diag(I_m \otimes \dot{A}, I_{(s-1)m^2})\dot{\chi}], \end{aligned} \tag{5.11}$$

where $\hat{F}_t^* = -(X_{t-1}X'_{t-1} \otimes \hat{V}_{t-1}^{-1}\hat{\varepsilon}_t\hat{\varepsilon}_t'\hat{V}_{t-1}^{-1})$. The following corollary gives the asymptotic distribution of LRT_G^*.

Corollary 5.3. *Suppose the assumptions in Lemma 5.3 hold. Then*

$$LRT_G^* \longrightarrow_{\mathcal{L}} \begin{pmatrix} w^* \\ \Phi \end{pmatrix}' \Omega^{*-1}\mathcal{A}_\perp^*(\mathcal{A}_\perp^{*'}\Omega^{*-1}\mathcal{A}_\perp)^{*-1}\mathcal{A}_\perp^{*'}\Omega^{*-1}\begin{pmatrix} w^* \\ \Phi \end{pmatrix}, \tag{5.12}$$

where Ω^* *is defined as*

$$\begin{pmatrix} z \otimes I_m & (l \otimes \Omega_1^{*-1/2})\Sigma\Omega_2^{*-1/2} \\ \Omega_2^{*-1/2}\Sigma'(l' \otimes \Omega_1^{*-1/2}) & I_m \end{pmatrix},$$

and all the other variables are defined above. $\square$

6 WLS versus QMLE: Efficiency and Computational Matters

7 Monte Carlo Experiments

8 An Empirical Example

9 Conclusions

Macroeconomic or financial data are often modelled with cointegration and GARCH. Noticeable examples include those studies of price discovery, in which stock prices of the same underlying asset are cointegrated and they exhibit multivariate GARCH. Modifying the asymptotic theories developed in Li, Ling and Wong (2001) and Sin and Ling (2004), this paper proposes a WLS (weighted least squares) for the parameters of an ECM (error-correction model). Apart from its computational simplicity, by construction, the consistency of WLS is insensitive to possible mis-specification in conditional variance. Further, asymmetrically distributed deflated error is allowed, at the expense of more deliberate estimation procedures. Efficiency loss relative to QMLE (quasi-maximum likelihood estimator) is discussed within the class of LABF (locally asymptotically Brownian functional) models. The insensitivity and efficiency of WLS in finite samples are examined through Monte Carlo experiments. We also apply the WLS to an empirical example of HSI (Hang Seng Index), HSIF (Hang Seng Index Futures) and TraHK (Hong Kong Tracker Fund).

A Appendix: Critical Values

TABLE A.1

Quantiles of the Limiting Distribution (5.3) or (5.5)

$d = 1$, no Constant Term

	$\alpha-$th simulated quantiles							
λ_1	.500	.750	.800	.850	.900	.950	.975	.990
0.0	0.602	1.550	1.891	2.343	2.995	4.153	5.357	7.018
0.1	0.575	1.539	1.869	2.315	2.978	4.140	5.365	6.941
0.2	0.553	1.511	1.850	2.308	2.964	4.138	5.362	6.939
0.3	0.533	1.489	1.824	2.282	2.941	4.108	5.305	6.921
0.4	0.515	1.462	1.800	2.254	2.914	4.083	5.286	6.929
0.5	0.499	1.441	1.770	2.223	2.883	4.043	5.242	6.895
0.6	0.490	1.414	1.743	2.197	2.845	4.013	5.225	6.824
0.7	0.481	1.385	1.718	2.171	2.811	3.963	5.174	6.839
0.8	0.470	1.364	1.693	2.139	2.782	3.920	5.097	6.774
0.9	0.461	1.354	1.674	2.105	2.746	3.867	5.047	6.718
1.0	0.455	1.326	1.649	2.078	2.711	3.827	5.068	6.633

The table values were computed from 100,000 simulations with $n = 2,000$.
λ_1 is the eigenvalue of Λ_1 in (5.3) or Λ_1^* in (5.5).

TABLE A.2

Quantiles of the Limiting Distribution (5.3) or (5.5)

$d = 2$, no Constant Term

		α−th simulated quantiles							
λ_1	λ_2	.500	.750	.800	.850	.900	.950	.975	.990
0.0	0.0	5.508	7.844	8.522	9.365	10.479	12.286	14.065	16.278
0.0	0.1	5.405	7.739	8.413	9.267	10.386	12.237	13.971	16.144
0.0	0.2	5.298	7.645	8.313	9.159	10.312	12.158	13.886	16.041
0.0	0.3	5.189	7.541	8.210	9.062	10.234	12.073	13.793	15.986
0.0	0.4	5.068	7.440	8.112	8.959	10.119	11.987	13.722	15.895
0.0	0.5	4.952	7.330	8.008	8.865	10.003	11.887	13.659	15.802
0.0	0.6	4.839	7.216	7.909	8.744	9.906	11.789	13.542	15.716
0.0	0.7	4.726	7.112	7.783	8.647	9.796	11.676	13.440	15.623
0.0	0.8	4.619	6.981	7.668	8.525	9.680	11.559	13.354	15.530
0.0	0.9	4.504	6.867	7.542	8.410	9.551	11.446	13.230	15.435
0.0	1.0	4.393	6.745	7.417	8.268	9.443	11.306	13.172	15.450
0.1	0.1	5.287	7.635	8.325	9.172	10.295	12.140	13.885	16.105
0.1	0.2	5.178	7.534	8.229	9.079	10.217	12.071	13.817	15.991
0.1	0.3	5.058	7.440	8.123	8.979	10.125	11.987	13.736	15.920
0.1	0.4	4.945	7.341	8.023	8.865	10.018	11.902	13.612	15.806
0.1	0.5	4.832	7.224	7.920	8.750	9.919	11.818	13.539	15.643
0.1	0.6	4.718	7.108	7.791	8.643	9.808	11.692	13.422	15.552
0.1	0.7	4.605	6.987	7.677	8.533	9.679	11.578	13.296	15.482
0.1	0.8	4.498	6.856	7.559	8.413	9.561	11.434	13.179	15.337
0.1	0.9	4.382	6.749	7.430	8.290	9.455	11.284	13.064	15.247
0.1	1.0	4.278	6.627	7.307	8.157	9.307	11.147	12.950	15.229
0.2	0.2	5.070	7.445	8.137	8.987	10.116	11.973	13.707	15.898
0.2	0.3	4.945	7.336	8.037	8.881	10.028	11.879	13.601	15.812
0.2	0.4	4.828	7.225	7.916	8.761	9.916	11.791	13.501	15.647
0.2	0.5	4.711	7.111	7.807	8.658	9.819	11.691	13.383	15.556
0.2	0.6	4.596	6.998	7.682	8.532	9.691	11.566	13.298	15.405
0.2	0.7	4.488	6.881	7.560	8.415	9.579	11.433	13.191	15.319
0.2	0.8	4.383	6.753	7.435	8.288	9.453	11.293	13.027	15.191
0.2	0.9	4.266	6.621	7.309	8.165	9.322	11.141	12.902	15.023
0.2	1.0	4.160	6.502	7.190	8.031	9.182	10.985	12.768	15.020
0.3	0.3	4.830	7.232	7.929	8.781	9.931	11.752	13.491	15.702
0.3	0.4	4.717	7.118	7.809	8.657	9.816	11.669	13.411	15.609
0.3	0.5	4.598	7.001	7.688	8.540	9.693	11.570	13.285	15.471

TABLE A.2 (Continued)

		$\alpha-$th simulated quantiles							
λ_1	λ_2	.500	.750	.800	.850	.900	.950	.975	.990
0.3	0.6	4.489	6.877	7.570	8.415	9.565	11.432	13.179	15.318
0.3	0.7	4.369	6.758	7.442	8.281	9.442	11.296	13.051	15.202
0.3	0.8	4.263	6.636	7.302	8.160	9.310	11.158	12.897	15.021
0.3	0.9	4.152	6.505	7.187	8.042	9.163	11.010	12.743	14.870
0.3	1.0	4.052	6.374	7.045	7.882	9.046	10.819	12.592	14.853
0.4	0.4	4.600	7.006	7.695	8.549	9.707	11.557	13.290	15.510
0.4	0.5	4.486	6.877	7.577	8.420	9.576	11.438	13.180	15.374
0.4	0.6	4.373	6.760	7.444	8.287	9.440	11.310	13.061	15.231
0.4	0.7	4.255	6.631	7.318	8.148	9.313	11.171	12.881	15.087
0.4	0.8	4.150	6.506	7.179	8.012	9.176	11.024	12.733	14.928
0.4	0.9	4.040	6.378	7.050	7.883	9.018	10.847	12.567	14.747
0.4	1.0	3.941	6.233	6.911	7.735	8.875	10.678	12.395	14.651
0.5	0.5	4.376	6.751	7.437	8.298	9.444	11.322	13.053	15.298
0.5	0.6	4.261	6.625	7.299	8.171	9.310	11.176	12.919	15.115
0.5	0.7	4.151	6.497	7.178	8.016	9.177	11.049	12.759	14.954
0.5	0.8	4.036	6.362	7.039	7.870	9.030	10.854	12.567	14.820
0.5	0.9	3.937	6.235	6.907	7.727	8.866	10.693	12.398	14.612
0.5	1.0	3.836	6.098	6.758	7.588	8.685	10.541	12.202	14.486
0.6	0.6	4.152	6.495	7.161	8.015	9.153	11.035	12.781	14.993
0.6	0.7	4.045	6.356	7.027	7.874	9.015	10.894	12.580	14.809
0.6	0.8	3.930	6.214	6.890	7.719	8.857	10.713	12.401	14.622
0.6	0.9	3.828	6.086	6.749	7.577	8.698	10.529	12.218	14.480
0.6	1.0	3.733	5.959	6.612	7.428	8.512	10.358	12.002	14.298
0.7	0.7	3.936	6.213	6.885	7.721	8.847	10.719	12.432	14.668
0.7	0.8	3.827	6.082	6.738	7.564	8.688	10.555	12.247	14.435
0.7	0.9	3.724	5.933	6.598	7.413	8.520	10.353	12.036	14.259
0.7	1.0	3.630	5.811	6.464	7.251	8.347	10.151	11.794	14.091
0.8	0.8	3.728	5.934	6.586	7.400	8.526	10.342	12.053	14.255
0.8	0.9	3.626	5.791	6.434	7.240	8.345	10.144	11.857	14.064
0.8	1.0	3.528	5.666	6.303	7.084	8.154	9.952	11.588	13.825
0.9	0.9	3.531	5.655	6.286	7.071	8.166	9.932	11.656	13.770
0.9	1.0	3.446	5.521	6.142	6.913	7.972	9.703	11.390	13.553
1.0	1.0	3.359	5.378	5.977	6.734	7.777	9.471	11.120	13.264

The table values were computed from $100,000$ simulations with $n = 2,000$.
$\lambda_1 \leq \lambda_2$ are the eigenvalues of Λ_2 in (5.3) or Λ_2^* in (5.5).

B Appendix: Technical Proofs

Lemma B.1. *Under the assumptions in Theorem 4.2, it follows that*

$$\begin{aligned}
&(a) \quad (\hat{B}\bar{B}')^{-1}(\dot{B}-\hat{B}) = O_p(n^{-1/2}),\\
&(b) \quad \hat{A}(\dot{B}\bar{B}') = \hat{A}(\hat{B}\bar{B}') + O_p(n^{-1/2}) = A + O_p(n^{-1/2}),\\
&(c) \quad (\dot{B}\bar{B}')^{-1}\hat{B}P_1 = (\hat{B}\bar{B}')^{-1}\hat{B}P_1 + O_p(n^{-3/2}) = BP_1 + O_p(n^{-1}),\\
&(d) \quad (\dot{B}\bar{B}')^{-1}\hat{B}P_2 = (\hat{B}\bar{B}')^{-1}\hat{B}P_2 + O_p(n^{-1/2}) = BP_2 + O_p(n^{-1/2}). \qquad \square
\end{aligned}$$

Proof. (a). We first denote $D_{\alpha_1} = diag(nI_{rd}, \sqrt{n}I_{r^2})$ and $\hat{Q}^{**} = \mathcal{Q}(I_m \otimes (\hat{B}\bar{B}')')$, with $\mathcal{Q} = (Q \otimes I_r)$. Also denote $\hat{\alpha}_1 = vec(\hat{B})$, $\check{\alpha}_1 = vec((\hat{B}\bar{B}')^{-1}\hat{B})$ and $\dot{\alpha}_1 = vec(\dot{B})$. $\hat{\alpha}_2$, $\check{\alpha}_2$ and $\dot{\alpha}_2$ are defined accordingly. $\hat{\alpha}$, $\check{\alpha}$ and $\dot{\alpha}$ are also defined accordingly. Since $\hat{Q}^{**\prime -1} = (P' \otimes I_r)(I_m \otimes (\hat{B}\bar{B}')^{-1})$, we have

$$\begin{aligned}
(I_m \otimes (\hat{B}\bar{B}')^{-1})(\dot{\alpha}_1 - \hat{\alpha}_1) &= \mathcal{Q}' D_{\alpha_1}^{-1} D_{\alpha_1} (P' \otimes I_r)(I_m \otimes (\hat{B}\bar{B}')^{-1})(\dot{\alpha}_1 - \hat{\alpha}_1)\\
&= \mathcal{Q}' D_{\alpha_1}^{-1} [D_{\alpha_1} \hat{Q}^{**\prime -1}(\dot{\alpha}_1 - \hat{\alpha}_1)].
\end{aligned}$$

As $\mathcal{Q}' D_{\alpha_1}^{-1} = O(n^{-1/2})$, it suffices to show $D_{\alpha_1}\hat{Q}^{**\prime -1}(\dot{\alpha}_1 - \hat{\alpha}_1) = O_p(1)$. By (4.9),

$$\begin{aligned}
&D_{\alpha_1}\hat{Q}^{**\prime -1}(\dot{\alpha}_1 - \hat{\alpha}_1)\\
&\quad = -[\sum_{t=1}^n D_{\alpha_1}^{-1}\hat{Q}^{**}(R_{1t}|_{\hat{\alpha},\dot{\delta}})\hat{Q}^{**\prime} D_{\alpha_1}^{-1}]^{-1}[\sum_{t=1}^n D_{\alpha_1}^{-1}\hat{Q}^{**}(r_{1t}|_{\hat{\alpha},\dot{\delta}})]\\
&\quad = -[\sum_{t=1}^n D_{\alpha_1}^{-1}\mathcal{Q}(R_{1t}|_{\check{\alpha},\dot{\delta}})\mathcal{Q}' D_{\alpha_1}^{-1}]^{-1}[\sum_{t=1}^n D_{\alpha_1}^{-1}\mathcal{Q}(r_{1t}|_{\check{\alpha},\dot{\delta}})].
\end{aligned}$$

By Theorem 4.1 and Theorem 3.1(c), $n(\check{\alpha}_1 - \alpha_1) = O_p(1)$, $\sqrt{n}(\check{\alpha}_2 - \alpha_2) = O_p(1)$, and $\sqrt{n}(\dot{\delta} - \delta) = O_p(1)$. Similar to the arguments for (4.7), it follows that:

$$\sum_{t=1}^n D_{\alpha_1}^{-1}\mathcal{Q}(R_{1t}|_{\check{\alpha},\dot{\delta}})\mathcal{Q}' D_{\alpha_1}^{-1} = \sum_{t=1}^n D_{\alpha_1}^{-1}\mathcal{Q}R_{1t}\mathcal{Q}' D_{\alpha_1}^{-1} + o_p(1). \qquad \text{(B. 1)}$$

On the other hand, by a Taylor's expansion and (B.1), with R_{1t}^* and r_{1t}^* being evaluated at a mid-point of $(\check{\alpha}, \dot{\delta})$ and (α, δ),

$$\sum_{t=1}^n D_{\alpha_1}^{-1}\mathcal{Q}(r_{1t}|_{\check{\alpha},\dot{\delta}})$$

$$= \sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q}r_{1t} + \sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q}(R_{1t}^*)(\check{\alpha}_1 - \alpha_1) + \sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q}(\nabla_{\alpha_2'} r_{1t}^*)(\check{\alpha}_2 - \alpha_2)$$
$$= \sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q}_1 r_{1t} + [\sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q}R_{1t}\mathcal{Q}'D_{\alpha_1}^{-1} + o_p(1)]\frac{1}{n}D_{\alpha_1}(P' \otimes I_r)[n(\check{\alpha}_1 - \alpha_1)]$$
$$+[\frac{1}{\sqrt{n}}\sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q}(\nabla_{\alpha_2'} r_{1t}^*)]\sqrt{n}(\check{\alpha}_2 - \alpha_2). \qquad \text{(B. 2)}$$

It is not difficult to show that $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q}(\nabla_{\alpha_2'} r_{1t}^*)$ is $O_p(1)$. So is the RHS of (B.2). By Lemmas 3.1(a)-(b), (B.1) and (B.2), (a) holds.

(b). By the $\sqrt{n}$-consistency of $\hat{A}(\hat{B}\bar{B}')$ for A, and (a) of this lemma,

$$\hat{A}(\dot{B}\bar{B}') = \hat{A}(\hat{B}\bar{B}') + \hat{A}(\hat{B}\bar{B}')(\hat{B}\bar{B}')^{-1}(\dot{B} - \hat{B})\bar{B}' = \hat{A}(\hat{B}\bar{B}') + O_p(1)O_p(n^{-1/2}).$$

Thus, (b) holds.

(c) and (d). Denote $\check{B} = (\hat{B}\bar{B}')^{-1}\hat{B}$.

$$(\dot{B}\bar{B}')^{-1}\hat{B} = [(\hat{B}\bar{B}')^{-1}\dot{B}\bar{B}']^{-1}(\hat{B}\bar{B}')^{-1}\hat{B} = [(\hat{B}\bar{B}')^{-1}\dot{B}\bar{B}']^{-1}\check{B}. \qquad \text{(B. 3)}$$

Using the formula $dF^{-1} = -F^{-1}(dF)F^{-1}$ for the $r \times r$ matrix F with $F(x) = [x\bar{B}]^{-1}$, and applying a Taylor's expansion to $[(\hat{B}\bar{B}')^{-1}\dot{B}\bar{B}']^{-1}$ around $\check{B}\bar{B}'$, we have

$$[(\hat{B}\bar{B}')^{-1}\dot{B}\bar{B}']^{-1} = [\check{B}\bar{B}']^{-1} - [B^*\bar{B}']^{-1}[(\hat{B}\bar{B}')^{-1}\dot{B} - \check{B}]\bar{B}'[B^*\bar{B}']^{-1},$$

where B^* lies between $(\hat{B}\bar{B}')^{-1}\dot{B}$ and $\check{B}$. Therefore, the RHS of (B.3) equals:

$$[(\hat{B}\bar{B}')^{-1}\hat{B}\bar{B}']^{-1}(\hat{B}\bar{B}')^{-1}\hat{B} - [B^*\bar{B}']^{-1}[(\hat{B}\bar{B}')^{-1}\dot{B} - \check{B}]\bar{B}'[B^*\bar{B}']^{-1}\check{B}$$
$$= (\hat{B}\bar{B}')^{-1}\hat{B} - [B^*\bar{B}']^{-1}[(\hat{B}\bar{B}')^{-1}\dot{B} - \check{B}]\bar{B}'[B^*\bar{B}']^{-1}\check{B}. \qquad \text{(B. 4)}$$

By (a) of this lemma, $(\hat{B}\bar{B}')^{-1}\dot{B} - \check{B} = O_p(n^{-1/2})$. From this, we can show that $[B^*\bar{B}']^{-1} = O_p(1)$. $\bar{B}$ and $\check{B}$ are also $O_P(1)$. By (B.4), (d) holds. By Theorem 4.1, $\check{B}P_1 = O_p(n^{-1})$ because $BP_1 = 0$. By (B.4),

$$[(\hat{B}\bar{B}')^{-1}\hat{B}\bar{B}']^{-1}(\hat{B}\bar{B}')^{-1}\hat{B}P_1 - [B^*\bar{B}']^{-1}[(\hat{B}\bar{B}')^{-1}\dot{B} - \check{B}]\bar{B}'[B^*\bar{B}']^{-1}\check{B}P_1$$
$$= (\hat{B}\bar{B}')^{-1}\hat{B}P_1 + O_p(n^{-3/2}).$$

Thus, (c) holds. This completes the proof. □

Proof of Theorem 4.2. Denote $\dot{Q}_1^{**}=(Q_1'\otimes I_r)(I_m\otimes(\dot{B}\bar{B}')')$, $\dot{Q}_2^{**}=diag((\dot{B}\bar{B}')^{-1}\otimes I_m, I_{(s-1)m^2})$, $\grave{\alpha}_1=vec((\dot{B}\bar{B}')^{-1}\hat{B})$, $\grave{\alpha}_2=vec[\hat{A}(\dot{B}\bar{B}'),\hat{\Phi}_1^*,\ldots,\hat{\Phi}_{s-1}^*]$, and $\grave{\alpha}=[\grave{\alpha}_1',\grave{\alpha}_2']'$. By Lemmas B.1(b)-(c), $(\grave{\alpha},\dot{\delta})\in\Xi_n$. Thus by (4.7),

$$
\begin{aligned}
n^{-2}\sum_{t=1}^{n}\dot{Q}_1^{**}(R_{1t}|_{\hat{\alpha},\dot{\delta}})\dot{Q}_1^{**\prime} &= n^{-2}\sum_{t=1}^{n}(Q_1'\otimes I_r)(R_{1t}|_{\grave{\alpha},\dot{\delta}})(Q_1\otimes I_r)\\
&= n^{-2}\sum_{t=1}^{n}(Q_1'\otimes I_r)R_{1t}(Q_1\otimes I_r)+o_p(1), \qquad (B.\ 5)\\
n^{-1}\sum_{t=1}^{n}\dot{Q}_2^{**}(R_{2t}|_{\hat{\alpha},\dot{\delta}})\dot{Q}_2^{**\prime} &= n^{-1}\sum_{t=1}^{n}(R_{2t}|_{\grave{\alpha},\dot{\delta}})=n^{-1}\sum_{t=1}^{n}R_{2t}+o_p(1). \qquad (B.\ 6)
\end{aligned}
$$

Refer to (4.6). Due to the block-diagonality of $\tilde{R}_t$, by (4.8),

$$
\begin{aligned}
&\frac{1}{n}\sum_{t=1}^{n}\dot{Q}_1^{**}(r_{1t}|_{\hat{\alpha},\dot{\delta}})=\frac{1}{n}\sum_{t=1}^{n}(Q_1'\otimes I_r)(r_{1t}|_{\grave{\alpha},\dot{\delta}})\\
=\ &\frac{1}{n}\sum_{t=1}^{n}(Q_1'\otimes I_r)r_{1t}+(\frac{1}{n}\sum_{t=1}^{n}(Q_1'\otimes I_r)R_{1t}(Q_1\otimes I_r))(P_1'\otimes I_r)(\grave{\alpha}_1-\alpha_1)+o_p(1), \quad (B.\ 7)\\
&\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\dot{Q}_2^{**}(r_{2t}|_{\hat{\alpha},\dot{\delta}})=\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(r_{2t}|_{\grave{\alpha},\dot{\delta}})\\
=\ &\frac{1}{\sqrt{n}}\sum_{t=1}^{n}r_{2t}+(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}R_{2t})(\grave{\alpha}_2-\alpha_2)+o_p(1). \quad (B.\ 8)
\end{aligned}
$$

(a). Recall that $\dot{Q}_1^{**\prime-1}\hat{\alpha}_1=(P_1'\otimes I_r)\grave{\alpha}_1$. By (4.9), (B.5) and (B.7),

$$
\begin{aligned}
n\dot{Q}_1^{**\prime-1}\dot{\alpha}_1 &= n\dot{Q}_1^{**\prime-1}\hat{\alpha}_1-[n^{-2}\sum_{t=1}^{n}\dot{Q}_1^{**}(R_{1t}|_{\hat{\alpha},\dot{\delta}})\dot{Q}_1^{**\prime}]^{-1}[n^{-1}\sum_{t=1}^{n}\dot{Q}_1^{**}(r_{1t}|_{\hat{\alpha},\dot{\delta}})]\\
&= n(P_1'\otimes I_r)\grave{\alpha}_1-[n^{-2}\sum_{t=1}^{n}(Q_1'\otimes I_r)R_{1t}(Q_1\otimes I_r)]^{-1}[n^{-1}\sum_{t=1}^{n}(Q_1'\otimes I_r)r_{1t}]\\
&\quad -n(P_1'\otimes I_r)(\grave{\alpha}_1-\alpha_1)+o_p(1)\\
&= n(P_1'\otimes I_r)\alpha_1-[\frac{1}{n^2}\sum_{t=1}^{n}(Q_1'\otimes I_r)R_{1t}(Q_1\otimes I_r)]^{-1}[\frac{1}{n}\sum_{t=1}^{n}(Q_1'\otimes I_r)r_{1t}]\\
&\quad +o_p(1). \qquad (B.\ 9)
\end{aligned}
$$

Note that $\dot{Q}_1^{**\prime-1}\dot{\alpha}_1-(P_1'\otimes I_r)\alpha_1=vec[((\dot{B}\bar{B}')^{-1}\dot{B}-B)P_1]$. By (B.9) and Lemma 3.1(a)-(b), (a) holds.

(b). By (4.10), (B.6) and (B.8),

$$
\sqrt{n}\dot{Q}_2^{**\prime-1}\dot{\alpha}_2 = \sqrt{n}\dot{Q}_2^{**\prime-1}\hat{\alpha}_2-[n^{-1}\sum_{t=1}^{n}\dot{Q}_2^{**}(R_{2t}|_{\hat{\alpha},\dot{\delta}})\dot{Q}_2^{**\prime}]^{-1}[n^{-1/2}\sum_{t=1}^{n}\dot{Q}_2^{**}(r_{2t}|_{\hat{\alpha},\dot{\delta}})]
$$

$$= \sqrt{n}\grave{\alpha}_2 - [n^{-1}\sum_{t=1}^{n} R_{2t}]^{-1}[n^{-1/2}\sum_{t=1}^{n} r_{2t}] - \sqrt{n}(\grave{\alpha}_2 - \alpha_2) + o_p(1)$$
$$= \sqrt{n}\alpha_2 - [n^{-1}\sum_{t=1}^{n} R_{2t}]^{-1}[n^{-1/2}\sum_{t=1}^{n} r_{2t}] + o_p(1). \qquad \text{(B. 10)}$$

By (B.10) and Lemma 3.1(a)-(b), (b) holds. This completes the proof. □

Proof of Lemma 5.1. Let $\dot{\varphi}^* = vec[CP_1, \dot{C}P_2, \dot{\Phi}_1^*, \cdots, \dot{\Phi}_{s-1}^*]$, and $l^*(\dot{\varphi}^*, \dot{\delta})$ be $l(\dot{\varphi}, \dot{\delta})$ with $\dot{C}P_1 Z_{1t-1}$ replaced by $CP_1 Z_{1t-1}$. By Lemma 3.1, Theorem 3.1 and a Taylor's expansion, we can show that

$$2[l(\dot{\varphi}, \dot{\delta}) - l^*(\dot{\varphi}^*, \dot{\delta})] = vec[n(\dot{C} - C)P_1]'[\frac{1}{n^2}\sum_{t=1}^{n} L_{1t}]vec[n(\dot{C} - C)P_1] + o_p(1), \text{(B. 11)}$$

where $L_{1t} = (Z_{1t-1}Z'_{1t-1} \otimes V_t^{-1}) + \sum_{l=1}^{t-1}[Z_{1t-l-1}Z'_{1t-l-1} \otimes ((\Gamma^{-1} \odot \Gamma + I_m) \odot \nu_l\nu_l' \odot \Pi_{lt})]$.

Denote $\ddot{A} = \dot{A}(\dot{B}\bar{B}')$ and $\ddot{B} = (\dot{B}\bar{B}')^{-1}\dot{B}$. Note $\dot{A}\dot{B} = \ddot{A}\ddot{B}$. Moreover,

$$\ddot{A}\ddot{B} - AB = (\ddot{A} - A)B + A(\ddot{B} - B) + (\ddot{A} - A)(\ddot{B} - B).$$

Recall that $BP_1 = 0$. By Theorem 4.2, $(\ddot{B} - B)P_1 = O_p(n^{-1})$ and $(\ddot{A} - A) = O_p(n^{-1/2})$ under H_0. Hence,

$$n(\ddot{A}\ddot{B} - AB)P_1 = n(\ddot{A} - A)BP_1 + nA(\ddot{B} - B)P_1 + (\ddot{A} - A)n(\ddot{B} - B)P_1$$
$$= nA(\ddot{B} - B)P_1 + O_p(n^{-1/2}). \qquad \text{(B. 12)}$$

Let $\dot{\alpha}^* = vec[ABP_1, \dot{A}\dot{B}P_2, \dot{\Phi}_1^*, \cdots, \dot{\Phi}_{s-1}^*]$, and $l^*(\dot{\alpha}^*, \dot{\delta})$ be $l(\dot{\alpha}, \dot{\delta})$ with $\dot{A}\dot{B}P_1 Z_{1t-1}$ replaced by $ABP_1 Z_{1t-1} = CP_1 Z_{1t-1}$. By Lemma 3.1, Theorem 4.2, a Taylor's expansion and (A.12), we can show that:

$$2[l(\dot{\alpha}, \dot{\delta}) - l^*(\dot{\alpha}^*, \dot{\delta})]$$
$$= vec[n(\ddot{A}\ddot{B} - AB)P_1]'[n^{-2}\sum_{t=1}^{n} L_{1t}]vec[n(\ddot{A}\ddot{B} - AB)P_1] + o_p(1)$$
$$= vec[nA(\ddot{B} - B)P_1]'[n^{-2}\sum_{t=1}^{n} L_{1t}]vec[nA(\ddot{B} - B)P_1] + o_p(1). \qquad \text{(B. 13)}$$

It is straightforward to show that $l^*(\dot{\varphi}^*, \dot{\delta}) - l^*(\dot{\alpha}^*, \dot{\delta}) = o_p(1)$. Furthermore, by (A.11), (A.13) and Lemma 3.1, it follows that

$$LR_G \longrightarrow_{\mathcal{L}} vec[\Omega_1^{-1}M^*]'[Z \otimes \Omega_1]vec[\Omega_1^{-1}M^*] - vec[DM^*]'[Z \otimes \Omega_1]vec[DM^*]$$

$$
\begin{aligned}
&= vec[\Omega_1^{-1}M^*]'vec[\Omega_1\Omega_1^{-1}M^*Z] - vec[DM^*]'vec[\Omega_1 DM^*Z] \\
&= tr[M^{*\prime}\Omega_1^{-1}M^*Z] - tr[M^{*\prime}D\Omega_1 DM^*Z] \\
&= tr[(\Omega_1^{-1} - A(A'\Omega_1 A)^{-1}A')M^*ZM^{*\prime}]. \qquad \text{(B. 14)}
\end{aligned}
$$

where $D \equiv A(A'\Omega_1 A)^{-1}A'$, $Z \equiv \psi_{11}\Omega_{a1}^{1/2}\int_0^1 B_d(u)B_d(u)'\Omega_{a1}^{1/2}\psi_{11}'$ and M^* is defined as in Theorem 3.1. Following the lines on p.359 of Reinsel and Ahn (1992), we can rewrite $\Omega_1^{-1} - A(A'\Omega_1 A)^{-1}A'$ as:

$$
\Omega_1^{-1}(\Omega_1 - \Omega_1 A(A'\Omega_1 A)^{-1}A'\Omega_1)\Omega_1^{-1} = \Omega_1^{-1}A_\perp(A_\perp'\Omega_1^{-1}A_\perp)^{-1}A_\perp'\Omega_1^{-1}.
$$

Therefore, we can rewrite the asymptotic distribution in (A.13) as:

$$
tr[(\int_0^1 B_d(u)dV_d^*(u)')'(\int_0^1 B_d(u)B_d(u)'du)^{-1}(\int_0^1 B_d(u)dV_d^*(u)')],
$$

where $V_d^*(u) \equiv (A_\perp'\Omega_1^{-1}A_\perp)^{-1/2}A_\perp'\Omega_1^{-1}W_m^*(u)$. Note $E[B_d(u)V_d^*(u)'] = u\Omega_{a1}^{-1/2}(A_\perp'\Omega_1^{-1}A_\perp)^{1/2} = u\Upsilon'$. Thus, we can rewrite $V_d^*(u)$ as a linear combination of two independent $d-$dimensional standard BMs:

$$
\Upsilon B_d(u) + [(A_\perp'\Omega_1^{-1}A_\perp)^{-1/2}A_\perp'\Omega_1^{-1}\Omega_1^*\Omega_1^{-1}A_\perp(A_\perp'\Omega_1^{-1}A_\perp)^{-1/2} - \Upsilon\Upsilon']^{1/2}V_d(u). \qquad \text{(B. 15)}
$$

The proof is complete. □

Proof of Theorem 5.1. When $\Omega_1^* = \Omega_1$, (A.15) in the proof of Lemma 5.1 can be simplified as $\Upsilon B_d(u) + [I_d - \Upsilon\Upsilon']^{1/2}V_d(u)$. Thus, the asymptotic distribution can be simplified as:

$$
\begin{aligned}
&tr\{[\int_0^1 \Upsilon B_d(u)dB_d(u)'\Upsilon' + \int_0^1 \Upsilon B_d(u)dV_d(u)'(I_d - \Upsilon\Upsilon')^{1/2}]' \\
&\cdot[\int_0^1 \Upsilon B_d(u)B_d(u)'\Upsilon' du]^{-1}[\int_0^1 \Upsilon B_d(u)dB_d(u)'\Upsilon' + \int_0^1 \Upsilon B_d(u)dV_d(u)'(I_d - \Upsilon\Upsilon')^{1/2}]\}.
\end{aligned}
$$

However, $\Upsilon B_d(u) \sim N(0, \Upsilon\Upsilon')$. Abusing the notation, we write $\Upsilon B_d(u)$ as $(\Upsilon\Upsilon')^{1/2}$ $B_d(u)$, where $B_d(u)$ is (another) $d-$dimensional standard BM independent of $V_d(u)$.

Therefore, cancelling some of the $(\Upsilon\Upsilon')^{1/2}$ terms, the asymptotic distribution can be expressed as:

$$
\begin{aligned}
&tr\{[\textstyle\int_0^1 B_d(u)dB_d(u)'(\Upsilon\Upsilon')^{1/2} + \int_0^1 B_d(u)dV_d(u)'(I_d - \Upsilon\Upsilon')^{1/2}]' \\
&[\textstyle\int_0^1 B_d(u)B_d(u)'du]^{-1}[\int_0^1 B_d(u)dB_d(u)'(\Upsilon\Upsilon')^{1/2} + \int_0^1 B_d(u)dV_d(u)'(I_d - \Upsilon\Upsilon')^{1/2}]\}.
\end{aligned}
$$

Since $(I_d - \Upsilon\Upsilon')$ is a real symmetric matrix, we can decompose it as $\Theta\Lambda_d\Theta'$, where Θ is an orthogonal matrix such that $\Theta'\Theta = I_d$. In view of $(\Upsilon\Upsilon')^{1/2} = \Theta(I_d - \Lambda_d)^{1/2}\Theta'$ and $(I_d - \Upsilon\Upsilon')^{1/2} = \Theta\Lambda_d^{1/2}\Theta'$ and due to the orthogonality of Θ, we can write the asymptotic distribution as:

$$
\begin{aligned}
tr\{&[\int_0^1 \Theta' B_d(u) dB_d(u)'\Theta(I_d - \Lambda_d)^{1/2}\Theta' + \int_0^1 \Theta' B_d(u) dV_d(u)'\Theta\Lambda_d^{1/2}\Theta']' \\
&\cdot[\int_0^1 \Theta' B_d(u) B_d(u)' du \Theta]^{-1} \\
&\cdot[\int_0^1 \Theta' B_d(u) dB_d(u)'\Theta(I_d - \Lambda_d)^{1/2}\Theta' + \int_0^1 \Theta' B_d(u) dV_d(u)'\Theta\Lambda_d^{1/2}\Theta']\}.
\end{aligned}
$$

Since $\Theta' B_d(u) \sim N(0, \Theta'\Theta) = N(0, I_d)$, similar to the previous arguments, and abusing the notation, we can write $\Theta' B_d(u)$ and $\Theta' V_d(u)$ as two independent standard BMs $B_d(u)$ and $V_d(u)$ respectively. Cancelling the orthogonal Θ, we have:

$$
\begin{aligned}
tr\{&[\int_0^1 B_d(u) dB_d(u)'(I_d - \Lambda_d)^{1/2} + \int_0^1 B_d(u) dV_d(u)'\Lambda_d^{1/2}]' \\
&\cdot[\int_0^1 B_d(u) B_d(u)' du]^{-1}[\int_0^1 B_d(u) dB_d(u)'(I_d - \Lambda_d)^{1/2} + \int_0^1 B_d(u) dV_d(u)'\Lambda_d^{1/2}]\} \\
= tr\{&[\zeta(I_d - \Lambda_d)^{1/2} + \Phi\Lambda_d^{1/2}]'[\zeta(I_d - \Lambda_d)^{1/2} + \Phi\Lambda_d^{1/2}]\}.
\end{aligned}
$$

This completes the proof. □

REFERENCES

Ahn, S. K. and Reinsel, G.C. (1990), Estimation for Partially Nonstationary Multivariate models, *Journal of American Statistical Association*, 85, 813-823.

Alexakis, P. and Apergis, N. (1996), ARCH Effects and Cointegration: Is the Foreign Exchange Market Efficient?, *Journal of Banking and Finance*, 20, 687-697.

Anderson, T.W. (1951), Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions, *Annals of Mathematical Statistics*, 22, 327-351. [Correction, *Annals of Statistics*, 8, 1980, p.1400.]

Anderson, T.W. (2002), Reduced Rank Regression in Cointegrated Models, *Journal of Econometrics*, 106, 203-216.

Bollerslev, T. (1990), Modeling the Coherence in the Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Approach, *Review of Economics and Statistics*, 72, 498-505.

Brenner, R.J. and Kroner, K.F. (1995), Arbitrage, Cointegration, and Testing the Unbiasedness Hypothesis in Financial Markets, *Journal of Financial and Quantitative Analysis*, 30, 23-42.

Chan, N.H. and Wei, C.Z. (1988), Limiting Distributions of Least Squares Estimates of Unstable Autoregressive Processes, *Annals of Statistics*, 16, 367-401.

Engle, R.F. and Granger, C.W.J. (1987), Cointegration and Error Correction: Representation, Estimation and Testing, *Econometrica*, 55, 251-276.

Engle, R.F., Patton, A.J. (2004), Impacts of Trades in an Error-Correction Model of Quote Prices. *Journal of Financial Markets*, 7, 125.

Franses, P.H., Kofman, P. and Moser. J. (1994), Garch Effects on a Test for Cointegration, *Review of Quantitative Finance and Accounting*, 4, 19-26.

Glosten, L.R., Jagannathan, R. and Runkle, D.E. (1993), On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance*, 48, 1779-1802.

Granger, C.W.J. (1983), Cointegrated Variables and Error Correction models, Discussion Paper, Department of Economics, University of California at San Diego.

Hasbrouck, J. (1995), One security, Many markets: Determining the Contributions to Price Discovery, *Journal of Finance*, 50, 1175-1199.

Hasbrouck, J. (2003), Intraday Price Formation in U.S. Equity Index Markets, *Journal of Finance*, 58, 2375-2399.

Jeantheau, T. (1998), Strong Consistency of Estimators for Multivariate ARCH Models, *Econometric Theory*, 14, 70-86.

Johansen, S. (1996), *Likelihood-Based Inference in Cointegrated Vector Autoregressive models.* Oxford: Oxford University Press.

Kroner, K.F. and Sultan, J. (1993), Time-Varying Distributions and Dynamic Hedging with Foreign Currency Futures, *Journal of Financial and Quantitative Analysis*, 28, 535-551.

Li, W.K., Ling, S. and Wong, H. (2001), Estimation for Partially Nonstationary Multivariate Autoregressive models with Conditional Heteroskedasticity, *Biometrika*, 88, 1135-1152.

Ling, S. and Li, W.K. (1998), Limiting Distributions of Maximum Likelihood Estimators for Unstable ARMA models with GARCH Errors, *Annals of Statistics*, 26, 84-125.

Ling, S., Li, W.K. and McAleer, M. (2003), Estimation and Testing for Unit Root Processes with GARCH(1,1) Errors: Theory and Monte Carlo Study, *Econometric Reviews*, 22, 179-202.

Ling, S. and McAleer, M. (2003a), Asymptotic Theory for a Vector ARMA-GARCH model, *Econometric Theory*, 19, 280-310.

Ling, S. and McAleer, M. (2003b), On Adaptive Estimation in Nonstationary ARMA models with GARCH Errors, *Annals of Statistics*, 31, 642-674.

Magnus, J.R. (1988), *Linear Structures.* New York: Oxford University Press.

Magnus, J.R. and Neudecker, H. (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics.* Chichester: John Wiley and Sons.

Phillips, P.C.B. and Durlauf, S.N. (1986), Multiple Time Series Regression with Integrated Processes, *Review of Economic Studies*, 53, 473-495.

Reinsel, G.C. and Ahn, S.K. (1992), Vector AR models with Unit Root and Reduced Rank Structure: Estimation, Likelihood Ratio Test, and Forecasting, *Journal of Time Series Analysis*, 13, 133-145.

Sin, C.-y. (2003), Estimating a Linear Exponential Density when the Weighting Matrix and Mean Parameter Vector are Functionally Related, in: Fomby, T.B., Hill, R.C. (Eds.), *Advances in Econometrics Volume 17 - Maximum Likelihood Estimation of Misspecified Models: Twenty Years Later.* Elsevier, Amsterdam, 177-197.

Sin, C.-y. and Ling, S. (2004), Estimation and Testing for Partially Nonstationary Vector Autoregressive models with GARCH, Discussion Paper, Hong Kong Baptist University, and Hong Kong University of Science and Technology.

Stock, J.H. and Watson, M.W. (1993), A Simple Estimator of Cointegrating Vectors in Higher Order Integrated System, *Econometrica*, 61, 783-820.

Tse, Y.K. and Tsui, A.K.C. (2002), A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-Varying Correlations, *Journal of Business and Economic Statistics*, 20, 351-362.