Estimation and Testing for Partially Nonstationary Vector Autoregressive Models with GARCH

Chor-yiu SIN
Department of Economics
Hong Kong Baptist University, Hong Kong

Shiqing LING
Department of Mathematics
Hong Kong University of Science and Technology, Hong Kong

This Version: October 1, 2002

Abstract

Macroeconomic or financial data are often modelled with cointegration and conditional heteroskedasticity, such as the generalized autoregressive conditional heteroskedasticity (GARCH). However, the statistical inference method and the asymptotic theory for the cointegration with GARCH errors have yet to be well developed. In this paper, we consider a partially nonstationary autoregressive model with GARCH. In addition, no prior knowledge of the reduced rank structure is assumed. We propose the full rank and the reduced rank quasi-maximum likelihood estimation for the model. The asymptotic distributions of the estimators are proved to be a functional of two correlated high-dimensional Brownian motions. These two estimators are used to construct a likelihood ratio (LR) test for the reduced rank, where asymptotic distribution is in turn a functional of a standard Brownian motion and a standard normal vector, with some unknown nuisance parameters. The critical values of the LR test are simulated via the Monte Carlo method. The performance of this test in finite samples is examined through Monte Carlo experiments. We also apply our approach to an empirical example of three interest rates.

Key Words: Autoregressive Model; Cointegration; Full rank estimation; Multivariate GARCH process; Partially nonstationary; Reduced rank estimation; Reduced rank structure.

Acknowledgments: We thank helpful comments from participants at the 1999 NBER/NSF Time Series Conference, the 2000 Econometric Society World Congress, and the 2002 North American Summer Meeting of the Econometric Society.

1 Introduction

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

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

$$\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{mt})', \tag{1.2}$$

$$\varepsilon_{it} = \eta_{it}\sqrt{h_{it}}, \;\; h_{it} = a_{i0} + \sum_{j=1}^{q} a_{ij}\varepsilon_{it-j}^2 + \sum_{k=1}^{p} b_{ik} h_{it-k}, \tag{1.3}$$

where Φ_j's are constant matrices, and $\det\{\Phi(z)\} = |I - \Phi_1 z - \cdots - \Phi_s z^s| = 0$ has $d \le m$ unit roots and $r = m - d$ roots outside the unit circle. $\eta_t = (\eta_{1t}, \ldots, \eta_{mt})'$ is a sequence of independently and identically distributed (i.i.d.) random vectors with zero mean and $E(\eta_t \eta_t') = \Gamma$, where

$$\Gamma = \begin{pmatrix} 1 & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & 1 & \sigma_{23} & \cdots \\ & & \cdots & \\ \sigma_{m,1} & \cdots & \sigma_{m,m-1} & 1 \end{pmatrix},$$

in which $\sigma_{ij} = \sigma_{ji}$. It is easy to see that $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$ and $E(\varepsilon_t\varepsilon_t'|\mathcal{F}_{t-1}) = V_t = D_t \Gamma D_t$, where $\mathcal{F}_t = \sigma\{\eta_s, s = t, t-1, \ldots\}$ and $D_t = diag(\sqrt{h_{1t}}, \ldots, \sqrt{h_{mt}})$. V_t is the time-varying covariance matrix with constant correlation. The process ε_t in (1.2)-(1.3) is the multivariate generalized autoregressive conditional heteroskedasticity (GARCH) process proposed by Bollerslev (1990) and has been widely used in the literature, for example by Tse (2000). We call Model (1.1)-(1.3) a partially nonstationary multivariate AR model with GARCH.

Assuming the ε_t's are i.i.d., Ahn and Reinsel (1990) (see also Johansen, 1988) show that, although some component series of $\{Y_t\}$ exhibit nonstationary behaviour, there are r linear combinations of $\{Y_t\}$ that are stationary. This phenomenon, which is called cointegration in the literature of economics, was first investigated in Granger (1983) (see also Engle and Granger, 1987). Numerous economic models such as consumption function, purchasing power parity, money demand function, hedging ratio of spot and futures exchange rates, and yield curves of different terms of maturities impose this restriction of cointegration. The partially nonstationary multivariate AR model or cointegration time series models without GARCH have been extensively discussed over the past twenty years. Noticeable examples include Phillips and Durlauf (1986), Johansen (1988,1995), Stock and Watson (1993), and Rahbek and Mosconi (1999).

Economic time series related to financial markets, such as exchange rates and interest rates, often exhibit time-varying variances such as GARCH. Recently, a considerable number of papers, including Kroner and Sultan (1993), Brenner and Kroner (1995), Alexakis and Apergis (1996), and Li, Ling and Wong (2001) (henceforth LLW (2001)), investigate multivariate time series that exhibit both cointegration and time-varying variances. On the other hand, while Franses, Kofman and Moser (1994) and Lee and Tse (1996) perform Monte Carlo experiments on different tests for cointegration which ignore the presence of GARCH, the statistical inference and the asymptotic theory have yet to be well developed. In LLW (2001), the heteroskedasticity part is the random coefficient autoregressive (RCAR) model proposed in Nicholls and Quinn (1982) and Tsay (1987). It excludes the terms h_{it-k} in Model (1.3). Moreover, LLW (2001) mainly consider the case with a diagonal Γ.

In contrast, the model of heteroskedasticity in this paper includes the h_{it-k} and it does not assume a diagonal Γ. These differences from LLW (2001)'s model are substantial in practice, and thus our model provides a much wider scope of applications. Furthermore, unlike Ahn and Reinsel (1990) and LLW (2001), our approach is similar to that of Anderson (1951) and Johansen (1988) and does not assume the unnecessary prior knowledge of the reduced rank structure. This is non-trivial when m is equal to or greater than 3.

In this paper, we first investigate the full rank and the reduced rank quasi-maximum likelihood estimations (QMLE) for Model (1.1)-(1.3). The asymptotic distribution of either estimator is proved to be a functional of two correlated high-dimensional Brownian motions. Using these two estimators, we construct an LR test for the reduced rank. To the best of our knowledge, such a test does not exist in the literature. Again, our test does not assume the prior knowledge of the reduced rank structure. We show that the asymptotic distribution of the LR test is a functional of a standard Brownian motion and a standard normal vector with d unknown nuisance parameters. The critical value of this LR test is thus simulated via the Monte Carlo method. It is worth mentioning that in the univariate case, Seo (1999) and Ling, Li and McAleer (2002) construct a unit root test with the MLE of a unit-root AR/ARMA-GARCH model. The Monte Carlo experiments in the two papers suggested that the unit root test based on the maximum likelihood estimation (MLE) is much more powerful that the Dickey-Fuller test based on the least squares estimation (LSE). It is expected that the LR test based on the MLE of Model (1.1)-(1.3) is more powerful than Johansen (1988)'s test or Reinsel and Ahn (1992)'s test which both ignore GARCH. This is confirmed with our Monte Carlo experiments.

This paper proceeds as follows. Section 2 discusses the structure of Model (1.1)-(1.3). Section 3 and Section 4 derive the asymptotic distribution of the full rank estimator and that of the reduced rank estimator, respectively. Section 5 devises a test for the reduced rank. Section 6 extends the previous results to a model with a constant term. We report the Monte Carlo experiments and an illustrative empirical example of three interest rates in Sections 7 and 8 respectively. Conclusions can be found in Section 9. All technical proofs are relegated to the Appendix.

2 Basic Properties of Models

First we reparameterize Model (1.1) as follows:

$$W_t = CY_{t-1} + \Phi_1^* W_{t-1} + \cdots + \Phi_{s-1}^* W_{t-s+1} + \varepsilon_t, \tag{2.1}$$

where $W_t = Y_t - Y_{t-1}$, $\Phi_j^* = -\sum_{k=j+1}^{s} \Phi_k$ and $C = -\Phi(1) = -(I_m - \sum_{j=1}^{s} \Phi_j)$. Following Ahn and Reinsel (1990), let $m \times m$ matrices P and $Q = P^{-1}$ be such that $Q(\sum_{j=1}^{s}\Phi_j)P = \text{diag}(I_d, \Gamma_r)$, the Jordan canonical form of $\sum_{j=1}^{s}\Phi_j$. Defining $Z_t = QY_t$, we obtain

$$Z_t = \text{diag}(I_d, \Gamma_r)Z_{t-1} + u_t \text{ and } u_t = Q[\Phi_1^* W_{t-1} + \cdots + \Phi_{s-1}^* W_{t-s-1} + \varepsilon_t]. \quad (2.2)$$

Furthermore, let $g(z) = (1-z)^{-d}det\{\Phi(z)\}$ and $H(z) = (1-z)^{-d+1}adj\{\Phi(z)\}$, we can rewrite u_t as

$$u_t = \{I_m + Q\sum_{k=1}^{s-1}\Phi_k^* g(B)^{-1}H(B)PB^k\}a_t = \Psi(B)a_t, \quad (2.3)$$

where $a_t = Q\varepsilon_t$ and $\Psi(B) = I_m + Q\sum_{k=1}^{s-1}\Phi_k^* g(B)^{-1}H(B)PB^k = \sum_{k=0}^{\infty}\Psi_k B^k$ in which $\Psi_0 = I_m$, $\Psi_k = O(\rho^k)$ and $\rho \in (0,1)$, as in Ahn and Reinsel (1990). Partition $Q' = [Q_1, Q_2]$ and $P = [P_1, P_2]$ such that Q_1 and P_1 are $m \times d$ matrices, and Q_2 and P_2 are $m \times r$ matrices. Further partition $u_t = [u_{1t}', u_{2t}']'$ such that u_{1t} is $d \times 1$ and u_{2t} is $r \times 1$. Define $Z_{1t} = Q_1'Y_t$ and $Z_{2t} = Q_2'Y_t$ so that

$$Z_{1t} = Z_{1,t-1} + u_{1t} \quad \text{and} \quad Z_{2t} = \Gamma_r Z_{2t-1} + u_{2t}. \quad (2.4)$$

Given Assumption (a) below, it can be shown that $\{Z_{1t}\}$ is a nonstationary random vector with d unit roots; while $\{Z_{2t}\}$ is stationary. The r columns of Q_2' are called cointegrated vectors in the literature of economics.

Now we make the following assumptions for Model (1.2)-(1.3).

Assumption (a). $a_{i0} > 0$, $a_{i1}, \ldots, a_{iq}, b_{i1}, \ldots, b_{ip} \geq 0$, and $\sum_{j=1}^{q} a_{ij} + \sum_{k=1}^{p} b_{ik} < 1$, where $i = 1, \ldots, m$.

Assumption (b). For $i = 1, \ldots, m$, all eigenvalues of $E(A_{it} \otimes A_{it})$ lie inside the unit circle, where $\otimes$ denotes the Kronecker product and

$$A_{it} = \begin{pmatrix} a_{i1}\eta_{it}^2 & \ldots & a_{iq}\eta_{it}^2 & b_{i1}\eta_{it}^2 & \ldots & b_{ip}\eta_{it}^2 \\ & I_{(q-1)\times(q-1)} & 0_{(q-1)\times 1} & & 0_{(q-1)\times p} & \\ a_{i1} & \ldots & a_{iq} & b_{i1} & \ldots & b_{ip} \\ & 0_{(p-1)\times q} & & & I_{(p-1)\times(p-1)} & 0_{(p-1)\times 1} \end{pmatrix}.$$

Assumption (c). η_t is symmetrically distributed.

Assumptions (a) and (b) are the necessary and sufficient conditions for the finite second and fourth moments of the GARCH error ε_t in Model (1.2)-(1.3), see Ling

(1999) and Ling and McAleer (2002a). Assumption (c) is for convenience, which renders (3.9) below and thus allows the parameters in the AR part and those in the GARCH part to be estimated separately without altering the asymptotic distributions.

3 Full Rank Estimation

This section considers a full rank estimation, which incorporates the GARCH specified in Model (1.2)-(1.3). The estimators in the GARCH part may be used in updating the reduced rank estimators which incorporate GARCH in Sub-section 4.2. We first rewrite Model (2.1) as

$$W_t = CP_1 Z_{1t-1} + CP_2 Z_{2t-1} + \sum_{j=1}^{s-1} \Phi_j^* W_{t-j} + \varepsilon_t. \qquad (3.1)$$

Let $X_{t-1} \equiv [Y'_{t-1}, W'_{t-1}, \ldots, W'_{t-s+1}]'$ (a $sm \times 1$ vector), $\varphi \equiv vec[C, \Phi_1^*, \ldots, \Phi_{s-1}^*]$ (a $m^2 s \times 1$ vector of parameters in the AR part) and $\delta \equiv [\delta_1', \delta_2']'$ (parameters in the GARCH part), where $\delta_1 \equiv [a_0', a_1', \ldots, a_q', b_1', \ldots, b_p']'$ (a $(1+q+p)m \times 1$ vector) ($a_j \equiv [a_{1j}, ..., a_{mj}]', j = 0, 1, ..., q;\ b_k \equiv [b_{1k}, ..., b_{mk}]', k = 1, ..., p.$) and $\delta_2 \equiv \tilde{v}(\Gamma)$ (note $\tilde{v}(\Gamma)$ is a $m(m-1)/2 \times 1$ vector which is obtained from $vec(\Gamma)$ by eliminating the supradiagonal and the diagonal elements of Γ. See, for instance, p.27 in Magnus, 1988).

3.1 Computational Procedure

Denote the full-rank estimates of φ and δ as $\hat{\varphi}$ and $\hat{\delta}$, which maximize the conditional log-likelihood function:

$$l = \sum_{t=1}^{n} l_t \ \text{ and } \ l_t = -\frac{1}{2}\varepsilon_t' V_t^{-1} \varepsilon_t - \frac{1}{2} \ln |V_t|, \qquad (3.2)$$

where $V_t = D_t \Gamma D_t$. For simplicity, we assume that the initial value $Y_s = 0$ for $s \leq 0$. Moreover, it should be noted that l_t, ε_t and V_t in (3.2) are functions of the unknown parameters φ and δ. Denote $h_t = (h_{1t}, \ldots, h_{mt})'$ and $\tilde{h}_t = (h_{1t}^{-1}, \ldots, h_{mt}^{-1})'$. Using the conventional definitions of gradients (see, for instance, Theorem 6 in Chapter 5 of Magnus and Neudecker, 1988), it follows that

$$\nabla_\varphi l_t = -\frac{1}{2}\nabla_\varphi h_t(\iota - w(\varepsilon_t\varepsilon_t' V_t^{-1})) \odot \tilde{h}_t + (X_{t-1}\otimes I_m)V_t^{-1}\varepsilon_t, \tag{3.3}$$

$$\nabla_\delta l_t = \begin{pmatrix} -\frac{1}{2}\nabla_{\delta_1} h_t(\iota - w(\varepsilon_t\varepsilon_t' V_t^{-1})) \odot \tilde{h}_t \\ -\tilde{\nu}(\Gamma^{-1} - \Gamma^{-1}D_t^{-1}\varepsilon_t\varepsilon_t' D_t^{-1}\Gamma^{-1}) \end{pmatrix}, \tag{3.4}$$

where $\odot$ is the Hadamard product, ι is the $m\times 1$ sum vector $(1,1,\ldots,1)'$, $w(\cdot)$ is a $m\times 1$ vector containing the diagonal elements of an $m\times m$ matrix. Moreover, for $i = 1,2,\ldots,m$, let $a^{(i)}(z)b^{(i)}(z)^{-1} = \sum_{j=1}^{\infty}\nu_{ij}z^j$, where $a^{(i)}(z) = \sum_{j=1}^{q} a_{ij}z^j$ and $b^{(i)}(z) = 1-\sum_{j=1}^{p} b_{ij}z^j$. Denote $\nu_j \equiv (\nu_{1j},\ldots,\nu_{mj})'$, $j = 1,2,\cdots$. Let $\tilde{\varepsilon}_t \equiv (\varepsilon_{1t}^2,\ldots,\varepsilon_{mt}^2)'$. It follows that

$$\begin{aligned}\nabla_\varphi h_t &= -2\sum_{l=1}^{q}(X_{t-l-1}\otimes I_m)diag(a_l\odot\varepsilon_{t-l}) + \sum_{l=1}^{p}(\nabla_\varphi h_{t-l})diag(b_l) \\ &= -2\sum_{l=1}^{t-1}(X_{t-l-1}\otimes I_m)diag(\nu_l\odot\varepsilon_{t-l}); \end{aligned}\tag{3.5}$$

$$\nabla_{a_0} h_t = I_m + \sum_{l=1}^{p}(\nabla_{a_0} h_{t-l})diag(b_l), \tag{3.6}$$

$$\nabla_{a_j} h_t = diag(\tilde{\varepsilon}_{t-j}) + \sum_{l=1}^{p}(\nabla_{a_j} h_{t-l})diag(b_l), j = 1,\ldots,q, \tag{3.7}$$

$$\nabla_{b_k} h_t = diag(h_{t-k}) + \sum_{l=1}^{p}(\nabla_{b_k} h_{t-l})diag(b_l); k = 1,\ldots,p. \tag{3.8}$$

Let $\beta = (\beta_1',\beta_2')'$, where $\beta_1 = vec(CP_1)$ and $\beta_2 = vec(CP_2,\Phi_1^*,\ldots,\Phi_{s-1}^*)$. Further denote $\hat{\beta}_1 = \text{vec}(\hat{C}P_1)$, $\hat{\beta}_2 = \text{vec}(\hat{C}P_2,\hat{\Phi}_1^*,\ldots,\hat{\Phi}_{s-1}^*)$ and $\bar{Q}^* = \text{diag}(Q\otimes I_m, I_{(s-1)m^2})$. Then $\bar{Q}^{*\prime -1}(\hat{\varphi}-\varphi) = [(\hat{\beta}_1-\beta_1)',(\hat{\beta}_2-\beta_2)']'$. Define $\bar{D}^* = \text{diag}(nI_{dm}, \sqrt{n}I_{rm+(s-1)m^2})$. Using Assumptions (a)-(c) and a method similar to the one used in Ling and Li (1998), we can show that

$$\begin{gathered}\bar{D}^{*-1}\bar{Q}^*(\sum_{t=1}^{n}\nabla^2_{\varphi\varphi'}l_t)\bar{Q}^{*\prime}\bar{D}^{*-1} = \sum_{t=1}^{n}\bar{D}^{*-1}\bar{Q}^*F_t\bar{Q}^{*\prime}\bar{D}^{*-1} + o_p(1), \\ \frac{1}{\sqrt{n}}\bar{D}^{*-1}\bar{Q}^*(\sum_{t=1}^{n}\nabla^2_{\varphi\delta'}l_t) = o_p(1) \text{ and } \frac{1}{n}\sum_{t=1}^{n}\nabla^2_{\delta\delta'}l_t = \frac{1}{n}\sum_{t=1}^{n}S_t + o_p(1),\end{gathered}\tag{3.9}$$

where $F_t = -(X_{t-1}X_{t-1}'\otimes V_t^{-1}) - (\nabla_\varphi h_t)D_t^{-2}(\Gamma^{-1}\odot\Gamma + I_m)D_t^{-2}(\nabla_\varphi' h_t)/4$, $S_t = (S_{ijt})_{2\times 2}$, $S_{11t} = -(\nabla_{\delta_1}h_t)D_t^{-2}(\Gamma^{-1}\odot\Gamma + I_m)D_t^{-2}(\nabla_{\delta_1}' h_t)/4$, $S_{12t} = -(\nabla_{\delta_1}h_t)D_t^{-2}\Psi_m (I_m\otimes\Gamma^{-1})N_m\tilde{L}_m'$, and $S_{22t} = -2\tilde{L}_m N_m[\Gamma^{-1}\otimes\Gamma^{-1}]N_m\tilde{L}_m'$, see Appendix A.

By (3.9), the φ and δ can be estimated separately without altering the asymptotic distributions. The estimators $\hat{\varphi}$ and $\hat{\delta}$ can be obtained by the iterative Newton-Raphson algorithm:

$$\hat{\varphi}^{(k+1)} = \hat{\varphi}^{(k)} - (\sum_{t=1}^{n} F_t)^{-1}(\sum_{t=1}^{n} \nabla_{\varphi} l_t)|_{\hat{\varphi}^{(k)},\hat{\delta}^{(k)}}, \tag{3.10}$$

$$\hat{\delta}^{(k+1)} = \hat{\delta}^{(k)} - (\sum_{t=1}^{n} S_t)^{-1}(\sum_{t=1}^{n} \nabla_{\delta} l_t)|_{\hat{\varphi}^{(k)},\hat{\delta}^{(k)}}; \tag{3.11}$$

where $\hat{\varphi}^{(k)}$ and $\hat{\delta}^{(k)}$ are the estimates at the k-th iteration. The initial estimator $\hat{\varphi}^{(0)}$ is the LSE of φ in Model (1.1). Using the residuals $\hat{\varepsilon}_t$ from $\hat{\varphi}^{(0)}$ as the artifical observations of Model (1.2)-(1.3), the QMLE of δ in Model (1.2)-(1.3) is used as the initial value $\hat{\delta}^{(0)}$. Given Assumptions (a)-(c), it is not difficult to show that $\bar{D}^{*-1}\bar{Q}^{*}(\hat{\varphi}^{(0)} - \varphi) = O_p(1)$ and $\hat{\delta}^{(0)}$ is $\sqrt{n}-$consistent for δ, using methods similar to those employed by LLW (2001) and Ling, Li and McAleer (2002).

3.2 Limiting Distributions

For any fixed positive constant K, first define $\Theta_n \equiv \{(\tilde{\varphi},\tilde{\delta}) : \|\bar{D}^{*}\bar{Q}^{*'-1}(\tilde{\varphi} - \varphi)\| \leq K \ and \ \|\sqrt{n}(\tilde{\delta}-\delta)\| \leq K\}$. Similar to the arguments for the univariate case in Ling, Li and McAleer (2002) and Ling and Li (2002), we can show that the following holds uniformly in the ball Θ_n:

$$\sum_{t=1}^{n} \bar{D}^{*-1}\bar{Q}^{*}(F_t|_{\tilde{\varphi},\tilde{\delta}} - F_t)\bar{Q}^{*'}\bar{D}^{*-1} = o_p(1) \text{ and } n^{-1}\sum_{t=1}^{n}(S_t|_{\tilde{\varphi},\tilde{\delta}} - S_t) = o_p(1); \tag{3.12}$$

$$\sum_{t=1}^{n} \bar{D}^{*-1}\bar{Q}^{*}(\nabla_{\varphi} l_t|_{\tilde{\varphi},\tilde{\delta}} - \nabla_{\varphi} l_t) = \sum_{t=1}^{n} \bar{D}^{*-1}\bar{Q}^{*}F_t(\tilde{\varphi} - \varphi) + o_p(1), \qquad \text{and}$$

$$n^{-1/2}\sum_{t=1}^{n}(\nabla_{\delta} l_t|_{\tilde{\varphi},\tilde{\delta}} - \nabla_{\delta} l_t) = n^{-1/2}\sum_{t=1}^{n} S_t(\tilde{\delta} - \delta) + o_p(1). \tag{3.13}$$

Thus, the estimator of (φ,δ) obtained by (3.10)-(3.11) satisfies $\bar{D}^{*}\bar{Q}^{*'-1}(\tilde{\varphi} - \varphi) = O_P(1)$ and $\sqrt{n}(\tilde{\delta} - \delta) = O_p(1)$ if the initial estimator satisfies these equations, too. As a result, given the initial estimator suggested in Sub-section 3.1 and the iterative algorithm (3.10)-(3.11), we obtain the asymptotic representations:

$$\bar{D}^{*}\bar{Q}^{*'-1}(\hat{\varphi} - \varphi) = -(\sum_{t=1}^{n} \bar{D}^{*-1}\bar{Q}^{*}F_t\bar{Q}^{*'}\bar{D}^{*-1})^{-1}(\sum_{t=1}^{n} \bar{D}^{*-1}\bar{Q}^{*}\nabla_{\varphi} l_t) + o_p(1), \tag{3.14}$$

$$\sqrt{n}(\hat{\delta}-\delta) = -(\sum_{t=1}^{n} n^{-1}S_t)^{-1}(\sum_{t=1}^{n} n^{-1/2}\nabla_\delta l_t) + o_p(1). \tag{3.15}$$

Partition $\bar{Q}^*(X_{t-j}\otimes I_m)$ into two parts corresponding to β_1 and β_2,

$$\bar{Q}^*(X_{t-j}\otimes I_m) = \begin{pmatrix} Z_{1t-j}\otimes I_m \\ U_{t-j}\otimes I_m \end{pmatrix},$$

where $U_t = [Z'_{2t}, W'_t, \ldots, W'_{t-s+2}]'$. In view of (3.3) and (3.5), we have:

$$\begin{aligned}
\sum_{t=1}^{n} \bar{D}^{*-1}\bar{Q}^*\nabla_\varphi l_t &= \begin{pmatrix} n^{-1}\sum_{t=1}^{n} N_{1t} \\ n^{-1/2}\sum_{t=1}^{n} N_{2t} \end{pmatrix}, \qquad (3.16)\\
N_{1t} &= (Z_{1t-1}\otimes I_m)V_t^{-1}\varepsilon_t + \sum_{j=1}^{t-1}(Z_{1t-j-1}\otimes I_m)(\nu_j\odot\varepsilon_{t-j})\odot\lambda_t,\\
N_{2t} &= (U_{t-1}\otimes I_m)V_t^{-1}\varepsilon_t + \sum_{j=1}^{t-1}(U_{t-j-1}\otimes I_m)(\nu_j\odot\varepsilon_{t-j})\odot\lambda_t,
\end{aligned}$$

in which $\lambda_t = (\iota - w(\varepsilon_t\varepsilon'_t V_t^{-1}))\odot\tilde{h}_t$. As demonstrated by Ling and Li (1998), it can be shown that $n^{-3/2}\sum_{t=1}^{n} Z_{1t-j}U'_{t-k} = o_p(1), j,k = 1,2,\ldots$, and hence the cross-product terms in $\sum_{t=1}^{n}\bar{D}^{*-1}\bar{Q}^*F_t\bar{Q}^*\bar{D}^{*-1}$ that involve Z_{1t-j} and U_{t-k} also converge to zero in probability. Denote $\Pi_{jt}\equiv(\varepsilon_{t-j}\varepsilon'_{t-j}\odot\tilde{h}_t\tilde{h}'_t)$. From (A.1) in Appendix A, we have:

$$\sum_{t=1}^{n}\bar{D}^{*-1}\bar{Q}^*F_t\bar{Q}^*\bar{D}^{*-1} = -diag(n^{-2}\sum_{t=1}^{n}L_{1t}, n^{-1}\sum_{t=1}^{n}L_{2t}) + o_p(1), \tag{3.17}$$

$$\begin{aligned}
L_{1t} &= [Z_{1t-1}Z'_{1t-1}\otimes V_t^{-1}] + \sum_{j=1}^{t-1}[Z_{1t-j-1}Z'_{1t-j-1}\otimes((\Gamma^{-1}\odot\Gamma + I_m)\odot\nu_j\nu'_j\odot\Pi_{jt})],\\
L_{2t} &= [U_{t-1}U'_{t-1}\otimes V_t^{-1}] + \sum_{j=1}^{t-1}[U_{t-j-1}U'_{t-j-1}\otimes((\Gamma^{-1}\odot\Gamma + I_m)\odot\nu_j\nu'_j\odot\Pi_{jt})].
\end{aligned}$$

To facilitate the discussion of the asymptotic distributions of $\hat{C}P_1$ and other parameters of concern, denote $(W'_m(u), \tilde{W}'_m(u))'$ as the $2m-$dimensional Brownian motion with the covariance matrix:

$$u\Omega \equiv u\begin{pmatrix} V_0 & I_m \\ I_m & \tilde{\Omega}_1 \end{pmatrix},$$

where $V_0 = EV_t$, $\tilde{\Omega}_1 = E(V_t^{-1}) + E(\tau_t\tau'_t)$, $\tau_t = (\sum_{j=1}^{\infty}\nu_j\odot\varepsilon_{t-j})\odot\lambda_t$, $E(\tau_t\tau'_t) = (\Delta - \iota\iota')\odot\sum_{j=1}^{\infty}(\nu_j\nu'_j\odot E(\Pi_{jt}))$, and $\Delta = E[w(\eta_t\eta'_t\Gamma^{-1})(w(\eta_t\eta'_t\Gamma^{-1}))']$. Further, define $B_d(u) = \Omega_{a_1}^{-1/2}[I_d,0]\ \Omega_a^{1/2}V_0^{-1/2}W_m(u)$, where $\Omega_a = E(a_t a_t')$ and $\Omega_{a_1} = [I_d,0]\Omega_a[I_d,0]'$. Then, it is easy to see that $B_d(u)$ is a standard $d-$dimensional Brownian motion.

Lemma 3.1. Suppose Assumptions (a)-(c) hold. Then

$$
\begin{aligned}
&(a) \qquad n^{-2}\sum_{t=1}^{n} L_{1t} \longrightarrow_{\mathcal{L}} \Psi_{11}\Omega_{a_1}^{1/2}\int_0^1 B_d(u)B_d(u)'\Omega_{a_1}^{1/2}\Psi_{11}' \otimes \Omega_1,\\
&(b) \qquad n^{-1}\sum_{t=1}^{n} N_{1t} \longrightarrow_{\mathcal{L}} vec[\{\int_0^1 B_d(u)d\tilde{W}_m(u)'\}'\Omega_{a_1}^{1/2}\Psi_{11}'];\\
&(c) \qquad n^{-1}\sum_{t=1}^{n} L_{2t} \longrightarrow_{p} \Omega_2,\\
&(d) \qquad n^{-1/2}\sum_{t=1}^{n} N_{2t} \longrightarrow_{\mathcal{L}} N(0,\tilde{\Omega}_2);\\
&(e) \qquad -n^{-1}\sum_{t=1}^{n} S_t \longrightarrow_{p} \Omega_\delta,\\
&(f) \qquad n^{-1/2}\sum_{t=1}^{n} \nabla_\delta l_t \longrightarrow_{\mathcal{L}} N(0,\tilde{\Omega}_\delta),
\end{aligned}
$$

where $\Psi_{11} \equiv [I_d,0](\sum_{k=1}^{\infty}\Psi_k)[I_d,0]'$, $\Omega_\delta \equiv -E(S_t)$, $\tilde{\Omega}_\delta \equiv E(\nabla_\delta l_t \nabla_\delta' l_t)$, and

$$
\begin{aligned}
\Omega_1 &\equiv E(V_t^{-1}) + (\Gamma^{-1}\odot\Gamma + I_m)\odot\sum_{j=1}^{\infty}(\nu_j\nu_j'\odot E(\Pi_{jt}));\\
\Omega_2 &\equiv E(U_{t-1}U_{t-1}'\otimes V_t^{-1}) + \sum_{j=1}^{\infty}E(U_{t-j-1}U_{t-j-1}'\otimes(\Gamma^{-1}\odot\Gamma + I_m)\odot\nu_j\nu_j'\odot\Pi_{jt}),\\
\tilde{\Omega}_2 &\equiv E(U_{t-1}U_{t-1}'\otimes V_t^{-1}) + \sum_{j=1}^{\infty}E(U_{t-j-1}U_{t-j-1}'\otimes(\Delta - \iota\iota')\odot\nu_j\nu_j'\odot\Pi_{jt}). \qquad \square
\end{aligned}
$$

The following theorem comes from Lemma 3.1.

Theorem 3.1. Under the assumptions in Lemma 3.1,

$$
\begin{aligned}
&(a) \qquad n(\hat{C}-C)P_1 \longrightarrow_{\mathcal{L}} \Omega_1^{-1}\tilde{M},\\
&(b) \qquad \sqrt{n}(\hat{\beta}_2 - \beta_2) \longrightarrow_{\mathcal{L}} N(0,\Omega_2^{-1}\tilde{\Omega}_2\Omega_2^{-1}),\\
&(c) \qquad \sqrt{n}(\hat{\delta} - \delta) \longrightarrow_{\mathcal{L}} N(0,\Omega_\delta^{-1}\tilde{\Omega}_\delta\Omega_\delta^{-1}),
\end{aligned}
$$

where $\tilde{M} = [\int_0^1 B_d(u)d\tilde{W}_m(u)']'[\int_0^1 B_d(u)B_d(u)'du]^{-1}\Omega_{a_1}^{-1/2}\Psi_{11}^{-1}$. $\square$

It is not difficult to show that when η_t's are normally distributed, $(\Delta - \iota\iota') = (\Gamma^{-1}\odot\Gamma + I_m)$ and thus $\tilde{\Omega}_1 = \Omega_1$, $\tilde{\Omega}_2 = \Omega_2$, and $\tilde{\Omega}_\delta = \Omega_\delta$. When all β_{ik} are zero and

Γ is diagonal, the asymptotic distribution $\tilde{M}$ is the same as that in LLW (2001) and further when h_{ti} is a constant, $\tilde{M}$ is the same as the asymptotic distribution of full rank LSE in Ahn and Reinsel (1990). It should be noted that, for Model (1.1)-(1.3), the LSE method in Ahn and Reinsel (1990) can still be used, but the estimator for C is not as efficient as the QMLE in Theorem 3.1 in the sense proposed by Ling and McAleer (2002).

4 Reduced Rank Estimation

In this section, we consider the reduced rank estimation. Rewrite Equation (3.1) in a reduced rank form, which is also known as the error-correction form in the literature of economics:

$$W_t = ABY_{t-1} + \sum_{j=1}^{s-1} \Phi_j^* W_{t-j} + \varepsilon_t, \tag{4.1}$$

where the matrix C in Equation (3.1) is expressed as AB. Both A and B are full rank matrices of dimensions $m \times r$ and $r \times m$, respectively.

As we argue in Sub-section 4.2 below, the full rank estimator for δ (parameters in the GARCH part) suggested in Section 3 can be used as the initial estimator for the reduced rank estimation that incorporates GARCH. However, we need an initial estimator for the parameters in the AR part, which is the subject for discussion in Sub-section 4.1.

We denote the parameters in the *reduced rank* AR part as $\alpha = [\alpha_1', \alpha_2']'$, where $\alpha_1 \equiv vec[B]$ and $\alpha_2 \equiv vec[A, \Phi_1^*, \ldots, \Phi_{s-1}^*]$.

4.1 Initial Estimator for Parameters in AR Part

In this sub-section, we first show the asymptotic properties of Johansen's estimator, which will be used as the initial estimator for the parameters in the AR part. Then we argue the asymptotic equivalence between Johansen's estimator and that suggested in Ahn and Reinsel (1990) (or LLW (2001)). The latter imposes some prior knowledge of the reduced rank structure.

Johansen's estimator is essentially the QMLE which ignores the possible GARCH. More precisely, the estimator maximizes the likelihood function in (3.2), with V_t replaced by V_0, a constant matrix.

Denote the Johansen's estimator as $\hat{\alpha} = [\hat{\alpha}_1', \hat{\alpha}_2']'$, where $\hat{\alpha}_1 = vec[\hat{B}]$ and $\hat{\alpha}_2 = vec[\hat{A}, \hat{\Phi}_1^*, \ldots, \hat{\Phi}_{s-1}^*]$. Following Johansen (1988), in Theorem 4.1 below, we consider the limiting distributions of the *normalized* estimators $\check{\alpha}_1 \equiv (I_m \otimes (\hat{B}\bar{B}')^{-1})\hat{\alpha}_1$ and $\check{\alpha}_2 \equiv diag((\hat{B}\bar{B}')' \otimes I_m, I_{(s-1)m^2})\hat{\alpha}_2$, where $\bar{B} = (BB')^{-1}B$. As in Sub-section 3.1, we define $U_t^\dagger \equiv [(BY_t)', W_t', \ldots, W_{t-s+2}']'$.

Theorem 4.1. Suppose the assumptions in Lemma 3.1 hold. Consider the *normalized* estimators $vec[\check{B}] \equiv vec[(\hat{B}\bar{B}')^{-1}\hat{B}] = \check{\alpha}_1$ and $vec[\hat{A}(\hat{B}\bar{B}'), \hat{\Phi}_1^*, \ldots, \hat{\Phi}_{s-1}^*] = \check{\alpha}_2$.

(*a*) $\quad n(\check{B} - B)P_1 \longrightarrow_{\mathcal{L}} (A'V_0^{-1}A)^{-1}A'V_0^{-1}PM,$ and thus $n(\check{B} - B) = O_p(1)$;

(*b*) $\quad \sqrt{n}(\check{\alpha}_2 - \alpha_2) \longrightarrow_{\mathcal{L}} N(0, \Sigma_2^{-1}\tilde{\Sigma}_2\Sigma_2^{-1}),$

where $M = \Omega_a^{1/2}[\int_0^1 B_d(u)dB_m(u)']'[\int_0^1 B_d(u)B_d(u)'du]^{-1}\Omega_{a_1}^{-1/2}\Psi_{11}^{-1}$, $B_m(u) = \Omega_a^{1/2}Q$ W_m, $\Sigma_2 = E(U_{t-1}^\dagger U_{t-1}^{\dagger\prime} \otimes I_m)$, $\tilde{\Sigma}_2 = E(U_{t-1}^\dagger U_{t-1}^{\dagger\prime} \otimes V_t)$, and the remaining variables are defined in Lemma 3.1. $\square$

From Theorem 4.1(a), one can see that the asymptotic distribution of $\check{B}$ is the same as that in Johansen (1988), regardless of the presence of GARCH. On the other hand, if there is no heteroskedasticity and $V_t = V_0$, from Theorem 4.1(b), $\Sigma_2^{-1}\tilde{\Sigma}_2\Sigma_2^{-1} = [E(U_{t-1}^\dagger U_{t-1}^{\dagger\prime})]^{-1} \otimes V_0$ and the asymptotic distribution of $\check{\alpha}_2$ is exactly the same as that derived in Johansen (1988) or Ahn and Reinsel (1990).

Although the normalization in Johansen's estimation involves the true parameter B, as shown in Chapter 7 by Johansen (1995), the above distribution is found useful for hypothesis testing. In the following two sections, we follow this line and use the distribution of a similar *normalized* statistic (which incorporates GARCH) to devise a test for the reduced rank.

We conclude this sub-section with a discussion about imposing some prior information on the reduced rank structure. Note $C = -P_2(I_r - \Gamma_r)Q_2' = AB$. Further, we can always write $ABY_{t-1} = AB\mathcal{R}^{-1}\mathcal{R}Y_{t-1}$, where $\mathcal{R}$ is a permutation matrix so that the last d components of the series $\mathcal{R}Y_t$ are purely nonstationary, see Section 5 of Ahn and Reinsel (1990). In this representation, we may write $B\mathcal{R}^{-1} \equiv [D_1, D_2]$, where D_1 is invertible. As a result, $ABY_{t-1} = AD_1[I_r, B_0]\mathcal{R}Y_{t-1}$, where $B_0 \equiv D_1^{-1}D_2$.

Corollary 4.1. Let $\breve{B}$ and $\breve{A} \equiv \hat{A}(\hat{B}\bar{B}')$ be defined as in Theorem 4.1, $\breve{B} \equiv D_1^{-1}\breve{B}\mathcal{R}^{-1} = [\breve{B}_1, \breve{B}_2]$ and $\breve{\alpha}_2 \equiv vec[\breve{A}D_1\breve{B}_1, \hat{\Phi}_1^*, \ldots, \hat{\Phi}_{s-1}^*]$. If the assumptions in Lemma 3.1 hold, then it follows that

$$\begin{aligned} (a)\ n(\breve{B}_1^{-1}\breve{B}_2 - B_0) &= nD_1^{-1}(\breve{B} - B)P_1R_{21}^{-1} + O_p(n^{-1/2}) \\ &\longrightarrow_{\mathcal{L}} D_1^{-1}(A'V_0^{-1}A)^{-1}A'V_0^{-1}PMR_{21}^{-1}; \\ (b)\ \sqrt{n}(\breve{\alpha}_2 - \alpha_2^\dagger) &\longrightarrow_{\mathcal{L}} N(0, \Sigma_2^{-1}\tilde{\Sigma}_2\Sigma_2^{-1}). \end{aligned}$$

where $\mathcal{R}P \equiv \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}$, and $\alpha_2^\dagger \equiv vec[AD_1, \Phi_1^*, \ldots, \Phi_{s-1}^*]$. □

If, as in Ahn and Reinsel (1990) and LLW (2001), we are able to arrange the components of Y_t so that the last d components are purely nonstationary, we can impose the structure $B = [I_r, B_0]$, $D_1 = I_r$, $\mathcal{R} = I_m$ and $R_{21} = P_{21}$. Then, $\hat{B}_1^{-1}\hat{B}_2 = \breve{B}_1^{-1}\breve{B}_2 = \breve{B}_1^{-1}\breve{B}_2$. $\breve{\alpha}_2 = vec[\hat{A}\hat{B}_1, \hat{\Phi}_1^*, \ldots, \hat{\Phi}_{s-1}^*]$. In accordance with Corollary 4.1,

$$n(\hat{B}_1^{-1}\hat{B}_2 - B_0) \longrightarrow_{\mathcal{L}} (A'V_0^{-1}A)^{-1}A'V_0^{-1}PMP_{21}^{-1}; \tag{4.2}$$

$$\sqrt{n}(\breve{\alpha}_2 - \alpha_2) \longrightarrow_{\mathcal{L}} N(0, \Sigma_2^{-1}\tilde{\Sigma}_2\Sigma_2^{-1}). \tag{4.3}$$

The distribution in (4.2) is exactly the same as that in Ahn and Reinsel (1990).

4.2 Reduced Rank Estimation that Incorporates GARCH

In this sub-section, using Johansen's estimator $\hat{\alpha} = [\hat{\alpha}_1', \hat{\alpha}_2]'$, where $\hat{\alpha}_1 = vec[\hat{B}]$ and $\hat{\alpha}_2 = vec[\hat{A}, \hat{\Phi}_1^*, \ldots, \hat{\Phi}_{s-1}^*]$ (see Sub-section 4.1) and the other estimator in the full rank estimation, $\hat{\delta}$ (see Section 3), we propose a new reduced rank estimation that incorporates GARCH.

For the error correction form (4.1), the log-likelihood function l_t is the same as that in (3.2), but now it is a function of parameters α and δ. Denote $U_t^* \equiv [(Y_t \otimes A')', (U_t^\dagger \otimes I_m)']'$. Similar to (3.3) and (3.5),

$$\begin{aligned} \nabla_\alpha l_t &= -\frac{1}{2}(\nabla_\alpha h_t)(\iota - w(\varepsilon_t \varepsilon_t' V_t^{-1})) \odot \tilde{h}_t + U_{t-1}^* V_t^{-1} \varepsilon_t, \text{where} \quad (4.4) \\ \nabla_\alpha h_t &= -2 \sum_{l=1}^q U_{t-l-1}^* diag(a_l \odot \varepsilon_{t-l}) + \sum_{l=1}^p (\nabla_\alpha h_{t-l}) diag(b_l). \end{aligned}$$

Denote the reduced rank estimators as $\dot{\alpha} = (\dot{\alpha}_1', \dot{\alpha}_2')'$ and $\dot{\delta}$, where $\dot{\alpha}_1 = vec[\dot{B}]$ and $\dot{\alpha}_2 = vec[\dot{A}, \dot{\Phi}_1^*, \ldots, \dot{\Phi}_{s-1}^*]$. Similar to Theorem 4.1, in Theorem 4.2 below, inter alia, we derive the limiting distributions of $n(\dot{B}\bar{B}')^{-1}\dot{B}P_1$ and $\sqrt{n}\dot{A}(\dot{B}\bar{B}')$. As a result, $(\dot{B}\bar{B}')^{-1}\dot{B} - B = O_p(n^{-1})$ (see, for instance, p.179 and Lemma 13.2 by Johansen, 1995). Although the true B is involved, this normalization is found useful for deriving the distribution of the rank test in the next section.

Define $\bar{D}^{**} \equiv diag[\bar{D}_1^{**}, \sqrt{n} I_{rm+(s-1)m^2}]$, where $\bar{D}_1^{**} = diag(nI_{rd}, \sqrt{n} I_{r^2})$. As in Section 3, we also define $\mathcal{Q} \equiv diag[\mathcal{Q}_1, I_{rm+(s-1)m^2}]$, where $\mathcal{Q}_1 = (Q \otimes I_r)$.

Using Assumptions (a)-(c) and a method similar to the one used by Ling and Li (1998), we can show that:

$$n^{-1/2} \bar{D}^{**-1} \mathcal{Q} (\sum_{t=1}^n \nabla_{\alpha\delta'}^2 l_t) = o_p(1). \quad (4.5)$$

As a result, α and δ can be estimated separately without loss of efficiency. The estimation procedure for δ and its asymptotic properties are the same as those given in Theorem 3.1. In the following, we confine our attention to the estimation of α.

Consider the *block-diagonal* terms in $\nabla_{\alpha\alpha'}^2 l_t$. With reference to the arguments in Appendix A, similar to (3.9),

$$\begin{aligned} \bar{D}_1^{**-1} \mathcal{Q}_1 \sum_{t=1}^n \nabla_{\alpha_1\alpha_1'}^2 l_t \mathcal{Q}_1' \bar{D}_1^{**-1} &= \bar{D}_1^{**-1} \mathcal{Q}_1 \sum_{t=1}^n R_{1t} \mathcal{Q}_1' \bar{D}_1^{**-1} + o_p(1), \quad (4.6) \\ n^{-1} \sum_{t=1}^n \nabla_{\alpha_2\alpha_2'}^2 l_t &= n^{-1} \sum_{t=1}^n R_{2t} + o_p(1), \quad (4.7) \end{aligned}$$

where

$$R_{1t} = -Y_{t-1}Y'_{t-1} \otimes A'V_t^{-1}A - \frac{1}{4}(\nabla_{\alpha_1}h_t)D_t^{-2}(\Gamma^{-1} \odot \Gamma + I_m)D_t^{-2}(\nabla'_{\alpha_1}h_t), \text{ and}$$
$$R_{2t} = -U^{\dagger}_{t-1}U^{\dagger\prime}_{t-1} \otimes V_t^{-1} - \frac{1}{4}(\nabla_{\alpha_2}h_t)D_t^{-2}(\Gamma^{-1} \odot \Gamma + I_m)D_t^{-2}(\nabla'_{\alpha_2}h_t).$$

Moreover, as $(Q'_1 \otimes I_r)(\sum_{t=1}^n \nabla^2_{\alpha_1\alpha'_2} l_t)$ involves terms with $Q'_1 Y_{t-j}$ and $U^{\dagger}_{t-k}$ $(j, k = 1, 2, \ldots)$, by the arguments similar to those around (3.17),

$$n^{-3/2}(Q'_1 \otimes I_r)(\sum_{t=1}^n \nabla^2_{\alpha_1\alpha'_2} l_t) = o_p(1). \tag{4.8}$$

All in all, confining the attention to the limiting distributions of estimators for $(P'_1 \otimes I_r)\alpha_1 = vec[BP_1]$ and α_2, we consider the iterative Newton-Raphson algorithm:

$$\dot{\alpha}_1^{(k+1)} = \dot{\alpha}_1^{(k)} - (\sum_{t=1}^n R_{1t})^{-1}(\sum_{t=1}^n \nabla_{\alpha_1} l_t)|_{\dot{\alpha}_1^{(k)}, \dot{\alpha}_2^{(k)}, \dot{\delta}^{(k)}}; \tag{4.9}$$
$$\dot{\alpha}_2^{(k+1)} = \dot{\alpha}_2^{(k)} - (\sum_{t=1}^n R_{2t})^{-1}(\sum_{t=1}^n \nabla_{\alpha_2} l_t)|_{\dot{\alpha}_1^{(k+1)}, \dot{\alpha}_2^{(k)}, \dot{\delta}^{(k)}}, \tag{4.10}$$

where $\dot{\alpha}^{(k)}$ and $\dot{\delta}^{(k)}$ are the estimates at the k-th iteration. The initial estimator $\dot{\alpha}^{(0)} = \hat{\alpha}$ is the Johansen's estimator described in the previous sub-section; while the initial estimator $\dot{\delta}^{(0)} = \hat{\delta}$ is the full rank estimator described in Sub-section 3.1. As the *un-normalized* $\dot{\alpha}_1^{(k+1)}$ may not converge to a fixed vector, the iterations (4.11)-(4.12) are terminated until the likelihood value in (3.2) ceases to increase.

Similar to Theorem 4.1, in Theorem 4.2, we derive the limiting distributions of the *normalized* estimators $\ddot{\alpha}_1 \equiv (I_m \otimes (\dot{B}\bar{B}')^{-1})\dot{\alpha}_1$ and $\ddot{\alpha}_2 \equiv diag((\dot{B}\bar{B}')' \otimes I_m, I_{(s-1)m^2})\dot{\alpha}_2$.

Again, using the arguments similar to those for the univariate case in Ling, Li and McAleer (2002) and Ling and Li (2002), we can show that the following holds uniformly in the ball $\Theta_n = \{(\tilde{\alpha}, \tilde{\delta}) : \|\bar{D}^{**}\mathcal{Q}'^{-1}(\tilde{\alpha} - \alpha)\| \leq K \ \ and \ \ \|\sqrt{n}(\tilde{\delta} - \delta)\| \leq K\}$ for any fixed positive constant K:

$$\sum_{t=1}^n \bar{D}_1^{**-1}\mathcal{Q}_1(R_{1t}|_{\tilde{\alpha},\tilde{\delta}} - R_{1t})\mathcal{Q}'_1\bar{D}_1^{**-1} = o_p(1), n^{-1}\sum_{t=1}^n(R_{2t}|_{\tilde{\alpha},\tilde{\delta}} - R_{2t}) = o_p(1); \tag{4.11}$$
$$\sum_{t=1}^n \bar{D}_1^{**-1}\mathcal{Q}_1(\nabla_{\alpha_1} l_t|_{\tilde{\alpha},\tilde{\delta}} - \nabla_{\alpha_1} l_t) = \sum_{t=1}^n \bar{D}_1^{**-1}\mathcal{Q}'_1 R_{1t}(\tilde{\alpha}_1 - \alpha_1) + o_p(1), \text{ and}$$

$$n^{-1/2}\sum_{t=1}^{n}(\nabla_{\alpha_2} l_t|_{\tilde{\alpha},\tilde{\delta}} - \nabla_{\alpha_2} l_t) = n^{-1/2}\sum_{t=1}^{n} R_{2t}(\tilde{\alpha}_2 - \alpha_2) + o_p(1). \quad (4.12)$$

The estimator $\dot{\alpha}$ obtained by (4.11)-(4.12), after normalization (the *normalized* estimator will be denoted as $\ddot{\alpha} = (\ddot{\alpha}_1', \ddot{\alpha}_2')'$), as one can see in the proof of Theorem 4.2, satisfies $\bar{D}^{**}\mathcal{Q}'^{-1}(\ddot{\alpha} - \alpha) = O_p(1)$, since the *normalized* initial estimator (denoted as $\check{\alpha} = (\check{\alpha}_1', \check{\alpha}_2')'$) does (see Theorem 4.1).

Theorem 4.2. Let $\ddot{B} = (\dot{B}\bar{B}')^{-1}\dot{B}$ and $\ddot{\alpha}_2 = vec[\dot{A}(\dot{B}\bar{B}'), \dot{\Phi}_1^*, \ldots, \dot{\Phi}_{s-1}^*]$. If the assumptions in Lemma 3.1 hold, then

$$(a) \qquad n(\ddot{B} - B)P_1 \longrightarrow_{\mathcal{L}} (A'\Omega_1 A)^{-1}A'\tilde{M}, \text{ and thus } n(\ddot{B} - B) = O_p(1);$$

$$(b) \qquad \sqrt{n}(\ddot{\alpha}_2 - \alpha_2) \longrightarrow_{\mathcal{L}} N(0, \Omega_2^{\dagger-1}\tilde{\Omega}_2^{\dagger}\Omega_2^{\dagger-1}),$$

where $\tilde{M}$ is defined in Theorem 3.1,

$$\Omega_2^{\dagger} \equiv E(U_{t-1}^{\dagger}U_{t-1}^{\dagger\prime} \otimes V_t^{-1}) + \sum_{j=1}^{\infty} E(U_{t-j-1}^{\dagger}U_{t-j-1}^{\dagger\prime} \otimes (\Gamma^{-1} \odot \Gamma + I_m) \odot \nu_j\nu_j' \odot \Pi_{jt}),$$

$$\tilde{\Omega}_2^{\dagger} \equiv E(U_{t-1}^{\dagger}U_{t-1}^{\dagger\prime} \otimes V_t^{-1}) + \sum_{j=1}^{\infty} E(U_{t-j-1}^{\dagger}U_{t-j-1}^{\dagger\prime} \otimes (\Delta - \iota\iota') \odot \nu_j\nu_j' \odot \Pi_{jt}),$$

and the remaining variables are defined as in Lemma 3.1. □

In order to prove Theorem 4.2, as in Section 3, we define $\bar{Q}^{**} = diag(\bar{Q}_1^{**}, \bar{Q}_2^{**})$, where

$$\bar{Q}_1^{**} \equiv (Q \otimes (\dot{B}\bar{B}')') = \mathcal{Q}_1(I_m \otimes (\dot{B}\bar{B}')'), \text{ and} \quad (4.13)$$

$$\bar{Q}_2^{**} \equiv diag((\dot{B}\bar{B}')^{-1} \otimes I_m, I_{(s-1)m^2}). \quad (4.14)$$

This choice of $\bar{Q}^{**}$ facilitates the derivation of the distribution of $\sqrt{n}(\dot{A}(\dot{B}\bar{B}') - A)$, and that of $n((\dot{B}\bar{B}')^{-1}\dot{B} - B)P_1$. $\bar{Q}_1^{**}$ and $\bar{Q}_2^{**}$ are also used in Lemmas B.4 and B.5, both of which are preliminaries for the Proof of Theorem 4.2.

Next consider estimating ABP_2 rather than A. Let $\alpha_3 = vec[ABP_2, \Phi_1^*, \ldots, \Phi_{s-1}^*]$ and its normalized estimator $\ddot{\alpha}_3 = vec[\dot{A}(\dot{B}\bar{B}')BP_2, \dot{\Phi}_1^*, \ldots, \dot{\Phi}_{s-1}^*]$. It is not difficult to see from Theorem 4.2(b) that:

$$\sqrt{n}(\ddot{\alpha}_3 - \alpha_3) \longrightarrow_{\mathcal{L}} N(0, \Omega_2^{-1}\tilde{\Omega}_2\Omega_2^{-1}),$$

which is exactly the same distribution of $\sqrt{n}(\hat{\beta}_2 - \beta_2)$ in Theorem 3.1(b). This result will be used in the next section.

We conclude this sub-section with a discussion about imposing some prior information on the reduced rank structure.

Corollary 4.2. Let $\ddot{A} = \dot{A}(\dot{B}\bar{B}')$, $\tilde{B} \equiv D_1^{-1}\ddot{B}\mathcal{R}^{-1} = [\tilde{B}_1, \tilde{B}_2]$ and $\tilde{\alpha}_2 \equiv vec[\ddot{A}D_1\tilde{B}_1, \dot{\Phi}_1^*, \ldots, \dot{\Phi}_{s-1}^*]$. Suppose the assumptions in Lemma 3.1 hold. Then:

$$\begin{aligned}
(a)\ \ n(\tilde{B}_1^{-1}\tilde{B}_2 - B_0) \quad &= \quad nD_1^{-1}(\ddot{B} - B)P_1R_{21}^{-1} + O_p(n^{-1/2}) \\
&\longrightarrow_{\mathcal{L}} \quad D_1^{-1}(A'\Omega_1 A)^{-1}A'\tilde{M}R_{21}^{-1}, \\
(b)\quad \sqrt{n}(\tilde{\alpha}_2 - \alpha_2^{\dagger}) \quad &\longrightarrow_{\mathcal{L}} \quad N(0, \Omega_2^{\dagger -1}\tilde{\Omega}_2^{\dagger}\Omega_2^{\dagger -1}),
\end{aligned}$$

where $\ddot{B}$ is defined as in Theorem 4.1, $\tilde{M}$ is defined in Theorem 3.1, and D_1, $\mathcal{R}$, R_{21} and $\alpha_2^{\dagger}$ are as defined in Corollary 4.1. $\square$

If, as in Ahn and Reinsel (1990) and LLW (2001), we are able to arrange the components of Y_t so that the last d components are purely nonstationary, then $\tilde{\alpha}_2 = vec[\dot{A}\dot{B}_1, \dot{\Phi}_1^*, \ldots, \dot{\Phi}_{s-1}^*]$. Following the arguments right after Corollary 4.1 and by Corollary 4.2,

$$n(\dot{B}_1^{-1}\dot{B}_2 - B_0) \quad \longrightarrow_{\mathcal{L}} \quad (A'\Omega_1 A)^{-1}A'\tilde{M}P_{21}^{-1}, \tag{4.15}$$

$$\sqrt{n}(\tilde{\alpha}_2 - \alpha_2) \quad \longrightarrow_{\mathcal{L}} \quad N(0, \Omega_2^{\dagger -1}\tilde{\Omega}_2^{\dagger}\Omega_2^{\dagger -1}). \tag{4.16}$$

The distribution in (4.15) is essentially the same as that in LLW (2001), with slightly different definitions of Ω_1 and $\tilde{W}_m(u)$, due to a different model of conditional heteroskedasticity.

5 Testing for the Reduced Rank

In this section, we apply the asymptotic distributions derived in Section 3 (that for the full rank estimation) and in Section 4 (that for the reduced rank estimation) to devise a test for the reduced rank. Both estimators incorporate GARCH.

Consider the null hypothesis,

$$H_0 : rank(C) = r < m \text{ vs } H_a : rank(C) = m. \tag{5.1}$$

Here we consider the LR test. To incorporate GARCH, instead of the ratio of the residual sum of squares, here we consider the general form:

$$LR_G \equiv 2l(\hat{\beta}_1) - 2l(\dot{\beta}_1), \tag{5.2}$$

where $l(.)$ is the log-likelihood function defined in (3.2). $\hat{\beta}_1 \equiv vec(\hat{C}P_1)$ is the full rank estimator. $\dot{\beta}_1 \equiv vec(\dot{A}\dot{B}P_1) = vec(\ddot{A}\ddot{B}P_1) \equiv \ddot{\beta}_1$ is the reduced rank estimator. Note in both estimations, the estimators for $\beta_2 = CP_2 = ABP_2$ and those for δ are unaffected (asymptotically) by whether or not the reduced rank is imposed on C. As a result, using a second-order Taylor expansion around the true parameter $\beta_1 = vec(CP_1) = vec(ABP_1)$,

$$\begin{aligned} LR_G &= (\hat{\beta}_1 - \beta_1)'(\sum_{t=1}^{n} L_{1t})(\hat{\beta}_1 - \beta_1) - (\dot{\beta}_1 - \beta_1)'(\sum_{t=1}^{n} L_{1t})(\dot{\beta}_1 - \beta_1) + o_p(1) \\ &= (\hat{\beta}_1 - \beta_1)'(\sum_{t=1}^{n} L_{1t})(\hat{\beta}_1 - \beta_1) - (\ddot{\beta}_1 - \beta_1)'(\sum_{t=1}^{n} L_{1t})(\ddot{\beta}_1 - \beta_1) + o_p(1) \end{aligned}$$

where L_{1t} is as defined in Section 3.

Using Lemma 3.1, Theorem 3.1 and Theorem 4.1, we derive the asymptotic distribution of LR_G, as follows.

Lemma 5.1. Suppose the assumptions in Lemma 3.1 hold. Under the null hypothesis in (5.1), the LR statistic in (5.2),

$$LR_G \longrightarrow_{\mathcal{L}} tr[(\int_0^1 B_d(u)d\tilde{V}_d(u)')'(\int_0^1 B_d(u)B_d(u)'du)^{-1}(\int_0^1 B_d(u)d\tilde{V}_d(u)')],$$

where $\tilde{V}_d(u) = \Upsilon B_d(u) + [(Q_1'\Omega_1^{-1}Q_1)^{-1/2}Q_1'\Omega_1^{-1}\tilde{\Omega}_1\Omega_1^{-1}Q_1(Q_1'\Omega_1^{-1}Q_1)^{-1/2} - \Upsilon\Upsilon']^{1/2} V_d(u)$, $\Upsilon = (Q_1'\Omega_1^{-1}Q_1)^{1/2}(Q_1'V_0Q_1)^{-1/2}$, and $(B_d'(u), V_d'(u))'$ is a $2d-$dimensional standard Brownian motion. □

When there is no heteroskedasticity, all the ν_j's defined in Section 3 are 0. Refer to (3.18), (3.19), and the definition of Ω_1 in Lemma 3.1. $\tilde{\Omega}_1 = \Omega_1 =$

V_0^{-1}, and thus $\tilde{V}_d(u) = B_d(u)$ The distribution in Lemma 5.1 can be simplified as $tr[(\int_0^1 B_d(u)dB_d(u)')'(\int_0^1 B_d(u)B_d(u)'du)^{-1}(\int_0^1 B_d(u)dB_d(u)')]$, which is exactly the same as that in Johansen (1988,1995), or that in Reinsel and Ahn (1992). On the other hand, when there is heteroskedasticity, but $\tilde{\Omega}_1 = \Omega_1$, the distribution can be simplified as that in the next theorem.

Theorem 5.1. Suppose the assumptions in Lemma 5.1 hold. If, in addition, $\tilde{\Omega}_1 = \Omega_1$,

$$LR_G \longrightarrow_{\mathcal{L}} tr\{[\zeta(I_d - \Lambda_d)^{1/2} + \Phi\Lambda_d^{1/2}]'[\zeta(I_d - \Lambda_d)^{1/2} + \Phi\Lambda_d^{1/2}]\}, \tag{5.3}$$

where Λ_d is a diagonal matrix containing the d eigenvalues of $(I_d - \Upsilon\Upsilon')$, $\Phi = [\int_0^1 B_d(u)B_d(u)'du]^{-1/2}\int_0^1 B_d(u)dV_d(u)'$ is a $d-$dimensional standard normal vector independent of $\zeta = [\int_0^1 B_d(u)B_d(u)'du]^{-1/2}\int_0^1 B_d(u)dB_d(u)'$. $\square$

When $\tilde{\Omega}_1 \neq \Omega_1$, alternatively we may consider the LR-type test statistic:

$$LR_G^* \equiv vec(\hat{C}^* - \dot{A}\dot{B}^*)'(-\sum_{t=1}^{n}\dot{F}_t^*)vec(\hat{C}^* - \dot{A}\dot{B}^*), \tag{5.4}$$

where $vec(\hat{C}^*) = (\sum_{t=1}^n \dot{F}_t^*)^{-1}(\sum_{t=1}^n \dot{F}_t)vec(\hat{C})$, $\dot{B}^* = (\dot{A}'\dot{\tilde{\Omega}}_1\dot{A})^{-1}(\dot{A}'\dot{\Omega}_1\dot{A})\dot{B}$. F_t^* is the F_t in (3.9) with $(\Gamma^{-1}\odot\Gamma + I_m)$ replaced by $(\Delta - \iota\iota')$. The asymptotic distribution is derived in the following corollary.

Corollary 5.1. Suppose the assumptions in Lemma 5.1 hold. The LR-type test statistic in (5.4),

$$LR_G^* \longrightarrow_{\mathcal{L}} tr\{[\zeta(I_d - \Lambda_d^*)^{1/2} + \Phi\Lambda_d^{*1/2}]'[\zeta(I_d - \Lambda_d^*)^{1/2} + \Phi\Lambda_d^{*1/2}]\}, \tag{5.5}$$

where Λ_d^* is a diagonal matrix containing the d eigenvalues of $(I_d - (Q_1'\tilde{\Omega}_1^{-1}Q_1)^{1/2}(Q_1'V_0 Q_1)^{-1}(Q_1'\tilde{\Omega}_1^{-1}Q_1)^{1/2})$. $\square$

As an illustration, in Tables C.10 and C.20, we tabulate the critical values for $d = 1$ and $d = 2$, respectively. More precisely, for each d and each set of d eigenvalues (refer to Λ_d in (5.3)), $100,000$ replications of the quantity $tr\{[\hat{\zeta}(I_d - \Lambda_d)^{1/2} + \hat{\Phi}\Lambda_d^{1/2}]'[\hat{\zeta}(I_d - \Lambda_d)^{1/2} + \hat{\Phi}\Lambda_d^{1/2}]\}$ are drawn, where

$$\hat{\zeta} = (n^{-2}\sum_{t=1}^{T} z_{1t-1}z_{1t-1}')^{-1/2}(n^{-1}\sum_{t=1}^{T} z_{1t-1}a_{1t}'),$$

$$\hat{\Phi} \;=\; (n^{-2}\sum_{t=1}^{T} z_{1t-1}z_{1t-1}')^{-1/2}(n^{-1}\sum_{t=1}^{T} z_{1t-1}a_{2t}'),$$

where $T = 2,000$, z_{1t} is a d-dimensional random walk process $z_{1t} = z_{1t-1} + a_{1t}$, and $[a_{1t}', a_{2t}']'$ is generated from an i.i.d. $N(0, I_{2d})$ process.

6 Inclusion of a Constant Term

The results in Sections 3-5 can be extended to Model (1.1)-(1.3) with a constant term. That is, we modify Equation (2.1) as:

$$W_t = CY_{t-1} + \Phi_1^* W_{t-1} + \cdots + \Phi_{s-1}^* W_{t-s+1} + \varepsilon_t + \mu_0, \tag{6.1}$$

where $\mu_0 \neq 0$. Similar to the analysis in Section 2, we consider $Z_t = QY_t$, where

$$Z_t = \text{diag}(I_d, \Gamma_r)Z_{t-1} + u_t. \tag{6.2}$$

However, in view of the drift term, u_t above has to be modified as:

$$u_t = Q[\Phi_1^* W_{t-1} + \cdots + \Phi_{s-1}^* W_{t-s-1} + \varepsilon_t + \mu_0].$$

On the other hand, it is not difficult to show that $E[u_t] = Q\mu_0$. Note that by the definition of Q, either $Q_1'\mu_0 \neq 0$ or $Q_2'\mu_0 \neq 0$, but not both. In the rest of the paper, we consider the latter case and leave the equally interesting former case to further research.

Since $Q_1'\mu_0 = 0$, as in Equation (2.6) above, $Z_{1t} - Z_{1t-1}$ has a zero mean and the Z_{1t} is defined as before. Moreover, as $Q_2'\mu_0 \neq 0$, $Z_{2t} \equiv Q_2'Y_t$ has a non-zero mean. Along the lines of Section 4 in Reinsel and Ahn (1992), we consider the following equation:

$$W_t - \bar{W}_0 = AB(Y_{t-1} - \bar{Y}_1)' + \Phi_1^*(W_{t-1} - \bar{W}_1) + \cdots + \Phi_{s-1}^*(W_{t-s+1} - \bar{W}_s) + \varepsilon_t, \tag{6.3}$$

where $\bar{W}_j$ and $\bar{Y}_j$ are the sample mean of W_{t-j} and that of Y_{t-j}, respectively.

The major theorems 3.1, 4.1, 4.3 and 5.1 still hold, with the d-dimensional standard Brownian motion $B_d(u)$ replaced by $[B_d(u) - \int_0^1 B_d(u)du]$. In particular, for the LR statistic in (5.3), we have the following theorem.

Theorem 5.1'. Suppose the assumptions in Lemma 5.1 hold for an AR model with a constant term. Consider the variables defined there. If, in addition, $\tilde{\Omega}_1 = \Omega_1$,

$$LR_G \longrightarrow_{\mathcal{L}} tr\{[\bar{\zeta}(I_d - \Lambda_d)^{1/2} + \bar{\Phi}\Lambda_d^{1/2}]'[\bar{\zeta}(I_d - \Lambda_d)^{1/2} + \bar{\Phi}\Lambda_d^{1/2}], \qquad (6.4)$$

where Λ_d is a diagonal matrix containing the d eigenvalues of $(I_d - \Upsilon\Upsilon')$, $\bar{\Phi} = [\int_0^1 \bar{B}_d(u)\bar{B}_d(u)'du]^{-1/2}\int_0^1 \bar{B}_d(u)dV_d(u)'$ is a $d-$dimensional standard normal vector independent of $\bar{\zeta} = [\int_0^1 \bar{B}_d(u)\bar{B}_d(u)'du]^{-1/2}\int_0^1 \bar{B}_d(u)dB_d(u)'$, $\bar{B}_d(u) \equiv [B_d(u) - \int_0^1 B_d(u)du]$ □

As an illustration, in Tables C.11 and C.21, we tabulate the critical values for $d = 1$ and $d = 2$, respectively. As in Section 5, for each d and each set of d eigenvalues (refer to Λ_d in (6.4)), $100,000$ replications of the quantity $tr\{[\hat{\bar{\zeta}}(I_d - \Lambda_d)^{1/2} + \hat{\bar{\Phi}}\Lambda_d^{1/2}]'[\hat{\bar{\zeta}}(I_d - \Lambda_d)^{1/2} + \hat{\bar{\Phi}}\Lambda_d^{1/2}]\}$ are drawn, where

$$\begin{aligned}\hat{\bar{\zeta}} &= (n^{-2}\sum_{t=1}^{T}[z_{1t-1} - \bar{z}_1][z_{1t-1} - \bar{z}_1]')^{-1/2}(n^{-1}\sum_{t=1}^{T}[z_{1t-1} - \bar{z}_1]a_{1t}'),\\ \hat{\bar{\Phi}} &= (n^{-2}\sum_{t=1}^{T}[z_{1t-1} - \bar{z}_1][z_{1t-1} - \bar{z}_1]')^{-1/2}(n^{-1}\sum_{t=1}^{T}[z_{1t-1} - \bar{z}_1]a_{2t}'),\end{aligned}$$

where $T = 2,000$, $\bar{z}_1$ is the sample mean of z_{1t}, where z_{1t} is a d-dimensional random walk process $z_{1t} = z_{1t-1} + a_{1t}$, and $[a_{1t}', a_{2t}']'$ is generated from an i.i.d. $N(0, I_{2d})$ process.

7 Monte Carlo Experiments

In this section, the finite-sample size and power of our LR test (LR_G) are examined with Monte Carlo experiments. Throughout, we consider testing the $H_0 : rank(C) = 1$ (see (5.1)). With $\Gamma = I_m$, the GARCH error ε_t (see (1.3)) is generated as:

$$\varepsilon_{it} = \eta_{it}\sqrt{h_{it}}, \;\; h_{it} = 0.1 + 0.3\varepsilon_{it-j}^2 + 0.6h_{it-k}, \;\; \eta_{it} \sim i.i.d.N(0,1).$$

In the first part of the experiments, we consider a bi-variate $AR(1)$ model (without a constant term). The matrix C in the error correction form (4.1) is:

$$DGP(1a) \qquad C = AB, \;\; A = \begin{pmatrix} -0.4 \\ 0.12 \end{pmatrix}, \;\; B = (1.0, -2.5).$$

$$DGP(1b) \qquad C = \kappa I_2, \;\; \kappa = 0.1.$$

$$DGP(1c) \qquad C = \kappa I_2, \;\; \kappa = 0.5.$$

In the second part of the experiments, we consider a tri-variate $AR(1)$ model (without a constant term). The matrix C is:

$$DGP(2a) \qquad C = AB, \;\; A = \begin{pmatrix} -0.4 \\ 0.12 \\ 0.12 \end{pmatrix}, \;\; B = (1.0, -2.5, 0.0).$$

$$DGP(2b) \qquad C = \kappa I_3, \;\; \kappa = 0.1.$$

$$DGP(2c) \qquad C = \kappa I_3, \;\; \kappa = 0.5.$$

Both DGP(1a) and DGP(2a) are the DGP under the null. However, the former DGP contains 1 unit root while the latter contains 2. DGP(1b) and DGP(1c) are the DGP under the alternative. The latter DGP is expected to reject the null more frequently, as the κ is larger. This is also the case with DGP(2b) and DGP(2c).

For each DGP, time series with different number of observations are generated. Reduced rank estimation with and without GARCH are done. The empirical means and standard deviations of the estimated A and B for DGP(1a) and DGP(2a) (the null models) are reported in Tables 7.1 and 7.2 respectively. In general, the biases of the estimators with GARCH are comparable to, if not smaller than, those without GARCH. The standard deviations and the mean squared errors are definitely smaller, even when the sample size is as small as 200.

TABLE 7.1

Empirical means and standard deviations of the estimated A and B

DGP(1a) A Bi-variate System

			$A_1 = -0.4$	$A_2 = 0.12$	$B_2 = -2.5$
n=200	*No GARCH*	Mean	-0.4047	0.1204	-2.5004
		SD	0.0310	0.0359	0.0506
	With GARCH	Mean	-0.4030	0.1199	-2.4993
		SD	0.0229	0.0262	0.0396
n=400	*No GARCH*	Mean	-0.4019	0.1210	-2.4999
		SD	0.0212	0.0269	0.0235
	With GARCH	Mean	-0.4012	0.1203	-2.4999
		SD	0.0146	0.0185	0.0181
n=800	*No GARCH*	Mean	-0.4008	0.1201	-2.5000
		SD	0.0161	0.0209	0.0116
	With GARCH	Mean	-0.4006	0.1200	-2.4998
		SD	0.0105	0.0124	0.0088

number of replications = 1,000.

TABLE 7.2

Empirical means and standard deviations of the estimated A and B

DGP(2a) A Tri-variate System

			$A_1 = -0.4$	$A_2 = 0.12$	$A_3 = 0.12$	$B_2 = -2.5$	$B_3 = 0.0$
n=200	*No GARCH*	Mean	-0.4026	0.1244	0.1219	-2.5038	-0.0007
		SD	0.0334	0.0343	0.0265	0.0717	0.0395
	With GARCH	Mean	-0.4019	0.1219	0.1209	-2.5001	-0.0010
		SD	0.0229	0.0265	0.0204	0.0589	0.0322
n=400	*No GARCH*	Mean	-0.4016	0.1228	0.1220	-2.5012	-0.0002
		SD	0.0225	0.0268	0.0179	0.0351	0.0188
	With GARCH	Mean	-0.4010	0.1211	0.1210	-2.5011	-0.0002
		SD	0.0149	0.0176	0.0137	0.0283	0.0134
n=800	*No GARCH*	Mean	-0.4009	0.1210	0.1207	-2.5007	0.0000
		SD	0.0165	0.0213	0.0126	0.0164	0.0095
	With GARCH	Mean	-0.4002	0.1203	0.1204	-2.5006	-0.0002
		SD	0.0103	0.0125	0.0094	0.0122	0.0065

number of replications = 1,000.

The LR_G is computed with the procedure described in Section 5. For comparison, we also compute the LR test (also known as the trace test in the literature) suggested in Johansen (1988,1995) (henceforth LR_{NG}). Note this test is (asymptotically) equivalent to the LR test suggested in Reinsel and Ahn (1992), if we impose the prior knowledge of the reduced rank structure. For LR_{NG}, the critical value can be found in Table I of Reinsel and Ahn (1992) or Table 15.1 in Johansen (1995). For LR_G, we first estimate the eigenvalue(s) of the matrix $I_d - \Upsilon\Upsilon'$ with the reduced rank estimation, the $p - value$ are then approximated with 100,000 simulations of the limiting distribution in (5.3), using the method described in Section 5 with T equals the actual number of observations.

Rejection frequencies are summarized in Tables 7.3 and 7.4, respectively. Both LR_{NG} and LR_G are of the reasonably correct finite-sample size, even when the number of observations is as small as 200. Both tests slightly over-reject when there are 2 unit roots and the sample size is 200 or 400, and the over-rejections are comparable. From the two tables, it is clear that LR_G is more powerful than LR_{NG}.

TABLE 7.3

Rejection Fequency of Testing for the Reduced Rank $H_0 : r = 1$

Part (1) A Bi-variate System

n	size = 0.05				size = 0.10			
	(DGP)	(1a)	(1b)	(1c)	(DGP)	(1a)	(1b)	(1c)
200	LR_{NG}	0.062	0.019	0.461	LR_{NG}	0.106	0.054	0.733
	LR_G	0.063	0.042	0.655	LR_G	0.097	0.113	0.824
400	LR_{NG}	0.048	0.049	0.972	LR_{NG}	0.095	0.142	0.994
	LR_G	0.061	0.184	0.992	LR_G	0.112	0.336	0.998
800	LR_{NG}	0.052	0.294	0.999	LR_{NG}	0.109	0.561	1.000
	LR_G	0.044	0.596	1.000	LR_G	0.109	0.784	1.000

number of replications = 1,000.

TABLE 7.4

Rejection Fequency of Testing for the Reduced Rank $H_0 : r = 1$
Part (2) A Tri-variate System

n	size = 0.05				size = 0.10			
	(DGP)	(2a)	(2b)	(2c)	(DGP)	(2a)	(2b)	(2c)
200	LR_{NG}	0.077	0.006	0.361	LR_{NG}	0.151	0.020	0.545
	LR_G	0.079	0.023	0.641	LR_G	0.135	0.076	0.767
400	LR_{NG}	0.071	0.022	0.972	LR_{NG}	0.127	0.062	0.995
	LR_G	0.078	0.114	0.997	LR_G	0.139	0.230	0.998
800	LR_{NG}	0.058	0.160	1.000	LR_{NG}	0.107	0.329	1.000
	LR_G	0.051	0.603	1.000	LR_G	0.107	0.774	1.000

number of replications = 1,000.

8 An Empirical Example

In this section, we fit our model to the logarithms of three US monthly interest rates, which are well-known of having GARCH effect. See, for instance, Bollerslev (1990). The series are the federal fund rate, the 90-day treasury bill rate, and the one-year treasury bill rate. For comparison with the results in Reinsel and Ahn (1992), we use the series from January 1960 to December 1979 and thus we have 240 observations.

As expected, the estimated AR parameters are comparable to those obtained if we ignore GARCH. However, unsurprisingly, our LR test for the reduced rank behaves differently from the one that ignores GARCH. Furthermore, it is well-known that this type of tests may be sensitive to the order of this model. Consequently, we try different s of an $AR(s)$ model, where $s = 1, ..., 6$, all with a constant term. The LR test results are summarized in Table 8.1, where the $p - value$ is approximated using the method described in Section 7.

TABLE 8.1

LR Test Statistics

	$H_0 : r = 0$		$H_0 : r = 1$		$H_0 : r = 2$	
s	LR_{NG}	LR_G	LR_{NG}	LR_G	LR_{NG}	LR_G
1	69.70(0.000)	71.17(0.000)	16.21(0.084)	53.97(0.000)	0.107(0.939)	42.048(0.000)
2	49.07(0.000)	29.08(0.003)	13.02(0.210)	16.32(0.009)	0.757(0.801)	1.519(0.259)
3	34.86(0.019)	11.89(0.381)	11.35(0.320)	8.10(0.147)	0.692(0.814)	4.317(0.056)
4	46.43(0.000)	52.25(0.000)	15.60(0.101)	31.01(0.000)	1.044(0.748)	33.044(0.000)
5	47.26(0.000)	56.77(0.000)	13.25(0.197)	19.54(0.003)	1.231(0.714)	18.168(0.000)
6	38.65(0.006)	56.77(0.000)	13.43(0.188)	19.54(0.003)	1.085(0.741)	18.168(0.000)

p-values are in brackets.

Table 8.1 clearly shows the hypothesis that $r = 0$ is rejected by both tests. Interestingly, while the LR_{NG} can hardly reject or only marginally rejects (when $s = 4$) the null of $r = 1$, our LR_G clearly rejects it, except possibly when $s = 3$. Note that unlike Reinsel and Ahn (1992), we have not made the finite-sample modification for LR_{NG}.

As with other empirical findings in the literature, the LR_{NG} does not reject the null of $r = 2$. In those studies, it is natural to conclude that the reduced rank equals 2 and, in other words, the interest rates are nonstationary and there exist 2 cointegrating vectors. Somewhat surprisingly, due to the high power of LR_G or other factors, it rejects the null of $r = 2$, except possibly when $s = 2$ and the $p-value$ may be unreliable due to the insufficient number of lags. (In fact, in both estimations, there is a substantial increase in the likelihood value, when s increases from 3 to 4.)

All in all, unlike LR_{NG}, the LR_G strongly rejects that the reduced rank is 1. In fact, judging from the LR_G, there is some evidence that the rank is 3, in other words, the interest rates are stationary.

9 Conclusions

Macroeconomic or financial data are often modelled with cointegration and conditional heteroskedasticity, such as GARCH. However, the asymptotic theory and the statistical inference method for the cointegration with GARCH errors have yet to be well developed in the literature. In this paper, we consider a partially nonstationary autoregressive model with GARCH. In addition, no prior knowledge of the reduced rank structure is assumed. We propose the full rank and the reduced rank quasi-maximum likelihood estimation for the model. The estimation is modified upon the conventional reduced rank regression. The asymptotic distributions of the estimators are derived to be a functional of two correlated high-dimensional Brownian motions. These results are used to construct a LR test for the reduced rank. It is shown that the asymptotic distribution of the LR test is a functional of the standard Brownian motion and the standard normal vector with some unknown nuisance parameters. The critical values of the LR test are tabulated. The performance of this test in finite-sample is examined through Monte Carlo experiments. We also apply our approach to an empirical example of three interest rates.

With GARCH in the data series generated in the Monte Carlo experiments, our test for the reduced rank shows substantial improvement upon the conventional trace test suggested in Johansen (1988,1995), which is asymptotically equivalent to Reinsel and Ahn (1992)'s LR test. On the other hand, in contrast to the empirical results in the existing literature, our LR test shows evidence that the US monthly interest rates are stationary. In fact, our empirical result is more in line with the common belief that the US interest rates are controllable under the stabilization mechanism of the US Federal Reserve Board. We conjecture that this conclusion is reached as our test becomes more powerful when heteroskedasticity is taken into account.

On the other hand, in the Monte Carlo experiments and the empirical study, we assume that the data are normally distributed. Though results can easily be robustified by considering the more involved Lemma 5.1 or Corollary 5.1, the crucial issue is: Will the power of our test be further improved if other distributional assumptions are taken into account? This challenging topic is undertaken along the line of adaptive estimation suggested in Ling and McAleer (2002b).

A Appendix to Sub-section 3.1

Given (3.3), considering the derivative of $\nabla_\varphi l_t$ w.r.t. φ, we obtain:

$$
\begin{aligned}
&\bar{D}^{*-1}\bar{Q}^*(\sum_{t=1}^n \nabla^2_{\varphi\varphi'} l_t)\bar{Q}^{*'}\bar{D}^{*-1} \\
&\quad = -\sum_{t=1}^n \bar{D}^{*-1}\bar{Q}^*[X_{t-1}X'_{t-1}\otimes V_t^{-1} + \frac{1}{4}(\nabla_\varphi h_t)\Lambda_t(\nabla'_\varphi h_t)]\bar{Q}^{*'}\bar{D}^{*-1} + o_p(1),
\end{aligned}
$$

where $\Lambda_t = (D_t^{-2}\varepsilon_t\varepsilon_t' D_t^{-2} \odot V_t^{-1}) + D_t^{-4} dg(\varepsilon_t\varepsilon_t' V_t^{-1})$, $dg(A)$ is a diagonal matrix containing the diagonal elements of A. Note $E[\Lambda_t|\mathcal{F}_{t-1}] = (D_t^{-1}\Gamma D_t^{-1}\odot V_t^{-1}) + D_t^{-4} = D_t^{-2}(\Gamma^{-1}\odot\Gamma + I_m)D_t^{-2}$. In view of this and (3.5), the F_t in (3.10) can be defined and expressed as:

$$
\begin{aligned}
F_t &\equiv -(X_{t-1}X'_{t-1}\otimes V_t^{-1}) - \frac{1}{4}(\nabla_\varphi h_t)D_t^{-2}(\Gamma^{-1}\odot\Gamma + I_m)D_t^{-2}(\nabla'_\varphi h_t) \qquad \text{(A. 1)} \\
&= -(X_{t-1}X'_{t-1}\otimes V_t^{-1}) - \sum_{j=1}^{t-1}(X_{t-j-1}X'_{t-j-1}\otimes(\Gamma^{-1}\odot\Gamma + I_m)\odot\nu_j\nu_j'\odot\Pi_{jt}),
\end{aligned}
$$

where Π_{jt} is defined around (3.17). Similarly,

$$
\begin{aligned}
n^{-1}(\nabla^2_{\delta_1\delta_1'} l) &= -n^{-1}\sum_{t=1}^n \frac{1}{4}(\nabla_{\delta_1} h_t)\Lambda_t(\nabla'_{\delta_1} h_t) + o_p(1) \\
&= -n^{-1}\sum_{t=1}^n \frac{1}{4}(\nabla_{\delta_1} h_t)D_t^{-2}(\Gamma^{-1}\odot\Gamma + I_m)D_t^{-2}(\nabla'_{\delta_1} h_t) + o_p(1). \qquad \text{(A. 2)}
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
n^{-1}(\nabla^2_{\delta_2\delta_1'} l) &= -n^{-1}\sum_{t=1}^n (\nabla_{\delta_1} h_t)\Psi_m(D_t^{-1}\varepsilon_t\varepsilon_t' D_t^{-1}\Gamma^{-1}\otimes D_t^{-2}\Gamma^{-1})N_m\tilde{L}'_m \\
&= -n^{-1}\sum_{t=1}^n (\nabla_{\delta_1} h_t)\Psi_m(I_m\otimes D_t^{-2}\Gamma^{-1})N_m\tilde{L}'_m + o_p(1). \qquad \text{(A. 3)}
\end{aligned}
$$

$$
\begin{aligned}
n^{-1}(\nabla^2_{\delta_2\delta_2'} l) &= -n^{-1}\sum_{t=1}^{n} \tilde{L_m} N_m[\Gamma^{-1}\otimes(4\Gamma^{-1}D_t^{-1}\varepsilon_t\varepsilon_t' D_t^{-1}\Gamma^{-1}-2\Gamma^{-1})]N_m\tilde{L_m'} \\
&= -n^{-1}\sum_{t=1}^{n} \tilde{L_m} N_m[\Gamma^{-1}\otimes 2\Gamma^{-1}]N_m\tilde{L_m'} + o_p(1). \qquad \text{(A. 4)}
\end{aligned}
$$

The definitions of Ψ_m (basis matrix for diagonality), $\tilde{L_m}$ (basis matrix for strict lower triangularity) and N_m (commutation matrix) can be found on p.109, p.96 and p.48 of Magnus (1988) respectively.

B Appendix: Technical Proofs

Lemma B.1. Let the process ε_t be defined as in model (1.1)-(1.3) and $\varepsilon_t^* = \sum_{l=1}^{t-1} diag(\nu_l \odot \varepsilon_{t-l})(\iota - w(\varepsilon_t\varepsilon_t' V_t^{-1}))\odot\tilde{h}_t + V_t^{-1}\varepsilon_t$. If Assumptions (a)-(c) are satisfied. Then

$$
\frac{1}{\sqrt{n}}\sum_{t=1}^{[n\tau]}\begin{pmatrix}\varepsilon_t\\ \varepsilon_t^*\end{pmatrix} \longrightarrow_{\mathcal{L}} \begin{pmatrix} W_m(\tau)\\ \tilde{W}_m(\tau)\end{pmatrix} \text{ in } D^{2m},
$$

where $(W_m'(\tau), \tilde{W}_m'(\tau))'$ are defined as in Lemma 3.1, and $D^n = D\times D\cdots\times D(n$ factors), with D denoting the space of functions on $[0,1]$ defined and equipped with the Skorokhod topology. □

Proof. Denote $\varepsilon_t^{**} = \sum_{l=1}^{\infty} diag(\nu_l\odot\varepsilon_{t-l})(\iota - w(\varepsilon_t\varepsilon_t' V_t^{-1}))\odot\tilde{h}_t + V_t^{-1}\varepsilon_t$. Since $\nu_l = O(\rho^l)$, it is easy to show that

$$
\frac{1}{\sqrt{n}}\sum_{t=1}^{n}|\varepsilon_t^{**}-\varepsilon_t^*| = o_p(1). \qquad \text{(B. 1)}
$$

Let λ_1 and λ_2 be constant $m\times 1$ vectors and $\lambda = (\lambda_1', \lambda_2')'$, with $\lambda\lambda' \neq 0$. Denote $\xi_t = \lambda_1'\varepsilon_t + \lambda_2'\varepsilon_t^{**}$ and $S_n = \sum_{t=1}^{n}\xi_t$. It is obvious that ξ_t is a martingale difference sequence with respect to $\mathcal{F}_t$. By Assumptions (a)-(b), $\sigma^{*2} = n^{-1}ES_n^2 = \lambda_1' E(V_t)\lambda_1 + 2\lambda_1'\lambda_2 + \lambda_2'\Omega_1\lambda_2 < \infty$. Using the invariance principle for martingales and a similar method as for Lemma A.1 in LLW (2001), we can show that

$$
\frac{1}{\sqrt{n}}\sum_{t=1}^{[n\tau]}\begin{pmatrix}\varepsilon_t\\ \varepsilon_t^{**}\end{pmatrix} \longrightarrow_{\mathcal{L}} \begin{pmatrix} W_m(\tau)\\ \tilde{W}_m(\tau)\end{pmatrix} \text{ in } D^{2m}. \qquad \text{(B. 2)}
$$

By (B.1) and (B.2), we complete the proof. □

Lemma B.2. Under the assumptions of Lemma 3.1, it follows that

$$(a) \qquad \frac{1}{\sqrt{n}}\sum_{k=1}^{[n\tau]} u_{1k} = \frac{1}{\sqrt{n}}[I_d, 0](\sum_{k=1}^{\infty}\Psi_k)\sum_{k=1}^{[n\tau]} a_k + O_p(1/\sqrt{n})$$
$$\longrightarrow_{\mathcal{L}} \Psi_{11}\Omega_{a_1}^{1/2}B_d(\tau) \text{ in } D^d,$$
$$(b) \qquad n^{-2}\sum_{t=1}^{n} Z_{1t-1}Z_{1t-1}' \longrightarrow_{\mathcal{L}} \Psi_{11}\Omega_{a_1}^{1/2}\int_0^1 B_d(u)B_d(u)'du\Omega_{a_1}^{1/2}\Psi_{11}',$$
$$(c) \qquad n^{-1}\sum_{t=1}^{n} a_t Z_{1t-1}' \longrightarrow_{\mathcal{L}} \Omega_a^{1/2}[\int_0^1 B_d(u)dB_m(u)']'\Omega_{a_1}^{1/2}\Psi_{11}',$$

where u_{1t} is defined as in (2.4), B_d and Ψ_{11} are defined as in Lemma 3.1, and B_m is defined as in Theorem 4.1. $\square$

Proof. We first consider (a). By (2.3), we have

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{k=1}^{[n\tau]} u_{1k} &= \frac{1}{\sqrt{n}}[I_d, 0]\sum_{k=1}^{[n\tau]}(\sum_{i=1}^{\infty}\Psi_i)a_{k-i} \\ &= \frac{1}{\sqrt{n}}[I_d, 0]\sum_{k=1}^{[n\tau]}(\sum_{i=1}^{k}\Psi_i)a_{k-i} + \frac{1}{\sqrt{n}}r_t \end{aligned} \qquad \text{(B. 3)}$$

where $r_t = [I_d, 0]\sum_{k=1}^{[n\tau]}(\sum_{i=k+1}^{\infty}\Psi_i)a_{k-i}$. Note that

$$E||r_t|| \leq \sum_{k=1}^{[n\tau]}(\sum_{i=k+1}^{\infty}||\Psi_i||)E||a_{k-i}|| = O(\sum_{k=1}^{[n\tau]}\sum_{i=k+1}^{\infty}\rho^i) = O(1), \qquad \text{(B. 4)}$$

$O(\cdot)$ holds uniformly in τ. Furthermore, we have

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{k=1}^{[n\tau]}(\sum_{i=1}^{k}\Psi_i)a_{k-i} &= \frac{1}{\sqrt{n}}\sum_{k=1}^{[n\tau]}\Psi_k(\sum_{i=1}^{[nt]-k} a_i) \\ &= \frac{1}{\sqrt{n}}(\sum_{k=1}^{[n\tau]}\Psi_k)(\sum_{i=1}^{[nt]} a_i) + \frac{1}{\sqrt{n}}r_{1t} \\ &= \frac{1}{\sqrt{n}}(\sum_{k=1}^{\infty}\Psi_k)(\sum_{i=1}^{[nt]} a_i) + \frac{1}{\sqrt{n}}r_{2t} + \frac{1}{\sqrt{n}}r_{1t} \end{aligned} \qquad \text{(B. 5)}$$

where $r_{1t} = (\sum_{k=1}^{[n\tau]}\Psi_k)(\sum_{i=[nt]-k}^{[nt]} a_i) = O_p(1)$, $r_{2t} = (\sum_{k=[nt]+1}^{\infty}\Psi_k)(\sum_{i=1}^{[nt]} a_i) = O_p(1)$, and $O_p(\cdot)$ holds uniformly in τ. By (B.3)-(B.5) and Lemma B.1, we have

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{[n\tau]} u_{1k} \longrightarrow_{\mathcal{L}} [I_d, 0](\sum_{k=1}^{\infty}\Psi_k)\Omega_a B_m(\tau).$$

As in Ahn and Reinsel (1992), $[I_d, 0]\sum_{i=1}^{\infty}\Psi_i = [\Psi_{11}, 0] = \Psi_{11}[I_d, 0]$. Thus, (a) holds. By (a) of this lemma and the continuity mapping theorem, we know that (b) holds.

(c) comes from Theorem 2.2 in Kurtz and Protter (1991), Lemma B.1 and (a) of this lemma. This completes the proof. □

Lemma B.3. Denote $s_{jkt} = 2\sum_{i=1}^{t-1}\nu_{ij}\nu_{ik}\varepsilon_{j,t-i}\varepsilon_{k,t-i}/h_{j,t}h_{k,t} + c(h_{jt}h_{kt})^{-1/2}$ for any constant c. Then, under the assumptions in Lemma 3.1, it follows that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[n\tau]}[s_{jkt} - E(s_{jkt})] \longrightarrow_{\mathcal{L}} \sigma_{jk}w_{jk}(\tau),$$

where σ_{jk} is a non-negative constant, w_{jk} is a standard Brownian motion, and $j, k = 1, \cdots, m$. □

Proof. Note that each ε_{it}, a component of ε_t, is generated by a univariate GARCH(p,q) model. Using Theorem 2.1 in Ling and Li (1997) and Lemma 3.3 in Ling and Li (1998), the proof is essentially the same as that of Theorem 3.4 in Ling and Li (1998). The details are omitted. This completes the proof. □

Proof of Lemma 3.1. Note that $Z_{1t-i-1} = \sum_{k=1}^{t-i-1}u_{1k} = Z_{1t-1} - \sum_{k=t-i-1}^{t-1}u_{1k} = Z_{1t} + r_{it}$, where $r_{it} = -\sum_{k=t-i-1}^{t-1}u_{1k}$. It is not difficult to show that

$$\frac{1}{n^2}\sum_{t=1}^{n}\sum_{i=1}^{t-1}\{(Z_{1t-1}r_{it}' + r_{it}Z_{1t-1} + r_{it}r_{it}') \otimes [(\Gamma^{-1}\odot\Gamma + I_m)\odot\nu_j\nu_j'\odot\Pi_{jt}]\} = o_p(1).$$

Thus, we have

$$\begin{aligned}
&\frac{1}{n^2}\sum_{t=1}^{n}[\sum_{i=1}^{t-1}\{Z_{1t-i-1}Z_{1t-i-1}' \otimes [(\Gamma^{-1}\odot\Gamma + I_m)\odot\nu_j\nu_j'\odot\Pi_{jt}]\} + Z_{1t-1}Z_{1t-1}'\otimes V_t^{-1}] \\
&\quad = n^{-2}\sum_{t=1}^{n}\Big[Z_{1t-1}Z_{1t-1}' \\
&\qquad \otimes\{\sum_{i=1}^{t-1}[(\Gamma^{-1}\odot\Gamma + I_m)\odot\nu_j\nu_j'\odot\Pi_{jt}] + V_t^{-1}\}\Big] + o_p(1). \qquad \text{(B. 6)}
\end{aligned}$$

By Theorem 2.1 in Ling and Li (1998), (B.6), Lemma B.2(b) and Lemma B.3, the LHS of (B.6) is given by

$$\begin{aligned}
&\frac{1}{n^2}\sum_{t=1}^{n}(Z_{1t-1}Z_{1t-1}'\otimes\Omega_1) + \frac{1}{n^2}\sum_{t=1}^{n}[Z_{1t-1}Z_{1t-1}'\otimes(\sum_{i=1}^{q}V_{it} + V_t^{-1} - \Omega_1)] + o_p(1) \\
&\qquad \longrightarrow_{\mathcal{L}} \Psi_{11}\Omega_{a_1}^{1/2}\int_0^1 B_d(u)B_d(u)'\Omega_{a_1}^{1/2}\Psi_{11}'\otimes\Omega_1. \qquad \text{(B. 7)}
\end{aligned}$$

By (B.6)-(B.7), (a) holds. Similarly, we can show that

$$\frac{1}{n}\sum_{t=1}^{n}N_{1t} = \frac{1}{n}\sum_{t=1}^{n}(Z_{1t-1}\otimes I_m)\varepsilon_t^* + o_p(1). \qquad \text{(B. 8)}$$

By Theorem 2.2 in Kurtz and Protter (1991), Lemma B.1 and (B.8), (b) holds. (c) and (e) can be proved using the ergodic theorem, while (d) and (f) can be proved by the standard martingale central limiting theorem. This completes the proof. □

Proof of Theorem 3.1. It comes directly from Lemma 3.1. □

Proof of Theorem 4.1. From the proof of Lemma 13.2 by Johansen (1995), in our notation,

$$n(\breve{B}-B)P_1 = (A'V_0^{-1}A)^{-1}A'V_0^{-1}(n^{-1}\sum_{t=1}^{n}\varepsilon_t Z'_{1t-1})(n^{-2}\sum_{t=1}^{n} Z_{1t-1}Z'_{1t-1})^{-1} + o_p(1).$$

Note that $\varepsilon_t = Pa_t$ by the definition of a_t. By Lemma B.2(c)

$$\begin{aligned} \frac{1}{n}\sum_{t=1}^{n}\varepsilon_t Z'_{1t-1} &= P(\frac{1}{n}\sum_{t=1}^{n} a_t Z'_{1t-1}) \\ &\longrightarrow_{\mathcal{L}} P\Omega_a^{1/2}[\int_0^1 B_d(u)dB_m(u)']'\Omega_{a_1}^{1/2}\Psi'_{11}. \end{aligned} \quad \text{(B. 9)}$$

Similarly, by Lemma B.2(b),

$$\frac{1}{n^2}\sum_{t=1}^{n} Z_{1t-1}Z'_{1t-1} \longrightarrow_{\mathcal{L}} \Psi_{11}\Omega_{a_1}^{1/2}[\int_0^1 B_d(u)B_d(u)'du]\Omega_{a_1}^{1/2}\Psi'_{11}. \quad \text{(B. 10)}$$

Combining (B.9) and (B.10), Part (a) is proved. The proof of Part (b) is straightforward and thus it is omitted. This completes the proof. □

Proof of Corollary 4.1. The proof essentially follows the lines in Lemma 13.3 and pp.179-180 by Johansen (1995).

From Theorem 4.1, $(\breve{B}-B) = O_p(n^{-1})$. As a result, $(\breve{B}-[I_r, B_0]) = O_p(n^{-1})$. Next, define an rxm matrix $T \equiv [I_r, 0_{rxd}]$ and a normalized estimator $\breve{B}_T \equiv [I_r, \breve{B}_1^{-1}\breve{B}_2]$. Algebra shows:

$$\breve{B}_T = (\breve{B}T')^{-1}\breve{B}.$$

Similar to the arguments on pp.179-180 by Johansen (1995), we take the first-order Taylor expansion of $(\breve{B}T')^{-1}\breve{B}\mathcal{R}P_1$ around $[I_r, B_0]$:

$$\begin{aligned} (\breve{B}T')^{-1}\breve{B}\mathcal{R}P_1 &= ([I_r, B_0]T')^{-1}[I_r, B_0]\mathcal{R}P_1 + ([I_r, B_0]T')^{-1}(\breve{B}-[I_r, B_0])\mathcal{R}P_1 \\ &\quad -(B^*T')^{-1}(\breve{B}-[I_r, B_0])T'(B^*T')^{-1}[I_r, B_0]\mathcal{R}P_1 \\ &\quad -(B^*T')^{-1}(\breve{B}-[I_r, B_0])T'(B^*T')^{-1}(\breve{B}-[I_r, B_0])\mathcal{R}P_1, \end{aligned}$$

where B^* lies between $\breve{B}$ and $[I_r, B_0]$.

Note that $[I_r, B_0]T' = I_r$ and $[I_r, B_0]\mathcal{R}P_1 = 0$, and the third term vanishes. Moreover, it is not difficult to see that the last term is $O_p(n^{-2})$. Therefore,

$$\begin{aligned}(\breve{B}_T - [I_r, B_0])\mathcal{R}P_1 &= (\breve{B} - [I_r, B_0])\mathcal{R}P_1 + O_p(n^{-2}) \\ &= D_1^{-1}(\breve{B} - B)P_1 + O_p(n^{-2}). \qquad \text{(B. 11)}\end{aligned}$$

However, given the definitions of R_{11} and R_{21},

$$\begin{aligned}(\breve{B}_T - [I_r, B_0])\mathcal{R}P_1 &= R_{11} + \breve{B}_1^{-1}\breve{B}_2 R_{21} - R_{11} - B_0 R_{21} \\ &= (\breve{B}_1^{-1}\breve{B}_2 - B_0)R_{21}.\end{aligned}$$

Therefore, by (B.11) and Theorem 4.1,

$$\begin{aligned}n(\breve{B}_1^{-1}\breve{B}_2 - B_0) &= D_1^{-1}n(\breve{B} - B)P_1 R_{21}^{-1} + O_p(n^{-1}) \\ &\longrightarrow_{\mathcal{L}} D_1^{-1}(A'V_0^{-1}A)^{-1}A'V_0^{-1}PMR_{21}^{-1}.\end{aligned}$$

The proof of Part (b) is straightforward and thus it is omitted. This completes the proof. $\square$

Lemma B.4. Under the assumptions of Theorem 4.2, it follows that

(*a*) $(\hat{B}\bar{B}')^{-1}(\dot{B} - \hat{B}) = O_p(n^{-1/2})$,

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

(*c*) $(\dot{B}\bar{B}')^{-1}\hat{B}P_1 = (\hat{B}\bar{B}')^{-1}\hat{B}P_1 + O_p(n^{-3/2}) = BP_1 + O_p(n^{-1})$,

(*d*) $(\dot{B}\bar{B}')^{-1}\hat{B}P_2 = (\hat{B}\bar{B}')^{-1}\hat{B}P_2 + O_p(n^{-1/2}) = BP_2 + O_p(n^{-1/2})$. $\square$

Proof. (a). We first note that $vec[(\hat{B}\bar{B}')^{-1}(\dot{B} - \hat{B})] = (I_m \otimes (\hat{B}\bar{B}')^{-1})(\dot{\alpha}_1 - \hat{\alpha}_1)$.

Define $\bar{Q}_1^{***} = (Q \otimes (\hat{B}\bar{B}')') = (Q \otimes I_r)(I_m \otimes (\hat{B}\bar{B}')')$. In other words, $\bar{Q}_1^{***}$ is the $\bar{Q}_1^{**}$ (see (4.13)) with $\dot{B}$ replaced by $\hat{B}$. Thus, $\bar{Q}_1^{***\prime -1} = (P' \otimes I_r)(I_m \otimes (\hat{B}\bar{B}')^{-1})$.

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

Therefore, it suffices to show that $\bar{D}_1^{**}\bar{Q}_1^{***\prime-1}(\dot{\alpha}_1-\hat{\alpha}_1)=O_p(1)$.

Set $k=0$ in (4.11) and pre-multiply the entire equation by $\bar{D}_1^{**}\bar{Q}_1^{***\prime-1}$.

$$
\begin{aligned}
&\bar{D}_1^{**}\bar{Q}_1^{***\prime-1}(\dot{\alpha}_1-\hat{\alpha}_1)\\
&=-[\sum_{t=1}^{n}(R_{1t}|_{\hat{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})\bar{Q}_1^{***\prime}\bar{D}_1^{**-1}]^{-1}[\sum_{t=1}^{n}(\nabla_{\alpha_1}l_t|_{\hat{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})]\\
&=-[\sum_{t=1}^{n}\bar{D}_1^{**-1}\bar{Q}_1^{***}(R_{1t}|_{\hat{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})\bar{Q}_1^{***\prime}\bar{D}_1^{**-1}]^{-1}[\sum_{t=1}^{n}\bar{D}_1^{**-1}\bar{Q}_1^{***}(\nabla_{\alpha_1}l_t|_{\hat{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})]\\
&=-[\sum_{t=1}^{n}\bar{D}_1^{**-1}\mathcal{Q}_1(R_{1t}|_{\check{\alpha}_1,\check{\alpha}_2,\hat{\delta}})\mathcal{Q}_1'\bar{D}_1^{**-1}]^{-1}[\sum_{t=1}^{n}\bar{D}_1^{**-1}\mathcal{Q}_1(\nabla_{\alpha_1}l_t|_{\check{\alpha}_1,\check{\alpha}_2,\hat{\delta}})].
\end{aligned}
$$

Recall from Theorem 4.1 and Theorem 3.1 that $n(\check{\alpha}_1-\alpha_1)=O_p(1)$, $\sqrt{n}(\check{\alpha}_2-\alpha_2)=O_p(1)$, and $\sqrt{n}(\hat{\delta}-\delta)=O_p(1)$. Using (4.11),

$$
\sum_{t=1}^{n}\bar{D}_1^{**-1}\mathcal{Q}_1(R_{1t}|_{\check{\alpha}_1,\check{\alpha}_2,\hat{\delta}})\mathcal{Q}_1'\bar{D}_1^{**-1}=\sum_{t=1}^{n}\bar{D}_1^{**-1}\mathcal{Q}_1R_{1t}\mathcal{Q}_1'\bar{D}_1^{**-1}+o_p(1). \qquad \text{(B. 12)}
$$

Similarly, using (4.12),

$$
\begin{aligned}
&\sum_{t=1}^{n}\bar{D}_1^{**-1}\mathcal{Q}_1(\nabla_{\alpha_1}l_t|_{\check{\alpha}_1,\check{\alpha}_2,\hat{\delta}})\\
&=\sum_{t=1}^{n}\bar{D}_1^{**-1}\mathcal{Q}_1\nabla_{\alpha_1}l_t+\sum_{t=1}^{n}\bar{D}_1^{**-1}\mathcal{Q}_1R_{1t}(\check{\alpha}_1-\alpha_1)+o_p(1)\\
&=\sum_{t=1}^{n}\bar{D}_1^{**-1}\mathcal{Q}_1\nabla_{\alpha_1}l_t+[\sum_{t=1}^{n}\bar{D}_1^{**-1}\mathcal{Q}_1R_{1t}\mathcal{Q}_1'\bar{D}_1^{**-1}]n^{-1}\bar{D}_1^{**}(P'\otimes I_r)[n(\check{\alpha}_1-\alpha_1)]+o_p(1)\\
&=O_p(1), \qquad \text{(B. 13)}
\end{aligned}
$$

where the last equation holds by Lemma 3.1. By (B.12) and (B.13), (a) is proved.

We now consider (b). By the 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.

For (c) and (d), note that:

$$
(\dot{B}\bar{B}')^{-1}\hat{B}=[(\hat{B}\bar{B}')^{-1}\dot{B}\bar{B}']^{-1}(\hat{B}\bar{B}')^{-1}\hat{B}=[(\hat{B}\bar{B}')^{-1}\dot{B}\bar{B}']^{-1}\check{B}. \qquad \text{(B. 14)}
$$

Using the formulas, $dF^{-1}=-F^{-1}(dF)F^{-1}$ (see, for instance, Theorem 3, Chapter 8 in Magnus and Neudecker, 1988) for any $r\times r$ squared matrix F, and applying a one-term 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.14) equals:

$$\begin{aligned} &[(\hat{B}\bar{B}')^{-1}\hat{B}\bar{B}']^{-1}(\hat{B}\bar{B}')^{-1}\hat{B} - [B^*\bar{B}']^{-1}[(\hat{B}\bar{B}')^{-1}\dot{B} - \check{B}]\bar{B}'[B^*\bar{B}']^{-1}\check{B} \\ &= (\hat{B}\bar{B}')^{-1}\hat{B} - [B^*\bar{B}']^{-1}[(\hat{B}\bar{B}')^{-1}\dot{B} - \check{B}]\bar{B}'[B^*\bar{B}']^{-1}\check{B}. \qquad \text{(B. 15)} \end{aligned}$$

By (a) of this lemma, $(\hat{B}\bar{B}')^{-1}\dot{B} - \check{B} = O_p(n^{-1/2})$. On the other hand, $[B^*\bar{B}']^{-1}$, $\bar{B}$ and $\check{B}$ are all $O_P(1)$. By (B.15), (d) holds.

By Theorem 4.1, $\check{B}P_1 = O_p(n^{-1})$ since $BP_1 = 0$. Post-multiply (B.15) by P_1,

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

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

Lemma B.5. Consider $\bar{Q}_1^{***} = (Q \otimes (\hat{B}\bar{B}')') = (Q \otimes I_r)(I_m \otimes (\hat{B}\bar{B}')')$, which is defined in the proof of Lemma B.4. Define $\grave{\alpha}_2 \equiv vec[\hat{A}(\dot{B}\bar{B}'), \hat{\Phi}_1^*, \ldots, \hat{\Phi}_{s-1}^*]$. That is, $\grave{\alpha}_2$ is $\check{\alpha}_2$ with $\hat{B}$ replaced by $\dot{B}$. It follows that

$$(a) \quad \sum_{t=1}^{n} \bar{D}_1^{**-1}\bar{Q}_1^{***}(R_{1t}|_{\hat{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})\bar{Q}_1^{**\prime}\bar{D}_1^{**-1} = \sum_{t=1}^{n} \bar{D}_1^{**-1}\mathcal{Q}_1(R_{1t}|_{\check{\alpha}_1,\check{\alpha}_2,\hat{\delta}})\mathcal{Q}_1'\bar{D}_1^{**-1} + o_p(1),$$

$$(b) \quad \sum_{t=1}^{n} \bar{D}_1^{**-1}\bar{Q}_1^{***}(\nabla_{\alpha_1} l_t|_{\hat{\alpha}_1,\hat{\alpha}_2,\hat{\delta}}) = \sum_{t=1}^{n} \bar{D}_1^{**-1}\mathcal{Q}_1(\nabla_{\alpha_1} l_t|_{\check{\alpha}_1,\check{\alpha}_2,\hat{\delta}}),$$

$$(c) \quad n^{-1}\sum_{t=1}^{n} \bar{Q}_2^{**}(R_{2t}|_{\dot{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})\bar{Q}_2^{**\prime} = n^{-1}\sum_{t=1}^{n}(R_{2t}|_{\ddot{\alpha}_1,\grave{\alpha}_2,\hat{\delta}}),$$

$$(d) \quad n^{-1/2}\sum_{t=1}^{n} \bar{Q}_2^{**}(\nabla_{\alpha_2} l_t|_{\dot{\alpha}_1,\hat{\alpha}_2,\hat{\delta}}) = n^{-1/2}\sum_{t=1}^{n}(\nabla_{\alpha_2} l_t|_{\ddot{\alpha}_1,\grave{\alpha}_2,\hat{\delta}}),$$

where all other variables are defined as in Sub-section 4.2. □

Proof. For (a), first note that $\sqrt{n}(\grave{\alpha}_2 - \check{\alpha}_2) = O_p(1)$ by Lemma B.4(b). Using the expression for R_{1t} around (4.6)-(4.7), it is not difficult to see that the LHS in (a) equals:

$$\sum_{t=1}^{n} \bar{D}_1^{**-1}\bar{Q}_1^{***}(R_{1t}|_{\hat{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})\bar{Q}_1^{***\prime}\bar{D}_1^{**-1} + o_p(1) = \sum_{t=1}^{n} \bar{D}_1^{**-1}\mathcal{Q}_1(R_{1t}|_{\check{\alpha}_1,\check{\alpha}_2,\hat{\delta}})\mathcal{Q}_1'\bar{D}_1^{**-1} + o_p(1).$$

(b), (c) and (d) are straightforward. This completes the proof. $\square$

Proof of Theorem 4.2. It suffices to consider the first iteration. To prove Theorem 4.2(a), recall that $\dot{\alpha}_1 = \dot{\alpha}_1^{(1)}$ and $\hat{\alpha}_1 = \dot{\alpha}_1^{(0)}$. Set $k = 0$ in (4.11),

$$\dot{\alpha}_1 = \hat{\alpha}_1 - (\sum_{t=1}^{n} R_{1t}|_{\hat{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})^{-1}(\sum_{t=1}^{n} \nabla_{\alpha_1} l_t|_{\hat{\alpha}_1,\hat{\alpha}_2,\hat{\delta}}). \qquad \text{(B. 16)}$$

Pre-multiply (B.16) by $\bar{D}_1^{**}\bar{Q}_1^{**\prime-1}$. By Lemmas B.5(a) and B.5(b),

$$\begin{aligned}
\bar{D}_1^{**}\bar{Q}_1^{**\prime-1}\dot{\alpha}_1 &= \bar{D}_1^{**}\bar{Q}_1^{**\prime-1}\hat{\alpha}_1 - [\sum_{t=1}^{n}(R_{1t}|_{\hat{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})\bar{Q}_1^{**\prime}\bar{D}_1^{**-1}]^{-1}[\sum_{t=1}^{n}(\nabla_{\alpha_1} l_t|_{\hat{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})] \\
&= \bar{D}_1^{**}\bar{Q}_1^{**\prime-1}\hat{\alpha}_1 - [\sum_{t=1}^{n} \bar{D}_1^{**-1}\bar{Q}_1^{***}(R_{1t}|_{\hat{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})\bar{Q}_1^{**\prime}\bar{D}_1^{**-1}]^{-1} \\
&\quad \cdot[\sum_{t=1}^{n} \bar{D}_1^{**-1}\bar{Q}_1^{***}(\nabla_{\alpha_1} l_t|_{\hat{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})] \\
&= \bar{D}_1^{**}\bar{Q}_1^{**\prime-1}\hat{\alpha}_1 - [\sum_{t=1}^{n} \bar{D}_1^{**-1}\mathcal{Q}_1(R_{1t}|_{\breve{\alpha}_1,\breve{\alpha}_2,\hat{\delta}})\mathcal{Q}_1'\bar{D}_1^{**-1}]^{-1} \\
&\quad \cdot[\sum_{t=1}^{n} \bar{D}_1^{**-1}\mathcal{Q}_1(\nabla_{\alpha_1} l_t|_{\breve{\alpha}_1,\breve{\alpha}_2,\hat{\delta}})] + o_p(1).
\end{aligned} \qquad \text{(B. 17)}$$

By (B.12) and (B.13), it follows that

$$\sum_{t=1}^{n} \bar{D}_1^{**-1}\mathcal{Q}_1(R_{1t}|_{\breve{\alpha}_1,\breve{\alpha}_2,\hat{\delta}})\mathcal{Q}_1'\bar{D}_1^{**-1} = \sum_{t=1}^{n} \bar{D}_1^{**-1}\mathcal{Q}_1 R_{1t}\mathcal{Q}_1'\bar{D}_1^{**-1} + o_p(1), \qquad \text{(B. 18)}$$

$$\begin{aligned}
\sum_{t=1}^{n} \bar{D}_1^{**-1}\mathcal{Q}_1(\nabla_{\alpha_1} l_t|_{\breve{\alpha}_1,\breve{\alpha}_2,\hat{\delta}}) &= \sum_{t=1}^{n} \bar{D}_1^{**-1}\mathcal{Q}_1\nabla_{\alpha_1} l_t + [\sum_{t=1}^{n} \bar{D}_1^{**-1}\mathcal{Q}_1 R_{1t}\mathcal{Q}_1'\bar{D}_1^{**-1}] \\
&\quad \cdot\bar{D}_1^{**}(P' \otimes I_r)(\breve{\alpha}_1 - \alpha_1) + o_p(1).
\end{aligned} \qquad \text{(B. 19)}$$

By (B.18)-(B.19), we can express (B.17) as:

$$\begin{aligned}
\bar{D}_1^{**}\bar{Q}_1^{**\prime-1}\dot{\alpha}_1 &= \bar{D}_1^{**}\bar{Q}_1^{**\prime-1}\hat{\alpha}_1 - [\sum_{t=1}^{n} \bar{D}_1^{**-1}\mathcal{Q}_1 R_{1t}\mathcal{Q}_1'\bar{D}_1^{**-1}]^{-1}[\sum_{t=1}^{n} \bar{D}_1^{**-1}\mathcal{Q}_1\nabla_{\alpha_1} l_t] \\
&\quad -\bar{D}_1^{**}(P' \otimes I_r)(\breve{\alpha}_1 - \alpha_1) + o_p(1).
\end{aligned} \qquad \text{(B. 20)}$$

However, note that

$$\begin{aligned}
\bar{Q}_1^{**\prime-1}\dot{\alpha}_1 &= [P' \otimes (\dot{B}\bar{B}')^{-1}]vec[\dot{B}] \\
&= vec[(\dot{B}\bar{B}')^{-1}\dot{B}P] = vec[(\dot{B}\bar{B}')^{-1}\dot{B}P_1, (\dot{B}\bar{B}')^{-1}\dot{B}P_2].
\end{aligned}$$

To prove Theorem 4.2(a), we only need to consider the first rd elements of $\bar{D}_1^{**}\bar{Q}_1^{**\prime -1}\dot{\alpha}_1$.

By (B.20) and Lemma B.4(c), the first rd elements of $\bar{D}_1^{**}\bar{Q}_1^{**\prime -1}\hat{\alpha}_1$ equal:

$$vec[n(\dot{B}\bar{B}')^{-1}\hat{B}P_1] = vec[n(\hat{B}\bar{B}')^{-1}\hat{B}P_1] + o_p(1). \quad \text{(B. 21)}$$

As $\sum_{t=1}^n \bar{D}_1^{**-1}\mathcal{Q}_1 R_{1t}\mathcal{Q}_1'\bar{D}_1^{**-1}$ converges in distribution to a block-diagonal matrix, the first rd elements of $[\sum_{t=1}^n \bar{D}_1^{**-1}\mathcal{Q}_1 R_{1t}\mathcal{Q}_1'\bar{D}_1^{**-1}]^{-1}[\sum_{t=1}^n \bar{D}_1^{**-1}\mathcal{Q}_1\nabla_{\alpha_1}l_t]$ equal:

$$[n^{-2}\sum_{t=1}^n (Q_1'\otimes I_r)R_{1t}(Q_1\otimes I_r)]^{-1}[n^{-1}\sum_{t=1}^n (Q_1'\otimes I_r)\nabla_{\alpha_1}l_t] + o_p(1). \quad \text{(B. 22)}$$

Furthermore, since $BP_1 = 0$, the first rd elements of $\bar{D}_1^{**}(P'\otimes I_r)(\check{\alpha}_1 - \alpha_1)$ equal:

$$\begin{aligned} n(P_1'\otimes I_r)vec[\check{B} - B] &= vec[n(\check{B}P_1 - BP_1)] \\ &= vec[n\check{B}P_1] = vec[n(\hat{B}\bar{B}')^{-1}\hat{B}P_1]. \quad \text{(B. 23)} \end{aligned}$$

By (B.21)-(B.23), the first rd elements in (B.20) can be expressed as:

$$\begin{aligned} & vec[n(\dot{B}\bar{B}')^{-1}\dot{B}P_1] \\ &= vec[n(\hat{B}\bar{B}')^{-1}\hat{B}P_1] \\ &\quad -[n^{-2}\sum_{t=1}^n (Q_1'\otimes I_r)R_{1t}(Q_1\otimes I_r)]^{-1}[n^{-1}\sum_{t=1}^n (Q_1'\otimes I_r)\nabla_{\alpha_1}l_t] \\ &\quad -vec[n(\hat{B}\bar{B}')^{-1}\hat{B}P_1] + o_p(1) \\ &= -[n^{-2}\sum_{t=1}^n (Q_1'\otimes I_r)R_{1t}(Q_1\otimes I_r)]^{-1}[n^{-1}\sum_{t=1}^n (Q_1'\otimes I_r)\nabla_{\alpha_1}l_t] + o_p(1). \text{(B. 24)} \end{aligned}$$

Thus, Theorem 4.2(a) is proved.

For Theorem 4.2(b), recall that $\dot{\alpha}_2 = \dot{\alpha}_2^{(1)}$ and $\hat{\alpha}_2 = \dot{\alpha}_2^{(0)}$. Set $k = 0$ in (4.12),

$$\dot{\alpha}_2 = \hat{\alpha}_2 - (\sum_{t=1}^n R_{2t}|_{\dot{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})^{-1}(\sum_{t=1}^n \nabla_{\alpha_2}l_t|_{\dot{\alpha}_1,\hat{\alpha}_2,\hat{\delta}}). \quad \text{(B. 25)}$$

Pre-multiply (B.25) by $\sqrt{n}\bar{Q}_2^{**\prime -1}$. By Lemmas B.5(c) and B.5(d), it follows that

$$\begin{aligned} \sqrt{n}\bar{Q}_2^{**\prime -1}\dot{\alpha}_2 &= \sqrt{n}\bar{Q}_2^{**\prime -1}\hat{\alpha}_2 \\ &\quad -[n^{-1}\sum_{t=1}^n \bar{Q}_2^{**}(R_{2t}|_{\dot{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})\bar{Q}_2^{**\prime}]^{-1}[n^{-1/2}\sum_{t=1}^n \bar{Q}_2^{**}(\nabla_{\alpha_2}l_t|_{\dot{\alpha}_1,\hat{\alpha}_2,\hat{\delta}})] \\ &= \sqrt{n}\bar{Q}_2^{**\prime -1}\hat{\alpha}_2 - [n^{-1}\sum_{t=1}^n R_{2t}|_{\ddot{\alpha}_1,\dot{\alpha}_2,\hat{\delta}}]^{-1}[n^{-1/2}\sum_{t=1}^n \nabla_{\alpha_2}l_t|_{\ddot{\alpha}_1,\dot{\alpha}_2,\hat{\delta}}], \text{(B. 26)} \end{aligned}$$

where $\grave{\alpha}_2$ is defined in Lemma B.5. By (a) of this theorem. $n(\ddot{\alpha}_1 - \alpha_1) = O_p(1)$. By Theorem 4.1(b) and Lemma B.4(b), $\sqrt{n}(\grave{\alpha}_2 - \alpha_2) = O_p(1)$. By Theorem 4.1(c), $\sqrt{n}(\hat{\delta} - \delta) = O_p(1)$. Thus, by (4.11), we have

$$n^{-1}\sum_{t=1}^{n}(R_{2t}|_{\ddot{\alpha}_1,\grave{\alpha}_2,\hat{\delta}}) = n^{-1}\sum_{t=1}^{n} R_{2t} + o_p(1). \qquad \text{(B. 27)}$$

Similarly, using (4.12),

$$\begin{aligned} & n^{-1/2}\sum_{t=1}^{n}(\nabla_{\alpha_2} l_t|_{\ddot{\alpha}_1,\grave{\alpha}_2,\hat{\delta}}) \\ &= n^{-1/2}\sum_{t=1}^{n}\nabla_{\alpha_2} l_t + [n^{-1}\sum_{t=1}^{n}\nabla^2_{\alpha_2\alpha_2'} l_t]\sqrt{n}(\grave{\alpha}_2 - \alpha_2) + o_p(1). \qquad \text{(B. 28)} \end{aligned}$$

Note that $\bar{Q}_2^{**\prime-1}\dot{\alpha}_2 = \ddot{\alpha}_2$ and $\bar{Q}_2^{**\prime-1}\hat{\alpha}_2 = \grave{\alpha}_2$. Combining (B.27) and (B.28), (B.26) can be expressed as:

$$\sqrt{n}\ddot{\alpha}_2 = \sqrt{n}\grave{\alpha}_2 - [n^{-1}\sum_{t=1}^{n} R_{2t}]^{-1}[n^{-1/2}\sum_{t=1}^{n}\nabla_{\alpha_2} l_t] - \sqrt{n}(\grave{\alpha}_2 - \alpha_2) + o_p(1).$$

Alternatively,

$$\sqrt{n}(\ddot{\alpha}_2 - \alpha_2) = -[n^{-1}\sum_{t=1}^{n} R_{2t}]^{-1}[n^{-1/2}\sum_{t=1}^{n}\nabla_{\alpha_2} l_t] + o_p(1). \qquad \text{(B. 29)}$$

Thus Theorem 4.2(b) is also proved. □

Proof of Corollary 4.2. Exactly the same as that of Corollary 4.1 with $\hat{B}$, $\check{B}$, and $(A'V_0^{-1}A)^{-1}A'V_0^{-1}PM$ replaced by $\dot{B}$, $\ddot{B}$, and $(A'\Omega_1 A)^{-1}A'\tilde{M}$, respectively. □

Proof of Lemma 5.1. First consider the first term in (5.2). By Theorem 3.1(a) and Lemma 3.1(a),

$$\begin{aligned} & (\hat{\beta}_1 - \beta_1)'(\sum_{t=1}^{n} L_{1t})(\hat{\beta}_1 - \beta_1) \\ &= vec[n(\hat{C} - C)P_1]'[n^{-2}\sum_{t=1}^{n} L_{1t}]vec[n(\hat{C} - C)P_1] \\ &\longrightarrow_{\mathcal{L}} vec[\Omega_1^{-1}\tilde{M}]'[Z \otimes \Omega_1]vec[\Omega_1^{-1}\tilde{M}] \\ &= vec[\Omega_1^{-1}\tilde{M}]'vec[\Omega_1\Omega_1^{-1}\tilde{M}Z] = tr[\tilde{M}'\Omega_1^{-1}\tilde{M}Z], \qquad \text{(B. 30)} \end{aligned}$$

where $Z \equiv \Psi_{11}\Omega_{a_1}^{1/2}\int_0^1 B_d(u)B_d(u)'\Omega_{a_1}^{1/2}\Psi_{11}'$ (see Section 3) and $\tilde{M}$ is as defined in Theorem 3.1. Next we consider the second term in (5.2). First note that:

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

Recall that $BP_1 = 0$. As argued in Theorem 4.1, $(\ddot{B} - B)P_1 = O_p(n^{-1})$ and $(\ddot{A} - A) = O_p(n^{-1/2})$ under H_0, and hence,

$$\begin{aligned} n(\ddot{A}\ddot{B} - AB)P_1 &= n(\ddot{A} - A)BP_1 + nA(\ddot{B} - B)P_1 + (\ddot{A} - A)n(\ddot{B} - B)P_1 \\ &= nA(\ddot{B} - B)P_1 + O_p(n^{-1/2}). \end{aligned}$$

Therefore, by Theorem 4.1(a) and Lemma 3.1(a),

$$\begin{aligned} &(\ddot{\beta}_1 - \beta_1)'(\sum_{t=1}^{n} L_{1t})(\ddot{\beta}_1 - \beta_1) \\ &= vec[n(\ddot{A}\ddot{B} - AB)P_1]'[n^{-2}\sum_{t=1}^{n} L_{1t}]vec[n(\ddot{A}\ddot{B} - AB)P_1] \\ &= vec[nA(\ddot{B} - B)P_1]'[n^{-2}\sum_{t=1}^{n} L_{1t}]vec[nA(\ddot{B} - B)P_1] + o_p(1) \\ &\longrightarrow_{\mathcal{L}} vec[D\tilde{M}]'[Z \otimes \Omega_1]vec[D\tilde{M}] \\ &= vec[D\tilde{M}]'vec[\Omega_1 D\tilde{M}Z] = tr[\tilde{M}'D\Omega_1 D\tilde{M}Z], \end{aligned} \qquad \text{(B. 31)}$$

where $D \equiv A(A'\Omega_1 A)^{-1}A'$. Combining (B.30) and (B.31),

$$\begin{aligned} LR_G &\longrightarrow_{\mathcal{L}} tr[\tilde{M}'(\Omega_1^{-1} - D\Omega_1 D)\tilde{M}Z] \\ &= tr[(\Omega_1^{-1} - A(A'\Omega_1 A)^{-1}A')\tilde{M}Z\tilde{M}']. \end{aligned}$$

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:

$$\Omega_1^{-1}(\Omega_1 - \Omega_1 A(A'\Omega_1 A)^{-1}A'\Omega_1)\Omega_1^{-1} = \Omega_1^{-1}Q_1(Q_1'\Omega_1^{-1}Q_1)^{-1}Q_1'\Omega_1^{-1}.$$

Therefore, we can rewrite the asymptotic distribution as:

$$tr[(\int_0^1 B_d(u)\tilde{V}_d(u)')'(\int_0^1 B_d(u)B_d(u)'du)^{-1}(\int_0^1 B_d(u)\tilde{V}_d(u)')],$$

where $\tilde{V}_d(u) \equiv (Q_1'\Omega_1^{-1}Q_1)^{-1/2}Q_1'\Omega_1^{-1}\tilde{W}_m(u)$. Note that

$$E[B_d(u)\tilde{V}_d(u)'] = u\Omega_{a1}^{-1/2}(Q_1'\Omega_1^{-1}Q_1)^{1/2} = u\Upsilon'.$$

Thus, we can rewrite $\tilde{V}_d(u)$ as a linear combination of two independent $d-$dimensional standard Brownian motions:

$$\Upsilon B_d(u) + [(Q_1'\Omega_1^{-1}Q_1)^{-1/2}Q_1'\Omega_1^{-1}\tilde{\Omega}_1\Omega_1^{-1}Q_1(Q_1'\Omega_1^{-1}Q_1)^{-1/2} - \Upsilon\Upsilon']^{1/2}V_d(u). \text{(B. 32)}$$

The proof is complete. $\square$

Proof of Theorem 5.1. When $\tilde{\Omega}_1 = \Omega_1$, (B.32) in the proof of Lemma 5.1 can be expressed as:

$$\Upsilon B_d(u) + [I_d - \Upsilon\Upsilon']^{1/2}V_d(u).$$

The limiting distribution can then be expressed as:

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

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 Brownian motion independent of $V_d(u)$.

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

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

Since $(I_d - \Upsilon\Upsilon')$ is a real symmetric matrix, we can decompose it as:

$$(I_d - \Upsilon\Upsilon') = \Theta\Lambda_d\Theta',$$

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 Brownian motions $B_d(u)$ and $V_d(u)$ respectively. Cancelling the orthogonal Θ, we have:

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

This completes the proof. □

Proof of Corollary 5.1. Similar to that of Theorem 5.1 and thus it is omitted.□

Proof of Theorem 5.1'. Similar to that of Theorem 5.1 and thus it is omitted.□

C Appendix: Simulated Critical Values

TABLE C.10

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 $T = 2,000$.
λ_1 is the eigenvalue of Λ_1 in (5.3) or Λ_1^* in (5.5).

TABLE C.11

Quantiles of the Limiting Distribution (6.4)

$d = 1$, with a Constant Term

	$\alpha-$th simulated quantiles							
λ_1	.500	.750	.800	.850	.900	.950	.975	.990
0.0	2.457	4.342	4.904	5.607	6.588	8.167	9.653	11.690
0.1	2.203	4.082	4.649	5.362	6.327	7.932	9.449	11.469
0.2	1.952	3.819	4.383	5.092	6.051	7.656	9.225	11.291
0.3	1.704	3.533	4.094	4.799	5.767	7.373	8.949	11.027
0.4	1.470	3.236	3.784	4.477	5.457	7.049	8.615	10.727
0.5	1.241	2.921	3.451	4.132	5.113	6.679	8.253	10.293
0.6	1.036	2.603	3.117	3.763	4.715	6.251	7.810	9.817
0.7	0.851	2.267	2.749	3.375	4.272	5.763	7.296	9.316
0.8	0.693	1.938	2.382	2.965	3.794	5.220	6.690	8.655
0.9	0.557	1.619	2.007	2.526	3.270	4.576	5.937	7.789
1.0	0.457	1.324	1.643	2.070	2.705	3.827	5.044	6.678

See the footnote in Table C.10.

TABLE C.20

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

$d = 2$, no Constant Term

		$\alpha-$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 C.20 (Continued)

		α−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 $T = 2,000$.

$\lambda_1 \leq \lambda_2$ are the eigenvalues of Λ_2 in (5.3) or Λ_2^* in (5.5).

TABLE C.21

Quantiles of the Limiting Distribution (6.4)

$d = 2$, with a Constant Term

		α−th simulated quantiles							
λ_1	λ_2	.500	.750	.800	.850	.900	.950	.975	.990
0.0	0.0	9.425	12.515	13.401	14.445	15.842	18.064	20.120	22.745
0.0	0.1	9.130	12.222	13.102	14.148	15.578	17.783	19.865	22.477
0.0	0.2	8.823	11.936	12.789	13.851	15.270	17.508	19.618	22.192
0.0	0.3	8.520	11.599	12.485	13.541	14.955	17.201	19.338	21.901
0.0	0.4	8.216	11.282	12.156	13.213	14.630	16.877	19.012	21.541
0.0	0.5	7.908	10.942	11.814	12.865	14.277	16.516	18.663	21.236
0.0	0.6	7.604	10.602	11.464	12.506	13.913	16.137	18.305	20.842
0.0	0.7	7.301	10.256	11.103	12.143	13.537	15.739	17.898	20.390
0.0	0.8	6.995	9.914	10.734	11.759	13.132	15.332	17.436	19.987
0.0	0.9	6.693	9.553	10.359	11.350	12.712	14.908	16.920	19.498
0.0	1.0	6.382	9.182	9.964	10.951	12.287	14.387	16.372	18.904
0.1	0.1	8.814	11.934	12.804	13.849	15.264	17.530	19.582	22.311
0.1	0.2	8.519	11.619	12.480	13.553	14.977	17.212	19.338	22.013
0.1	0.3	8.208	11.304	12.172	13.232	14.650	16.897	19.051	21.670
0.1	0.4	7.889	10.976	11.838	12.909	14.326	16.564	18.724	21.346
0.1	0.5	7.582	10.636	11.500	12.568	13.984	16.206	18.361	20.978
0.1	0.6	7.273	10.299	11.157	12.198	13.616	15.833	17.947	20.528
0.1	0.7	6.963	9.960	10.792	11.832	13.219	15.439	17.508	20.137
0.1	0.8	6.665	9.600	10.424	11.450	12.808	15.004	17.070	19.640
0.1	0.9	6.365	9.236	10.037	11.054	12.393	14.561	16.565	19.157
0.1	1.0	6.065	8.863	9.665	10.640	11.952	14.014	16.001	18.530
0.2	0.2	8.207	11.319	12.182	13.233	14.661	16.917	19.005	21.739
0.2	0.3	7.900	11.008	11.849	12.933	14.341	16.583	18.703	21.389
0.2	0.4	7.574	10.673	11.515	12.591	13.985	16.246	18.381	21.052
0.2	0.5	7.262	10.337	11.182	12.244	13.630	15.894	18.007	20.695
0.2	0.6	6.946	9.984	10.832	11.886	13.259	15.495	17.593	20.250
0.2	0.7	6.636	9.623	10.477	11.502	12.861	15.094	17.145	19.815
0.2	0.8	6.337	9.269	10.114	11.114	12.457	14.650	16.720	19.281
0.2	0.9	6.039	8.901	9.714	10.707	12.034	14.178	16.204	18.757
0.2	1.0	5.738	8.522	9.323	10.289	11.585	13.665	15.612	18.155
0.3	0.3	7.589	10.678	11.539	12.600	14.000	16.253	18.369	21.093
0.3	0.4	7.265	10.342	11.202	12.246	13.658	15.906	18.007	20.787
0.3	0.5	6.944	9.996	10.859	11.903	13.310	15.539	17.642	20.360

TABLE C.21 (Continued)

		α−th simulated quantiles							
λ_1	λ_2	.500	.750	.800	.850	.900	.950	.975	.990
0.3	0.6	6.625	9.653	10.503	11.533	12.919	15.147	17.224	19.921
0.3	0.7	6.319	9.297	10.132	11.149	12.515	14.736	16.763	19.441
0.3	0.8	6.016	8.930	9.762	10.762	12.087	14.277	16.341	18.904
0.3	0.9	5.710	8.551	9.355	10.337	11.666	13.806	15.836	18.329
0.3	1.0	5.415	8.174	8.955	9.916	11.200	13.272	15.228	17.725
0.4	0.4	6.953	10.001	10.857	11.915	13.306	15.556	17.672	20.385
0.4	0.5	6.630	9.644	10.499	11.547	12.961	15.153	17.233	20.004
0.4	0.6	6.319	9.300	10.143	11.160	12.572	14.747	16.856	19.548
0.4	0.7	6.016	8.945	9.765	10.796	12.152	14.345	16.402	19.072
0.4	0.8	5.705	8.579	9.387	10.387	11.709	13.888	15.889	18.549
0.4	0.9	5.394	8.191	8.993	9.955	11.264	13.372	15.399	17.907
0.4	1.0	5.092	7.817	8.571	9.525	10.783	12.869	14.789	17.292
0.5	0.5	6.323	9.303	10.136	11.186	12.595	14.741	16.868	19.545
0.5	0.6	6.013	8.935	9.770	10.796	12.191	14.352	16.438	19.122
0.5	0.7	5.700	8.575	9.384	10.397	11.774	13.919	15.986	18.639
0.5	0.8	5.386	8.191	9.002	9.993	11.309	13.470	15.485	18.047
0.5	0.9	5.080	7.820	8.591	9.557	10.850	12.958	14.958	17.429
0.5	1.0	4.790	7.421	8.172	9.100	10.340	12.409	14.340	16.794
0.6	0.6	5.697	8.569	9.390	10.407	11.780	13.924	16.018	18.673
0.6	0.7	5.384	8.190	9.000	10.004	11.356	13.484	15.555	18.199
0.6	0.8	5.079	7.806	8.604	9.580	10.897	13.018	15.024	17.607
0.6	0.9	4.776	7.419	8.190	9.139	10.412	12.479	14.464	16.943
0.6	1.0	4.486	7.022	7.756	8.681	9.892	11.922	13.828	16.291
0.7	0.7	5.071	7.809	8.595	9.584	10.921	13.031	15.056	17.617
0.7	0.8	4.770	7.423	8.195	9.155	10.456	12.553	14.497	17.093
0.7	0.9	4.471	7.029	7.758	8.689	9.956	11.975	13.937	16.376
0.7	1.0	4.182	6.613	7.328	8.221	9.415	11.389	13.272	15.693
0.8	0.8	4.470	7.029	7.772	8.711	9.963	12.045	13.915	16.478
0.8	0.9	4.174	6.616	7.341	8.240	9.466	11.429	13.363	15.794
0.8	1.0	3.904	6.207	6.888	7.750	8.914	10.816	12.668	15.007
0.9	0.9	3.892	6.219	6.905	7.759	8.942	10.861	12.674	15.178
0.9	1.0	3.620	5.794	6.438	7.260	8.386	10.211	11.956	14.235
1.0	1.0	3.350	5.377	5.983	6.737	7.800	9.527	11.218	13.381

See the footnote in Table C.20.

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.

Bollerslev, T. (1990), Modelling 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.

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

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.

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

Johansen, S. (1988), Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control*, 12, 231-254.

Johansen, S. (1995), *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.

Kurtz, T.G. and Protter, P. (1991), Weak Limit Theorems to Stochastic Integrals and Stochastic Differential Equations, *Annals of Probability*, 19, 1035-1070.

Lee, T.H. and Tse, Y. (1996), Cointegration Tests with Conditional Heteroskedasticity. *Journal of Econometrics*, 73, 401-410.

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. (1999), On the Probabilistic Properties of a Double Threshold ARMA Conditional Heteroskedastic Model, *Journal of Applied Probability*, 36, 688-705.

Ling, S. and Li, W.K. (1997), On Fractionally Integrated Autoregressive Moving-Average Time Series Models with Conditional Heteroskedasticity, *Journal of the American Statistical Association*, 92, 1184-1194.

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. and Li, W.K. (2002), Asymptotic Inference for Unit Root Processes with GARCH(1,1) Errors, Forthcoming in *Econometric Theory*.

Ling, S., Li, W.K. and McAleer, M. (2002), Estimation and Testing for Unit Root Processes with GARCH(1,1) Errors: Theory and Monte Carlo Study, Forthcoming in *Econometric Reviews*.

Ling, S. and McAleer, M. (2002a), Asymptotic Theory for a New Vector ARMA-GARCH Model, Forthcoming in *Econometric Theory*.

Ling, S. and McAleer, M. (2002b), On Adaptive Estimation in Nonstationary ARMA Models with GARCH Errors, Forthcoming in *Annals of Statistics*.

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.

Nicholls, D.F. and Quinn, B.G. (1982) *Random Coefficient Autoregressive Model: An Introduction* Lecture Notes in Statistics, 11. New York: Springer.

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

Rahbek, A. and Mosconi, R. (1999), Cointegration Rank Inference with Stationary Regressions in VAR Models, *Econometrics Journal*, 2, 76-91.

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.

Seo, B. (1999), Distribution Theory for Unit Root Tests with Conditional Heteroskedasticity, *Journal of Econometrics*, 91, 113-144.

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

Tsay, R.S. (1987), Conditional Heteroscedastic Time Series Models, *Journal of the American Statistical Association*, 82, 590-604.

Tse, Y.K. (2000), A Test for Constant Correlations in a Multivariate GARCH Model, *Journal of Econometrics*, 98, 107-127.