ASYMPTOTICS FOR LS, GLS, AND FEASIBLE GLS STATISTICS IN AN AR(1) MODEL WITH CONDITIONAL HETEROSKEDATICITY

By

Donald W.K. Andrews and Patrik Guggenberger

June 2008

COWLES FOUNDATION DISCUSSION PAPER NO. 1665

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

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

Asymptotics for LS, GLS, and Feasible GLS Statistics in an AR(1) Model with Conditional Heteroskedaticity

Donald W. K. Andrews*

Cowles Foundation for Research in Economics

Yale University

Patrik Guggenberger

Department of Economics

UCLA

April 2008

* Andrews gratefully acknowledges the research support of the National Science Foundation via grant number SES-0417911. Guggenberger gratefully acknowledges research support the National Science Foundation via grant number SES-0748922.

Abstract

This paper considers a first-order autoregressive model with conditionally heteroskedastic innovations. The asymptotic distributions of least squares (LS), infeasible generalized least squares (GLS), and feasible GLS estimators and t statistics are determined. The GLS procedures allow for misspecification of the form of the conditional heteroskedasticity and, hence, are referred to as quasi-GLS procedures. The asymptotic results are established for drifting sequences of the autoregressive parameter and the distribution of the time series of innovations. In particular, we consider the full range of cases in which the autoregressive parameter ρ_n satisfies (i) $n(1-\rho_n) \to \infty$ and (ii) $n(1-\rho_n) \to h_1 < \infty$ as $n \to \infty$, where n is the sample size. Results of this type are needed to establish the uniform asymptotic properties of the LS and quasi-GLS statistics.

Keywords: Asymptotic distribution, autoregression, conditional heteroskedasticity, generalized least squares, least squares.

JEL Classification Numbers: C22.

1 Introduction

This paper establishes the asymptotic distributions of quasi-GLS statistics in an AR(1) model with intercept and conditional heteroskedasticity. The statistics considered include infeasible and feasible quasi-GLS estimators, heteroskedasticity-consistent (HC) standard error estimators, and the t statistics formed from these estimators. The paper considers the cases where the autoregressive parameter ρ_n satisfies (i) $n(1-\rho_n) \to \infty$ and (ii) $n(1-\rho_n) \to h_1 < \infty$ as $n \to \infty$. In case (i), the quasi-GLS t statistic is shown to have a standard normal asymptotic distribution. In case (ii), its asymptotic distribution is shown to be that of a convex linear combination of a random variable with a "demeaned near unit-root distribution" and an independent standard normal random variable. The weights on the two random variables depend on the correlation between the innovation, say U_i, and the innovation rescaled by the quasi-conditional variance, say U_i/ϕ_i^2. Here ϕ_i^2 is the (possibly misspecified) conditional variance used by the GLS estimator. In the case of LS, we have $\phi_i^2 = 1$, the correlation between U_i and U_i/ϕ_i^2 is one, and the asymptotic distribution is a demeaned near unit-root distribution (based on an Ornstein-Uhlenbeck process).

An AR(1) model with conditional heteroskedasticity and $\rho = 1$, which falls within case (ii) above, has been considered by Seo (1999) and Guo and Phillips (2001). The results given here make use of ideas in these two papers. For an AR(1) model without conditional heteroskedasticity, case (i) is studied by Park (2002), Giraitis and Phillips (2006), and Phillips and Magdalinos (2007). Case (ii) is the "near integrated" case that has been studied in AR models without conditional heteroskedasticity by Bobkowski (1983), Cavanagh (1985), Chan and Wei (1987), Phillips (1987), Elliott (1999), Elliott and Stock (2001), and Müller and Elliott (2003). The latter three papers consider the situation that also is considered here in which the initial condition yields a stationary process.

As noted above, in the present paper, we consider a heteroskedasticity-consistent (HC) standard error estimator. Such an estimator is needed in order for the quasi-GLS t statistic to have a standard normal asymptotic distribution in case (i) when the form of the conditional heteroskedasticity is misspecified.

The paper provides high-level conditions under which infeasible and feasible quasi-GLS estimators are asymptotically equivalent. The high-level conditions are verified for cases in which the GLS estimator employs a parametric model, with parameter π, for the form of the conditional heteroskedasticity. For technical reasons, we take the estimator of π to be a discretized estimator and we require the parametric form of the conditional heteroskedasticity to be such that the conditional variance depends upon a finite number of lagged squared innovations. Neither of these conditions is particularly restrictive because (a) the grid size for the discretized estimator can be defined such that there is little difference between the discretized and non-discretized versions of the estimator of π, (b) the parametric model for the conditional heteroskedasticity may be misspecified, and (c) any parametric model with stationary conditional heteroskedasticity, such as a GARCH(1,1) model, can be approximated arbitrarily well by a model with a large finite number of lags.

The results of this paper are used in Andrews and Guggenberger (2005) to show that symmetric two-sided subsampling confidence intervals (based on the quasi-GLS t statistic described above) have correct asymptotic size in an AR(1) model with conditional heteroskedasticity. (Here "asymptotic size" is defined to be the limit as the sample size n goes to infinity of the exact, i.e., finite-sample, size.) This result requires uniformity in the asymptotics and, hence, relies on asymptotic results in which the autoregressive parameter and the innovation distribution may depend on n. In addition, Andrews and Guggenberger (2005) shows that upper and lower one-sided and symmetric and equal-tailed two-sided hybrid-subsampling confidence intervals have correct asymptotic size. No other confidence intervals in the literature, including those in Stock (1991), Andrews (1993), Andrews and Chen (1994), Nankervis and Savin (1996), Hansen (1999), Chen and Deo (2007), and Mikusheva (2007), have correct asymptotic size in an AR(1) model with conditional heteroskedasticity.

The remainder of the paper is organized as follows. Section 2 introduces the model and statistics considered. Section 3 gives the assumptions, normalization constants, and asymptotic results. Section 4 provides proofs of the results.

2 Model, Estimators, and t Statistic

We use the unobserved components representation of the AR(1) model. The observed time series $\{Y_i : i = 0, ..., n\}$ is based on a latent no-intercept AR(1) time series $\{Y_i^* : i = 0, ..., n\}$:

$$\begin{aligned} Y_i &= \alpha + Y_i^*, \\ Y_i^* &= \rho Y_{i-1}^* + U_i, \text{ for } i = 1, ..., n, \end{aligned} \tag{2.1}$$

where $\rho \in [-1 + \varepsilon, 1]$ for some $0 < \varepsilon < 2$, $\{U_i : i = ..., 0, 1, ...\}$ are stationary and ergodic with conditional mean 0 given a σ-field $\mathcal{G}_{i-1}$ defined below, conditional variance $\sigma_i^2 = E(U_i^2|\mathcal{G}_{i-1})$, and unconditional variance $\sigma_U^2 \in (0, \infty)$. The distribution of Y_0^* is the distribution that yields strict stationarity for $\{Y_i^* : i \leq n\}$ when $\rho < 1$, i.e., $Y_0^* = \sum_{j=0}^{\infty} \rho^j U_{-j}$, and is arbitrary when $\rho = 1$.

The model can be rewritten as

$$Y_i = \widetilde{\alpha} + \rho Y_{i-1} + U_i, \text{ where } \widetilde{\alpha} = \alpha(1 - \rho), \tag{2.2}$$

for $i = 1, ..., n$.[1]

We consider a feasible quasi-GLS (FQGLS) estimator of ρ and a t statistic based on it. The FQGLS estimator depends on estimators $\{\widehat{\phi}_{n,i}^2 : i \leq n\}$ of the conditional

[1] By writing the model as in (2.1), the case $\rho = 1$ and $\widetilde{\alpha} \neq 0$ is automatically ruled out. Doing so is desirable because when $\rho = 1$ and $\widetilde{\alpha} \neq 0$, Y_i is dominated by a deterministic trend and the LS estimator of ρ converges at rate $n^{3/2}$.

variances $\{\sigma_i^2 : i \leq n\}$. The estimators $\{\widehat{\phi}_{n,i}^2 : i \leq n\}$ may be from a parametric specification of the conditional heteroskedasticity, e.g., a GARCH(1, 1) model, or from a nonparametric estimator, e.g., one based on q lags of the observations. We do not assume that the conditional heteroskedasticity estimator is consistent. For example, we allow for incorrect specification of the parametric model in the former case and conditional heteroskedasticity that depends on more than q lags in the latter case. The estimated conditional variances $\{\widehat{\phi}_{n,i}^2 : i \leq n\}$ are defined such that they approximate a stationary $\mathcal{G}_{i-1}$-adapted sequence $\{\phi_i^2 : i \leq n\}$ in the sense that certain normalized sums have the same asymptotic distribution whether $\widehat{\phi}_{n,i}^2$ or ϕ_i^2 appears in the sum. This is a typical property of feasible and infeasible GLS estimators.

As an example, the results allow for the case where (i) $\{\widehat{\phi}_{n,i}^2 : i \leq n\}$ are from a GARCH(1,1) parametric model estimated using LS residuals with GARCH and LS parameter estimators $\widetilde{\pi}_n$ and $(\widetilde{\alpha}_n, \widetilde{\rho}_n)$, respectively, (ii) $(\widetilde{\alpha}_n, \widetilde{\rho}_n)$ have probability limit given by the true values $(\widetilde{\alpha}_0, \rho_0)$, see (2.2), (iii) $\widetilde{\pi}_n$ has a probability limit given by the "pseudo-true" value π_0, (iv) $\widehat{\phi}_{n,i}^2 = \phi_{i,1}^2(\widetilde{\alpha}_n, \widetilde{\rho}_n, \widetilde{\pi}_n)$, where $\phi_{i,1}^2(\widetilde{\alpha}, \rho, \pi)$ is the i-th GARCH conditional variance based on a start-up at time 1 and parameters $(\widetilde{\alpha}, \rho, \pi)$, and (v) $\phi_{i,-\infty}^2(\widetilde{\alpha}, \rho, \pi)$ is the GARCH conditional variance based on a start-up at time $-\infty$ and parameters $(\widetilde{\alpha}, \rho, \pi)$. In this case, $\phi_i^2 = \phi_{i,-\infty}^2(\widetilde{\alpha}_0, \rho_0, \pi_0)$. Thus, ϕ_i^2 is just $\widehat{\phi}_{n,i}^2$ with the estimation error and start-up truncation eliminated.

Under the null hypothesis that $\rho = \rho_n$, the studentized t statistic is

$$T_n^*(\rho_n) = \frac{n^{1/2}(\widehat{\rho}_n - \rho_n)}{\widehat{\sigma}_n}, \tag{2.3}$$

where $\widehat{\rho}_n$ is the LS estimator from the regression of $Y_i/\widehat{\phi}_{n,i}$ on $Y_{i-1}/\widehat{\phi}_{n,i}$ and $1/\widehat{\phi}_{n,i}$, and $\widehat{\sigma}_n^2$ is the $(1, 1)$ element of the standard heteroskedasticity-robust variance estimator for the LS estimator in the preceding regression.

To define $T_n^*(\rho_n)$ more explicitly, let Y, U, X_1, and X_2 be n-vectors with ith elements given by $Y_i/\widehat{\phi}_{n,i}$, $U_i/\widehat{\phi}_{n,i}$, $Y_{i-1}/\widehat{\phi}_{n,i}$, and $1/\widehat{\phi}_{n,i}$, respectively. Let Δ be the diagonal $n \times n$ matrix with ith diagonal element given by the ith element of the residual vector $M_X Y$, where $X = [X_1 : X_2]$ and $M_X = I_n - X(X'X)^{-1}X'$. That is, $\Delta = Diag(M_X Y)$. Then, by definition,

$$\begin{aligned} \widehat{\rho}_n &= (X_1' M_{X_2} X_1)^{-1} X_1' M_{X_2} Y, \text{ and} \\ \widehat{\sigma}_n^2 &= \left(n^{-1} X_1' M_{X_2} X_1\right)^{-1} \left(n^{-1} X_1' M_{X_2} \Delta^2 M_{X_2} X_1\right) \left(n^{-1} X_1' M_{X_2} X_1\right)^{-1}. \end{aligned} \tag{2.4}$$

By assumption, $\{(U_i, \phi_i^2) : i \geq 1\}$ are stationary and strong mixing. We define $\mathcal{G}_i$ to be some non-decreasing sequence of σ-fields for $i \geq 1$ for which $(U_j, \phi_{j+1}^2) \in \mathcal{G}_i$ for all $j \leq i$.

3 Asymptotic Results

3.1 Assumptions

We let F denote the distribution of $\{(U_i, \phi_i^2) : i = ..., 0, 1, ...\}$. Our asymptotic results below are established under drifting sequences $\{(\rho_n, F_n) : n \geq 1\}$ of autoregressive parameters ρ_n and distributions F_n. In particular, we provide results for the cases $n(1-\rho_n) \to \infty$ and $n(1-\rho_n) \to h_1 < \infty$. When F_n depends on n, $\{(U_i, \phi_i^2) : i \leq n\}$ for $n \geq 1$ form a triangular array of random variables and $(U_i, \phi_i^2) = (U_{n,i}, \phi_{n,i}^2)$. We now specify assumptions on $(U_{n,i}, \phi_{n,i}^2)$. The assumptions place restrictions on the drifting sequence of distributions $\{F_n : n \geq 1\}$ that are considered.

The statistics $\widehat{\rho}_n$, $\widehat{\sigma}_n$, and $T_n^*(\rho_n)$ are invariant to the value of α. Hence, without loss of generality, from now on we take $\alpha = 0$ and $Y_{n,i} = Y_{n,i}^*$.

Assumption INNOV. (i) For each $n \geq 1$, $\{(U_{n,i}, \phi_{n,i}^2) : i = ..., 0, 1, ...\}$ are stationary and strong mixing with $E(U_{n,i}|\mathcal{G}_{n,i-1}) = 0$ a.s., $E(U_{n,i}^2|\mathcal{G}_{n,i-1}) = \sigma_{n,i}^2$ a.s. where $\mathcal{G}_{n,i}$ is some non-decreasing sequence of σ-fields for $i = ..., 1, 2, ...$ for $n \geq 1$ for which $(U_{n,j}, \phi_{n,j+1}^2) \in \mathcal{G}_{n,i}$ for all $j \leq i$, (ii) the strong-mixing numbers $\{\alpha_n(m) : m \geq 1\}$ satisfy $\alpha(m) = \sup_{n \geq 1} \alpha_n(m) = O(m^{-3\zeta/(\zeta-3)})$ as $m \to \infty$ for some $\zeta > 3$, (iv) $\sup_{n,i,s,t,u,v,A} E_{F_n}|\prod_{a \in A} a|^{\zeta} < \infty$, where $0 \leq i, s, t, u, v < \infty$, $n \geq 1$, and A is any non-empty subset of $\{U_{n,i-s}, U_{n,i-t}, U_{n,i+1}^2, U_{n,-u}, U_{n,-v}, U_{n,1}^2\}$, (v) $\phi_{n,i}^2 \geq \delta > 0$ *a.s.*, (vi) $\lambda_{\min} E(X^1 X^{1\prime} U_{n,1}^2/\phi_{n,1}^2) \geq \delta > 0$, where $X^1 = (Y_{n,0}^*/\phi_{n,1}, \phi_{n,1}^{-1})'$, and (vii) the following limits exist and are positive: $h_{2,1} = \lim_{n\to\infty} EU_{n,i}^2$, $h_{2,2} = \lim_{n\to\infty} E(U_{n,i}^2/\phi_{n,i}^4)$, $h_{2,3} = \lim_{n\to\infty} E(U_{n,i}^2/\phi_{n,i}^2)$, $h_{2,4} = \lim_{n\to\infty} E\phi_{n,i}^{-1}$, $h_{2,5} = \lim_{n\to\infty} E\phi_{n,i}^{-2}$, and $h_{2,6} = \lim_{n\to\infty} E\phi_{n,i}^{-4}$.

Given that $\phi_{n,i}$ is bounded away from zero by Assumption INNOV(v), Assumption INNOV(iv) implies that $\sup_{n,i,s,t,u,v,A^*} E_{F_n}|\prod_{a\in A^*} a|^{\zeta} < \infty$, where $0 \leq i, s, t, u, v < \infty$, $n \geq 1$, and A^* is a non-empty subset of $\{U_{n,i-s}, U_{n,i-t}, U_{n,i+1}^2/\phi_{n,i+1}^4, U_{n,-u}, U_{n,-v}, U_{n,1}^2/\phi_{n,1}^4\}$ or a subset of $\{U_{n,i-s}, U_{n,i-t}, \phi_{n,i+1}^{-k}, U_{n,-u}, U_{n,-v}, \phi_{n,1}^{-k}\}$ for $k = 2, 3, 4$. The uniform bound on these expectations is needed in the proofs of Lemmas 7 and 8 below.

If $\rho_n = 1$, the initial condition $Y_{n,0}^*$ is arbitrary. If $\rho_n < 1$, then the initial condition satisfies the following assumption:

Assumption STAT. $Y_{n,0}^* = \sum_{j=0}^{\infty} \rho_n^j U_{n,-j}$.

We determine the asymptotic distributions $\widehat{\rho}_n$, $\widehat{\sigma}_n^2$, and $T_n^*(\rho_n)$ under sequences $\{(\rho_n, F_n) : n \geq 1\}$ such that (a) Assumption INNOV holds and if $\rho_n < 1$ Assumption STAT also holds, and

$$\text{(b) } n(1-\rho_n) \to h_1 \text{ for (i) } h_1 = \infty \text{ and (ii) } 0 \leq h_1 < \infty. \tag{3.1}$$

The asymptotic distributions of $\widehat{\rho}_n$ and $\widehat{\sigma}_n^2$ are shown to depend on the parameters h_1, $h_{2,1}$, and $h_{2,2}$ (where $h_{2,1}$ and $h_{2,2}$ are defined in Assumption INNOV(vi)) and the parameter $h_{2,7}$, which is defined by

$$h_{2,7} = \frac{h_{2,3}}{(h_{2,1}h_{2,2})^{1/2}} = \lim_{n\to\infty} Corr_{F_n}(U_{n,i}, U_{n,i}/\phi_{n,i}^2)). \tag{3.2}$$

The asymptotic distribution of $T_n^*(\rho_n)$ is shown to depend only on h_1 and $h_{2,7}$.

Define

$$\begin{aligned} h_2 &= (h_{2,1}, ..., h_{2,7})' \text{ and} \\ h &= (h_1, h_2')' \in H = R_{+,\infty} \times H_2, \end{aligned} \tag{3.3}$$

where $R_+ = \{x \in R : x \geq 0\}$, $R_{+,\infty} = R_+ \cup \{\infty\}$, and $H_2 \subset (0,\infty)^6 \times (0,1]$.

For notational simplicity, we index the asymptotic distributions of $\widehat{\rho}_n$, $\widehat{\sigma}_n^2$, and $T_n^*(\rho_n)$ by h below (even though they only depend on a subvector of h).

3.2 Normalization Constants

The normalization constants a_n and d_n used to obtain the asymptotic distributions of $\widehat{\rho}_n$ and $\widehat{\sigma}_n^2$ depend on (ρ_n, F_n) and are denoted $a_n(\rho_n, F_n)$ and $d_n(\rho_n, F_n)$. They are defined as follows. Let $\{\rho_n : n \geq 1\}$ be a sequence for which $n(1-\rho_n) \to \infty$ or $n(1-\rho_n) \to h_1 < \infty$. Define the 2-vectors

$$\begin{aligned} X^1 &= (Y_{n,0}^*/\phi_{n,1}, \phi_{n,1}^{-1})' \text{ and} \\ Z &= (1, -E_{F_n}(Y_{n,0}^*/\phi_{n,1}^2)/E_{F_n}(\phi_{n,1}^{-2}))'. \end{aligned} \tag{3.4}$$

Define

$$\begin{aligned} a_n &= a_n(\rho_n, F_n) = n^{1/2} d_n(\rho_n, F_n) \text{ and} \\ d_n &= d_n(\rho_n, F_n) = \begin{cases} \frac{E_{F_n}(Y_{n,0}^{*2}/\phi_{n,1}^2) - (E_{F_n}(Y_{n,0}^*/\phi_{n,1}^2))^2/E_{F_n}(\phi_{n,1}^{-2})}{(Z' E_{F_n}(X^1 X^{1\prime} U_{n,1}^2/\phi_{n,1}^2) Z)^{1/2}} & \text{if } n(1-\rho_n) \to \infty \\ n^{1/2} & \text{if } n(1-\rho_n) \to h_1 < \infty. \end{cases} \end{aligned} \tag{3.5}$$

Note that the normalization constant for the t statistic $T_n^*(\rho_n)$ is $a_n(\rho_n, F_n)/d_n(\rho_n, F_n)$ $= n^{1/2}$.

In certain cases, the normalization constants simplify. In the case where $n(1-\rho_n) \to \infty$ and $\rho_n \to 1$, the constants a_n and d_n in (3.5) simplify to

$$a_n = n^{1/2} \frac{E_{F_n}(Y_{n,0}^{*2}/\phi_{n,1}^2)}{(E_{F_n}(Y_{n,0}^{*2} U_{n,1}^2/\phi_{n,1}^4))^{1/2}} \text{ and } d_n = \frac{E_{F_n}(Y_{n,0}^{*2}/\phi_{n,1}^2)}{(E_{F_n}(Y_{n,0}^{*2} U_{n,1}^2/\phi_{n,1}^4))^{1/2}} \tag{3.6}$$

up to lower order terms. This holds because by Lemma 6 below

$$\begin{aligned} & Z' E_{F_n}(X^1 X^{1\prime} U_{n,1}^2/\phi_{n,1}^2) Z \\ &= E_{F_n}(Y_{n,0}^{*2} U_{n,1}^2/\phi_{n,1}^4) - 2E_{F_n}(Y_{n,0}^* U_{n,1}^2/\phi_{n,1}^4) E_{F_n}(Y_{n,0}^*/\phi_{n,1}^2)/E_{F_n}(\phi_{n,1}^{-2}) \\ &\quad + (E_{F_n}(Y_{n,0}^*/\phi_{n,1}^2))^2 E_{F_n}(U_{n,1}^2/\phi_{n,1}^4)/(E_{F_n}(\phi_{n,1}^{-2}))^2 \\ &= E_{F_n}(Y_{n,0}^{*2} U_{n,1}^2/\phi_{n,1}^4)(1 + O(1-\rho_n)) \end{aligned} \tag{3.7}$$

and

$$E_{F_n}(Y_{n,0}^{*2}/\phi_{n,1}^2) - (E_{F_n}(Y_{n,0}^{*}/\phi_{n,1}^2))^2/E_{F_n}(\phi_{n,1}^{-2}) = E_{F_n}(Y_{n,0}^{*2}/\phi_{n,1}^2)(1+O(1-\rho_n)). \quad (3.8)$$

If, in addition, $\{U_{n,i} : i = ..., 0, 1, ...\}$ are i.i.d. with mean 0, variance $\sigma_{U,n}^2 \in (0, \infty)$, and distribution F_n and $\phi_{n,i}^2 = 1$, then the constants a_n and d_n simplify to

$$a_n = n^{1/2}(1-\rho_n^2)^{-1/2} \text{ and } d_n = (1-\rho_n^2)^{-1/2}. \quad (3.9)$$

This follows because in the present case $\phi_{n,i}^2 = 1$, $E_{F_n}Y_{n,0}^{*2} = \sum_{j=0}^{\infty} \rho_n^{2j} E_{F_n} U_{n,-j}^2 = (1-\rho_n^2)^{-1}\sigma_{U,n}^2$, and $E_{F_n}(Y_{n,0}^{*2}U_{n,1}^2/\phi_{n,1}^2) = (1-\rho_n^2)^{-1}\sigma_{n,U}^4$. The expression for a_n in (3.9) is as in Giraitis and Phillips (2006).

3.3 Results for LS and Infeasible QGLS

In this section, we provide results for the (infeasible) QGLS estimator based on $\{\phi_{n,i}^2 : i \le n\}$ rather than $\{\widehat{\phi}_{n,i}^2 : i \le n\}$. Conditions under which feasible and infeasible QGLS estimators are asymptotically equivalent are given in Section 3.4 below. The LS estimator is covered by the results of this section by taking $\phi_{n,i}^2 = 1$ for all n, i.

Let $W(\cdot)$ and $W_2(\cdot)$ be independent standard Brownian motions on $[0, 1]$. Let Z_1 be a standard normal random variable that is independent of $W(\cdot)$ and $W_2(\cdot)$. By definition,

$$\begin{aligned}
I_h(r) &= \int_0^r \exp(-(r-s)h_1)dW(s), \\
I_h^*(r) &= I_h(r) + \frac{1}{\sqrt{2h_1}}\exp(-h_1 r)Z_1 \text{ for } h_1 > 0 \text{ and } I_h^*(r) = W(r) \text{ for } h_1 = 0, \\
I_{D,h}^*(r) &= I_h^*(r) - \int_0^1 I_h^*(s)ds, \text{ and} \\
Z_2 &= \left(\int_0^1 I_{D,h}^*(r)^2 dr\right)^{-1/2} \int_0^1 I_{D,h}^*(r)dW_2(r). \quad (3.10)
\end{aligned}$$

As defined, $I_h(r)$ is an Ornstein-Uhlenbeck process. Note that the conditional distribution of Z_2 given $W(\cdot)$ and Z_1 is standard normal. Hence, its unconditional distribution is standard normal and it is independent of $W(\cdot)$ and Z_1.

The asymptotic distribution of the infeasible QGLS estimator and t statistic are given in the following Theorem.

Theorem 1 *Suppose* (i) *Assumption* INNOV *holds,* (ii) *Assumption* STAT *holds when* $\rho_n < 1$, (iii) $\rho_n \in [-1+\varepsilon, 1]$ *for some* $0 < \varepsilon < 2$, *and* (iv) $\rho_n = 1 - h_{n,1}/n$ *and*

$h_{n,1} \to h_1 \in [0, \infty]$. *Then, the infeasible QGLS estimator* $\widehat{\rho}_n$ *and* t *statistic* $T_n^*(\rho_n)$ *(defined in* (2.3) *and* (2.4) *with* $\phi_{n,i}$ *in place of* $\widehat{\phi}_{n,i}$*) satisfy*

$$a_n(\widehat{\rho}_n - \rho_n) \to_d V_h, \;\; d_n\widehat{\sigma}_n \to_d Q_h, \;\; \text{and } T_n^*(\rho_n) = \frac{n^{1/2}(\widehat{\rho}_n - \rho_n)}{\widehat{\sigma}_n} \to_d J_h,$$

where a_n, d_n, V_h, Q_h, *and* J_h *are defined as follows.*
(a) *For* $h_1 \in [0, \infty)$, $a_n = n$, $d_n = n^{1/2}$, V_h *is the distribution of*

$$h_{2,7}\frac{\int_0^1 I_{D,h}^*(r)dW(r)}{h_{2,2}^{1/2}h_{2,1}^{1/2}\int_0^1 I_{D,h}^*(r)^2dr} + (1-h_{2,7}^2)^{1/2}\frac{\int_0^1 I_{D,h}^*(r)dW_2(r)}{h_{2,2}^{1/2}h_{2,1}^{1/2}\int_0^1 I_{D,h}^*(r)^2dr}, \qquad (3.11)$$

Q_h *is the distribution of*

$$h_{2,2}^{-1/2}h_{2,1}^{-1/2}\left[\int_0^1 I_{D,h}^*(r)^2dr\right]^{-1/2}, \qquad (3.12)$$

and J_h *is the distribution of*

$$h_{2,7}\frac{\int_0^1 I_{D,h}^*(r)dW(r)}{\left(\int_0^1 I_{D,h}^*(r)^2dr\right)^{1/2}} + (1-h_{2,7}^2)^{1/2}Z_2. \qquad (3.13)$$

(b) *For* $h_1 = \infty$, a_n *and* d_n *are defined as in* (3.5), V_h *is a* $N(0,1)$ *distribution,* Q_h *is the distribution of the constant one, and* J_h *is a* $N(0,1)$ *distribution.*

Comments. 1. Theorem 1 shows that the asymptotic distribution of the QGLS t statistic is a standard normal distribution when $n(1-\rho_n) \to \infty$ and a mixture of a standard normal distribution and a "demeaned near unit-root distribution" when $n(1-\rho_n) \to h_1 < \infty$. In the latter case, the mixture depends on $h_{2,7}$, which is the asymptotic correlation between the innovation $U_{n,i}$ and the rescaled innovation $U_{n,i}/\phi_{n,i}^2$. When the LS estimator is considered (which corresponds to $\phi_{n,i}^2 = 1$), we have $h_{2,7} = 1$ and the asymptotic distribution is a "demeaned near unit-root distribution."

2. The asymptotic results of Theorem 1 apply to a first-order AR model. They should extend without essential change to a p-th order autoregressive model in which ρ equals the "sum of the AR coefficients." Of course, the proofs will be more complex. We do not provide them here.

3. Theorem 1 is used in the AR(1) example of Andrews and Guggenberger (2005) to verify Assumptions BB(i) and (iii) for the (infeasible) QGLS estimator (with Q_h playing the role of W_h in Assumption BB). In turn, the results of Andrews and Guggenberger (2005) show that whether or not conditional heteroskedasticity is present: (i) the symmetric two-sided subsampling confidence interval for ρ has correct asymptotic size (defined to be the limit as $n \to \infty$ of exact size) and (ii) upper and lower one-sided and symmetric and equal-tailed two-sided hybrid-subsampling confidence intervals for ρ have correct asymptotic size. These results hold even if the form of the conditional heteroskedasticity is misspecified.

3.4 Asymptotic Equivalence of Feasible and Infeasible QGLS

Here we provide sufficient conditions for the feasible and infeasible QGLS statistics to be asymptotically equivalent. In particular, we give conditions under which Theorem 1 holds when $\widehat{\rho}_n$ is defined using the feasible conditional heteroskedasticity estimators $\{\widehat{\phi}_{n,i} : i \leq n\}$.

We assume that the conditional heteroskedasticity estimators (CHE) $\{\widehat{\phi}^2_{n,i} : i \leq n\}$ satisfy the following assumption.

Assumption CHE. (i) For some $\varepsilon > 0$, $\widehat{\phi}^2_{n,i} \geq \varepsilon$ a.s. for all $i \leq n$, $n \geq 1$. (ii) For random variables $\{(U_{n,i}, \phi^2_{n,i}) : i = ..., 0, 1, ...\}$ for $n \geq 1$ that satisfy Assumption INNOV and for $Y_{n,i} = \alpha + Y^*_{n,i}$, $Y^*_{n,i} = \rho_n Y^*_{n,i-1} + U_{n,i}$, with $\alpha = 0$, that satisfies Assumption STAT when $\rho_n < 1$ and $n(1-\rho_n) \to h_1 \in [0, \infty]$, we have (a) when $h_1 \in [0, \infty)$, $n^{-1/2} \sum_{i=1}^n (n^{-1/2} Y^*_{n,i-1})^j U_{n,i} (\widehat{\phi}^{-2}_{n,i} - \phi^{-2}_{n,i}) = o_p(1)$ for $j = 0, 1$, (b) when $h_1 \in [0, \infty)$, $n^{-1} \sum_{i=1}^n |U_{n,i}|^d |\widehat{\phi}^{-j}_{n,i} - \phi^{-j}_{n,i}| = o_p(1)$ for $(d, j) = (0, 1), (1, 2)$, and $(2, 2)$, (c) when $h_1 = \infty$, $n^{-1/2} \sum_{i=1}^n ((1-\rho_n)^{1/2} Y^*_{n,i-1})^j U_{n,i} (\widehat{\phi}^{-2}_{n,i} - \phi^{-2}_{n,i}) = o_p(1)$ for $j = 0, 1$, and (d) when $h_1 = \infty$, $n^{-1} \sum_{i=1}^n |U_{n,i}|^k \, |\widehat{\phi}^{-j}_{n,i} - \phi^{-j}_{n,i}|^d = o_p(1)$ for $(d, j, k) = (1, 2, 0), (2, 2, 0)$, and $(2, 4, k)$ for $k = 0, 2, 4$.

Assumption CHE(i) is not restrictive. For example, if $\widehat{\phi}_{n,i}$ is obtained by specifying a parametric model for the conditional heteroskedasticity, then Assumption CHE(i) holds provided the specified parametric model (which is user chosen) consists of an intercept that is bounded away from zero plus a non-negative random component (as in (3.14) below). Most parametric models in the literature have this form and it is always possible to use one that does. Typically, Assumptions CHE(ii)(a) and (c) are more difficult to verify than Assumptions CHE(ii)(b) and (d) because they have the scale factor $n^{-1/2}$ rather than n^{-1}.

Theorem 2 *Suppose* (i) *Assumptions* CHE *and* INNOV *hold,* (ii) *Assumption* STAT *holds when* $\rho_n < 1$, (iii) $\rho_n \in [-1+\varepsilon, 1]$ *for some* $0 < \varepsilon < 2$, *and* (iv) $\rho_n = 1 - h_{n,1}/n$ *and* $h_{n,1} \to h_1 \in [0, \infty]$. *Then, the feasible QGLS estimator* $\widehat{\rho}_n$ *and t statistic* $T^*_n(\rho_n)$ (*defined in* (2.3) *and* (2.4) *using* $\widehat{\phi}_{n,i}$) *satisfy*

$$a_n(\widehat{\rho}_n - \rho_n) \to_d V_h, \; d_n \widehat{\sigma}_n \to_d Q_h, \; \text{and } T^*_n(\rho_n) = \frac{n^{1/2}(\widehat{\rho}_n - \rho_n)}{\widehat{\sigma}_n} \to_d J_h,$$

where a_n, d_n, V_h, Q_h, *and* J_h *are defined as in Theorem* 1 (*that is, with* a_n *and* d_n *defined using* $\phi_{n,i}$, *not* $\widehat{\phi}_{n,i}$).

Comment. Theorem 2 shows that the infeasible and feasible QGLS statistics have the same asymptotic distributions under Assumption CHE.

We now provide sufficient conditions for Assumption CHE. Suppose $\{\widehat{\phi}^2_{n,i} : i \leq n\}$ are based on a parametric model with conditional heteroskedasticity parameter π

estimated using residuals. Let $\widetilde{\pi}_n$ be the estimator of π and let $(\widetilde{\alpha}_n, \widetilde{\rho}_n)$ be the estimators of $(\widetilde{\alpha}, \rho)$ used to construct the residuals, where $\widetilde{\alpha}$ is the intercept when the model is written in regression form, see (2.2). For example, $\widetilde{\pi}_n$ may be an estimator of π based on residuals in place of the true errors and $(\widetilde{\alpha}_n, \widetilde{\rho}_n)$ may be the LS estimators (whose properties are covered by the asymptotic results given in Theorem 1 by taking $\phi_{n,i} = 1$). In particular, suppose that

$$\begin{aligned} \widehat{\phi}_{n,i}^2 &= \phi_{n,i}^2(\widetilde{\alpha}_n, \widetilde{\rho}_n, \widetilde{\pi}_n), \text{ where} \\ \phi_{n,i}^2(\widetilde{\alpha}, \rho, \pi) &= \omega + \sum_{j=1}^{L_i} \mu_j(\pi) \widehat{U}_{n,i-j}^2(\widetilde{\alpha}, \rho), \\ \widehat{U}_{n,i}(\widetilde{\alpha}, \rho) &= Y_{n,i} - \widetilde{\alpha} - \rho Y_{n,i-1}, \end{aligned} \tag{3.14}$$

$L_i = \min\{i-1, L\}$, and ω is an element of π. Here $L < \infty$ is a bound on the maximum number of lags allowed. Any model with stationary conditional heteroskedasticity (bounded away from the nonstationary region), such as a GARCH(1,1) model, can be approximated arbitrarily well by taking L sufficiently large. Hence, the restriction to finite lags is not overly restrictive. The upper bound L_i, rather than L, on the number of lags in the sum in (3.14) takes into account the truncation at 1 that naturally occurs because one does not observe residuals for $i < 1$.

The parameter space for π is Π, which is a bounded subset of R^{d_π}, for some $d_\pi > 0$. Let $\widehat{\pi}_n \in \Pi$ be an n^{δ_1}-consistent estimator of π for some $\delta_1 > 0$. For technical reasons, we base $\widehat{\phi}_{n,i}^2$ on an estimator $\widetilde{\pi}_n$ that is a discretized version of $\widehat{\pi}_n$ that takes values in a finite set Π_n ($\subset \Pi$) for $n \geq 1$, where Π_n consists of points on a uniform grid with grid size that goes to zero as $n \to \infty$ and hence the number of elements of Π_n diverges to infinity as $n \to \infty$. The reason for considering a discretized estimator is that when the grid size goes to zero more slowly than $n^{-\delta_1}$, then wp$\to$ 1 the estimators $\{\widetilde{\pi}_n : n \geq 1\}$ take values in a sequence of finite sets $\{\Pi_{n,0} : n \geq 1\}$ whose numbers of elements is bounded as $n \to \infty$. The latter property makes it easier to verify Assumption CHE(ii). The set Π_n can be defined such that there is very little difference between $\widehat{\pi}_n$ and $\widetilde{\pi}_n$ in a finite sample of size n.

We employ the following sufficient condition for the FQGLS estimator to be asymptotically equivalent to the (infeasible) QGLS estimator.

Assumption CHE2. (i) $\widehat{\phi}_{n,i}^2$ satisfies (3.14) with $L < \infty$ and $\mu_j(\cdot) \geq 0$ for all $j = 1, ..., L$, (ii) $\phi_{n,i}^2 = \omega_n + \sum_{j=1}^{L} \mu_j(\pi_n) U_{n,i-j}^2$ and $\pi_n \to \pi_0$ for some $\pi_0 \in \Pi$ (and π_0 may depend on the sequence), where ω_n is an element of π_n, (iii) $a_n(\widetilde{\rho}_n - \rho_n) = O_p(1)$, $n^{1/2}\widetilde{\alpha}_n = O_p(1)$, and $n^{\delta_1}(\widehat{\pi}_n - \pi_n) = o_p(1)$ for some $\delta_1 > 0$ under any sequence $(U_{n,i}, \phi_{n,i}^2)$ that satisfies Assumption INNOV and for $Y_{n,i}$ defined as in Assumption CHE with $\alpha = \beta = 0$ satisfying Assumption STAT when $\rho_n < 1$, and with $\rho = \rho_n$ that satisfies $n(1-\rho_n) \to h_1 \in [0, \infty]$, where a_n is defined in (3.5), (iv) $\widetilde{\pi}_n$ minimizes $||\pi - \widehat{\pi}_n||$ over $\pi \in \Pi_n$ for $n \geq 1$, where Π_n ($\subset \Pi$) consists of points on a uniform grid with grid size $Cn^{-\delta_2}$ for some $0 < \delta_2 < \delta_1$ and $0 < C < \infty$, (v) Π bounds the intercept ω away from zero, and (vi) $\mu_j(\pi)$ is continuous on Π for $j = 1, ..., L$.

The part of Assumption CHE2(iii) concerning $\widetilde{\rho}_n$ holds for the LS estimator by Theorem 1(a) (by taking $\phi_{n,i} = 1$), the part concerning $\widetilde{\alpha}_n$ holds for the LS estimator by similar, but simpler, arguments, and typically the part concerning $\widehat{\pi}_n$ holds for all $\delta_1 < 1/2$. Assumptions CHE2(iv)-(vi) can always be made to hold by choice of $\widehat{\pi}_n$, Π, and $\mu_j(\pi)$.

Lemma 1 *Assumption* CHE2 *implies Assumption* CHE.

Comment. The use of a discretized estimator $\widetilde{\pi}_n$ and a finite bound L on the number of lags in Assumption CHE2 are made for technical convenience. Undoubtedly, they are not necessary for the Lemma to hold (although other conditions may be needed in their place).

4 Proofs

This section provides proofs of Theorems 1 and 2 and Lemma 1. Section 4.1.1 states Lemmas 2-9, which are used in the proof of Theorem 1. Section 4.1.2 proves Theorem 1. Section 4.1.3 proves Lemmas 2-9. Section 4.2 proves Theorem 2. Section 4.3 proves Lemma 1.

To simplify notation, in the remainder of the paper we omit the subscript F_n on expectations.

4.1 Proof of Theorem 1

4.1.1 Lemmas 2-9

The proof of Theorem 1 uses eight lemmas that we state in this section. The first four lemmas deal with the case of $h_1 \in [0, \infty)$. The last four deal with the case of $h_1 = \infty$.

In integral expressions below, we often leave out the lower and upper limits zero and one, the argument r, and dr to simplify notation when there is no danger of confusion. For example, $\int_0^1 I_h(r)^2 dr$ is typically written as $\int I_h^2$. By " $\Rightarrow$ " we denote weak convergence as $n \to \infty$.

Lemma 2 *Suppose Assumptions* INNOV *and* STAT *hold,* $\rho_n \in (-1, 1)$ *and* $\rho_n = 1 - h_{n,1}/n$ *where* $h_{n,1} \to h_1 \in [0, \infty)$ *as* $n \to \infty$. *Then,*

$$(2h_{n,1}/n)^{1/2} Y_{n,0}^* / \lambda_{n,1}^{1/2} \to_d Z_1 \sim N(0, 1).$$

Define $h_{n,1}^* \geq 0$ by $\rho_n = \exp(-h_{n,1}^*/n)$. As shown in the proof of Lemma 2, $h_{n,1}^*/h_{n,1} \to 1$ when $h_1 \in [0, \infty)$. By recursive substitution, we have

$$\begin{aligned} Y_{n,i}^* &= \widetilde{Y}_{n,i} + \exp(-h_{n,1}^* i/n) Y_{n,0}^*, \text{ where} \\ \widetilde{Y}_{n,i} &= \textstyle\sum_{j=1}^{i} \exp(-h_{n,1}^*(i-j)/n) U_{n,j}. \end{aligned} \tag{4.1}$$

Let $BM(\Omega)$ denote a bivariate Brownian motion on $[0, 1]$ with variance matrix Ω. The next lemma is used to establish the simplified form of the asymptotic distribution that appears in Theorem 1(a).

Lemma 3 *Suppose* $(h_{2,1}^{1/2} W(r), M(r))' = BM(\Omega)$, *where*

$$\Omega = \begin{bmatrix} h_{2,1} & h_{2,3} \\ h_{2,3} & h_{2,2} \end{bmatrix}.$$

Then, $M(r)$ *can be written as* $M(r) = h_{2,2}^{1/2} \left(h_{2,7} W(r) + (1 - h_{2,7}^2)^{1/2} W_2(r) \right)$, *where* $(W(r), W_2(r))' = BM(I_2)$ *and* $h_{2,7} = h_{2,3}/(h_{2,1} h_{2.2})^{1/2}$ *is the correlation that arises in the variance matrix* Ω.

The following Lemma states some general results on weak convergence of certain statistics to stochastic integrals. It is proved using Theorems 4.2 and 4.4 of Hansen (1992) and Lemma 2 above. Let $\otimes$ denote the Kronecker product.

Lemma 4 *Suppose* $\{v_{n,i} : i \leq n, n \geq 1\}$ *is a triangular array of row-wise strictly-stationary strong-mixing random* d_v*-vectors with* (i) *strong-mixing numbers* $\{\alpha_n(m) : m \geq 1, n \geq 1\}$ *that satisfy* $\alpha(m) = \sup_{n \geq 1} \alpha_n(m) = O(m^{-\zeta\tau/(\zeta-\tau)})$ *as* $m \to \infty$ *for some* $\zeta > \tau > 2$, *and* (ii) $\sup_{n \geq 1} ||v_{n,i}||_\zeta < \infty$. *Suppose* $n^{-1}EV_nV_n' \to \Omega_0$ *as* $n \to \infty$, *where* $V_n = \sum_{i=1}^n v_{n,i}$, *and* Ω_0 *is some* $d_v \times d_v$ *variance matrix. Let* $X_{n,i} = \rho_n X_{n,i-1} + v_{n,i}$, *where* $n(1-\rho_n) \to h_1 \in [0, \infty)$. *If* $h_1 > 0$, *the first element of* $X_{n,i}$ *has a stationary initial condition and all of the other elements have zero initial conditions. If* $h_1 = 0$, *all of the elements of* $X_{n,i}$ *have zero initial conditions, i.e.,* $X_{n,0} = 0$. *Let* $\Lambda = \lim_{n\to\infty} n^{-1} \sum_{i=1}^n \sum_{j=i+1}^n E v_{n,i} v_{n,j}'$. *Let* $K_h(r) = \int_0^r \exp((r-s)h_1) dB(s)$, *where* $B(\cdot)$ *is a* d_v*-vector* $BM(\Omega_0)$ *on* $[0,1]$. *If* $h_1 > 0$, *let* $K_h^*(r) = K_h(r) + e_1(2h_1)^{-1/2}\exp(-h_1 r)\Omega_{0,1,1}^{1/2} Z_1$, *where* $Z_1 \sim N(0,1)$ *is independent of* $B(\cdot)$, $e_1 = (1, 0, ..., 0)' \in R^{d_v}$, *and* $\Omega_{0,1,1}$ *denotes the* $(1,1)$ *element of* Ω_0. *If* $h_1 = 0$, *let* $K_h^*(r) = K_h(r)$. *Then,*
(a) $n^{-1/2} X_{n,[nr]} \Rightarrow K_h^*(r)$,
(b) $n^{-1} \sum_{i=1}^n X_{n,i-1} v_{n,i}' \to_d \int K_h^* dB' + \Lambda$, *and*
(c) *for* $\tau \geq 3$, $n^{-3/2} \sum_{i=1}^n (X_{n,i-1} \otimes X_{n,i-1}) v_{n,i}' \to_d \int (K_h^* \otimes K_h^*) dB' + \left(\Lambda \otimes \int K_h^*\right) + \left(\int K_h^* \otimes \Lambda\right)$.

We now use Lemma 4 to establish the following results which are key in the proof of Theorem 1(a). Let $[a]$ denote the integer part of a.

Lemma 5 *Suppose Assumptions* INNOV *and* STAT *hold,* $\rho_n \in (-1, 1]$, $\rho_n = 1 - h_{n,1}/n$ *where* $h_{n,1} \to h_1 \in (0, \infty)$. *Then, the following results* (a)-(k) *hold jointly,*
(a) $n^{-1/2} Y_{n,[nr]}^* \Rightarrow h_{2,1}^{1/2} I_h^*(r)$,
(b) $n^{-1} \sum_{i=1}^n \phi_{n,i}^{-j} \to_p \lim_{n\to\infty} E\phi_{n,i}^{-j} = h_{2,(j+3)}$ *for* $j = 1, 2, 4$,
(c) $n^{-1} \sum_{i=1}^n U_{n,i}/\phi_{n,i}^4 \to_p \lim_{n\to\infty} E(U_{n,i}/\phi_{n,i}^4) = 0$,
(d) $n^{-1} \sum_{i=1}^n U_{n,i}^2/\phi_{n,i}^4 \to_p \lim_{n\to\infty} E(U_{n,i}^2/\phi_{n,i}^4) = h_{2,2}$,
(e) $n^{-1/2} \sum_{i=1}^n U_{n,i}/\phi_{n,i}^2 \to_d M(1) = \int dM = h_{2,2}^{1/2} \int d[h_{2,7}W(r) + (1-h_{2,7}^2)^{1/2} W_2(r)]$,
(f) $n^{-3/2} \sum_{i=1}^n Y_{n,i-1}^*/\phi_{n,i}^2 = n^{-3/2} \sum_{i=1}^n Y_{n,i-1}^* E\phi_{n,1}^{-2} + O_p(n^{-1/2}) \to_d h_{2,5} h_{2,1}^{1/2} \int I_h^*$,
(g) $n^{-1} \sum_{i=1}^n Y_{n,i-1}^* U_{n,i}/\phi_{n,i}^2 \to_d h_{2,1}^{1/2} \int I_h^* dM = h_{2,2}^{1/2} h_{2,1}^{1/2} \int I_h^* d[h_{2,7}W(r) +$
$(1-h_{2,7}^2)^{1/2} W_2(r)]$,
(h) $n^{-2} \sum_{i=1}^n Y_{n,i-1}^{*2}/\phi_{n,i}^2 = n^{-2} \sum_{i=1}^n Y_{n,i-1}^{*2} E\phi_{n,1}^{-2} + O_p(n^{-1/2}) \to_d h_{2,5} h_{2,1} \int I_h^{*2}$,
(i) $n^{-3/2} \sum_{i=1}^n Y_{n,i-1}^* U_{n,i}^2/\phi_{n,i}^4 = n^{-3/2} \sum_{i=1}^n Y_{n,i-1}^* E(U_{n,1}^2/\phi_{n,1}^4) + O_p(n^{-1/2})$
$\to_d h_{2,2} h_{2,1}^{1/2} \int I_h^*$,
(j) $n^{-2} \sum_{i=1}^n Y_{n,i-1}^{*2} U_{n,i}^2/\phi_{n,i}^4 = n^{-2} \sum_{i=1}^n Y_{n,i-1}^{*2} E(U_{n,1}^2/\phi_{n,1}^4) + O_p(n^{-1/2})$
$\to_d h_{2,2} h_{2,1} \int I_h^{*2}$,

(k) $n^{-1-\ell_1/2}\sum_{i=1}^n Y_{n,i-1}^{*\ell_1}U_{n,i}^{\ell_2}/\phi_{n,i}^4 = o_p(n)$ *for* $(\ell_1,\ell_2) = (1,0)$, $(1,1)$, $(2,0)$, $(2,1)$, $(3,0)$, $(3,1)$, *and* $(4,0)$, *and*
(l) *when* $h_1 = 0$, *parts* (a) *and* (f)-(k) *hold with* $Y_{n,i-1}^*$ *replaced by* $\widetilde{Y}_{n,i-1}$.

In the proof of Theorem 1(b), we use the following well-known strong-mixing covariance inequality, see e.g. Doukhan (1994, Thm. 3, p. 9). Let X and Y be strong-mixing random variables with respect to σ-fields $\mathcal{F}_i^j$ (for integers $i \leq j$) such that $X \in \mathcal{F}_{-\infty}^n$ and $Y \in \mathcal{F}_{n+m}^\infty$ with strong-mixing numbers $\{\alpha(m) : m \geq 1\}$. For $p,q > 0$ such that $1-p^{-1}-q^{-1} > 0$, let $||X||_p = (E|X|^p)^{1/p}$ and $||Y||_q = (E|Y|^q)^{1/q}$. Then, the following inequality holds

$$Cov(X,Y) \leq 8||X||_p||Y||_q\alpha(k)^{1-p^{-1}-q^{-1}}. \tag{4.2}$$

The proof of Theorem 1(b) uses the following technical Lemmas. The Lemmas make repeated use of the mixing inequality (4.2) applied with $p = q = \zeta > 3$, where ζ appears in Assumption INNOV.

Lemma 6 *Suppose* $n(1-\rho_n) \to \infty$, $\rho_n \to 1$, *and Assumptions* INNOV *and* STAT *hold, then we have*

$$\begin{aligned} E(Y_{n,0}^{*2}U_{n,1}^2/\phi_{n,1}^4) - (1-\rho_n^2)^{-1}(EU_{n,1}^2)^2/\phi_{n,1}^4 &= O(1), \\ E(Y_{n,0}^{*2}/\phi_{n,1}^2) - (1-\rho_n^2)^{-1}EU_{n,1}^2E\phi_{n,1}^{-2} &= O(1), \\ E(Y_{n,0}^{*}/\phi_{n,1}^2) &= O(1), \text{ and} \\ E(Y_{n,0}^{*}U_{n,1}^2/\phi_{n,1}^4) &= O(1). \end{aligned}$$

Lemma 7 *Suppose* $n(1-\rho_n) \to \infty$, $\rho_n \to 1$ *and Assumptions* INNOV *and* STAT *hold, then we have*

$$E\left(\sum_{i=1}^n [E\zeta_{n,i}^2 - E(\zeta_{n,i}^2|\mathcal{G}_{n,i-1})]\right)^2 \to 0, \text{ where } \zeta_{n,i} \equiv n^{-1/2}\frac{Y_{n,i-1}^*U_{n,i}/\phi_{n,i}^2}{(E(Y_{n,0}^{*2}U_{n,1}^2/\phi_{n,1}^4))^{1/2}}.$$

Lemma 8 *Suppose* $n(1-\rho_n) \to \infty$, $\rho_n \to 1$, *and Assumptions* INNOV *and* STAT *hold, then we have*
(a) $n^{-1}(1-\rho_n)^{1/2}X_1'X_2 = o_p(1)$,
(b) $E(Y_{n,0}^{*2}/\phi_{n,1}^2)^{-1}n^{-1}X_1'X_1 \to_p 1$,
(c) $(E(Y_{n,0}^{*2}U_{n,1}^2/\phi_{n,1}^4))^{-1}n^{-1}\sum_{i=1}^n(Y_{n,i-1}^{*2}U_{n,i}^2/\phi_{n,i}^4) \to_p 1$,
(d) $(X'X)^{-1}X'U = (O_p((1-\rho_n)^{1/2}n^{-1/2}), O_p(n^{-1/2}))'$,
(e) $(E(Y_{n,0}^{*2}U_{n,1}^2/\phi_{n,1}^4))^{-1}n^{-1}X_1'\Delta^2X_1 \to_p 1$,
(f) $(1-\rho_n)^{1/2}n^{-1}(X_2'\Delta^2X_1) = O_p(1)$, *and*
(g) $n^{-1}(X_2'\Delta^2X_2) = O_p(1)$.

Lemma 9 *Suppose* $n(1-\rho_n) \to \infty$, $\rho_n \to 1$, *and Assumptions* INNOV *and* STAT *hold, we have* $\sum_{i=1}^n E(\zeta_{n,i}^2 1(|\zeta_{n,i}| > \delta)|\mathcal{G}_{n,i-1}) \to_p 0$ *for any* $\delta > 0$.

4.1.2 Proof of Theorem 1

To simplify notation, in the remainder of the paper we often leave out the subscript n. For example, instead of $\rho_n, \sigma_{U,n}^2, Y_{n,i}^*, U_{n,i}, \phi_{n,i}, \widehat{\phi}_{n,i}$, and $\zeta_{n,i}$, we write $\rho, \sigma_U^2, Y_i^*, U_i$, $\phi_i, \widehat{\phi}_i$, and ζ_i. We do not drop n from $h_{n,1}$ because $h_{n,1}$ and h_1 are different quantities. As above, we omit the subscript F_n on expectations.

In the proofs of Theorem 1 and Lemmas 2-9 below, X_1, X_2, U, Δ, and Y are defined as in the paragraph containing (2.4), but with ϕ_i in place of $\widehat{\phi}_{n,i}$.

Proof of Theorem 1. First we prove part (a) of the Theorem when $h_1 > 0$. In this case, $a_n = n$ and $d_n = n^{1/2}$. We can write

$$\begin{aligned} n(\widehat{\rho}_n - \rho) &= \left(n^{-2}X_1'M_{X_2}X_1\right)^{-1} n^{-1}X_1'M_{X_2}U \text{ and} \\ n\widehat{\sigma}_n^2 &= \left(n^{-2}X_1'M_{X_2}X_1\right)^{-1} \left(n^{-2}X_1'M_{X_2}\Delta^2 M_{X_2}X_1\right) \left(n^{-2}X_1'M_{X_2}X_1\right)^{-1}. \end{aligned} \tag{4.3}$$

We consider the terms in (4.3) one at a time. First, we have

$$\begin{aligned} & n^{-2}X_1'M_{X_2}X_1 \\ &= n^{-2}\sum_{i=1}^n \left(Y_{i-1}^*/\phi_i - \left(\sum_{j=1}^n Y_{j-1}^*/\phi_j^2\right)\left(\sum_{j=1}^n \phi_j^{-2}\right)^{-1}\phi_i^{-1}\right)^2 \\ &= n^{-2}\sum_{i=1}^n Y_{i-1}^{*2}/\phi_i^2 - \left(n^{-3/2}\sum_{j=1}^n Y_{j-1}^*/\phi_j^2\right)^2 \left(n^{-1}\sum_{j=1}^n \phi_j^{-2}\right)^{-1} \\ &\to_d h_{2,5}h_{2,1}\int I_h^{*2} - \left(h_{2,5}h_{2,1}^{1/2}\int I_h^*\right)^2 h_{2,5}^{-1} = h_{2,5}h_{2,1}\int I_{D,h}^{*2}, \end{aligned} \tag{4.4}$$

where the first two equalities hold by definitions and some algebra, and the convergence holds by Lemma 5(b), (f), and (h) with $j = 2$ in part (b).

Similarly, we have

$$\begin{aligned} & n^{-1}X_1'M_{X_2}U \\ &= n^{-1}\sum_{i=1}^n \left(Y_{i-1}^*/\phi_i - \left(\sum_{j=1}^n Y_{j-1}^*/\phi_j^2\right)\left(\sum_{j=1}^n \phi_j^{-2}\right)^{-1}\phi_i^{-1}\right) U_i/\phi_i \\ &= n^{-1}\sum_{i=1}^n Y_{i-1}^*U_i/\phi_i^2 - \left(n^{-3/2}\sum_{j=1}^n Y_{j-1}^*/\phi_j^2\right)\left(n^{-1}\sum_{j=1}^n \phi_j^{-2}\right)^{-1} n^{-1/2}\sum_{i=1}^n U_i/\phi_i^2 \\ &\to_d h_{2,1}^{1/2}\int I_h^* dM - h_{2,1}^{1/2}\int I_h^* \int dM = h_{2,1}^{1/2}\int I_{D,h}^* dM, \end{aligned} \tag{4.5}$$

where the first two equalities hold by definitions and some algebra, and the convergence holds by Lemma 5(b) and (e)-(g) with $j = 2$ in part (b).

To determine the asymptotic distribution of $n^{-2}X_1'M_{X_2}\Delta^2 M_{X_2}X_1$, we make the following preliminary calculations. Let $\widehat{U}_i/\phi_i$ denote the ith element of $M_XY = M_XU$. That is,

$$\begin{aligned}
\widehat{U}_i/\phi_i &= U_i/\phi_i - A_n'B_n^{-1}\begin{pmatrix} n^{-1/2}\phi_i^{-1} \\ n^{-1}Y_{i-1}^*/\phi_i \end{pmatrix}, \text{ where} \\
A_n &= \begin{pmatrix} n^{-1/2}\sum_{j=1}^n U_j/\phi_j^2 \\ n^{-1}\sum_{j=1}^n Y_{j-1}^*U_j/\phi_j^2 \end{pmatrix} \text{ and} \\
B_n &= \begin{pmatrix} n^{-1}\sum_{j=1}^n \phi_j^{-2} & n^{-3/2}\sum_{j=1}^n Y_{j-1}^*/\phi_j^2 \\ n^{-3/2}\sum_{j=1}^n Y_{j-1}^*/\phi_j^2 & n^{-2}\sum_{j=1}^n Y_{j-1}^{*2}/\phi_j^2 \end{pmatrix}. \qquad (4.6)
\end{aligned}$$

Using (4.6), we have

$$\begin{aligned}
& n^{-2}\sum_{i=1}^n Y_{i-1}^{*2}\widehat{U}_i^2/\phi_i^4 = n^{-2}\sum_{i=1}^n Y_{i-1}^{*2}U_i^2/\phi_i^4 - 2n^{-1}A_n'B_n^{-1}\begin{pmatrix} n^{-3/2}\sum_{i=1}^n Y_{i-1}^{*2}U_i/\phi_i^4 \\ n^{-2}\sum_{i=1}^n Y_{i-1}^{*3}U_i/\phi_i^4 \end{pmatrix} \\
& +n^{-1}A_n'B_n^{-1}\begin{pmatrix} n^{-2}\sum_{i=1}^n Y_{i-1}^{*2}/\phi_i^4 & n^{-5/2}\sum_{i=1}^n Y_{i-1}^{*3}/\phi_i^4 \\ n^{-5/2}\sum_{i=1}^n Y_{i-1}^{*3}/\phi_i^4 & n^{-3}\sum_{i=1}^n Y_{i-1}^{*4}/\phi_i^4 \end{pmatrix} B_n^{-1}A_n \\
& = n^{-2}\sum_{i=1}^n Y_{i-1}^{*2}U_i^2/\phi_i^4 + o_p(1), \qquad (4.7)
\end{aligned}$$

where the second equality holds using Lemma 5(k) with $(\ell_1, \ell_2) = (2, 1), (3, 1), (2, 0),$ $(3, 0)$, and $(4, 0)$ and to show that A_n and B_n^{-1} are $O_p(1)$ we use Lemma 5(b) and (e)-(h) with $j = 2$ in part (b).

Similarly to (4.7) but with Y_{i-1}^* in place of Y_{i-1}^{*2}, and then with Y_{i-1}^{*2} deleted, we have

$$\begin{aligned}
n^{-3/2}\sum_{i=1}^n Y_{i-1}^*\widehat{U}_i^2/\phi_i^4 &= n^{-3/2}\sum_{i=1}^n Y_{i-1}^*U_i^2/\phi_i^4 + o_p(1) \text{ and} \\
n^{-1}\sum_{i=1}^n \widehat{U}_i^2/\phi_i^4 &= n^{-1}\sum_{i=1}^n U_i^2/\phi_i^4 + o_p(1) \qquad (4.8)
\end{aligned}$$

using Lemma 5 as above to show that A_n and B_n^{-1} are $O_p(1)$, using Lemma 5(k) with $(\ell_1, \ell_2) = (1, 1), (2, 1), (1, 0), (2, 0)$, and $(3, 0)$ for the first result, and using Lemma 5(k) with $(\ell_1, \ell_2) = (1, 1), (1, 0)$, and $(2, 0)$, Lemma 5(b) with $j = 4$, and Lemma 5(c) for the second result.

We now have

$$\begin{aligned}
& n^{-2}X_1'M_{X_2}\Delta^2 M_{X_2}X_1 \\
= \; & n^{-2}\sum_{i=1}^n \left(\widehat{U}_i^2/\phi_i^2\right)\left(Y_{i-1}^*/\phi_i - \left(\sum_{j=1}^n Y_{j-1}^*/\phi_j^2\right)\left(\sum_{j=1}^n \phi_j^{-2}\right)^{-1}\phi_i^{-1}\right)^2
\end{aligned}$$

$$
\begin{aligned}
&= n^{-2}\sum_{i=1}^{n} Y_{i-1}^{*2}\widehat{U}_i^2/\phi_i^4 - 2\left(n^{-3/2}\sum_{j=1}^{n} Y_{j-1}^{*}/\phi_j^2\right)\left(n^{-1}\sum_{j=1}^{n}\phi_j^{-2}\right)^{-1} n^{-3/2}\sum_{i=1}^{n} Y_{i-1}^{*}\widehat{U}_i^2/\phi_i^4 \\
&\quad + \left(n^{-3/2}\sum_{j=1}^{n} Y_{j-1}^{*}/\phi_j^2\right)^2\left(n^{-1}\sum_{j=1}^{n}\phi_j^{-2}\right)^{-2} n^{-1}\sum_{i=1}^{n}\widehat{U}_i^2\phi_i^{-4} \\
&= n^{-2}\sum_{i=1}^{n} Y_{i-1}^{*2}U_i^2/\phi_i^4 - 2\left(n^{-3/2}\sum_{j=1}^{n} Y_{j-1}^{*}/\phi_j^2\right)\left(n^{-1}\sum_{j=1}^{n}\phi_j^{-2}\right)^{-1} n^{-3/2}\sum_{i=1}^{n} Y_{i-1}^{*}U_i^2/\phi_i^4 \\
&\quad + \left(n^{-3/2}\sum_{j=1}^{n} Y_{j-1}^{*}/\phi_j^2\right)^2\left(n^{-1}\sum_{j=1}^{n}\phi_j^{-2}\right)^{-2} n^{-1}\sum_{i=1}^{n} U_i^2/\phi_i^4 + O_p(n^{-1}) \\
&\to_d h_{2,2}h_{2,1}\int I_h^{*2} - 2h_{2,1}^{1/2}\int I_h^{*}\cdot\left(h_{2,2}h_{2,1}^{1/2}\int I_h^{*}\right) + \left(h_{2,1}^{1/2}\int I_h^{*}\right)^2 h_{2,2} \\
&= h_{2,2}h_{2,1}\int\left(I_h^{*} - \int I_h^{*}\right)^2 = h_{2,2}h_{2,1}\int I_{D,h}^{*2}, \qquad (4.9)
\end{aligned}
$$

where the first two equalities follow from definitions and some algebra, the third equality holds by (4.7), (4.8), and Lemma 5(b), (d), (f), (i), and (j) with $j = 2$ in part (b), and the convergence holds by the same parts of Lemma 5.

Putting the results of (4.3), (4.4), (4.5), (4.9), and Lemma 3 together gives

$$
\begin{aligned}
T_n^{*}(\rho_n) &\to_d \frac{h_{2,1}^{1/2}\int I_{D,h}^{*}dM}{\left(h_{2,2}h_{2,1}\int I_{D,h}^{*2}\right)^{1/2}} \\
&= \frac{h_{2,2}^{1/2}\int I_{D,h}^{*}d\left(h_{2,7}W + (1-h_{2,7}^2)^{1/2}W_2\right)}{h_{2,2}^{1/2}\left(\int I_{D,h}^{*2}\right)^{1/2}} \\
&= h_{2,7}\left(\int I_{D,h}^{*2}\right)^{-1/2}\int I_{D,h}^{*}dW + (1-h_{2,7}^2)^{1/2}Z_2, \qquad (4.10)
\end{aligned}
$$

where the last equality uses the definition of Z_2 in (3.10). This completes the proof of part (a) of the Theorem when $h_1 > 0$.

Next, we consider the case where $h_1 = 0$. In this case, (4.3)-(4.10) hold except that the convergence results in (4.4), (4.5), and (4.9) only hold with Y_{i-1}^{*} replaced by $\widetilde{Y}_{i-1}$ because Lemma 5(l) only applies to random variables based on a zero initial condition when $h_1 = 0$. Hence, we need to show that the difference between the second last line of (4.4) with Y_{i-1}^{*} appearing and with $\widetilde{Y}_{i-1}$ appearing is $o_p(1)$ and that analogous results hold for (4.5) and (4.9).

For $h_1 = 0$, by a mean value expansion, we have

$$
\begin{aligned}
&\max_{0\le j\le 2n}|1-\rho^j| = \max_{0\le j\le 2n}|1-\exp(-h_{n,1}^{*}j/n)| = \max_{0\le j\le 2n}|1-(1-h_{n,1}^{*}j\exp(m_j)/n)| \\
&\le 2h_{n,1}^{*}\max_{0\le j\le 2n}|\exp(m_j)| = O(h_{n,1}^{*}), \qquad (4.11)
\end{aligned}
$$

for $0 \le |m_j| \le h^*_{n,1} j/n \le 2h^*_{n,1} \to 0$, where $h^*_{n,1}$ is defined just above (4.1).

Using the decomposition in (4.1), we have $Y^*_{i-1} = \widetilde{Y}_{i-1} + \rho^{i-1} Y^*_0$. To show the desired result for (4.4), we write the second last line of (4.4) as

$$\begin{aligned}
& n^{-2} \sum_{i=1}^{n} \left(Y^*_{i-1}/\phi_i - \left(\sum_{j=1}^{n} Y^*_{j-1}/\phi_j^2 \right) \left(\sum_{j=1}^{n} \phi_j^{-2} \right)^{-1} \phi_i^{-1} \right)^2 \\
&= n^{-2} \sum_{i=1}^{n} \left(\widetilde{Y}_{i-1}/\phi_i + \rho^{i-1} Y^*_0/\phi_i - \left(\sum_{j=1}^{n} \widetilde{Y}_{j-1}/\phi_j^2 + \rho^{j-1} Y^*_0/\phi_j^2 \right) \left(\sum_{j=1}^{n} \phi_j^{-2} \right)^{-1} \phi_i^{-1} \right)^2 \\
&= n^{-2} \sum_{i=1}^{n} \left(\widetilde{Y}_{i-1}/\phi_i - \left(\sum_{j=1}^{n} \widetilde{Y}_{j-1}/\phi_j^2 \right) \left(\sum_{j=1}^{n} \phi_j^{-2} \right)^{-1} \phi_i^{-1} + O_p(h^*_{n,1} Y^*_0)/\phi_i \right)^2 \qquad (4.12) \\
&= n^{-2} \sum_{i=1}^{n} \left(\widetilde{Y}_{i-1}/\phi_i - \left(\sum_{j=1}^{n} \widetilde{Y}_{j-1}/\phi_j^2 \right) \left(\sum_{j=1}^{n} \phi_j^{-2} \right)^{-1} \phi_i^{-1} \right)^2 + O_p(n^{-1/2} h^*_{n,1} Y^*_0),
\end{aligned}$$

where the second equality holds because $\rho^{i-1} = 1 + O(h^*_{n,1})$ uniformly in $i \le n$ by (4.11), and the third equality holds using Lemma 5. Next, Lemma 2 and $h^*_{n,1}/h_{n,1} \to 1$ (which is established at the beginning of the proof of Lemma 2) show that $n^{-1/2} h^*_{n,1} Y^*_0$ $= O_p(h^{*1/2}_{n,1}) = o_p(1)$. This completes the proof of the desired result for (4.4) when $h_1 = 0$. The proofs for (4.5) and (4.9) are similar. This completes the proof of part (a) of the Theorem.

It remains to consider the case where $h_1 = \infty$, i.e., part (b) of the Theorem. The results in part (b) generalize the results in Giraitis and Phillips (2006) in the following ways: (i) from a no-intercept model to a model with an intercept, (ii) to a case in which the innovation distribution depends on n, (iii) to allow for conditional heteroskedasticity in the error distribution, (iv) to cover a quasi-GLS estimator in place of the LS estimator, and (v) to cover the standard deviation estimator as well as the GLS/LS estimator itself.

It is enough to consider the two cases $\rho \to \rho^* < 1$ and $\rho \to 1$. First, assume $\rho \to 1$ and $n(1-\rho) \to \infty$. In this case, the sequences a_n and d_n are equal to the expressions in (3.6) up to lower order terms. We first prove $a_n(\widehat{\rho}_n - \rho) \to_d N(0,1)$. Note that

$$a_n(\widehat{\rho}_n - \rho) = \left(n^{-1} \frac{X_1' M_{X_2} X_1}{E(Y_0^{*2}/\phi_1^2)} \right)^{-1} \frac{n^{-1/2} X_1' M_{X_2} U}{(E(Y_0^{*2} U_1^2/\phi_1^4))^{1/2}} \equiv \nu_n \xi_n, \qquad (4.13)$$

where ν_n and ξ_n have been implicitly defined. We now show $\nu_n \to_p 1$ and $\xi_n \to_d N(0,1)$.

To show the latter, define the martingale difference sequence

$$\zeta_i \equiv n^{-1/2} \frac{Y^*_{i-1} U_i/\phi_i^2}{(E(Y_0^{*2} U_1^2/\phi_1^4))^{1/2}}. \qquad (4.14)$$

We show that

$$\frac{n^{-1/2}X_1'P_{X_2}U}{(E(Y_0^{*2}U_1^2/\phi_1^4))^{1/2}} \to_p 0 \text{ and } \sum_{i=1}^n \zeta_i \to_d N(0,1). \tag{4.15}$$

To show the first result, note that $n^{-1/2}X_2'U = n^{-1/2}\sum_{i=1}^n U_i/\phi_i^2 = O_p(1)$ by a CLT for a triangular array of martingale difference random variables U_i/ϕ_i^2 for which $E|U_i/\phi_i^2|^3 < \infty$ and $n^{-1}\sum_{i=1}^n(U_i^2/\phi_i^4 - EU_i^2/\phi_i^4) \to_p 0$. The latter convergence in probability condition holds by Lemma 5(d). Furthermore, $(n^{-1}X_2'X_2)^{-1} = O_p(1)$ by Lemma 5(d) and Assumption INNOV(vii). Finally, $n^{-1}(1-\rho)^{1/2}X_1'X_2 = n^{-1}(1-\rho)^{1/2}\sum_{i=1}^n Y_{i-1}^*/\phi_i^2 = o_p(1)$ by Lemma 8(a). The first result in (4.15) then follows because $E(Y_0^{*2}U_1^2/\phi_1^4) = O((1-\rho)^{-1})$ by Lemma 6.

To show the latter we adjust the proof of Lemma 1 in Giraitis and Phillips (2006). It is enough to prove the analogue of equations (11) and (12) in Giraitis and Phillips (2006), namely the Lindeberg condition $\sum_{i=1}^n E(\zeta_i^2 1(|\zeta_i| > \delta)|\mathcal{G}_{i-1}) \to_p 0$ for any $\delta > 0$ and $\sum_{i=1}^n E(\zeta_i^2|\mathcal{G}_{i-1}) \to_p 1$. Lemma 9 shows the former and Lemma 7 implies the latter, because by stationarity (within rows) we have $\sum_{i=1}^n E\zeta_i^2 = 1$.

By Lemma 8(b) and Lemma 6

$$\frac{n^{-1}X_1'X_1}{E(Y_0^{*2}/\phi_1^2)} \to_p 1 \text{ and } \frac{n^{-1}X_1'P_{X_2}X_1}{E(Y_0^{*2}/\phi_1^2)} \to_p 0 \tag{4.16}$$

which imply $\nu_n \to_p 1$.

We next show that $d_n\widehat{\sigma}_n \to_p 1$. By (4.16) it is enough to show that

$$\frac{n^{-1}X_1'M_{X_2}\Delta^2 M_{X_2}X_1}{E(Y_0^{*2}U_1^2/\phi_1^4)} \to_p 1. \tag{4.17}$$

Lemma 8(e)-(g) shows that $(E(Y_0^{*2}U_1^2/\phi_1^4))^{-1}n^{-1}X_1'\Delta^2X_1 \to_p 1$, $(1-\rho)^{1/2} \times n^{-1}(X_2'\Delta^2X_1) = O_p(1)$, and $n^{-1}(X_2'\Delta^2X_2) = O_p(1)$. These results combined with Lemma 6, $(n^{-1}X_2'X_2)^{-1} = O_p(1)$, and $n^{-1}(1-\rho)^{1/2}X_1'X_2 = o_p(1)$ imply (4.17).

In the case $\rho \to \rho^* < 1$, Theorem 1(b) follows by using appropriate CLTs for martingale difference sequences and weak laws of large numbers. For example, the analogue to the expression in parentheses in (4.13) satisfies

$$\frac{n^{-1}X_1'M_{X_2}X_1}{E(Y_0^{*2}/\phi_1^2) - (E(Y_0^*/\phi_1^2))^2/E(\phi_1^{-2})} \to_p 1. \tag{4.18}$$

This follows by a weak law of large numbers for triangular arrays of mean zero, $L^{1+\delta}$ bounded (for some $\delta > 0$), near-epoch dependent random variables. Andrews (1988, p.464) shows that the latter conditions imply that the array is a uniformly integrable L^1 mixingale for which a WLLN holds, see Andrews (1988, Thm. 2). For example, to show $n^{-1}X_1'X_1 - E(Y_0^{*2}/\phi_1^2) \to_p 0$, note that $Y_{i-1}^{*2}/\phi_1^2 - EY_0^{*2}/\phi_1^2$ is near-epoch dependent with respect to the σ-field $\mathcal{G}_i$ using the moment conditions in Assumption INNOV(iv), $\sum_{j=0}^\infty \rho^{*j} = (1-\rho^*)^{-1} < \infty$, and $\rho \to \rho^* < 1$. $\square$

4.1.3 Proof of Lemmas 2-9

Proof of Lemma 2. We have: $\rho_n = 1 - h_{n,1}/n$ and $h_{n,1} = O(1)$ implies that $\rho_n \to 1$. Hence, $\exp(-h_{n,1}^*/n) = \rho_n \to 1$ and $h_{n,1}^* = o(n)$. By a mean-value expansion of $\exp(-h_{n,1}^*/n)$ about 0,

$$0 = \rho_n - \rho_n = \exp(-h_{n,1}^*/n) - (1 - h_{n,1}/n) = h_{n,1}/n - \exp(-h_{n,1}^{**}/n)h_{n,1}^*/n, \quad (4.19)$$

where $h_{n,1}^{**} = o(n)$ given that $h_{n,1}^* = o(n)$. Hence, $h_{n,1} - (1+o(1))h_{n,1}^* = 0$, $h_{n,1}^*/h_{n,1} \to 1$, and it suffices to prove the result with $h_{n,1}^*$ in place of $h_{n,1}$.

Let $\{m_n : n \geq 1\}$ be a sequence such that $m_n h_{n,1}^*/n \to \infty$. By Assumption STAT (which holds because $\rho_n < 1$), we can write $(2h_{n,1}^*/n)^{1/2} Y_0^*/\lambda_{n,1}^{1/2} = A_{1n} + A_{2n}$ for $A_{1n} = (2h_{n,1}^*/n)^{1/2} \sum_{j=0}^{m_n} \rho_n^j U_{-j} \lambda_{n,1}^{1/2}$ and $A_{2n} = (2h_{n,1}^*/n)^{1/2} \sum_{j=m_n+1}^{\infty} \rho_n^j U_{-j} \lambda_{n,1}^{1/2}$. Note that $EA_{2n} = 0$ and

$$\begin{aligned} & var(A_{2n}) = (2h_{n,1}^*/n) \textstyle\sum_{j=m_n+1}^{\infty} \rho_n^{2j} = (2h_{n,1}^*/n)\rho_n^{2(m_n+1)}/(1-\rho_n^2) \quad (4.20) \\ & = (2h_{n,1}^*/n)\rho_n^{2(m_n+1)}/((2h_{n,1}^*/n)(1+o(1))) = O(\exp(-2(m_n+1)h_{n,1}^*/n)) = o(1), \end{aligned}$$

where the third equality holds because $\rho_n^2 = \exp(-2h_{n,1}^*/n) = 1 - (2h_{n,1}^*/n)(1+o(1))$ by a mean value expansion and the last equality holds because $m_n h_{n,1}^*/n \to \infty$ by assumption. Therefore, $A_{2n} \to_p 0$.

The result now follows from $A_{1n} \to_d Z_1$ which holds by the CLT in Corollary 3.1 in Hall and Heyde (1980) with their $X_{n,i}$ being equal to $(2h_{n,1}^*/n)^{1/2} \rho_n^i U_{-i}/\lambda_{n,1}^{1/2}$. To apply their Corollary 3.1 we have to verify their (3.21), a Lindeberg condition, and a conditional variance condition. For all i=..., 0, 1, ... set $\mathcal{F}_{0,i} = \varnothing$ and define recursively $\mathcal{F}_{n+1,i} = \sigma(\mathcal{F}_{n,i} \cup \sigma(U_{n+1,j} : j = 0, -1, ..., -i))$ for $n \geq 1$. Then, (3.21) in Hall and Heyde (1980) holds automatically. To check the remaining two conditions, note first that $\sum_{i=0}^{m_n} E(X_{n,i}^2 | \mathcal{F}_{n,i-1}) = \sum_{i=0}^{m_n} EX_{n,i}^2 = 2h_{n,1}^* \sum_{i=0}^{m_n} \rho_n^{2i}/n \to 1$ which holds because $\sum_{i=0}^{m_n} \rho_n^{2i} = (1-\rho_n^{2(m_n+1)})/(1-\rho_n^2)$, $\rho_n^{2(m_n+1)} = \exp(-2h_{n,1}^*(m_n+1)/n) \to 0$, and

$$n(1-\rho_n^2) = n(1-\rho_n)(1+\rho_n) = h_{n,1}(1+\rho_n) \to 2h. \quad (4.21)$$

Secondly, for $\varepsilon > 0$,

$$\begin{aligned} & \textstyle\sum_{i=0}^{m_n} E(X_{n,i}^2 I(|X_{ni}| > \varepsilon) | \mathcal{F}_{n,i-1}) \\ = & \textstyle\sum_{i=0}^{m_n} EX_{n,i}^2 I(|X_{n,i}| > \varepsilon) \\ \leq & (2h_{n,1}^*/n) \textstyle\sum_{i=0}^{m_n} \rho_n^{2i} E((U_{-i}^2/\lambda_{n,1}) I(2h_{n,1}^* U_{-i}^2/(n\lambda_{n,1}) > \varepsilon^2)) \\ = & (2h_{n,1}^*/n)[\textstyle\sum_{i=0}^{m_n} \rho_n^{2i}] E((U_0^2/\lambda_{n,1}) I(2h_{n,1}^* U_0^2/\lambda_{n,1} > n\varepsilon^2)) \\ = & O(1)o(1), \end{aligned} \quad (4.22)$$

where the second equality holds because the U_{-i} have identical distributions. For the last equality, write $W_n = (U_0^2/\lambda_{n,1})$. For any $\nu > 0$, $W_n I((2h_{n,1}^* W_n/(n\varepsilon^2))^\nu > 1) \leq$

$W_n^{1+\nu}(2h_{n,1}^*/(n\varepsilon^2))^\nu$ and the result follows from Assumption INNOV which implies that $(2h_{n,1}^*/(n\varepsilon^2))^\nu \ EW_n^{1+\nu} = O(n^{-\nu})$. $\square$

Proof of Lemma 3. We decompose $M(r)$ into the sum of two independent Brownian motions, one of which is $W(r)$. (The decomposition is as in Guo and Phillips (2001) but with the added complication that $\phi_i^2 \neq \sigma_i^2$.) Let

$$\Sigma = A\Omega A' = \begin{bmatrix} h_{2,1} & 0 \\ 0 & h_{2,2}h_{2,3}^{-2} - h_{2,1}^{-1} \end{bmatrix}, \text{ where } A = \begin{bmatrix} 1 & 0 \\ -h_{2,1}^{-1} & h_{2,3}^{-1} \end{bmatrix}. \tag{4.23}$$

Hence, $W(r)$ and $W_2^0(r)$ are independent Brownian motions, where $W_2^0(r)$ is defined by

$$\begin{pmatrix} h_{2,1}^{1/2}W(r) \\ W_2^0(r) \end{pmatrix} = A \begin{pmatrix} h_{2,1}^{1/2}W(r) \\ M(r) \end{pmatrix} = \begin{pmatrix} h_{2,1}^{1/2}W(r) \\ h_{2,3}^{-1}M(r) - h_{2,1}^{-1/2}W(r) \end{pmatrix} = BM(\Sigma). \tag{4.24}$$

Let

$$W_2(r) = (h_{2,2}h_{2,3}^{-2} - h_{2,1}^{-1})^{-1/2}W_2^0(r). \tag{4.25}$$

As defined, W_2 is a standard univariate Brownian motion on $[0,1]$. We have

$$\begin{aligned} M(r) &= h_{2,3}\left(h_{2,1}^{-1/2}W(r) + W_2^0(r)\right) \\ &= h_{2,3}\left(h_{2,1}^{-1/2}W(r) + (h_{2,2}h_{2,3}^{-2} - h_{2,1}^{-1})^{1/2}W_2(r)\right) \\ &= h_{2,3}(h_{2,2}^{1/2}h_{2,3}^{-1})\left(h_{2,7}W(r) + (1-h_{2,7}^2)^{1/2}W_2(r)\right) \\ &= h_{2,2}^{1/2}\left(h_{2,7}W(r) + (1-h_{2,7}^2)^{1/2}W_2(r)\right). \end{aligned} \tag{4.26}$$

This concludes the proof. $\square$

Proof of Lemma 4. Parts (a) and (b) of the Lemma follow from Theorem 4.4 of Hansen (1992) when $X_{n,i}$ is defined with zero initial conditions and $\{v_{n,i} : i \leq n, n \geq 1\}$ is a sequence rather than a triangular array. Part (c) of the Lemma follows from a combination of Theorems 4.2 and 4.4 of Hansen (1992) under the same conditions as just stated. (Note that the same argument as in Hansen (1992) can be used when the random variables form a triangular array as when they form a sequence, given the conditions of the Lemma.) Hence, parts (a)-(c) of the Lemma hold when $h_1 = 0$.

When $h_1 > 0$, the first element of $X_{n,i}$ is based on a stationary initial condition. In this case, (4.1) applies with Y_i^* and U_i denoting the first element of $X_{n,i}$ and $v_{n,i}$, respectively. By the proof of Lemma 2, the result of Lemma 2 holds with Y_0^* denoting the first element of $X_{n,0}$ and with $\lambda_{n,1}$ replaced by $\Omega_{0,1,1}$. In consequence, we have

$$\begin{aligned} n^{-1/2}Y_{[nr]}^* &= n^{-1/2}\widetilde{Y}_{[nr]} + \exp(-h_{n,1}^*[nr]/n)(2h_n)^{-1/2}(2h_n/n)^{1/2}Y_0^* \\ &\Rightarrow \int_0^r \exp((r-s)h_1)dB_1(s) + (2h_1)^{-1/2}\exp(-h_1 r)\Omega_{0,1,1}^{1/2}Z_1, \end{aligned} \tag{4.27}$$

where $\widetilde{Y}_{[nr]} = \widetilde{Y}_{n,[nr]}$ and $h^*_{n,1}$ are defined as in the paragraph containing (4.1), $B_1(s)$ denotes the first element of $B(s)$, the first summand converges by Thm. 4.4 of Hansen (1992), the second summand converges by the result of Lemma 2 and the convergence of $\exp(-h^*_{n,1}[nr]/n)$ to $\exp(-h_1 r)$, which holds uniformly over $r \in [0,1]$, and the convergence of the two summands holds jointly. The limit random quantities $B_1(\cdot)$ and Z_1 are independent due to the strong-mixing assumption. In addition, the convergence of the first element of $X_{n,i}$, given in (4.27), holds jointly with the convergence of the remaining elements, whose weak limit is that stated in the Lemma by Thm. 4.4 of Hansen (1992). This concludes the proof of part (a) when $h_1 > 0$.

When $h_1 > 0$, the effect of the stationary initial condition of the first element of $X_{n,i}$ on the limit distribution in parts (b) and (c) of the Lemma is established in a similar way to that given above for part (a). $\square$

Proof of Lemma 5. Part (a) holds by applying Lemma 4(a) with $h_1 > 0$, $d_v = 1$, $v_{n,i} = U_i$, Ω_0 $(= \Omega_{0,1,1}) = h_{2,1}$, $K^*_h(r) = h^{1/2}_{2,1} I^*_h(r)$, and $\tau = 3$, using Assumptions INNOV and STAT.

Parts (b)-(d) hold by a weak law of large numbers for triangular arrays of $L^{1+\delta}$-bounded strong-mixing random variables for $\delta > 0$, e.g., see Andrews (1988), using the moment conditions in Assumption INNOV(iv).

The convergence in parts (e) and (g) holds by applying Lemma 4 with $h_1 > 0$, $d_v = 2$, $v_{n,i} = (U_i, U_i/\phi_i^2)'$, $\Omega_0 = \Omega$ (where Ω is defined in Lemma 3), $\Lambda = 0$ (because $\{(U_i, U_i/\phi_i^2) : i \leq n\}$ is a martingale difference array) and $\tau = 3$, using Assumptions INNOV and STAT. In particular, we use $\sup_{n\geq 1}[E|U_i|^\zeta + E|U_i/\phi_i^2|^\zeta] < \infty$ for some $\zeta > 3$ by Assumption INNOV(iv). Let $B(r) = (h^{1/2}_{2,1} W(r), M(r))' = BM(\Omega)$. Then, the first element of $K^*_h(r)$ can be written as $h^{1/2}_{2,1} I^*_h(r)$ and the second element of $B(r)$ equals $M(r)$. The convergence in part (e) holds by the convergence to $M(1)$ of the second element of $X_{n,[nr]}$ in Lemma 4(a) with $r = 1$. The convergence in part (g) holds by the convergence of the $(1,2)$ element of $n^{-1}\sum_{i=1}^n X_{n,i-1} v'_{n,i}$ in Lemma 4(b). The $(1,2)$ element of $\int K^*_h dB'$ equals $h^{1/2}_{2,1} \int I^*_h dM$. The last equalities of parts (e) and (g) hold by Lemma 3.

The equality in part (f) holds by applying Lemma 4(b) with $v_{n,i} = (U_i, \phi_i^{-2} - E\phi_1^{-2})$ and $\tau = 3$ because the appropriate element of this vector result gives $n^{-1}\sum_{i=1}^n Y^*_{i-1}(\phi_i^{-2} - E\phi_1^{-2}) = O_p(1)$. The equality in part (h) holds by applying Lemma 4(c) with $v_{n,i} = (U_i, \phi_i^{-2} - E\phi_1^{-2})$ and $\tau = 3$ because the appropriate element of this matrix result gives $n^{-3/2}\sum_{i=1}^n Y^{*2}_{i-1}(\phi_i^{-2} - E\phi_1^{-2}) = O_p(1)$. This result uses the assumption that $\sup_{n\geq 1} E\phi_1^{-2\zeta} < \infty$ for some $\zeta > 3$ in Assumption INNOV(iv). The equality in parts (i) and (j) holds by applying Lemma 4(b) and (c), respectively, with $v_{n,i} = (U_i, U_i^2/\phi_i^4 - E(U_1^2/\phi_1^4))$. This result uses the assumption that $\sup_{n\geq 1} E|U_i/\phi_i^2|^{2\zeta} < \infty$ for some $\zeta > 3$ in Assumption INNOV(iv).

The convergence in parts (f) and (h)-(j) holds by Assumption INNOV(vii) and by part (a) of the current Lemma combined with the continuous mapping theorem using standard arguments (e.g., see the proof of Lemma 1 of Phillips (1987)) which gives $n^{-3/2}\sum_{i=1}^n Y_{i-1}^* \to_d \int I_h^*$ and $n^{-2}\sum_{i=1}^n Y_{i-1}^{*2} \to_d \int I_h^{*2}$.

This concludes the proof of parts (a)-(j).

Part (k) holds because

$$|n^{-1-\ell_1/2}\textstyle\sum_{i=1}^n Y_{i-1}^{*\ell_1}U_i^{\ell_2}/\phi_i^4| \leq \sup_{i\leq n}|n^{-1/2}Y_{i-1}^*|^{\ell_1}\cdot n^{-1}\textstyle\sum_{i=1}^n |U_i|^{\ell_2}/\phi_i^4 = O_p(1), \quad (4.28)$$

where the equality holds by part (a) of the current Lemma combined with the continuous mapping theorem and the weak law of large numbers (referred to above).

Part (l) holds by the same argument as given above for parts (a)-(k), but without the extra detail needed to cover the case of a non-zero initial condition. □

Proof of Lemma 6. Using $Y_i^* = \sum_{j=0}^\infty \rho^j U_{i-j}$ and stationarity (within the rows of the triangular array),

$$\begin{aligned}
&E(Y_0^{*2}U_1^2/\phi_1^4)\\
&= \sum_{u=0}^\infty\sum_{v=0}^\infty \rho^{u+v}EU_{-u}U_{-v}U_1^2/\phi_1^4\\
&= \sum_{u=0}^\infty \rho^{2u}EU_{-u}^2U_1^2/\phi_1^4 + 2\sum_{u=0}^\infty\sum_{v=0}^{u-1}\rho^{u+v}EU_{-u}U_{-v}U_1^2/\phi_1^4\\
&= (1-\rho^2)^{-1}EU_1^2EU_1^2/\phi_1^4 + O(1). \qquad (4.29)
\end{aligned}$$

The last equality holds by the following argument. First, for $u > v$, we have $EU_{-u}U_{-v}U_1^2/\phi_1^4 = Cov(U_{-u}, U_{-v}U_1^2/\phi_1^4) = Cov(U_{-u}U_{-v}, U_1^2/\phi_1^4)$. This, α-mixing, (4.2), and Assumption INNOV(iv) give

$$\begin{aligned}
&EU_{-u}U_{-v}U_1^2/\phi_1^4\\
&= O(1)\max\{||U_{-u}U_{-v}||_\zeta||U_1^2/\phi_1^4||_\zeta, ||U_{-u}||_\zeta||U_{-v}U_1^2/\phi_1^4||_\zeta\}\times\\
&\quad {\alpha_n}^{1-2\zeta^{-1}}(\max\{(u-v),(1+v)\})\\
&= O((\max\{(u-v),(1+v)\})^{-3-\varepsilon}) \qquad (4.30)
\end{aligned}$$

for some $\varepsilon > 0$ because $\alpha_n^{1-2\zeta^{-1}}(m) = O(m^{-3\zeta(1-2\zeta^{-1})/(\zeta-3)}) = O(m^{-3-\varepsilon})$. Therefore,

$$\begin{aligned}
&\sum_{u=0}^\infty\sum_{v=0}^{u-1}\rho^{u+v}EU_{-u}U_{-v}U_1^2/\phi_1^4\\
&= O(1)\sum_{u=0}^\infty\sum_{v=0}^{u-1}\rho^{u+v}\min\{(u-v)^{-3-\varepsilon},(1+v)^{-3-\varepsilon}\}\\
&= O(1)\sum_{u=0}^\infty\sum_{v=0}^{u/2}(u-v)^{-3-\varepsilon} + O(1)\sum_{u=0}^\infty\sum_{v=u/2+1}^{u}v^{-3-\varepsilon}\\
&= O(1)\sum_{u=0}^\infty u^{-2-\varepsilon} + O(1)\sum_{u=0}^\infty u^{-2-\varepsilon}\\
&= O(1). \qquad (4.31)
\end{aligned}$$

Second,

$$\begin{aligned}\sum_{u=0}^{\infty}\rho^{2u}EU_{-u}^2U_1^2/\phi_1^4 &= \sum_{u=0}^{\infty}\rho^{2u}[Cov(U_{-u}^2,U_1^2/\phi_1^4)+EU_{-u}^2EU_1^2/\phi_1^4] \\ &= O(1)\sum_{u=0}^{\infty}u^{-3-\varepsilon}+\sum_{u=0}^{\infty}\rho^{2u}EU_1^2EU_1^2/\phi_1^4 \\ &= O(1)+(1-\rho^2)^{-1}EU_1^2EU_1^2/\phi_1^4. \qquad (4.32)\end{aligned}$$

The other statements in the Lemma are proven analogously. For example, for the last statement note that $E(Y_0^* U_1^2/\phi_1^4) = \sum_{u=0}^{\infty}\rho^u EU_{-u}U_1^2/\phi_1^4$ and $|EU_{-u}U_1^2/\phi_1^4| = O((u+1)^{-3-\varepsilon})$ give the desired result. □

Proof of Lemma 7. Using $Y_i^* = \sum_{j=0}^{\infty}\rho^j U_{i-j}$ and stationarity (within rows), we have

$$\begin{aligned}&E\left(\sum_{i=1}^{n}[E\zeta_i^2 - E(\zeta_i^2|\mathcal{G}_{i-1})]\right)^2 = \frac{E(\sum_{i=1}^n(Y_{i-1}^{*2}\sigma_i^2/\phi_i^4 - EY_0^{*2}U_1^2/\phi_1^4))^2}{n^2(E(Y_0^{*2}U_1^2/\phi_1^4))^2} \\ &= \frac{\sum_{i,j=1}^n Cov(Y_{i-1}^{*2}\sigma_i^2/\phi_i^4, Y_{j-1}^{*2}\sigma_j^2/\phi_j^4)}{n^2(E(Y_0^{*2}U_1^2/\phi_1^4))^2} = \frac{\sum_{i=1}^n(n-i+1)Cov(Y_{i-1}^{*2}\sigma_i^2/\phi_i^4, Y_0^{*2}\sigma_1^2/\phi_1^4)}{n^2(E(Y_0^{*2}U_1^2/\phi_1^4))^2} \\ &= \frac{\sum_{i=1}^n(n-i+1)\sum_{s,t=0}^{\infty}\rho^{s+t}\sum_{u,v=0}^{\infty}\rho^{u+v}Cov(U_{i-1-s}U_{i-1-t}U_i^2/\phi_i^4, U_{-u}U_{-v}U_1^2/\phi_1^4)}{n^2(E(Y_0^{*2}U_1^2/\phi_1^4))^2}.\end{aligned} \qquad (4.33)$$

The key portion of the proof is to bound the covariance term $C(i,s,t,u,v) = Cov(U_{i-1-s}U_{i-1-t}U_i^2/\phi_i^4, U_{-u}U_{-v}U_1^2/\phi_1^4)$ using strong mixing. However, it is not enough to use the strong-mixing inequality (4.2) in the case where $i-1-s$ and $i-1-t$ are both strictly positive and to exploit $C(i,s,t,u,v) = O((\max\{i-1-s, i-1-t\})^{-3\zeta(1-2/\zeta)/(\zeta-3)})$ in this case. The trick is to consider disjoint sets A and B such that $A\cup B = \{U_{i-1-s}, U_{i-1-t}, U_i^2/\phi_i^4, U_{-u}, U_{-v}, U_1^2/\phi_1^4\}$ and to note that

$$\begin{aligned}&|C(i,s,t,u,v)| \\ &\le |EU_{i-1-s}U_{i-1-t}(U_i^2/\phi_i^4)U_{-u}U_{-v}U_1^2/\phi_1^4| + |EU_{i-1-s}U_{i-1-t}U_i^2/\phi_i^4\cdot EU_{-u}U_{-v}U_1^2/\phi_1^4| \\ &\le |E\prod_{a\in A}aE\prod_{b\in B}b + Cov(\prod_{a\in A}a, \prod_{b\in B}b)| + |EU_{i-1-s}U_{i-1-t}U_i^2/\phi_i^4\cdot EU_{-u}U_{-v}U_1^2/\phi_1^4|.\end{aligned} \qquad (4.34)$$

Note that if $A\in\{\{U_{i-1-s}, U_{i-1-t}, U_i^2/\phi_i^4\}, \{U_{-u}, U_{-v}, U_1^2/\phi\}\}$ then the simpler bound $|C(i,s,t,u,v)| \le |Cov(\prod_{a\in A}a, \prod_{b\in B}b)|$ applies. We will pick the partition $A\cup B$ such that $E\prod_{a\in A}a\cdot E\prod_{b\in B}b = 0$ and then apply the strong-mixing inequality (4.2) to bound $Cov(\prod_{a\in A}a, \prod_{b\in B}b)$ and also $|EU_{i-1-s}U_{i-1-t}U_i^2/\phi_i^4\cdot EU_{-u}U_{-v}U_1^2/\phi_1^4|$. In fact, $E\prod_{a\in A}a\cdot E\prod_{b\in B}b = 0$ holds true for any partition, unless 1 is the largest subindex in one group A or B.

First we show that we can assume that all the subindices $i-1-s$, $i-1-t$, i, $-u$, $-v$, 1 that appear in the covariance expression (4.33) are different because the sum of all summands, where at least two of these subindices are equal, is of order $o(1)$. To see this, consider first the case where there is more than one pair of subindices that coincides, e.g. when $i-1-s=i-1-t=-u$ or when $i-1-t=1$ and $-u=-v$. For example, assume $i-1-s=-u$ and $i-1-t=1$ (the other cases are proven analogously). Then $i=-u+s+1=t+2$ and the numerator in (4.33) equals

$$O(1)\sum_{i=1}^{n}(n-i+1)\sum_{s,t=0}^{\infty}\rho^{s+t}\sum_{u,v=0}^{\infty}\rho^{u+v}=O(n)\sum_{u,v=0}^{\infty}\rho^{u+v}\sum_{s=0}^{\infty}\rho^{s}\rho^{-u+s-1}=O(n(1-\rho)^{-3}). \tag{4.35}$$

Because $E(Y_0^{*2}U_1^2/\phi_1^4)$ is of order $(1-\rho)^{-1}$ by Lemma 6 and $n(1-\rho)\to\infty$, the result follows. We can therefore assume there is *exactly* one pair of subindices that coincides, for example, $i-1-s=1$. (The other cases are proven analogously.) Then the numerator in (4.33) is bounded by

$$2n\sum_{u=0}^{\infty}\sum_{v=0}^{u-1}\rho^{u+v}\sum_{t=0}^{\infty}\sum_{s=0}^{n-2}\rho^{s+t}|Cov(U_1U_{s+1-t}U_{s+2}^2/\phi_{s+2}^4,U_{-u}U_{-v}U_1^2/\phi_1^4)|, \tag{4.36}$$

where the summations are such that all subindices $1, s+1-t, -u, -v$ are different. There are four cases to consider: (i) $1<s+1-t$, (ii) $-v<s+1-t<1$, (iii) $-u<s+1-t<-v$, and (iv) $s+1-t<-u$. In case (i), we use (4.34) with $A=\{U_{-u},U_{-v}\}$ and $B=\{U_1,U_{s+1-t},U_{s+2}^2/\phi_{s+2}^4,U_1^2/\phi_1^4\}$. This leads to

$$\begin{aligned}
&|C(s+2,s,t,u,v)|\\
&\le |Cov(\prod_{a\in A}a,\prod_{b\in B}b)|+|EU_1U_{s+1-t}U_{s+2}^2/\phi_{s+2}^4|\cdot|EU_{-u}U_{-v}U_1^2/\phi_1^4|\\
&\le (v+1)^{-3\zeta(1-2\zeta^{-1})/(\zeta-3)}+(\max\{v+1,u-v\})^{-3\zeta(1-2\zeta^{-1})/(\zeta-3)}\\
&\le (v+1)^{-3-\varepsilon}+(\max\{v+1,u-v\})^{-3-\varepsilon}
\end{aligned} \tag{4.37}$$

for some $\varepsilon>0$, where in the second to last inequality we use Assumption INNOV(iv) and (4.2) and apply an argument analogous to (4.34) to the expectation $EU_{-u}U_{-v}U_1^2/\phi_1^4$, namely, $EU_{-u}U_{-v}U_1^2/\phi_1^4=Cov(U_{-u},U_{-v}U_1^2/\phi_1^4)=Cov(U_{-u}U_{-v},$ $U_1^2/\phi_1^4)$. In the last inequality we use the fact that $-3\zeta(1-2\zeta^{-1})/(\zeta-3)<-3-\varepsilon$ for some $\varepsilon>0$. Picking $A=\{U_{-u}\}$ and $B=\{U_{-v},U_1,U_{s+1-t},U_{s+2}^2/\phi_{s+2}^4,U_1^2/\phi_1^4\}$ the same argument can be used to show that $|C(s+2,s,t,u,v)|\le(u-v)^{-3-\varepsilon}+$ $(\max\{v+1,u-v\})^{-3-\varepsilon}$. Therefore, $|C(s+2,s,t,u,v)|\le 2(\max\{v+1,u-v\})^{-3-\varepsilon}$. Thus, the summands in (4.36) over case (i) are bounded by

$$\begin{aligned}
&4n\sum_{u=0}^{\infty}\sum_{v=0}^{u-1}\rho^{u+v}\sum_{t=0}^{\infty}\sum_{s=t+1}^{n-2}\rho^{s+t}(\max\{v+1,u-v\})^{-3-\varepsilon}\\
&\le O(n)\sum_{u=0}^{\infty}\sum_{v=0}^{u-1}(\max\{v+1,u-v\})^{-3-\varepsilon}\sum_{t=0}^{\infty}\sum_{s=t+1}^{n-2}\rho^{s+t}
\end{aligned}$$

$$
\begin{aligned}
&\leq O(n(1-\rho)^{-2}) \sum_{u=0}^{\infty} \left(\sum_{v=0}^{[u/2]} (u/2)^{-3-\varepsilon} + \sum_{v=[u/2]}^{u-1} (u/2)^{-3-\varepsilon} \right) \\
&= O(n(1-\rho)^{-2}) \sum_{u=0}^{\infty} (u/2)^{-2-\varepsilon} \\
&= O(n(1-\rho)^{-2}). \qquad (4.38)
\end{aligned}
$$

Because by Lemma 6 the denominator is of order $n^2(1-\rho)^{-2}$ the result follows. Cases (ii)–(iv) are handled analogously.

From now on, we can therefore assume that all the subindices $i-1-s$, $i-1-t$, i, $-u$, $-v$, 1 that appear in the covariance expression in (4.33) are different. From now on, all summations are subject to this restriction without explicitly stating it.

We now show that the second summand in (4.34), i.e., $|EU_{i-1-s}U_{i-1-t}U_i^2/\phi_i^4 \times EU_{-u}U_{-v}U_1^2/\phi_1^4|$, is negligible when substituted into (4.33). Note that $EU_{-u}U_{-v}U_1^2/\phi_1^4 = Cov(U_{-u}, U_{-v}U_1^2/\phi_1^4) = Cov(U_{-v}, U_{-u}U_1^2/\phi_1^4)$ whenever $u \neq v$. Therefore, for some $\varepsilon > 0$, (4.2) and Assumption INNOV(iv) yield

$$
\begin{aligned}
EU_{-u}U_{-v}U_1^2/\phi_1^4 &= O(\max\{|v-u|, 1+v, 1+u\}^{-3\zeta(1-2\zeta^{-1})/(\zeta-3)}) \\
&= O(\max\{|v-u|, 1+v, 1+u\}^{-3-\varepsilon}) \qquad (4.39)
\end{aligned}
$$

if $u \neq v$ and likewise for the term $EU_{i-1-s}U_{i-1-t}U_i^2/\phi_i^4$. Therefore, in the numerator of (4.33), the contribution of the second summand of (4.34) is

$$
\begin{aligned}
&\sum_{i=1}^{n}(n-i+1) \sum_{s,t=0}^{\infty} \rho^{s+t} \sum_{u,v=0}^{\infty} \rho^{u+v} |EU_{i-1-s}U_{i-1-t}U_i^2/\phi_i^4 \cdot EU_{-u}U_{-v}U_1^2/\phi_1^4| \\
&= O(n^2) \sum_{s,t=0}^{\infty} \sum_{u,v=0}^{\infty} \max\{|v-u|, 1+v, 1+u\}^{-3-\varepsilon} \max\{|s-t|, s+1, t+1\}^{-3-\varepsilon} \\
&= O(n^2) \left(\sum_{u,v=0}^{\infty} \max\{|v-u|, 1+v, 1+u\}^{-3-\varepsilon} \right)^2. \qquad (4.40)
\end{aligned}
$$

By symmetry in u, v, the latter equals

$$
\begin{aligned}
&O(n^2) \left(\sum_{u=0}^{\infty} \sum_{v=0}^{u-1} \max\{u-v, 1+u\}^{-3-\varepsilon} \right)^2 \\
&= O(n^2) \left(\sum_{u=0}^{\infty} \sum_{v=0}^{[u/2]} (u/2)^{-3-\varepsilon} + \sum_{u=0}^{\infty} \sum_{v=[u/2]+1}^{u} (u/2)^{-3-\varepsilon} \right)^2 \\
&= O(n^2) \sum_{u=0}^{\infty} (u/2)^{-2-\varepsilon} \\
&= O(n^2). \qquad (4.41)
\end{aligned}
$$

Because the denominator $n^2(E(Y_0^{*2}U_1^2/\phi_1^4))^2$ in (4.33) is of order $n^2(1-\rho)^{-2}$ by Lemma 6, we have shown that the summands $|EU_{i-1-s}U_{i-1-t}U_i^2/\phi_i^4 \cdot EU_{-u}U_{-v}U_1^2/\phi_1^4|$ in (4.33) are negligible.

We are now left to show that the sum of all summands in the last line of (4.33) is $o(1)$ when all the subindices $i-1-s$, $i-1-t$, i, $-u$, $-v$, 1 that appear in the covariance expression (4.33) are different. We can assume $u > v$ and $s > t$. We can also impose the bound $|C(i,s,t,u,v)| \leq |E\prod_{a\in A} aE\prod_{b\in B} b + Cov(\prod_{a\in A} a, \prod_{b\in B} b)|$ because we have shown that the contributions of the last summand in (4.34) are negligible. We only consider partitions A and B where 1 is not the largest subindex in any of the two sets A or B in which case we have $|C(i,s,t,u,v)| \leq |Cov(\prod_{a\in A} a, \prod_{b\in B} b)|$. There are ten different cases to consider regarding the order of $i-1-s$ and $i-1-t$ relative to $1, -u$, and $-v$. In case (1) $i-1-s > 1$ (which implies $i-1-t > 1$ because we assume $s > t$), (2) $1 > i-1-s > -v$ (which implies $i-1-s > -u$ because $u > v$) and $i-1-t > 1$, (3) $-v > i-1-s > -u$ and $i-1-t > 1$, (4) $-u > i-1-s$ and $i-1-t > 1$, (5) $-v < i-1-s < 1$ and $-v < i-1-t < 1$, (6) $-u < i-1-s < -v$ and $-v < i-1-t < 1$, (7) $-u < i-1-s$ and $-v < i-1-t < 1$, (8) $-u < i-1-s < -v$ and $-v < i-1-t < -u$, (9) $-u < i-1-s$ and $-u < i-1-t < -u$, and (10) $-u < i-1-s$ and $-u < i-1-t < -u$. We will only deal with the two cases (1) and (2), the other cases can be handled analogously.

Case (1). Consider the partitions A and B of $\{U_{-u}, U_{-v}, U_1^2/\phi_1^4, U_{i-1-s}, U_{i-1-t}, U_i^2/\phi_i^4\}$, where $A = \{U_{-u}\}$, $A = \{U_{-u}, U_{-v}\}$, and $A = \{U_{-u}, U_{-v}, U_1^2/\phi_1^4, U_{i-1-s}, U_{i-1-t}\}$. The strong-mixing covariance inequality implies that

$$|C(i,s,t,u,v)| \leq |Cov(\textstyle\prod_{a\in A} a, \prod_{b\in B} b)| \leq (\max\{u-v, v+1, t+1\})^{-3-\varepsilon}. \quad (4.42)$$

Therefore,

$$\frac{\sum_{i=1}^n (n-i+1)\sum_{s>t=0}^\infty \rho^{s+t}\sum_{u>v=0}^\infty \rho^{u+v}|C(i,s,t,u,v)|}{n^2(E(Y_0^{*2}U_1^2/\phi_1^4))^2}$$
$$= O(1-\rho)\sum_{t=0}^\infty \sum_{u>v=0}^\infty (\max\{u-v, v+1, t+1\})^{-3-\varepsilon}, \quad (4.43)$$

where we use $\sum_{s=0}^\infty \rho^s = (1-\rho)^{-1}$, (4.42), and Lemma 6. We now consider three subcases 1(i) $t+1 > u-v$ and $t+1 > v+1$, 1(ii) $u-v > t+1$ and $u-v > v+1$, 1(iii) $v+1 > t+1$ and $v+1 > u-v$. In case 1(i), the sum over s,t,u,v in (4.43) can be bounded by

$$\sum_{t=0}^\infty \sum_{v=0}^{t-1} \sum_{u=v}^{t+1+v} (t+1)^{-3-\varepsilon} \leq \sum_{t=0}^\infty \sum_{v=0}^{t-1} (t+1)^{-2-\varepsilon} = \sum_{t=0}^\infty (t+1)^{-1-\varepsilon} = O(1). \quad (4.44)$$

In case 1(ii), the sum over s,t,u,v in (4.43) can be bounded by

$$\sum_{u=1}^\infty \sum_{v=0}^{[u/2]} \sum_{t=0}^{u-v-1} (u-v)^{-3-\varepsilon} \leq \sum_{u=1}^\infty \sum_{v=0}^{[u/2]} (u-v)^{-2-\varepsilon} \leq \sum_{u=1}^\infty (u/2)^{-1-\varepsilon} = O(1). \quad (4.45)$$

In case 1(iii), the sum over s, t, u, v in (4.43) can be bounded by

$$\sum_{u=1}^{\infty}\sum_{t=0}^{u-1}\sum_{v=\max(t+1,[(u-1)/2])}^{u-1}(v+1)^{-3-\varepsilon} \leq \sum_{u=1}^{\infty}\sum_{t=0}^{u-1}\sum_{v=[(u-1)/2]}^{u-1}(v+1)^{-3-\varepsilon} \leq \sum_{u=1}^{\infty}\sum_{t=0}^{u-1}(u/2)^{-2-\varepsilon} = O(1). \quad (4.46)$$

This proves case (1). We next deal with case (2).

Case (2). Consider the partitions A and B of $\{U_{-u}, U_{-v}, U_1^2/\phi_1^4, U_{i-1-s}, U_{i-1-t}, U_i^2/\phi_i^4\}$, where $A = \{U_{-u}, U_{-v}\}$, $A = \{U_{-u}, U_{-v}, U_{i-1-s}\}$, or $A = \{U_{-u}, U_{-v}, U_{i-1-s}, U_1^2/\phi_1^4, U_{i-1-t}\}$. The strong-mixing covariance inequality implies that

$$|C(i,s,t,u,v)| \leq |Cov(\textstyle\prod_{a\in A} a, \prod_{b\in B} b)| \leq (\max\{i-1-s+v, 2-i+s, t+1\})^{-3-\varepsilon}. \quad (4.47)$$

We consider several subcases. In case 2(i) suppose that $i-1-s+v < t+1$. Then,

$$\frac{\sum_{i=1}^{n}(n-i+1)\sum_{s>t=0}^{\infty}\rho^{s+t}\sum_{u>v=0}^{\infty}\rho^{u+v}|C(i,s,t,u,v)|}{n^2(E(Y_0^{*2}U_1^2/\phi_1^4))^2} = O\left(n^{-1}(1-\rho)^2\sum_{u>v=0}^{\infty}\rho^{u+v}\sum_{s>t=0}^{\infty}\rho^{s+t}\sum_{i=s-v+1}^{t+s-v+2}(t+1)^{-3+\varepsilon}\right), \quad (4.48)$$

where the restrictions on the summation over i result from $i-1-s > -v$ and $i-1-s+v < t+1$. The expression in (4.48) is of order $O(n^{-1}\sum_{s=0}^{\infty}\rho^s\sum_{t=0}^{\infty}(t+1)^{-2+\varepsilon})$ because of Lemma 6 and $t+s-v+2-(s-v+1) = t+1$. But the latter expression is $o(1)$ because $n(1-\rho) \to \infty$ and $\sum_{t=0}^{\infty}(t+1)^{-2+\varepsilon} = O(1)$.

In case 2(ii) suppose that $t+1 > 2-i+s$. Therefore,

$$\frac{\sum_{i=1}^{n}(n-i+1)\sum_{s>t=0}^{\infty}\rho^{s+t}\sum_{u>v=0}^{\infty}\rho^{u+v}|C(i,s,t,u,v)|}{n^2(E(Y_0^{*2}U_1^2/\phi_1^4))^2} = O(n^{-1}(1-\rho)^{-2}\sum_{u>v=0}^{\infty}\rho^{u+v}\sum_{s>t=0}^{\infty}\rho^{s+t}\sum_{i=s-t+1}^{s+2}(t+1)^{-3+\varepsilon}), \quad (4.49)$$

where the restrictions on the summation over i result from $t+1 > 2-i+s$ and $1 > i-1-s$. The expression in (4.49) is of order $O(n^{-1}\sum_{s=0}^{\infty}\rho^s\sum_{t=0}^{\infty}(t+1)^{-2+\varepsilon})$. The latter expression is $o(1)$ as in case 2(i).

Finally consider the case 2(iii) where $i-1-s+v > t+1$ and $t+1 < 2-i+s$. Assume first that $i-1-s+v < 2-i+s$. This implies that $i < -v/2+s+3/2$. Therefore,

$$\frac{\sum_{i=1}^{n}(n-i+1)\sum_{s>t=0}^{\infty}\rho^{s+t}\sum_{u>v=0}^{\infty}\rho^{u+v}|C(i,s,t,u,v)|}{n^2(E(Y_0^{*2}U_1^2/\phi_1^4))^2} = O(n^{-1}(1-\rho)^2\sum_{s>t=0}^{\infty}\rho^{s+t}\sum_{u>v=0}^{\infty}\rho^{u+v}\sum_{i=-v+s-1}^{[-v/2+s+3/2]}(2-i+s)^{-3+\varepsilon}). \quad (4.50)$$

The expression in (4.50) is of order $O(n^{-1}\sum_{u=0}^{\infty}\rho^u\sum_{v=0}^{\infty}(v/2)(v/2)^{-3+\varepsilon})$. The latter expression is $o(1)$ as in case 2(i).

The subcase $i-1-s+v \geq 2-i+s$ of case 2(iii) can be handled using the same steps. That completes the proof of case (2). □

Proof of Lemma 8. To prove (a), by Markov's inequality it is enough to show that $n^{-2}(1-\rho)E(X_1'X_2)^2 = o(1)$. Note that

$$E(X_1'X_2)^2 = \sum_{i,k=1}^{n}\sum_{j,l=0}^{\infty}\rho^{j+l}EU_{i-1-j}\phi_i^{-2}U_{k-1-l}\phi_k^{-2}. \tag{4.51}$$

The contribution of the summands where $i=k$ is $\sum_{i=1}^{n}\sum_{j,l=0}^{\infty}\rho^{j+l}EU_{i-1-j}\phi_i^{-4}U_{i-1-l}$ which is of order $O(n(1-\rho)^{-2})$ and thus negligible because $n(1-\rho)\to\infty$. It is therefore enough to study the sum $\sum_{i>k=1}^{n}\sum_{j,l=0}^{\infty}\rho^{j+l}EU_{i-1-j}\phi_i^{-2}U_{k-1-l}\phi_k^{-2}$. We have to consider several subcases, namely, (1) $i-1-j<k-1-l$, (2) $k-1-l\leq i-1-j<k$, and (3) $k\leq i-1-j$. In case (1), the sum in (4.51) can be bounded by

$$\sum_{j,l=0}^{\infty}\rho^{j+l}\sum_{k=1}^{n}\sum_{i=k+1}^{k+j-l}\max\{l+1,k-l-i+j\}^{-3-\varepsilon} \tag{4.52}$$

noting that $EU_{i-1-j}\phi_i^{-2}U_{k-1-l}\phi_k^{-2} = Cov(U_{i-1-j},U_{k-1-l}\phi_i^{-2}\phi_k^{-2}) = Cov(U_{i-1-j}U_{k-1-l},$ $\phi_i^{-2}\phi_k^{-2})$ and using (4.2) and Assumption INNOV(iv). The sum in (4.52) can be bounded by

$$\begin{aligned}
&\sum_{j,l=0}^{\infty}\rho^{j+l}\sum_{k=1}^{n}\left[\sum_{i=k+1}^{k-2l+j-1}(k-l-i+j)^{-3-\varepsilon}+\sum_{i=k-2l+j}^{k+j-l}(l+1)^{-3-\varepsilon}\right]\\
&\leq\sum_{j,l=0}^{\infty}\rho^{j+l}\sum_{k=1}^{n}[l^{-3-\varepsilon}\max\{j-2l,0\}+(l+1)^{-2-\varepsilon}]\\
&=O(n(1-\rho)^{-2}+n(1-\rho)^{-1}),
\end{aligned} \tag{4.53}$$

where the last equality holds because

$$\sum_{j,l=0}^{\infty}\rho^{j+l}\sum_{k=1}^{n}l^{-3-\varepsilon}\max\{j-2l,0\}\leq n\sum_{l=0}^{\infty}\rho^l l^{-3-\varepsilon}\sum_{j=2l}^{\infty}\rho^j(j-2l)=O(n)\sum_{j=0}^{\infty}\rho^j j \tag{4.54}$$

and $\sum_{j=0}^{\infty}\rho^j j=\rho(1-\rho)^{-2}$. This proves case (1). Cases (2) and (3) can be proved analogously.

Next, we prove part (b) of the Lemma. It is enough to show that

$$E\left(\frac{n^{-1}X_1'X_1-E(Y_0^{*2}/\phi_1^2)}{E(Y_0^{*2}/\phi_1^2)}\right)^2\to 0. \tag{4.55}$$

By Lemma 6, this holds if

$$((1-\rho)^2/n^2)\sum_{i,j=1}^{n} Cov(Y_{i-1}^{*2}/\phi_i^2, Y_{j-1}^{*2}/\phi_j^2) = o(1). \tag{4.56}$$

The latter can be established using the same approach as was used in (4.33) to establish that $(E(Y_0^{*2}U_1^2/\phi_1^4))^{-2}\ n^{-2}\sum_{i,j=1}^{n} Cov(Y_{i-1}^{*2}\sigma_i^2/\phi_i^4,\ Y_{j-1}^{*2}\sigma_j^2/\phi_j^4) = o(1)$.

We can show part (c) by proceeding as in part (b).

Next, we prove part (d) of the Lemma. Note that

$$\begin{aligned}
&(n^{-1}X'X)^{-1}n^{-1}X'U = \det{}^{-1}(T_1, T_2)', \text{ where}\\
\det &= n^{-1}\sum_{i=1}^{n}(Y_{i-1}^{*2}/\phi_i^2)n^{-1}\sum_{i=1}^{n}\phi_i^{-2} - \left(n^{-1}\sum_{i=1}^{n}Y_{i-1}^{*}/\phi_i^2\right)^2, \qquad (4.57)\\
T_1 &= \left(n^{-1}\sum_{i=1}^{n}\phi_i^{-2}\right)\left(n^{-1}\sum_{i=1}^{n}Y_{i-1}^{*}U_i/\phi_i^2\right) - \left(n^{-1}\sum_{i=1}^{n}Y_{i-1}^{*}/\phi_i^2\right)n^{-1}\sum_{i=1}^{n}U_i/\phi_i^2, \text{ and}\\
T_2 &= -n^{-1}\sum_{i=1}^{n}(Y_{i-1}^{*}/\phi_i^2)\left(n^{-1}\sum_{i=1}^{n}Y_{i-1}^{*}U_i/\phi_i^2\right) + n^{-1}\sum_{i=1}^{n}(Y_{i-1}^{*2}/\phi_i^2)n^{-1}\sum_{i=1}^{n}U_i/\phi_i^2.
\end{aligned}$$

Using parts (a) and (b) of the Lemma, (4.15), $n^{-1}\sum_{i=1}^{n}U_i/\phi_i^2 = O_p(n^{-1/2})$, and Lemma 6, it follows that $\det^{-1} = O_p(1-\rho)$, $T_1 = O_p((n(1-\rho))^{-1/2})$, and $T_2 = O_p((1-\rho)^{-1}n^{-1/2})$, which proves the claim.

Next we prove part (e). Note that since $\Delta = Diag(M_XY) = Diag(M_XU)$ we have

$$X_1'\Delta^2X_1 = \sum_{i=1}^{n}(Y_{i-1}^{*2}/\phi_i^2)\{U_i/\phi_i - (Y_{i-1}^{*}/\phi_i, \phi_i^{-1})(X'X)^{-1}X'U\}^2. \tag{4.58}$$

By part (c), we are left to show that

$$\begin{aligned}
(E(Y_0^{*2}U_1^2/\phi_1^4))^{-1}n^{-1}\sum_{i=1}^{n}(Y_{i-1}^{*2}/\phi_i^2)(U_i/\phi_i)(Y_{i-1}^{*}/\phi_i, \phi_i^{-1})(X'X)^{-1}X'U \to_p 0 \text{ and}\\
(E(Y_0^{*2}U_1^2/\phi_1^4))^{-1}n^{-1}\sum_{i=1}^{n}(Y_{i-1}^{*2}/\phi_i^2)[(Y_{i-1}^{*}/\phi_i, \phi_i^{-1})(X'X)^{-1}X'U]^2 \to_p 0.
\end{aligned} \tag{4.59}$$

Part (d) and Lemma 6 imply that it is sufficient to show that

$$O_p((1-\rho)n^{-1})\sum_{i=1}^{n}(Y_{i-1}^{*2}U_i/\phi_i^4)O_p(n^{-1/2}) = o_p(1),$$

$$O_p((1-\rho)n^{-1})\sum_{i=1}^n (Y_{i-1}^{*2}/\phi_i^4) O_p(n^{-1}) = o_p(1),$$

$$O_p((1-\rho)n^{-1})\sum_{i=1}^n (Y_{i-1}^{*3}U_i/\phi_i^4) O_p((1-\rho)^{1/2}n^{-1/2}) = o_p(1),$$

$$O_p((1-\rho)n^{-1})\sum_{i=1}^n (Y_{i-1}^{*3}/\phi_i^4) O_p((1-\rho)^{1/2}n^{-1}) = o_p(1), \text{ and}$$

$$O_p((1-\rho)n^{-1})\sum_{i=1}^n (Y_{i-1}^{*4}/\phi_i^4) O_p((1-\rho)n^{-1}) = o_p(1). \quad (4.60)$$

The first and second conditions follow by proofs as for parts (c) and (b), respectively. The other conditions can be proven along the same lines as above. For example, one can establish that

$$(1-\rho)^{3/2}n^{-2}\sum_{i=1}^n (Y_{i-1}^{*3}/\phi_i^4) = o_p(1) \quad (4.61)$$

by using Markov's inequality and methods as in Lemma 7.

Finally, for the proofs of parts (f) and (g) note that

$$X_1'\Delta^2 X_2 = \sum_{i=1}^n (Y_{i-1}^*/\phi_i^2)[U_i/\phi_i - (Y_{i-1}^*/\phi_i, \phi_i^{-1})(X'X)^{-1}X'U]^2 \text{ and}$$

$$X_2'\Delta^2 X_2 = \sum_{i=1}^n \phi_i^{-2}[U_i/\phi_i - (Y_{i-1}^*/\phi_i, \phi_i^{-1})(X'X)^{-1}X'U]^2. \quad (4.62)$$

Therefore the desired results are implied by showing that

$$(1-\rho)^{1/2}n^{-1}\sum_{i=1}^n (Y_{i-1}^*/\phi_i^2)(U_i^2/\phi_i^2) = O_p(1),$$

$$(1-\rho)^{1/2}n^{-1}\sum_{i=1}^n (Y_{i-1}^*/\phi_i^2)(U_i/\phi_i)(Y_{i-1}^*/\phi_i, \phi_i^{-1})(X'X)^{-1}X'U = O_p(1),$$

$$(1-\rho)^{1/2}n^{-1}\sum_{i=1}^n (Y_{i-1}^*/\phi_i^2)[(Y_{i-1}^*/\phi_i, \phi_i^{-1})(X'X)^{-1}X'U]^2 = O_p(1), \quad (4.63)$$

and

$$n^{-1}\sum_{i=1}^n \phi_i^{-2}(U_i^2/\phi_i^2) = O_p(1),$$

$$n^{-1}\sum_{i=1}^n \phi_i^{-2}(U_i/\phi_i)(Y_{i-1}^*/\phi_i, \phi_i^{-1})(X'X)^{-1}X'U = O_p(1),$$

$$n^{-1}\sum_{i=1}^n \phi_i^{-2}((Y_{i-1}^*/\phi_i, \phi_i^{-1})(X'X)^{-1}X'U)^2 = O_p(1). \quad (4.64)$$

All of the statements in (4.63) and (4.64) follow from earlier parts of the Lemma or by arguments used in earlier parts of the Lemma. For example, $(1-\rho)^{1/2}n^{-1}\sum_{i=1}^{n}(Y_{i-1}^{*}U_i^2/\phi_i^4) = O_p(1)$ and $(1-\rho)^{1/2}n^{-1}\sum_{i=1}^{n}(Y_{i-1}^{*}U_i/\phi_i^4) = O_p(1)$ are proven as part (a) of the Lemma. To show $(1-\rho)n^{-3/2}\sum_{i=1}^{n}(Y_{i-1}^{*2}U_i/\phi_i^4) = O_p(1)$, one can use a proof as for part (c). □

Proof of Lemma 9. It is enough to show that $\sum_{i=1}^{n} E(\zeta_i^2 1(|\zeta_i| > \delta)) \to 0$ for any $\delta > 0$. We have

$$\sum_{i=1}^{n} E(\zeta_i^2 1(|\zeta_i| > \delta)) \le \delta^{-2}\sum_{i=1}^{n} E(\zeta_i^4) = n\delta^{-2}E(\zeta_1^4) = O(n^{-1}(1-\rho)^2)E(Y_0^{*}U_1/\phi_1^2)^4, \tag{4.65}$$

by stationarity (within rows) and Lemma 6. Furthermore,

$$E(Y_{i-1}^{*}U_i/\phi_i^2)^4 = \sum_{u,v,s,t=0}^{\infty} \rho^{u+v+s+t} EU_{-u}U_{-v}U_{-s}U_{-t}U_1^4/\phi_1^8. \tag{4.66}$$

The contributions of all summands for which at least two of the indices u, v, s, t are the same is $o(n(1-\rho)^{-2})$. For example, suppose $u = v$. Note that $\sum_{u,s,t=0}^{\infty} \rho^{2u+s+t} EU_{-u}^2 U_{-s}U_{-t}U_1^4/\phi_1^8 = O((1-\rho)^{-3})$ which is indeed $o(n(1-\rho)^{-2})$ because $n(1-\rho) \to \infty$. We can therefore restrict attention in the sum in (4.66) to the case where all indices are different and by symmetry, we can even restrict summation to the cases where $v = \min\{u, t, s, v\}$. Using the strong-mixing inequality as above, we have

$$E(Y_{i-1}^{*}U_i/\phi_i^2)^4 \le O(\sum_{u,t,s} \rho^{u+s+t} \sum_{v}(v-1)^{-3-\varepsilon}) = O((1-\rho)^{-3}), \tag{4.67}$$

which is $o(n(1-\rho)^{-2})$ as shown above. □

4.2 Proof of Theorem 2

Proof of Theorem 2. Suppose $h_1 < \infty$. Inspection of the proof of Theorem 1 shows that is suffices to show that Lemma 5 holds with $\widehat{\phi}_i$ in place of ϕ_i. The difference between the lhs quantity in Lemma 5(b) with $j = 1$ and the corresponding quantity with $\widehat{\phi}_i$ in place of ϕ_i is $o_p(1)$ by Assumption CHE(ii)(b) with $(d, j) = (0, 1)$. The same result holds for $j = 2$ because

$$\begin{aligned} &|n^{-1}\textstyle\sum_{i=1}^{n} \widehat{\phi}_i^{-2} - \phi_i^{-2}| \\ \le\ & n^{-1}\textstyle\sum_{i=1}^{n} \widehat{\phi}_i^{-1}|\widehat{\phi}_i^{-1} - \phi_i^{-1}| + n^{-1}\textstyle\sum_{i=1}^{n} \phi_i^{-1}|\widehat{\phi}_i^{-1} - \phi_i^{-1}| \\ \le\ & 2\varepsilon^{-1/2}n^{-1}\textstyle\sum_{i=1}^{n} |\widehat{\phi}_i^{-1} - \phi_i^{-1}| = o_p(1), \end{aligned} \tag{4.68}$$

where the first inequality holds by the triangle inequality, the second inequality holds by Assumption CHE(i), and the equality holds by Assumption CHE(ii)(b) with $(d, j) = (0, 1)$. For $j = 4$, the same result holds by the same argument as just given with 4 in place of 2 in the first line and 2 in place of 1 in the second and third lines.

The differences between the lhs quantities in Lemma 5(c) and (d) and the corresponding quantities with $\widehat{\phi}_i$ in place of ϕ_i are $o_p(1)$ by the same argument as in (4.68) (with 4 in place of 2 in the first line and 2 in place of 1 in the second and third lines) using Assumption CHE(ii)(b) with $(d, j) = (1, 2)$ and $(2, 2)$, respectively.

The differences between the lhs quantities in Lemma 5(e) and (g) and the corresponding quantities with $\widehat{\phi}_i$ in place of ϕ_i are $o_p(1)$ by Assumption CHE(ii)(a) with $j = 0$ and $j = 1$, respectively.

The difference between the lhs quantity in Lemma 5(f) and the corresponding quantity with $\widehat{\phi}_i$ in place of ϕ_i is $o_p(1)$ because

$$\begin{aligned} & |n^{-3/2} \textstyle\sum_{i=1}^n Y_{i-1}^* (\widehat{\phi}_i^{-2} - \phi_i^{-2})| \\ \leq & \sup_{i \leq n, n \geq 1} |n^{-1/2} Y_{i-1}^*| \cdot n^{-1} \textstyle\sum_{i=1}^n |\widehat{\phi}_i^{-2} - \phi_i^{-2}| = o_p(1), \end{aligned} \tag{4.69}$$

where the equality holds by (4.68) and $\sup_{i \leq n, n \geq 1} |n^{-1/2} Y_{i-1}^*| = O_p(1)$, which holds by Lemma 5(a) and the continuous mapping theorem. Analogous results hold for Lemma 5(h)-(j) using Assumption CHE(ii)(b) with $(d, j) = (2, 2)$ for parts (i) and (j).

Next, we show that the lhs quantity in Lemma 5(k) with $\widehat{\phi}_i$ in place of ϕ_i is $o_p(n)$. We have

$$\begin{aligned} & |n^{-1-\ell_1/2} \textstyle\sum_{i=1}^n Y_{i-1}^{*\ell_1} U_i^{\ell_2} / \widehat{\phi}_i^4| \\ \leq & \varepsilon^{-2} \sup_{i \leq n, n \geq 1} |n^{-1/2} Y_{i-1}^*|^{\ell_1} \cdot n^{-1} \textstyle\sum_{i=1}^n |U_i| = O_p(1), \end{aligned} \tag{4.70}$$

using Assumption CHE(i), $\sup_{i \leq n, n \geq 1} |n^{-1/2} Y_{i-1}^*| = O_p(1)$, and a WLLN for strong-mixing triangular arrays of $L^{1+\delta}$-bounded random variables, see Andrews (1988), which relies on Assumption INNOV(iv). The results in Lemma 5(l) hold by the same arguments as given above.

Next, suppose $h_1 = \infty$. Lemma 6 shows that $E(Y_0^{*2}/\phi_1^2) = O((1-\rho)^{-1})$ and $E(Y_0^{*2} U_1^2 / \phi_1^4) = O((1-\rho)^{-1})$, where $O((1-\rho)^{-1}) = O(1)$ in the case where $\rho \to \rho^* < 1$. Inspection of the proof of Theorem 1 then shows that it suffices to show that the equivalent of (4.15)-(4.17) holds when ϕ_i is replaced by $\widehat{\phi}_i$. More precisely, by Lemma 6, for (4.16) it is sufficient to show that

$$\text{(i) } n^{-1}(1-\rho) \sum_{i=1}^n (Y_{i-1}^*)^2 (\widehat{\phi}_i^{-2} - \phi_i^{-2}) = o_p(1), \tag{4.71}$$

(ii) $n^{-1}(1-\rho)^{1/2}\sum_{i=1}^n Y_{i-1}^*(\widehat{\phi}_i^{-2}-\phi_i^{-2}) = o_p(1)$, and (iii) $n^{-1}\sum_{i=1}^n(\widehat{\phi}_i^{-2}-\phi_i^{-2}) = o_p(1)$. In addition, for (4.15), it is sufficient to show that (iv) $n^{-1/2}\sum_{i=1}^n((1-\rho)^{1/2}Y_{i-1}^*)^jU_i \times(\widehat{\phi}_i^{-2}-\phi_i^{-2}) = o_p(1)$ for $j = 0, 1$. To show (4.17), it is enough to show that in addition $n^{-1}(1-\rho)X_1'\Delta^2X_1 \to_p 1$, $n^{-1}(1-\rho)^{1/2}(X_2'\Delta^2X_1) = O_p(1)$, and $n^{-1}(X_2'\Delta^2X_2) = O_p(1)$ hold (with X_1, X_2, and Δ defined with $\widehat{\phi}_i$, not ϕ_i). Inspecting the proof of Lemma 8(e)-(g) carefully, it follows that to show the latter three conditions, it is enough to show that in addition to (i)-(iv), we have (v) $n^{-1}(1-\rho)\sum_{i=1}^n(Y_{i-1}^*)^2 U_i^2(\widehat{\phi}_i^{-4}-\phi_i^{-4}) = o_p(1)$ and (vi) $n^{-r_1}(1-\rho)^{r_2}\sum_{i=1}^n(Y_{i-1}^*)^{r_3}\, U_i^{r_4}(\widehat{\phi}_i^{-4}-\phi_i^{-4}) = o_p(1)$ for $(r_1,...,r_4) = (3/2,1,2,1)$, $(2,1,2,0)$, $(3/2,3/2,3,1)$, $(2,3/2,3,0)$, and $(2,3/2,4,0)$. These conditions come from the proof of Lemma 8 in (4.60).

Conditions (iii) and (iv) are assumed in Assumption CHE(ii)(c) and (d). Immediately below we prove (i) in (4.71) using Assumption CHE(ii)(d) with $(d,j,k) = (2,2,0)$; (ii), (v), and (vi) can be shown using exactly the same approach by applying Assumption CHE(ii)(d) with $(d,j,k) = (1,2,0)$, $(2,4,0)$, $(2,4,2)$, and $(2,4,4)$, respectively.

We now prove (i) in (4.71). Note that by the Cauchy-Schwarz inequality we have

$$\begin{aligned} & n^{-1}(1-\rho)\sum_{i=1}^n(Y_{i-1}^*)^2(\widehat{\phi}_i^{-2}-\phi_i^{-2}) \\ \leq\ & \left(n^{-1}(1-\rho)^2\sum_{i=1}^n(Y_{i-1}^*)^4\right)^{1/2}\left(n^{-1}\sum_{i=1}^n(\widehat{\phi}_i^{-2}-\phi_i^{-2})^2\right)^{1/2} \end{aligned} \tag{4.72}$$

and therefore by Assumption CHE(ii)(d) it is enough to show that $n^{-1}(1-\rho)^2\sum_{i=1}^n (Y_{i-1}^*)^4 = O_p(1)$. By Markov's inequality, we have

$$P\left(n^{-1}(1-\rho)^2\sum_{i=1}^n(Y_{i-1}^*)^4 > M\right) \leq M^{-2}n^{-2}(1-\rho)^4\sum_{i,j=1}^n E(Y_{i-1}^*Y_{j-1}^*)^4. \tag{4.73}$$

Thus, it is enough to show that for

$$E_{ijstuvabcd} = E(U_{i-1-s}U_{i-1-t}U_{i-1-u}U_{i-1-v}U_{j-1-a}U_{j-1-b}U_{j-1-c}U_{j-1-d}), \tag{4.74}$$

we have

$$n^{-2}(1-\rho)^4\sum_{i,j=1}^n\sum_{s,t,u,v=0}^{\infty}\sum_{a,b,c,d=0}^{\infty}\rho^{a+b+c+d+s+t+u+v}E_{ijstuvabcd} = O(1). \tag{4.75}$$

In the case where $\rho \to \rho^* < 1$, (4.75) holds by Assumption INNOV(iv). Next consider the case when $\rho \to 1$. Note that when the largest subindex $i-1-s, ..., j-1-d$ in (4.75) appears only once in $E_{ijstuvabcd}$, then the expectation equals zero because U_i is a martingale difference sequence. As in earlier proofs, one can then show that it is enough to consider the case where the largest subindex appears twice and all other subindices are different from each other. As in earlier proofs, one has to consider different subcases regarding the order of the subindices. We consider only one case here, namely the case where $i-1-s < i-1-t < ... < j-1-b < j-1-c = j-1-d$ and thus $c = d$. The other cases are handled using an analogous approach. We make use of the mixing inequality in (4.2) and apply Assumption INNOV(iv). Note that

$$\begin{aligned}
& n^{-2}(1-\rho)^4 \sum_{i,j=1}^{n} \sum_{s>t>u>v=0}^{\infty} \sum_{a>b>c=0}^{\infty} \rho^{a+b+2c+s+t+u+v} E_{ijstuvabcc} \\
&= O(n^{-2}(1-\rho)^4) \sum_{i,j=1}^{n} \sum_{s>t>u>v=0}^{\infty} \sum_{a>b>c=0}^{\infty} \rho^{a+b+2c+s+t+u+v} (\max\{s-t, t-u, b-c\})^{-3-\varepsilon} \\
&= O(n^{-2}(1-\rho)^3) \sum_{i,j=1}^{n} \sum_{s>t=0}^{\infty} \rho^s (s-t)^{-1-\varepsilon/3} \sum_{u>v=0}^{\infty} \rho^s (s-t)^{-1-\varepsilon/3} \sum_{b>c=0}^{\infty} \rho^b (b-c)^{-1-\varepsilon/3} \\
&= O(1), \qquad (4.76)
\end{aligned}$$

where the last equality holds because $\sum_{b>c=0}^{\infty} \rho^b (b-c)^{-1-\varepsilon/3} = \sum_{c=0}^{\infty} \rho^c \sum_{b=1}^{\infty} \rho^b b^{-1-\varepsilon/3}$ $= O((1-\rho)^{-1})$. This completes the proof of (i) in (4.71). □

4.3 Proof of Lemma 1

Proof of Lemma 1. Assumption CHE(i) holds by Assumption CHE2(i) and (v). We verify Assumption CHE(ii)(a) (which applies when $h_1 < \infty$) for $j = 1$. The proof for $j = 0$ is similar. We need to show that

$$n^{-1/2} \sum_{i=1}^{n} (n^{-1/2} Y_{i-1}^*) U_i [\widehat{\phi}_i^{-2} - \phi_i^{-2}] = o_p(1). \qquad (4.77)$$

To do so, we need to take account of the fact that under Assumption CHE2, $\widehat{\phi}_i^2$ differs from ϕ_i^2 in three ways. First, $\widehat{\phi}_i^2$ is based on the estimated conditional heteroskedasticity parameter $\widetilde{\pi}_n$, not the pseudo-true value π_n; second, $\widehat{\phi}_i^2$ is based on residuals, i.e., it uses $(\widetilde{\alpha}_n, \widetilde{\rho}_n)$, not the true values $(0, \rho_n)$; and third $\widehat{\phi}_i^2$ is defined using the truncated-at-time-period-one value L_i, not L.

Assumption CHE2(iii) and (iv) implies that $||\widehat{\pi}_n - \pi_n|| \leq Cn^{-\delta_2}$ wp→ 1 for some constant $C < \infty$. Hence, $\widetilde{\pi}_n \in \Pi_{n,0} = \Pi_n \cap B(\pi_n, Cn^{-\delta_2})$ wp→ 1 (where $B(\pi, \delta)$ denotes a ball with center at π and radius δ). The set $\Pi_{n,0}$ contains a finite number of elements and the number is bounded over $n \geq 1$. Without loss of generality, we can assume that $\Pi_{n,0}$ contains $K < \infty$ elements for each $n \geq 1$. We order the elements in each set $\Pi_{n,0}$ and call them $\pi_{n,k}$ for $k = 1, ..., K$. This yields K sequences $\{\pi_{n,k} : n \geq 1\}$ for $k = 1, ..., K$.

To show (4.77), we use the following argument. Suppose for some random variables $\{(Z_{n,0}, Z_n(\pi_{n,1}), ..., Z_n(\pi_{n,K}))' : n \geq 1\}$ and Z, we have

$$(Z_{n,0}, Z_n(\pi_{n,1}), ..., Z_n(\pi_{n,K}))' \to_d (Z, ..., Z)' \tag{4.78}$$

as $n \to \infty$. In addition, suppose $\widetilde{\pi}_n \in \{\pi_{n,1}, ..., \pi_{n,K}\}$ wp$\to$ 1. Then, by the continuous mapping theorem,

$$\begin{aligned}
&\min_{k \leq K} Z_n(\pi_{n,k}) - Z_{n,0} \to_d \left(\min_{k \leq K} Z\right) - Z = 0, \\
&\max_{k \leq K} Z_n(\pi_{n,k}) - Z_{n,0} \to_d \left(\max_{k \leq K} Z\right) - Z = 0, \\
&Z_n(\widetilde{\pi}_n) - Z_{n,0} \in [\min_{k \leq K} Z_n(\pi_{n,k}) - Z_{n,0}, \max_{k \leq K} Z_n(\pi_{n,k}) - Z_{n,0}] \text{ wp} \to 1, \text{ and hence,} \\
&Z_n(\widetilde{\pi}_n) - Z_{n,0} \to_d 0. \qquad (4.79)
\end{aligned}$$

Since convergence in distribution to zero is equivalent to convergence in probability to zero, this gives $Z_n(\widetilde{\pi}_n) - Z_{n,0} \to_p 0$. We apply this argument with

$$\begin{aligned}
Z_{n,0} &= n^{-1/2} \sum_{i=1}^n (n^{-1/2} Y_{i-1}^*) U_i \phi_i^{-2} \text{ and} \\
Z_n(\pi_{n,k}) &= n^{-1/2} \sum_{i=1}^n (n^{-1/2} Y_{i-1}^*) U_i \phi_i^{-2}(\widetilde{\alpha}_n, \widetilde{\rho}_n, \pi_{n,k}) \qquad (4.80)
\end{aligned}$$

for $k = 1, ..., K$.

Hence, it suffices to show (4.78), where $\{\pi_{n,k} : n \geq 1\}$ is a fixed sequence such that $\pi_{n,k} \to \pi_0$ for $k = 1, ..., K$. To do so, we show below that

$$\begin{aligned}
&Z_n(\pi_{n,k}) - \overline{Z}_n(\pi_{n,k}) = o_p(1), \text{ where} \\
&\overline{Z}_n(\pi_{n,k}) = n^{-1/2} \sum_{i=1}^n (n^{-1/2} Y_{i-1}^*) U_i \phi_i^{-2}(0, \rho_n, \pi_{n,k}) \qquad (4.81)
\end{aligned}$$

(By definition, $\overline{Z}_n(\pi_{n,k})$ is the same as $Z_n(\pi_{n,k})$ except that it is defined using the true parameters $(0, \rho_n)$ rather than the estimated parameters $(\widetilde{\alpha}_n, \widetilde{\rho}_n)$.) It is then enough to show that (4.78) holds with $\overline{Z}_n(\pi_{n,k})$ in place of $Z_n(\pi_{n,k})$.

For the case $h_1 < \infty$ considered here, we do the latter by applying Lemma 4 with

$$v_{n,i} = (U_i, U_i \phi_i^{-2}, U_i \phi_i^{-2}(0, \rho_n, \pi_{n,1}), ..., U_i \phi_i^{-2}(0, \rho_n, \pi_{n,K}))'. \tag{4.82}$$

Conditions (i) and (ii) of Lemma 4 hold by Assumptions INNOV and CHE2(v) (which guarantees that $\widehat{\phi}_i^{-2}$ and $\phi_i^{-2}(0, \rho_n, \pi_{n,k})$ are uniformly bounded above). In addition,

$\Lambda = 0$ because $\{(v_{n,i}, \mathcal{G}_{n,i-1}) : i = ..., 0, 1, ...; n \geq 1\}$ is a martingale difference triangular array. Using Assumption CHE2(vi), for all $k_1, k_2, k_3, k_4 = 0, ..., K$, we have

$$\begin{aligned}
&\lim_{n\to\infty} n^{-1} E V_{n,k_1} V_{n,k_2}' = \lim_{n\to\infty} n^{-1} E V_{n,k_3} V_{n,k_4}', \text{ where} \\
&V_{n,0} = \sum_{i=1}^{n} U_i \phi_i^{-2} = \sum_{i=1}^{n} U_i \left(\omega_n + \sum_{j=1}^{L} \mu_j(\pi_n) U_{i-j}^2 \right) \text{ and} \\
&V_{n,k} = \sum_{i=1}^{n} U_i \phi_i^{-2}(0, \rho_n, \pi_{n,k}) = \sum_{i=1}^{n} U_i \left(\omega_{n,k} + \sum_{j=1}^{L_i} \mu_j(\pi_{n,k}) U_{i-j}^2 \right) \quad (4.83)
\end{aligned}$$

for $k = 1, ..., K$. In consequence, the matrix Ω_0 in Lemma 4 has all elements that are not in the first row or column equal to each other. For this reason, the elements in the limit random vector in (4.78) are equal to each other. We conclude that (4.78) holds when $\overline{Z}_n(\pi_{n,k})$ appears in place of $Z_n(\pi_{n,k})$ by Lemma 4(b). In this case, $Z = h_{2,1}^{1/2} \int I_h^* dM$, see Lemma 5(g) and its proof. The verification of Assumption CHE(ii)(a) when $j = 0$ is the same as that above because one of the elements of X_{i-1} in Lemma 4(b) can be taken to equal 1 and the latter result still holds with the corresponding element of K_h^* being equal to 1, see Hansen (1992, Thm. 3.1).

It remains to show (4.81) holds in the case $h_1 < \infty$ considered here. We only deal with the case $j = 1$. The case $j = 0$ can be handled analogously. To evaluate $\phi_i^{-2}(\widetilde{\alpha}_n, \widetilde{\rho}_n, \pi_{n,k}) - \phi_i^{-2}(0, \rho_n, \pi_{n,k})$, we use the Taylor expansion

$$(x + \delta)^{-1} = x^{-1} - x^{-2}\delta + x_*^{-3}\delta^2, \quad (4.84)$$

where x_* is between $x+\delta$ and x, applied with $x+\delta = \phi_i^2(\widetilde{\alpha}_n, \widetilde{\rho}_n, \pi_{n,k})$, $x = \phi_i^2(0, \rho_n, \pi_{n,k})$, and

$$\delta = \delta_i = \phi_i^2(\widetilde{\alpha}_n, \widetilde{\rho}_n, \pi_{n,k}) - \phi_i^2(0, \rho_n, \pi_{n,k}). \quad (4.85)$$

Thus, to show Assumption CHE(ii)(a), it suffices to show that

$$n^{-1/2} \sum_{i=1}^{n} (n^{-1/2} Y_{i-1}^*) U_i (\phi_i^{-4}(0, \rho_n, \pi_{n,k})\delta - x_*^{-3}\delta^2) = o_p(1). \quad (4.86)$$

Note that in the Taylor expansion, x^{-2} and x_*^{-3} are both bounded above (uniformly in i) because both $x+\delta$ and x are bounded away from zero by Assumption CHE2(v). Simple algebra gives

$$\begin{aligned}
\delta = &\sum_{t=1}^{L_i} \mu_t(\pi_{n,k})[-2U_{i-t}\widetilde{\alpha}_n - 2Y_{i-t-1}^* U_{i-t}(\widetilde{\rho}_n - \rho_n) \\
&+ \widetilde{\alpha}_n^2 + 2Y_{i-t-1}^*(\widetilde{\rho}_n - \rho_n)\widetilde{\alpha}_n + Y_{i-t-1}^{*2}(\widetilde{\rho}_n - \rho_n)^2]. \quad (4.87)
\end{aligned}$$

The effect of truncation by L_i rather than L only affects the finite number of summands with $i \leq L$ and hence its effect is easily seen to be asymptotically negligible and hence without loss of generality we can set $L_i = L$ for the rest of the proof.

We first deal with the contributions from $\phi_i^{-4}(0, \rho_n, \pi_{n,k})\delta$ in (4.86). Rather than considering the sum $\sum_{t=1}^{L_i}$ in (4.87) when showing (4.86), it is enough to show that for every fixed $t = 1, ..., L$ the resulting expression in (4.86) is $o_p(1)$. Fix $t \in \{1, ..., L\}$ and set $b_i = \phi_i^{-4}(0, \rho, \pi_{n,k})$. It is enough to show that

$$n^{-1/2} \sum_{i=1}^{n} (n^{-1/2} Y_{i-1}^*) U_i b_i c_{it} = o_p(1), \tag{4.88}$$

where c_{it} equals

$$\begin{aligned} &\text{(i) } U_{i-t}\widetilde{\alpha}_n, \text{ (ii) } Y_{i-t-1}^* U_{i-t}(\widetilde{\rho}_n - \rho), \text{ (iii) } \widetilde{\alpha}_n^2, \\ &\text{(iv) } Y_{i-t-1}^*(\widetilde{\rho}_n - \rho)\widetilde{\alpha}_n, \text{ or (v) } Y_{i-t-1}^{*2}(\widetilde{\rho}_n - \rho)^2. \end{aligned} \tag{4.89}$$

By Assumption CHE2(iii) and because $h_1 < \infty$, we have (1) $\widetilde{\alpha}_n = O_p(n^{-1/2})$ and $\widetilde{\rho}_n - \rho = O_p(n^{-1})$. Terms of the form (2) $n^{-1} \sum_{i=1}^n Y_{i-1}^* U_i b_i U_{i-t}$ and $n^{-3/2} \sum_{i=1}^n Y_{i-1}^* Y_{i-t-1}^* \times U_i U_{i-t} b_i$ are $O_p(1)$ by Lemma 4(b) and (c) applied with $v_{n,i} = (U_i, U_{i-t}, U_i U_{i-t} b_i)'$. Note here that b_i is an element of the σ-field $\sigma(U_{i-L}, ..., U_{i-1})$ by definition of $\phi_i^2(0, \rho, \pi_{n,k})$ in (3.14) and by Assumption CHE2(i) and (v), (3) $\sup_{i \le n, n \ge 1} |n^{-1/2} Y_{i-1}^*| = O_p(1)$ by Lemma 5(a), (4) terms of the form $n^{-1} \sum_{i=1}^n |U_i U_{i-1}^j|$ for $j = 1, 2$ are $O_p(1)$ by a WLLN for strong-mixing triangular arrays, see Andrews (1988), and (5) the b_i are $O_p(1)$ uniformly in i. The result in (4.88) for cases (i)-(ii) of (4.89) follows from (2). Cases (iii)-(v) are established by $|n^{-1/2} \sum_{i=1}^n (n^{-1/2} Y_{i-1}^*) U_i b_i c_{it}| \le \sup_{i \le n, n \ge 1} |n^{-1/2} Y_{i-1}^*| n^{-3/2} \sum_{i=1}^n |U_i| = o_p(1)$ using (1) and (3)-(5).

Next, we deal with the contributions from $x_*^{-3}\delta^2$ in (4.86). Because x_*^{-3} and $\mu_t(\pi_{n,k})$ are both $O_p(1)$ uniformly in i, it is enough to show that

$$n^{-1/2} \sum_{i=1}^{n} |n^{-1/2} Y_{i-1}^* U_i c_{ij_1} d_{ij_2}| = o_p(1), \tag{4.90}$$

where c_{ij} and $d_{ij} \in \{U_{i-j}\widetilde{\alpha}_n, Y_{i-j-1}^* U_{i-j}(\widetilde{\rho}_n - \rho), \widetilde{\alpha}_n^2, Y_{i-j-1}^*(\widetilde{\rho}_n - \rho)\widetilde{\alpha}_n, Y_{i-j-1}^{*2}(\widetilde{\rho}_n - \rho)^2\}$ and $j_1, j_2 \in \{1, ..., L_i\}$. Conditions (1), (3), and (4) then imply (4.90). This completes the proof of Assumption CHE(ii)(a).

Next, we verify Assumption CHE(ii)(b) (which applies when $h_1 < \infty$). For the cases of $(d, j) = (0, 2)$, $(1, 2)$, and $(2, 2)$, the proof is similar to that given below for Assumption CHE(ii)(d) but with $a_n = O(n^{1/2}(1-\rho)^{-1/2})$ replaced by $a_n = n$ and using the results above that (i) $\sup_{i \le n, n \ge 1} |n^{-1/2} Y_{i-1}^*| = O_p(1)$ and (ii) terms of the form $n^{-1} \sum_{i=1}^n |U_i^{j_1} U_{i-1}^{j_2}|$ for $j_1 = 1, 2$ and $j_2 = 1, 2$ are $O_p(1)$, which holds using Assumption INNOV(iv). (Note that the case of $(d, j) = (0, 2)$ is not needed for Assumption CHE(ii) but is used in the verification of Assumption CHE(ii)(b) for the case where $(d, j) = (0, 1)$, which follows.)

We now verify Assumption CHE(ii)(b) for $(d, j) = (0, 1)$. We have

$$\begin{aligned} &n^{-1} \textstyle\sum_{i=1}^n |\widehat{\phi}_i^{-1} - \phi_i^{-1}| = n^{-1} \sum_{i=1}^n |\widehat{\phi}_i - \phi_i| / (\widehat{\phi}_i \phi_i) \\ &\le \varepsilon^{-1} n^{-1} \textstyle\sum_{i=1}^n |\widehat{\phi}_i - \phi_i| \le \varepsilon^{-3/2} n^{-1} \sum_{i=1}^n |\widehat{\phi}_i^2 - \phi_i^2|, \end{aligned} \tag{4.91}$$

where the first inequality holds because $\widehat{\phi}_i^2$ and ϕ_i^2 are bounded away from zero by some $\varepsilon > 0$ by Assumption CHE2(i), (ii), and (v) and the second inequality holds by the mean-value expansion $(x+\delta)^{1/2} = x^{1/2} + (1/2)x_*^{-1/2}\delta$, where x_* lies between $x+\delta$ and x, applied with $x+\delta = \widehat{\phi}_i^2$, $x = \phi_i^2$, $\delta = \widehat{\phi}_i^2 - \phi_i^2$, and $x_*^{-1/2} = \phi_{i,*}^{-1} \leq \varepsilon^{-1/2}$ using Assumption CHE2(v), where $\phi_{i,*}^2$ lies between $\widehat{\phi}_i^2$ and ϕ_i^2. The rhs of (4.91) is $o_p(1)$ by the result above that Assumption CHE(ii)(b) holds for $(d,j) = (0,2)$.

Next, we verify Assumption CHE(ii)(c) (which applies when $h_1 = \infty$). We only show the case $j = 1$, the case $j = 0$ is handled analogously. We use a very similar approach to the one in the proof of Assumption CHE(ii)(a). We show that (4.81) holds when $h_1 = \infty$ and that

$$Z_{n,0} - \overline{Z}_n(\pi_{n,k}) = o_p(1) \tag{4.92}$$

for every $k = 1, ..., K$, where

$$\begin{aligned} Z_{n,0} &= n^{-1/2}\sum_{i=1}^{n}((1-\rho)^{1/2}Y_{i-1}^*)U_i\phi_i^{-2}, \\ Z_n(\pi_{n,k}) &= n^{-1/2}\sum_{i=1}^{n}((1-\rho)^{1/2}Y_{i-1}^*)U_i\phi_i^{-2}(\widetilde{\alpha}_n, \widetilde{\rho}_n, \pi_{n,k}), \text{ and} \\ \overline{Z}_n(\pi_{n,k}) &= n^{-1/2}\sum_{i=1}^{n}((1-\rho)^{1/2}Y_{i-1}^*)U_i\phi_i^{-2}(0, \rho, \pi_{n,k}). \end{aligned} \tag{4.93}$$

We first show (4.81). By (4.84),

$$\begin{aligned} & n^{-1/2}\sum_{i=1}^{n}((1-\rho)^{1/2}Y_{i-1}^*)U_i(\phi_i^{-2}(\widetilde{\alpha}_n, \widetilde{\rho}_n, \pi_{n,k}) - \phi_i^{-2}(0, \rho, \pi_{n,k})) \\ &= n^{-1/2}\sum_{i=1}^{n}((1-\rho)^{1/2}Y_{i-1}^*)U_i(-\phi_i^{-4}(0, \rho, \pi_{n,k})\delta + x_*^{-3}\delta^2), \end{aligned} \tag{4.94}$$

where δ is defined in (4.87) and x_* in (4.84). Hence, it suffices to show that the expression in the second line of (4.94) is $o_p(1)$. First, we deal with the contributions from $-\phi_i^{-4}(0, \rho, \pi_{n,k})\delta$ in (4.94). Rather than considering the sum $\sum_{j=1}^{L_i}$ in (4.87) when showing (4.94), it is enough to show that for every fixed $j = 1, ..., L_i$ the expression in the second line of (4.94) is $o_p(1)$. Fix $j \in \{1, ..., L_i\}$, set $b_i = \phi_i^{-4}(0, \rho, \pi_{n,k})$, and note that $\mu_j(\pi_{n,k})$ is bounded by Assumption CHE2(vi). It is enough to show that

$$n^{-1/2}\sum_{i=1}^{n}((1-\rho)^{1/2}Y_{i-1}^*)U_ib_ic_{ij} = o_p(1), \tag{4.95}$$

where c_{ij} equals

$$\begin{aligned} &\text{(i) } U_{i-j}\widetilde{\alpha}_n, \text{ (ii) } Y_{i-j-1}^*U_{i-j}(\widetilde{\rho}_n - \rho), \text{ (iii) } \widetilde{\alpha}_n^2, \\ &\text{(iv) } Y_{i-j-1}^*(\widetilde{\rho}_n - \rho)\widetilde{\alpha}_n, \text{ or (v) } Y_{i-j-1}^{*2}(\widetilde{\rho}_n - \rho)^2. \end{aligned} \tag{4.96}$$

In case (i) of (4.96), we use Assumption CHE2(iii) which implies $\widetilde{\alpha}_n = O_p(n^{-1/2})$. By Markov's inequality and Assumption STAT, we have

$$\begin{aligned}
&P(|n^{-1}(1-\rho)^{1/2}\sum_{i=1}^{n} Y_{i-1}^{*} U_i b_i U_{i-j}| > \varepsilon) \\
&= O(n^{-2}(1-\rho)) \sum_{i,k=1}^{n} E b_i b_k Y_{i-1}^{*} Y_{k-1}^{*} U_i U_{i-j} U_k U_{k-j} \\
&= O(n^{-2}(1-\rho)) \sum_{i,k=1}^{n} \sum_{s,t=0}^{\infty} \rho^{s+t} E b_i b_k U_{i-s-1} U_{k-t-1} U_i U_{i-j} U_k U_{k-j}. \qquad (4.97)
\end{aligned}$$

Note that b_i is an element of the σ-field $\sigma(U_{i-L}, ..., U_{i-1})$. The latter holds by definition of $\phi_i^2(0, \rho, \pi_{n,k})$ in (3.14) and by Assumption CHE2(i) and (v). To show that the last expression in (4.97) is $o(1)$ we have to distinguish several subcases. As in several proofs above, we can assume that all subindices $i-s-1$, $k-t-1, ...,$ $k-j$ are different. We only consider the case $i-s-1 < k-t-1 < i-j < k-j$. The other cases can be dealt with using an analogous approach. By Assumption INNOV(iv) and the mixing inequality in (4.2), we have

$$\begin{aligned}
&\sum_{k=1}^{n} \sum_{s,t=0}^{\infty} \sum_{i=1}^{n} \rho^{s+t} E b_i b_k U_{i-s-1} U_{k-t-1} U_i U_{i-j} U_k U_{k-j} \\
&= O(1) \sum_{k=1}^{n} \sum_{s,t=0}^{\infty} \sum_{i=1}^{k-t+s-1} \rho^{s+t} (k-t-i+s)^{-3-\varepsilon} \\
&= O(1) \sum_{k=1}^{n} \sum_{s,t=0}^{\infty} \rho^{s+t} \sum_{i=1}^{k-t+s-1} i^{-3-\varepsilon} \\
&= O(n(1-\rho)^{-2}), \qquad (4.98)
\end{aligned}$$

where in the third line we do the change of variable $i \mapsto k-t-i+s$. This implies that the expression in (4.97) is $o(1)$ because $n(1-\rho) \to \infty$.

In case (ii) of (4.96), using $\widetilde{\rho}_n - \rho = O_p(n^{-1/2}(1-\rho)^{1/2})$ by Assumption CHE2(iii), (3.6), and Lemma 6, and using Markov's inequality as for case (i), it is enough to show that

$$\sum_{i,k=1}^{n} \sum_{s,t=0}^{\infty} \sum_{u,v=0}^{\infty} \rho^{s+t+u+v} E b_{ij} b_{kj} U_{i-s-1} U_{i-j-1-t} U_i U_{i-j} U_{k-u-1} U_{k-j-1-v} U_k U_{k-j} \qquad (4.99)$$

is $o(n^2(1-\rho)^{-2})$. Again, one has to separately examine several subcases regarding the order of the subindices $i-s-1, ..., k-j$ on the random variables U_i. We can assume that all subindices are different. We only study the case $i-s-1 < i-j-1-t < k-u-1 < k-j-1-v < i-j$. The other cases can be handled analogously. By

Assumption INNOV(iv), boundedness of b_i, and the mixing inequality in (4.2), the expression in (4.99) is of order

$$
\begin{aligned}
&O(1)\sum_{k=1}^{n}\sum_{s,t=0}^{\infty}\sum_{u,v=0}^{\infty}\sum_{i=k-v}^{n}\rho^{s+t+u+v}\max(s-t-j,i-k+v+1)^{-3-\varepsilon}\\
&=O(1)\sum_{u,v=0}^{\infty}\rho^{u+v}\sum_{k=1}^{n}\sum_{i=k-v}^{n}(i-k+v+1)^{-3/2}\sum_{s,t=0}^{\infty}\rho^{s+t}(s-t-j)^{-3/2}\\
&=O((1-\rho)^{-3}n), \qquad (4.100)
\end{aligned}
$$

where in the first line we use $k-1-v<i$ and in the last line we use $\sum_{i=k-v}^{n}(i-k+v+1)^{-3/2}=\sum_{i=1}^{n-k+v+1}i^{-3/2}=O(1)$. The desired result then follows because $n(1-\rho)\to\infty$ implies $O((1-\rho)^{-3}n)=o(n^2(1-\rho)^{-2})$.

Cases (iii)-(v) of (4.96) can be handled analogously.

Next, we show that the contribution from $x_*^{-3}\delta^2$ in (4.94) is $o_p(1)$. Noting that x_*^{-3} and $\mu_j(\pi_{n,k})$ are $O_p(1)$ uniformly in i by Assumption CHE2(ii), (v), and (vi), it is enough to show that $n^{-1/2}(1-\rho)^{1/2}\sum_{i=1}^{n}|Y_{i-1}^*U_ic_{ij_1}d_{ij_2}|=o_p(1)$, where c_{ij} and $d_{ij}\in\{U_{i-j}\widetilde{\alpha}_n, Y_{i-j-1}^*U_{i-j}(\widetilde{\rho}_n-\rho), \widetilde{\alpha}_n^2, Y_{i-j-1}^*(\widetilde{\rho}_n-\rho)\widetilde{\alpha}_n, Y_{i-j-1}^{*2}(\widetilde{\rho}_n-\rho)^2\}$ and $j_1,j_2\in\{1,...,L_i\}$. Using $\widetilde{\alpha}_n=O_p(n^{-1/2})$ and $\widetilde{\rho}_n-\rho=O_p(n^{-1/2}(1-\rho)^{1/2})$ the latter follows easily from Markov's inequality. For example,

$$
\begin{aligned}
&P(n^{-1/2}(1-\rho)^{1/2}\sum_{i=1}^{n}|Y_{i-1}^*U_i(U_{i-j_1}\widetilde{\alpha}_n)(U_{i-j_2}\widetilde{\alpha}_n)|>\varepsilon)\\
&=O(n^{-3}(1-\rho))\sum_{i,k=1}^{n}\sum_{s,t=0}^{\infty}\rho^{s+t}E|U_{i-1-s}U_iU_{i-j_1}U_{i-j_2}U_{k-1-t}U_kU_{k-j_1}U_{k-j_2}|\\
&=O(n^{-3}(1-\rho))(1-\rho)^{-2}n^2\\
&=o(1) \qquad (4.101)
\end{aligned}
$$

by Assumption INNOV(iv) and $n(1-\rho)\to\infty$.

Next we show that (4.92) holds. We have

$$
\begin{aligned}
&Z_{n,0}-\overline{Z}_n(\pi_{n,k})\\
&=n^{-1/2}\sum_{i=1}^{n}((1-\rho)^{1/2}Y_{i-1}^*)U_i(\phi_i^{-2}-\phi_i^{-2}(0,\rho,\pi_{n,k}))\\
&=n^{-1/2}(1-\rho)^{1/2}\sum_{i=1}^{n}Y_{i-1}^*U_i(\phi_i^2(0,\rho,\pi_{n,k})-\phi_i^2)(\phi_i^{-2}\phi_i^{-2}(0,\rho,\pi_{n,k}))\\
&=n^{-1/2}(1-\rho)^{1/2}\sum_{i=1}^{n}Y_{i-1}^*U_i\left(\omega_n-\omega_{n,k}+\sum_{j=1}^{L}(\mu_j(\pi_n)-\mu_j(\pi_{n,k}))U_{i-j}^2\right)\\
&\quad\times(\phi_i^{-2}\phi_i^{-2}(0,\rho,\pi_{n,k}))+o_p(1), \qquad (4.102)
\end{aligned}
$$

where ω_n is defined in Assumption CHE2(ii). Thus, it is enough to show that

$$D_1 = n^{-1/2}(1-\rho)^{1/2}\sum_{i=1}^{n} Y_{i-1}^* U_i(\omega_n - \omega_{n,k})(\phi_i^{-2}\phi_i^{-2}(0,\rho,\pi_{n,k})) \text{ and}$$
$$D_{2j} = n^{-1/2}(1-\rho)^{1/2}\sum_{i=1}^{n} Y_{i-1}^* U_i((\mu_j(\pi_n) - \mu_j(\pi_{n,k}))U_{i-j}^2)(\phi_i^{-2}\phi_i^{-2}(0,\rho,\pi_{n,k})) \tag{4.103}$$

are $o_p(1)$ for $j = 1, ..., L$. We can prove $D_{2j} = o_p(1)$ along the same lines as $D_1 = o_p(1)$ and we therefore only prove $D_1 = o_p(1)$. By Assumption CHE2(ii) and $\pi_{n,k} \to \pi_0$, we have $\omega_n - \omega_{n,k} \to 0$. Thus, by Markov's inequality and Assumption STAT,

$$\begin{aligned} &P(|D_1| > \varepsilon) \\ &= o(n^{-1}(1-\rho))\sum_{i,v=1}^{n}\sum_{s,t=0}^{\infty}\rho^{s+t}EU_{i-1-s}U_iU_{v-1-t}U_v \\ &\quad\times\phi_i^{-2}\phi_i^{-2}(0,\rho,\pi_{n,k})\phi_v^{-2}\phi_v^{-2}(0,\rho,\pi_{n,k}). \end{aligned} \tag{4.104}$$

The random variable $e_{iv} = (\phi_i^{-2}\phi_i^{-2}(0,\rho,\pi_{n,k}))(\phi_v^{-2}\phi_v^{-2}(0,\rho,\pi_{n,k}))$ is an element of the σ-field $\sigma(U_{\min\{i,v\}-L}, ..., U_{\max\{i,v\}})$ by definition of $\phi_i^2(0,\rho,\pi_{n,k})$ in (3.14) and by Assumption CHE2(i) and (v). To prove that the rhs in (4.104) is $o_p(1)$ we have to study several subcases. We only examine the subcase where all subindices $i-1-s, i, v-1-t, v$ are different and where $i-1-s < i < v-1-t < v$. The other cases can be dealt with analogously. By Assumption INNOV(iv), boundedness of e_{iv}, and the mixing inequality in (4.2), the rhs in (4.104) for the particular subcase is of order

$$\begin{aligned} &o(n^{-1}(1-\rho))\sum_{i,v=1}^{n}\sum_{s,t=0}^{\infty}\rho^{s+t}(s+1)^{-3/2}(v-1-t-i)^{-3/2} \\ &= o(n^{-1}(1-\rho))\sum_{s,t=0}^{\infty}\rho^{s+t}(s+1)^{-3/2}\sum_{v=1}^{n}\sum_{i=1}^{v-2-t}(v-1-t-i)^{-3/2} \\ &= o(n^{-1}(1-\rho))O((1-\rho)^{-1})O(n) \\ &= o(1), \end{aligned} \tag{4.105}$$

where in the third line a change of variable $i \to -i-t-1+v$ was used. This completes the verification of Assumption CHE(ii)(c).

Finally, we show that Assumption CHE(ii)(d) holds. First, note that Assumptions CHE2(i), (ii), and (v) imply $\widehat{\phi}_i^{-j}\phi_i^{-j} = O_p(1)$ uniformly in i. Therefore, writing $|\widehat{\phi}_i^{-j} - \phi_i^{-j}|^d$ as $|(\phi_i^j - \widehat{\phi}_i^j)/(\widehat{\phi}_i^j\phi_i^j)|^d$ we have

$$n^{-1}\sum_{i=1}^{n}|U_i^k(\widehat{\phi}_i^{-j} - \phi_i^{-j})^d| = O_p(1)n^{-1}\sum_{i=1}^{n}|U_i^k|\cdot|\widehat{\phi}_i^j - \phi_i^j|^d. \tag{4.106}$$

We need to show that the quantity in (4.106) is $o_p(1)$. Note that by the definition of $\widehat{\phi}_i^2$ in (3.14) and ϕ_i^2 in Assumption CHE2(ii) we have

$$|\widehat{\phi}_i^j - \phi_i^j|^d = \left| \left(\widetilde{\omega}_n + \sum_{v=1}^{L_i} \mu_v(\widetilde{\pi}_n) \widehat{U}_{i-v}^2(\widetilde{\alpha}_n, \widetilde{\rho}_n) \right)^{j/2} - \left(\omega_n + \sum_{v=1}^{L} \mu_v(\pi_n) U_{i-v}^2 \right)^{j/2} \right|^d \tag{4.107}$$

with $\widehat{U}_{i-v}^2(\widetilde{\alpha}_n, \widetilde{\rho}_n) = (-(\widetilde{\rho}_n - \rho)Y_{i-v-1}^* - \widetilde{\alpha}_n + U_{i-v})^2$. It can be shown that the additional terms in (4.106), that arise if we replace L_i by L in (4.107), are of order $o_p(1)$. We first study the case where $j = 2$. Multiplying out in (4.107), it follows that when $d = 1$, $\widehat{\phi}_i^2 - \phi_i^2$ can be bounded by a finite sum of elements in $S = \{|\widetilde{\omega}_n - \omega_n|$, $|\mu_v(\widetilde{\pi}_n) - \mu_v(\pi_n)|U_{i-v}^2$, $(\widetilde{\rho}_n - \rho)^2 Y_{i-v-1}^{*2}$, $\widetilde{\alpha}_n^2$, $|(\widetilde{\rho}_n - \rho)Y_{i-v-1}^* \widetilde{\alpha}_n|$, $|(\widetilde{\rho}_n - \rho)Y_{i-v-1}^* U_{i-v}|$, $\widetilde{\alpha}_n U_{i-v}$: for $v = 1, ..., L\}$. When $d = 2$, $(\widehat{\phi}_i^j - \phi_i^j)^2$ can be bounded by a finite sum of elements given as products of two terms in S. By Assumption CHE2(iii) and $a_n = O(n^{1/2}(1-\rho)^{-1/2})$, we have $\widetilde{\rho}_n - \rho = O_p(n^{-1/2}(1-\rho)^{1/2})$, $\widetilde{\alpha}_n = O_p(n^{-1/2})$, and $\widetilde{\omega}_n - \omega_n = O_p(n^{-\delta_2})$. To show the quantity in (4.106) is $o_p(1)$, it is enough to verify that $n^{-1} \sum_{i=1}^n |U_i^k s_{i1} s_{i2}| = o_p(1)$ where for $d = 1$, $s_{i1} \in S$ and $s_{i2} = 1$ and for $d = 2$, $s_{i1}, s_{i2} \in S$. We only show this for one particular choice of s_{i1}, s_{i2}, namely, $s_{i1} = s_{i2} = |\mu_v(\widetilde{\pi}_n) - \mu_v(\pi_n)|U_{i-v}^2$; the other cases can be handled analogously. In that case, we have $|\mu_v(\widetilde{\pi}_n) - \mu_v(\pi_n)|^2 n^{-1} \sum_{i=1}^n |U_i^k U_{i-v}^2| = o_p(1)$ because $|\mu_v(\widetilde{\pi}_n) - \mu_v(\pi_n)|^2 = o(1)$ by Assumption CHE2(iii), (iv), and (vi), and $n^{-1} \sum_{i=1}^n |U_i^k U_{i-v}^2| = O_p(1)$ by a weak law of large numbers for triangular arrays of $L^{1+\delta}$-bounded strong-mixing random variables for $\delta > 0$, see Andrews (1988), using the moment conditions in Assumption INNOV(iv).

The case $j = 4$ can be proved analogously. □

References

Andrews, D. W. K., 1988. Laws of large numbers for dependent non-identically distributed random variables. Econometric Theory 4, 458-467.

Andrews, D. W. K., 1993. Exactly median-unbiased estimation of first-order autoregressive/unit root models. Econometrica 61, 139-165.

Andrews, D. W. K., Chen, H.-Y., 1994. Approximately median-unbiased estimation of autoregressive models. Journal of Business and Economic Statistics 12, 187–204.

Andrews, D. W. K., Guggenberger, P., 2005. Hybrid and size-corrected subsampling methods. Unpublished working paper, Cowles Foundation, Yale University.

Bobkowski, M. J., 1983. Hypothesis testing in nonstationary time series. Unpublished Ph. D. Thesis, University of Wisconsin, Madison.

Cavanagh, C., 1985. Roots local to unity. Unpublished manuscript, Harvard University.

Chan, N. H., Wei, C. Z., 1987. Asymptotic inference for nearly nonstationary AR(1) processes. Annals of Statistics 15, 1050-1063.

Chen, W., Deo R., 2007. A smooth transition to the unit root distribution via the chi-square distribution with interval estimation for nearly integrated autoregressive processes. Unpublished manuscript, Stern School of Business, New York University.

Doukhan P., 1994. Mixing properties and examples. Lecture Notes in Statistics 85. Springer, New York.

Elliott, G., 1999. Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution. International Economic Review 40, 767-783.

Elliott, G., Stock, J. H., 2001. Confidence intervals for autoregressive coefficients near one. Journal of Econometrics 103, 155-181.

Giraitis, L., Phillips, P.C.B., 2006. Uniform limit theory for stationary autoregression. Journal of Time Series Analysis, 27, 51-60. Corrigendum, forthcoming.

Guo, B., Phillips, P.C.B., 2001. Testing for autocorrelation and unit roots in the presence of conditional heteroskedasticity of unknown form. Unpublished working paper, Department of Economics, UC Santa Cruz.

Hall, P., Heyde, C. C., 1980. Martingale limit theory and its application. Academic Press, New York.

Hansen, B. E., 1992. Convergence to stochastic integrals for dependent heterogeneous processes. Econometric Theory 8, 489-500.

Hansen, B. E., 1999. The grid bootstrap and the autoregressive model. Review of Economic Studies 81, 594-607.

Kim, K., Schmidt, P., 1993. Unit root tests with conditional heteroskedasticity. Journal of Econometrics 59, 287-300.

Mikusheva, A., 2007. Uniform inferences in autoregressive models. Econometrica 75, 1411-1452.

Müller, U. K., Elliott, G., 2003. Tests for unit roots and the initial condition. Econometrica 71, 1269-1286.

Nankervis, J. C., Savin, N. E., 1996. The level and power of the bootstrap t-test in the trend model with AR(1) errors. Journal of Business and Economic Statistics 14, 161-168.

Park, J. Y., 2002. Weak unit roots. Unpublished working paper, Department of Economics, Rice University.

Phillips, P. C. B., 1987. Towards a unified asymptotic theory for autoregression. Biometrika 74, 535-547.

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

Seo, B., 1999. Distribution theory for unit root tests with conditional heteroskedasticity. Journal of Econometrics 91, 113-144.

Stock, J. H., 1991. Confidence intervals for the largest autoregressive root in U. S. macroeconomic time series. Journal of Monetary Economics 28, 435-459.

White H., 1984. Asymptotic theory for Econometricians. Academic Press, New York.