A Heteroskedasticity Robust Breusch-Pagan Test for Contemporaneous Correlation in Dynamic Panel Data Models

Andreea Halunga[a], Chris D. Orme[b], Takashi Yamagata[c,*]
[a] Department of Economics, University of Bath
[b] Economics, University of Manchester
[c] Department of Economics and Related Studies, University of York

March 17, 2015

Abstract

This paper proposes a heteroskedasticity-robust Breusch-Pagan test of the null hypothesis of zero cross-section (or contemporaneous) correlation in linear panel data models, without necessarily assuming independence of the cross-sections. The procedure allows for either fixed, strictly exogenous and/or lagged dependent regressor variables, as well as quite general forms of both non-normality and heteroskedasticity in the error distribution. The asymptotic validity of the test procedure is predicated on the number of time series observations, T, being large relative to the number of cross-section units, N, in that: (i) either N is fixed as $T \to \infty$; or, (ii) $N^2/T \to 0$, as both T and N diverge, jointly, to infinity. Given this, it is not expected that asymptotic theory would provide an adequate guide to finite sample performance when T/N is "small". Because of this we also propose, and establish asymptotic validity of, a number of wild bootstrap schemes designed to provide improved inference when T/N is small. Across a variety of experimental designs, a Monte Carlo study suggests that the predictions from asymptotic theory do, in fact, provide a good guide to the finite sample behaviour of the test when T is large relative to N. However, when T and N are of similar orders of magnitude, discrepancies between the nominal and empirical significance levels occur as predicted by the first-order asymptotic analysis. On the other hand, for all the experimental designs, the proposed wild bootstrap approximations do improve agreement between nominal and empirical significance levels, when T/N is small, with a recursive-design wild bootstrap scheme performing best, in general, and providing quite close agreement between the nominal and empirical significance levels of the test even when T and N are of similar size. Moreover, in comparison with the wild bootstrap "version" of the original Breusch-Pagan test (Godfrey and Yamagata, 2011) our experiments indicate that the corresponding version of the heteroskedasticity-robust Breusch-Pagan test appears reliable. As an illustration, the proposed tests are applied to a dynamic growth model for a panel of 20 OECD countries.

1 Introduction

In a linear panel data model, with exogenous regressors and Zellner's (1962) Seemingly Unrelated Regression Equation (SURE) structure, a Lagrange multiplier (LM) test to

*Corresponding author. Department of Economics & Related Studies, University of York, Heslington, York YO10 5DD, United Kingdom. Email: takashi.yamagata@york.ac.uk.

detect cross-sectional dependence was proposed by Breusch and Pagan (1980) and is now a commonly employed diagnostic tool of applied workers. This test is based on the average of the squared pair-wise sample correlation coefficients of the residuals and is applicable when N is fixed and $T \to \infty$; i.e., when N is small relative to a large T. However, as pointed out in, for example, Pesaran (2004) and Pesaran, Ullah, and Yamagata (2008), the LM (henceforth, Breusch-Pagan) test based upon asymptotic critical values from the relevant χ^2 distribution can suffer from serious size distortion when N/T is not small.

In view of this, one area of research has focused on cross-section dependence tests for large T and/or N panels. Frees (1995) has proposed a "distribution free" version of the Breusch-Pagan test based on squared pair-wise Spearman sample rank correlation coefficients of the regression residuals. Pesaran (2004, 2012) proposes a, so-called, CD test based on the average pair-wise sample correlations of residuals across the different cross-section units. The CD test statistic has very good finite sample performance under a wide class of panel data model designs. However, it will lack power when the population *average* pair-wise correlations is zero, even though underlying *individual* population pair-wise correlations are non-zero. Adopting a different strategy, Pesaran *et al* (2008) make use of analytical adjustments for each squared pair-wise sample correlation in order to correct the bias of the Breusch-Pagan statistic. These analytical adjustments are derived under the same assumptions as the original Breusch-Pagan test; i.e., normality, regressor exogeneity and homoskedasticity within cross-sections. In a similar vein, Baltagi, Feng, and Kao (2012) have proposed an (asymptotic) bias-correction of Breusch-Pagan test statistic, based on the $\sqrt{NT}$ consistent Fixed Effect estimator and present Monte Carlo results which suggest that their test behaves well even when T is smaller than N; Juhl (2011) considers a similar approach. Relaxing normality and regressor exogeneity, Sarafidis, Yamagata, and Robertson (2009) propose a test for cross-sectional dependence based on Sargan's difference test for over-identifying restrictions in a dynamic panel data model, but again assuming homoskedasticity within each cross section and under a slope homogeneity assumption. Relaxing the within cross-section homoskedasticity assumption, but still maintaining exogenous regressors, Godfrey and Yamagata (2011) recently advocated a wild bootstrap[1] version of the original Breusch-Pagan test in order to address the large N/T (small T/N) issue. The Monte Carlo evidence presented by Godfrey and Yamagata (2011) suggests that such a test can provide quite reliable inferences.

However, the slope homogeneity assumption of Sarafidis *et al.* (2009), Baltagi *et al.* (2010) and Juhl (2011) can be restrictive in macroeconometric applications: see Haque, Pesaran, and Sharma (1999), Bassanini and Scarpetta (2002), amongst others. For the case of a dynamic panel data model, Pesaran and Smith (1995) demonstrate that ignorance of the heterogeneity, in general, renders equation by equation OLS regression inconsistent when regressors are serially correlated, even when N and T are large. Therefore, it is important to allow slope heterogeneity in this type of models, unless there is evidence of slope homogeneity; see Swamy (1970), Pesaran and Yamagata (2008) for the slope heterogeneity tests.

In some situations, the OLS estimator may not be consistent under cross-sectional correlation. Suppose that the cross section correlation is stemmed from error factor structure, such that $y_{it} = \alpha_i + x'_{it}\beta_i + \lambda_i y_{it-1} + u_{it}$, $u_{it} = \gamma_i f_t + \varepsilon_{it}$, where $\varepsilon_{it} \sim i.n.d.(0, \sigma_t^2)$, so that $E(u_{it}u_{jt}) = \gamma_i\gamma_j E(f_t^2)$. If x_{it} is a linear function of f_t, x_{it} can be correlated with u_{it}. If f_t is serially correlated, it can be $E(y_{it-1}u_{it}) \neq 0$. Therefore, it is important to detect cross-sectional correlation. Even when the OLS estimator is consistent, under

[1] See, for example, Wu (1986), Liu (1988), Mammen (1993), Davidson and Flachaire (2008), in the context of the classical linear regression model.

time-series heteroskedasticity, the relative efficiency of the conventional feasible generalized least square (FGLS) estimator (of SUR approach) over the OLS estimator may not be guaranteed. Furthermore, under time-series heteroskedasticity, inference based on FGLS estimation may not be reliable. Therefore, under potential time-series heteroskedasticity, it would be recommended to use a robust cross section dependence test, which is proposed in this paper. To our knowledge there is no such test available in the literature to date. If the null is not rejected by the test, it would be more confidently concluded that the rejection is not due to the heteroskedasticity, and the OLS estimation would be preferred. If the null is rejected, then suitable estimation procedure should be pursued.

This paper makes two contributions which are distinct from Godfrey and Yamagata (2011). First, it proposes new asymptotically pivotal heteroskedasticity robust Breusch-Pagan tests that allow for fixed, strictly exogenous and lagged dependent regressor variables as well as quite general forms of both non-normality and heteroskedasticity, in the linear model error distribution. (Juhl, 2011, proposes an alternative test which allows for cross-section heteroskedasticity, but requiring time-series homoskedasticity.) The last point is particularly pertinent because the modern approach in applied research is to implement inference by employing some heteroskedasticity robust variance-covariance estimator. It emerges from this analysis that the original Breusch-Pagan test and its standardised version suggested by Pesaran (2004, 2012) will asymptotically over reject, under the null, in the presence of heteroskedasticity, except when the squared errors are (asymptotically) contemporaneously uncorrelated. Our Monte Carlo study reveals rejection rates of 100%, under the null, even when T is large. The asymptotic distribution of the new statistic is first derived under the assumption that $T \to \infty$ with N fixed and, then, an asymptotically valid normalised statistic is also developed when both T and N jointly diverge to infinity, but requiring $N^2/T \to 0$ in order to eliminate an asymptotic bias in the resultant limiting distribution. However, as is well known, asymptotic theory can provide a poor approximation to actual finite sample behaviour; specifically in this case, and as noted previously, when N/T is not small, and our Monte Carlo study does indeed reveal severe size distortions when T and N are of comparable magnitude.

Second, this paper describes three asymptotically valid wild bootstrap procedure schemes which are employed in order to provide closer agreement between the desired nominal and the empirical significance level of a test procedure. For all experiments, the recursive-design wild bootstrap performs the best among the bootstrap schemes even when T and N are of similar magnitude. Moreover, in comparison with the wild bootstrap "version" of the (normalised) original Breusch-Pagan test (Godfrey and Yamagata, 2011) the corresponding (normalised) version of the heteroskedasticity-robust Breusch-Pagan test is more reliable with this wild bootstrap scheme, performing the best under the null in all experiments. Note also that the recursive-design wild bootstrap, employed in this paper, is asymptotically justified under less restrictive assumptions than those imposed by Goncalves and Kilian (2004) and Godfrey and Tremayne (2005), which rule out certain asymmetric conditional heteroskedastic error processes. The reason being that Goncalves and Kilian (2004) wish to show that the recursive wild bootstrap provides consistent estimates of heteroskedasticity-robust standard errors. However, the additional restrictive assumption they employ is not required to *directly* prove the asymptotic validity of the recursive design wild bootstrap when used in conjunction with heteroskedasticity-robust t-ratios (see Halunga (2005)). Thus, our assumptions still provide the basis for asymptotically valid inferences for regression parameters, by employing this wild bootstrap scheme, under zero cross section correlation.

Finally, it has been traditional when developing tests for cross-section dependence that the actual null hypothesis under test is one of zero contemporaneous correlation among cross sections (i.e., individuals, households, firms, countries, etc.) the failure of which, of course, is consistent with contemporaneous dependence; see, for example, the survey by Moscone and Tosetti (2009). However, zero contemporaneous correlation does not, necessarily, imply contemporaneous independence. Nonetheless, virtually all previous tests of this null hypothesis that have been proposed in the literature have maintained the stronger assumption of independence. In this paper, such independence is not assumed.

The rest of the paper is organised as follows, with all proofs relegated to the Appendix. Section 2 introduces the notation and assumptions which afford the subsequent asymptotic analysis. Section 3 establishes the limit distribution of the new test statistic when $T \to \infty$ and N is fixed and Section 4 establishes the limit distribution of the new statistic when both $(N, T) \to \infty$. Section 5 describes the wild bootstrap tests, which are applicable to both the new heteroskedasticity robust Breusch-Pagan test and the original version. Section 6 reports the results of a small Monte Carlo study designed to shed light on the finite sample reliability of the various test procedures and Section 7 provides a simple empirical application. Finally, Section 8 concludes.

2 The Model, Notation & Assumptions

In this paper, we allow for an Autoregressive Distributed Lag (ADL) heterogeneous panel data model structure. In particular, if i indexes the cross-section observations and t the time series observations, then the following model is assumed

$$\phi_i(L)y_{it} = w'_{it}\theta_i + u_{it}, \;\; i = 1, ..., N, \; t = 1, ..., T, \tag{1}$$

where $\{y_{i,-p+1}, ..., y_{i0}, y_{i1}, ..., y_{iT}, w_{i1}, ..., w_{iT}\}$, $i = 1, ..., N$, are the sample data and $\phi_i(L) = 1 - \phi_{i1}L - \phi_{i2}L^2 - ... - \phi_{ip}L^p$, $\phi_{ip} \neq 0$, has all roots lying outside the unit circle, for all i, with p, the lag length, known, finite and common across i, and $\|\theta_i\| < \infty$. The M regressors, $w'_{it} = \{w_{itl}\}$, $l = 1, ...M$, are strictly exogenous, with $w_{it1} = 1$, for all i and t; the errors, u_{it}, have zero mean for all i and t; and, $\{w'_{it}, u_{it}\}$ satisfy the regularity conditions discussed below.

Stacking the observations, $t = 1, ..., T$, per cross-section we write (1) as

$$y_i = X_i\beta_i + u_i \tag{2}$$

$\beta'_i = (\theta'_i, \phi'_i)$, $\phi'_i = (\phi_{i1}, ..., \phi_{ip})$, where $y_i = \{y_{it}\}$, $(T \times 1)$, $X_i = (W_i, Y_i)$ is $(T \times M + p)$ and has rows x'_{it}, W_i has rows $w'_{it} = \{w_{itl}\}$, Y_i has rows $Y'_{i,t-1} = \{y_{i,t-q}\}$, $q = 1, ..., p$, and $u_i = \{u_{it}\}$, $(T \times 1)$. The Ordinary Least Squares estimator of β_i, in (2), is given by

$$\hat{\beta}_i = (X'_iX_i)^{-1} X'_iy_i, \;\; i = 1, ..., N.$$

Zero contemporaneous (or cross-section) correlation is equivalent to the null hypothesis of $H_0 : E[u_iu'_j] = 0$, for all $i \neq j$, or $H_0 : E[u_{it}u_{jt}] = 0$ for all $t = 1, ..., T$ and all $i \neq j$. It is common practice, in the literature, for tests of $H_0 : E[u_{it}u_{jt}] = 0$ to be constructed under the stronger assumption of contemporaneous independence. The asymptotic validity of the test procedure proposed in this paper does not rely on such a strong assumption. Rather, a weaker set of conditions are invoked which specify various quantities of interest to be martingale differences.

The following assumptions are made in which $\mathcal{F}_{NT,t-1}$ is the sigma field generated by: (i) lagged values of y_{it} (i.e., $\{y_{i,t-k}\}$, $i = 1, ..., N$, $k = 1, 2, ...$); and, (ii) current and lagged values of any strictly exogenous variables, $i = 1, ..., N$, including $w_{i,t-k}$, $k = 0, 1, 2, ...$, and possibly other strictly exogenous variables as well; see, for example, White (2001, p.59).

Uniformly over $i = 1, ..., N$, the following hold:

Assumption A1: $\{w'_{it}\}$ is a mixing sequence, with either ϕ of size $-\eta/(2\eta-1)$, $\eta \geq 1$, or α of size $-\eta/(\eta-1)$, $\eta > 1$.

Assumption A2:

(i) $E[u_{it}w_{i,t+k}|\mathcal{F}_{NT,t-1}] = 0$, almost surely, for any $k \geq 0$ and all t;

(ii) $E\left[u_{it}^2|\mathcal{F}_{NT,t-1}\right] = \sigma_{it}^2$, almost surely, for all t;

(iii) $\text{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}\left\{\sigma_{it}^2 - E[u_{it}^2]\right\} = 0$;

(iv) $E|w_{itl}|^{2\kappa+\delta} \leq \Delta < \infty$, where $\kappa = \max[2, \eta]$, for some $\delta > 0$, and all $t = 1, ..., T$, $l = 1, ..., M$;

(v) $E|u_{it}|^{4+\delta} \leq \Delta < \infty$ for some $\delta > 0$, and all $t = 1, ..., T$.

Assumption A3:

(i) $E(W_i'W_i/T) = \frac{1}{T}\sum_{t=1}^{T}E[w_{it}w'_{it}]$ is uniformly positive definite (i.e., positive definite for all T sufficiently large).

(ii) $E(u_i'u_i/T) = \frac{1}{T}\sum_{t=1}^{T}E[u_{it}^2]$ is uniformly positive.

For all $1 \leq i < j = 2, ..., N$ the following holds:

Assumption A4:

(i) $E[u_{it}u_{jt}|\mathcal{F}_{NT,t-1}] = 0$, almost surely, for all t;

(ii) $E[u_{it}^2u_{jt}^2|\mathcal{F}_{NT,t-1}] = \tau_{ijt}^2$, almost surely, for all t;

(iii) $\text{plim}_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\left\{\tau_{ijt}^2 - E[u_{it}^2u_{jt}^2]\right\} = 0$;

(iv) $\omega_{ij,T} = \frac{1}{T}\sum_{t=1}^{T}E[u_{it}^2u_{jt}^2]$ is uniformly positive, such that for T sufficiently large $\inf_{i,j}\omega_{ij,T} > K > 0$;

(v) $E[u_{it}u_{jt}u_{ht}u_{kt}|\mathcal{F}_{NT,t-1}] = 0$, almost surely, for $i < j < k$, $i \leq h < k$, and for all t.

Note that the above entail uniform bounds (in both i and t) on certain moments of u_{it} and w_{it}. In addition, Assumption A1 allows w_{it} to contain fixed or random (but strictly exogenous) regressors. Assumption A2 is somewhat weaker than allowing the errors to be serially independent (although they are still uncorrelated). Assumption A2(i) follows from the strict exogeneity assumption on w_{it} and, together with Assumption A2(v) and the fact that $w_{it1} = 1$ for all t, it implies that $\{u_{it}, \mathcal{F}_t\}$ is a martingale difference sequence (m.d.s).[2] Assumptions A2(ii) and (iii) also allow for general (conditional or unconditional) heteroskedasticity (with σ_{it}^2 possibly varying across cross-sections and through time). A wide class of models for the variance are allowed that include cross-sectional heterogeneity, volatility that evolves over time such as GARCH type models, trending volatility, break and smooth transition shifts in variance. Notice, that we do not need asymptotic normality

[2] This formulation is similar to that employed, for example, by Weiss (1986).

of $\sqrt{T}(\hat{\beta}_i - \beta_i)$ in order to justify the asymptotic validity of the test procedure in this paper; in contrast to the assumption of Godfrey and Yamagata (2011). Assumption A4 permits the derivation of the robust test procedure for cross-section correlation (Lemma 1 and Theorem 1 below). Assumption A4(i) states that u_{it} and u_{jt} are uncorrelated, $i \neq j$, whilst A4(v) requires that all distinct pairs $\{u_{it}u_{jt}\}$ and $\{u_{ht}u_{kt}\}$ are uncorrelated, $i \neq j$ and $h \neq k$. These two assumptions could be replaced by the much stronger assumption that the $\{u_{it}\}$ are independent, which we wish to resist.

3 Test Statistics and Limit Distributions: $T \to \infty$, fixed N

The commonly used Breusch-Pagan test statistic is

$$BP_T = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij,T}^2 \tag{3}$$

where[3]

$$\hat{\rho}_{ij,T} = \frac{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \hat{u}_{it}\hat{u}_{jt}}{\sqrt{\left\{\frac{1}{T}\sum_{t=1}^{T} \hat{u}_{it}^2\right\}\left\{\frac{1}{T}\sum_{t=1}^{T} \hat{u}_{jt}^2\right\}}}.$$

As noted, for example, by Moscone and Tosetti (2009), under (1), cross-section independence, but homoskedasticity across the time dimension, it can be shown that $BP_T \xrightarrow{d} \chi_{\upsilon}^2$, for fixed N, as $T \to \infty$, where $\upsilon = \frac{1}{2}N(N-1)$. Given Theorem 1, below, and under Assumption A4(i) and (v), rather than full independence, this remains true. However, this will not be the case, in general, when there is heteroskedasticity across the time dimension. In these circumstances, the use of BP_T could lead to asymptotically invalid inferences. (This was also recently pointed out by Godfrey and Yamagata (2011), but in the context of a static heterogeneous panel.) Therefore the availability of a test procedure that is robust to more general heteroskedasticity would appear desirable. Such a robust statistic is defined as

$$RBP_T = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\gamma}_{ij,T}^2 \tag{4}$$

where

$$\hat{\gamma}_{ij,T} = \frac{\sum_{t=1}^{T} \hat{u}_{it}\hat{u}_{jt}}{\sqrt{\sum_{t=1}^{T} \hat{u}_{it}^2\hat{u}_{jt}^2}} = \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \hat{u}_{it}\hat{u}_{jt}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T} \hat{u}_{it}^2\hat{u}_{jt}^2}}. \tag{5}$$

Allowing for heteroskedasticity across both the cross-section and time dimension, Assumption A4(iv) and a straightforward application of White (2001, Corollary 5.26, p.135), yields

$$\gamma_{ij,T} = \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_{it}u_{jt}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T} E[u_{it}^2 u_{jt}^2]}} \xrightarrow{d} N(0,1),$$

a result which motivates the construction of (robust test statistic) RBP_T given in (4).

We are now in a position to establish the following Theorem, which justifies the construction of a robust version of BP_T, as detailed in the subsequent Corollary.

[3] We have dropped the T subscript on $\hat{\rho}_{ij}$ for notational simplicity.

Theorem 1 *Under Assumptions A1-A4, we have, for all $i \neq j$, and as $T \to \infty$, and fixed N, $\hat{\gamma}_{ij,T} - \gamma_{ij,T} = o_p(1)$, so that*

$$\hat{\gamma}_{ij,T} \xrightarrow{d} N(0,1).$$

Corollary 1 *Under Assumptions A1-A4, and as $T \to \infty$, and fixed N,*

$$RBP_T = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\gamma}_{ij,T}^2 \xrightarrow{d} \chi_\upsilon^2, \quad \upsilon = \frac{1}{2} N(N-1).$$

From Theorem 1 the asymptotic behaviour of BP_T can be inferred, under certain forms of heteroskedasticity. In particular, under cross-sectional heteroskedasticity only, it is easily verified that $\hat{\rho}_{ij,T} - \hat{\gamma}_{ij,T} = o_p(1)$, so that BP_T remains asymptotically valid, as noted earlier. However, in general, we have (under our assumptions)

$$\begin{aligned} \hat{\rho}_{ij,T} &= \left\{ \sqrt{\frac{\frac{1}{T}\sum_{t=1}^T \hat{u}_{it}^2 \hat{u}_{jt}^2}{\frac{1}{T}\sum_{t=1}^T \hat{u}_{it}^2 \frac{1}{T}\sum_{t=1}^T \hat{u}_{jt}^2}} \right\} \hat{\gamma}_{ij,T} \\ &= \left\{ \sqrt{\frac{\frac{1}{T}\sum_{t=1}^T E[u_{it}^2 u_{jt}^2]}{\frac{1}{T}\sum_{t=1}^T E[u_{it}^2] \frac{1}{T}\sum_{t=1}^T E[u_{jt}^2]}} \right\} \hat{\gamma}_{ij,T} + o_p(1), \end{aligned}$$

so that, asymptotically at least, $\hat{\rho}_{ij,T} - \hat{\gamma}_{ij,T} = o_p(1)$ if and only if u_{it}^2 and u_{jt}^2 are (asymptotically) contemporaneously uncorrelated. For illustrative purposes, suppose $u_{it} = \sigma_{it}\varepsilon_{it}$, where the ε_{it} are zero mean and unit variance, independently and identically distributed (i.i.d.), random variables. In this context, for example, with a one-break-in-volatility model which specifies $\sigma_{it}^2 = \sigma_{i1}^2$ for $t = 1, ..., T_1 < T$ and $\sigma_{it}^2 = \sigma_{i2}^2 > \sigma_{i1}^2$ for $t = T_1 + 1, ..., T$, u_{it}^2 and u_{jt}^2 will be (asymptotically), positively contemporaneously correlated, so that $\hat{\rho}_{ij} > \hat{\gamma}_{ij,T}$, in probability. Under the null hypothesis of $H_0 : E[u_{it}u_{jt}] = 0$, this will lead to over-rejection, asymptotically, for a test procedure which employs BP_T in conjunction with χ_υ^2 critical values. A qualitatively similar conclusion emerges for a trending volatility model ("Model 2" in Cavaliere and Taylor, 2008), where $\sigma_{it} = \sigma_{i0} - (\sigma_{i1} - \sigma_{i0}) \left(\frac{t-1}{T-1}\right)$, $\sigma_{i1} > \sigma_{i0}$, since, again, u_{it}^2 and u_{jt}^2 will be (asymptotically), positively contemporaneously correlated. However, for conditional heteroskedasticity in which $\sigma_{it}^2 = E\left[u_{it}^2 | \mathcal{F}_{NT,t-1}\right]$ is a stationary process (for example, a GARCH error process) then, due to the independence of the ε_{it}, u_{it}^2 and u_{jt}^2 are (asymptotically) contemporaneously uncorrelated so that the use of BP_T with χ_υ^2 critical values is asymptotically valid. The tests designed by Juhl (2011), Baltagi, Feng and Kao (2011) and Pesaran, Ullah and Yamagata (2008) might lead to missleading inference in a similar fashion as BP_T.

Thus, there will be situations in which BP_T remains asymptotically robust. In general, though, it seems prudent to use a procedure based on a statistic, such as RBP_T, that is robust under quite general forms of (unknown) heteroskedasticity.

4 Test Statistic and Limit Distributions: $(N, T) \to \infty$

Pesaran (2004, 2012) proposed a standardised version of the BP_T test as

$$NBP_{NT} = \frac{1}{\sqrt{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\hat{\rho}_{ij}^2 - 1\right) \tag{6}$$

and under (1), cross-section independence but homoskedasticity across the time dimension, $NBP_{NT} \xrightarrow{d} N(0,1)$ as $T \to \infty$ first, followed by $N \to \infty$.

Allowing for heteroskedasticity across both the cross-section and time dimension, a standardised version of RBT_T proposed in the previous section is defined as

$$NRBP_{NT} = \frac{1}{\sqrt{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\hat{\gamma}_{ij,T}^2 - 1\right). \tag{7}$$

The limiting distribution of the test defined in (7) is obtained by still maintaining zero cross-section correlation under the null rather than the stronger assumption of cross-section independence as it is commonly assumed in the current literature. Specifically, the following assumptions are in addition to/or strengthen the previous Assumption A, and are made in order to derive the $O(1)$ limiting distribution of the new statistic in (7):

Assumption B1: $(N,T) \to \infty$ jointly, such that $N^2/T \to 0$.

Assumption B2: For some $\delta > 0$,

(i) $\sup_{i \neq j} E\left|\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(u_{it}^2 u_{jt}^2 - E[u_{it}^2 u_{jt}^2]\right)\right| \leq \Delta < \infty$;

(ii) $\sup_{t,i} E\left|u_{it}\right|^{8+\delta} \leq \Delta < \infty$;

(iii) $\sup_{i \neq j} \frac{1}{T} \sum_{t=2}^{T} \sum_{s \neq r}^{t-1} \left|cov\left(u_{it}^2 u_{jt}^2, u_{is}u_{js}u_{ir}u_{jr}\right)\right| \leq \Delta < \infty$;

(iv) $\sup_{i \neq j} \frac{1}{T} \sum_{t \neq s}^{T} \left|cov\left(u_{it}^2 u_{jt}^2, u_{is}^2 u_{js}^2\right)\right| \leq \Delta < \infty$;

(v) $E\left[u_{it}u_{jt}u_{lt}u_{mt}u_{pt}u_{qt}u_{ht}u_{nt}|\mathcal{F}_{N,t-1}\right] = k_{ijlmpqhn}$ does not depend on t and $N^{-4} \sum_{i<j} \sum_{l<m} \sum_{p<q} \sum_{h<n} |k_{ijlmpqhn}| \leq \Delta < \infty$

(vi) $\frac{1}{T^2} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \sum_{r=1}^{t-1} \left(u_{is}u_{js}u_{ir}u_{jr} E\left[u_{it}^2 u_{jt}^2 | F_{N,t-1}\right] - E\left[u_{it}^2 u_{jt}^2 u_{is}u_{js}u_{ir}u_{jr}\right]\right) = o_p(1)$ uniformly in i, j.

Assumption B1 ensures that an asymptotic bias in the limiting distribution of $NRBP_{NT}$ in 7) disappears as T and N diverge jointly to infinity. A CLT for U-statistic of Hall (1984) or martingale difference arrays of Hall and Heyde (1980, Corollary 3.1) applies under Assumption B2. Assumption B2(iii), (iv) and (v) restrict the cross-section dependence resembling similar assumptions as in Bai (2009) with Assumption B2(iii) and (iv) being employed to establish that the asymptotic variance of $NRBP_{NT}$ is one. Assumption B2(vi) is the same as Assumption A(v) in Goncalves and Kilian (2004) for $\varepsilon_{ij,t} = u_{it}u_{jt}$ and allows for the difference between the conditional and unconditional variance of the test statistic to be asymptotically negligible.

Theorem 2 *Under Assumption A1-A4 combined with Assumptions B1-B2*

$$NRBP_{NT} = \frac{1}{\sqrt{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\hat{\gamma}_{ij,T}^2 - 1\right) \xrightarrow{d} N(0,1).$$

The method of proof is in two stages. The first stage requires the following Central Limit Theorem:

Lemma 1 *Under Assumptions A2(v), A4(i), (iv) and (v) combined with/or strengthened by Assumptions B1-B2*

$$Z_{NT} = \frac{1}{\sqrt{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\gamma_{ij,T}^2 - 1\right) \xrightarrow{d} N(0,1),$$

where

$$\gamma_{ij,T} = \frac{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_{it} u_{jt}}{\sqrt{\omega_{ij,T}}}$$

and $\omega_{ij,T} = \frac{1}{T} \sum_{t=1}^{T} E[u_{it}^2 u_{jt}^2]$.

In the second stage, Lemma 2, below, establishes that the asymptotic bias which appears in the limiting distribution of $NRBP_{NT}$ disappears as $N^2/T \to 0$. This implies that the standard normal limiting distribution approximates the limiting distribution of the statistic $NRBP_{NT}$ when T is large relative to N. Thus, in the case when $(N,T) \to \infty$ jointly, the chi-square version of the test RBP_T should do as well as its standardised version, $NRBP_{NT}$.

Lemma 2 *Under Assumptions A1-A4 combined with/or strengthened by Assumptions B1 and B2(i)*

$$NRBP_{NT} = Z_{NT} + o_p(1)$$

Armed with Lemmas 1 and 2, Theorem 2 follows immediately.

Although, Theorem 1 and Theorem 2 show that the chi-square version of the new statistic, RBP_T, and its standardised version, $NRBP_{NT}$, are asymptotically robust to general forms of heteroskedasticity, it might be anticipated that improved sampling behaviour, in finite samples, will be afforded by employing a wild bootstrap scheme. Indeed, Godfrey and Yamagata (2011) proposed the use of a wild bootstrap scheme in order to control the significance levels of the BP_T test procedure, in the presence of non-normality and unknown heteroskedasticity, under both large T and large N asymptotics. Their analysis, however, is limited to the static heterogeneous panel data model and is not based on an asymptotic pivot. In the next section, asymptotic validity of three wild bootstrap schemes is established in a dynamic heterogenous panel data model under non-normality and unknown heteroskedasticity.

5 Wild Bootstrap Procedures

The wild bootstrap tests based on either the chi-square version of RBP_T (resp., BP_T) for fixed N or the standardised version of $NRBP_{NT}$ (resp., NBP_{NT}), proposed as (N,T) diverge jointly to infinity, will deliver the same empirical size and power results, since it does not matter which asymptotic distribution is employed for the bootstrap. As a consequence, the wild bootstrap procedures considered in this section are based only on the standardised normal statistics, i.e. $NRBP_{NT}$ and NBP_{NT}, respectively.

We consider three wild bootstrap procedures, as follows.

5.1 Wild Bootstrap 1 (WB1)

This is a recursive design wild bootstrap scheme, implemented using the following steps:

1. Estimate the model by OLS to get $\hat{u}_{it}$, $i = 1, ..., N$, and construct test statistics $NRBP_T$ and NBP_T

2. (which is repeated B times)

   (a) Generate $u_{it}^* = \varepsilon_{it}\hat{u}_{it}$, where the ε_{it} are i.i.d., over i and t, with zero mean and unit variance.

   (b) Construct

   $$y_{it}^* = \hat{\beta}_i' x_{it}^* + u_{it}^*. \tag{8}$$

   Here, x_{it}^* is generated recursively, from (8), given initial values y_{it}^*, $t \leq 0$ for any regressors which are lagged dependent variables (these could be zero or sample values). Sample values of the regressors are employed in this wild bootstrap scheme for any strictly exogenous variables. Thus, for example, if $x_{it}' = (w_{it}', y_{i,t-1})$, where w_{it} is strictly exogenous, then $w_{it}^* = w_{it}$, for all i and t, $\beta_i' = (\theta_i', \phi_i)$ and choosing $y_{i0}^* = y_{i0}$ bootstrap data are generated according to

   $$\begin{aligned} y_{i1}^* &= \hat{\theta}_i' w_{i1} + \hat{\phi}_i y_{i0} + u_{i1}^* \\ y_{it}^* &= \hat{\theta}_i' w_{it} + \hat{\phi}_i y_{i,t-1}^* + u_{it}^*, \ t = 2, ..., T. \end{aligned}$$

   (c) Construct the bootstrap test statistics (B simulations)

   $$NRBP_{NT}^* = \frac{1}{\sqrt{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\hat{\gamma}_{ij,T}^{*2} - 1\right), \quad \hat{\gamma}_{ij,T}^* = \frac{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \hat{u}_{it}^* \hat{u}_{jt}^*}{\sqrt{\frac{1}{T} \sum_{t=1}^{T} \hat{u}_{it}^{*2} \hat{u}_{jt}^{*2}}} \tag{9}$$

   $$NBP_{NT}^* = \frac{1}{\sqrt{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\hat{\rho}_{ij,T}^{*2} - 1\right), \ \hat{\rho}_{ij,T}^* = \frac{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \hat{u}_{it}^* \hat{u}_{jt}^*}{\sqrt{\frac{1}{T} \sum_{t=1}^{T} \hat{u}_{it}^{*2} \frac{1}{T} \sum_{t=1}^{T} \hat{u}_{jt}^{*2}}} \tag{10}$$

   where $\hat{u}_{it}^* = y_{it}^* - x_{it}^{*\prime}\hat{\beta}_i^*$ is the OLS residual from (8).

3. Calculate the proportion of bootstrap test statistics, $NRBP_{NT}^*$ (resp., NBP_{NT}^*), from the B repetitions of Step 2c that are at least as large as the actual value of $NRBP_{NT}$ (resp., NBP_{NT}). Let this proportion be denoted by $\hat{p}$ and the desired significance level be denoted by α. The asymptotically valid rejection rule is that H_0 is rejected if $\hat{p} \leq \alpha$.

5.2 Wild Bootstrap 2 (WB2)

This is a fixed design wild bootstrap scheme which replaces (8) in the recursive design scheme with

$$y_{it}^* = \hat{\beta}_i' x_{it} + u_{it}^*$$

at stage 2b.

5.3 Wild Bootstrap 3 (WB3)

Note, from Theorem 1, $\hat{\gamma}_{ij,T} - \gamma_{ij} = o_p(1)$; i.e., $\hat{\gamma}_{ij,T}$ has the same limit distribution as it would have if β_i were known. This suggests that the following wild bootstrap procedure should work (asymptotically) at least.

1. As for WB1.
2. (which is repeated B times)
   (a) Generate $u_{it}^* = \varepsilon_{it}\hat{u}_{it}$, as in WB1(but omit step 2b in WB1).
   (b) Construct the bootstrap test statistics

$$NRBP_{NT}^* = \frac{1}{\sqrt{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\tilde{\gamma}_{ij,T}^{*2} - 1\right), \quad \tilde{\gamma}_{ij,T}^* = \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_{it}^* u_{jt}^*}{\sqrt{\frac{1}{T}\sum_{t=1}^{T} u_{it}^{*2} u_{jt}^{*2}}},$$

$$NBP_{NT}^* = \frac{1}{\sqrt{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\tilde{\rho}_{ij}^{*2} - 1\right), \; \tilde{\rho}_{ij,T}^* = \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_{it}^* u_{jt}^*}{\sqrt{\frac{1}{T}\sum_{t=1}^{T} \hat{u}_{jt}^{*2} \frac{1}{T}\sum_{t=1}^{T} \hat{u}_{jt}^{*2}}}$$

3. Calculate the proportion of bootstrap test statistics, $NRBP_{NT}^*$ (resp., NBP_{NT}^*), from the B repetitions of Step 2b that are at least as large as the actual value of $NRBP_{NT}$ (resp., NBP_{NT}). Let this proportion be denoted by $\tilde{p}$ and the desired significance level be denoted by α. The asymptotically valid rejection rule is that H_0 is rejected if $\tilde{p} \leq \alpha$.

The asymptotic validity of these wild bootstrap schemes is established in the theorem below[4] under the strengthened assumption:

Assumption B3 $E\|w_{it}\|^{4k+\delta} \leq \Delta < \infty$ where $\kappa = \max[2,\eta]$, for some $\delta > 0$, and all $i = 1,..,N, t = 1,...,T$ and $l = 1,...,M$;

Theorem 3 *Under Assumptions A1-A4 combined with/or strengthened by Assumptions B1-B3, and for all three wild bootstrap designs, WB1, WB2 and WB3,*

$$\sup_x |P^*(NRBP_{NT}^* \leq x) - P(NRBP_{NT} \leq x)| \xrightarrow{p} 0$$
$$\sup_x |P^*(NBP_{NT}^* \leq x) - P(NBP_{NT} \leq x)| \xrightarrow{p} 0$$

where P^* is the probability measure induced by the wild bootstrap conditional on the sample data.

Note that, even when allowing for conditional heteroskedasticity, we do not require the restrictive Assumption A' (iv') of Goncalves and Kilian (2004) to justify the recursive-design WB1, since our test criteria are asymptotically independent of $\hat{\beta}_i$. Specifically, the class of conditionally heteroskedastic autoregressive models is not restricted to the symmetric ones as in Goncalves and Kilian (2004).

Henceforth, a test procedure which employs $NRBP_{NT}$ (resp., NBP_{NT}) in conjunction with asymptotic critical values will be called an "asymptotic test", whilst the one that

[4] In the Appendix, we verify this for the recursive wild bootstrap scheme (WB1) only and, following Davidson and Flachaire (2008), with $u_{it}^* = \varepsilon_{it}\hat{u}_{it}$ where the ε_{it} are independently and identically distributed for all i and t taking the discrete values ± 1 with an equal probability of 0.5.

employs either of WB1, WB2 or WB3 will be referred to as a "bootstrap test". In order to shed light on the relevance of the preceding asymptotic analysis as an approximation to actual finite sample behaviour, the next section describes, and reports the results of, a small Monte Carlo study which investigates the sampling behaviour of the test statistics considered above under a variety of heteroskedastic error distributions, and (N, T) combinations.

6 Monte Carlo Study

Three data generating processes (DGPs) are considered: Panel autoregressive and distributed lag (ADL) models, with strictly exogenous regressors, and pure panel autoregressive (AR) models.

6.1 Monte Carlo Design

6.1.1 DGP1

The first data generating process considered is a dynamic panel $ADL(1,0)$ model, which is specified by

$$\begin{aligned} y_{it} &= \theta_{i1} + \theta_{i2} z_{it} + \phi_i y_{i,t-1} + u_{it} \\ &= \theta_i' w_{it} + \phi_i y_{i,t-1} + u_{it}, \quad i = 1, 2, ..., N \text{ and } t = -49, -48, ..., T \end{aligned} \tag{11}$$

with $\theta_{i1} \sim$ i.i.d. $N(0,1)$, $\theta_{i2} = 1 - \phi_i$, $\phi_i \sim$ i.i.d. Uniform$[0.4, 0.6]$, and the z_{it} are generated for $(N = 5, T = 25)$ as independent random draws from the standard lognormal distribution. This block of regressor values is then reused as necessary to build up data for the other combinations of (N, T). $y_{i,-50} = 0$, and first 49 values are discarded. The error term is generated as

$$u_{it} = \sigma_{it} \varepsilon_{it}, \; i = 1, 2, ..., N \text{ and } t = -49, -48, ..., T \tag{12}$$

and

$$\varepsilon_{it} = \sqrt{1 - \rho^2} \xi_{it} + \rho \zeta_t \tag{13}$$

where $\xi_{it} \sim$ i.i.d. $(0,1)$ independently of $\zeta_t \sim$ i.i.d. $(0,1)$. Thus, $corr\,(u_{it}, u_{jt}) = \rho$, a constant in this case. For estimating significance levels, the value of ρ is set to zero, whilst power is investigated using $\rho = 0.2$, which provides a useful range of experimental results. Two distributions are used to obtain the i.i.d. standardised errors for ξ_{it} and ζ_t: the standard normal distribution and the chi-square distribution with six degrees of freedom (χ^2_6), with the latter being employed to provide evidence on the effects of skewness. In particular, with a coefficient of skewness greater than 1, it is heavily skewed, according to the arguments of Ramberg, Tadikamalla, Dudewicz, and Mykytka (1979).

Five models for σ_{it} are considered, all of which satisfy, in particular, Assumption A2(v). First, there is homoskedasticity, denoted HET0, with $\sigma_{it} = 1$ for all t. Second, a one-break-in-volatility model, henceforth HET1, is employed with $\sigma_{it} = 0.8$ for $t = 1, 2, ..., m = \lfloor T/2 \rfloor$ and $\sigma_{it} = 1.2$ for $t = m, m+1, ..., T$, where $\lfloor A \rfloor$ is the largest integer part of A. Third, HET2 is a trending volatility model, with $\sigma_{it} = \sigma_0 - (\sigma_1 - \sigma_0)\left(\frac{t-1}{T-1}\right)$; see "Model 2" in Cavaliere and Taylor (2008), where $\sigma_0 = 0.8$ and $\sigma_1 = 1.2$. Fourth, HET3 is a conditional heteroskedasticity scheme, with $\sigma_{it} = \sqrt{\exp(cz_{it})}$, $t = 1, ..., T$; this sort of skedastic function is discussed in Lima, Souza, Cribari-Neto, and Fernandes (2009). The value of c in HET3 is chosen to be 0.4; so that $\max(\sigma^2_{it}) / \min(\sigma^2_{it})$, which is a well-known measure of the strength of heteroskedasticity, is 7.9. For HET0-HET3, $\sigma_{it} = 1$ for $t = -49, ..., 0$. Finally, we consider a generalized autoregressive conditional heteroskedasticity, GARCH(1,1) model, denoted HET4, where

$$\sigma_{it}^2 = \delta + \alpha_1 u_{i,t-1}^2 + \alpha_2 \sigma_{i,t-1}^2,\ t = -49, -48, ..., T. \tag{14}$$

Following Godfrey and Tremayne (2005), the value of parameters are chosen to be $\delta = 1$, $\alpha_1 = 0.1$ and $\alpha_2 = 0.8$.

6.1.2 DGP2

The second data generating process considered is a model with strictly exogenous regressors, specified by

$$y_{it} = \beta_{i1} + \beta_{i2} z_{it} + u_{it} \tag{15}$$

$$= \beta_i' w_{it} + u_{it},\ i = 1, 2, ..., N \text{ and } t = 1, 2, ..., T, \tag{16}$$

where $\beta_{i1} \sim$ i.i.d. $N(0, 1)$, $\beta_{i2} \sim$ i.i.d. Uniform$[0.9, 1.1]$ and the z_{it} are generated for $(N = 5, T = 25)$ as independent random draws from the standard lognormal distribution. Again, this block of regressor values is then reused as necessary to build up data for the other combinations (N, T).

The error term in (15) is written as

$$u_{it} = \sigma_{it} \varepsilon_{it},\ i = 1, 2, ..., N \text{ and } t = 1, 2, ..., T. \tag{17}$$

The three distributions of ε_{it} and the five models for σ_{it} are considered as before.

6.1.3 DGP3

The third data generating process considered is a dynamic panel $AR(1)$ model, which is specified by

$$y_{it} = \theta_i (1 - \phi_i) + \phi_i y_{it-1} + u_{it}, i = 1, 2, ..., N \text{ and } t = -49, -48, ..., T. \tag{18}$$

with $\theta_i \sim$ i.i.d. $N(0, 1)$, $\phi_i \sim$ i.i.d. Uniform$[0.4, 0.6]$, $y_{i,-49} = 0$, and first 49 values are discarded. The error term is written as

$$u_{it} = \sqrt{1 - \phi_i^2} \sigma_{it} \varepsilon_{it},\ i = 1, 2, ..., N \text{ and } t = -49, -48, ..., T. \tag{19}$$

The three distributions of ε_{it} and the five models for σ_{it} are considered as before.

All combinations of $N = 5, 10, 25, 50, 100$ and $T = 25, 50, 100, 200$ are considered. The sampling behaviour of the tests are investigated using 2000 replications of sample data and 200 bootstrap samples, employing a nominal 5% significance level.

6.2 Monte Carlo Results

Before looking at the results from the Monte Carlo study, it is important to define criteria to evaluate the performance of the different tests considered. Given the large number of replications performed, the standard asymptotic test for proportions can be used to test the null hypotheses that the true significance level is equal to its nominal value. In these experiments, this null hypothesis is accepted (at the 5% level) for estimated rejection frequencies in the range 4% to 6%. In practice, however, what is important is not that the significance level of the test is identical to the chosen nominal level, but rather that the true and nominal rejection frequencies stay reasonably close, even when the test is only approximately valid. Following Cochran's (1952) suggestion, we shall regard a test as being robust, relative to a nominal value of 5%, if its actual significance level is between 4.5% and 5.5%. Considering the number of replications used in these experiments, estimated rejection frequencies within the range 3.6% to 6.5% are viewed as providing evidence consistent with the robustness of the test, according to this definition.

To economize on space, and as the results for three DGPs are qualitatively similar, the discussion below focuses on the results in the case of dynamic $ADL(1,0)$ model (DGP1), since this nests the other two models and can thus be regarded as the most general one. The experimental results, in this case, under the various heteroskedastic schemes and error distributions are reported in Tables 1 to 5. We summarise, first, the finite sample behaviour of the asymptotic tests before reporting that of the bootstrap tests.

[INSERT Table 1 HERE]

Under the null, with homoskedastic errors (reported in Table 1, $H_0 : E[u_{it}u_{jt}] = 0$), the rejection frequencies of the asymptotic RBP_T and BP_T tests and the normalised versions, $NRBP_{NT}$ and NBP, respectively, are in the main close to the nominal significance level of 5% when N/T is "small", less than 0.5, although BP_T and NBP_{NT} are slightly oversized when $N = 10$ and $T = 25$. Under standard normal errors, with the exception of the case when the empirical size of RBP_T is 3.3% for $N = 5$ and $T = 25$, RBP_T performs slightly better than the normalised test $NRBP_{NT}$. It can also be noted that the BP_T chi-square test performs better than its standardised version NBP_{NT} under both types of errors. When $N/T = 0.5$, slight over-rejections occur for all tests with the empirical sizes being in the range $7.4\% - 9\%$. However, when N/T is not "small", being greater than 0.5, severe distortions can occur. For example, when $N = 100$, BP_T rejection rates are 86.8% and 36.1% for $T = 25$ and $T = 50$, respectively. The possibility of such size distortions, when N/T is not "small", has been pointed out by Pesaran *et al* (2008). Even the normalised tests, NBP_{NT} and $NRBP_{NT}$, do suffer from such distortions since these tests require that $N^2/T \to 0$ in order for an asymptotic bias in their limiting distribution to disappear. Similar patterns are revealed under asymmetric errors as well. A comparison of their rejection frequencies under $H_A : E[u_{it}u_{jt}] = 0.2$, reveals similar power properties under homoskedastic normal and χ^2_6 errors. However, the power of the asymptotic RBP_T and $NRBP_{NT}$ tests is slightly lower than that of the corresponding asymptotic BP_T and NBP_{NT} tests under χ^2_6 errors. For example, with $N = 5$ (resp., $N = 10$) and $T = 100$, the empirical power of $NRBP_{NT}$ is 18.1% (resp., 36.4%) compared with 24.6% (resp., 45.7%) for NBP_T.

[INSERT Tables 2 - 5 ABOUT HERE]

The results obtained when the errors are heteroskedastic (Tables 2 - 5), show that the asymptotic RBP_T and $NRBP_{NT}$ tests again exhibit close agreement, between nominal and empirical significance levels across both error distributions, when N/T is small. The chi-square test RBP_T performs in general better than $NRBP_{NT}$ when N/T is small, except for the case when $N = 5$ and $T = 25$, when RBP_T is slightly undersized. In fact, the results are qualitatively similar to those obtained with homoskedastic errors, with severe distortions apparent when N/T is not small. By contrast, and consistent with the analyses in Sections 3 and 4, the asymptotic BP_T and NBP_{NT} tests tend to over-reject the null hypothesis significantly even when N/T is small, except for GARCH errors (Table 5). Moreover, for all results in Tables 2 - 5, the rejection rates for BP_T are less than those for NBP_{NT}.

For example, when $T = 200$, and under the one-break-in-volatility heteroskedastic scheme (HET1, reported in Table 2) the rejection frequencies for the asymptotic NBP_T (resp., BP_T) tests, under normal errors, range from $12.1\% - 100\%$ ($9.5\% - 100\%$) whereas for the $NRBP_{NT}$ (resp., RBP_T) range from $6.0\% - 10.3\%$ ($4.7\% - 10.1\%$). Similar pattern across the tests is revealed for the χ^2_6 errors. For the trending volatility model, Table 3, the corresponding ranges are: $8.3\% - 86.1\%$ (resp., $6.6\% - 85.8\%$) for NBP_T (resp., BP_T) and $6.5\% - 7.5\%$ (resp., $4.4\% - 9\%$) for $NRBP_{NT}$ (resp., RBP_T). For the HET3 scheme (Table 4), these ranges are $7.5\% - 86.8\%$ (resp., $5.3\% - 86.3\%$) and $6.8\% - 13.8\%$ (resp., $5.0\% - 13.7\%$), for NBP_T (resp., BP_T) and $NRBP_{NT}$ (resp., RBP_T), respectively. There is significantly less over-rejection in the latter when N/T is small, where $\sigma^2_{it} = \exp(cz_{it})$, since the z_{it} are generated as i.i.d. random variables but held fixed in repeated samples, yielding a low (but positive) contemporaneous correlation measure between the squared errors. Under GARCH(1,1) errors, where σ^2_{it} is a stationary process, BP_T (resp., NBP_{NT}) remains asymptotically justified and exhibits close agreement, in general, between nominal and empirical significance levels across all error distributions, when N/T is small, although with more pronounced distortions, than that of RBP_T.

Turning our attention to the wild bootstrap tests, both procedures, employing $NRBP^*_{NT}$ and NBP^*_{NT}, control the significance levels much better than their asymptotic counterparts, across all models and wild bootstrap schemes. Under $H_0 : E[u_{it}u_{jt}] = 0$ and over the 135 different models investigated, the recursive-design wild bootstrap scheme WB1 performs the best among all bootstrap schemes and across all models. Specifically, when $N = 5$ and 10, there is not much to choose between the bootstrap schemes but when N increases, WB1 clearly dominates the other bootstrap schemes WB2 and WB3. Under homoskedasticity and employing WB1, NBP^*_{NT} performs slightly better than $NRBP^*_{NT}$ under χ^2_6 errors when N and T are large, as NBP^*_{NT} is more efficient. Nevertheless, under heteroskedasticity, the bootstrap heteroskedasticity-robust test $NRBP^*_{NT}$ performs better than NBP^*_{NT} across all bootstrap schemes and across all models, except for GARCH errors (Table 5) when both $NRBP^*_{NT}$ and NBP^*_{NT} are comparable. In particular, empirical size distortions occur for NBP^*_{NT} when N is large and T is small. For example, for HET1 and WB1, there is hardly any evidence of distortion in the empirical significance level, with *two* cases, for $NRBP^*_{NT}$ across both error distributions, whereas there are *thirteen* times when empirical rejections of the non-robust test NBP^*_{NT} fall outside the acceptable interval of [3.6%, 6.5%]. For WB2 under normal errors, only *once* does the empirical rejection rate fall outside the acceptable interval for $NRBP^*_{NT}$ given HET1 and HET2, whereas for NBP^*_{NT} *eight* times for HET1 and *five* times for HET2. Under HET2 with normal errors, the rejection rate for $NRBP^*_{NT}$ is 7% when $N = 100$ and $T = 25$, whereas rejection rate for NBP^*_{NT} is 11.1% for this combination of N and T. Higher rejection rates are revealed under HET3, i.e. the rejection rate for $NRBP^*_{NT}$ is 9.7% when $N = 100$ and $T = 50$, whereas the rejection rate for NBP^*_{NT} is 26.8%. Such results for NBP^*_{NT} are consistent with those found by Godfrey and Yamagata (2011), although their experiments only considered a static (not dynamic) heterogeneous panel data mode. Thus, the bootstrap test $NRBP^*_{NT}$, employing WB1, exhibits good agreement between nominal and empirical significance levels and appears more reliable than NBP^*_{NT}. With regard to power comparisons, for WB1, between $NRBP^*_{NT}$ and NBP^*_{NT}, there is not a significant difference, except that NBP^*_{NT} appears consistently more powerful under χ^2_6 errors. Note that these are not size-adjusted power results and NBP^*_{NT} has revealed higher distortions under the null. Qualitatively, the results are similar across all schemes but, as an illustration, under one-break-in-volatility model with correlated errors (Table 2), under χ^2_6 errors and for $N = 25$, the rejection rates for NBP^*_{NT} are approximately

26%, 49%, 82% and 99%, respectively for $T = 25, 50, 100$ and 200, for the recursive-design resampling scheme (WB1), whilst those of $NRBP^*_{NT}$ are 18%, 36%, 72% and 98%.

7 An empirical application

In this section we examine error cross section correlation in a dynamic growth equation following Bond *et al.* (2010). Two variables, real GDP per worker and the share of total gross investment in GDP are obtained from Penn World Table Version 7.0 (PWT 7.0). Our sample consists of 20 OECD countries ($N = 20$) with annual data covering the period 1955-2004 (50 data points).[5] In order to factor out common trending components, we transformed the log of output per worker ($lgdpw_{it}$) and the log of the investment share (lk_{it}) to the deviations from the cross section mean: namely, $\widetilde{lgdpw}_{it} = lgdpw_{it} - N^{-1}\sum_{i=1}^{N} lgdpw_{it}$ and $\widetilde{lk}_{it} = lk_{it} - N^{-1}\sum_{i=1}^{N} lk_{it}$. We statistically checked the order of integration of these variables, and the evidence suggests that $\widetilde{lgdpw}_{it}$ $lgdpw_{it}$ are $I(1)$ but $\widetilde{lk}_{it}$ are $I(0)$, which is consistent with the results given by Bond et al (2010, Table I(b)).[6]

Allowing the slope coefficients to differ across countries, the dynamic specification of the growth equation is adopted from Bond et al. (equation 10):

$$\Delta\widetilde{lgdpw}_{it} = \theta_{1i} + \theta_{2i}\widetilde{lk}_{it} + \theta_{3i}\Delta\widetilde{lk}_{it} + \theta_{4i}\Delta\widetilde{lk}_{it-1} + \phi_{1i}\Delta\widetilde{lgdpw}_{i,t-1} + \phi_{2i}\Delta\widetilde{lgdpw}_{i,t-2} + u_{it}, \tag{20}$$

$i = 1, 2, ..., N = 20$ and $t = 1, 2, ..., T = 47$. In line with our notation, this model can be written as $y_{it} = x'_{it}\beta_i + u_{it}$, where $y_{it} = \Delta\widetilde{lgdpw}_{it}$, $x'_{it} = (y_{it-1}, y_{it-2}, w'_{it})$ with $w'_{it} = (1, \widetilde{lk}_{it}, \Delta\widetilde{lk}_{it}, \Delta\widetilde{lk}_{it-1})$, and $\beta_i = (\theta_{1i}, \theta_{2i}, \theta_{3i}, \theta_{4i}, \phi_{1i}, \phi_{2i})'$.

Firstly, we applied a (time-varying) heteroskedasticity-robust version of Lagrange multiplier (LM) test for error serial correlation for each country regression, as discussed in Godfrey and Tremayne (2005). The test statistic for m^{th}-order serial correlation is defined by

$$RLM_{T,i} = \hat{u}'_i\hat{U}_i\left(\hat{U}'_iM_{xi}\hat{\Lambda}_iM_{ix}\hat{U}_i\right)^{-1}\hat{U}'_i\hat{u}_i \tag{21}$$

where $\hat{u}_i = (\hat{u}_{i1}, \hat{u}_{i2}, ..., \hat{u}_{iT})'$ is a $(T \times 1)$ residual vector, $\hat{U}_i = (\hat{u}_{i,-1}, \hat{u}_{i,-2}, ..., \hat{u}_{i,-m})$ which is a $(T \times m)$ matrix with $\hat{u}_{i,-\ell} = (\hat{u}_{i,1-\ell}, \hat{u}_{i,2-\ell}, ..., \hat{u}_{i,T-\ell})'$ being a $(T \times 1)$ vector but $\hat{u}_{i,t-\ell} \equiv 0$ for $t - \ell < 1$, $\ell = 1, 2, ..., m$, $M_{ix} = I_T - X_i\left(X'_iX_i\right)^{-1}X'_i$ with t^{th} row vector of X_i being x'_{it}, and $\hat{\Lambda}_i = diag(\hat{u}^2_{it})$. Under the null hypothesis of no error serial correlation, $RLM_{T,i}$ is asymptotically distributed as χ^2_m. The finite sample experimental results in Godfrey and Tremayne (2005) show that the use of asymptotic critical value can be unreliable but that recursive resampling wild bootstrap (our WB1) approach is reliable with good control over finite sample significance levels.[7]

[5] These OECD countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Netherland, Norway, Spain, Sweden, Switzerland, United Kingdom and United States.

[6] The values of t-bar statistics, which are the cross-sectional averages of country ADF(2) statistics with a linear trend for $\widetilde{lgdpw}_{it}$ is -1.55, and the exact 5% critical values reported Im *et al.* (2003; table 2) for $N = 20$ and $T = 50$ is -2.47. The values of similar t-bar statistics but with an intercept only for $\Delta\widetilde{lgdpw}_{it}$, $\widetilde{lk}$ and $\Delta\widetilde{lk}$ are -3.45, -2.00 and -4.71, respectively, and the exact 5% critical value is -1.85.

[7] They considered a Hausman-type test and a modified version of the LM test, but based on the finite sample results the bootstrap $RLM_{T,i}$ test or a bootstrap modified LM test is recommended. We consider the WB1 bootstrap $RLM_{T,i}$ test only, since the reported performance of these two tests by Godfrey and Tremayne (2005) was very similar and the former is computationally simpler. Note, however, that these procedures require more restrictive assumptions than those imposed in this paper.

We have applied the WB1 bootstrap $RLM_{T,i}$ test for second-order serial correlation ($m = 2$) to the model (20) and the results show that the null hypothesis of no error serial correlation cannot be rejected at the 5% significance level for all 20 OECD countries. Therefore, there is no strong evidence against a claim of no error serial correlation for all 20 OECD countries.[8]

[INSERT Table 6 HERE]

Now let us turn our attention to error cross section correlation tests. Table 6 reports the asymptotic and various bootstrap p-values of the tests. As can be seen, the asymptotic NBP_{NT} test rejects the null hypothesis at the 5% level, but our asymptotic $NRBP_{NT}$ test does not. When the bootstrap methods are applied to these tests, both have similar p-values, ranging between 10.7% to 12.8%. Therefore, based on our proposed testing approach, there is no strong evidence of contemporaneous error cross section correlation.

8 Conclusion

The paper has developed heteroskedasticity robust Breusch-Pagan tests for the null hypothesis of zero-cross section correlation in dynamic panel data models under the assumption that the number of time series observations, T, is large relative to the number of cross sections, N, with either N fixed or both N and T large; but not under an assumption cross section independence. The procedures can be employed with fixed, strictly exogenous and/or lagged dependent regressors and are (asymptotically) robust to quite general forms of non-normality and heteroskedasticity, in the error distribution, across both time and cross-section. However, when N/T is not small, the asymptotic tests reveal severe size distortions in line with the qualitative predictions from first order asymptotic theory. Wild bootstrap schemes can be used to improve the finite sample behaviour of the tests, with the recursive-design wild bootstrap scheme performing the best among the bootstrap procedures employed in our Monte Carlo study. By allowing conditional heteroskedasticity with asymmetric errors, these wild bootstrap schemes are all asymptotically valid under less restrictive assumptions than those imposed by, say, Goncalves and Kilian (2004). Across all combinations of error distributions and types of heteroskedasticity, considered in our study, the recursive-design wild bootstrap version of the new robust standardised Breusch-Pagan test ($NRBP^*_{NT}$) provides quite reliable finite sample inferences, even when N/T is not small, as hoped would be the case. Furthermore, the $NRBP^*_{NT}$ seems to be as powerful as its asymptotic counterpart $NRBP_{NT}$ (except when T is small and N is large, but $NRBP_{NT}$ is severely oversized in this case) under homoskedasticity and therefore there appears to be no penalty attached to using these wild bootstrap schemes even if the errors are homoskedastic. An intetesting feature, perhaps, is that the Breusch-Pagan wild bootstrap tests also provide significant improvements over first-order asymptotic theory but appeared less reliable than $NRBP^*_{NT}$. Thus the use $NRBP^*_{T}$ in conjunction with a recursive-design bootstrap scheme recommends itself as an additional useful test procedure for applied workers.

[8] Full test results are available upon request. Only the p-value of Norway was on the borderline, being 5.1%. However, assuming all country specific errors are cross-sectionally independent, then the serial correlation test statistics are also independent over countries. Thus, the result that the proportion of the rejections, at (about) the 5% significance level and over 20 countries, is 5% is consistent with the hypothesis of no error serial correlation.

References

Bai, J. (2009). 'Panel data models with interactive fixed effects', *Econometrica*, 77(4): 1229–1279.

Baltagi, B., Feng, Q., and Kao, C. (2012). 'A Lagrange multiplier test for cross-sectional dependence in Fixed Effects panel data models', *Journal of Econometrics*, 170: 164–177.

Bassanini, A., and Scarpetta, S. (2002). 'Does human capital matter for growth in OECD countries? A pooled mean-group approach', *Economics Letters*, 74(3): 399–405.

Bond, S., Leblebicioglu, A., and Schiantarelli, F. (2010). 'Capital accumulation and growth: a new look at the empirical evidence', *Journal of Applied Econometrics*, 25(7): 1073–1099.

Breusch, T., and Pagan, A. (1980). 'The Lagrange multiplier test and its applications to model specification in econometrics', *The Review of Economic Studies*, 47(1): 239–253.

Bühlmann, P. (1995). 'Moving-average representation of autoregressive approximations', *Stochastic processes and their applications*, 60(2): 331–342.

Cavaliere, G., and Taylor, A. (2008). 'Testing for a change in persistence in the presence of non-stationary volatility', *Journal of Econometrics*, 147(1): 84–98.

Cochran, W. (1952). 'The chisquare test of goodness of fit', *Annals of Mathematical Statistics*, 23(3): 315–345.

Davidson, R., and Flachaire, E. (2008). 'The wild bootstrap, tamed at last', *Journal of Econometrics*, 146(1): 162–169.

Frees, E. (1995). 'Assessing cross-sectional correlation in panel data', *Journal of Econometrics*, 69(2): 393–414.

Godfrey, L., and Tremayne, A. (2005). 'The wild bootstrap and heteroskedasticity-robust tests for serial correlation in dynamic regression models', *Computational statistics & data analysis*, 49(2): 377–395.

Godfrey, L., and Yamagata, T. (2011). 'A note on a robust test for error cross-section correlation in panel data models', *Mimeo. University of York.*

Goncalves, S., and Kilian, L. (2004). 'Bootstrapping autoregressions with conditional heteroskedasticity of unknown form', *Journal of Econometrics*, 123(1): 89–120.

Hall, P. (1984). 'A Central limit theorem for integrated square error of multivariate nonparametric density estimators', *Journal of Multivariate Analysis*, 14: 1–16.

Hall, P., and Heyde, C. (1980). *Martingale limit theory and its application.* Academic Press.

Halunga, A. (2005). *Misspecification testing when modelling financial time series.* PhD thesis, University of Manchester.

Haque, N., Pesaran, M., and Sharma, S. (1999). 'Neglected heterogeneity and dynamics in cross-country savings regressions', *IMF Working paper, WP/99/128.*

Im, K., Pesaran, M., and Shin, Y. (2003). 'Testing for unit roots in heterogenous panels', *Journal of Econometrics*, 115: 53–74.

Juhl, T. (2011). 'A direct test for cross-sectional correlation in panel data models', *Mimeo. University of Kansas.*

Kuersteiner, G. (2001). 'Optimal instrumental variables estimation for ARMA models', *Journal of Econometrics*, 104(2): 359–405.

Lima, V., Souza, T., Cribari-Neto, F., and Fernandes, G. (2009). 'Heteroskedasticity-Robust Inference in Linear Regressions', *Communications in Statistics-Simulation and Computation*, 39(1): 194–206.

Liu, R. (1988). 'Bootstrap procedures under some non-iid models', *The Annals of Statistics*, 16(4): 1696–1708.

Magnus, J., and Neudecker, H. (1999). *Matrix differential calculus with applications in statistics and econometrics. Revised edition.* Wiley, Chichester, England.

Mammen, E. (1993). 'Bootstrap and wild bootstrap for high dimensional linear models', *The Annals of Statistics*, 21(1): 255–285.

Moscone, F., and Tosetti, E. (2009). 'A Review and Comparison of Tests of Cross-Section Independence in Panels', *Journal of Economic Surveys*, 23(3): 528–561.

Pesaran, M. (2004). 'General diagnostic tests for cross section dependence in panels', *CESifo Working Paper Series No. 1229; IZA Discussion Paper No. 1240.*

——— (2012). 'Testing weak cross-sectional dependence in large panels', *IZA Discussion Paper No. 6432.*

Pesaran, M., and Smith, R. (1995). 'Estimating long-run relationships from dynamic heterogenous panels', *Journal of Econometrics*, 68: 79–113.

Pesaran, M., Ullah, A., and Yamagata, T. (2008). 'A bias-adjusted LM test of error cross section independence', *Econometrics Journal*, 11(1): 105–127.

Pesaran, M., and Yamagata, T. (2008). 'Testing slope homogeneity in large panels', *Journal of Econometrics*, 142: 50–93.

Ramberg, J., Tadikamalla, P., Dudewicz, E., and Mykytka, E. (1979). 'A probability distribution and its uses in fitting data', *Technometrics*, 21(2): 201–214.

Sarafidis, V., Yamagata, T., and Robertson, D. (2009). 'A test of cross section dependence for a linear dynamic panel model with regressors', *Journal of Econometrics*, 148(2): 149–161.

Swamy, P. (1970). 'Efficient inference in a random coefficient regression model', *Econometrica*, 38: 311–323.

Weiss, A. (1986). 'Asymptotic theory for ARCH models: estimation and testing', *Econometric theory*, pp. 107–131.

White, H. (2001). *Asymptotic Theory for Econometricians.* Academic Press Orlando, Florida.

Wu, C. (1986). 'Jackknife, bootstrap and other resampling methods in regression analysis', *The Annals of Statistics*, 14(4): 1261–1295.

Zellner, A. (1962). ‘An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias’, *Journal of the American statistical Association*, 57(298): 348–368.

Appendix A

In what follows $\|A\| = \sqrt{\sum_i \sum_j a_{ij}^2}$ denotes the Euclidean norm of a matrix $A = \{a_{ij}\}$ and $\mathbb{N}$ the set of positive integers.

Asymptotic Validity of RBP_T

We first present some preliminary results which are employed in the Proof of Theorem 1. The proofs of these intermediate results exploit the fact that, following Kuersteiner (2001) and Goncalves and Kilian (2004), (1) can be written as $y_{it} = \sum_{k=0}^{\infty} \psi_{ik} r_{i,t-k}$, $r_{it} = w_{it}'\theta_i + u_{it}$ where ψ_{ik} is a function of the true parameter vector ϕ_i, satisfying the recursion $\psi_{is} - \phi_{i1}\psi_{i,s-1} - ... - \phi_{ip}\psi_{i,s-p} = 0$, for all $s > 0$, with $\psi_{i0} = 1$ and $\psi_{ik} = 0$, $k < 0$, for all i, implying that $\sum_{k=1}^{\infty} k\,|\psi_{ik}| < \infty$ for all i (see Bühlmann, 1995). Furthermore, we can write $Y_{i,t-1} = \sum_{k=1}^{\infty} c_{ik} r_{i,t-k}$ where $c_{ik} = \left(\psi_{i,k-1}, ..., \psi_{i,k-p}\right)'$ and $\sum_{k=1}^{\infty} \|c_{ik}\| \leq \Delta < \infty$, for all $i = 1, ..., N$.

Proposition 1 *Under Assumption A2(i),(iv),(v), and uniformly in $i, j = 1, ..., N$:*

(a) $E\,\|x_{it}\|^{4+\delta} \leq \Delta < \infty$ *for some* $\delta > 0$ *and all* t*;*

(b) $\{x_{it}u_{jt}, \mathcal{F}_t\}$ *is a vector m.d.s.*

Lemma 3 *Consider a sequence of scalar random variables denoted $\bar{Z}_{T,k}$, indexed by $k \in \mathbb{N}$, such that: (i) $E\left|\bar{Z}_{T,k}\right| \leq \Delta < \infty$ uniformly in k and T; and, (ii) $\bar{Z}_{T,k} \xrightarrow{p} 0$, as $T \to \infty$, for each fixed $k \in \mathbb{N}$. Define $\bar{S}_T = \sum_{k=1}^{\infty} \xi_k \bar{Z}_{T,k}$, where $\sum_{k=1}^{\infty} |\xi_k| < \infty$. Then, $\bar{S}_T \xrightarrow{p} 0$.*

The following Lemma exploits Lemma 3 and is central to the proof of Theorem 1.

Lemma 4 *Under Assumptions A1, A2, A3 and A4(i), and uniformly in $i, j = 1, ..., N$:*

(a) $\frac{1}{T}\sum_{t=1}^{T}\left(x_{it}x_{jt}' - E[x_{it}x_{jt}']\right) = o_p(1)$*, where* $E[x_{it}x_{jt}'] \leq \Delta < \infty$ *for all* t*;*

(b) $\frac{1}{T}\sum_{t=1}^{T} E[x_{it}x_{it}']$ *is uniformly positive definite;*

(c) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} x_{it}u_{jt} = O_p(1)$.

Proof of Theorem 1. It is shown that $\hat{\gamma}_{ij,T} - \gamma_{ij,T} = o_p(1)$ and the result follows.
1. First, define $H_i = X_i\left(X_i'X_i\right)^{-1}X_i'$. Then,

$$\sum_{t=1}^{T} \hat{u}_{it}\hat{u}_{jt} = \sum_{t=1}^{T} u_{it}u_{jt} - u_i'H_iu_j - u_i'H_ju_j + u_i'H_iH_ju_j$$

It follows from Lemma 4 that $u_i'H_iu_j$, $u_i'H_ju_j$ and $u_i'H_iH_ju_j$ are all $O_p(1)$ with $T^{-1}X_i'X_i$, in particular, being uniformly positive definite with probability one. Thus $T^{-1/2}\sum_{t=1}^{T} \hat{u}_{it}\hat{u}_{jt} = T^{-1/2}\sum_{t=1}^{T} u_{it}u_{jt} + O_p(T^{-1/2})$.
2. We now show that $\frac{1}{T}\sum_{t=1}^{T} \hat{u}_{it}^2\hat{u}_{jt}^2 - \frac{1}{T}\sum_{t=1}^{T} u_{it}^2u_{jt}^2 = o_p(1)$, and the result follows. Making the substitution $\hat{u}_{it} = u_{it} - x_{it}'(\hat{\beta}_i - \beta_i)$ we get

$$\hat{u}_{it}^2 = u_{it}^2 - 2u_{it}x_{it}'(\hat{\beta}_i - \beta_i) + (\hat{\beta}_i - \beta_i)'x_{it}x_{it}'(\hat{\beta}_i - \beta_i),$$

so that, writing $\delta_i = \hat{\beta}_i - \beta_i = O_p(T^{-1/2})$,

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\hat{u}_{it}^2\hat{u}_{jt}^2 - \frac{1}{T}\sum_{t=1}^{T}u_{it}^2u_{jt}^2 &= 4\delta_i'\left(\frac{1}{T}\sum_{t=1}^{T}u_{it}u_{jt}x_{it}x_{jt}'\right)\delta_j \\
&\quad -2\delta_i'\frac{1}{T}\sum_{t=1}^{T}u_{jt}^2u_{it}x_{it} - 2\delta_j'\frac{1}{T}\sum_{t=1}^{T}u_{it}^2u_{jt}x_{jt} \\
&\quad +\delta_i'\left(\frac{1}{T}\sum_{t=1}^{T}u_{jt}^2x_{it}x_{it}'\right)\delta_i + \delta_j'\left(\frac{1}{T}\sum_{t=1}^{T}u_{it}^2x_{jt}x_{jt}'\right)\delta_j \\
&\quad +\delta_i'\left(\frac{1}{T}\sum_{t=1}^{T}x_{it}x_{it}'\delta_i\delta_j'x_{jt}x_{jt}'\right)\delta_j \\
&\quad -2\delta_i'\left(\frac{1}{T}\sum_{t=1}^{T}u_{jt}x_{jt}'\delta_j x_{it}x_{it}'\right)\delta_i \\
&\quad -2\delta_j'\left(\frac{1}{T}\sum_{t=1}^{T}u_{it}x_{it}'\delta_i x_{jt}x_{jt}'\right)\delta_j \\
&= \sum_{q=1}^{8}R_{qT}, \text{ say.}
\end{aligned}$$

By Markov's inequality, Assumption A2(v), Proposition 1(a) and repeated application of Cauchy-Schwartz, it can be shown that $R_{qT} = o_p(1)$, $q = 1, ..., 8$, and the result follows.
For example, consider $R_{1T} = 4\delta_i'\left(\frac{1}{T}\sum_{t=1}^{T}u_{it}u_{jt}x_{it}x_{jt}'\right)\delta_j$. By Cauchy-Schwartz

$$E\left|u_{it}u_{jt}x_{itl}x_{jtm}\right| \leq \sqrt{E\left|u_{it}x_{itl}\right|^2 E\left|u_{jt}x_{jtm}\right|^2} \leq \Delta < \infty,$$

and $E\left|u_{it}x_{itl}\right|^2 \leq E\left|u_{it}\right|^4 E\left|x_{itl}\right|^4 \leq \Delta < \infty$, by Assumption A2(v) and Proposition 1(a). Thus, by Markov's equality, $R_{1T} = O_p(T^{-1})$ uniformly in (i,j). Similar reasoning gives $R_{qT} = O_p(T^{-1/2})$, $q = 2, 3$, and $R_{qT} = O_p(T^{-1})$, for $q = 4, 5$.
For $R_{6T} = \delta_i'\left(\frac{1}{T}\sum_{t=1}^{T}x_{it}x_{it}'\delta_i\delta_j'x_{jt}x_{jt}'\right)\delta_j$, note that $vec(ABC) = (C' \otimes A)\, vec\,(B)$, yielding

$$vec\left(\frac{1}{T}\sum_{t=1}^{T}x_{it}x_{it}'\delta_i\delta_j'x_{jt}x_{jt}'\right) = \frac{1}{T}\sum_{t=1}^{T}\left(x_{jt}x_{jt}' \otimes x_{it}x_{it}'\right) vec\left(\delta_i\delta_j'\right)$$

where elements of $\left(x_{jt}x_{jt}' \otimes x_{it}x_{it}'\right)$ are $x_{jth}x_{jtl}x_{itm}x_{itn}$, with

$$E\left|x_{jth}x_{jtl}x_{itm}x_{itn}\right| \leq \sqrt{E\left|x_{jth}x_{jtl}\right|^2 E\left|x_{itm}x_{itn}\right|^2} \leq \Delta^2 < \infty,$$

implying that $R_{6T} = O_p(T^{-2})$. Again, similar reasoning gives $R_{qT} = O_p(T^{-3/2})$, $q = 7, 8$, and this completes the proof.
3. We show that $\text{plim}_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\left\{u_{it}^2u_{jt}^2 - E[u_{it}^2u_{jt}^2]\right\} = 0$. Note that, with $\tau_{ijt}^2 = E[u_{it}^2u_{jt}^2|\mathcal{F}_{NT,t-1}]$

$$\frac{1}{T}\sum_{t=1}^{T}\left\{u_{it}^2u_{jt}^2 - E[u_{it}^2u_{jt}^2]\right\} = \frac{1}{T}\sum_{t=1}^{T}\left\{u_{it}^2u_{jt}^2 - \tau_{ijt}^2\right\} + \frac{1}{T}\sum_{t=1}^{T}\left\{\tau_{ijt}^2 - E[u_{it}^2u_{jt}^2]\right\}$$

where the second term is $o_p(1)$ by Assumption A4(iii). The first term is $o_p(1)$ by a Law of Large Numbers for the heterogeneous m.d.s., $\left\{u_{it}^2u_{jt}^2 - \tau_{ijt}^2, \mathcal{F}_t\right\}$, since $E\left|u_{it}^2u_{jt}^2\right|^{1+\delta} < \infty$. ■

Proof of Corolary 1. Since $\hat{\gamma}_{ij,T} - \gamma_{ij,T} = o_p(1)$ and $\gamma_{ij,T} \xrightarrow{d} N(0,1)$, $\hat{\gamma}_{ij,T}^2 \xrightarrow{d} \chi_1^2$. Furthermore, by asymptotic normality of $\gamma_{ij,T}$, verifying that $E\left[u_{it}u_{jt}u_{ks}u_{ms}\right] = 0$, for pairs $(i,j) \neq (k,m)$ and all t, s establishes the asymptotic independence of the $\hat{\gamma}_{ij,T}$ and the result follows. Firstly, note by Assumption A4(i), $E\left[u_{it}u_{jt}|\mathcal{F}_{NT,t-1}\right] = 0$ so we need only consider $t = s$. Now, without loss of generality, we can assume $i < j$ and $k < m$, with $i \leq k < m$ so that $E\left[u_{it}u_{jt}u_{kt}u_{mt}\right]$ gives the covariance between all possible distinct products $\{u_{it}u_{jt}\}$, $i < j$, and $\{u_{kt}u_{mt}\}$, $k < m$. But this is zero by Assumption A4(v)and we are done. ■

Proof of Proposition 1. The proof is omitted as follows directly under our Assumptions and repeated applications of Minkowski's inequality for (a) and the triangle and Cauchy-Schwartz inequalities for (b). ■

Proof of Lemma 3. Let $\bar{S}_T^n = \sum_{k=1}^n \xi_k \bar{Z}_{T,k}$, for fixed n. Firstly, it is clear that $\bar{S}_T^n \xrightarrow{p} 0$, as $T \to \infty$ for fixed n. Secondly, by Markov's inequality and $\sum_{k=1}^{\infty} |\xi_k| < \infty$, $\bar{S}_T \xrightarrow{p} 0$. ■

Proof of Lemma 4. (a) Consider the corresponding conformable partitions of $\frac{1}{T}\sum_{t=1}^T x_{it}x_{jt}'$, where $x_{it}' = (w_{it}', Y_{i,t-1}')$. First, by Assumption A1, $w_{it}w_{jt}$ is mixing and Assumption A2(iv) implies that $E\|w_{it}w_{jt}\|^{\eta+\delta} \leq \Delta < \infty$ uniformly in i,j,t, by an application of the Cauchy-Schwartz inequality. Thus, $\frac{1}{T}\sum_{t=1}^T w_{i,t-h}w_{j,t-k}' - \frac{1}{T}\sum_{t=1}^T E[w_{i,t-h}w_{j,t-k}'] = o_p(1)$, uniformly in i,j, by a Law of Large Numbers (e.g., White (2001, Corollary 4.48)), so that $\frac{1}{T}\sum_{t=1}^T w_{i,t-h}w_{j,t-k}' = O_p(1)$, for all fixed $h,k \in \mathbb{N}$, uniformly in i,j. In particular, these results hold for $h=k=0$ and $i=j$.

Second, for any $\mu \in \mathbb{R}^M$ and any $\lambda \in \mathbb{R}^p$ such that $\|\mu\| = \|\lambda\| = 1$

$$\mu'\left(\frac{1}{T}\sum_{t=1}^T w_{it}Y_{j,t-1}'\right)\lambda = \frac{1}{T}\sum_{t=1}^T\sum_{k=1}^{\infty}\xi_{jk}v_{it}r_{j,t-k}$$

where $\xi_{jk} = c_{jk}'\lambda$, $v_{it} = \mu' w_{it}$. Since $E\|v_{it}r_{j,t-k}\| \leq E\|w_{it}r_{j,t-k}\| \leq \Delta < \infty$, by Assumption A2(iv),(v), Proposition 1 and Cauchy-Schwartz, we can write

$$\mu'\frac{1}{T}\sum_{t=1}^T\left(w_{it}Y_{j,t-1}' - E[w_{it}Y_{j,t-1}']\right)\lambda = \sum_{k=1}^{\infty}\xi_{jk}\bar{Z}_{T,k}^{(i,j)}$$

where

$$\begin{aligned}
\bar{Z}_{T,k}^{(i,j)} &= \frac{1}{T}\sum_{t=1}^T\left(v_{it}r_{j,t-k} - E\left[v_{it}r_{j,t-k}\right]\right) \\
&= \mu'\left\{\frac{1}{T}\sum_{t=1}^T\left(w_{it}w_{j,t-k}' - E[w_{it}w_{j,t-k}']\right)\right\}\theta_j \\
&\quad + \mu'\frac{1}{T}\sum_{t=1}^T w_{it}u_{j,t-k},
\end{aligned}$$

and satisfies $E\left|\bar{Z}_{T,k}^{(i,j)}\right| \leq \Delta < \infty$. Moreover, Assumptions A2(i),(iv) and (v) imply that $\{w_{it}u_{j,t-k}, \mathcal{F}_{t-k}\}$ is a vector m.d.s. satisfying $\frac{1}{T}\sum_{t=1}^T w_{it}u_{j,t-k} = o_p(1)$ for all fixed $k \in \mathbb{N}$. As noted above, $\frac{1}{T}\sum_{t=1}^T\left(w_{it}w_{j,t-k}' - E[w_{it}w_{j,t-k}']\right) = o_p(1)$, uniformly in i,j, so that $\bar{Z}_{T,k}^{(i,j)} \xrightarrow{p} 0$ for all $\mu \in \mathbb{R}^M$, $\|\mu\| = 1$. Since $\sum_{k=1}^{\infty}\left|\xi_{jk}\right| \leq \Delta < \infty$, uniformly in j. Lemma 3 gives $\frac{1}{T}\sum_{t=1}^T\left(w_{it}Y_{j,t-1}' - E[w_{it}Y_{j,t-1}']\right) = o_p(1)$, uniformly in i,j.

Similarly, apply Lemma 3 repeatedly to

$$\lambda'\left\{\frac{1}{T}\sum_{t=1}^T\left(Y_{i.t-1}Y_{j,t-1}' - E[Y_{i,t-1}Y_{j,t-1}']\right)\right\}\lambda = \frac{1}{T}\sum_{t=1}^T\sum_{k=1}^{\infty}\sum_{h=1}^{\infty}\xi_{ik}\xi_{jh}\bar{Z}_{T,k,h}^{(i,j)}$$

for any $\lambda \in \mathbb{R}^p$ and again writing $\xi_{jk} = c_{jk}'\lambda$, where

$$\begin{aligned}
\bar{Z}_{T,k,h}^{(i,j)} &= \frac{1}{T}\sum_{t=1}^T\left(r_{i,t-k}r_{j,t-h} - E[r_{i,t-k}r_{j,t-h}]\right) \\
&= \theta_i'\left\{\frac{1}{T}\sum_{t=1}^T\left(w_{i,t-k}w_{j,t-h}' - E[w_{i,t-k}w_{j,t-h}]\right)\right\}\theta_j \\
&\quad + \theta_i'\frac{1}{T}\sum_{t=1}^T w_{i,t-k}u_{j,t-h} + \theta_j'\frac{1}{T}\sum_{t=1}^T w_{j,t-h}u_{i,t-k} \\
&\quad + \frac{1}{T}\sum_{t=1}^T\left(u_{i,t-k}u_{j,t-h} - E\left[u_{i,t-k}u_{j,t-h}\right]\right),
\end{aligned}$$

satisfying $E\left|\bar{Z}_{T,k,h}^{(i,j)}\right| \leq \Delta < \infty$ by Cauchy-Schwartz and Assumption A2(iv),(v) and Proposition 1. Similar to before, and for all fixed $h,k \in \mathbb{N}$, the first three terms in the expression for $\bar{Z}_{T,k,h}^{(i,j)}$ are all $o_p(1)$. For the final term, consider first $k \neq h$, so that $\{u_{i,t-k}u_{j,t-h}, \mathcal{F}_{t-g}\}$, $g = \min(k,h)$, is a m.d.s. and Assumption A2(v) ensures that $\frac{1}{T}\sum_{t=1} u_{i,t-k}u_{j,t-h} = o_p(1)$, for all fixed $h,k \in \mathbb{N}$. Now, for $k = h$, and $i \neq j$, $\{u_{i,t-k}u_{j,t-k}, \mathcal{F}_{t-k}\}$ is a m.d.s. by Assumption A4(i) and $\frac{1}{T}\sum_{t=1}^{T} u_{i,t-k}u_{j,t-k} \xrightarrow{p} 0$, for fixed $k \in \mathbb{N}$. For $k = h$ *and* $i = j$, we have, by Assumption A2(ii) and (iii)

$$\frac{1}{T}\sum_{t=1}^{T}\left(u_{i,t-k}^2 - E[u_{i,t-k}^2]\right) = \frac{1}{T}\sum_{t=1}^{T}\left(u_{i,t-k}^2 - \sigma_{i,t-k}^2\right) + o_p(1)$$

By Assumption 2(v), $\left\{u_{i,t-k}^2 - \sigma_{i,t-k}^2, \mathcal{F}_{t-k}\right\}$ is a m.d.s., and $\frac{1}{T}\sum_{t=1}^{T}\left(u_{i,t-k}^2 - \sigma_{i,t-k}^2\right) \xrightarrow{p} 0$, also. Thus for fixed $h \in \mathbb{N}$ and $k \in \mathbb{N}$, $\bar{Z}_{T,k,h}^{(i,j)} = o_p(1)$. Repeated application of Lemma 3 establishes the desired result.
(b) By part (a), for $i = j$, $\frac{1}{T}\sum_{t=1}^{T} x_{it}x_{it}' - Q_{iT} = o_p(1)$, uniformly in i, where $Q_{iT} = \frac{1}{T}\sum_{t=1}^{T} E[x_{it}x_{it}']$. Writing $z_{it} = \sum_{k=1}^{\infty} c_{ik}\left(w_{i,t-k}'\theta_i\right)$, $(p \times 1)$, Q_{iT} can be expressed as

$$Q_{iT} = \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} E[w_{it}w_{it}'] & E[w_{it}z_{it}'] \\ E[z_{it}w_{it}'] & E\left[z_{it}z_{it}'\right] + \sum_{k=1}^{\infty} c_{ik}c_{ik}'E\left[u_{i,t-k}^2\right] \end{bmatrix}.$$

Now, by Assumption A3(i), $\frac{1}{T}\sum_{t=1}^{T} E\left[w_{it}w_{it}'\right]$ is uniformly positive definite so that its inverse exists for large enough T. Then, exploiting, for example, Magnus and Neudecker (1999, Theorem 27, p.23), Q_{iT} is uniformly positive definite if and only if

$$A_T = \frac{1}{T}\sum_{t=1}^{T} E\left[\tilde{z}_{it}\tilde{z}_{it}'\right] + \sum_{k=1}^{\infty} c_{ik}c_{ik}'\frac{1}{T}\sum_{t=1}^{T} E\left[u_{i,t-k}^2\right]$$

is uniformly positive definite where

$$\tilde{z}_{it} = z_{it} - \frac{1}{T}\sum_{t=1}^{T} E\left[z_{it}w_{it}'\right]\left\{\frac{1}{T}\sum_{t=1}^{T} E\left[w_{it}w_{it}'\right]\right\}^{-1} w_{it}.$$

Now, for all non-zero $\lambda \in \mathbb{R}^p$

$$\begin{aligned} \lambda' A_T \lambda &= \frac{1}{T}\sum_{t=1}^{T} E\left|\lambda'\tilde{z}_{it}\right|^2 + \sum_{k=1}^{\infty}\left|\lambda' c_{ik}\right|^2\left\{\frac{1}{T}\sum_{t=1-k}^{T-k} E\left[u_{it}^2\right]\right\} \\ &\geq \sum_{k=1}^{p}\left|\lambda' c_{ik}\right|^2\left\{\frac{1}{T}\sum_{t=1}^{T-k} E\left[u_{it}^2\right]\right\} \end{aligned}$$

and the right hand side is uniformly positive, because $\frac{1}{T}\sum_{t=1}^{T-k} E\left[u_{it}^2\right]$ is uniformly positive by Assumption A3(ii), for any $k \leq p$, and $\sum_{k=1}^{p}\left|\lambda' c_{ik}\right|^2 > 0$, uniformly in i, for all non-zero $\lambda \in \mathbb{R}^p$. Therefore $A_T > 0$ for sufficiently large T (uniformly positive) and the result follows.
(c) It suffices to show that $var\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} x_{it}u_{jt}\right] = O(1)$ uniformly in i, j. By Proposition 1(b), $\{\lambda' x_{it}u_{jt}, \mathcal{F}_t\}$ is a m.d.s. for any $\lambda \in \mathbb{R}^{p+M}$, such that $\|\lambda\| = 1$, so

$$var\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\lambda' x_{it}u_{jt}\right] = \lambda'\left(\frac{1}{T}\sum_{t=1}^{T} E[u_{jt}^2 x_{it}x_{it}']\right)\lambda.$$

By Assumption A2(v) and Proposition 1(a), and a repeated application of Cauchy-Schwartz, it can be shown that $\sup_{i,t} E\left\|u_{it}^2 x_{it}x_{it}'\right\| = O(1)$, and the result follows. This completes the proof. ■

Proof of Lemma 1. Firstly, write

$$\begin{aligned} Z_{NT} &= \frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\left(\frac{\frac{1}{T}\left(\sum_{t=1}^{T} u_{it}^2 u_{jt}^2\right)^2 - \frac{1}{T}\sum_{t=1}^{T} u_{it}^2 u_{jt}^2}{\omega_{ij,T}}\right) \\ &\quad + \frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\left(\frac{\frac{1}{T}\sum_{t=1}^{T}\left(u_{it}^2 u_{jt}^2 - E\left[u_{it}^2 u_{jt}^2\right]\right)}{\omega_{ij,T}}\right) \\ &= 2\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{1}{T}\frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{u_{it}u_{jt}u_{is}u_{js}}{\omega_{ij,T}} \\ &\quad + \frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\left(\frac{\frac{1}{T}\sum_{t=1}^{T}\left(u_{it}^2 u_{jt}^2 - E\left[u_{it}^2 u_{jt}^2\right]\right)}{\omega_{ij,T}}\right) \\ &= Z_{1,NT} + Z_{2,NT} \end{aligned}$$

and $Z_{2,NT} = o_p(1)$ by Markov's Inequality and Assumption A4(iv), B1 and B2(i) because (for T sufficiently large)

$$\begin{aligned} E\left|Z_{2,NT}\right| &\leq K^{-1}\frac{1}{\sqrt{TN(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\left(E\left|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(u_{it}^2u_{jt}^2-E[u_{it}^2u_{jt}^2]\right)\right|\right) \\ &\leq \frac{\Delta}{K}\sqrt{\frac{N(N-1)}{T}}\to 0. \end{aligned}$$

Now, for the first term, $Z_{1,NT}$, we have

$$Z_{1,NT}=\sum_{t=2}^{T}\sum_{s=1}^{t-1}H_T\left(\underline{u}_t,\underline{u}_s\right)$$

where $\underline{u}_t=(u_{1t},..,u_{Nt})'$ and

$$H_T\left(\underline{u}_t,\underline{u}_s\right)=2\frac{1}{T}\frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{u_{it}u_{jt}u_{is}u_{js}}{\omega_{ij,T}}$$

Let $W_{Tt}=\sum_{s=1}^{t-1}H_T\left(\underline{u}_t,\underline{u}_s\right)$ so that

$$E\left[W_{Tt}|\mathcal{F}_{NT,t-1}\right]=0 \text{ a.s.}$$

by Assumption A4(i) since $\omega_{ij,T}=\frac{1}{T}\sum_{t=1}^{T}E\left[u_{it}^2u_{jt}^2\right]$ is measurable with respect to $\mathcal{F}_{NT,t-1}$ and thus W_{Tt} is a mds array with respect to $\mathcal{F}_{NT,t-1}$.
Therefore, we can apply the CLT for U-statistic of Hall (1984) or martingale difference arrays of Hall and Heyde (1980, Corollary 3.1), where $T=g(N)$ and $(N,T)\to\infty$. The following conditions for CLT for mds have to be satisfied as $(T,N)\to\infty$:
(i) $s_T^2\to 1$
where

$$s_T^2=E\left[\left(\sum_{t=2}^{T}W_{Tt}\right)^2\right]$$

(ii) $\left(s_T^2\right)^{-1}\sum_{t=2}^{T}E\left[W_{Tt}^2 1\left(|W_{Tt}|>\alpha s_T\right)|\mathcal{F}_{NT,t-1}\right]\to 0$ all $\alpha>0$
(iii) $\left(s_T^2\right)^{-1}V_T\to 1$, where $V_T=\sum_{t=2}^{T}E\left[W_{Tt}^2|\mathcal{F}_{NT,t-1}\right]$
Then

$$Z_{1,NT}=\sum_{t=2}^{T}W_{Tt}\xrightarrow{d}N(0,1)$$

as $(T,N)\to\infty$.
For **(i)**,

$$\begin{aligned} s_T^2 &= E\left[\sum_{t=2}^{T}\sum_{t'=2}^{T}W_{Tt}W_{Tt'}\right] \\ &= E\left[\sum_{t=2}^{T}W_{Tt}^2\right] \\ &= \sum_{t=2}^{T}E\left[\sum_{s=1}^{t-1}\sum_{r=1}^{t-1}H_T\left(\underline{u}_t,\underline{u}_s\right)H_T\left(\underline{u}_t,\underline{u}_r\right)\right] \\ &= \sum_{t=2}^{T}\sum_{s=1}^{t-1}E\left[H_T^2\left(\underline{u}_t,\underline{u}_s\right)\right]+O\left(T^{-1}\right) \end{aligned}$$

where the second line follows by the m.d.s. property of W_{Tt} and the last line follows since

$$E\left[H_T\left(\underline{u}_t,\underline{u}_s\right)H_T\left(\underline{u}_t,\underline{u}_r\right)\right]=\frac{4}{T^2N(N-1)}E\left[\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\sum_{l=1}^{N-1}\sum_{m=l+1}^{N}\frac{u_{it}u_{jt}u_{is}u_{js}}{\omega_{ij,T}}\frac{u_{lt}u_{mt}u_{lr}u_{mr}}{\omega_{lm,T}}\right]$$

and for $s \neq r$

$$
\begin{aligned}
E\left[\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\sum_{l=1}^{N-1}\sum_{m=l+1}^{N} u_{it}u_{jt}u_{is}u_{js}u_{lt}u_{mt}u_{lr}u_{mr}\right] &= \sum_{i<j} E\left[u_{it}^2 u_{jt}^2 u_{is}u_{js}u_{ir}u_{jr}\right] \\
&+ \sum_{i<j\neq l<m} E\left[u_{it}u_{jt}u_{is}u_{js}u_{lt}u_{mt}u_{lr}u_{mr}\right] \\
&+3\sum_{i<j<m} E\left[u_{it}u_{jt}^2 u_{is}u_{js}u_{mt}u_{jr}u_{mr}\right]
\end{aligned}
$$

where the last two terms are zero by Assumption A4(v), whereas for the first term we can write

$$
\begin{aligned}
& \frac{4}{T^2 N(N-1)}\sum_{i<j}\sum_{t=2}^{T}\sum_{s\neq r}^{t-1} E\left[u_{it}^2 u_{jt}^2 u_{is}u_{js}u_{ir}u_{jr}\right] \\
= & \frac{4}{T^2 N(N-1)}\sum_{i<j}\sum_{t=2}^{T}\sum_{s\neq r}^{t-1} cov\left(u_{it}^2 u_{jt}^2, u_{is}u_{js}u_{ir}u_{jr}\right) \\
& + \frac{4}{T^2 N(N-1)}\sum_{i<j}\sum_{t=2}^{T}\sum_{s\neq r}^{t-1} E\left[u_{it}^2 u_{jt}^2\right] E\left[u_{is}u_{js}u_{ir}u_{jr}\right] \\
= & O\left(T^{-1}\right)
\end{aligned}
$$

by Assumption B2(iii) and since $E\left[u_{is}u_{js}u_{ir}u_{jr}\right]=0$ by Assumption A4(i) for $s\neq r$ and $i\neq j$. Moreover,

$$
\begin{aligned}
2\frac{1}{T^2}\sum_{t=2}^{T}\sum_{s=1}^{t-1} E\left[u_{it}^2 u_{jt}^2 u_{is}^2 u_{js}^2\right] &= \frac{1}{T^2}\sum_{t\neq s} E\left[u_{it}^2 u_{jt}^2 u_{is}^2 u_{js}^2\right] \\
&= \frac{1}{T^2}\sum_{t\neq s} cov\left(u_{it}^2 u_{jt}^2, u_{is}^2 u_{js}^2\right) + \frac{1}{T^2}\sum_{t\neq s} E\left[u_{it}^2 u_{jt}^2\right] E\left[u_{is}^2 u_{js}^2\right] \\
&= \frac{1}{T^2}\sum_{t\neq s} E\left[u_{it}^2 u_{jt}^2\right] E\left[u_{is}^2 u_{js}^2\right] + O\left(T^{-1}\right) \\
&= \omega_{ijT}^2 + O\left(T^{-1}\right)
\end{aligned}
$$

where the third line follows by Assumption B2(iv) and the last line follows by Assumption A4(iv) and the fact that $T^{-2}\sum_{t=1}^{T}\left(E\left[u_{it}^2 u_{jt}^2\right]\right)^2 = O\left(T^{-1}\right)$ by Cauchy-Schwartz inequality and Assumption B2(ii). Thus

$$
\begin{aligned}
s_T^2 &= 2\frac{1}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{\omega_{ijT}^2}{\omega_{ijT}^2} + O\left(T^{-1}\right) \\
&= 1 + O\left(T^{-1}\right).
\end{aligned}
$$

To establish **(ii)** it is sufficient to show that

$$
\left(s_T^2\right)^{-1-\delta}\sum_{t=2}^{T} E\left|W_{Tt}\right|^{2+\delta} \to 0 \text{ as } (T,N)\to\infty \text{ for some } \delta>0. \tag{22}
$$

Since $s_T^2 = O(1)$, (22) is established for $\delta=2$ if

$$
\sum_{t=2}^{T} E\left[W_{Tt}^4\right] \to 0 \text{ as } (T,N)\to\infty
$$

where

$$\begin{aligned}
E\left[W_{Tt}^4\right] &= E\left[\left(\sum_{s=1}^{t-1} H_T\left(\underline{u}_t, \underline{u}_s\right)\right)^4\right] \\
&= \sum_{s=1}^{t-1} E\left[H_T^4\left(\underline{u}_t, \underline{u}_s\right)\right] + 3\sum_{s\neq r}^{t-1}\sum^{t-1} E\left[H_T^2\left(\underline{u}_t, \underline{u}_s\right) H_T^2\left(\underline{u}_t, \underline{u}_r\right)\right] \\
&\quad + \sum_{s\neq r\neq s'\neq r'}^{t-1} \sum^{t-1}\sum^{t-1}\sum^{t-1} E\left[H_T\left(\underline{u}_t, \underline{u}_s\right) H_T\left(\underline{u}_t, \underline{u}_r\right) H_T\left(\underline{u}_t, \underline{u}_{s'}\right) H_T\left(\underline{u}_t, \underline{u}_{r'}\right)\right] \\
&\quad + 6\sum_{s\neq r\neq s'}^{t-1} \sum^{t-1}\sum^{t-1} E\left[H_T^2\left(\underline{u}_t, \underline{u}_s\right) H_T\left(\underline{u}_t, \underline{u}_r\right) H_T\left(\underline{u}_t, \underline{u}_{s'}\right)\right] \\
&\quad + 4\sum_{s\neq r}^{t-1}\sum^{t-1} E\left[H_T^3\left(\underline{u}_t, \underline{u}_s\right) H_T\left(\underline{u}_t, \underline{u}_r\right)\right] \qquad (23)
\end{aligned}$$

The first term in (23) is

$$\begin{aligned}
E\left[H_T^4\left(\underline{u}_t, \underline{u}_s\right)\right] &= 16T^{-4}\left[\left(N\left(N-1\right)\right)\right]^{-2} E\left[\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \frac{u_{it}u_{jt}u_{is}u_{js}}{\omega_{ij,T}}\right)^4\right] \\
&\leq 16K^{-4}T^{-4}\left[\left(N\left(N-1\right)\right)\right]^{-2} E\left[\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} u_{it}u_{jt}u_{is}u_{js}\right)^4\right]
\end{aligned}$$

by Assumption A4(iv) and where

$$\begin{aligned}
& E\left[\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} u_{it}u_{jt}u_{is}u_{js}\right)^4\right] \\
&= E\left[\sum_{i<j}\sum_{l<m}\sum_{p<q}\sum_{h<n} u_{it}u_{jt}u_{is}u_{js}u_{lt}u_{mt}u_{ls}u_{ms}u_{pt}u_{qt}u_{ps}u_{qs}u_{ht}u_{nt}u_{hs}u_{ns}\right] \\
&= \sum_{i<j}\sum_{l<m}\sum_{p<q}\sum_{h<n} k_{ijlmpqhn} E\left[u_{is}u_{js}u_{ls}u_{ms}u_{ps}u_{qs}u_{hs}u_{ns}\right] \\
&\leq \Delta \sum_{i<j}\sum_{l<m}\sum_{p<q}\sum_{h<n} k_{ijlmpqhn}
\end{aligned}$$

and the last line follows since $E\left[u_{is}u_{js}u_{ls}u_{ms}u_{ps}u_{qs}u_{hs}u_{ns}\right] \leq \sup_t E\left|u_{it}^8\right| \leq \Delta < \infty$ by repeated application of Cauchy-Schwartz inequality and Assumption B2(v). Thus

$$\begin{aligned}
E\left[H_T^4\left(\underline{u}_t, \underline{u}_s\right)\right] &\leq 16K^{-4}\Delta T^{-4}N^{-4}\sum_{i<j}\sum_{l<m}\sum_{p<q}\sum_{h<n} k_{ijlmpqhn} \\
&= O\left(T^{-4}\right)
\end{aligned}$$

by Assumptions A4(iv) and B2(v). For the second term in (23)

$$\begin{aligned}
E\left[H_T^2\left(\underline{u}_t, \underline{u}_s\right) H_T^2\left(\underline{u}_t, \underline{u}_r\right)\right] &\leq \left\{E\left[H_T^4\left(\underline{u}_t, \underline{u}_s\right)\right] E\left[H_T^4\left(\underline{u}_t, \underline{u}_r\right)\right]\right\}^{1/2} \\
&= O\left(T^{-4}\right)
\end{aligned}$$

where the first line follows by Cauchy-Schwartz inequality and the second line by the previous result. By Assumption A4(iv), the third term in (23) can be written further as

$$\begin{aligned}
& E\left[H_T\left(\underline{u}_t, \underline{u}_s\right) H_T\left(\underline{u}_t, \underline{u}_r\right) H_T\left(\underline{u}_t, \underline{u}_{s'}\right) H_T\left(\underline{u}_t, \underline{u}_{r'}\right)\right] \\
\leq\ & 16K^{-4}\frac{1}{T^4}\frac{1}{N^2\left(N-1\right)^2}\sum_{i<j}\sum_{l<m}\sum_{p<q}\sum_{h<n} E\left[u_{it}u_{jt}u_{is}u_{js}u_{lt}u_{mt}u_{lr}u_{mr}u_{pt}u_{qt}u_{ps'}u_{qs'}u_{ht}u_{nt}u_{hr'}u_{nr'}\right] \\
=\ & 16K^{-4}\frac{1}{T^4}\frac{1}{N^2\left(N-1\right)^2}\sum_{i<j}\sum_{l<m}\sum_{p<q}\sum_{h<n} k_{ijlmpqhn} E\left[u_{is}u_{js}u_{lr}u_{mr}u_{ps'}u_{qs'}u_{hr'}u_{nr'}\right] \\
=\ & 0
\end{aligned}$$

where the last line follows by Assumptions A4(i) and B2(v). Analogously, the fourth term is

$$\begin{aligned} & E\left[H_T^2\left(\underline{u}_t, \underline{u}_s\right) H_T\left(\underline{u}_t, \underline{u}_r\right) H_T\left(\underline{u}_t, \underline{u}_{s'}\right)\right] \\ \leq\ & 16K^{-4}\frac{1}{T^4}\frac{1}{N^2\left(N-1\right)^2}\sum_{i<j}\sum_{l<m}\sum_{p<q}\sum_{h<n} E\left[u_{it}u_{jt}u_{lt}u_{mt}u_{pt}u_{qt}u_{ht}u_{nt}u_{is}u_{js}u_{ls}u_{ms}u_{pr}u_{qr}u_{hs'}u_{ns'}\right] \\ =\ & 16K^{-4}\frac{1}{T^4}\frac{1}{N^2\left(N-1\right)^2}\sum_{i<j}\sum_{l<m}\sum_{p<q} k_{ijlmpqhn} E\left[u_{is}u_{js}u_{ls}u_{ms}u_{pr}u_{qr}u_{hs'}u_{ns'}\right] \\ =\ & 0 \end{aligned}$$

where $E\left[u_{is}u_{js}u_{ls}u_{ms}u_{pr}u_{qr}u_{hs'}u_{ns'}\right] = 0$ by Assumption A4(i) when $s < \max(r, s')$, whereas when $s > \max\left(r, s'\right)$ and $i = l < j = m \neq p < q$ the expectation is zero by Assumptions B2(iii) and Assumption A4(i), otherwise the expectation is zero by Assumption A4(v). The last line then follows by Assumption B2(v). By Assumption A4(iv), the fifth term in (23) is

$$\begin{aligned} & E\left[H_T^3\left(\underline{u}_t, \underline{u}_s\right) H_T\left(\underline{u}_t, \underline{u}_r\right)\right] \\ \leq\ & 16K^{-4}\frac{1}{T^4}\frac{1}{N^2\left(N-1\right)^2}\sum_{i<j}\sum_{l<m}\sum_{p<q}\sum_{h<n} E\left[u_{it}u_{jt}u_{lt}u_{mt}u_{pt}u_{qt}u_{ht}u_{nt}u_{is}u_{js}u_{ls}u_{ms}u_{ps}u_{qs}u_{hr}u_{nr}\right] \\ =\ & 16K^{-4}\frac{1}{T^4}\frac{1}{N^2\left(N-1\right)^2}\sum_{i<j}\sum_{l<m}\sum_{p<q}\sum_{h<n} k_{ijlmpqhn} E\left[u_{is}u_{js}u_{ls}u_{ms}u_{ps}u_{qs}u_{hr}u_{nr}\right] \\ =\ & O(T^{-4}) \end{aligned}$$

where the third and last line by Assumption B2(ii), (v) for $s > r$ since $E\left[u_{is}u_{js}u_{ls}u_{ms}u_{ps}u_{qs}u_{hs}u_{ns}\right] \leq \sup_{t,i} E\left|u_{it}^8\right| \leq \Delta < \infty$ by repeated application of Cauchy-Schwartz inequality, whereas for $s < r$, $E\left[u_{is}u_{js}u_{ls}u_{ms}u_{ps}u_{qs}u_{hr}u_{nr}\right] = 0$ by Assumption A4(i). Thus,

$$\begin{aligned} E\left[W_{Tt}^4\right] &= E\left[\sum_{s=1}^{t-1} H_T^4\left(\underline{u}_t, \underline{u}_s\right)\right] + 3\sum_{s=1}^{t-1}\sum_{r=1, s\neq r}^{t-1} E\left[H_T^2\left(\underline{u}_t, \underline{u}_s\right) H_T^2\left(\underline{u}_t, \underline{u}_r\right)\right] \\ &\quad + 4\sum_{s=1}^{t-1}\sum_{r=1, s\neq r}^{t-1} E\left[H_T^3\left(\underline{u}_t, \underline{u}_s\right) H_T\left(\underline{u}_t, \underline{u}_r\right)\right] \\ &= O(T^{-2}) \end{aligned}$$

which yields that

$$\left(s_T^2\right)^{-2}\sum_{t=2}^{T} E\left[W_{Tt}^4\right] = O\left(T^{-1}\right)$$

which establishes the result.

(iii) We have

$$\begin{aligned} V_T &= \sum_{t=2}^{T} E\left[W_{Tt}^2 | \mathcal{F}_{NT,t-1}\right] \\ &= \sum_{t=2}^{T} E\left[\sum_{s=1}^{t-1}\sum_{r=1}^{t-1} H_T\left(\underline{u}_t, \underline{u}_s\right) H_T\left(\underline{u}_t, \underline{u}_r\right) | \mathcal{F}_{NT,t-1}\right] \\ &= \frac{4}{N\left(N-1\right)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{1}{\omega_{ij,T}^2}\frac{1}{T^2}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\sum_{r=1}^{t-1} u_{is}u_{js}u_{ir}u_{jr} E\left[u_{it}^2 u_{jt}^2 | \mathcal{F}_{NT,t-1}\right] \end{aligned}$$

using similar arguments as established in **(i)**. Further

$$\begin{aligned} V_T - s_T^2 &= \frac{4}{N\left(N-1\right)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{1}{\omega_{ij,T}^2}\frac{1}{T^2}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\sum_{r=1}^{t-1}\Big(u_{is}u_{js}u_{ir}u_{jr} E\left[u_{it}^2 u_{jt}^2 | \mathcal{F}_{NT,t-1}\right] \\ &\quad - E\left[u_{it}^2 u_{jt}^2 u_{is}u_{js}u_{ir}u_{jr}\right]\Big) + O\left(T^{-1}\right) \\ &= o_p\left(1\right) \end{aligned}$$

by Assumption B2(vi). ■

Proof of Lemma 2. Define

$$\hat{Z}^{\dagger}_{NT} = \frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\left(\hat{\gamma}^{\dagger 2}_{ij,T}-1\right)$$
$$\hat{\gamma}^{\dagger}_{ij,T} = \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\hat{u}_{it}\hat{u}_{jt}}{\sqrt{\omega_{ij,T}}}.$$

then we have

$$NRBP_{NT}-Z_{NT}=(\hat{Z}^{\dagger}_{NT}-Z_{NT})+(NRBP_{NT}-\hat{Z}^{\dagger}_{NT})$$

and we show that:
(i) First we show that $\left|\hat{Z}^{\dagger}_{NT}-Z_{NT}\right|=O_p(N/\sqrt{T})$, as $(N,T)\to\infty$.
(ii) Second, that $\left|NRBP_{NT}-\hat{Z}^{\dagger}_{NT}\right|=O_p(N/\sqrt{T})$, as $(N,T)\to\infty$.
and the result follows by Assumption B1.
Step (i): Firstly, it can be noted that from Proof of Theorem 1

$$\hat{\gamma}^{\dagger}_{ij,T}=\gamma_{ij,T}+\frac{1}{\sqrt{\omega_{ij,T}}}\frac{1}{\sqrt{T}}b_{ij,T}=\gamma_{ij,T}+c_{ij,T},\text{ say}$$

where

$$b_{ij,T}=u_i'H_iH_ju_j-u_i'H_iu_j-u_i'H_ju_j=\sum_{k=1}^{3}a_{ijk,T} \tag{24}$$

which is $O_p(1)$, for all i and j, where $H_i=X_i\left(X_i'X_i\right)^{-1}X_i'$ with analogous notation for H_j. Thus

$$\begin{aligned}\left|\hat{Z}^{\dagger}_{NT}-Z_{NT}\right| &= \frac{1}{\sqrt{N(N-1)}}\left|\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\left(\hat{\gamma}^{\dagger 2}_{ij,T}-\gamma^2_{ij,T}\right)\right|\\ &\le \frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\left|\hat{\gamma}^{\dagger 2}_{ij,T}-\gamma^2_{ij,T}\right|\\ &\le \frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\left|c^2_{ij,T}+2\gamma_{ij,T}c_{ij,T}\right|\\ &= A_{NT},\text{ say,}\end{aligned}$$

and we show that $A_{NT}=O_p(N/\sqrt{T})$. Then, by Cauchy-Schwartz inequality,

$$\begin{aligned}|A_{NT}| &\le \frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}c^2_{ij,T}+2\sqrt{\frac{1}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\gamma^2_{ij,T}\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}c^2_{ij,T}\right)}\\ &= \frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}c^2_{ij,T}+O_p(1)\sqrt{\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}c^2_{ij,T}\right)}\end{aligned}$$

where the second line follows because $G_{NT}=\frac{1}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\gamma^2_{ij,T}=O_p(1)$ since by Markov's inequality $\sup_{i,,j}E\left|\gamma^2_{ij,T}\right|\le\Delta<\infty$. Thus, to show that $A_{NT}=O_p(N/\sqrt{T})$, all that is required is to establish that $\sup_{i,j}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}c^2_{ij,T}=O_p(N^2/T)$. But since $c_{ij,T}=\frac{1}{\sqrt{\omega_{ij,T}}}\frac{1}{\sqrt{T}}b_{ij,T}$ and $\inf_{i,j}\omega_{ij,T}\ge K>0$ for T sufficiently large by Assumption A4(iv), we have

$$\begin{aligned}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}c^2_{ij,T} &\le K^{-1}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{1}{T}b^2_{ij,T}\\ &= K^{-1}\frac{N(N-1)}{T}\left\{\frac{1}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}b^2_{ij,T}\right\}\end{aligned}$$

Therefore, if $\sup_{i,j}\frac{1}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}b^2_{ij,T}=O_p(1)$, then $\left|\hat{Z}^{\dagger}_{NT}-Z_{NT}\right|\le|A_{NT}|=O_p(N/\sqrt{T})$, and we are done. In a similar vein to previous reasoning, in order to show this, it is sufficient that

$$\sup_{i,j}\frac{1}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}a^2_{ijk,T}=O_p(1)$$

for $k=1,2,3$ as $(N,T)\longrightarrow\infty$, where $a_{ijk,T}$ are defined in (24).
Consider $M_{NT,k}=\frac{1}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}a_{ijk,T}^2$, $k=1,2,3$. This is $O_p(1)$ by Markov's inequality, if $\sup_{i,j}E\left|a_{ijk,T}\right|^2\leq\Delta<\infty$, for $k=1,2,3$.For $a_{ij1,T}=u_i'H_iH_ju_j$, we have $\left|a_{ij1,T}\right|^2\leq\left\|H_iu_i\right\|^2\left\|H_ju_j\right\|^2$ and by Cauchy-Schwartz then yields

$$
\begin{aligned}
E\left|a_{ij1,T}\right|^2 &\leq E\left(\left\|H_iu_i\right\|^2E\left\|H_ju_j\right\|^2\right)\\
&\leq \sqrt{E\left\|H_iu_i\right\|^4E\left\|H_ju_j\right\|^4}
\end{aligned}
$$

where

$$E\left\|H_iu_i\right\|^4=E\left|u_i'H_iu_i\right|^2\leq\Delta<\infty$$

uniformly in i, since by Lemma 4 $u_i'H_iu_i=\frac{u_i'X_i}{\sqrt{T}}\left(\frac{X_i'X_i}{T}\right)^{-1}\frac{X_i'u_i}{\sqrt{T}}=O_p(1)$, uniformly in i. Similar arguments hold for $a_{ij2,T}$ and $a_{ij3,T}$. This establishes Step (i).

Step (ii): Here we show that $\left|RBP_{NT}-\hat{Z}_{NT}^{\dagger}\right|=O_p(N/\sqrt{T})$, as $(N,T)\rightarrow\infty$. Defining $\hat{\omega}_{ij,T}=\frac{1}{T}\sum_{t=1}^{T}\hat{u}_{it}^2\hat{u}_{jt}^2$, we have

$$
\begin{aligned}
\left|RBP_{NT}-\hat{Z}_{NT}^{\dagger}\right| &\leq \frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\left|\hat{\gamma}_{ij,T}^2-\hat{\gamma}_{ij,T}^{\dagger 2}\right|\\
&= \frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\hat{\gamma}_{ij,T}^2\left|1-\frac{\omega_{ij,T}}{\hat{\omega}_{ij,T}}\right|\\
&= \sqrt{\frac{N(N-1)}{T}}\frac{1}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\hat{\gamma}_{ij,T}^2\left|\frac{\sqrt{T}\left(\hat{\omega}_{ij,T}-\omega_{ij,T}\right)}{\hat{\omega}_{ij,T}}\right|\\
&= \sqrt{\frac{N(N-1)}{T}}B_{NT},\ \text{say.}
\end{aligned}
$$

Now in order to show that $B_{NT}=O_p(1)$, it is sufficient to have $\sup_{i,j}\hat{\gamma}_{ij,T}^2\left|\frac{\sqrt{T}\left(\hat{\omega}_{ij,T}-\omega_{ij,T}\right)}{\hat{\omega}_{ij,T}}\right|=O_p(1)$. Further, since $\sup_{i,j}\hat{\gamma}_{ij,T}^2=O_p(1)$ by Corollary 1, it is sufficient to establish that $\sup_{i,j}\left|\frac{\sqrt{T}\left(\hat{\omega}_{ij,T}-\omega_{ij,T}\right)}{\hat{\omega}_{ij,T}}\right|=O_p(1)$. Therefore, consider

$$\left|\frac{\sqrt{T}\left(\hat{\omega}_{ij,T}-\omega_{ij,T}\right)}{\hat{\omega}_{ij,T}}\right|\leq\frac{1}{K}\left|\sqrt{T}\left(\hat{\omega}_{ij,T}-\omega_{ij,T}\right)\right|$$

since for T sufficiently large $\inf\hat{\omega}_{ij,T}\geq K>0$ a.s. Further, we can write

$$\sqrt{T}\left(\hat{\omega}_{ij,T}-\omega_{ij,T}\right)=A_{ij,T}+B_{ij,T}$$

where

$$
\begin{aligned}
A_{ij,T} &= \sqrt{T}\left(\frac{1}{T}\sum_{t=1}^{T}\left\{\hat{u}_{it}^2\hat{u}_{jt}^2-u_{it}^2u_{jt}^2\right\}\right)\\
B_{ij,T} &= \sqrt{T}\left(\frac{1}{T}\sum_{t=1}^{T}u_{it}^2u_{jt}^2-\omega_{ij,T}\right)
\end{aligned}
$$

Now, $\sup_{i,j}\left|B_{ij,T}\right|=O_p(1)$ by Assumption B2(i), whereas for the term $A_{ij,T}$ from the proof of Theorem 1, $\frac{1}{T}\sum_{t=1}^{T}\left\{\hat{u}_{it}^2\hat{u}_{jt}^2-u_{it}^2u_{jt}^2\right\}=O_p\left(T^{-1/2}\right)$ uniformly in i,j, so that $\sup_{i,j}\left|A_{ij,T}\right|=O_p(1)$. Thus, it follows that $\sup_{i,j}\left|\frac{\sqrt{T}\left(\hat{\omega}_{ij,T}-\omega_{ij,T}\right)}{\hat{\omega}_{ij,T}}\right|=O_p(1)$. ∎

Asymptotic Validity of the Wild Bootstrap

We verify this for the recursive wild bootstrap scheme (WB1) only and, following Davidson and Flachaire (2008), with $u_{it}^*=\varepsilon_{it}\hat{u}_{it}$ where the ε_{it} are i.i.d for all i and t taking the discrete values ±1 with an equal probability of 0.5. With slight amendments, the proofs remain valid for any ε_{it} which are i.i.d mean zero and unit variance and the derivations for the other two bootstrap schemes are straightforward. Finally, and for simplicity, $y_{is}^* = 0$, for all $s \leq 0$, although the proofs can be adapted for the case of $y_{is}^* = y_{is}$, for all $s = -p+1, ..., 0$, so that from (8),

$$y_{it}^* = \sum_{k=0}^{t-1} \hat{\psi}_{ik} r_{i,t-k}^*$$

where $r_{it}^* = \hat{\theta}_i' w_{it} + u_{it}^*$ and $\hat{\psi}_{ik} = \psi_{ik}(\hat{\phi})$, satisfying $\hat{\psi}_{is} - \hat{\phi}_{i1}\hat{\psi}_{i,s-1} - ... - \hat{\phi}_{ip}\hat{\psi}_{i,s-p} = 0$, for all $s > 0$, with $\hat{\psi}_{i0} = 1$ and $\hat{\psi}_{ik} = 0$, $k < 0$, for all i. Furthermore, for $t = 1, ..., T$, $Y_{i,t-1}^*$ can be expressed as

$$\begin{aligned} Y_{i,t-1}^* &= \sum_{k=1}^{t-1} \hat{c}_{ik} r_{i,t-k}^* \\ &= \sum_{k=1}^{T-1} \hat{c}_{ik} b_{i,t-k}^* \end{aligned}$$

where $\hat{c}_{ik} = c_{ik}(\hat{\phi}) = \left(\hat{\psi}_{i,k-1}, ..., \hat{\psi}_{i,k-p}\right)'$, $b_{it}^* = \mathbf{1}(t > 0)\, r_{it}^*$, where $\mathbf{1}(.)$ is the usual binary indicator function since $r_{it}^* = 0$ for all $t \leq 0$.

We exploit the following definitions (as in Goncalves and Kilian, 2004). For any bootstrap statistic, S_T^*, we write $S_T^* = o_{p^*}(1)$, in probability, if for any $\delta > 0$, $P^*(\|S_T^*\| > \delta) = o_p(1)$, where P^* is the probability measure induced by the wild bootstrap conditional on the sample data. Similarly, $S_T^* = O_{p^*}(1)$, in probability, if for some $r > 0$ and all $\lambda > 0$, $P^*(\|S_T^*\| > \lambda) \leq M_T/\lambda^r$, and $M_T = E^*[\|S_T^*\|^r] = O_p(1)$, at most, where $E^*[.]$ denotes expectations induced by the wild bootstrap conditional on the sample data. Finally, $S_T^* \xrightarrow{d^*} \mathcal{D}$, in probability, for any distribution $\mathcal{D}$, when weak convergence under the bootstrap probability measure occurs in a set with probability converging to one; i.e., if the proposed limit distribution is $\mathcal{D}(x)$ then, $\sup_{x\in\mathbb{R}} |P^*(S_T^* \leq x) - \mathcal{D}(x)| = o_p(1)$.

The proof of the asymptotic validity for the wild bootstrap procedures for the heteroskedasticity robust statistic $NRBP_{NT}$ in Theorem 3 is based on the following lemmas, since $NRBP_{NT}$ has an asymptotic standard normal distribution as established in Theorem 2.

Lemma 5 *Under Assumptions A1-A4(i) and (iv) combined with/or strengthened by Assumptions* $B2(ii)$ *and* $B3$ *as* $(N,T) \xrightarrow{j} \infty$

$$\begin{aligned} Z_{NT}^* &= \frac{1}{\sqrt{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\gamma_{ij,T}^{*2} - 1\right) \qquad (25) \\ &= \frac{1}{\sqrt{\frac{1}{2}N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\frac{\gamma_{ij,T}^{*2} - 1}{\sqrt{2}}\right) \\ &\xrightarrow{d^*} N(0,1), \end{aligned}$$

where

$$\gamma_{ij,T}^* = \frac{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_{it}^* u_{jt}^*}{\sqrt{\omega_{ij,T}^*}}$$

with $\omega_{ij,T}^* = \frac{1}{T}\sum_{t=1}^{T} u_{it}^{*2} u_{jt}^{*2} = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_{it}^2 \hat{u}_{jt}^2$ *by construction of the wild bootstrap errors using the Rachemacher distribution and notice that* $E^*\left[\gamma_{ij,T}^{*2} - 1\right] = 0$.

Lemma 6 *Under Assumptions A1-A4 combined with Assumption B1*

$$NRBP_{NT}^* = Z_{NT}^* + o_{p*}(1)$$

as $(N,T) \to \infty$, $N^2/T \to 0$ *for the wild bootstrap designs WB1 and WB2. For the wild bootstrap design WB3, the limiting distribution of* $NRBP_{NT}^*$ *follows directly from Lemma 5 since* $NRBP_{NT}^* = Z_{NT}^*$.

Corresponding results apply for the wild bootstrap procedure based on the statistic NBP_{NT}. Specifically, $\hat{\rho}_{ij} = \sqrt{v_T^{ij}}\hat{\gamma}_{ij,T} + o_p(1)$ where (the scalar)

$$v_T^{ij} = \frac{\frac{1}{T}\sum_{t=1}^{T} E[u_{it}^2 u_{jt}^2]}{\frac{1}{T}\sum_{t=1}^{T} E[u_{it}^2]\ \frac{1}{T}\sum_{t=1}^{T} E[u_{jt}^2]} = O(1)$$

and is strictly positive for T sufficiently large, by Assumptions A3(ii) and A4(iv). Furthermore, for $\hat{\rho}_{ij}^*$ defined at (10), and by Lemma 8 (c) below, it is also true that $\hat{\rho}_{ij}^* = \sqrt{v_T^{ij}\gamma_{ij}^*} + o_{p^*}(1)$, in probability, since by the Davidson and Flachaire (2008) wild bootstrap scheme, $u_{it}^{*2} = \hat{u}_{it}^2$. Therefore, the result for NBP_{NT}^* in Theorem 3 follows from the following analysis for $NRBP_{NT}^*$.

In what follows, let $\mathcal{F}_{N,t}^*$ be the sigma field generated by current and lagged values of ε_{it} in the bootstrap sample (i.e., $\{\varepsilon_{i,t-p}\}$, $i = 1, ..., N$, $p = 0, 1, 2, ..., t-1$).

The following preliminary Lemmas inform the proofs of Lemmas 5 and 6 and thus of Theorem 3, with the first two lemmas being the bootstrap counterparts of Lemmas 3 and 4:

Lemma 7 *Consider a sequence of scalar bootstrap random variables denoted $\bar{Z}_{T,k}^*$ and a sequence of scalars, $\bar{\mu}_{T,k}$, indexed by $k \in \mathbb{N}$, such that: (i) $E^*\left|\bar{Z}_{T,k}^*\right| \leq M_T = O_p(1)$ uniformly in k, as $T \to \infty$; (ii) $\bar{Z}_{T,k}^* - \bar{\mu}_{T,k} = o_{p^*}(1)$, in probability, as $T \to \infty$, for each fixed $k \in \mathbb{N}$; and, (iii) $\left|\bar{\mu}_{T,k}\right| \leq \Delta < \infty$, uniformly in k and T. Define $\bar{S}_T^* = \sum_{k=1}^{T-1}\hat{\xi}_k\bar{Z}_{T,k}^* - \sum_{k=1}^{\infty}\xi_k\bar{\mu}_{T,k}$, where the $\hat{\xi}_k$ are scalar functions of the parameter estimators, such that, for each $k \in \mathbb{N}$, $\hat{\xi}_k - \xi_k = o_p(1)$, and $\sum_{k=1}^{\infty}|\xi_k| < \infty$. Then, $\bar{S}_T^* = o_{p^*}(1)$, in probability.*

Lemma 8 *Under Assumptions A1-A4, and uniformly in $i, j = 1, ..., N$:*

(a) $T^{-1}\sum_{t=1}^{T}\left(x_{it}^*x_{jt}^{*\prime} - E[x_{it}x_{jt}']\right) = o_{p^*}(1)$, *in probability;*

(b) $T^{-1/2}\sum_{t=1}^{T}x_{it}^*u_{jt}^* = O_{p^*}(1)$, *in probability;*

(c) $\frac{1}{T}\sum_{t=1}^{T}\left(\hat{u}_{it}^{*2}\hat{u}_{jt}^{*2} - E[u_{it}^2u_{jt}^2]\right) = o_{p^*}(1)$, *in probability;*

(d) $\hat{\gamma}_{ij,T}^{2*} = O_p^*(1)$, *in probability.*

Proof of Lemma 7. Write

$$\bar{S}_T^* = \bar{S}_T^{*n} + R_T^*,$$

where $\bar{S}_T^{*n} = \sum_{k=1}^{n-1}\hat{\xi}_k\bar{Z}_{T,k}^* - \sum_{k=1}^{n-1}\xi_k\bar{\mu}_{T,k}$, for any fixed $n < T$, and $R_T^* = \sum_{k=n}^{T-1}\hat{\xi}_k\bar{Z}_{T,k}^* - \sum_{k=n}^{\infty}\xi_k\bar{\mu}_{T,k}$. Consider $\bar{S}_T^{*n}$, which can be expressed

$$\begin{aligned}\bar{S}_T^{*n} &= \sum_{k=1}^{n-1}\xi_k(\bar{Z}_{T,k}^* - \bar{\mu}_{T,k}) + \sum_{k=1}^{n-1}(\hat{\xi}_k - \xi_k)\bar{Z}_{T,k}^* \\ &= S_{1T}^{*n} + S_{2T}^{*n}.\end{aligned}$$

First, since, $\bar{Z}_{T,k}^* - \bar{\mu}_{T,k} = o_{p^*}(1)$, in probability, for each $k \in \mathbb{N}$, $S_{1T}^{*n} = o_{p^*}(1)$, in probability. Second, $E^*\left|S_{2T}^{*n}\right| \leq M_T\sum_{k=1}^{n-1}\left|\hat{\xi}_k - \xi_k\right| = o_p(1)$, so by Markov's Inequality $S_{2T}^{*n} = o_{p^*}(1)$, in probability. It then suffices to show that for any $\delta > 0$, $\lim_{n\to\infty}\limsup_{T\to\infty}P^*\left(|R_T^*| > \delta\right) = 0$, in probability, or $\lim_{n\to\infty}\limsup_{T\to\infty}E^*\left(|R_T|\right) = 0$, in probability. To show this, note that

$$\begin{aligned}E^*\left(|R_T|\right) &\leq \sum_{k=n}^{T-1}\left|\hat{\xi}_k\right|E^*\left|\bar{Z}_{T,k}^*\right| + \sum_{k=n}^{\infty}|\xi_k|\left|\bar{\mu}_{T,k}\right| \\ &\leq M_T\sum_{k=n}^{\infty}\left|\hat{\xi}_k\right| + \Delta\sum_{k=n}^{\infty}|\xi_k|\end{aligned}$$

where $M_T = O_p(1)$ and $\Delta = O(1)$. Since $\hat{\xi}_k - \xi_k = o_p(1)$, and $\sum_{k=1}^{\infty}|\xi_k| < \infty$, there exists a T_1 such that $\sup_{T\geq T_1}\sum_{k=1}^{\infty}\left|\hat{\xi}_k\right| < \infty$, in probability (c.f. Bühlmann, 1995, Lemma 2.2) which implies that $\sup_{T\geq T_1}\sum_{k=n}^{\infty}\left|\hat{\xi}_k\right| = o_p(1)$ as $n \to \infty$. Thus

$$\lim_{n\to\infty}\limsup_{T\to\infty}E^*\left(|R_T|\right) = o_p(1)$$

which completes the proof. ■

Proof of Lemma 8. **(a)** Consider the corresponding conformable partitions of $\frac{1}{T}\sum_{t=1}^{T}x_{it}^*x_{jt}^{*\prime}$. Since we already have that $T^{-1}\sum_{t=1}^{T}\left(w_{it}w_{jt}' - E[w_{it}w_{jt}']\right) = o_p(1)$, it suffices to show that uniformly in i, j:
(i) $T^{-1}\sum_{t=1}^{T}\left(w_{it}Y_{j,t-1}^{*\prime} - E[w_{it}Y_{j,t-1}']\right) = o_{p^*}(1)$, in probability,and,

(ii) $T^{-1}\sum_{t=1}^{T}\left(Y_{i,t-1}^{*}Y_{j,t-1}^{\prime *}-E[Y_{i,t-1}Y_{j,t-1}^{\prime}]\right)=o_{p^*}(1)$, in probability

For **(i)**, exploiting $b_{it}^{*}=\mathbf{1}\,(t>0)\,(\hat{\theta}_i^{\prime}w_{it}+u_{it}^{*})$, we can write for any $\mu\in\mathbb{R}^M$ and any $\lambda\in\mathbb{R}^p$ such that $\|\mu\|=\|\lambda\|=1$,

$$\mu\left\{\frac{1}{T}\sum_{t=1}^{T}\left(w_{it}Y_{j,t-1}^{\prime *}-E[w_{it}Y_{j,t-1}^{\prime}]\right)\right\}\lambda=\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{k=1}^{T-1}\hat{\xi}_{jk}v_{it}b_{j,t-k}^{*}-\sum_{k=1}^{\infty}\xi_{jk}E[v_{it}r_{j,t-k}]\right)$$

where $v_{it}=\mu^{\prime}w_{it}$ and $\hat{\xi}_{jk}=\hat{c}_{jk}^{\prime}\lambda$, $\xi_{jk}=c_{jk}^{\prime}\lambda$, such that $\hat{\xi}_{jk}-\xi_{jk}=o_p(1)$, uniformly in j, and $\sum_{k=1}^{\infty}|\xi_{jk}|\leq\Delta<\infty$, for all j. Thus,

$$\begin{aligned}\mu^{\prime}\left\{\frac{1}{T}\sum_{t=1}^{T}\left(w_{it}Y_{j,t-1}^{\prime *}-E[w_{it}Y_{j,t-1}^{\prime}]\right)\right\}\lambda &= \sum_{k=1}^{T-1}\hat{\xi}_{jk}\bar{Z}_{T,k}^{*(i,j)}-\sum_{k=1}^{\infty}\xi_{jk}\bar{\mu}_{T,k}^{(i,j)}\\ &= \bar{S}_{T}^{*(i,j)},\text{ say,}\end{aligned}$$

where

$$\begin{aligned}\bar{Z}_{T,k}^{*(i,j)} &= \mu^{\prime}\frac{1}{T}\sum_{t=1}^{T}w_{it}b_{j,t-k}^{*}\\ &= \mu^{\prime}\left\{\frac{1}{T}\sum_{t=k+1}^{T}\left(w_{it}w_{j,t-k}^{\prime}\right)\right\}\hat{\theta}_j+\mu^{\prime}\frac{1}{T}\sum_{t=k+1}^{T}w_{it}u_{j,t-k}^{*}\end{aligned}$$

and

$$\begin{aligned}\bar{\mu}_{T,k}^{(i,j)} &= \mu^{\prime}\frac{1}{T}\sum_{t=1}^{T}E[w_{it}r_{j,t-k}]\\ &= \mu^{\prime}\left\{\frac{1}{T}\sum_{t=1}^{T}E[w_{it}w_{j,t-k}^{\prime}]\right\}\theta_j.\end{aligned}$$

Now apply Lemma 7 to $\bar{S}_{T}^{*(i,j)}$. First, to establish that $\bar{Z}_{T,k}^{*(i,j)}-\bar{\mu}_{T,k}^{(i,j)}=o_{p^*}(1)$, in probability, uniformly in i,j, note that for any fixed $k\in\mathbb{N}$,

$$\begin{aligned}\bar{Z}_{T,k}^{*(i,j)}-\bar{\mu}_{T,k}^{(i,j)} &= \mu^{\prime}\left\{\frac{1}{T}\sum_{t=k+1}^{T}\left(w_{it}w_{j,t-k}^{\prime}-E[w_{it}w_{j,t-k}^{\prime}]\right)\right\}\theta_j+\mu^{\prime}\frac{1}{T}\sum_{t=k+1}^{T}w_{it}u_{j,t-k}^{*}\\ &\quad+\mu^{\prime}\left\{\frac{1}{T}\sum_{t=1}^{k}E[w_{it}w_{j,t-k}^{\prime}]\right\}(\hat{\theta}_j-\theta_j),\end{aligned}$$

so that

$$\begin{aligned}\bar{Z}_{T,k}^{*(i,j)}-\bar{\mu}_{T,k}^{(i,j)} &= \mu^{\prime}\frac{1}{T}\sum_{t=k+1}^{T}w_{it}u_{j,t-k}^{*}+o_{p^*}(1)\\ &= \mu^{\prime}\frac{1}{T}\sum_{t=1}^{T-k}w_{i,t+k}u_{jt}^{*}+o_{p^*}(1).\end{aligned}$$

It follows that, conditional on the original sample,

$$\begin{aligned}E^{*}\left[\mu^{\prime}w_{i,t+k}u_{jt}^{*}|\mathcal{F}_{NT,t-1}^{*}\right] &= \mu^{\prime}w_{i,t+k}E^{*}\left[\varepsilon_{jt}\hat{u}_{jt}|\mathcal{F}_{NT,t-1}^{*}\right]\\ &= \mu^{\prime}w_{i,t+k}\hat{u}_{jt}E^{*}\left[\varepsilon_{jt}|\mathcal{F}_{NT,t-1}^{*}\right]\\ &= 0\end{aligned}$$

so that $\left\{\mu^{\prime}w_{i,t+k}u_{jt}^{*},\mathcal{F}_{N,t}^{*}\right\}$ is a m.d.s. and, by Cauchy-Schwartz,

$$\begin{aligned}var^{*}\left[\mu^{\prime}\frac{1}{T}\sum_{t=1}^{T-k}w_{i,t+k}u_{jt}^{*}\right] &\leq \frac{1}{T^2}\sum_{t=1}^{T-k}\hat{u}_{jt}^{2}\|w_{i,t+k}\|^{2}\\ &\leq \frac{1}{T}\sqrt{\frac{1}{T}\sum_{t=1}^{T-k}\hat{u}_{jt}^{4}\frac{1}{T}\sum_{t=1}^{T-k}\|w_{i,t+k}\|^{4}}\\ &= O_p(T^{-1})\end{aligned}$$

uniformly in i, j, since by Minkowski's inequality

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}|\hat{u}_{it}|^q &\leq \left(\left\{\frac{1}{T}\sum_{t=1}^{T}|u_{it}|^q\right\}^{1/q}+\left\{\frac{1}{T}\sum_{t=1}^{T}\left|x_{it}'\left(\hat{\beta}_i-\beta_i\right)\right|^q\right\}^{1/q}\right)^q \\
&\leq \left(\left\{\frac{1}{T}\sum_{t=1}^{T}|u_{it}|^q\right\}^{1/q}+\left\|\hat{\beta}_i-\beta_i\right\|\left\{\frac{1}{T}\sum_{t=1}^{T}\|x_{it}\|^q\right\}^{1/q}\right)^q \qquad (26)
\end{aligned}
$$

which is $O_p(1)$, uniformly in i, for $q=4$, under our Assumptions. Specifically, since $x_{it}'=(w_{it}', Y_{i,t-1}')$ we only need to show that $\sup_i E\|Y_{i,t-1}\|^q \leq \Delta < \infty$ given Assumption 2(iv). Applying Minkowski's inequality, we can write

$$
E\|Y_{i,t-1}\|^q \leq \left(\sum_{k=1}^{\infty}\|c_{ik}\|\left(E|r_{i,t-k}|^q\right)^{\frac{1}{q}}\right)^q
$$

and by another application of Minkowski's inequality

$$
E|r_{it}|^q \leq \left(\|\theta_i\|\left(E\|w_{it}\|^q\right)^{\frac{1}{q}}+\left(E|u_{it}|^q\right)^{\frac{1}{q}}\right)^q < \infty.
$$

Therefore, by Chebyshev's inequality $\bar{Z}_{T,k}^{*(i,j)}-\bar{\mu}_{T,k}^{(i,j)}=o_{p^*}(1)$, uniformly in i, j.
Second, by the triangle inequality and noting that $|\varepsilon_{j,t-k}|=1$,

$$
\begin{aligned}
E^*\left|\bar{Z}_{T,k}^{*(i,j)}\right| &\leq \left\|\hat{\theta}_j\right\|\left\|\frac{1}{T}\sum_{t=k+1}^{T}w_{it}w_{j,t-k}\right\|+E^*\left\|\frac{1}{T}\sum_{t=k+1}^{T}w_{it}u_{j,t-k}^*\right\| \\
&\leq \left\|\hat{\theta}_j\right\|\left\|\frac{1}{T}\sum_{t=k+1}^{T}w_{it}w_{j,t-k}\right\|+\frac{1}{T}\sum_{t=k+1}^{T}\|w_{it}\hat{u}_{j,t-k}\|,
\end{aligned}
$$

where,

$$
T^{-1}\sum_{t=k+1}^{T}\|w_{it}\hat{u}_{j,t-k}\| \leq \left\{\left(T^{-1}\sum_{t=1}^{T}\|w_{it}\|^2\right)\left(T^{-1}\sum_{t=1}^{T}|\hat{u}_{jt}|^2\right)\right\}^{1/2},
$$

and thus $O_p(1)$, uniformly in i, j, k. Similarly, $\left\|\hat{\theta}_j\right\|\left\|\frac{1}{T}\sum_{t=k+1}^{T}w_{it}w_{j,t-k}\right\|$ is $O_p(1)$, uniformly in i, j, k. Thus $E^*\left|\bar{Z}_{T,k}^{*(i,j)}\right| \leq M_T = O_p(1)$ uniformly in k..
Third, $\left|\bar{\mu}_{T,k}^{(i,j)}\right| \leq \|\theta_j\|\frac{1}{T}\sum_{t=1}^{T}E\|w_{it}\|^2 \leq \Delta < \infty$, uniformly in i, j, k, by the triangle inequality, Assumption A2(iv), and Cauchy-Schwartz, and we are done.
For **(ii)**, we can write, for any $\lambda \in \mathbb{R}^p$ such that $\|\lambda\|=1$,

$$
\begin{aligned}
\lambda'\frac{1}{T}\sum_{t=1}^{T}\left(Y_{i,t-1}^{*}Y_{j,t-1}^{*\prime}-E[Y_{i,t-1}Y_{j,t-1}']\right)\lambda &= \sum_{k=1}^{T-1}\sum_{h=1}^{T-1}\hat{\xi}_{ik}\hat{\xi}_{jh}\bar{Z}_{T,k,h}^{*(i,j)}-\sum_{k=1}^{\infty}\sum_{h=1}^{\infty}\xi_{ik}\xi_{jh}\bar{\mu}_{T,k,h}^{(i,j)} \\
&= \bar{S}_T^{*(i,j)}, \text{ say,}
\end{aligned}
$$

where

$$
\begin{aligned}
\bar{Z}_{T,k,h}^{*(i,j)} &= \hat{\theta}_i'\left\{\frac{1}{T}\sum_{t=\max(k,h)+1}^{T}w_{i,t-k}w_{j,t-h}'\right\}\hat{\theta}_j \\
&\quad +\hat{\theta}_i'\frac{1}{T}\sum_{t=\max(k,h)+1}^{T}w_{i,t-k}u_{j,t-h}^*+\hat{\theta}_j'\frac{1}{T}\sum_{t=\max(k,h)+1}^{T}w_{j,t-h}u_{i,t-k}^* \\
&\quad +\frac{1}{T}\sum_{t=\max(k,h)+1}^{T}u_{i,t-k}^*u_{j,t-h}^*,
\end{aligned}
$$

and

$$
\begin{aligned}
\bar{\mu}_{T,k,h}^{(i,j)} &= \theta_i'\left\{\frac{1}{T}\sum_{t=1}^{T}E[w_{i,t-k}w_{j,t-h}']\right\}\theta_j \\
&\quad +\frac{1}{T}\sum_{t=1}^{T}E\left[u_{i,t-k}u_{j,t-h}\right].
\end{aligned}
$$

Again, we apply Lemma 7 (twice), to $\bar{S}_T^{*(i,j)}$. First, and by arguments similar to those used above, $\sup_{i,j} E^* \left| \bar{Z}_{T,k,h}^{*(i,j)} \right| \leq M_T = O_p(1)$, uniformly in k and h, noting that

$$E^* \left| \frac{1}{T} \sum_{t=\max(k,h)+1}^{T} u_{i,t-k}^* u_{j,t-h}^* \right| \leq \frac{1}{T} \sum_{t=1}^{T} \hat{u}_{it}^2.$$

Second

$$\begin{aligned} \left| \bar{\mu}_{T,k,h}^{(i,j)} \right| &\leq \|\theta_i\| \|\theta_j\| \frac{1}{T} \sum_{t=1}^{T} E \|w_{it}\|^2 + \frac{1}{T} \sum_{t=1}^{T} E |u_{it}|^2 \\ &\leq \Delta < \infty. \end{aligned}$$

Finally, we can write

$$\bar{Z}_{T,k,h}^{*(i,j)} - \bar{\mu}_{T,k,h}^{(i,j)} = \frac{1}{T} \sum_{t=\max(k,h)+1}^{T} \left(u_{i,t-k}^* u_{j,t-h}^* - E\left[u_{i,t-k} u_{j,t-h} \right] \right) + o_p(1),$$

where the $o_p(1)$ term incorporates: $\frac{1}{T} \sum_{t=\max(k,h)+1}^{T} \left(w_{i,t-k} w_{j,t-h}' - E\left[w_{i,t-k} w_{j,t-h}' \right] \right) = o_p(1)$, $\frac{1}{T} \sum_{t=1}^{\max(k,h)} E[w_{i,t-k} w_{j,t-h}'] = O(T^{-1})$, $\frac{1}{T} \sum_{t=1}^{\max(k,h)} E[u_{i,t-k} u_{j,t-h}'] = O(T^{-1})$, uniformly in i,j, and similar arguments to before show that, for example, $\frac{1}{T} \sum_{t=\max(k,h)+1}^{T} w_{i,t-k} u_{j,t-h}^* = o_{p^*}(1)$, uniformly in i,j, for all fixed $k,h \in \mathbb{N}$. For the remaining term, consider first $i \neq j$. Then for all fixed $k,h \in \mathbb{N}$, $E\left[u_{i,t-k} u_{j,t-h} \right] = 0$ by Assumption A4(i), $\left\{ u_{i,t-k}^* u_{j,t-h}^*, \mathcal{F}_{N,t-g}^* \right\}$, $g = \min(k,h)$, is a m.d.s. and it can be shown that $\frac{1}{T} \sum_{t=\max(k,h)+1}^{T} u_{i,t-k}^* u_{j,t-h}^* = o_{p^*}(1)$. In a similar fashion, for $i = j$ and $k \neq h$, $E\left[u_{i,t-k} u_{i,t-h} \right] = 0$ by Assumption A2(i) and $\frac{1}{T} \sum_{t=\max(k,h)+1}^{T} u_{i,t-k}^* u_{i,t-h}^* = o_{p^*}(1)$. Now, for $i = j$ and $k = h$, we have $u_{i,t-k}^{*2} = \hat{u}_{i,t-k}^2$, and we have previously argued that $\frac{1}{T} \sum_{t=k+1}^{T} \left(\hat{u}_{i,t-k}^2 - E[u_{i,t-k}^2] \right) = o_p(1)$. Thus, $\bar{Z}_{T,k,h}^{*(i,j)} - \bar{\mu}_{T,k,h}^{(i,j)} = o_{p^*}(1)$, uniformly in i,j, and we are done.

(b) First consider $d_{1T}^{*(i,j)} = T^{-1/2} \sum_{t=1}^{T} w_{it} u_{jt}^*$. Now, for any $\mu \in \mathbb{R}^M$ such that $\|\mu\| = 1$, $\left\{ \mu' w_{it} u_{jt}^*, \mathcal{F}_{N,t}^* \right\}$ is a m.d.s, and similar arguments to before show that $var^* \left[\mu' d_{1T}^{*(i,j)} \right] = O_p(1)$ and the result follows from Chebyshev's inequality.

Second, for any $\lambda \in \mathbb{R}^p$ such $\|\lambda\| = 1$, consider

$$\begin{aligned} \lambda' d_{2T}^{*(i,j)} &= T^{-1/2} \sum_{t=1}^{T} \lambda' Y_{i,t-1}^* u_{jt}^* \\ &= T^{-1/2} \sum_{t=1}^{T} \sum_{k=1}^{t-1} \hat{\xi}_{ik} \hat{r}_{i,t-k}^* u_{jt}^* \\ &= T^{-1/2} \sum_{t=1}^{T} \sum_{k=1}^{T-1} \hat{\xi}_{ik} b_{i,t-k}^* u_{jt}^*. \end{aligned}$$

where $\hat{\xi}_{ik} = \lambda' \hat{c}_{ik}$, $b_{it}^* = \mathbf{1}(t > 0) \hat{r}_{it}^*$, and it suffices to show that $T^{-1/2} \sum_{t=1}^{T} \sum_{k=1}^{T-1} \hat{\xi}_{ik} b_{i,t-k}^* u_{jt}^* = O_{p^*}(1)$, in probability,uniformly in i,j,. We have

$$T^{-1/2} \sum_{t=1}^{T} \sum_{k=1}^{T-1} \hat{\xi}_{ik} b_{i,t-k}^* u_{jt}^* = T^{-1/2} \sum_{t=1}^{T} u_{jt}^* \kappa_{it}^*$$

where $\kappa_{it}^* = \sum_{k=1}^{T-1} \hat{\xi}_{ik} b_{i,t-k}^* = \sum_{k=1}^{t-1} \hat{\xi}_{ik} r_{i,t-k}^*$, is simply a function of $\hat{\theta}_i' w_{is} + u_{is}^*$, $s = 1, ..., t-1$. Therefore, $\left\{ u_{jt}^* \kappa_{it}^*, \mathcal{F}_{N,t}^* \right\}$ is a m.d.s. and, since $|\varepsilon_{it}| = 1$,

$$\begin{aligned} var^* \left[T^{-1/2} \sum_{t=1}^{T} \sum_{k=1}^{T-1} \hat{\xi}_{ik} b_{i,t-k}^* u_{jt}^* \right] &= \frac{1}{T} \sum_{t=1}^{T} E^*[u_{jt}^{*2} \kappa_{it}^{*2}] \\ &= \frac{1}{T} \sum_{t=1}^{T} \hat{u}_{it}^2 E^* \left[\left\{ \sum_{k=1}^{T-1} \hat{\xi}_{ik} b_{i,t-k}^* \right\}^2 \right] \end{aligned}$$

and it suffices to show that this is $O_p(1)$. By the triangle inequality and Cauchy-Schwartz we can write

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\hat{u}_{it}^2E^*\left[\left\{\sum_{k=1}^{T-1}\hat{\xi}_{ik}b_{i,t-k}^*\right\}^2\right] &\leq \sum_{k=1}^{T-1}\sum_{h=1}^{T-1}\left|\hat{\xi}_{ik}\hat{\xi}_{ih}\right|\ E^*\left\{\frac{1}{T}\sum_{t=1}^{T}\hat{u}_{it}^2\left|b_{i,t-k}^*b_{i,t-h}^*\right|\right\}\\ &\leq \sum_{k=1}^{T-1}\sum_{h=1}^{T-1}\left|\hat{\xi}_{ik}\hat{\xi}_{ih}\right|\ E^*\sqrt{\frac{1}{T}\sum_{t=1}^{T}\hat{u}_{it}^4\ \frac{1}{T}\sum_{t=1}^{T}\left|b_{i,t-k}^*b_{i,t-h}^*\right|^2}\\ &\leq \left\{\sum_{k=1}^{\infty}\left|\hat{\xi}_{ik}\right|\right\}^2\ E^*\sqrt{\frac{1}{T}\sum_{t=1}^{T}\hat{u}_{it}^4\ \frac{1}{T}\sum_{t=1}^{T}\left|r_{it}^*\right|^4},\end{aligned}$$

Now, we have previously shown that $\sup_i \frac{1}{T}\sum_{t=1}^{T}\hat{u}_{it}^4 = O_p(1)$. Using this, and noting that $|r_{it}^*| \leq ||\hat{\theta}_i||\,\|w_{it}\| + |\hat{u}_{it}|$, it can similarly be shown by Minkowski's inequality that $\sup_i \frac{1}{T}\sum_{t=1}^{T}|r_{it}^*|^4 \leq M_T = O_p(1)$. Finally, there exists a T_1 such that $\sup_{i,T\geq T_1}\sum_{k=1}^{\infty}\left|\hat{\xi}_{ik}\right| = O_p(1)$. These results are sufficient to show that $\sup_{i,j} d_{2T}^{*(i,j)} = O_{p^*}(1)$, in probability.

(c) Since $\hat{\beta}_i^* - \hat{\beta}_i = o_{p^*}(1)$, in probability, uniformly in i, it is sufficient to show that $\frac{1}{T}\sum_{t=1}^{T}y_{i,t-h}^{*4} = O_{p^*}(1)$, in probability, uniformly in i, since then $\frac{1}{T}\sum_{t=1}^{T}\left(\hat{u}_{it}^{*2}\hat{u}_{jt}^{*2} - \hat{u}_{it}^2\hat{u}_{jt}^2\right) = o_{p^*}(1)$, in probability, uniformly in i,j, given $\varepsilon_{it}^2 = 1$. The result follows by the triangle inequality and $\frac{1}{T}\sum_{t=1}^{T}\left(\hat{u}_{it}^2\hat{u}_{jt}^2 - E[u_{it}^2u_{jt}^2]\right) \xrightarrow{p} 0$. Briefly, let e_h be the $(p\times 1)$ unit vector with 1 at position h and zeros elsewhere and let $\hat{\xi}_{ik} = \hat{c}_{ik}'e_h$. Then

$$\begin{aligned}\left|\frac{1}{T}\sum_{t=1}^{T}y_{i,t-h}^{*4}\right| &\leq \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{k=1}^{T-1}\left|\hat{\xi}_{ik}b_{i,t-k}^*\right|\right)^4\\ &= \sum_{k_1=1}^{T-1}\sum_{k_2=1}^{T-1}\sum_{k_3=1}^{T-1}\sum_{k_4=1}^{T-1}\left|\hat{\xi}_{ik_1}\hat{\xi}_{ik_2}\hat{\xi}_{ik_3}\hat{\xi}_{ik_4}\right|\\ &\quad\times\frac{1}{T}\sum_{t=1}^{T}\left|b_{i,t-k_1}^*b_{i,t-k_2}^*b_{i,t-k_3}^*b_{i,t-k_4}^*\right|\\ &\leq \left\{\sum_{k=1}^{\infty}\left|\hat{\xi}_{ik}\right|\right\}^4\left\{\frac{1}{T}\sum_{t=1}^{T}\left|r_{it}^*\right|^4\right\}^2.\end{aligned}$$

But, $T^{-1}\sum_{t=1}^{T}|\hat{r}_{it}^*|^4 \leq M_T = O_p(1)$, so $\frac{1}{T}\sum_{t=1}^{T}y_{i,t-h}^{*4} = O_{p^*}(1)$, in probability, uniformly in i.

(d) Firstly, we show that

$$\hat{\gamma}_{ij}^* = \gamma_{ij}^* + o_{p^*}(1), \tag{27}$$

Note that

$$T^{-1/2}\sum_{t=1}^{T}\hat{u}_{it}^*\hat{u}_{jt}^* = T^{-1/2}\sum_{t=1}^{T}u_{it}^*u_{jt}^* - T^{-1/2}\left\{u_i^{*\prime}H_i^*u_j^* - u_i^{*\prime}H_j^*u_j^* + u_i^{*\prime}H_i^*H_j^*u_j^*\right\}.$$

It is immediate from (a) and (b), and Lemma 4(b), that the terms $u_i^{*\prime}H_i^*u_j^*$, $u_i^{*\prime}H_j^*u_j^*$ and $u_i^{*\prime}H_i^*H_j^*u_j^*$ are all $O_{p^*}(1)$, in probability, uniformly in $i,j,$. Furthermore, since $\sup_{i,j}\frac{1}{T}\sum_{t=1}^{T}\left(\hat{u}_{it}^2\hat{u}_{jt}^2 - E[u_{it}^2u_{jt}^2]\right) = o_p(1)$, the result in (c) and the triangle inequality gives $\frac{1}{T}\sum_{t=1}^{T}\hat{u}_{it}^{*2}\hat{u}_{jt}^{*2} - \frac{1}{T}\sum_{t=1}^{T}\hat{u}_{it}^2\hat{u}_{jt}^2 = o_{p^*}(1)$, in probability, uniformly in $i,j,$. The result in (27) follows immediately. Moreover, by Markov's inequality $\sup_{i,j}\gamma_{ij}^{*2} = O_{p^*}(1)$, in probability, since $\sup_{i,j}E^*\left|\gamma_{ij}^{*2}\right| = 1$ by (conditional) independence and $u_{it}^{*2} = \hat{u}_{it}^2$. ■

Proof of Lemma 5. Firstly write

$$\begin{aligned}Z_{NT}^* &= 2\frac{1}{T}\frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{u_{it}^*u_{jt}^*u_{is}^*u_{js}^*}{\omega_{ij,T}^*}\\ &\quad+\frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\left(\frac{\frac{1}{T}\sum_{t=1}^{T}u_{it}^{*2}u_{jt}^{*2}}{\omega_{ij,T}^*} - 1\right)\\ &= Z_{1,NT}^* + Z_{2,NT}^*\end{aligned}$$

For the second term, using the Rademacher distribution for generating the bootstrap errors, $\frac{1}{T}\sum_{t=1}^{T} u_{it}^{*2} u_{jt}^{*2} = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_{it}^{2} \hat{u}_{jt}^{2}$ and thus $Z_{2,NT}^{*} = 0$.
Consider now the first term

$$\begin{aligned} Z_{1,NT}^{*} &= \sum_{t=2}^{T}\sum_{s=1}^{t-1} H_{T}^{*}\left(\underline{u}_{t}^{*}, \underline{u}_{s}^{*}\right) \\ &= \sum_{t=2}^{T} W_{Tt}^{*} \end{aligned}$$

where $\underline{u}_{t}^{*} = (u_{1t}^{*}, ..., u_{Nt}^{*})'$ and

$$H_{T}^{*}\left(\underline{u}_{t}^{*}, \underline{u}_{s}^{*}\right) = 2\frac{1}{T}\frac{1}{\sqrt{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{u_{it}^{*}u_{jt}^{*}u_{is}^{*}u_{js}^{*}}{\omega_{ij,T}^{*}}.$$

Note that $E^{*}\left[W_{Tt}^{*}|\mathcal{F}_{NT,t-1}^{*}\right] = E^{*}\left[W_{Tt}^{*}\right] = 0$ due to (conditional) independence where $E^{*}[\cdot]$ denotes the expectation induced by the wild bootstrap conditional on the sample data. Therefore, we apply the CLT theorem for U-statistic for (conditionally) independent but heterogenous data, for which it suffices to check as $(T, N) \to \infty$:
(i) $s_{T}^{*2} \to 1$, where

$$s_{T}^{*2} = E^{*}\left[\left(\sum_{t=2}^{T} W_{Tt}^{*}\right)^{2}\right]$$

(ii) $s_{T}^{*-2}\sum_{t=2}^{T} E^{*}\left[W_{Tt}^{*2} 1\left(|W_{Tt}^{*}| > \delta s_{T}^{*}\right)|\mathcal{F}_{NT,t-1}^{*}\right] \to 0$ for all $\delta > 0$
Then

$$Z_{1,NT}^{*} = \sum_{t=2}^{T} W_{Tt}^{*} \xrightarrow{d} N(0,1)$$

as $(T, N) \to \infty$.
For **(i)**,

$$\begin{aligned} s_{T}^{*2} &= E^{*}\left[\sum_{t=2}^{T}\sum_{t'=2}^{T} W_{Tt}^{*}W_{Tt'}^{*}\right] \\ &= E^{*}\left[\sum_{t=2}^{T} W_{Tt}^{*2}\right] \\ &= \sum_{t=2}^{T} E^{*}\left[\sum_{s=1}^{t-1}\sum_{r=1}^{t-1} H_{T}^{*}\left(\underline{u}_{t}^{*}, \underline{u}_{s}^{*}\right) H_{T}^{*}\left(\underline{u}_{t}^{*}, \underline{u}_{r}^{*}\right)\right] \\ &= \sum_{t=2}^{T}\sum_{s=1}^{t-1} E^{*}\left[H_{T}^{*2}\left(\underline{u}_{t}^{*}, \underline{u}_{s}^{*}\right)\right] \end{aligned}$$

where the second and third lines follow by m.d.s. and conditional independence of the bootstrap and the last line follows since

$$E^{*}\left[H_{T}^{*}\left(\underline{u}_{t}^{*}, \underline{u}_{s}^{*}\right) H_{T}^{*}\left(\underline{u}_{t}^{*}, \underline{u}_{r}^{*}\right)\right] = \frac{4}{T^{2}N(N-1)} E^{*}\left[\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\sum_{l=1}^{N-1}\sum_{m=l+1}^{N}\frac{u_{it}^{*}u_{jt}^{*}u_{is}^{*}u_{js}^{*}}{\omega_{ij,T}^{*}}\frac{u_{lt}^{*}u_{mt}^{*}u_{lr}^{*}u_{mr}^{*}}{\omega_{lm,T}^{*}}\right]$$

and for $s \neq r$

$$\begin{aligned} E^{*}\left[\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\sum_{l=1}^{N-1}\sum_{m=l+1}^{N} u_{it}^{*}u_{jt}^{*}u_{is}^{*}u_{js}^{*}u_{lt}^{*}u_{mt}^{*}u_{lr}^{*}u_{mr}^{*}\right] &= \sum_{i=l\neq j=m} E^{*}\left[u_{it}^{*2}u_{jt}^{*2}u_{is}^{*}u_{js}^{*}u_{ir}^{*}u_{jr}^{*}\right] \\ &\quad + \sum_{i\neq j\neq l\neq m} E^{*}\left[u_{it}^{*}u_{jt}^{*}u_{is}^{*}u_{js}^{*}u_{lt}^{*}u_{mt}^{*}u_{lr}^{*}u_{mr}^{*}\right] \\ &\quad +3\sum_{i\neq j=l\neq m} E^{*}\left[u_{it}^{*}u_{jt}^{*2}u_{is}^{*}u_{js}^{*}u_{mt}^{*}u_{jr}^{*}u_{mr}^{*}\right] \\ &= 0 \end{aligned}$$

by conditional independence and construction of the bootstrap. Thus

$$
\begin{aligned}
s_T^{*2} &= \frac{4}{T^2 N(N-1)} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{E^*\left[u_{it}^{*2} u_{jt}^{*2} u_{is}^{*2} u_{js}^{*2}\right]}{\omega_{ij,T}^{*2}} \\
&= \frac{4}{T^2 N(N-1)} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\hat{u}_{it}^2 \hat{u}_{jt}^2 \hat{u}_{is}^2 \hat{u}_{js}^2}{\omega_{ij,T}^{*2}}
\end{aligned}
$$

and we can write further that

$$
2\frac{1}{T^2} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \hat{u}_{it}^2 \hat{u}_{jt}^2 \hat{u}_{is}^2 \hat{u}_{js}^2 = \frac{1}{T^2} \left(\sum_{t=1}^{T} \hat{u}_{it}^2 \hat{u}_{jt}^2 \right)^2 - \frac{1}{T^2} \sum_{t=1}^{T} \hat{u}_{it}^4 \hat{u}_{jt}^4
$$

Now $\frac{1}{T}\sum_{t=1}^{T} \hat{u}_{it}^4 \hat{u}_{jt}^4 = O_p(1)$, uniformly in i, j, by Cauchy-Schwartz inequality and Assumption B2(ii), since $\hat{u}_{it} = u_{it} - x_{it}'\left(\hat{\beta}_i - \beta_i\right)$. Specifically

$$
\begin{aligned}
\frac{1}{T} \sum_{t=1}^{T} \hat{u}_{it}^4 \hat{u}_{jt}^4 &\leq \sqrt{\frac{1}{T} \sum_{t=1}^{T} \hat{u}_{it}^8 \frac{1}{T} \sum_{t=1}^{T} \hat{u}_{jt}^8} \\
&= O_p(1)
\end{aligned}
$$

uniformly in i, j, by Assumption B2(ii) and B3 following arguments in (26) for $q = 8$. Therefore

$$
\begin{aligned}
2\frac{1}{T^2} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \hat{u}_{it}^2 \hat{u}_{jt}^2 \hat{u}_{is}^2 \hat{u}_{js}^2 &= \left(\frac{1}{T} \sum_{t=1}^{T} \hat{u}_{it}^2 \hat{u}_{jt}^2 \right)^2 + O_p\left(T^{-1}\right) \\
&= \omega_{ij,T}^{*2} + O_p\left(T^{-1}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
s_T^{*2} &= \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\omega_{ij,T}^{*2}}{\omega_{ij,T}^{*2}} + O_p\left(T^{-1}\right) \\
&= 1 + o_p(1).
\end{aligned}
$$

To establish **(ii)** it is sufficient to show that $\sum_{t=2}^{T} E^*\left[W_{Tt}^{*4}\right] = o_p(1)$, where

$$
\begin{aligned}
E^*\left[W_{Tt}^{*4}\right] &= E^*\left[\left(\sum_{s=1}^{t-1} H_T^*\left(\underline{u}_t^*, \underline{u}_s^*\right) \right)^4 \right] \\
&= \sum_{s=1}^{t-1} E^*\left[H_T^{*4}\left(\underline{u}_t^*, \underline{u}_s^*\right)\right] + 3 \sum_{s \neq r}^{t-1} \sum^{t-1} E^*\left[H_T^{*2}\left(\underline{u}_t^*, \underline{u}_s^*\right) H_T^{*2}\left(\underline{u}_t^*, \underline{u}_r^*\right)\right] \\
&\quad + \sum_{s \neq r \neq s' \neq r'}^{t-1} \sum^{t-1} \sum^{t-1} \sum^{t-1} E^*\left[H_T^*\left(\underline{u}_t^*, \underline{u}_s^*\right) H_T^*\left(\underline{u}_t^*, \underline{u}_r^*\right) H_T^*\left(\underline{u}_t^*, \underline{u}_{s'}^*\right) H_T^*\left(\underline{u}_t^*, \underline{u}_{r'}^*\right)\right] \\
&\quad + 6 \sum_{s \neq r}^{t-1} \sum^{t-1} \sum^{t-1} E^*\left[H_T^{*2}\left(\underline{u}_t^*, \underline{u}_s^*\right) H_T^*\left(\underline{u}_t^*, \underline{u}_r^*\right) H_T^*\left(\underline{u}_t^*, \underline{u}_{s'}^*\right)\right] \\
&\quad + 4 \sum_{s \neq r}^{t-1} \sum^{t-1} E^*\left[H_T^{*3}\left(\underline{u}_t^*, \underline{u}_s^*\right) H_T^*\left(\underline{u}_t^*, \underline{u}_r^*\right)\right]. \qquad (28)
\end{aligned}
$$

The first term in (28) is

$$
E^*\left[H_T^{*4}\left(\underline{u}_t^*, \underline{u}_s^*\right)\right] = 16 T^{-4} \left[(N(N-1))\right]^{-2} E^*\left[\left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{u_{it}^* u_{jt}^* u_{is}^* u_{js}^*}{\omega_{ij,T}^*} \right)^4 \right]
$$

where

$$
\begin{aligned}
& E^*\left[\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} u_{it}^* u_{jt}^* u_{is}^* u_{js}^*\right)^4\right] \\
= \; & E^*\left[\sum_{i<j}\sum_{l<m}\sum_{p<q}\sum_{h<n} u_{it}^* u_{jt}^* u_{is}^* u_{js}^* u_{lt}^* u_{mt}^* u_{ls}^* u_{ms}^* u_{pt}^* u_{qt}^* u_{ps}^* u_{qs}^* u_{ht}^* u_{nt}^* u_{hs}^* u_{ns}^*\right] \\
= \; & \sum_{i<j} \hat{u}_{it}^4 \hat{u}_{jt}^4 \hat{u}_{is}^4 \hat{u}_{js}^4 + \sum_{i<j}\sum_{l<m} \hat{u}_{it}^2 \hat{u}_{jt}^2 \hat{u}_{is}^2 \hat{u}_{js}^2 \hat{u}_{lt}^2 \hat{u}_{mt}^2 \hat{u}_{ls}^2 \hat{u}_{ms}^2
\end{aligned}
$$

For the second term in (28)

$$
E^*\left[H_T^{*2}\left(\underline{u}_t^*, \underline{u}_s^*\right) H_T^{*2}\left(\underline{u}_t^*, \underline{u}_r^*\right)\right] \leq \left\{E^*\left[H_T^{*4}\left(\underline{u}_t^*, \underline{u}_s^*\right)\right] E^*\left[H_T^{*4}\left(\underline{u}_t^*, \underline{u}_r^*\right)\right]\right\}^{1/2}
$$

where the first line follows by Cauchy-Schwartz inequality whereas the second line by the previous result. The other terms in (28) are zero by the conditional independence and construction of the bootstrap errors. Since for T sufficiently large $\inf_{i,f} \omega_{ij,T}^* \geq K > 0$, letting $C = 48K^{-4}$

$$
\begin{aligned}
\sum_{t=2}^{T} E^*\left[W_{Tt}^{*4}\right] \leq \; & CT \sum_{t=2}^{T}\sum_{s=1}^{t-1} E^*\left[H_T^{*4}\left(\underline{u}_t^*, \underline{u}_s^*\right)\right] \\
= \; & CN^{-2}(N-1)^{-2} \sum_{i<j} T^{-3} \sum_{t=2}^{T}\sum_{s=1}^{t-1} \hat{u}_{it}^4 \hat{u}_{jt}^4 \hat{u}_{is}^4 \hat{u}_{js}^4 \\
& + CN^{-2}(N-1)^{-2} \sum_{i<j}\sum_{l<m} T^{-3} \sum_{t=2}^{T}\sum_{s=1}^{t-1} \hat{u}_{it}^2 \hat{u}_{jt}^2 \hat{u}_{is}^2 \hat{u}_{js}^2 \hat{u}_{lt}^2 \hat{u}_{mt}^2 \hat{u}_{ls}^2 \hat{u}_{ms}^2 \\
= \; & R_{1,NT} + R_{2,NT}
\end{aligned}
$$

where

$$
\begin{aligned}
R_{1,NT} \leq \; & C \frac{1}{2} \frac{1}{N^2 (N-1)^2} \sum_{i<j} \frac{1}{T}\left(\frac{1}{T}\sum_{t=2}^{T} \hat{u}_{it}^4 \hat{u}_{jt}^4\right)^2 \\
= \; & O_p\left(N^{-2}T^{-1}\right)
\end{aligned}
$$

uniformly in i, j, since $\sup_{i,j} \frac{1}{T}\sum_{t=2}^{T} \hat{u}_{it}^4 \hat{u}_{jt}^4 = O_p(1)$ as established previously. Furthermore

$$
\begin{aligned}
R_{2,NT} \leq \; & C \frac{1}{2} \frac{1}{N^2 (N-1)^2} \sum_{i<j}\sum_{l<m} \frac{1}{T}\left(\frac{1}{T}\sum_{t=2}^{T} \hat{u}_{it}^2 \hat{u}_{jt}^2 \hat{u}_{lt}^2 \hat{u}_{mt}^2\right)^2 \\
\leq \; & \frac{1}{2} \frac{1}{N^2 (N-1)^2} \sum_{i<j}\sum_{l<m} \frac{1}{T}\left(\frac{1}{T}\sum_{t=2}^{T} \hat{u}_{it}^4 \hat{u}_{jt}^4\right)\left(\frac{1}{T}\sum_{t=2}^{T} \hat{u}_{lt}^4 \hat{u}_{mt}^4\right) \\
= \; & O_p\left(T^{-1}\right),
\end{aligned}
$$

uniformly in i, j, where the second line follows by the Cauchy-Schwartz inequality. Therefore, condition **(ii)** holds as $(N, T) \to \infty$. ■

Proof of Lemma 6. Secondly, for the wild bootstrap designs WB1 and WB2 in order to establish that

$$
NRBP_{NT}^* = Z_{NT}^* + o_p(1)
$$

as $(N, T) \to \infty, N^2/T \to 0$ for, we show that

(i) $\left|\tilde{Z}_{NT}^* - Z_{NT}^*\right| = O_{p^*}(N/\sqrt{T})$, in probability as $(N, T) \to \infty$, where

$$
\begin{aligned}
\tilde{Z}_{NT}^* = \; & \frac{1}{\sqrt{N(N-1)}} \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \left(\tilde{\gamma}_{ij,T}^{*2} - 1\right) \\
\tilde{\gamma}_{ij,T}^* = \; & \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \hat{u}_{it}^* \hat{u}_{jt}^*}{\sqrt{\omega_{ij,T}^*}}
\end{aligned}
$$

(ii) $\left|NRBP^{*}_{NT}-\tilde{Z}^{*}_{NT}\right|=O_{p^{*}}(N/\sqrt{T})$, as $(N,T)\to\infty$.
For **(i)**, similar to the first-order asymptotic analysis, define

$$\tilde{\gamma}^{*}_{ij}=\gamma^{*}_{ij,T}+\frac{1}{\sqrt{\hat{\omega}_{ij,T}}}\frac{1}{\sqrt{T}}b^{*}_{ij,T}=\gamma^{*}_{ij,T}+c^{*}_{ij,T},\text{ say}$$

where note that $\omega^{*}_{ij,T}=\hat{\omega}_{ij,T}=\frac{1}{T}\sum_{t=1}^{T}\hat{u}^{2}_{it}\hat{u}^{2}_{jt}$ and

$$b^{*}_{ij,T}=u^{*\prime}_{i}H^{*}_{i}H^{*}_{j}u^{*}_{j}-u^{*\prime}_{i}H^{*}_{i}u^{*}_{j}-u^{*\prime}_{i}H^{*}_{j}u^{*}_{j}=\sum_{k=1}^{3}a^{*}_{ijk,T} \tag{29}$$

where $H^{*}_{i}=X^{*}_{i}\left(X^{*\prime}_{i}X^{*}_{i}\right)^{-1}X^{*\prime}_{i}$, X^{*}_{i} has rows x^{*}_{it}, with $u^{*}_{i}=(u^{*}_{i1},...,u^{*}_{iT})'$ and note that Lemma 8 and Assumption A3(i), ensures that $\left(X^{*\prime}_{i}X^{*}_{i}/T\right)^{-1}$ exists for sufficiently large T and is $O_{p^{*}}(1)$, in probability, uniformly in i. Further, since for T sufficiently large $\inf_{i,j}\hat{\omega}_{ij,T}\geq K>0$

$$\begin{aligned}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}c^{*2}_{ij,T} &\leq K^{-1}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{1}{T}b^{*2}_{ij,T}\\ &= K^{-1}\frac{N(N-1)}{T}\left\{\frac{1}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}b^{*2}_{ij,T}\right\}\end{aligned}$$

and we need to show that $\sup_{i,j}\frac{1}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}b^{*2}_{ij,T}=O^{*}_{p}(1)$ and thus

$$\sup_{i,j}\frac{1}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}a^{*2}_{ijk,T}=O_{p^{*}}(1)$$

in probability for $k=1,2,3$ as $(N,T)\to\infty$. For $a^{*}_{ij\,1,T}=u^{*\prime}_{i}H^{*}_{i}H^{*}_{j}u^{*}_{j}$

$$\begin{aligned}\left|a^{*}_{ij1,T}\right|^{2} &\leq \left\|H^{*}_{i}u^{*}_{i}\right\|^{2}\left\|H^{*}_{j}u^{*}_{j}\right\|^{2}\\ &= \left|u^{*\prime}_{i}H^{*}_{i}u^{*}_{i}\right|^{2}\left|u^{*\prime}_{j}H^{*}_{j}u^{*}_{j}\right|^{2}\\ &= O_{p^{*}}(1)\end{aligned}$$

in probability, uniformly in i,j, from Lemma 8 (a)-(b) and Lemma 4 (b). Similar arguments hold for $a^{*}_{ij2,T}$ and $a^{*}_{ij3,T}$.
Now for step **(ii)**,

$$\begin{aligned}\left|RBP^{*}_{NT}-\tilde{Z}^{*}_{NT}\right| &\leq \frac{1}{\sqrt{N(N-1)}}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\left|\hat{\gamma}^{*2}_{ij,T}-\tilde{\gamma}^{*2}_{ij,T}\right|\\ &= \sqrt{\frac{N(N-1)}{T}}\frac{1}{N(N-1)}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\hat{\gamma}^{*2}_{ij,T}\left|\frac{\sqrt{T}\left(\hat{\omega}^{*}_{ij,T}-\omega^{*}_{ij,T}\right)}{\hat{\omega}^{*}_{ij,T}}\right|\\ &= \sqrt{\frac{N(N-1)}{T}}B^{*}_{NT}\end{aligned}$$

and given that $\sup_{i,j}\hat{\gamma}^{*2}_{ij,T}=O^{*}_{p}(1)$, in probability by Lemma 8 (d), we only need to establish that $\sup_{i,j}\left|\frac{\sqrt{T}\left(\hat{\omega}^{*}_{ij,T}-\omega^{*}_{ij,T}\right)}{\hat{\omega}^{*}_{ij,T}}\right|=O^{*}_{p}(1)$, in probability. We have that

$$\frac{\sqrt{T}\left(\hat{\omega}^{*}_{ij,T}-\omega^{*}_{ij,T}\right)}{\hat{\omega}^{*}_{ij,T}}\leq K^{-1}T^{-1/2}\sum_{t=1}^{T}\left(\hat{u}^{*2}_{it}\hat{u}^{*2}_{jt}-\hat{u}^{2}_{it}\hat{u}^{2}_{jt}\right)$$

since T sufficiently large $\inf_{i,j}\hat{\omega}_{ij,T}\geq K>0$ a.s. and where $\omega^{*}_{ij,T}=E^{*}\left(\frac{1}{T}\sum_{t=1}^{T}u^{*2}_{it}u^{*2}_{jt}\right)=\frac{1}{T}\sum_{t=1}^{T}\hat{u}^{2}_{it}\hat{u}^{2}_{jt}$ given that for Rademacher distribution $\varepsilon^{2}_{it}=\varepsilon^{2}_{jt}=1$. Further, $T^{-1/2}\sum_{t=1}^{T}\left(\hat{u}^{*2}_{it}\hat{u}^{*2}_{jt}-\hat{u}^{2}_{it}\hat{u}^{2}_{jt}\right)=O^{*}_{p}\left(T^{-1/2}\right)$, in probability, uniformly in i,j, by Lemma 8 (c) and triangle inequality which yields that $\sup_{i,j}\left|\frac{\sqrt{T}\left(\hat{\omega}^{*}_{ij,T}-\omega^{*}_{ij,T}\right)}{\hat{\omega}^{*}_{ij,T}}\right|=O^{*}_{p}(1)$, in probability. This establishes that $B^{*}_{NT}=O^{*}_{p}(1)$, in probability. ■

Table 1: Rejection frequencies of the asymptotic and various Wild Bootstrap RBP and BP tests in panel ADL(1,0) models under homoskedastic errors (HET0).

	$H_0 : E[u_{it}u_{jt}] = 0$										$H_A : E[u_{it}u_{jt}] = 0.2$									
	SN					χ^2_6					SN					χ^2_6				
N	5	10	25	50	100	5	10	25	50	100	5	10	25	50	100	5	10	25	50	100
	Asymptotic critical values										Asymptotic critical values									
T	RBP_T										RBP_T									
25	3.3	5.6	12.2	27.2	76.7	3.5	4.5	9.7	23.0	69.0	6.2	11.1	30.5	69.3	98.1	4.4	8.4	26.9	58.9	92.2
50	4.4	5.6	7.4	13.4	30.3	3.9	4.4	7.0	12.0	32.5	9.1	18.0	54.8	88.8	99.6	8.1	15.6	43.7	76.1	97.0
100	4.9	5.6	6.4	8.4	12.6	4.3	5.3	7.0	9.8	19.3	17.9	40.3	87.6	99.7	100.0	14.1	33.7	79.6	98.5	99.9
200	4.6	5.1	4.5	6.4	8.2	5.9	4.2	7.0	8.1	12.7	35.9	76.0	99.8	100.0	100.0	33.8	70.2	99.1	100.0	100.0
	BP_T										BP_T									
25	5.0	8.5	14.8	36.6	86.8	5.7	7.1	15.6	38.5	86.3	7.9	14.4	39.0	78.2	99.3	8.0	15.6	41.6	77.1	98.8
50	4.8	5.9	8.4	16.3	36.1	5.7	6.3	8.5	15.3	36.3	9.9	21.2	60.3	91.0	99.9	12.1	23.9	59.0	89.3	99.6
100	5.0	6.0	6.1	8.6	14.1	4.8	6.0	7.3	8.7	15.0	19.3	42.4	89.4	99.8	100.0	20.3	43.3	88.0	99.8	100.0
200	5.0	5.3	4.7	6.7	8.4	5.5	4.7	6.7	7.2	8.8	37.3	77.2	99.9	100.0	100.0	37.4	76.4	99.7	100.0	100.0
T	$NRBP_{NT}$										$NRBP_{NT}$									
25	5.0	6.8	13.4	27.9	77.0	4.8	5.9	10.5	23.7	69.4	8.0	12.8	31.7	69.7	98.2	6.7	10.6	28.2	59.5	92.3
50	6.1	6.6	8.3	13.9	30.9	5.3	5.3	7.7	12.7	32.9	11.2	20.3	56.6	89.1	99.6	10.4	17.1	44.9	76.7	97.1
100	6.3	6.5	6.7	8.6	12.9	6.4	6.5	7.8	10.2	19.6	21.0	43.4	88.2	99.8	100.0	18.1	36.4	80.2	98.5	99.9
200	6.4	6.1	5.0	6.8	8.6	7.7	5.1	7.2	8.7	13.0	41.2	78.3	99.8	100.0	100.0	37.7	72.8	99.1	100.0	100.0
	NBP_{NT}										NBP_{NT}									
25	6.7	9.4	16.3	37.4	87.1	7.4	8.6	16.6	39.5	86.5	10.1	16.2	39.9	78.5	99.3	10.1	17.3	43.1	77.5	98.8
50	6.4	7.3	9.0	16.5	36.5	6.7	7.5	9.0	16.0	36.8	12.5	24.0	61.2	91.0	99.9	14.5	26.0	60.2	89.5	99.6
100	6.7	7.0	6.7	8.8	14.2	6.8	6.9	8.0	9.0	15.3	22.5	44.9	90.1	99.8	100.0	24.6	45.7	88.6	99.8	100.0
200	7.0	6.6	5.1	7.0	8.4	7.4	5.6	7.0	7.5	9.2	42.0	79.7	100.0	100.0	100.0	41.0	78.6	99.7	100.0	100.0
	WB 1: Recursive resampling										WB 1: Recursive resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	5.0	5.1	5.6	5.3	6.7	4.5	5.1	4.2	5.2	5.1	5.5	9.9	21.1	40.6	70.6	6.3	9.4	17.6	33.8	54.4
50	4.4	4.8	4.8	5.9	4.7	5.6	4.9	4.8	6.1	6.8	9.4	18.1	46.4	81.5	98.0	8.4	16.0	40.6	67.8	91.1
100	4.9	4.5	5.2	5.2	5.8	4.6	5.1	6.6	6.4	8.0	15.8	38.7	86.4	99.5	100.0	17.2	32.5	78.4	96.7	99.9
200	5.9	5.0	6.0	4.9	5.5	4.1	5.0	5.2	6.5	8.3	35.6	72.9	99.9	100.0	100.0	32.5	68.8	98.7	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.4	5.0	4.8	4.9	5.7	4.4	4.8	4.3	5.7	5.9	6.4	9.4	22.0	40.8	70.7	6.7	10.6	23.7	44.8	69.8
50	4.8	4.8	4.8	5.5	4.6	5.5	4.5	4.3	4.7	5.5	9.8	19.3	47.2	82.3	98.2	10.8	19.4	50.0	80.7	96.8
100	5.1	4.8	4.5	5.4	5.8	4.9	5.1	5.7	4.5	5.2	16.4	39.2	86.9	99.6	100.0	18.8	39.2	85.6	99.3	100.0
200	5.8	4.9	5.7	4.5	5.6	4.3	4.4	4.6	4.7	5.1	35.5	73.2	99.9	100.0	100.0	34.9	73.7	99.4	100.0	100.0
	WB 2: Fixed-design resampling										WB 2: Fixed-design resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	5.0	5.3	6.1	6.5	10.6	4.7	5.3	4.9	6.7	7.9	5.9	9.8	21.8	43.8	74.6	6.6	9.7	18.9	36.0	58.9
50	4.9	5.0	4.9	6.3	5.8	5.5	4.7	5.1	6.6	7.6	9.0	18.0	46.8	81.8	98.0	9.0	16.1	40.6	68.7	91.7
100	5.0	4.3	5.0	5.5	6.0	4.5	5.3	6.5	6.3	9.2	15.6	38.0	86.8	99.5	100.0	17.1	32.5	78.3	96.8	99.9
200	6.0	5.2	5.8	4.7	5.8	4.2	4.9	5.2	6.4	8.5	35.6	73.3	99.9	100.0	100.0	32.7	69.0	98.7	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.8	5.1	5.6	6.6	9.6	4.4	5.0	5.1	7.3	9.0	6.7	10.0	22.8	44.9	74.6	7.1	11.3	25.3	48.1	73.6
50	4.9	4.9	5.0	6.4	5.1	5.7	4.4	4.8	5.4	6.6	9.4	19.2	47.7	82.9	98.4	11.1	19.9	50.6	81.6	97.1
100	5.0	4.8	4.9	5.4	5.8	4.7	5.5	5.9	4.4	5.4	16.1	38.2	86.9	99.4	100.0	18.5	39.1	85.6	99.2	100.0
200	6.0	5.0	5.7	4.8	5.6	4.3	4.4	4.4	4.5	5.2	36.1	72.9	99.9	100.0	100.0	35.3	74.5	99.4	100.0	100.0
	WB 3: Direct resampling										WB 3: Direct resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	4.9	5.7	6.9	9.2	16.4	4.5	5.6	5.3	8.5	12.2	5.8	10.6	24.0	47.5	80.3	6.7	10.1	20.8	40.1	64.6
50	4.5	5.0	5.2	7.2	8.0	5.7	4.8	5.3	7.2	10.9	9.1	18.5	48.6	83.4	98.4	8.8	16.9	42.0	70.4	92.7
100	4.9	5.0	5.5	5.9	7.0	4.4	5.3	7.1	6.9	10.1	15.7	38.3	86.9	99.5	100.0	17.1	33.8	79.0	97.0	99.9
200	6.1	5.1	6.2	5.1	6.4	4.3	5.1	5.2	6.4	8.9	36.0	73.3	99.9	100.0	100.0	32.4	69.2	98.7	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.7	5.5	6.7	9.4	17.3	4.6	5.8	6.3	9.8	16.9	6.7	9.8	25.5	50.7	82.1	7.2	11.8	27.5	53.4	81.7
50	4.9	4.9	5.9	7.3	7.9	5.3	5.1	5.6	6.8	9.2	9.7	19.6	49.5	84.2	98.7	11.2	20.2	51.4	83.1	97.9
100	5.2	4.7	5.1	5.9	7.1	4.6	5.4	5.9	4.9	6.7	16.3	39.0	87.5	99.6	100.0	19.0	39.3	85.9	99.3	100.0
200	5.9	4.9	5.8	4.9	6.0	4.7	4.6	4.9	4.9	5.7	36.3	72.9	99.9	100.0	100.0	34.9	74.3	99.4	100.0	100.0

Notes: The data generating process considered is $y_{it} = \theta_{i1} + \theta_{i2}z_{it} + \phi_i y_{i,t-1} + u_{it}$, $i = 1, 2, ..., N$ and $t = -49, -48, ..., T$.with $\theta_{i1} \sim$ i.i.d. $N(0, 1)$, $\theta_{i2} = 1 - \phi_i$, $\phi_i \sim$ i.i.d. Uniform$[0.4, 0.6]$, and the z_{it} are generated for $(N = 5, T = 25)$ as independent random draws from the standard lognormal distribution. This block of regressor values is then reused as necessary to build up data for the other combinations (N, T). $y_{i,-49} = 0$, and first 49 values are discarded. The error term is written as $u_{it} = \sigma_{it}\varepsilon_{it}$, $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$. There is homoskedasticity under scheme HET0, with $\sigma_{it} = 1$ for all t. The term ε_{it} is generated as $\varepsilon_{it} = \sqrt{1-\rho^2}\xi_{it} + \rho\zeta_t$where $\xi_{it} \sim$ i.i.d.

Table 2: Rejection frequencies of the asymptotic and various Wild-Bootstrap RBP and BP tests in panel ADL(1,0) models under one-break-in-volatility heteroskedastic scheme (HET1).

	$H_0: E[u_{it}u_{jt}]=0$										$H_A: E[u_{it}u_{jt}]=0.2$									
	SN					χ^2_6					SN					χ^2_6				
N	5	10	25	50	100	5	10	25	50	100	5	10	25	50	100	5	10	25	50	100
	Asymptotic critical values										Asymptotic critical values									
T	RBP_T										RBP_T									
25	3.5	5.4	12.6	29.3	82.8	3.3	4.6	9.9	25.2	72.9	5.8	9.8	29.5	69.3	98.3	4.0	8.2	26.7	59.0	93.9
50	4.5	5.2	8.2	14.0	34.3	3.6	4.5	6.3	12.7	32.9	8.2	15.9	50.0	85.9	99.3	6.9	14.4	40.5	74.2	96.4
100	5.3	5.5	5.7	8.4	15.2	4.4	5.0	6.8	8.7	18.0	15.9	35.9	82.6	99.4	100.0	13.1	29.8	75.8	96.9	99.9
200	4.7	5.5	4.9	6.6	10.1	5.2	4.7	6.4	8.0	12.8	31.9	68.7	99.5	100.0	100.0	29.5	63.2	98.3	100.0	100.0
	BP_T										BP_T									
25	8.1	18.1	57.0	96.2	100.0	8.7	16.6	53.7	95.6	100.0	11.5	26.1	75.9	99.4	100.0	11.7	25.1	74.1	99.2	100.0
50	9.5	18.6	53.6	96.8	100.0	9.4	17.8	50.9	94.1	100.0	15.5	37.1	87.9	99.9	100.0	16.5	37.6	86.8	99.9	100.0
100	10.2	18.7	51.1	96.0	100.0	10.2	16.9	52.4	95.3	100.0	25.7	60.4	98.8	100.0	100.0	25.6	57.8	98.1	100.0	100.0
200	9.5	17.5	54.0	95.3	100.0	9.5	15.9	53.6	96.2	100.0	44.4	85.9	100.0	100.0	100.0	42.6	84.5	100.0	100.0	100.0
T	$NRBP_{NT}$										$NRBP_{NT}$									
25	5.5	7.0	13.2	30.1	83.4	4.7	5.2	10.7	26.0	73.1	8.1	11.6	31.3	70.3	98.3	5.9	10.2	27.8	60.1	93.9
50	6.3	6.4	8.5	14.6	34.8	5.2	6.0	7.0	13.3	33.4	10.6	18.4	51.1	86.3	99.3	8.7	16.8	41.9	74.8	96.4
100	7.1	6.5	6.3	8.7	15.4	6.1	5.8	7.5	9.0	18.4	19.5	38.3	83.8	99.4	100.0	16.5	32.6	76.8	97.1	99.9
200	6.0	6.7	5.5	7.0	10.3	6.3	5.8	6.9	8.3	13.0	35.9	71.6	99.5	100.0	100.0	34.0	65.9	98.4	100.0	100.0
	NBP_{NT}										NBP_{NT}									
25	11.2	20.7	58.8	96.4	100.0	11.1	19.0	55.4	95.6	100.0	14.1	28.8	77.3	99.4	100.0	14.3	27.3	75.4	99.3	100.0
50	12.2	20.4	55.3	96.9	100.0	11.2	20.2	52.4	94.5	100.0	18.5	40.4	88.3	99.9	100.0	20.2	40.2	87.6	100.0	100.0
100	12.5	21.7	53.3	96.1	100.0	12.9	19.3	54.3	95.5	100.0	29.5	62.5	98.8	100.0	100.0	29.4	60.7	98.1	100.0	100.0
200	12.3	19.9	55.7	95.5	100.0	12.1	17.6	55.1	96.4	100.0	48.7	87.3	100.0	100.0	100.0	47.8	85.8	100.0	100.0	100.0
	WB 1: Recursive resampling										WB 1: Recursive resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	4.8	5.2	5.7	5.2	6.9	3.9	4.5	4.1	4.0	4.5	5.7	8.3	18.1	36.0	66.3	6.0	8.8	17.6	31.2	51.2
50	4.9	4.5	4.4	5.8	5.2	4.9	4.8	4.7	4.8	5.6	8.2	16.7	41.3	76.3	97.0	8.7	14.6	36.5	64.5	88.3
100	4.9	4.8	4.5	5.2	7.0	4.6	5.0	6.1	5.6	6.8	15.0	33.1	81.2	98.8	100.0	15.0	29.6	72.2	94.8	99.8
200	4.5	4.7	5.9	4.9	5.7	4.9	5.2	5.3	5.9	7.6	31.6	65.9	99.4	100.0	100.0	28.8	62.5	97.6	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.5	5.7	6.6	9.6	20.1	4.2	5.6	6.6	8.8	18.3	6.7	9.5	22.1	48.0	79.3	7.0	12.1	26.5	50.6	79.5
50	4.6	4.7	6.1	9.5	14.1	5.9	5.1	6.0	8.2	13.9	8.6	17.9	46.9	81.9	98.3	10.8	18.5	48.9	81.4	98.0
100	5.2	5.3	6.0	7.5	11.7	4.6	5.8	6.8	6.6	10.6	16.0	34.2	83.5	99.4	100.0	18.1	35.3	81.9	98.5	100.0
200	5.1	5.2	6.2	5.7	7.7	4.8	4.8	5.6	5.2	7.3	32.4	66.5	99.5	100.0	100.0	31.0	68.1	98.7	100.0	100.0
	WB 2: Fixed-design resampling										WB 2: Fixed-design resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	5.5	5.5	6.0	6.3	10.7	4.5	5.4	4.8	5.3	7.0	6.1	9.1	19.6	39.0	71.2	6.6	9.4	18.8	33.9	56.1
50	4.6	4.7	5.1	6.3	6.1	5.1	4.6	4.8	5.0	6.4	8.4	16.1	42.4	76.8	97.3	8.9	15.0	37.2	65.5	89.1
100	4.7	4.7	4.9	5.4	6.7	4.7	4.7	6.4	5.6	7.2	14.6	33.0	81.3	98.9	100.0	15.1	30.0	72.2	94.8	99.8
200	4.8	5.0	5.9	5.1	5.7	4.8	5.6	5.8	6.0	7.7	31.7	65.6	99.5	100.0	100.0	28.4	63.0	97.7	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.9	5.9	8.4	13.1	31.8	4.4	6.1	8.3	11.8	29.3	6.5	10.0	24.1	53.7	84.6	7.3	12.7	28.5	55.0	84.8
50	4.6	4.9	6.5	10.7	16.3	5.6	5.2	6.6	9.0	15.4	8.5	18.2	47.8	82.5	98.6	10.8	19.1	49.4	82.1	98.2
100	4.9	5.4	6.3	7.9	12.0	4.9	5.9	6.9	6.3	11.4	15.4	33.6	83.4	99.4	100.0	18.3	35.4	81.8	98.8	100.0
200	5.1	5.0	6.4	5.7	7.7	4.8	5.1	5.7	5.6	7.3	32.4	66.6	99.6	100.0	100.0	30.8	68.7	98.7	100.0	100.0
	WB 3: Direct resampling										WB 3: Direct resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	5.2	5.7	7.2	8.3	17.3	4.1	5.4	5.6	7.2	10.8	5.9	9.2	21.0	43.8	77.0	6.6	9.7	20.3	38.4	62.9
50	4.7	5.0	5.5	7.3	8.5	4.9	4.8	5.2	6.3	9.4	8.7	16.2	42.7	79.0	97.6	9.6	15.5	37.9	66.8	91.0
100	4.8	4.6	4.6	5.9	8.4	4.8	5.0	6.7	6.5	8.5	15.2	33.2	82.0	99.0	100.0	15.5	30.3	73.0	95.4	99.9
200	4.7	4.7	5.9	5.2	6.3	4.7	5.1	5.6	6.1	8.9	31.7	65.6	99.5	100.0	100.0	28.9	62.5	97.5	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.5	5.9	8.3	13.3	30.9	4.2	6.6	8.6	12.3	29.8	7.2	10.2	24.0	52.5	85.2	7.4	12.6	28.8	55.6	85.2
50	4.4	5.0	6.0	9.1	13.0	5.8	5.1	6.1	7.9	13.4	8.4	17.6	46.5	81.1	98.4	10.6	18.6	48.1	81.2	97.9
100	5.2	5.1	5.6	7.0	10.0	4.9	5.6	6.3	5.8	9.1	14.9	33.9	82.8	99.1	100.0	17.9	35.2	81.2	98.4	100.0
200	5.3	4.9	6.3	5.1	6.9	4.9	4.5	5.1	5.3	6.4	31.9	66.4	99.5	100.0	100.0	30.6	68.3	98.7	100.0	100.0

Notes: The data generating process is identical to those used for Table 1 except that $\sigma_{it} = 0.8$ for $t = 1, 2, ..., m = \lfloor T/2 \rfloor$ and $\sigma_{it} = 1.2$ for $t = m, m+1, ..., T$, where $\lfloor A \rfloor$ is the largest integer part of A.

Table 3: Rejection frequencies of the asymptotic and various wild-bootstrap RBP and BP tests in panel ADL(1,0) models under trending volatility heteroskedastic scheme (HET2).

	$H_0 : E[u_{it}u_{jt}] = 0$										$H_A : E[u_{it}u_{jt}] = 0.2$									
	SN					χ^2_6					SN					χ^2_6				
N	5	10	25	50	100	5	10	25	50	100	5	10	25	50	100	5	10	25	50	100
Asymptotic critical values											Asymptotic critical values									
T	RBP_T										RBP_T									
25	3.4	5.2	12.1	27.0	78.6	3.5	4.1	9.6	24.7	70.7	5.6	10.1	30.1	69.8	98.1	4.3	8.1	26.8	59.2	92.7
50	4.3	5.3	8.0	13.6	32.1	3.7	4.6	6.8	12.7	33.2	8.5	17.3	52.7	87.8	99.4	7.5	15.0	43.0	75.5	96.6
100	5.1	5.1	6.2	9.0	13.1	4.3	5.0	7.1	9.2	18.0	17.0	38.2	85.5	99.6	100.0	14.1	32.2	78.7	97.9	100.0
200	4.4	5.6	5.1	6.8	9.0	5.3	4.5	6.1	8.0	13.2	35.0	73.7	99.7	100.0	100.0	31.4	68.0	98.7	100.0	100.0
	BP_T										BP_T									
25	6.4	10.8	28.1	68.5	99.5	6.0	9.5	27.6	68.2	99.3	9.4	17.6	52.3	92.4	100.0	9.3	18.5	53.9	90.7	100.0
50	6.4	10.1	21.2	52.4	95.3	6.7	10.5	20.3	48.9	94.2	12.1	26.7	71.9	98.1	100.0	13.7	28.7	70.9	96.9	100.0
100	6.8	9.1	17.0	43.2	89.7	6.5	8.4	19.3	41.6	88.5	21.4	48.0	93.9	100.0	100.0	22.8	48.3	93.0	99.9	100.0
200	6.6	8.7	16.7	39.5	85.8	7.0	8.2	17.2	39.3	83.6	39.2	81.2	99.9	100.0	100.0	38.6	79.4	100.0	100.0	100.0
T	$NRBP_{NT}$										$NRBP_{NT}$									
25	5.0	6.4	13.1	28.0	79.1	4.6	5.1	10.7	25.1	71.1	8.1	12.0	31.7	70.5	98.1	6.1	10.0	28.0	60.1	92.8
50	6.1	6.6	8.3	14.1	32.6	5.1	5.9	7.3	13.4	33.7	11.0	20.4	53.7	88.1	99.4	9.7	17.6	43.8	76.1	96.7
100	7.1	6.1	6.6	9.3	13.3	5.7	5.9	7.8	9.5	18.5	20.7	42.3	86.6	99.6	100.0	17.1	34.4	79.4	97.9	100.0
200	6.5	6.5	5.6	7.0	9.2	7.5	5.4	6.8	8.3	13.3	39.5	75.9	99.7	100.0	100.0	35.9	70.5	98.7	100.0	100.0
	NBP_{NT}										NBP_{NT}									
25	8.2	12.4	29.5	69.1	99.5	8.0	10.9	28.7	68.9	99.5	11.4	20.0	53.8	92.7	100.0	12.1	20.4	55.0	91.1	100.0
50	8.4	11.7	22.1	53.0	95.5	8.6	11.7	21.2	49.7	94.2	14.6	28.9	73.0	98.1	100.0	16.7	31.9	72.3	96.9	100.0
100	8.6	10.9	18.0	44.4	89.7	8.9	9.8	20.5	42.5	88.7	25.4	51.4	94.4	100.0	100.0	26.3	51.0	93.3	99.9	100.0
200	8.3	10.1	17.7	40.1	86.1	9.3	9.4	18.8	40.5	83.8	45.7	82.8	99.9	100.0	100.0	42.9	81.3	100.0	100.0	100.0
WB 1: Recursive resampling											WB 1: Recursive resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	5.3	5.7	5.3	5.7	7.0	4.9	5.1	5.3	4.9	5.1	7.1	9.4	21.6	40.0	67.7	6.6	9.4	17.3	32.1	53.7
50	3.6	4.5	5.0	5.7	6.2	4.2	4.8	4.9	5.5	6.1	8.3	16.9	45.7	79.9	96.8	9.6	15.3	37.7	68.3	88.8
100	5.0	5.6	6.0	4.1	5.3	5.1	4.4	6.5	5.9	7.5	16.8	37.5	85.6	99.5	100.0	16.1	31.9	77.5	96.6	99.9
200	4.4	5.0	4.7	5.3	4.8	4.5	6.5	5.1	5.7	7.4	32.2	71.9	99.9	100.0	100.0	30.4	67.4	98.3	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	5.1	5.2	6.5	7.2	11.1	4.7	4.8	5.5	6.7	11.2	6.3	9.9	23.5	46.0	76.6	6.6	10.6	24.3	45.7	75.7
50	4.0	4.7	5.1	6.5	9.8	4.5	4.5	5.2	5.4	8.1	8.8	17.3	48.1	82.6	97.9	11.3	19.1	49.6	81.1	97.6
100	4.5	5.9	6.1	4.9	7.3	5.3	4.9	6.6	5.0	6.2	16.4	38.7	86.5	99.6	100.0	17.3	37.8	86.5	99.0	100.0
200	4.8	5.2	4.6	6.0	6.1	4.0	5.8	5.2	4.5	5.5	32.5	71.7	99.9	100.0	100.0	33.1	72.6	99.1	100.0	100.0
WB 2: Fixed-design resampling											WB 2: Fixed-design resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	5.0	5.3	6.1	6.1	10.3	4.4	4.7	4.6	6.1	7.4	6.1	9.2	21.3	41.1	72.4	6.6	9.4	18.5	34.9	56.9
50	4.7	5.0	4.7	6.2	5.3	5.2	4.9	4.9	6.0	6.9	8.9	18.1	45.5	80.0	97.6	9.2	15.8	39.6	68.0	90.5
100	5.2	4.8	4.6	5.5	6.3	4.9	5.2	6.9	5.9	7.8	15.9	35.3	85.4	99.3	100.0	16.1	31.8	77.0	96.0	99.8
200	5.2	4.7	6.1	4.8	6.1	4.5	5.1	5.4	6.0	8.0	34.0	70.4	99.8	100.0	100.0	30.8	66.3	98.1	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.8	5.4	6.4	8.3	15.0	4.6	5.0	6.0	8.2	14.7	6.6	9.9	23.4	46.0	78.3	7.4	11.1	26.1	50.3	77.8
50	4.3	5.0	5.3	7.5	8.1	5.4	4.5	5.3	5.9	8.8	9.4	18.3	48.7	82.5	98.5	11.4	19.2	50.3	81.4	97.5
100	5.1	5.4	4.8	5.8	8.3	4.4	5.5	6.6	5.2	7.6	16.0	36.8	86.2	99.4	100.0	17.9	38.0	84.0	98.9	100.0
200	5.9	4.7	6.4	5.1	6.5	4.4	4.9	5.1	5.2	6.2	34.2	71.3	99.8	100.0	100.0	33.9	72.0	99.1	100.0	100.0
WB 3: Direct resampling											WB 3: Direct resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	5.3	5.5	7.0	8.6	16.5	4.5	4.7	5.2	8.1	11.6	6.1	9.6	22.7	45.9	79.1	6.7	10.0	20.3	38.3	63.5
50	4.8	5.2	5.3	6.8	8.3	4.9	4.8	5.1	7.2	9.8	9.2	18.0	47.0	81.5	98.1	9.1	16.3	40.5	69.7	92.0
100	5.3	4.8	4.9	5.9	7.2	4.7	5.7	6.7	6.8	9.6	15.9	35.9	85.5	99.4	100.0	16.1	32.1	77.5	96.5	99.9
200	5.7	4.5	6.6	5.0	6.3	4.5	5.1	5.6	6.3	8.7	34.4	70.5	99.9	100.0	100.0	30.4	66.5	98.3	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.5	5.7	7.1	9.7	21.9	4.4	6.2	7.0	9.8	21.3	6.7	10.8	24.6	50.8	83.1	7.4	11.8	28.2	54.5	82.5
50	4.4	5.1	5.7	8.3	9.8	5.6	5.1	5.7	7.2	10.4	9.1	18.8	49.0	83.1	98.5	11.0	19.2	50.6	82.3	97.8
100	4.7	5.4	5.2	5.7	8.3	4.8	5.3	6.6	5.4	7.6	16.0	36.8	85.9	99.5	100.0	17.9	37.3	84.0	99.1	100.0
200	5.9	4.6	6.6	4.9	6.6	4.8	4.9	4.9	5.2	6.1	34.4	71.7	99.8	100.0	100.0	33.9	71.7	99.1	100.0	100.0

Notes: The data generating process is identical to those used for Table 1 except that $\sigma_{it} = \sigma_0 - (\sigma_1 - \sigma_0)\left(\frac{t-1}{T-1}\right)$ with $\sigma_0 = 0.8$ and $\sigma_1 = 1.2$.

Table 4: Rejection frequencies of the asymptotic and various Wild Bootstrap RBP and BP tests in panel ADL(1,0) models under conditional heteroskedasticity depending on a regressor (HET3).

	$H_0 : E[u_{it}u_{jt}] = 0$										$H_A : E[u_{it}u_{jt}] = 0.2$									
	SN					χ^2_6					SN					χ^2_6				
N	5	10	25	50	100	5	10	25	50	100	5	10	25	50	100	5	10	25	50	100
	Asymptotic critical values										Asymptotic critical values									
T	RBP_T										RBP_T									
25	4.0	5.7	13.4	35.9	88.1	3.1	5.2	11.1	29.9	80.1	5.5	10.5	33.7	75.1	99.1	5.1	9.4	28.9	63.7	95.8
50	4.1	4.9	9.2	20.8	53.5	4.1	4.8	7.9	18.8	51.6	8.3	18.5	53.7	90.4	99.9	7.3	15.5	47.2	79.6	98.2
100	3.9	4.8	7.4	12.8	23.3	3.6	5.2	7.0	9.7	25.0	14.9	35.1	86.6	99.7	100.0	14.3	30.1	78.3	97.5	100.0
200	5.0	5.0	7.2	7.1	13.7	4.2	4.8	5.6	8.7	15.5	33.3	68.9	99.6	100.0	100.0	29.3	63.8	98.7	100.0	100.0
	BP_T										BP_T									
25	5.4	8.3	20.0	51.4	96.6	5.1	8.2	21.0	51.5	95.7	7.4	13.5	45.0	85.7	99.8	8.6	16.1	46.9	84.5	99.6
50	4.5	7.1	17.9	43.2	87.9	5.1	7.4	15.8	41.0	86.3	9.5	23.5	65.9	96.1	100.0	12.4	24.4	66.0	95.1	100.0
100	4.9	7.3	16.9	42.0	85.9	4.5	7.9	16.1	36.4	82.4	15.4	40.6	92.0	100.0	100.0	18.0	40.8	89.9	99.9	100.0
200	5.3	7.1	16.1	38.7	86.6	4.5	6.8	14.8	36.4	84.8	33.1	72.2	100.0	100.0	100.0	33.7	72.2	99.5	100.0	100.0
T	$NRBP_{NT}$										$NRBP_{NT}$									
25	6.1	6.8	14.4	37.3	88.2	5.1	6.4	12.0	30.8	80.3	7.4	12.1	35.1	75.9	99.1	7.0	11.2	30.3	64.2	95.8
50	5.2	6.6	10.2	21.6	53.7	5.6	5.5	8.8	19.4	51.9	10.9	20.6	55.4	90.6	99.9	9.6	17.3	48.4	80.0	98.2
100	5.9	6.0	8.0	13.2	23.6	5.4	6.4	7.3	10.2	25.4	17.7	38.2	87.2	99.7	100.0	18.0	32.9	79.0	97.6	100.0
200	6.8	6.3	7.9	7.4	13.8	5.9	5.7	5.9	8.9	15.7	38.3	71.3	99.6	100.0	100.0	33.4	67.0	98.8	100.0	100.0
	NBP_{NT}										NBP_{NT}									
25	7.4	9.5	21.6	52.4	96.7	7.2	9.2	22.2	52.0	95.8	10.3	15.7	46.6	86.2	99.8	11.0	18.3	48.4	84.6	99.6
50	5.9	8.2	19.2	43.7	88.1	6.9	8.7	16.9	41.9	86.5	12.2	25.7	67.0	96.4	100.0	14.9	26.5	67.1	95.3	100.0
100	6.3	8.3	17.9	42.8	86.2	6.4	9.1	16.9	37.1	82.7	18.4	43.5	92.4	100.0	100.0	21.6	43.8	90.5	99.9	100.0
200	7.5	8.2	17.1	39.7	86.8	5.9	8.4	16.1	36.8	85.1	37.9	74.6	100.0	100.0	100.0	38.4	74.9	99.6	100.0	100.0
	WB 1: Recursive resampling										WB 1: Recursive resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	5.2	5.3	5.8	6.4	10.4	4.2	5.2	4.1	6.3	7.3	6.6	9.6	21.4	42.0	72.3	6.3	8.7	18.4	34.4	57.5
50	4.3	5.1	5.3	7.8	9.7	4.8	5.1	5.2	6.7	10.0	8.9	17.5	46.0	81.0	98.1	8.5	16.1	41.4	68.8	91.2
100	4.4	4.3	5.8	6.1	6.3	4.8	5.6	5.9	5.6	8.7	15.3	34.8	83.5	99.4	100.0	15.8	30.0	74.7	95.7	99.9
200	5.7	4.7	6.1	4.5	6.1	4.4	5.0	5.2	6.2	7.7	33.1	69.0	99.6	100.0	100.0	30.6	63.8	98.2	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.5	5.0	6.0	8.1	12.1	4.9	5.0	5.7	8.2	12.1	6.9	9.0	22.7	44.9	77.2	6.8	10.1	25.2	47.3	74.6
50	3.9	5.6	8.1	13.8	26.8	5.0	5.1	6.3	12.6	25.2	8.7	18.6	50.5	86.7	99.1	10.3	18.9	52.7	84.6	98.8
100	4.2	5.4	7.7	11.0	19.8	4.7	5.8	6.5	8.3	18.3	14.8	35.2	85.5	99.6	100.0	17.3	35.7	85.1	99.2	100.0
200	5.6	5.3	7.5	6.9	13.2	4.5	5.3	5.2	6.9	10.6	33.6	67.7	99.6	100.0	100.0	33.0	68.5	99.1	100.0	100.0
	WB 2: Fixed-design resampling										WB 2: Fixed-design resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	5.5	5.3	6.7	8.1	14.4	4.4	5.6	5.1	8.0	11.0	6.3	9.8	22.7	44.6	76.8	6.5	9.6	20.4	37.7	61.3
50	4.1	4.8	5.3	8.3	10.6	5.1	4.8	5.5	7.4	11.8	9.0	18.0	46.3	81.4	98.3	8.3	15.4	41.6	69.6	91.7
100	4.3	4.7	5.7	6.4	6.7	4.1	5.6	5.8	5.5	8.6	14.9	34.7	83.7	99.4	100.0	15.6	30.8	75.0	95.8	99.9
200	5.7	4.9	6.3	4.9	6.4	4.5	5.1	5.1	6.3	7.9	33.8	68.7	99.5	100.0	100.0	30.5	64.0	98.3	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.4	5.2	6.2	9.8	17.5	5.0	5.2	6.6	9.9	19.0	7.1	9.5	23.7	48.3	80.4	7.0	10.6	26.8	50.0	79.0
50	4.3	5.7	8.0	14.9	29.6	4.9	4.7	6.8	12.7	27.7	8.6	19.3	51.4	86.7	99.3	10.4	19.3	53.1	85.4	98.9
100	4.1	5.5	8.1	11.4	20.4	4.2	5.5	6.8	8.5	19.3	14.4	34.9	85.2	99.6	100.0	17.5	35.3	85.2	99.2	100.0
200	5.8	5.1	7.8	7.0	13.4	4.8	5.2	5.3	7.5	10.5	33.0	68.0	99.7	100.0	100.0	32.5	68.6	99.2	100.0	100.0
	WB 3: Direct resampling										WB 3: Direct resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	5.2	5.4	7.5	11.7	26.8	4.7	5.6	6.4	11.7	21.4	6.6	10.2	25.6	50.6	85.3	6.5	9.6	22.5	43.3	71.5
50	4.5	5.5	7.4	11.5	21.1	4.7	5.0	7.0	10.9	22.2	9.2	18.6	48.8	85.1	99.0	8.6	16.3	43.7	73.7	95.0
100	4.1	4.5	6.7	8.3	11.8	4.7	5.2	6.7	7.9	15.0	15.3	35.3	85.1	99.5	100.0	15.5	31.8	76.9	96.9	100.0
200	5.5	5.0	6.8	6.0	9.6	4.5	5.2	5.7	7.9	11.2	33.6	68.8	99.6	100.0	100.0	30.8	64.5	98.3	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.8	5.6	7.9	13.7	31.4	5.1	6.0	8.5	14.4	32.9	7.2	10.3	26.9	55.1	88.0	7.1	11.3	29.4	57.3	86.8
50	4.3	5.7	8.8	16.2	34.9	5.3	4.9	7.5	15.4	34.3	8.7	19.5	51.8	89.1	99.5	10.8	20.1	54.5	87.3	99.3
100	4.5	5.2	7.6	11.6	20.9	4.9	5.9	7.1	9.0	20.6	15.5	35.7	85.4	99.6	100.0	17.7	35.9	85.3	99.2	100.0
200	5.8	5.1	8.0	7.0	13.2	5.1	5.3	5.0	7.4	10.8	33.5	68.1	99.6	100.0	100.0	32.7	68.6	99.2	100.0	100.0

Notes: The data generating process is identical to those used for Table 1 except that $\sigma_{it} = \sqrt{\exp\{cz_{it}\}}$, $t = 1, ..., T$.

Table 5: Rejection frequencies of the asymptotic and various Wild-Bootstrap RBP and BP tests in panel ADL(1,0) models under conditional heteroskedasticity, GARCH(1,1) (HET4).

	$H_0 : E[u_{it}u_{jt}] = 0$										$H_A : E[u_{it}u_{jt}] = 0.2$									
	SN					χ^2_6					SN					χ^2_6				
N	5	10	25	50	100	5	10	25	50	100	5	10	25	50	100	5	10	25	50	100
	Asymptotic critical values										Asymptotic critical values									
T	RBP_T										RBP_T									
25	4.5	5.8	12.5	31.6	83.2	3.3	5.8	10.9	27.7	76.5	5.1	10.9	33.9	73.6	98.7	5.1	10.2	29.4	63.4	94.4
50	4.0	5.1	7.5	14.9	33.1	4.9	4.7	7.5	13.6	32.6	8.8	19.0	53.2	89.3	99.6	8.0	16.0	45.5	78.0	97.2
100	4.5	4.6	6.5	8.8	14.4	3.9	5.0	7.2	8.5	17.5	15.9	38.7	88.0	99.6	100.0	15.8	31.9	80.9	97.9	100.0
200	5.6	5.2	6.7	5.8	9.3	4.1	4.3	5.6	7.6	11.6	35.5	73.0	100.0	100.0	100.0	30.6	68.1	98.9	100.0	100.0
	BP_T										BP_T									
25	5.1	7.6	16.6	40.8	91.3	5.8	8.4	17.6	42.5	91.0	7.7	14.6	40.9	80.9	99.5	8.6	16.2	44.2	81.3	99.4
50	5.2	5.7	9.5	17.7	37.9	6.0	5.7	9.0	16.8	38.8	10.0	21.4	58.0	91.5	99.8	12.4	23.5	60.1	89.8	99.7
100	5.5	5.3	7.3	9.5	16.6	4.8	6.0	8.2	8.9	15.5	17.1	40.3	88.9	99.9	100.0	20.1	41.0	87.6	99.4	100.0
200	5.7	5.1	6.6	5.9	9.9	4.7	4.7	5.5	6.8	9.4	36.5	74.4	99.9	100.0	100.0	35.4	73.7	99.4	100.0	100.0
T	$NRBP_{NT}$										$NRBP_{NT}$									
25	6.1	7.0	13.2	32.3	83.6	5.4	7.0	11.7	28.5	77.0	7.1	12.7	34.8	74.2	98.8	7.1	12.5	30.3	64.1	94.5
50	5.3	6.1	8.2	15.4	33.7	6.4	5.6	8.0	14.2	33.1	11.5	21.0	54.9	89.6	99.6	10.5	18.1	46.9	78.8	97.2
100	5.9	5.6	7.0	9.1	14.8	5.5	6.1	7.4	8.9	18.0	19.3	41.1	88.5	99.7	100.0	19.2	34.9	81.5	98.0	100.0
200	6.9	5.9	7.1	6.3	9.5	5.5	5.5	5.9	8.1	11.7	40.1	74.8	100.0	100.0	100.0	35.2	70.4	98.9	100.0	100.0
	NBP_{NT}										NBP_{NT}									
25	7.7	8.8	17.2	41.7	91.5	7.4	9.8	18.6	43.4	91.2	9.9	16.4	42.2	81.2	99.5	11.5	18.5	46.2	81.7	99.4
50	6.1	7.3	10.2	18.4	38.3	7.9	7.0	9.9	17.5	39.2	12.8	23.5	59.7	91.6	99.8	15.3	25.9	61.4	90.2	99.7
100	7.0	6.3	7.7	9.7	16.7	6.7	7.4	8.7	9.2	15.6	20.6	43.1	89.2	99.9	100.0	23.4	43.8	88.3	99.4	100.0
200	7.5	6.3	7.0	6.4	10.0	6.2	5.9	5.8	7.1	9.5	40.7	76.0	100.0	100.0	100.0	39.0	76.0	99.4	100.0	100.0
	WB 1: Recursive resampling										WB 1: Recursive resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	5.0	4.7	5.8	5.6	6.8	4.5	5.7	4.5	5.7	6.5	6.1	9.9	20.5	40.0	71.0	6.3	9.5	18.8	35.0	55.8
50	4.4	4.8	5.5	5.8	4.7	5.6	4.8	5.1	5.9	6.4	9.0	18.4	46.6	81.1	97.6	8.9	16.5	41.5	68.8	91.3
100	4.8	4.4	5.4	5.3	5.5	4.7	4.7	6.4	5.8	6.9	16.6	37.8	86.5	99.3	100.0	16.7	32.7	78.6	97.0	99.9
200	5.7	4.7	5.5	4.9	5.3	4.4	4.6	5.2	6.4	7.6	35.9	72.6	99.9	100.0	100.0	31.7	67.8	98.7	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.4	5.1	4.6	5.5	6.0	4.4	4.9	4.9	5.9	6.6	6.4	10.3	21.4	40.6	71.0	7.2	11.2	24.2	45.6	70.7
50	4.5	5.0	5.0	5.8	4.3	5.7	4.5	4.9	4.9	6.1	9.5	18.4	47.7	81.7	98.1	10.6	19.2	49.9	81.1	96.9
100	5.1	4.4	5.3	5.1	5.9	5.2	5.7	6.0	4.5	5.3	16.7	38.0	86.8	99.4	100.0	19.0	38.9	85.4	99.1	100.0
200	5.7	4.6	5.8	4.7	5.2	4.7	4.3	4.5	5.2	4.9	36.1	72.8	100.0	100.0	100.0	34.5	73.7	99.4	100.0	100.0
	WB 2: Fixed-design resampling										WB 2: Fixed-design resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	5.1	5.1	6.7	7.6	12.6	4.8	6.3	5.3	7.7	10.2	5.9	10.5	22.6	44.4	76.6	6.8	10.2	20.9	39.5	62.0
50	4.5	4.8	5.6	7.4	5.9	5.5	4.9	5.7	6.9	8.4	9.3	18.1	47.2	81.7	98.1	9.0	16.3	42.0	70.1	92.2
100	5.2	4.5	5.5	5.2	5.9	4.7	5.3	6.4	5.9	7.8	16.3	37.9	86.6	99.3	100.0	17.1	32.7	78.8	97.1	99.9
200	5.9	4.9	5.9	4.7	5.4	4.2	4.5	4.9	6.0	7.9	35.4	72.4	99.9	100.0	100.0	31.7	67.7	98.7	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.9	4.9	5.8	7.3	11.3	4.8	6.0	6.0	8.3	11.7	6.5	11.0	23.5	46.2	77.9	7.0	11.7	25.9	50.4	76.9
50	4.8	5.0	5.3	6.9	5.6	5.9	4.4	5.3	5.6	7.0	9.4	18.7	48.3	82.5	98.4	10.9	19.7	50.6	82.1	97.3
100	4.8	4.8	5.2	5.1	6.0	5.4	5.5	6.0	4.7	6.0	16.5	37.8	86.6	99.4	100.0	18.8	39.1	85.7	99.2	100.0
200	6.1	4.7	5.9	4.8	5.3	4.5	4.3	4.4	5.1	5.0	35.2	72.9	99.9	100.0	100.0	34.2	73.8	99.3	100.0	100.0
	WB 3: Direct resampling										WB 3: Direct resampling									
T	$NRBP^*_{NT}$										$NRBP^*_{NT}$									
25	5.0	5.3	7.4	9.9	19.8	4.9	6.4	6.0	10.2	14.7	6.0	10.5	25.2	48.3	82.0	6.6	10.7	21.6	42.3	68.4
50	4.3	5.0	6.0	7.9	8.4	5.4	4.9	6.0	7.8	10.9	9.4	18.7	48.1	83.0	98.6	9.2	17.1	42.6	71.2	93.0
100	5.1	4.7	5.7	5.9	7.2	4.9	5.3	6.7	6.3	8.9	16.4	38.0	86.7	99.3	100.0	17.2	32.9	79.2	97.2	99.9
200	5.8	4.8	5.9	5.0	5.9	4.4	4.6	5.1	6.4	8.5	35.4	72.8	100.0	100.0	100.0	31.8	68.2	98.7	100.0	100.0
	NBP^*_{NT}										NBP^*_{NT}									
25	4.7	5.4	6.8	10.1	20.4	4.7	6.3	7.6	11.1	20.6	6.8	11.1	25.3	52.2	84.2	7.4	12.3	28.7	55.1	83.5
50	4.8	4.9	6.0	8.1	8.6	5.7	4.5	5.8	7.0	9.8	9.8	19.4	49.9	83.8	98.6	11.0	19.8	51.8	83.8	97.9
100	4.8	4.8	5.8	5.7	7.3	5.2	5.6	6.3	5.2	7.2	16.6	37.7	87.3	99.5	100.0	18.7	38.7	86.2	99.2	100.0
200	5.9	4.5	5.8	4.6	6.1	4.7	4.4	4.8	5.2	5.8	35.5	72.2	99.9	100.0	100.0	34.1	73.7	99.2	100.0	100.0

Notes: The data generating process is identical to those used for Table 1 except that $\sigma^2_{it} = \delta + \alpha_1 u^2_{i,t-1} + \alpha_2 \sigma^2_{i,t-1}$, $t = -49, -48, ..., T$. The value of parameters are chosen to be $\delta = 1$, $\alpha_1 = 0.1$ and $\alpha_2 = 0.8$.

Table 6: p-values of cross section correlation tests in dynamic empirical growth models, 20 OECD countries, annual data 1955-2004

p-values	$NRBP_{NT}$	NBP_{NT}
asymptotic	0.089	0.017*
wild bootstrap 1	0.118	0.115
wild bootstrap 2	0.112	0.107
wild bootstrap 3	0.108	0.128

Note: The dynamic model estimated is $\Delta\widetilde{lgdpw}_{it} = \theta_{1i} + \theta_{2i}\widetilde{lk}_{it} + \theta_{3i}\Delta\widetilde{lk}_{it} + \theta_{4i}\Delta\widetilde{lk}_{it-1} + \phi_{1i}\Delta\widetilde{lgdpw}_{i,t-1} + \phi_{2i}\Delta\widetilde{lgdpw}_{i,t-2} + u_{it},, i = 1, 2, ..., 20$ and $t = 1, 2, ..., 47$, where $\widetilde{lgdpw}_{it}$ is cross section demeaned log of output per worker and $\widetilde{lk}_{it}$ is cross section demeaned log of the investment share. "*" signifies the null hypothesis being rejected at the 5% level. Asymptotic p-values are obtained referring the value of the statistics to standard normal distribution (one-sided). Bootstrap p-values are based on 5000 bootstrap resampling. Three wild bootstrap schemes are explained in the previous section. For the wild bootstrap scheme 1, $\widetilde{lk}_{it}$, $\Delta\widetilde{lk}_{it}$ and $\Delta\widetilde{lk}_{it-1}$ are treated as fixed.