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 Economics, University of Exeter Business School
[b] Economics, University of Manchester
[c] Department of Economics and Related Studies, University of York

August 16, 2011

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. 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. Whilst the asymptotic validity of the test procedure, under the null, is predicated on the number of time series observations, T, being large relative to the number of cross-section units, N, independence of the cross-sections is not assumed. Across a variety of experimental designs, a Monte Carlo study suggests that, in general (but not always), the predictions from asymptotic theory provide a good guide to the finite sample behaviour of the test. In particular, with skewed errors and/or when N/T is not small, discrepancies can occur. However, for all the experimental designs, any one of three asymptotically valid wild bootstrap approximations (that are considered in this paper) gives very close agreement between the nominal and empirical significance levels of the test. Moreover, in comparison with wild bootstrap "version" of the original Breusch-Pagan test (Godfrey and Yamagata, 2011) the corresponding version of the heteroskedasticity-robust Breusch-Pagan test is more 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 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.

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

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) proposes a, so-called, CD test based on 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 (2010) 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. However, the slope homogeneity assumption of Sarafidis *et al* (2009), Baltagi *et al* (2010) and Juhl (2011) can be restrictive in empirical work. For example, a growing body of literature suggests that a slope homogeneity assumption may not be relevant in macroeconometric applications: see Haque, Pesaran, and Sharma (1999), Bassanini and Scarpetta (2002), amongst others. 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 issue. The Monte Carlo evidence presented by Godfrey and Yamagata (2011) suggests that such a test can provide quite reliable inferences.

This paper makes two contributions which are distinct from Godfrey and Yamagata (2011). First, it proposes a new asymptotically pivotal heteroskedasticity robust Breusch-Pagan test, under the assumption that $T \to \infty$ and N is fixed, that allows 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. 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 will asymptotically over reject, under the null, if and only if the squared errors are (asymptotically) contemporaneously uncorrelated.

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. Second, this paper describes three asymptotically valid wild bootstrap procedure schemes, allowing for lagged dependent regressors, which might be employed in order to provide closer agreement between the desired nominal and the empirical significance level of a test procedure when N/T is not small. In particular, the recursive-design wild bootstrap is asymptotically justified under less restrictive assumptions than those imposed by Goncalves and Kilian (2004) and Godfrey and Tremayne (2005), which rule

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

out certain asymmetric conditional heteroskedastic error processes. In addition, 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 and Section 4 describes the wild bootstrap tests, which are applicable to both the new heteroskedasticity robust Breusch-Pagan test and the original version. Section 5 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 6 provides a simple empirical application. Finally, Section 7 concludes.

2 The Model, Notation & Assumptions

In this paper, we allow for a 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

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

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

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

$\beta_i' = \left(\theta_i', \phi_i'\right)$, $\phi_i' = \left(\phi_{i1}, ..., \phi_{ip}\right)$, 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

$$\hat{\beta}_i = \left(X_i'X_i\right)^{-1} X_i'y_i, \quad i = 1, ..., N.$$

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; see, *inter alia*, Moscone and Tosetti (2009) and Pesaran et al (2008). 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 asymptotic analysis keeps N fixed whilst $T \to \infty$. In addition, the following assumptions are made in which $\mathcal{F}_{t-1}$ is the sigma field generated by: (i) current and 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).

For all $i = 1, ..., N$, the following hold:

Assumption 1: $\{w'_{it}\}$ 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 2:

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

(ii) $E\left[u_{it}^2 | \mathcal{F}_{t-1}\right] = \sigma_{it}^2$, 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 3:

(i) $E(W_i' W_i / T) = \frac{1}{T} \sum_{t=1}^{T} E[w_{it} w_{it}']$ is uniformly positive definite;

(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 4:

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

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

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

(iv) $\omega_{ijT} = \frac{1}{T} \sum_{t=1}^{T} E[u_{it}^2 u_{jt}^2]$ is uniformly positive;

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

Assumption 1 allows w_{it} to contain fixed or random (but strictly exogenous) regressors. Assumption 2 is somewhat weaker than allowing the errors to be serially independent (although they are still uncorrelated). Assumption 2(i) follows from the strict exogeneity assumption on w_{it} and, together with Assumption 2(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 2(ii) and (iii) also allow for general (conditional or unconditional) heteroskedasticity (with σ_{it}^2 possibly varying across within 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

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

shifts in variance. Assumptions 1, 2 and 3 also ensure that, for each i, j, $\frac{X_i' u_j}{\sqrt{T}}$ and $\frac{X_i' X_j}{T} = \frac{1}{T}\sum_{t=1}^{T} x_{it}x_{jt}'$ are both $O_p(1)$, $\left(\frac{1}{T}\sum_{t=1}^{T} x_{it}x_{it}'\right)^{-1}$ exists, in probability, and is $O_p(1)$ and that, consequently, $\hat{\beta}_i - \beta_i = O_p(T^{-1/2})$. Notice, that additional assumptions are required to establish asymptotic normality for $\sqrt{T}(\hat{\beta}_i - \beta_i)$; specifically, these will be sufficient to ensure that $\frac{1}{T}\sum_{t=1}^{T} u_{it}^2 x_{it}x_{it}' - \frac{1}{T}\sum_{t=1}^{T} E\left[u_{it}^2 x_{it}x_{it}'\right] \xrightarrow{p} 0$, with $\frac{1}{T}\sum_{t=1}^{T} E\left[u_{it}^2 x_{it}x_{it}'\right]$ being uniformly positive definite. However, we do not need asymptotic normality 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 4 permits the derivation of the robust test procedure, for cross-section correlation (Lemma 1 and Theorem 1 below). Assumption 4(i) states that u_{it} and u_{jt} are uncorrelated, $i \neq j$, whilst 4(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

The commonly used Breusch-Pagan test statistic is

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

where[3]

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

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 4(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 statistic is defined as

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

where

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

Allowing for heteroskedasticity across both the cross-section and time dimension, we have the following preliminary Lemma which motivates the construction of RBP_T, given in (4):

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

Lemma 1 *Under Assumptions 4(i)-(iv) and 2(v), we have, for all $i \neq j$, and as $T \to \infty$, and fixed N,*

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

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

Theorem 1 *Under Assumptions 1-4, we have, for all $i \neq j$, and as $T \to \infty$, and fixed N*

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

Finally, we have the following Corollary which details the limit distribution of RBP_T.

Corollary 1 *Under Assumptions 1-4, and as $T \to \infty$, with N fixed,*

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

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 $\hat{\rho}_{ij} - \hat{\gamma}_{ij} = o_p(1)$, so that BP_T remains asymptotically valid, as noted earlier. However, in general, we have (under our assumptions)

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

so that, asymptotically at least, $\hat{\rho}_{ij} - \hat{\gamma}_{ij} = o_p(1)$ 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 $\hat{\rho}_{ij} > \hat{\gamma}_{ij}$, 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 $\sigma_{it}^2 = E\left[u_{it}^2 | \mathcal{F}_{t-1}\right]$ 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.

Thus, there will be situations in which BP_T remains asymptotically robust. In general, though, it seems prudent to use a procedure that is robust under quite general forms of (unknown) heteroskedasticity. Although, Theorem 1 shows that the statistic RBP_T is

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, the asymptotic validity ($T \to \infty$, N fixed) of three wild bootstrap schemes is established when applied to both RBP_T and BP_T in a dynamic heterogenous panel data model under non-normality and unknown heteroskedasticity.

4 Wild Bootstrap Procedures

We consider three wild bootstrap procedures, as follows.

4.1 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 RBP_T and BP_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
   $$y^*_{it} = \hat{\beta}'_i x^*_{it} + u^*_{it}. \tag{6}$$
   Here, x^*_{it} is generated recursively, from (6), 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, $\beta'_i = \left(\theta'_i, \phi_i\right)$ 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)
   $$RBP^*_T = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\gamma}^{*2}_{ij}, \quad \hat{\gamma}^*_{ij} = \frac{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \hat{u}^*_{it}\hat{u}^*_{jt}}{\sqrt{\frac{1}{T} \sum_{t=1}^{T} \hat{u}^{*2}_{it}\hat{u}^{*2}_{jt}}} \tag{7}$$
   where $\hat{u}^*_{it} = y^*_{it} - x^{*\prime}_{it}\hat{\beta}^*_i$ is the OLS residual from (6), and
   $$BP^*_T = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}^{*2}_{ij}, \quad \hat{\rho}^*_{ij} = \frac{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \hat{u}^*_{it}\hat{u}^*_{jt}}{\sqrt{\frac{1}{T} \sum_{t=1}^{T} \hat{u}^{*2}_{it} \frac{1}{T} \sum_{t=1}^{T} \hat{u}^{*2}_{jt}}}. \tag{8}$$

3. Calculate the proportion of bootstrap test statistics, RBP_T^* (resp., BP_T^*), from the B repetitions of Step 2c that are at least as large as the actual value of RBP_T (resp., BP_T). 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 $\hat{p} \leq \alpha$.

4.2 Wild Bootstrap 2 (WB2)

This is a fixed design wild bootstrap scheme which replaces (6) in the recursive design scheme with

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

at stage 2b.

4.3 Wild Bootstrap 3 (WB3)

Note, from Theorem 1, $\hat{\gamma}_{ij} - \gamma_{ij} = o_p(1)$; i.e., $\hat{\gamma}_{ij}$ 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 boostrap test statistics

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

   and

$$BP_T^* = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \tilde{\rho}_{ij}^{*2}, \quad \tilde{\rho}_{ij}^* = \frac{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_{it}^* u_{jt}^*}{\sqrt{\frac{1}{T} \sum_{t=1}^{T} \hat{u}_{jt}^{*2} \frac{1}{T} \sum_{t=1}^{T} \hat{u}_{jt}^{*2}}}.$$

3. Calculate the proportion of bootstrap test statistics, RBP_T^* (resp., BP_T^*), from the B repetitions of Step 2b that are at least as large as the actual value of RBP_T (resp., BP_T). 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 following Theorem states the asymptotic validity of these wild bootstrap procedures:[4]

Theorem 2 *Under Assumptions 1-4, and for all three wild bootstrap designs, WB1, WB2 and WB3,*

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

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

[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 ± 0.5 with an equal probability of 0.5.

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 assumptions of Goncalves and Kilian (2004) to justify the resursive-design WB1, since our test criteria are asymptotically independent of $\hat{\beta}_i$.

Henceforth, a test procedure which employs RBP_T (resp., BP_T) in conjunction with asymptotic critical values will be called an "asymptotic test", whilst that 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.

5 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.

5.1 Monte Carlo Design

5.1.1 DGP1

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

$$\begin{aligned} y_{it} &= \theta_{i1} + \theta_{i2} z_{it} + \phi_i y_{i,t-1} + u_{it} \\ &= \theta_i' w_{it} + \phi_i y_{i,t-1} + u_{it}, \;\; i = 1, 2, ..., N \text{ and } t = -49, -48, ..., T \end{aligned} \tag{9}$$

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

$$u_{it} = \sigma_{it} \varepsilon_{it}, \;\; i = 1, 2, ..., N \text{ and } t = -49, -48, ..., T \tag{10}$$

and

$$\varepsilon_{it} = \sqrt{1 - \rho^2} \xi_{it} + \rho \zeta_t \tag{11}$$

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. Three distributions are used to obtain the i.i.d. standardized errors for ξ_{it} and ζ_t: the standard normal distribution; the t-distribution with five degrees of freedom (t_5); and the chi-square distribution with six degrees of freedom (χ^2_6). The t_5 distribution satisfies the restrictions placed on the moments of u_{it} (Assumption 2(v)), whilst the χ^2_6 distribution is 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 2(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 $\max(\sigma_{it}^2)/\min(\sigma_{it}^2)$, 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

$$\sigma_{it}^2 = \delta + \alpha_1 u_{i,t-1}^2 + \alpha_2 \sigma_{i,t-1}^2,\ t = -49, -48, ..., T. \tag{12}$$

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

5.1.2 DGP2

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

$$\begin{aligned} y_{it} &= \beta_{i1} + \beta_{i2} z_{it} + u_{it} && (13) \\ &= \beta_i' w_{it} + u_{it},\ i = 1, 2, ..., N \text{ and } t = 1, 2, ..., T, && (14) \end{aligned}$$

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 (13) is written as

$$u_{it} = \sigma_{it}\varepsilon_{it},\ i = 1, 2, ..., N \text{ and } t = 1, 2, ..., T. \tag{15}$$

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

5.1.3 DGP3

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

$$y_{it} = \theta_i (1 - \phi_i) + \phi_i y_{it-1} + u_{it}, i = 1, 2, ..., N \text{ and } t = -49, -48, ..., T. \tag{16}$$

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

$$u_{it} = \sqrt{1 - \phi_i^2}\sigma_{it}\varepsilon_{it},\ i = 1, 2, ..., N \text{ and } t = -49, -48, ..., T. \tag{17}$$

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

All combinations of $N = 5, 10, 25$ and $T = 50, 100, 200$ 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.

5.2 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.[5]

[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 both the asymptotic, RBP_T and BP_T, tests are in the main close to the nominal significance level of 5%, although there is slight distortion in both when $N = 25$, and smaller values of T, with BP_T being the more sensitive across all error distributions. For example, with normal errors and $N = 25$, BP_T rejection rates are 9.1% and 7.5%, respectively, for $T = 50$ and $T = 100$, but acceptable at 5.3% when $T = 200$. The possibility of such size distortion, when N/T is not "small", has been pointed out Pesaran *et al* (2008). The results indicate that RBP_T also suffers in these circumstances, as might be expected, but the results suggest that this is to a lesser degree. Bearing in mind the general close agreement between nominal and actual significance levels of the asymptotic RBP_T and BP_T tests, a comparison of their rejection frequencies under $H_A : E[u_{it}u_{jt}] = 0.2$, reveals similar power properties under homoskedastic normal and t_5 errors. However, the power of the asymptotic RBP_T test is noticeably lower than that of the asymptotic BP_T test under χ^2_6 errors. For example, with $N = 5$ (resp., $N = 10$) and $T = 100$, the empirical power of RBP_T is 16% (resp., 32%) compared with 24% (resp., 43%) for BP_T. This feature is also a characteristic of the bootstrap tests under all heteroskedasticity schemes considered.

[INSERT Tables 2 - 5 ABOUT HERE]

The results obtained when the errors are heteroskedastic (Tables 2 - 5), show that the asymptotic RBP_T test again exhibits close agreement, in general, between nominal and empirical significance levels across all error distributions. In fact, the results are qualitatively similar to those obtained with homoskedastic errors, with slight distortions apparent when $N = 25$, and for smaller values of T; although, as before, these disappear at $T = 200$.

[5] A full set of results can, of course, be obtained from the authors upon request.

By contrast, and consistent with the analysis at the end of Section 3, the asymptotic BP_T test tends to overreject the null hypothesis significantly, except for GARCH errors (Table 5). 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 BP_T test, across the three error distributions, range from $8.9\% - 11.7\%$, $17.2\% - 18.2\%$ and $51.0\% - 54.9\%$, for $N = 5, 10$ and 25, respectively. For the trending volatility model, Table 3, and the HET3 scheme (Table 4) the corresponding ranges are: $6.0\% - 7.6\%$, $8.4\% - 9.5\%$, $16.0\% - 17.9\%$ and $4.9\% - 6.0\%$, $7.0\% - 7.3\%$, $12.2\% - 14.5\%$, respectively. There is significantly less over-rejection in the latter, where $\sigma_{it}^2 = \exp(cz_{it})$, 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_{it}^2 is a stationary process, BP_T remains asymptotically justified and exhibits close agreement , in general, between nominal and empirical significance levels across all error distributions, although with more pronounced distortions, than that of RBP_T, when $N = 25$ and for smaller values of T.

Turning our attention to the wild bootstrap tests, both procedures, employing RBP_T^* and BP_T^*, control the significance levels much better than their asymptotic counterparts, across models and wild bootstrap schemes. Indeed, under $H_0 : E[u_{it}u_{jt}] = 0$ and over the 135 different models investigated, for each wild bootstrap scheme, there is hardly any evidence of distortion in the empirical significance level for RBP_T^*. Only once, for WB1, and twice, for WB3, do the empirical rejection rate fall outside the acceptable interval of [3.6%, 6.5%], and these all occur under χ_6^2 errors with $N = 25$ and $T = 100$: under HET2 and WB3,and under HET4, WB1 and WB3. In contrast, the empirical rejection rate for BP_T^* falls outside of this interval *four* times, for WB1, and *five* times for each of WB2 and WB3. All of these occur only when $N = 25$ and $T \leq 100$, but with the majority being under the HET3 scheme. Such results for BP_T^* are consistent with those found by Godfrey and Yamagata (2011), although their experiments only considered a static (not dynamic) heterogeneous panel data mode. Thus, both bootstrap tests, RBP_T^* and BP_T^*, exhibit good agreement between nominal and empirical significance levels, although the former appears more reliable than the latter, especially when $N = 25$. With regard to power comparisons, between RBP_T^* and BP_T^*, there is little difference except (as noted under homoskedastic errors) that BP_T^* appears consistently more powerful under χ_6^2 errors. Qualitatively, the results are similar across all schemes but, as an illustration, under GARCH(1,1) correlated errors (Table 5), and for $N = 10$, the rejection rates for BP_T^* are approximately 19%, 39% and 74%, respectively for $T = 50, 100$ and 200, for all wild bootstrap schemes, whilst those of RBP_T^* are 17%, 33% and 69%.

Finally, there appears little to choose between the differing wild bootstrap schemes: WB1, WB2 and WB3. However, the direct resampling wild bootstrap (WB3) has clear advantage of being less computationally costly over other schemes, since it does not require to estimate the model using bootstrap sample.

6 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).[6] 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)).[7]

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

$$\Delta\widetilde{lgdpw}_{it} = \theta_{1i} + \theta_{2i}\widetilde{lk}_{it} + \theta_{3i}\Delta\widetilde{lk}_{it} + \theta_{4i}\Delta\widetilde{lk}_{it-1} + \phi_{1i}\Delta\widetilde{lgdpw}_{i,t-1} + \phi_{2i}\Delta\widetilde{lgdpw}_{i,t-2} + u_{it}, \tag{18}$$

$i = 1, 2, ..., N = 20$ 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

$$RLM_{T,i} = \hat{u}'_i\hat{U}_i\left(\hat{U}'_i M_{xi}\hat{\Lambda}_i M_{ix}\hat{U}_i\right)^{-1}\hat{U}'_i\hat{u}_i \tag{19}$$

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 $t - \ell < 1$, $\ell = 1, 2, ..., m$, $M_{ix} = I_T - X_i\left(X'_i X_i\right)^{-1} X'_i$ 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.[8]

We have applied the WB1 bootstrap $RLM_{T,i}$ test for second-order serial correlation ($m = 2$) to the model (18) 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.[9]

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[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 BP_T test rejects the null hypothesis at the 5% level, but our asymptotic RBP_T 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.

7 Conclusion

The paper has developed a heteroskedasticity robust Breusch-Pagan test 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; but not on the independence of the cross sections. The procedure can be employed with fixed, strictly exogenous and/or lagged dependent regressors and is (asymptotically) robust to quite general forms of non-normality and heteroskedasticity, in the error distribution, across both time and cross-section. One of three wild bootstrap schemes can be used to improve the finite sample behaviour of the test. 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). A Monte Carlo study examines the performance of the new test procedure and its wild bootstrap version in relation to the original Breusch-Pagan test and its wild bootstrap version. Across all combinations of error distributions and types of heteroskedasticity, considered, the wild bootstrap version of the new robust Breusch-Pagan test (RBP_T^*) provided quite reliable finite sample inferences; especially when N/T is not small, as hoped would be the case. Furthermore, the RBP_T^* seems to be as powerful as its asymptotic counterpart, RBP_T, under homoskedasticity and therefore there is no penalty attached to using these wild bootstrap schemes even if the errors are homoskedastic. Surprisingly, perhaps, the Breusch-Pagan wild bootstrap tests also provides significant improvements over first-order asymptotic theory but proved less reliable that RBP_T^*. Thus RBP_T^* recommends itself as an additional useful test procedure for applied workers. Additionally, there is little to be chosen between the different bootstrap schemes presented but direct resampling wild bootstrap scheme is computationally less costly than the other schemes.

References

Baltagi, B., Feng, Q., and Kao, C. (2010). 'A Largrange multiplier test for cross-sectional dependence in Fixed Effects panel data models', *Mimeo. Syracuse University.*

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.

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

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.*

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

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.

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

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

Proof of Lemma 1

By Assumptions 2(v) and 4(i), $\{u_{it}u_{jt}, \mathcal{F}_t\}$ is a m.d.s., with $E\left|u_{it}u_{jt}\right|^{2+\delta} < \infty$ and, by Assumptions 4 (ii) and (iii)

$$\operatorname*{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}\left\{u_{it}^2u_{jt}^2 - E[u_{it}^2u_{jt}^2]\right\} = 0. \tag{20}$$

To verify (20) note that, with $\tau_{ijt}^2 = E[u_{it}^2u_{jt}^2|\mathcal{F}_{t-1}]$,

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

and the second term is $o_p(1)$ by Assumption 4(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 $E\left|u_{it}^2u_{jt}^2\right|^{1+\delta} < \infty$.
Then Assumption 4(iv) and a straightforward application of White (2001, Corollary 5.26, p.135), yields

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

The result then follows by (20). ■

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 $\sum_{k=1}^{\infty} k\left|\psi_{ik}\right| < \infty$ 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 $\sum_{k=1}^{\infty}\|c_{ik}\| < \infty$, for all $i = 1, ..., N$.

Proposition 1 *Under Assumption 2(i),(iv),(v), and for all $i, j = 1, ..., N$:*

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

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

Lemma 2 *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 $\sum_{k=1}^{\infty}\left|\xi_k\right| < \infty$. Then, $\bar{S}_T \xrightarrow{p} 0$.*

The following Lemma exploits Lemma 2 and is central to the proof of Theorem 1.

Lemma 3 *Under Assumptions 1, 2, 3 and 4(i), and for all $i, j = 1, ..., N$:*

(a) $\frac{1}{T}\sum_{t=1}^{T}\left(x_{it}x_{jt}' - E[x_{it}x_{jt}']\right) = o_p(1)$, *where $E[x_{it}x_{jt}'] \leq \Delta < \infty$ uniformly in i, j and 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 $\hat{\gamma}_{ij} - \gamma_{ij} = o_p(1)$ and the result follows.

1. First, define $M_i = I_N - H_i$, $H_i = X_i \left(X_i' X_i\right)^{-1} X_i'$. Then,

$$\begin{aligned}
\sum_{t=1}^{T} \hat{u}_{it}\hat{u}_{jt} &= \hat{u}_i'\hat{u}_j \\
&= u_i' M_i M_j u_j \\
&= u_i' u_j - u_i' H_i u_j - u_i' H_j u_j + u_i' H_i H_j u_j \\
&= \sum_{t=1}^{T} u_{it}u_{jt} - u_i' H_i u_j - u_i' H_j u_j + u_i' H_i H_j u_j
\end{aligned}$$

It follows from Lemma 3 that $u_i' H_i u_j$, $u_i' H_j u_j$ and $u_i' H_i H_j u_j$ are all $O_p(1)$ with $T^{-1} X_i' X_i$, in particular, being uniformly positive definite with probability one.
Thus $T^{-1/2} \sum_{t=1}^{T} \hat{u}_{it}\hat{u}_{jt} = T^{-1/2} \sum_{t=1}^{T} u_{it}u_{jt} + O_p(T^{-1/2})$ and so, by Lemma 1, $\frac{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \hat{u}_{it}\hat{u}_{jt}}{\sqrt{\frac{1}{T} \sum_{t=1}^{T} u_{it}^2 u_{jt}^2}} \xrightarrow{d} N(0,1)$.

2. We now show that $\frac{1}{T} \sum_{t=1}^{T} \hat{u}_{it}^2 \hat{u}_{jt}^2 - \frac{1}{T} \sum_{t=1}^{T} u_{it}^2 u_{jt}^2 = o_p(1)$, 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 $\delta_i = \hat{\beta}_i - \beta_i = O_p(T^{-1/2})$,

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

By Markov's inequality, Assumption 2(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 2(v) and Proposition 1(a). Thus, by Markov's equality, $R_{1T} = O_p(T^{-1})$. Similarly, 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. ■

Proof of Corollary 1

Since $\hat{\gamma}_{ij} - \gamma_{ij} = o_p(1)$ and $\gamma_{ij} \xrightarrow{d} N(0,1)$, $\hat{\gamma}_{ij}^2 \xrightarrow{d} \chi_1^2$. Furthermore, by asymptotic normality of γ_{ij}, verifying that $E\left[u_{it}u_{jt}u_{ks}u_{ms}\right] = 0$, for pairs $(i,j) \neq (p,q)$ and all t, s establishes the asymptotic independence of the $\hat{\gamma}_{ij}$ and the result follows. Firstly, note by Assumption 4(i), $E\left[u_{it}u_{jt}|\mathcal{F}_{t-1}\right] = 0$ 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 4(v) and we are done. ■

Proof of Proposition 1

(a) Since $x'_{it} = (w'_{it}, Y'_{i,t-1})$ we only need to show that $E\left\|Y_{i,t-1}\right\|^{4+\delta} \leq \Delta < \infty$, given Assumption 2(iv). Applying Minkowski's inequality, with $q = 4 + \delta$, we can write

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

and by another application of Minkowski's inequality

$$E\left|r_{it}\right|^q \leq \left(\left\|\theta_i\right\| \left(E\left\|w_{it}\right\|^q\right)^{\frac{1}{q}} + \left(E\left|u_{it}\right|^q\right)^{\frac{1}{q}}\right)^q < \infty \tag{21}$$

by Assumption 2(iv) and (v). This latter bound, (21), will also be exploited in subsequent proofs.

(b) We verify that $E\left|\lambda' x_{it}u_{jt}\right| < \infty$ and $E\left[\lambda' x_{it}u_{jt}|\mathcal{F}_{t-1}\right] = 0$, for all $\lambda \in \mathbb{R}^{p+M}$ such that $\lambda'\lambda = 1$. First, by the triangle inequality and Cauchy-Schwartz

$$\begin{aligned} E\left|\lambda' x_{it}u_{jt}\right| &\leq \sqrt{E\left\|x_{it}\right\|^2 E\left|u_{jt}\right|^2} \\ &< \infty \end{aligned}$$

from (a) and Assumption 2(v). Second, since $\{u_{it}, \mathcal{F}_t\}$ is a m.d.s., $E\left[w_{it}u_{jt}|\mathcal{F}_{t-1}\right] = w_{it}E\left[u_{jt}|\mathcal{F}_{t-1}\right] = 0$, almost surely, for all t and $E\left[Y_{i,t-1}u_{jt}|\mathcal{F}_{t-1}\right] = \sum_{k=1}^{\infty} c_{ik}E\left[r_{i,t-k}u_{jt}|\mathcal{F}_{t-1}\right] = 0$, for $i, j = 1, ..., N$. Thus, $E\left[\lambda' x_{it}u_{jt}|\mathcal{F}_{t-1}\right] = 0$. ■

Proof of Lemma 2

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, for any $\lambda > 0$,

$$\begin{aligned} \lim_{n\to\infty} \limsup_{T\to\infty} \Pr\left(\left|\bar{S}_T - \bar{S}_T^n\right| > \lambda\right) &< \lim_{n\to\infty} \limsup_{T\to\infty} \frac{1}{\lambda} E\left|\sum_{k>n}^{\infty} \xi_k \bar{Z}_{T,k}\right| \\ &< \lim_{n\to\infty} \limsup_{T\to\infty} \frac{1}{\lambda} \sum_{k>n}^{\infty} \left|\xi_k\right| E\left|\bar{Z}_{T,k}\right| \\ &\leq \frac{\Delta}{\lambda} \lim_{n\to\infty} \sum_{k>n}^{\infty} \left|\xi_k\right| = 0 \end{aligned}$$

since $\sum_{k=1}^{\infty} \left|\xi_k\right| < \infty$. Thus $\bar{S}_T \xrightarrow{p} 0$. ■

Proof of Lemma 3

(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 1, $w_{it}w_{jt}$ is mixing and Assumption 2(iv) implies that $E\left\|w_{it}w_{jt}\right\|^{\eta+\delta} \leq \Delta < \infty$, by an application of the Cauchy-Schwartz inequality. Thus, $\frac{1}{T}\sum_{t=1}^{T} w_{i,t-h}w_{j,t-k}' - \frac{1}{T}\sum_{t=1}^{T} E[w_{i,t-h}w_{j,t-k}'] = o_p(1)$, 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 $h, k \in \mathbb{N}$. 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 $E\left\|v_{it}r_{j,t-k}\right\| \leq E\left\|w_{it}r_{j,t-k}\right\| \leq \Delta < \infty$, by Assumption 2(iv), (21) 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 2(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, $\frac{1}{T}\sum_{t=1}^{T}\left(w_{it}w_{j,t-k}' - E[w_{it}w_{j,t-k}']\right) = o_p(1)$, so that $\bar{Z}_{T,k}^{(i,j)} \xrightarrow{p} 0$ for all $\mu \in \mathbb{R}^M$, $\|\mu\| = 1$. Since $\sum_{k=1}^{\infty}\left|\xi_{jk}\right| < \infty$, for all $j = 1, ..., N$, Lemma 2 gives $\frac{1}{T}\sum_{t=1}^{T}\left(w_{it}Y_{j,t-1}' - E[w_{it}Y_{j,t-1}']\right) = o_p(1)$.
Finally, for any $\lambda \in \mathbb{R}^p$ and again writing $\xi_{jk} = c_{jk}'\lambda$,

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

In order to show that $\lambda'\left\{\frac{1}{T}\sum_{t=1}^{T}\left(Y_{i,t-1}Y_{j,t-1}' - E[Y_{i,t-1}Y_{j,t-1}']\right)\right\}\lambda = o_p(1)$, we apply Lemma 2 repeatedly. Thus, we can write

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

where $\bar{Z}_{T,k}^{(i,j)} = \sum_{h=1}^{\infty}\xi_{jh}\bar{Z}_{T,k,h}^{(i,j)}$ and

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

satisfying $E\left|\bar{Z}_{T,k,h}^{(i,j)}\right| \leq \Delta < \infty$, since $E\left|r_{i,t-k}r_{j,t-h}\right| \leq \Delta < \infty$ by Cauchy-Schwartz and (21), which in turn implies $E\left|\bar{Z}_{T,k}^{(i,j)}\right| \leq \Delta < \infty$. 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 2(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 4(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 2(ii)

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

by Assumptions 2(ii) and (iii). 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 $k \in \mathbb{N}$, $\bar{Z}_{T,k,h}^{(i,j)} = o_p(1)$. An application of Lemma 2 establishes first that $\bar{Z}_{T,k}^{(i,j)} = \sum_{h=1}^{\infty}\xi_{jh}\bar{Z}_{T,k,h}^{(i,j)} = o_p(1)$, for fixed $k \in \mathbb{N}$. A second application yields $\sum_{k=1}^{\infty}\xi_{ik}\bar{Z}_{T,k}^{(i,j)} = o_p(1)$, the desired result.

(b) By part (a), for $i = j$, $\frac{1}{T}\sum_{t=1}^{T} x_{it}x_{it}' - Q_{iT} = o_p(1)$, 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 3(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 3(ii), for any $k \leq p$, and $\sum_{k=1}^{p}\left|\lambda' c_{ik}\right|^2 > 0$, 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 $var\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} x_{it}u_{jt}\right] = O(1)$. 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 2(v) and Proposition 1(a), and a repeated application of Cauchy-Schwartz, it can be shown that $E\left\|u_{it}^2 x_{it}x_{it}'\right\| = O(1)$, uniformly in t, and the result follows.

This completes the proof. ■

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 ± 0.5 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 (6),

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

where $r_{it}^* = \hat{\theta}_i' w_{it} + u_{it}^*$. 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 $b_{it}^* = \mathbf{1}(t > 0) r_{it}^*$, 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 boostrap probability measure occurs in a set with probability converging to one; i.e., if the proposed limit distribution is $\mathcal{D}(x)$ then, $\sup_{x\in\mathbb{R}} |P^*(S_T^* \leq x) - \mathcal{D}(x)| = o_p(1)$.

Furthermore, in what follows, let $\mathcal{F}_t^*$ 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 informs the proof of Theorem 2 and are the bootstrap counterparts of Lemmas 2 and 3:

Lemma 4 *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 5 *Under Assumptions 1,2 and 4(i),(ii) and (iii), and for all $i, j = 1, ..., N$:*

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

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

Proof of Theorem 2

Consider first RBP_T^*. For $\hat{\gamma}_{ij}^*$ defined at (7), we first show that

$$\hat{\gamma}_{ij}^* = \gamma_{ij}^* + o_{p^*}(1), \tag{22}$$

in probability; and, second, that

$$\gamma_{ij}^* \equiv \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_{it}^* u_{jt}^*}{\sqrt{\frac{1}{T}\sum_{t=1}^{T} \hat{u}_{it}^2 \hat{u}_{jt}^2}} \xrightarrow{d^*} N(0,1), \tag{23}$$

in probability. In particular, the variance estimator employed in the construction of $\hat{\gamma}_{ij}^*$ is asymptotically equivalent to the variance estimator employed in the construction of $\hat{\gamma}_{ij}$; c.f., Goncalves and Killian (2004, Corollary 3.1). It follows immediately that $\hat{\gamma}_{ij}^* \xrightarrow{d^*} N(0,1)$, in probability. Thus, by Theorem 1 and continuity of the normal distribution,

$$\sup_x \left| P^*(\hat{\gamma}_{ij}^* \leq x) - P\left(\hat{\gamma}_{ij} \leq x\right) \right| \xrightarrow{p} 0$$

as $T \to \infty$ and for fixed N, and the result follows since $E^*\left[u_{it}^* u_{jt}^* u_{hs}^* u_{ks}^*\right] = 0$, for distinct pairs (i,j) and (h,k). (Note, in passing, that the asymptotic validity of WB3 follows immediately from (22).)

Step 1: First, define $H_i^* = X_i^* \left(X_i^{*\prime} X_i^*\right)^{-1} X_i^{*\prime}$, where X_i^* has rows x_{it}^*, with $u_i^* = (u_{i1}^*, ..., u_{iT}^*)$. Note that Lemma 5 and Assumption 3(i), ensures that $\left(X_i^{*\prime} X_i^*/T\right)^{-1}$ exists for sufficiently large T and is $O_{p^*}(1)$, in probability. Then,

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

It is immediate from Lemma 5 (a) and (b), and Lemma 3(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. Furthermore, since $\frac{1}{T}\sum_{t=1}^T \left(\hat{u}_{it}^2 \hat{u}_{jt}^2 - E[u_{it}^2 u_{jt}^2]\right) = o_p(1)$, Lemma 5 (d) 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. The result in (22) follows immediately.

Step 2: Write

$$\gamma_{ij}^* = \frac{\frac{1}{\sqrt{T}} \sum_{t=1}^T \varepsilon_{it} \varepsilon_{jt} \hat{u}_{it} \hat{u}_{jt}}{\sqrt{\frac{1}{T} \sum_{t=1}^T \hat{u}_{it}^2 \hat{u}_{jt}^2}} = \frac{1}{\sqrt{T}} \sum_{t=1}^T \zeta_{ijt,T}^*,$$

where $\zeta_{ijt,T}^* = \frac{\varepsilon_{it}\varepsilon_{jt}\hat{u}_{it}\hat{u}_{jt}}{\sqrt{\frac{1}{T}\sum_{t=1}^T \hat{u}_{it}^2 \hat{u}_{jt}^2}}$. Now, $E^*[\zeta_{ijt,T}^*] = 0$ and, due to (conditional) independence,

$$var^*\left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \zeta_{ijt,T}^*\right] \equiv \frac{1}{T} \sum_{t=1}^T \zeta_{ijt,T}^{*2} = 1.$$

To apply a (triangular array) Central Limit Theorem for (conditionally) independent, but heterogeneous data, it suffices to check that the Liapounov condition[10]

$$T^{-(1+\delta)} \sum_{t=1}^T E^* \left|\zeta_{ijt,T}^*\right|^{2(1+\delta)} = o_p(1).$$

But this is true because

$$\frac{1}{T} \sum_{t=1}^T E^* \left|\zeta_{ijt,T}^*\right|^{2(1+\delta)} = \left\{ \frac{1}{T} \sum_{t=1}^T \hat{u}_{it}^2 \hat{u}_{jt}^2 \right\}^{-(1+\delta)} \frac{1}{T} \sum_{t=1}^T \left|\hat{u}_{it}\hat{u}_{jt}\right|^{2(1+\delta)} = O_p(1)$$

since, by Cauchy-Schwartz inequality,

$$\frac{1}{T} \sum_{t=1}^T \left|\hat{u}_{it}\hat{u}_{jt}\right|^{2(1+\delta)} \leq \sqrt{\frac{1}{T} \sum_{t=1}^T \left|\hat{u}_{it}\right|^{4(1+\delta)} \; \frac{1}{T} \sum_{t=1}^T \left|\hat{u}_{jt}\right|^{4(1+\delta)}}$$

and $T^{-1}\sum_{t=1}^T \left|\hat{u}_{it}\right|^{4(1+\delta)} = O_p(1)$, under our assumptions, since $\hat{u}_{it} = u_{it} - x_{it}'(\hat{\beta}_i - \beta_i)$. To see the latter, write $q = 4(1+\delta)$ and apply Minkowski's inequality, which yields

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \left|\hat{u}_{jt}\right|^q &\leq \left(\left\{ \frac{1}{T} \sum_{t=1}^T \left|u_{jt}\right|^q \right\}^{1/q} + \left\{ \frac{1}{T} \sum_{t=1}^T \left|x_{jt}'\hat{\beta}_j\right|^q \right\}^{1/q} \right)^q \\ &\leq \left(\left\{ \frac{1}{T} \sum_{t=1}^T \left|u_{jt}\right|^q \right\}^{1/q} + \left\|\hat{\beta}_j\right\| \left\{ \frac{1}{T} \sum_{t=1}^T \left\|x_{jt}\right\|^q \right\}^{1/q} \right)^q. \end{aligned}$$

[10] Here the stronger Liapounov condition replaces the Lindeberg condition of, for example, White (2001, p.117).

Since $\hat{\beta}_j = O_p(1)$, the right hand side is $O_p(1)$ by Markov's inequality applied to $\frac{1}{T}\sum_{t=1}^{T}|u_{jt}|^q$ and $\frac{1}{T}\sum_{t=1}^{T}\|x_{jt}\|^q$, exploiting Assumptions 2(iv) and (v).

Thus, $\gamma_{ij}^* = T^{-1/2}\sum_{t=1}^{T}\zeta_{ijt,T}^* \xrightarrow{d^*} N(0,1)$, in probability, and this completes the proof for RBP_T^*.

Now consider BP_T^*. Firstly, since our assumptions ensure that $\frac{1}{T}\sum_{t=1}^{T}(\hat{u}_{it}^2 - E[u_{it}^2]) = o_p(1)$ and $\frac{1}{T}\sum_{t=1}^{T}(\hat{u}_{it}^2\hat{u}_{jt}^2 - E[u_{it}^2u_{jt}^2]) = o_p(1)$, $\hat{\rho}_{ij} = \sqrt{v_T^{ij}}\hat{\gamma}_{ij} + o_p(1)$ where (the scalar)

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

and is strictly positive for T sufficiently large, by Assumptions 3(ii) and 4(iv). Furthermore, for $\hat{\rho}_{ij}^*$ defined at (8), and by Lemma 5 (c) 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, we can write

$$\sup_x \left|P^*(\hat{\rho}_{ij}^* \le x) - P\left(\hat{\rho}_{ij} \le x\right)\right| = \sup_x \left|P^*(\gamma_{ij}^* \le x/\sqrt{v_T^{ij}}) - P(\hat{\gamma}_{ij} \le x/\sqrt{v_T^{ij}})\right| + o_{p^*}(1),$$

in probability, as $T \to \infty$ and for fixed N. The result then follows from the analysis for RBP_T^*, above. ■

Proof of Lemma 4

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, $E^*\left|S_{2T}^{*n}\right| \le M_T\sum_{k=1}^{n-1}(\hat{\xi}_k - \xi_k) = o_p(1)$, 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

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

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 $\sup_{T\ge T_1}\sum_{k=1}^{\infty}\left|\hat{\xi}_k\right| < \infty$, in probability (c.f. Bühlmann, 1995, Lemma 2.2) which implies that $\sup_{T\ge T_1}\sum_{k=n}^{\infty}\left|\hat{\xi}_k\right| = o_p(1)$ as $n \to \infty$. Thus

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

which completes the proof. ■

Proof of Lemma 5

(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:
(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) $T^{-1}\sum_{t=1}^{T}\left(Y_{i,t-1}^*Y_{j,t-1}'^* - E[Y_{i,t-1}Y_{j,t-1}']\right) = o_{p^*}(1)$, in probability.

For (i), exploiting $b_{it}^* = \mathbf{1}(t>0)(\hat{\theta}_i' w_{it} + u_{it}^*)$, we can write for any $\mu \in \mathbb{R}^M$ and any $\lambda \in \mathbb{R}^p$ such that $\|\mu\| = \|\lambda\| = 1$,

$$\mu \left\{ \frac{1}{T} \sum_{t=1}^{T} \left(w_{it} Y_{j,t-1}'^{*} - E[w_{it} Y_{j,t-1}'] \right) \right\} \lambda = \frac{1}{T} \sum_{t=1}^{T} \left(\sum_{k=1}^{T-1} \hat{\xi}_{jk} v_{it} b_{j,t-k}^* - \sum_{k=1}^{\infty} \xi_{jk} E[v_{it} r_{j,t-k}] \right)$$

where $v_{it} = \mu' w_{it}$ and $\hat{\xi}_{jk} = \hat{c}_{jk}'\lambda$, $\xi_{jk} = c_{jk}'\lambda$, such that $\hat{\xi}_k - \xi_k = o_p(1)$, and $\sum_{k=1}^{\infty} |\xi_k| < \infty$. Thus,

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

where

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

and

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

Now apply Lemma 4 to $\bar{S}_T^{*(i,j)}$. First, by the triangle inequality and noting that $|\varepsilon_{j,t-k}| = 1$,

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

and,

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

which is also $O_p(1)$. Thus $E^* \left| \bar{Z}_{T,k}^{*(i,j)} \right| \leq M_T = O_p(1)$ uniformly in k.
Second, $\left| \bar{\mu}_{T,k}^{(i,j)} \right| \leq \|\theta_j\| \frac{1}{T} \sum_{t=1}^{T} E \|w_{it}\|^2 \leq \Delta < \infty$, by the triangle inequality, Assumption 2(iv), and Cauchy-Schwartz.
Third, to establish that $\bar{Z}_{T,k}^{*(i,j)} - \bar{\mu}_{T,k}^{(i,j)} = o_{p^*}(1)$, in probability, note that for any fixed $k \in \mathbb{N}$,

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

so that

$$\begin{aligned} \bar{Z}_{T,k}^{*(i,j)} - \bar{\mu}_{T,k}^{(i,j)} &= \mu' \frac{1}{T} \sum_{t=k+1}^{T} w_{it} u_{j,t-k}^* + o_{p^*}(1) \\ &= \mu' \frac{1}{T} \sum_{t=1}^{T-k} w_{i,t+k} u_{jt}^* + o_{p^*}(1). \end{aligned}$$

It follows that, conditional on the original sample,

$$\begin{aligned} E^*\left[\mu' w_{i,t+k} u^*_{jt} | \mathcal{F}^*_{t-1}\right] &= \mu' w_{i,t+k} E^*\left[\varepsilon_{jt}\hat{u}_{jt} | \mathcal{F}^*_{t-1}\right] \\ &= \mu' w_{i,t+k} \hat{u}_{jt} E^*\left[\varepsilon_{jt} | \mathcal{F}^*_{t-1}\right] \\ &= 0 \end{aligned}$$

so that $\left\{\mu' w_{i,t+k} u^*_{jt}, \mathcal{F}^*_t\right\}$ is a m.d.s. and, by Cauchy-Schwartz,

$$\begin{aligned} var^*\left[\mu' \frac{1}{T}\sum_{t=1}^{T-k} w_{i,t+k} u^*_{jt}\right] &\leq \frac{1}{T^2}\sum_{t=1}^{T-k} \hat{u}^2_{jt} \left\|w_{i,t+k}\right\|^2 \\ &\leq \frac{1}{T}\sqrt{\frac{1}{T}\sum_{t=1}^{T-k} \hat{u}^2_{jt} \frac{1}{T}\sum_{t=1}^{T-k} \left\|w_{i,t+k}\right\|^2} \\ &\leq \frac{1}{T}\sqrt{\frac{1}{T}\sum_{t=1}^{T} \hat{u}^2_{jt} \frac{1}{T}\sum_{t=1}^{T} \left\|w_{it}\right\|^2} \\ &= O_p(T^{-1}) \end{aligned}$$

because both $\frac{1}{T}\sum_{t=1}^{T} \hat{u}^2_{jt}$ and $\frac{1}{T}\sum_{t=1}^{T} \left\|w_{it}\right\|^2$ are $O_p(1)$. Therefore, by Chebyshev's inequality $\bar{Z}^{*(i,j)}_{T,k} - \bar{\mu}^{(i,j)}_{T,k} = o_{p^*}(1)$ and we are done.
For (ii), we can write, for any $\lambda \in \mathbb{R}^p$ such that $\|\lambda\| = 1$,

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

where

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

and

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

Again, we apply Lemma 4 (twice), to $\bar{S}^{*(i,j)}_T$. First, and by arguments similar to those used above, $E^*\left|\bar{Z}^{*(i,j)}_{T,k,h}\right| \leq M_T = O_p(1)$, uniformly in k and h, noting that

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

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: $\frac{1}{T}\sum_{t=\max(k,h)+1}^{T}\left(w_{i,t-k}w_{j,t-h}'-E\left[w_{i,t-k}w_{j,t-h}'\right]\right)=o_p(1)$, $\frac{1}{T}\sum_{t=1}^{\max(k,h)}E[w_{i,t-k}w_{j,t-h}']=O(T^{-1})$, $\frac{1}{T}\sum_{t=1}^{\max(k,h)}E[u_{i,t-k}u_{j,t-h}']=O(T^{-1})$, and similar arguments to before show that, for example, $\frac{1}{T}\sum_{t=\max(k,h)+1}^{T}w_{i,t-k}u_{j,t-h}^{*}=o_{p^*}(1)$, 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 4(i), $\left\{u_{i,t-k}^{*}u_{j,t-h}^{*},\mathcal{F}_{t-g}^{*}\right\}$, $g=\min(k,h)$, 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 2(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, $\bar{Z}_{T,k,h}^{*(i,j)}-\bar{\mu}_{T,k,h}^{(i,j)}=o_{p^*}(1)$, 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 $\|\mu\|=1$, $\left\{\mu'w_{it}u_{jt}^{*},\mathcal{F}_t^{*}\right\}$ 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 $\hat{\xi}_{ik}=\lambda'\hat{c}_{ik}$, $b_{it}^{*}=\mathbf{1}\left(t>0\right)\hat{r}_{it}^{*}$, 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. 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, $\left\{u_{jt}^{*}\kappa_{it}^{*},\mathcal{F}_t^{*}\right\}$ 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

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

since

$$\frac{1}{T}\sum_{t=1}^{T}\left|b_{i,t-k}^{*}b_{i,t-h}^{*}\right|^{2} \leq \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left|b_{i,t-k}^{*}\right|^{4}\;\frac{1}{T}\sum_{t=1}^{T}\left|b_{i,t-h}^{*}\right|^{4}} \leq \frac{1}{T}\sum_{t=1}^{T}\left|r_{it}^{*}\right|^{4}.$$

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

(c) Write $\hat{u}_{it}^{*} = u_{it}^{*} - x_{it}^{*\prime}(\hat{\beta}_i^{*} - \hat{\beta}_i)$. From (a) and (b) above, $\hat{\beta}_i^{*} - \hat{\beta}_i = O_{p^*}(T^{-1})$, in probability, so that $\frac{1}{T}\sum_{t=1}^{T}\left(\hat{u}_{it}^{*2} - \hat{u}_{it}^{2}\right) = o_{p^*}(1)$, since $\varepsilon_{it}^{2} = 1$. Moreover, our assumptions ensure that $\frac{1}{T}\sum_{t=1}^{T}\left(\hat{u}_{it}^{2} - E[u_{it}^{2}]\right) \xrightarrow{p} 0$, and the result follows by the triangle inequality.

(d) Since $\hat{\beta}_i^{*} - \hat{\beta}_i = o_{p^*}(1)$, in probability, it is sufficient to show that $\frac{1}{T}\sum_{t=1}^{T}y_{i,t-h}^{*4} = O_{p^*}(1)$, in probability, 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, given $\varepsilon_{it}^{2} = 1$. The result follows by the triangle inequality and $\frac{1}{T}\sum_{t=1}^{T}\left(\hat{u}_{it}^{2}\hat{u}_{jt}^{2} - E[u_{it}^{2}u_{jt}^{2}]\right) \xrightarrow{p} 0$. Briefly, let e_h be the $(p\times 1)$ unit vector with 1 at position h and zeros elsewhere and let $\hat{\xi}_{ik} = \hat{c}_{ik}^{\prime}e_h$. 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. ■

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			t_5			χ^2_6			SN			t_5			χ^2_6		
N	5	10	25	5	10	25	5	10	25	5	10	25	5	10	25	5	10	25
	Asymptotic critical values									Asymptotic critical values								
T	RBP_T									RBP_T								
50	5.1	5.8	7.7	4.5	3.5	6.4	3.7	3.9	7.0	9.6	17.2	53.3	9.8	20.2	55.2	8.2	16.1	42.9
100	4.3	5.7	6.8	4.0	5.1	5.6	4.6	3.9	7.0	18.0	40.0	89.0	17.7	42.1	87.0	15.6	31.7	80.3
200	4.6	5.0	5.5	4.6	5.1	4.8	4.4	5.4	5.9	35.1	75.6	99.9	37.4	77.5	99.6	30.1	69.1	99.0
	BP_T									BP_T								
50	5.1	5.6	9.1	5.4	5.5	8.4	4.9	5.2	8.4	10.5	20.0	59.9	11.7	24.5	59.5	12.1	24.3	57.9
100	5.1	6.3	7.5	5.2	6.8	7.1	4.9	4.9	7.1	19.2	41.5	89.9	19.8	44.0	87.0	21.1	42.9	88.7
200	4.8	5.2	5.3	6.0	5.9	5.6	4.2	5.8	5.4	36.2	76.0	99.9	37.9	77.2	99.3	35.8	74.9	99.5
	WB 1: Recursive resampling									WB 1: Recursive resampling								
T	RBP^*_T									RBP^*_T								
50	5.8	5.7	4.7	5.3	4.6	5.0	4.7	4.1	4.5	10.3	17.3	48.0	11.3	20.9	51.4	9.4	16.1	37.7
100	4.5	5.4	5.6	4.8	5.7	5.0	5.2	4.9	6.3	18.5	39.3	87.2	19.6	43.2	86.3	17.1	32.7	79.1
200	4.8	4.7	4.9	4.8	5.7	4.9	4.9	5.8	5.2	34.8	74.6	99.9	38.8	77.4	99.6	31.6	69.5	98.9
	BP^*_T									BP^*_T								
50	5.2	4.9	4.0	5.2	4.4	4.6	4.8	3.9	4.3	10.5	16.7	48.8	10.7	20.8	49.7	11.4	19.9	48.2
100	4.6	5.7	5.7	4.7	5.9	5.1	4.6	4.6	5.5	18.7	40.1	87.5	18.3	42.0	84.5	19.8	39.5	86.6
200	4.5	4.8	4.5	5.3	5.5	5.1	4.3	5.3	4.8	35.2	74.9	99.9	36.4	75.8	99.3	34.3	74.1	99.6
	WB 2: Fixed design resampling									WB 2: Fixed design resampling								
T	RBP^*_T									RBP^*_T								
50	5.6	5.7	5.2	5.6	4.3	5.2	4.7	4.1	4.9	10.7	17.2	47.6	11.7	20.8	51.6	10.1	15.7	38.2
100	4.6	5.5	5.9	5.0	5.5	5.0	5.3	4.8	6.1	18.1	39.1	86.9	20.0	43.6	86.3	17.7	32.8	79.4
200	4.8	4.8	4.9	5.0	5.9	4.8	4.9	5.6	5.1	34.6	74.5	99.9	38.5	77.7	99.6	31.7	69.2	99.0
	BP^*_T									BP^*_T								
50	5.7	5.4	4.7	5.3	4.7	5.1	4.9	4.1	4.2	10.2	16.9	48.8	10.8	20.9	50.0	11.1	19.6	49.1
100	4.8	5.5	5.9	4.7	5.9	5.2	5.0	4.2	5.3	18.8	40.0	87.5	18.4	41.8	84.8	19.2	39.2	86.9
200	4.6	5.1	5.0	5.1	5.5	4.8	4.2	5.7	4.7	34.8	74.8	99.9	36.6	75.6	99.3	34.7	74.0	99.6
	WB 3: Direct error resampling									WB 3: Direct error resampling								
T	RBP^*_T									RBP^*_T								
50	5.8	5.9	5.3	5.4	4.9	5.1	4.8	3.9	5.1	10.3	17.5	48.6	12.2	21.8	52.9	10.0	16.7	39.5
100	4.3	5.5	6.3	4.9	5.9	5.2	5.4	4.7	6.5	18.3	39.8	87.3	19.5	43.3	86.8	17.2	32.8	79.6
200	4.8	4.8	5.1	4.7	5.3	5.0	4.4	5.7	5.1	34.5	74.7	99.9	38.7	77.8	99.5	31.4	69.5	99.0
	BP^*_T									BP^*_T								
50	5.5	5.2	5.1	5.2	4.6	5.6	4.9	4.2	4.5	10.5	17.4	50.4	11.3	21.8	51.1	11.2	20.2	50.1
100	4.5	5.7	5.9	4.7	5.9	5.2	4.7	4.4	6.0	18.5	40.2	88.0	18.4	42.1	84.5	19.5	39.8	87.3
200	4.9	4.9	4.5	5.1	5.7	5.0	4.0	5.3	4.7	34.9	75.0	99.9	36.7	75.6	99.3	34.3	74.1	99.6

Notes: The first 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. $(0, 1)$ and $\zeta_t \sim$ i.i.d. $(0, 1)$, which are independent of each other. For estimating significance levels, $\rho = 0.0$. Power is investigated using $\rho = 0.2$. The i.i.d. standardized errors for ξ_{it} and ζ_t are drawn from: the standard normal distribution (SN); the t-distribution with five degrees of freedom (t_5); and the chi-square distribution with six degrees of freedom (χ^2_6). The RBP_T test signifies the proposed robust cross sectional correlation test, and the BP_T test is the LM test of Breusch and Pagan (1980). Three wild bootstrap procedures are explained in the earlier section in details. The sampling behaviour of the tests are investigated using 2000 replications of sample data and 200 bootstrap samples, employing a nominal 5% significance level.

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			t_5			χ^2_6			SN			t_5			χ^2_6		
N	5	10	25	5	10	25	5	10	25	5	10	25	5	10	25	5	10	25
	Asymptotic critical values									Asymptotic critical values								
T	RBP_T									RBP_T								
50	4.7	4.5	7.2	3.7	4.1	6.4	3.9	4.0	7.1	8.6	16.1	49.1	8.6	18.3	52.6	7.6	14.0	40.6
100	4.2	5.5	6.7	4.0	4.4	5.8	4.4	4.0	6.4	14.8	34.3	84.8	15.2	39.2	82.5	14.0	28.7	75.0
200	4.7	5.1	5.0	4.2	5.0	4.9	4.1	5.5	5.5	29.7	67.8	99.9	33.7	72.4	99.3	26.7	62.1	97.8
	BP_T									BP_T								
50	9.5	17.4	56.2	9.4	16.3	52.8	10.1	15.7	53.2	16.2	35.7	89.8	17.5	37.7	87.0	18.4	37.4	88.8
100	8.5	16.9	52.9	9.7	17.4	50.2	9.8	16.7	52.9	25.4	57.4	98.4	25.2	59.8	97.8	27.4	56.4	98.5
200	9.3	17.2	54.9	11.7	18.2	51.8	8.9	18.1	51.0	43.2	84.4	100.0	44.4	85.2	99.9	42.5	84.1	100.0
	WB 1: Recursive resampling									WB 1: Recursive resampling								
T	RBP^*_T									RBP^*_T								
50	5.2	4.6	4.3	4.6	4.5	4.7	4.5	4.2	5.0	9.7	15.9	42.3	9.7	19.8	49.1	8.7	14.6	35.8
100	4.5	5.8	5.8	5.0	5.1	5.0	5.5	4.4	5.6	16.1	34.3	82.1	17.0	40.3	81.5	15.9	30.0	73.4
200	5.1	5.4	5.1	5.5	5.4	4.9	4.3	5.9	5.0	29.7	67.2	99.8	34.8	72.1	99.3	28.0	63.0	97.7
	BP^*_T									BP^*_T								
50	4.7	5.1	6.3	5.3	5.0	6.1	5.6	4.3	5.5	10.0	16.1	47.3	9.5	20.0	49.5	10.7	18.8	47.8
100	4.4	5.9	6.7	4.7	5.9	5.7	5.2	4.4	5.9	16.3	35.5	83.3	15.8	39.4	80.6	17.7	36.5	83.7
200	5.1	5.5	5.6	5.3	6.3	5.9	4.7	5.3	5.2	30.8	67.9	99.8	33.5	69.5	98.9	31.4	68.6	99.0
	WB 2: Fixed design resampling									WB 2: Fixed design resampling								
T	RBP^*_T									RBP^*_T								
50	5.3	4.3	4.4	4.8	4.8	5.0	4.6	4.2	4.9	9.8	16.0	43.4	10.4	19.8	49.5	9.4	14.7	35.8
100	4.4	5.8	5.8	4.8	5.1	5.3	5.1	4.3	5.7	15.9	34.3	82.4	16.7	40.6	81.7	16.0	29.8	73.5
200	4.9	5.1	4.7	5.5	5.5	4.9	4.5	5.9	4.9	29.8	67.3	99.7	34.8	72.2	99.2	27.9	63.0	97.6
	BP^*_T									BP^*_T								
50	4.8	5.3	6.3	5.3	5.3	7.1	6.1	4.6	5.9	9.8	16.2	47.5	10.2	20.2	49.6	10.3	19.0	49.1
100	4.4	6.1	6.7	4.4	5.9	6.1	4.7	4.4	6.0	16.5	35.7	83.3	16.2	39.2	81.0	17.4	35.9	83.8
200	4.9	5.5	5.4	5.5	6.0	5.8	4.4	5.4	5.3	30.9	67.7	99.9	34.1	69.3	98.9	31.4	68.9	99.0
	WB 3: Direct error resampling									WB 3: Direct error resampling								
T	RBP^*_T									RBP^*_T								
50	5.5	4.9	4.8	5.0	5.1	5.8	4.6	4.0	5.4	9.6	16.1	44.2	10.3	19.9	49.8	9.4	14.7	37.2
100	4.2	5.9	5.7	4.9	5.1	5.6	5.3	4.5	5.9	16.4	34.0	82.3	17.4	40.8	82.3	15.9	30.3	73.8
200	5.1	5.5	4.9	5.2	5.2	4.9	4.3	6.1	5.1	29.9	67.6	99.8	35.0	72.7	99.3	28.3	62.9	97.7
	BP^*_T									BP^*_T								
50	4.7	4.7	5.5	5.1	5.0	6.1	5.2	4.4	5.9	9.8	15.5	46.9	10.0	19.5	49.5	10.1	18.6	47.7
100	4.3	6.1	6.1	4.4	5.8	5.5	4.7	4.3	5.9	16.0	35.1	82.9	15.9	38.8	80.1	17.4	35.4	83.4
200	4.6	5.5	5.2	5.2	6.0	5.5	4.5	5.3	5.1	30.5	67.4	99.8	33.9	69.7	98.9	31.1	68.1	99.0

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			t_5			χ^2_6			SN			t_5			χ^2_6		
N	5	10	25	5	10	25	5	10	25	5	10	25	5	10	25	5	10	25
	Asymptotic critical values									Asymptotic critical values								
T	RBP_T									RBP_T								
50	5.2	4.9	7.4	4.0	3.9	6.4	4.1	3.8	7.3	8.8	16.6	51.4	9.1	19.6	55.1	8.2	14.9	41.6
100	4.0	5.4	6.4	4.0	4.4	5.2	4.5	3.4	6.9	16.8	38.8	87.4	17.0	40.8	85.7	14.4	30.5	78.9
200	4.7	4.8	5.2	4.2	5.2	5.1	4.2	5.2	5.5	32.9	72.5	99.9	37.1	75.5	99.5	29.0	66.8	98.5
	BP_T									BP_T								
50	6.8	9.5	21.3	6.3	8.2	21.1	6.4	8.5	20.0	12.4	24.2	72.3	13.3	29.6	71.3	14.1	28.9	71.0
100	5.5	9.4	19.5	6.1	9.0	18.4	6.0	8.1	19.5	21.4	47.5	94.5	21.0	49.2	91.7	23.4	47.9	93.6
200	6.3	8.4	17.9	7.6	9.5	17.2	6.0	9.1	16.0	38.6	79.2	100.0	39.8	80.3	99.7	39.0	79.2	99.9
	WB 1: Recursive resampling									WB 1: Recursive resampling								
T	RBP^*_T									RBP^*_T								
50	5.9	5.1	4.5	5.3	4.3	4.7	5.0	4.2	4.9	10.0	16.6	45.9	11.0	20.3	50.2	9.2	15.2	37.6
100	4.1	5.6	5.4	4.7	5.1	5.0	5.2	3.9	6.4	17.3	38.2	85.5	18.3	42.8	84.9	16.3	32.1	77.1
200	5.1	4.8	4.8	4.8	5.7	5.6	4.7	5.8	5.1	32.5	71.1	99.9	37.8	75.6	99.5	30.6	67.0	98.6
	BP^*_T									BP^*_T								
50	5.2	5.1	5.0	5.1	5.1	5.3	5.0	4.1	4.7	9.9	17.2	48.0	10.2	20.3	50.4	11.2	19.5	48.5
100	4.4	5.9	6.2	4.5	5.8	4.9	4.8	3.6	6.1	17.5	38.6	86.2	17.7	40.2	83.1	18.3	38.7	86.1
200	4.8	5.3	5.0	5.2	5.8	5.6	4.3	5.6	4.8	32.7	71.8	99.9	35.9	73.1	99.3	33.1	71.9	99.3
	WB 2: Fixed design resampling									WB 2: Fixed design resampling								
T	RBP^*_T									RBP^*_T								
50	5.7	4.8	4.8	5.0	4.3	4.9	5.0	4.3	5.4	10.1	16.8	45.7	11.0	20.8	51.2	9.7	15.9	37.9
100	4.1	5.1	5.4	4.4	5.1	4.9	5.1	4.0	6.1	17.5	38.6	85.4	18.5	42.5	85.0	16.7	32.3	77.4
200	5.4	4.9	4.7	4.9	5.9	5.3	4.8	6.1	5.1	32.9	71.0	99.9	38.0	75.9	99.5	30.3	67.3	98.5
	BP^*_T									BP^*_T								
50	5.5	5.1	5.3	5.0	4.9	5.9	5.1	4.1	4.9	10.1	17.2	47.8	10.6	20.4	51.1	11.4	19.4	49.1
100	4.2	5.7	6.3	4.3	5.8	5.1	4.8	3.9	6.0	17.6	38.7	86.2	17.7	40.6	83.0	18.0	38.7	86.2
200	5.0	5.0	4.8	5.2	5.7	5.3	4.3	5.5	5.1	32.9	71.6	100.0	36.1	72.9	99.3	33.8	71.9	99.1
	WB 3: Direct error resampling									WB 3: Direct error resampling								
T	RBP^*_T									RBP^*_T								
50	5.7	5.1	4.9	4.8	5.0	5.1	5.1	4.3	5.9	10.5	17.0	46.7	10.9	21.0	52.1	9.7	15.6	39.0
100	4.3	5.4	5.6	4.4	5.7	4.9	5.0	4.2	6.7	18.0	38.5	85.9	18.8	42.2	85.2	16.6	32.0	77.0
200	4.9	5.1	4.9	5.1	5.9	5.6	4.7	6.2	5.1	32.7	71.4	99.9	37.9	76.0	99.5	30.1	66.9	98.5
	BP^*_T									BP^*_T								
50	5.2	5.0	5.1	5.5	5.3	5.9	5.2	4.3	5.0	10.1	16.9	48.6	10.2	20.4	51.2	11.0	19.2	49.3
100	4.7	5.8	6.1	4.6	5.7	5.3	5.0	3.9	5.8	17.7	38.5	86.0	18.0	40.2	82.9	18.3	38.4	86.0
200	4.9	5.0	4.9	5.3	5.9	5.6	4.3	5.4	5.1	32.9	71.5	99.9	35.8	72.9	99.3	33.3	71.8	99.3

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			t_5			χ^2_6			SN			t_5			χ^2_6		
N	5	10	25	5	10	25	5	10	25	5	10	25	5	10	25	5	10	25
	Asymptotic critical values									Asymptotic critical values								
T	RBP_T									RBP_T								
50	4.9	5.1	8.3	4.1	4.3	6.5	3.5	4.5	7.0	8.7	16.8	53.4	9.4	19.5	54.0	7.8	16.0	43.8
100	4.4	5.5	7.6	4.3	4.5	5.7	3.8	5.3	6.7	15.7	37.1	87.2	16.7	40.9	85.5	15.2	31.1	78.5
200	4.6	4.8	5.5	4.5	5.1	4.9	4.1	5.6	5.5	32.6	71.2	99.8	35.9	74.1	99.5	28.3	65.8	98.7
	BP_T									BP_T								
50	5.5	7.4	15.0	5.4	6.7	14.6	5.2	6.8	14.5	10.7	20.5	63.9	11.0	24.3	62.4	12.0	25.3	63.0
100	4.9	7.9	14.2	5.6	8.6	13.9	5.4	7.3	13.8	18.3	41.9	92.2	19.0	45.4	88.9	20.2	43.2	90.8
200	4.9	7.0	14.5	6.0	7.1	13.6	4.9	7.3	12.2	34.3	75.4	100.0	36.6	76.9	99.6	33.8	75.1	99.9
	WB 1: Recursive resampling									WB 1: Recursive resampling								
T	RBP^*_T									RBP^*_T								
50	5.3	5.0	5.3	5.1	4.3	4.5	4.2	4.3	4.8	9.3	15.7	46.0	11.2	20.1	49.8	9.0	16.5	38.7
100	4.8	5.9	6.1	5.0	5.1	4.9	4.9	5.9	5.6	16.0	36.9	85.1	18.3	41.6	84.3	16.4	31.8	75.8
200	4.8	4.5	4.9	5.1	5.4	4.7	4.4	5.9	5.0	32.3	71.0	99.8	36.3	74.6	99.5	29.3	66.7	98.4
	BP^*_T									BP^*_T								
50	5.8	5.1	6.7	5.0	4.7	6.1	4.3	4.7	5.8	10.0	16.7	49.4	10.0	19.7	50.0	10.2	19.7	50.1
100	4.7	6.0	7.5	5.2	6.0	5.9	4.5	5.7	6.6	17.2	37.6	86.5	17.8	41.0	83.1	18.3	37.9	85.9
200	5.0	5.1	5.7	5.0	5.7	5.9	4.2	5.5	5.2	32.7	70.8	99.8	34.3	72.6	99.1	31.6	71.4	99.7
	WB 2: Fixed design resampling									WB 2: Fixed design resampling								
T	RBP^*_T									RBP^*_T								
50	5.8	4.9	5.4	5.2	4.3	4.8	4.0	4.5	5.1	10.0	15.9	46.5	10.5	20.0	50.4	9.0	16.5	38.9
100	4.7	5.8	6.3	5.2	4.8	5.2	4.5	5.9	5.7	16.6	36.8	85.1	18.3	41.8	83.9	16.1	31.8	75.7
200	4.8	4.8	4.6	5.3	5.5	4.7	4.3	6.1	4.9	32.6	70.6	99.8	36.4	74.9	99.4	29.5	66.9	98.5
	BP^*_T									BP^*_T								
50	5.8	5.3	6.6	5.1	5.1	6.5	4.7	5.0	6.3	10.2	17.0	50.3	10.1	19.2	50.7	10.4	19.8	50.8
100	4.7	6.0	7.6	5.3	6.2	5.8	4.8	5.7	6.8	16.9	37.9	86.0	17.4	41.0	83.4	18.0	38.1	86.0
200	4.9	5.0	5.5	5.2	5.6	5.4	4.3	5.5	5.6	33.4	70.5	99.8	34.1	72.7	99.2	32.0	71.6	99.6
	WB 3: Direct error resampling									WB 3: Direct error resampling								
T	RBP^*_T									RBP^*_T								
50	5.4	5.2	6.4	5.2	5.2	5.6	4.5	5.0	5.9	9.7	16.8	48.5	11.0	20.9	51.8	9.2	16.5	40.9
100	4.9	5.8	6.8	5.3	5.1	5.6	4.9	6.0	6.1	16.3	36.8	86.1	19.0	42.5	84.6	16.7	32.4	77.2
200	4.8	4.8	5.1	5.2	5.7	5.1	4.4	6.2	5.7	32.5	70.9	99.8	36.7	75.3	99.5	29.4	66.8	98.5
	BP^*_T									BP^*_T								
50	6.0	5.3	6.8	5.0	4.8	6.6	4.9	5.1	6.8	10.0	17.3	50.3	10.0	19.8	51.4	10.2	20.7	50.8
100	4.7	5.9	7.1	5.0	6.3	6.0	4.4	5.5	6.8	16.9	36.8	86.1	18.1	41.1	83.6	18.4	38.2	85.9
200	4.7	5.1	5.1	5.2	5.2	5.4	4.3	5.2	5.4	32.5	70.7	99.8	34.0	72.0	99.2	31.7	71.1	99.4

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			t_5			χ^2_6			SN			t_5			χ^2_6		
N	5	10	25	5	10	25	5	10	25	5	10	25	5	10	25	5	10	25
	Asymptotic critical values									Asymptotic critical values								
T	RBP_T									RBP_T								
50	4.9	5.5	7.5	4.3	3.8	6.5	3.7	4.0	6.8	9.6	16.9	53.3	10.2	20.5	55.6	8.3	16.6	43.6
100	4.6	5.5	6.7	3.9	4.5	5.1	4.2	4.0	7.1	17.8	38.8	88.5	18.3	41.7	87.0	15.3	32.6	80.8
200	4.6	5.1	5.2	4.6	4.9	4.9	4.5	5.3	5.9	34.5	74.0	100.0	37.5	76.6	99.6	30.1	68.0	98.9
	BP_T									BP_T								
50	5.2	6.4	9.6	5.5	5.8	8.4	5.0	5.2	8.6	10.9	19.3	58.8	11.8	25.0	59.3	12.1	23.4	58.2
100	5.0	6.2	7.5	5.1	5.9	7.0	4.8	5.1	7.5	19.1	40.9	89.2	19.7	44.5	86.8	20.5	41.9	88.7
200	5.0	5.2	5.5	6.3	5.9	5.9	4.4	5.5	5.7	36.1	75.2	99.9	37.6	76.7	99.4	34.5	73.9	99.5
	WB 1: Recursive resampling									WB 1: Recursive resampling								
T	RBP^*_T									RBP^*_T								
50	5.1	5.2	5.1	5.5	4.7	4.8	4.2	4.2	4.5	10.3	16.2	47.6	11.3	20.9	51.2	9.5	16.6	39.4
100	4.8	5.6	5.8	4.3	5.0	4.9	5.0	4.4	6.6	18.1	38.7	87.0	19.8	43.3	86.2	16.8	33.3	79.4
200	4.9	4.7	4.9	5.1	5.8	4.9	4.8	5.6	5.5	33.6	73.5	99.9	38.0	77.2	99.6	31.0	68.9	98.9
	BP^*_T									BP^*_T								
50	5.1	5.5	4.8	5.0	4.5	4.8	4.5	4.3	4.0	10.5	16.0	48.2	11.3	21.1	49.6	10.8	19.3	48.2
100	4.5	5.6	5.1	5.0	5.7	5.0	4.8	4.4	5.8	18.5	38.9	86.8	18.5	41.4	84.2	19.4	38.8	86.0
200	4.7	5.0	5.1	4.9	5.7	4.7	4.1	5.1	4.8	34.6	73.9	99.9	36.8	74.6	99.2	33.5	73.7	99.5
	WB 2: Fixed design resampling									WB 2: Fixed design resampling								
T	RBP^*_T									RBP^*_T								
50	5.3	5.6	5.6	5.2	4.5	5.5	4.4	4.7	5.0	10.3	16.6	47.7	11.8	21.2	51.8	9.7	16.7	39.2
100	4.7	5.5	5.5	4.8	5.2	4.8	5.1	4.5	6.3	18.4	38.7	87.1	19.5	43.1	86.1	17.1	32.5	79.2
200	4.8	4.7	4.5	4.9	5.7	5.0	5.0	5.4	5.3	34.4	73.3	100.0	38.0	77.2	99.6	31.0	69.1	98.9
	BP^*_T									BP^*_T								
50	5.4	5.7	5.0	5.3	4.3	5.1	4.4	4.4	4.7	9.9	16.3	49.3	11.0	21.0	50.3	10.8	19.2	49.3
100	4.5	5.5	5.5	4.8	5.2	5.0	4.9	4.7	5.9	18.6	38.7	87.0	18.3	41.8	84.2	18.8	39.3	86.4
200	4.5	5.0	5.1	5.2	5.3	4.6	4.1	5.0	4.8	34.3	73.9	99.9	37.0	74.7	99.3	33.6	73.8	99.5
	WB 3: Direct error resampling									WB 3: Direct error resampling								
T	RBP^*_T									RBP^*_T								
50	5.6	5.8	5.5	5.2	4.7	5.6	4.5	4.5	5.1	10.2	16.6	48.8	11.9	21.7	53.0	9.7	17.1	39.9
100	4.5	5.8	5.7	4.9	5.2	5.0	5.0	4.5	6.7	18.4	38.9	87.4	20.0	43.5	86.3	17.2	33.4	79.4
200	5.0	4.6	4.7	5.4	5.8	4.9	5.0	5.7	5.7	34.0	73.4	99.9	38.0	77.5	99.5	31.1	68.8	98.9
	BP^*_T									BP^*_T								
50	4.8	5.3	5.6	5.3	4.5	5.8	4.9	4.5	4.9	10.0	16.6	50.0	11.4	21.6	51.5	10.8	20.1	50.5
100	4.6	5.6	5.4	4.9	5.3	5.1	4.6	4.6	6.2	18.7	39.1	87.5	18.9	41.6	84.6	19.6	39.4	86.5
200	4.9	5.0	5.0	5.1	5.7	4.9	4.0	5.1	5.1	34.6	74.1	99.9	37.2	75.0	99.3	34.0	73.7	99.5

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	RBP_T	BP_T
asymptotic	0.092	0.022*
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 χ^2_{190} distribution. 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.