HONG KONG INSTITUTE FOR MONETARY RESEARCH

STABILITY TESTS FOR HETEROGENEOUS PANEL DATA

Felix Chan, Tommaso Mancini-Griffoli and Laurent L. Pauwels

HKIMR Working Paper No.09/2008

May 2008

Stability Tests for Heterogeneous Panel Data

Felix Chan*
School of Economics and Finance
Curtin University of Technology

and

Tommaso Mancini-Griffoli*
Swiss National Bank
Paris School of Economics (PSE)
CEPREMAP

and

Laurent L. Pauwels*†
Hong Kong Monetary Authority
Graduate Institute of International Studies, Geneva

May 2008

Abstract

This paper proposes a new test for structural stability in panels by extending the testing procedure proposed in the seminal work of Andrews (2003) originally developed for time series. The test is robust to non-normal, heteroskedastic and serially correlated errors, and, importantly, allows for the number of post break observations to be small. Moreover, the test accommodates the possibility of a break affecting only some - and not all - individuals of the panel. Under mild assumptions the test statistic is shown to be asymptotically normal, thanks to the cross sectional dimension of panel data. This greatly facilitates the calculation of critical values with respect to the test's time series counterpart. Monte Carlo experiments show that the test has good size and power under a wide range of circumstances. Finally, the test is illustrated in practice, in a brief study of the euro's effect on trade.

Keywords: Structural Change, Instability, Cross Sectionally Dependent Errors, Heterogeneous Panels, Monte Carlo, Euro Effect on Trade.
JEL Classification: C23, C52

---

* The authors express their sincere thanks and gratitude for constructive comments to Jushan Bai, In Choi, Hans Genberg, Jaya Krishnakumar, Michael McAleer, Hashem Pesaran, Michael Salemi and Charles Wyplosz, as well as participants at the Third Symposium on Econometric Theory and Applications (SETA), the 14th International Conference on Panel Data, 2007 International Congress on Modelling and Simulation and seminars at HKUST and HKIMR. This paper was written while Laurent L. Pauwels was a visiting fellow at the School of Economics and Finance at Curtin University of Technology, Perth, Australia. The hospitality of the school is gracefully acknowledged. The views expressed in this paper are those of the authors, and do not necessarily reflect those of the Hong Kong Monetary Authority or Swiss National Bank.

† Corresponding author: Laurent Pauwels, Hong Kong Monetary Authority, Research Department, 55/F Two International Finance Centre, 8 Finance Street, Central, Hong Kong. Tel: +852 28781664, Fax: +852 28781897, Email: lpauwels@hkma.gov.hk.

1. Introduction

This paper proposes a new test for structural instability among only some individuals in a panel regression model, and allows for this instability to occur at the very end of a sample. Most tests for structural breaks were developed specifically for time-series, like the popular Chow (1960) tests, and those for unknown or multiple break dates in Andrews (1993), Andrews and Ploberger (1994) and Bai and Perron (1998). The distribution of the corresponding test statistic is suitably found using asymptotics in which the number of observations before and after the break point go to infinity. However, it is often at the end of a sample that researchers and policy-makers alike are interested in testing for instability.

Andrews (2003) proposes a test for structural break specifically designed for few post-break observations. Monte Carlo results suggest that the test has reasonable size and power even when the number of post-break observation is 1. Unlike the well known Predictive Failure test of Chow (1960), the critical values of Andrews' (2003) test statistic are calculated using parametric sub-sampling methods making the test robust to non-normal, heteroskedastic and serially correlated errors. The extension of the test to panel data, under the assumption of cross sectional independence, is relatively straightforward as shown in Mancini-Griffoli and Pauwels (2006). This extension assumes the alternative hypothesis that all individuals exhibit a break, as in other relevant tests in the panel literature, like in Han and Park (1989) which extends the CUSUM tests, or Emerson and Kao (2001, 2002), Kao et al. (2005) and De Wachter and Tzavalis (2004) which build on Andrews (1993) and Andrews and Ploberger (1994). Yet, this approach does not address the interesting alternative that only some – and not all – individuals are affected by a break. This is the more general question, but also likely to be the more prominent in applied work, as shocks rarely affect all individuals equally, if at all. This is the question addressed by this paper.

This paper proposes a test for heterogeneous breaks in panels based on the Andrews (2003) end of sample stability test. In particular, this paper introduces a standardized Z statistic calculated by taking a weighted average of Andrews's (2003) statistics for each individual. Methodologically, this is similar to the approach in Im et al. (2003) which, while focusing on the different question of unit root tests, also considers an average of separate statistics. This paper derives the asymptotic distribution of the proposed test statistic using the Lindeberg-Feller Central Limit Theorem (LF-CLT). The test statistic is shown to follow a normal distribution as the number of individuals goes to infinity. This greatly simplifies the computation of the critical values with respect to Andrews (2003). As in Andrews (2003), though, the proposed statistic is robust to non-normal, heteroskedastic, serially correlated errors and when the instability occurs at the end of a given sample. In addition, the test allows for parameter heterogeneity pre- and post-instability.

Although the asymptotic results are derived under the assumption of cross sectional independence, this does not severely restrict the applicability of the test. The asymptotic results still hold in the case of cross sectional dependence as long as it can be "filtered out" using appropriate estimators. This paper provides an example of how this can be accomplished by modifying the proposed test statistics using the Common Correlated Effects (CCE) estimator proposed in Pesaran (2006).

A series of Monte Carlo experiments show that the proposed structural break test performs very well in finite samples. The experiments accommodate serial correlation in the errors with a mixture of different distributions for the innovations. The results show that even under these circumstances the distribution of the test is close to a standardised normal. Furthermore, Monte Carlo results indicate that the test has good size and power with relatively few observations over time and moderate serial correlation within cross sections. For high levels of serial correlation, the performance of the test improves as the sample size increases. Lastly, the test has good power and size even when instabilities are of a small magnitude.

Finally, this paper considers an empirical application of the test, to demonstrate its usefulness in a real-world setting: did the introduction of the euro increase intra-Eurozone trade? The question has been at the center of lively debates in academic and policy circles alike. However, the papers that have tackled the issue have not provided strong empirical evidence in support of the presumed effect. This is largely due to two empirical issues: the few datapoints available after the euro's introduction and the heterogeneity of the trade effect over different countries. Given both of these characteristics, the test introduced in this paper is particularly well suited. Results show a break at the 10% significance level in Eurozone trade starting in 1999, thereby supporting the presumption commonly expressed in the literature.

The paper is organised as follows. Section 2 introduces the panel data stability test for heterogeneous breaks. A solution to the issue of cross sectional dependence is also discussed. Section 3 follows with a derivation of relevant asymptotic results. Section 4 investigates the test's finite sample properties with Monte Carlo simulations. And finally, section 5 illustrates how the test can be put to use to answer the question of the euro's effect on intra-Eurozone trade.

2. Heterogeneous Panel Data Stability Tests

2.1 Setup

Consider the following baseline model for panel data,

$$y_{it} = \mathbf{\Theta}_i^{\prime}\mathbf{x}_{it} + u_{it} \tag{1}$$

$$u_{it} = \gamma_i^{\prime}\mathbf{f}_t + \varepsilon_{it} \quad i = 1, \dots, N, \quad t = 1, \dots, T \tag{2}$$

where y_{it} is the dependent variable, $\mathbf{x}_{it} = (x_{it}^{(1)}, ..., x_{it}^{(d)})'$ is the d x 1 vector of explanatory variables including intercepts and/or seasonal dummies, $\mathbf{\Theta}_i$ is the d x 1 vector of coefficients. Moreover, ε_{it} are the idiosyncratic shocks specific to each individual and assumed to be uncorrelated to $\mathbf{x}_{it}$ and have zero mean, $\mathbf{f}_t$ is the l x 1 vector of unobserved common effects and γ_i are the factor loadings associated with $\mathbf{f}_t$. For the purposes of deriving the test statistics, γ_i is assumed to be 0 for all i = 1, ..., N. The more practically relevant case in which $\gamma_i \neq 0$ is discussed in Section 2.4.

Under equation (1) with $\gamma_i = 0$, the hypothesis of structural instability implies,

$$y_{it} = \begin{cases} \mathbf{\Theta}_{0i}'\mathbf{x}_{it} + u_{it} & t = 1, \ldots, T_i \\ (\mathbf{\Theta}_{0i}' + \delta_i')\mathbf{x}_{it} + u_{it} & t = T_i + 1, \ldots, T_i + m_i \end{cases} \tag{3}$$

for $i = 1, \ldots, N$, and where T_i are the presumed break dates, which can differ for each individual i. Thus, m_i, the number of post-break observations, can be different for each i. $\mathbf{\Theta}_{0i}$ is the parameter vector before the break and δ_i is the difference between the post- and pre-break coefficient vectors. Thus, the hypothesis of structural stability is,

$$H_0 : \delta_i = 0 \quad \forall\, i = 1, \ldots, N \text{ and } \forall\, t$$
$$H_1 : \delta_i \neq 0 \quad \exists\, i \in \{1, \ldots, N\} \text{ and for } t > T_i$$

with $t = 1, \ldots, T_i, T_i + 1, \ldots, T_i + m_i$. $\mathbf{\Theta}_{0i}$ can be estimated heterogeneously for each individual i by OLS. In this case, the consistency of $\mathbf{\Theta}_{0i}$ relies on large T only for all i, whereas if $\mathbf{\Theta}$ is homogeneous, its consistency can rely on either large N or large T, as is standard in the panel literature.

Let $N = N_0 + N_1$, where N_0 is the number of individuals for whom $\delta_i = 0$ and N_1 is the number of individuals that exhibit a break ($\delta_i \neq 0$). The null hypothesis states that there are no structural breaks across all N individuals, whereas the alternative states that a proportion of individuals experience a structural break. The alternative requires that the proportion of individuals who experience a break relative to N tends to a non-zero positive constant as $N \to \infty$. Mathematically, this implies $\lim_{N\to\infty}(\frac{N_1}{N}) = c$, where $0 < c \leq 1$ as introduced in Choi (2001) and used again in Im et al. (2003). This assumption ensures the asymptotic validity of the test.

When the null hypothesis of structural stability is rejected, the exact proportion of individuals who experience a break can be found by conducting the Andrews (2003) test on each individual separately. However, conducting multiple Andrews tests is not a good replacement for the panel test proposed in this paper for at least two reasons: (i) the computation cost of conducting multiple Andrews (2003) tests is extremely high relative to the panel test, especially when N is large. This is because critical values in the Andrews (2003) test are calculated by constructing empirical distributions; (ii) if $\Theta_{0i} = \Theta_0$ for all $i = 1, .., N$ prior to the break, that is, if the panel is homogeneous before the break, then a panel estimation benefitting from cross sectional variation yields more precise estimation results. This is particularly important if the panel has small T and large N.

2.2 Motivating and Defining the Individual Statistics

The proposed statistic, to test for heterogeneous instability in panel data models, essentially amounts to comparing two average statistics taken from both the pre-break and post-break samples. These averages are based on test statistics for each individual in the panel, computed as in Andrews (2003). The section below first motivates these statistics, then defines them explicitly.

Let $\mathbf{Y}_{i,p}^{q} = (y_{i,p}, y_{i,p+1}, ..., y_{i,q})'$ be a $(q-p+1)\times 1$ vector and

$$\mathbf{X}_{i,p}^{q} = \begin{pmatrix} x_{i,p}^{(1)} & x_{i,p}^{(2)} & ... & x_{i,p}^{(d)} \\ x_{i,p+1}^{(1)} & x_{i,p+1}^{(2)} & ... & x_{i,p+1}^{(d)} \\ \vdots & \vdots & \dots & \vdots \\ x_{i,q}^{(1)} & x_{i,q}^{(2)} & ... & x_{i,q}^{(d)} \end{pmatrix}$$

be a $(q-p+1)\times d$ matrix such that $p, q \in \mathbb{Z}^+$ with $q > p$, $\forall i$ = 1, ..., N. $\mathbf{Y}_{i,p}^{q}$ contains the values of the endogenous variable, y, and $\mathbf{X}_{i,p}^{q}$ contains the values of all the explanatory variables for the individual i over the sample period spanning from p to q. Therefore, equation (3) can be rewritten in terms of the data as

$$\mathbf{Y}_{i,1}^{\bar{T}} = \mathbf{X}_{i,1}^{\bar{T}}\mathbf{\Theta}_{0i} + \begin{pmatrix} 0_{T_i\times d} \\ \mathbf{X}_{i,T_i+1}^{\bar{T}} \end{pmatrix}\delta_i + \mathbf{U}_{i,1}^{\bar{T}} \tag{4}$$

where $\bar{T} = T_i{+}m_i$, $0_{T_i\times d}$ is a $T_i{\times}d$ null matrix and $\mathbf{U}_{i,p}^{q} = (u_{i,p}, u_{i,p+1}..., u_{i,q})'$ is a (q - p + 1) x 1 vector containing the residuals for the individual, i, over the sample period spanning from p to q.

From equation (4), it is clear that the OLS estimator for δ_i is

$$\hat{\delta}_i = \left[(\mathbf{X}_{i,T_i+1}^{\bar{T}})'\mathbf{X}_{i,T_i+1}^{\bar{T}}\right]^{-1}(\mathbf{X}_{i,T_i+1}^{\bar{T}})'(\mathbf{Y}_{i,T_i+1}^{\bar{T}} - \mathbf{X}_{i,T_i+1}^{\bar{T}}\mathbf{\Theta}_{0i}) \tag{5}$$

and therefore the estimated residual for the post-break observations can be calculated as

$$\begin{aligned}\mathbf{U}_{i,T_i+1}^{\bar{T}} &= (\mathbf{Y}_{i,T_i+1}^{\bar{T}} - \mathbf{X}_{i,T_i+1}^{\bar{T}}\mathbf{\Theta}_{0i}) - \mathbf{X}_{i,T_i+1}^{\bar{T}}\hat{\delta}_i \\ &= (I_{m_i} - \mathbf{P}_{\mathbf{X}_{i,T_i+1}^{\bar{T}}})(\mathbf{Y}_{i,T_i+1}^{\bar{T}} - \mathbf{X}_{i,T_i+1}^{\bar{T}}\mathbf{\Theta}_{0i})\end{aligned}$$

where I_{m_i} is the m_i x m_i identity matrix and

$$\mathbf{P}_{\mathbf{X}_{i,T_i+1}^{\bar{T}}} = \mathbf{X}_{i,T_i+1}^{\bar{T}}\left[\left(\mathbf{X}_{i,T_i+1}^{\bar{T}}\right)'\mathbf{X}_{i,T_i+1}^{\bar{T}}\right]^{-1}\left(\mathbf{X}_{i,T_i+1}^{\bar{T}}\right)'$$

is the well known projection matrix. Therefore, the (unrestricted) sum squares residuals, $SSR_{i,U} = (\hat{\mathbf{U}}_{i,T_i+1}^{\bar{T}})'\hat{\mathbf{U}}_{i,T_i+1}^{\bar{T}}$, can be written as

$$\left(\mathbf{Y}_{i,T_i+1}^{\bar{T}} - \mathbf{X}_{i,T_i+1}^{\bar{T}}\hat{\mathbf{\Theta}}_{i,1}^{T_i}\right)'\left(I_{m_i} - \mathbf{P}_{\mathbf{X}_{i,T_i+1}^{\bar{T}}}\right)\left(\mathbf{Y}_{i,T_i+1}^{\bar{T}} - \mathbf{X}_{i,T_i+1}^{\bar{T}}\hat{\mathbf{\Theta}}_{i,1}^{T_i}\right) \tag{6}$$

where $\hat{\mathbf{\Theta}}_{i,1}^{T_i}$ is the least squares estimate of $\mathbf{\Theta}_{0i}$ using the sample spanning over the pre-break sample from 1 to T_i , $\forall i$ = 1, ..., N. Under the null hypothesis (δ_i = 0), the (restricted) sum squares residuals for the post-break period is defined to be

$$SSR_{i,R} = \left(\mathbf{Y}_{i,T_i+1}^{\bar{T}} - \mathbf{X}_{i,T_i+1}^{\bar{T}}\hat{\mathbf{\Theta}}_{i,1}^{T_i}\right)'\left(\mathbf{Y}_{i,T_i+1}^{\bar{T}} - \mathbf{X}_{i,T_i+1}^{\bar{T}}\hat{\mathbf{\Theta}}_{i,1}^{T_i}\right) \tag{7}$$

Obviously, equation (6) collapses to equation (7) under the null hypothesis, as δ_i = 0 for all i. Therefore, if the null is true, then the test statistics,

$$\begin{aligned} S_{i,T_i+1}^{\bar{T}} &= SSR_{i,R} - SSR_{i,U} \\ &= \left(\mathbf{Y}_{i,T_i+1}^{\bar{T}} - \mathbf{X}_{i,T_i+1}^{\bar{T}}\hat{\mathbf{\Theta}}_{i,1}^{T_i}\right)' \mathbf{P}_{\mathbf{X}_{i,T_i+1}^{\bar{T}}} \left(\mathbf{Y}_{i,T_i+1}^{\bar{T}} - \mathbf{X}_{i,T_i+1}^{\bar{T}}\hat{\mathbf{\Theta}}_{i,1}^{T_i}\right) \end{aligned}$$

would be centered around 0. Thus, the further is $S_{i,T_i+1}^{\bar{T}}$ from 0, the more evidence there is to reject the null of structural stability, $\forall i$ = 1, ..., N. The power of the test can be increased by introducing a positive definite weighting matrix, $\mathbf{\Sigma}_i^{-1/2}$, such that $Var(\mathbf{U}_i) = \mathbf{\Sigma}_i$. In this case, the $S_{i,T_i+1}^{\bar{T}}$ statistic can be written as

$$\left(\mathbf{Y}_{i,T_i+1}^{\bar{T}} - \mathbf{X}_{i,T_i+1}^{\bar{T}}\hat{\mathbf{\Theta}}_{i,1}^{T_i}\right)' \mathbf{\Sigma}_i^{-1} \bar{\mathbf{P}}_{\mathbf{X}_{i,T_i+1}^{\bar{T}}} \mathbf{\Sigma}_i^{-1} \left(\mathbf{Y}_{i,T_i+1}^{\bar{T}} - \mathbf{X}_{i,T_i+1}^{\bar{T}}\hat{\mathbf{\Theta}}_{i,1}^{T_i}\right) \quad (8)$$

where $\bar{\mathbf{P}}_{\mathbf{X}_i} = \mathbf{X}_i(\mathbf{X}_i'\mathbf{\Sigma}_i^{-1}\mathbf{X}_i)^{-1}\mathbf{X}_i'$

Following the intuition above, the fundamental test statistic for each individual i is defined, as in Andrews (2003), to be

$$S_{i,p}^{q}(\mathbf{\Theta}_i, \mathbf{\Sigma}_i) = A_{i,p}^{q}(\mathbf{\Theta}_i, \mathbf{\Sigma}_i)' [V_{i,p}^{q}(\mathbf{\Sigma}_i)]^{-1} A_{i,p}^{q}(\mathbf{\Theta}_i, \mathbf{\Sigma}_i) \quad (9)$$

$$A_{i,p}^{q}(\mathbf{\Theta}_i, \mathbf{\Sigma}_i) = \mathbf{X}_{i,p}^{\prime\, q}\, \mathbf{\Sigma}_i^{-1} \left(\mathbf{Y}_{i,p}^{q} - \mathbf{X}_{i,p}^{q}\mathbf{\Theta}_i\right) \quad (10)$$

$$V_{i,p}^{q}(\mathbf{\Sigma}_i) = \mathbf{X}_{i,p}^{\prime\, q}\, \mathbf{\Sigma}_i^{-1}\, \mathbf{X}_{i,p}^{q} \quad (11)$$

for all i = 1, ..., N.

There are two specific variants of $S_{i,p}^{q}$ that are used in calculating the standardised Z statistic essential to this paper:

$$S_i^0 = S_{i,1}^{m_i}\left(\widehat{\mathbf{\Theta}}_{i,1}^{T_i}, \widehat{\mathbf{\Sigma}}_{i,1}^{\bar{T}}\right) \quad (12)$$

$$S_i^1 = S_{i,T_i+1}^{\bar{T}}\left(\widehat{\mathbf{\Theta}}_{i,1}^{\bar{T}}, \widehat{\mathbf{\Sigma}}_{i,1}^{\bar{T}}\right) \quad (13)$$

Both sets of statistics are computed using m_i observations. The post-break sample statistics, S_i^1 , are computed for the sample spanning from p = T_i + 1 to q = $\bar{T}$, whereas the pre-break sample statistics, S_i^0, are calculated over m_i observations anywhere in the pre-break sample so long as $q \leq T_i$.

The estimated time-series covariance matrix derived in Andrews (2003) is used as a weight matrix for each individual i. The covariance matrix is

$$\widehat{\mathbf{\Sigma}}_{i,1}^{\bar{T}} = (T_i+1)^{-1}\sum_{r=1}^{T_i+1}\left(\widehat{\mathbf{U}}_{i,r}^{r+m_i-1}\widehat{\mathbf{U}}_{i,r}^{\prime\, r+m_i-1}\right)$$

and $\widehat{\mathbf{U}}_{i,r}^{\,r+m_i-1}$ is individual i's m_i x 1 estimated residual vector resulting from the i^{th} time-series regression

$$\widehat{\mathbf{U}}_{i,r}^{\,r+m_i-1} = \left(\mathbf{Y}_{i,r}^{\,r+m_i-1} - \mathbf{X}_{i,r}^{\,r+m_i-1}\widehat{\mathbf{\Theta}}_{i,1}^{\,\bar{T}}\right)$$

The coefficient vector $\widehat{\mathbf{\Theta}}_{i,1}^{\,\bar{T}}$ is the least square estimates of $\mathbf{\Theta}$ for individual i over the full temporal sample.

If $m_i \leq d$, the $S_{i,p}^{\,q}(\mathbf{\Theta}_i, \mathbf{\Sigma}_i)$ statistic simplifies to the following:

$$P_{i,p}^{\,q}(\mathbf{\Theta}_i, \mathbf{\Sigma}_i) = \left(\mathbf{Y}_{i,p}^{\,q} - \mathbf{X}_{i,p}^{\,q}\mathbf{\Theta}_i\right)' \ \mathbf{\Sigma}_i^{-1} \ \left(\mathbf{Y}_{i,p}^{\,q} - \mathbf{X}_{i,p}^{\,q}\mathbf{\Theta}_i\right) \tag{14}$$

where details are given in Andrews (2003).

2.3 The Z Statistic

This paper defines the Z statistic to test for heterogeneous breaks in panels as follows

$$Z = \frac{\left(\bar{S}^1 - \bar{S}^0\right)}{\sqrt{\widehat{Var}\left(\bar{S}^1 - \bar{S}^0\right)}} \tag{15}$$

where $\bar{S}^\nu = N^{-1}\sum_{i=1}^{N} S_i^\nu,\ \nu = 0, 1$ are the average statistics for the pre- and post-break sample respectively. Intuitively, if the null hypothesis were true, Z would be centered around 0. Therefore, the further from 0 is the Z statistics, the more evidence there is to reject the null hypothesis in favor of the alternative. Since the variances of the individual statistics are unknown, we use the estimated variance of the difference of the average statistics.[1]

A point of practical importance is waranted. It is recommended to use the first, or earliest possible, m_i observations to estimate $\bar{S}^0$ in order to maximise the distance between the subsamples and thereby minimise the potential impact of serial correlation in the errors. This is essentially an empirical issue, and the problem of serial correlation should diminish as the gap between the two subsamples increases, as implied by ergodicity. The computation of the S_i statistics can be simplified in the case of a homogeneous panel where Θ_0 can be estimated once to construct the test and does not need to be estimated for each cross section.

1 Let $\widehat{Var}\left(\bar{S}^1 - \bar{S}^0\right)$ be the consistent estimate of $Var\left(\bar{S}^1 - \bar{S}^0\right)$ such that $\widehat{Var}\left(\bar{S}^1 - \bar{S}^0\right) \xrightarrow{p} Var\left(\bar{S}^1 - \bar{S}^0\right)$. The estimated variance can be simplified to $\widehat{Var}\left(\bar{S}^1 - \bar{S}^0\right) = N^{-1}\sum_{i=1}^{N}\left[(S_i^1)^2 + (S_i^0)^2\right]$, $\because\ E[\bar{S}^1] = E[\bar{S}^0]$ under H_0, and $\because\ Cov(\bar{S}^0, \bar{S}^1)$ = 0, The $Cov(\bar{S}^0, \bar{S}^1)$ = 0 comes from the independence of the individual statistics. All that is required for a consistent estimate of the variance is $N^{-1}\sum_{i=1}^{N}(S_i^\nu)^2 \xrightarrow{p} E[(\bar{S}^\nu)^2]$.

2.4 Cross Sectional Dependence

The test derived in the previous sub-section, though robust to serial correlation and heteroskedasticity, assumes cross sectional independence. If this assumption is not valid, appropriate estimators can be used to "filter out" the cross sectional dependence. The recent panel data literature has proposed several such solutions. Since the focus of the paper is not related to cross sectional dependence, this sub-section only provides an illustration of how this paper's results can be extended to allow for cross sectional dependence. The bottom line is that the asymptotic results supporting this paper's panel test (presented in the next section) still hold if it is possible to obtain consistent estimates of $\mathbf{\Theta}_i$ and ε_{it} $\forall i, t$.

The Common Correlated Effect (CCE) estimator proposed by Pesaran (2006) is particularly convenient for this paper's purposes and the asymptotic results derived in the following section continue to hold with minimal modifications to the assumptions.[2] Although this section provides the CCE estimator as an example, other estimators such as the Principal Component Estimator proposed in Coakley et al. (2002), Bai and Ng (2002) and Bai (2005) can also be used with suitable modifications.

The CCE estimator is defined to be

$$\tilde{\mathbf{\Theta}}_{i,p}^{\ q} = \left[\left(\mathbf{X}_{i,p}^{\ q}\right)' M_{w,p}^{\ q} \mathbf{X}_{i,p}^{\ q}\right]^{-1} \left(\mathbf{X}_{i,p}^{\ q}\right)' M_{w,p}^{\ q} \mathbf{Y}_{i,p}^{\ q} \tag{16}$$

where $M_{w,p}^{\ q} = (I_{(q-p+1)\times 1} - w_p^q[(w_p^q)'(w_p^q)]^{-1}(w_p^q)')$, with $w_p^q = (\bar{\mathbf{y}}_{t,p}^{\ q}, \bar{\mathbf{x}}_{t,p}^{\ q})$ a $(q-p+1)\times(d+1)$ matrix, such that $\bar{\mathbf{y}}_{t,p}^{\ q}$ is a $(q-p+1)\times 1$ vector containing the cross sectional averages of the endogenous variable from the sample spanning $t = p$, ..., q and $\bar{\mathbf{x}}_{t,p}^{\ q} = (\bar{x}_{t,p}^{(1)q}, \bar{x}_{t,p}^{(2)q}, ..., \bar{x}_{t,p}^{(d)q})$ is a $(q-p+1)\times d$ matrix consisting of d columns of $(q-p+1)\times 1$ vectors, with $\bar{x}_{t,p}^{(j)q}$ containing the cross sectional averages of the jth explanatory variable from the sample spanning $t = p$, ..., q.

The idea of the CCE estimator is to use the cross sectional averages of the endogenous and explanatory variables as proxies for the common factors, f_t. With this, the effect from the common factors can be "filtered out" using the residual maker, M_w. Pesaran (2006) shows that such a proxy is consistent under certain regularity conditions. Therefore, it is possible to replace the OLS estimator with the CCE estimator in the proposed test in the presence of cross sectional dependence. The asymptotic properties of the modified test will be discussed more carefully in the next section. However, the additional assumption that $T_i = T$ for all i = 1, ..., N is required in order to adopt the CCE estimator in the proposed test statistic. This assumption restricts all the individuals to share the same break date. This is necessary given the construction of w_p^q which contains the cross sectional averages from both the pre-break sample and the post-break sample. Without a common break date, it would be unclear how to compute these averages across individuals and whether the consistency results from Pesaran (2006) would hold.

[2] f_t is defined to include both observable and unobservable common effects in this paper, whereas f_t is defined to be the unobservable common effect only in Pesaran (2006).

With the CCE estimator, as defined in equation (16), the basic test statistic can be rewritten as follows:

$$\tilde{S}_{i,p}^{\ q}\big(\mathbf{\Theta}_i, \Sigma_i\big) = \tilde{A}_{i,p}^{\ q}\big(\mathbf{\Theta}_i, \Sigma_i\big)'\Big[\tilde{V}_{i,p}^{\ q}\big(\Sigma_i\big)\Big]^{-1}\tilde{A}_{i,p}^{\ q}\big(\mathbf{\Theta}_i, \Sigma_i\big) \tag{17}$$

$$\tilde{A}_{i,p}^{\ q}\big(\mathbf{\Theta}_i, \Sigma_i\big) = \mathbf{X}_{i,p}^{\ q} M_{w,p}^{\ q} \Sigma_i^{-1} M_{w,p}^{\ q}\big(\mathbf{Y}_{i,p}^{\ q} - \mathbf{X}_{i,p}^{\ q}\mathbf{\Theta}_i\big) \tag{18}$$

$$\tilde{V}_{i,p}^{\ q}\big(\Sigma_i\big) = \big(\mathbf{X}_{i,p}^{\ q}\big)' M_{w,p}^{\ q} \Sigma_i^{-1} M_{w,p}^{\ q} \mathbf{X}_{i,p}^{\ q} \tag{19}$$

and

$$\tilde{S}_i^0 = \tilde{S}_{i,1}^{\ m}\big(\tilde{\mathbf{\Theta}}_{i,1}^{\ \bar{T}}, \tilde{\Sigma}_{i,1}^{\ T}\big) \tag{20}$$

$$\tilde{S}_i^1 = \tilde{S}_{i,T+1}^{\ \bar{T}}\big(\tilde{\mathbf{\Theta}}_{i,1}^{\ \bar{T}}, \tilde{\Sigma}_{i,1}^{\ \bar{T}}\big) \tag{21}$$

Likewise, $\hat{\Sigma}_{i,1}^{\ \bar{T}}$ can be computed following the same procedure as $\hat{\Sigma}_{i,1}^{\ \bar{T}}$ in section 2, that is,

$$\tilde{\Sigma}_{i,1}^{\ \bar{T}} = (T+1)^{-1}\sum_{r=1}^{T+1}\bigg(\Lambda_{i,r}^{\ r+m-1}\Lambda_{i,r}^{\prime\ r+m-1}\bigg)$$

with

$$\Lambda_{i,r}^{\ r+m-1} = M_{w,1}^{\ \bar{T}}\bigg(\mathbf{Y}_{i,r}^{\ r+m-1} - \mathbf{X}_{i,r}^{\ r+m-1}\tilde{\mathbf{\Theta}}_{i,1}^{\ \bar{T}}\bigg)$$

Moreover, the average test statistics become,

$$\bar{\tilde{S}}^0 = N^{-1}\sum_{i=1}^{N}\tilde{S}_i^0 \tag{22}$$

$$\bar{\tilde{S}}^1 = N^{-1}\sum_{i=1}^{N}\tilde{S}_i^1 \tag{23}$$

Lastly, the test of structural stability using the CCE estimator can be defined as

$$\tilde{Z} = \frac{\bar{\tilde{S}}^1 - \bar{\tilde{S}}^0}{\sqrt{\widehat{Var}(\bar{\tilde{S}}^1 - \bar{\tilde{S}}^0)}} \tag{24}$$

3. Asymptotic Results

3.1 Assumptions

This section provides the asymptotic properties of the proposed test. Define the data set as the outcomes of a sequence of random variables $\{\mathbf{W}_{0,it}\}$ where $\{\mathbf{Y}_i, \mathbf{X}_i\} \subset \{\mathbf{W}_{0,it}\}$. Under H_0, the data are $\mathbf{W}_{0,it}$ for $t = 1, \ldots, T_i + m_i$ and $i \in \{1, ..., N\}$,while under H_1 the data are $\mathbf{W}_{it} = \mathbf{W}_{0,it}$ for $t = 1, \ldots, T_i$ and $\mathbf{W}_{it} = \mathbf{W}_{T_i,it}$ for $t = T_i, \ldots, T_i + m_i$, where $\{\mathbf{W}_{T_i,it} : t = T_i, \ldots, T_i + m_i\}$ are some random variables with a joint distribution different from $\{\mathbf{W}_{0,it} : t = T_i, \ldots, T_i + m_i\}$. Assume also that the distribution of $\{\mathbf{W}_{0,it} : t = T_i, \ldots, T_i + m_i\}$ is independent of T_i. Note that under H_1 the data are from a triangular array since the breakpoint is changing with $T \to \infty$.

Let $B(\mathbf{\Theta}_{0i}, \epsilon_T)$ be a ball centered around $\mathbf{\Theta}_{0i}$ with radius $\epsilon_T > 0$ as in Andrews (2003). For $m_i > d$, the following assumptions underlying the asymptotic properties of S_i^ν, $\nu = 0$, 1 are:

Assumption 1

$\{\mathbf{W}_{0,it} : t \geq 1\}$ $\forall$ i, *is stationary and ergodic.*

Assumption 2

(a) $\|\widehat{\mathbf{\Theta}}_{i,1}^{T_i+m_i} - \mathbf{\Theta}_{0i}\| \xrightarrow{p} 0$, *with* $(T_i, N) \to \infty$ *with* m_i *fixed under* H_0 *and* H_1.

(b) $\sup_{\mathbf{\Theta} \in B(\mathbf{\Theta}_{0i}, \epsilon_T)} \|\widehat{\mathbf{\Sigma}}_{i,1}^{T_i+m_i} - \mathbf{\Sigma}_{0i}\| \xrightarrow{p} 0$ *with* $(T_i, N) \to \infty$ *with* m_i *fixed under* H_0 *and* H_1, *for some nonsingular matrix* $\mathbf{\Sigma}_{0i}$, *for all sequences of constant* $\{\epsilon_{T,N} : T_i \geq 1, N \geq 1\}$ *and* $\epsilon_{T,N} \to 0$ *as* $(T_i, N) \to \infty$.

Assumption 3

(a) $S_i^\nu(\mathbf{\Theta}_i, \mathbf{\Sigma}_i)$, $\nu = 0, 1$, *is continuously differentiable in a neigbourhood of* $(\mathbf{\Theta}_{0i}, \mathbf{\Sigma}_{0i})$ *with probability one under* H_0 *and* H_1, *where* $\mathbf{\Sigma}_{0i}$ *is as in assumption 2(b).*

(b) Let $(\partial/\partial(\mathbf{\Theta}_i, \mathbf{\Sigma}_i^{-1}))$ *denote the partial differentiation with respect to* $\mathbf{\Theta}_i$ *and the non redundant elements of* $\mathbf{\Sigma}_i^{-1}$. S_i^ν *is bounded as*

$$E \sup_{\mathbf{\Theta} \in B(\mathbf{\Theta}_{0i}, \epsilon_T), \mathbf{\Sigma}_i \in N(\mathbf{\Sigma}_{0i})} \|(\partial/\partial(\mathbf{\Theta}_i, \mathbf{\Sigma}_i^{-1}))S_i^\nu(\widehat{\mathbf{\Theta}}_i, \widehat{\mathbf{\Sigma}}_i)\| < \infty$$

for $\nu = 0, 1$ *and for some* $\epsilon_T > 0$, *where* $\mathbf{\Sigma}_{0i}$ *is as in assumption 2(b).* $N(\mathbf{\Sigma}_{0i})$ *denotes some neighbourhood of* $\mathbf{\Sigma}_{0i}$.

(c) The distribution function of $S_i^\nu(\widehat{\mathbf{\Theta}}_{0i}, \widehat{\mathbf{\Sigma}}_{0i})$, $\nu = 0, 1$, *is continuous and increasing at its* $1 - \alpha$ *quantile, when* $m > d$.

Assumption 4

(a) $E\left[U_{it}\mathbf{X}_{it}\right]=0, \forall\ i$ *and* $t\geq 1$.

(b) $E\left[U_{it}^{2}\right]<\infty$ *and* $E\|\mathbf{X}_{it}\|^{2+\delta}<\infty$ *for some* $\delta>0$ *and* $\forall\ i$ *and* $t\geq 1$.

(c) $E\left[\mathbf{X}_{it}\mathbf{X}_{it}^{\prime}\right]$ *and* $\mathbf{\Sigma}_{0}=E\left[\mathbf{U}_{i,1}^{m_i}\mathbf{U}_{i,1}^{\prime\ m_i}\right]$ *are positive definite,* $\forall\ i$ *and* $t\geq 1$.

Assumptions (1), (3) and (4) are identical to that of Andrews (2003). The assumptions also hold for $P_i(\mathbf{\Theta}_i, \mathbf{\Sigma}_i)$ when $m_i\leq d$. The first assumption allows for both weakly dependent processes and long memory processes, as well as conditional variation in all moments, including conditional heteroskedasticity. Assumption 3 is required to ensure that the empirical distribution of the S statistics converge to the true distribution as derived in Andrews (2003). Furthermore, Assumption 3 ensures that the distribution of the S statistics are differentiable and finite. Assumption 2 is required to ensure the consistency of the estimators for both the coefficient vector and the variance-covariance matrix; it is a panel extension to Assumption 2 in Andrews (2003). The assumption covers estimators whose consistency properties rely on large T and N, as is the case in the presence of cross sectional dependence. Obviously, this assumption also covers estimators whose consistency properties rely on just a large T or N alone. Assumptions (1) - (3) are sufficient for all the asymptotic results that follow. However, these assumptions can be simplified further if the parameter vector is estimated by Ordinary Least Squares. In such a case, assumptions (1) and (4) are sufficient for assumptions (2) and (3) to hold, as shown in Lemma (1) (see also Andrews, 2003).

In the event of cross sectional dependence it is possible to use the Common Correlated Effect (CCE) estimator as proposed by Pesaran (2006) and discussed in section 2.4. Such an estimator would require slight modifications to the above assumptions for the asymptotic results to hold. These are:

Assumption CCE 1

(a) All individuals share the same break date, that is, $T_i=T$ *and* $m_i=m\ \forall i=1,...,N$.

(b) f_t *is covariance stationary with absolute summable autocovariance, distributed independently* u_{it} *and* $\mathbf{x}_{it}$, $\forall i,t$.

(c) The unobservable factor loadings, γ_i *are independently and identically distributed across* i *and independent to* u_{it} *and* f_t $\forall i,t$, *with fixed mean* γ, *and finite variance.*

(d) $T^{-1}[(\mathbf{X}_{i,1}^{T})^{\prime}M_{w,1}^{T}\mathbf{X}_{i,1}^{T}]$, $m^{-1}[(\mathbf{X}_{i,T+1}^{\bar{T}})^{\prime}M_{w,T+1}^{\bar{T}}\mathbf{X}_{i,T+1}^{\bar{T}}]$, $T^{-1}[(\mathbf{X}_{i,1}^{T})^{\prime}G_{1}^{T}\mathbf{X}_{i,1}^{T}]$ *and* $m^{-1}[(\mathbf{X}_{i,T+1}^{\bar{T}})^{\prime}G_{T+1}^{\bar{T}}\mathbf{X}_{i,T+1}^{\bar{T}}]$ *are non-singular, where* $G_{p}^{q}=(f_{p},...,f_{q})^{\prime}$ *is a* $(q-p+1)\times l$ *matrix containing the values of common factors from the sample spanning* $t=p,...,q$.

(e) All other necessary assumptions required by Pesaran (2006) to ensure the consistency of the CCE estimator.

Assumptions CCE 1 (b) - (e) are sufficient for the consistency of the CCE estimator (see Lemma 5 and Pesaran, 2006).

3.2 Results and Comments

This sub-section derives the asymptotic distribution for the Z (and $\tilde{Z}$) statistic and defines the properties of the tests.

Lemma 1 *Assumptions 1 and 4 imply that Assumptions 2 and 3 hold for the regression model estimated using OLS.*

Proof. See Appendix.

Remark 1 Lemma 1 is useful for reducing the number of assumptions. Assumption 4, in its current formulation, is made strictly for the Least Squares estimation procedure. For other estimators, such as IV or GMM, the conditions in Assumption 4 must be modified accordingly. These, however, need not guarantee the result in Lemma 1. Therefore Assumptions 2 and 3 are still required for the remaining results of this paper to hold when different estimators are used. For the appropriate modifications to Assumption 4 for IV or GMM see Andrews (2003).

Lemma 2 *Let $S^{\nu}_{i,\infty}$ be a random variable with the same distribution as $S^{\nu}_i(\mathbf{\Theta}_{0i}, \mathbf{\Sigma}_{0i})$, $\nu = 0, 1$. Under Assumptions 1-3 and Theorem 1 in Andrews (2003), as $T \to \infty$:*

(a) $S^{\nu}_i(\widehat{\mathbf{\Theta}}_i, \widehat{\mathbf{\Sigma}}_i) \xrightarrow{d} S^{\nu}_{i,\infty}\ \forall\ i = 1, \ldots, N$ and $\nu = 0, 1$

(b) Let F^{ν}_i be the distribution of $S^{\nu}_i\ \forall\ i$, then F^{ν}_i is a well defined distribution with finite mean and variance.

Proof. See Appendix.

Lemma 3 *Under Lemma 2,*

$$\lim_{N\to\infty} (\bar{\sigma}^2_{\nu} N)^{-1} \sum_{i=1}^{N} \int_{(S^{\nu}_i - E(S^{\nu}_i))^2 \geq \varepsilon N \bar{\sigma}^2_{\nu}} ((S^{\nu}_i - E(S^{\nu}_i))^2 dF_i(S^{\nu}_i) = 0$$

where $\bar{\sigma}^2_{\nu} = N^{-1} \sum_{i=1}^{N} Var(S^{\nu}_i)\ \forall i = 1, ..., N$, $\nu = 0, 1$ and $\forall \epsilon > 0$.

Proof. See Appendix.

Lemma 4 *Under Lemma 3 the asymptotic distribution of the $\bar{S}^{\nu}$ statistic is*

$$\sqrt{N}\, \frac{\bar{S}^{\nu} - E(\bar{S}^{\nu})}{\sqrt{Var(\bar{S}^{\nu})}} \overset{A}{\sim} N(0, 1)$$

with $E(\bar{S}^{\nu}) = N^{-1} \sum_{i=1}^{N} E(S^{\nu}_i)$, $Var(\bar{S}^{\nu}) = N^{-1} \sum_{i=1}^{N} Var(S^{\nu}_i)$

Proof. See Appendix.

Theorem 1 *Under Lemma 4, the Z statistic as described in equation (15)*

$$Z = \frac{\left(\bar{S}^1 - \bar{S}^0\right)}{\sqrt{\widehat{Var}\left(\bar{S}^1 - \bar{S}^0\right)}}$$

has an asymptotic distribution

$$\sqrt{N}\ Z \overset{A}{\sim} N\left(0,1\right)$$

Proof. See Appendix.

Remark 2 Lemma 2 shows that each S_i^ν converges to a well defined distribution with finite mean and variance. This is an important result as it is a necessary condition for Lemma 3 and 4 to hold, which subsequently lead to the proof of asymptotic normality for the arithmetic average of S_i^ν (namely, $\bar{S}^\nu$). The asymptotic normality of the proposed test statistics removes the need of using sub-sampling techniques to calculate the critical values as proposed in Andrews (2003).

Remark 3 Lemma 3 shows that Assumptions 1 - 4 are sufficient to satisfy the Lindeberg condition required by the LF-CLT. This is particularly important as the $\bar{S}^\nu$ statistic is the average of the S_i^ν statistics computed for every individual in the panel. The earlier assumptions imply that the variance of the $\bar{S}^\nu$ statistics is not dominated by the variance of the S_i^ν statistics from any particular individual.

Remark 4 Although $\bar{S}^\nu$ converges to a normal distribution asymptotically, the mean and the variance of the statistics are still unknown. Hence, it is not possible to draw statistical inference on $\bar{S}^\nu$ alone. Under the null hypothesis, however, the mean of $\bar{S}^\nu$ is the same for $\nu = 0, 1$ and therefore the Z statistic – which is based on the difference between $\bar{S}^1$ and $\bar{S}^0$ – will have mean 0. Furthermore, the variance of $(\bar{S}^0 - \bar{S}^1)$ can be estimated from S_i^ν for $i = 1, \ldots, N$ and $\nu = 0, 1$ defined earlier. It ensues that the Z statistic will converge to a standard normal distribution in which valid inference can be obtained.

Remark 5 Theorem 1 and Lemmas 1 - 4 hold under Assumptions 1 - 4 for $P_i^\nu(\mathbf{\Theta}_i, \mathbf{\Sigma}_i)$ as defined under equation (14), when $m_i \leq d$.

Lemma 5 *Under Assumptions CCE (1) (a) - (e), the CCE estimator, $\tilde{\mathbf{\Theta}}_i$ as defined in equation (16) is consistent. Moreover, if $\sqrt{T}/N \rightarrow 0$ as $(N, T) \rightarrow \infty$, then $\tilde{\mathbf{\Theta}}_i$ is asymptotically normal.*

Proof. See Appendix.

Theorem 2 *Under Assumptions (1), (4) and the Assumptions under Lemma 5*

$$\sqrt{N}\tilde{Z} \overset{A}{\sim} N(0,1)$$

Proof. See Appendix.

Remark 6 Lemma 5 is required to obtain consistency for the CCE estimator and Theorem 2 is homologous to Theorem 1 when using a CCE estimator to tackle cross sectional dependence.

4. Simulations

4.1 Monte Carlo Design

This section aims to provide some benchmark Monte Carlo results in order to investigate the normality, size and power of the proposed test for heterogeneous breaks in panels. The experiment uses the following linear regression model

$$y_{it} = \mathbf{\Theta}_i^{'}\mathbf{x}_{it} + u_{it} \quad t = 1, \ldots, \bar{T} \quad i = 1, \ldots, N$$

where the number of regressors in $\mathbf{x}_{it}$ is set to d = 5 including a constant. All regressors are calculated as a trigonometric function of a set of random normal variables, which are independent and identical.

$$\mathbf{x}_{it} = \tan(\mathbf{w}_{it})$$

where $\mathbf{w}_{it}$ is the vector of the random normal iid variables. The regression's error term, u_{it} is generated with an AR(1) process

$$u_{it} = \rho u_{it-1} + \epsilon_{it} \quad t = 1, \ldots, \bar{T}, \quad i = 1, \ldots, N$$

with the following autoregressive parameters: ρ = 0.4 and 0.95 which is common to all individuals (in other words, all individuals' errors have the same ρ). Four different types of iid distributions for the innovation of the error term are considered: standard N(0, 1), a recentered and rescaled χ_5^2 and t_5 with mean zero and variance one, and an uniform distribution with support [0, 1]. More formally,

$$\epsilon_{it} \sim \begin{cases} N(0,1) & \text{for } i = 1, \ldots, \frac{N}{4} \\ \chi_5^2 & \text{for } i = \frac{N}{4} + 1, \ldots, \frac{N}{2} \\ t_5 & \text{for } i = \frac{N}{2} + 1, \ldots, \frac{3N}{4} \\ U[0,1] & \text{for } i = \frac{3N}{4} + 1, \ldots, N \end{cases}$$

Thus different individuals have different innovation processes, such that the four distributions are intermixed evenly in the panel.[3]

[3] Similar Monte Carlo experiments allowing cross sectional dependence have also been conducted using the CCE version of the Z test statistic as defined in equation (24). The results are comparable to those presented here. Since cross sectional dependence is not the focus of the paper, the Monte Carlo results are not presented but are available upon request.

The results from the simulation exercises are presented in two parts. Firstly, the set of Monte Carlo experiments simulate the null in order to analyse the size of the test. Moreover, a discussion of the properties of the distributions of the test under the LF-CLT is provided. The null hypothesis is simulated over the full sample $\bar{T}$ using the coefficient vector $\mathbf{\Theta}_{0i} = 0, \forall\, i$. Secondly, the power properties of the test are examined. The alternative hypothesis of partial instability is simulated, allowing only a limited number N_1 of individuals to experience a structural break. The ratio $\frac{N_1}{N} = c$ is gradually changed from 0.10, 0.50, 0.65, 0.80, to 1, in order to allow for a larger proportion of the individuals to experience a structural break. The alternative hypothesis featuring a partial structural break is simulated using $\mathbf{\Theta}_{0i} = 0$ and $\delta_i = \frac{1}{10} \times (1,1,1,1,1)'$, for some i and $\delta_j = 0$, for some $j \neq i$. Moreover, all results use $\|\delta_i\| = \sqrt{0.05}$, where $\|x\|$ denotes the Euclidean norm for the vector x. Note that when $\frac{N_1}{N} = 1$, the coefficient vector is homogeneous across i's, implying that all individuals experience a structural break.

Moreover, Monte Carlo experiments are conducted with the following settings: $m = m_i = \frac{1}{10} \times \bar{T}$, $\bar{T}$= 30, 50, 100, N= 20, 40, 60, 80, 100, where, as earlier, $\bar{T} = T + m = T_i + m_i$. For simplicity, the break dates T_i and post-break observations m_i are known and identical for all individuals. The distribution property, size and power of the test are also investigated when the number of post-break observations are increased to $m = \frac{1}{20} \times \bar{T}$ for $\bar{T}$ = 30, 50. The number of replications is 2000. All simulations are carried out using Ox 4.02.[4]

4.2 Monte Carlo Results

4.2.1 Size

The first results look at the probabilities of a type I error with significance level of 0.05. The main results can be summarised as follows:

1. Overall the Monte Carlo experiments reveal that the test statistic is close to normal with 2000 replications showing that the LF-CLT holds with moderate serial correlation and both relatively small time and cross-sectional dimensions. The Jarque-Bera test statistics show strong evidence of normality at the 5% level of significance. The results are shown in Table 1 and 2, where Table 2 presents results when the number of post-break observations are increased to 20% of $\bar{T}$ instead of 10%.

2. As shown in Table 3, the size of the test is relatively close to the desired value of 0.05 for $N > 20$ with $\bar{T}$ = 100 and m = 10. These results hold even in the presence of moderate serial correlation (ρ = 0.4). The test has reasonable size when the time horizon is decreased to $\bar{T}$ = 50 and $\bar{T}$ = 30, especially when $N \geq 40$ and 80 respectively. Size is relatively unaffected if the number of post-break observations are increased to 20% of $\bar{T}$.

[4] The programming code is available upon request.

3. The normality of the distribution worsens in the presence of extreme serial correlation (ρ = 0.95) as shown in Table 1. The Jarque-Bera test for normality is rejected at the 5% level of significance.

4. The size, on the other hand, deteriorates with extreme serial correlation especially as the number of individuals increases. This result is expected as all individuals exhibit the same high degree of serial correlation. Under these conditions, increasing $\bar{T}$ from 100 to 250 observations improves the size, as implied by ergodicity.

In sum, the test has reasonable size even in small temporal and cross-sectional samples with moderate serial correlation. However, under extreme serial correlation the size of the test deteriorates substantially, especially as the cross-sectional dimension increases.

4.2.2 Power

Overall the test has good power. The power of the test is analysed for the significance level of 0.05. Results are shown in Table 4. The most important results of the Monte Carlo experiments are as follows:

1. The power of the test remains satisfactory even with a relatively limited time dimension, except when $\bar{T}$ is very small ($\bar{T} \leq 50$). A larger N, though, counterbalances the effect of diminishing $\bar{T}$; this underlines the advantage of working with a panel structure. When $\bar{T}$ = 50, for example, the test remains powerful when N = 80 and c = 0.80 (power is 0.65). The power of the test improves if the number of post-break observations are increased to 20% of $\bar{T}$. For example, when $\bar{T}$ = 50, c = 0.80 and N = 60, the power is 0.72, instead of 0.44 when m is 10% of $\bar{T}$.

2. The test gains power as either N, c or $\bar{T}$ increases. For instance, the power of the test is above 0.95 when $N \geq$ 80 and $c \geq$ 0.65, for $\bar{T}$ = 100. But even when both c and N are of medium size, as when c = 0.65 and N = 60 (and $\bar{T}$ = 100), the power is 0.85. Moreover, the power of the test is still good (0.70) when N is high (100) and c is low (0.50) with $\bar{T}$ = 100. The reverse is also true: when N = 40 and c = 0.80, the power is 0.91.

3. The power of the test is quite robust to serial correlation, especially when N and c are large. Even in extreme cases when serial correlation is 0.95, for instance, the test has power of 0.71 when N = 100 and c = 0.80.

Overall, the power of the test is good given the data generating process. Power increases with N, c , $\bar{T}$ and m. Power is better when serial correlation is moderate, but remains robust even to very high levels of serial correlation.

5. Empirical Example

This section provides an empirical application to demonstrate the usefulness of the proposed test, focusing on the question of whether the euro has increased intra-Eurozone trade. The question has recently been at the forefront of the empirical trade literature, revived after the seminal contribution of Rose (2000), and has been discussed actively in policy circles. However, empirical evidence has been clouded by econometric techniques somewhat ill-suited for the very few available data.

Most papers in the literature, of which the most prominent are Micco et al. (2003) and Flam and Nordström (2003),[5] introduce various flavors of dummy variables in their regressions to capture the new currency's introduction. Furthermore, the use of F-tests employed to evaluate the significance of the dummy coefficients rest on highly restrictive assumptions in finite samples: normal, homoskedastic and iid errors. These are particularly bold in light of the macro data typically used in these exercises, where heteroskedasticity and autocorrelated errors are commonplace. Lastly, Andrews (2003) shows that F-tests exhibit large size distortions when testing for parameter instability at the end of sample. Given these limitations, some authors like Micco et al. (2003) avoid, in part, the use of explicit tests and rely on eye-balling the size of the coefficients on the euro-dummies.

The test developed in this paper allows for a very different and more rigorous approach, better adapted for the question of the euro's effect on trade. First, the test is residual based and does not require the estimation of coefficients on dummy variables to capture the effect of the euro. Second, the test requires very few regularity conditions. It remains asymptotically valid despite non-normal, heteroskedastic and/or autocorrelated errors. Third, the test is explicitly designed for few datapoints following a presumed break and makes no distributional assumptions on any individual specific test-statistic; only the cross sectional average statistic is shown to be asymptotically normal as warranted by the panel's cross sectional dimension. Fourth, the test explicitly allows for some individuals, and not all, to exhibit a break.

This last characteristic, allowing for heterogeneous instability, is particularly well suited for the example at hand. For instance, while it was clear that Germany was going to play a central role in the euro from its inception, uncertainty over whether Italy would meet the strict accession requirements loomed almost until the euro's introduction. It would therefore seem natural that each country's trade pattern would have responded differently, if at all, to the new currency's introduction.

The test for the euro's effect on trade is rooted in a standard trade gravity equation, used in various flavors in all the above-mentioned papers, and whose microfoundations are discussed at some length in Mancini-Griffoli and Pauwels (2006). The regression used here is:

$$\begin{aligned} V_{i,jt} &= \alpha_{i,j} + \psi_1 Y_{it} + \psi_2 Y_{jt} + \psi_3 \xi_{i,jt} + \epsilon_{it} \\ \epsilon_{it} &= \gamma_i f_t + \nu_{it} \end{aligned} \tag{25}$$

5 See also Bun and Klaassen (2002), Nitsch (2002), De Sousa (2002), Barr et al. (2003), De Nardis and Vicarelli (2003), Piscitelli (2003), Nitsch and Berger (2005), and Baldwin (2006) (which offers a particularly nice summary of the literature).

where $V_{i,jt}$ is the value of imports from country j to country i, Y_{jt} and Y_{it} are nominal GDP at time t for country j and country i, respectively, to control for demand and country size effects, $\xi_{i,jt}$ is the real exchange rate between the two countries engaged in trade, capturing relative price effects as well as changes in relative demand for tradables, and $\alpha_{i,j}$ is a pair-specific fixed effect to control for variables of type common border, language, history, legal system, distance and others traditionally shown to matter in gravity equations. Also, f_t includes observed and unobserved common effects, including time effects (responsible for any cross sectional correlation of the errors). Finally, ν_{it} is the individual specific idiosyncratic shock.

Several modifications to the above regression are necessary, though, in order to carry out proper estimation. First, all variables fail to reject the null of a unit root. Mancini-Griffoli and Pauwels (2006) present these results along with a discussion. Here, the most straightforward solution is adopted: that of taking all variables in first-differences. The test therefore becomes one for a break in the relation between the growth of trade and the growth of its explanatory variables. Other solutions to the problem of non-stationary data are considered in Mancini-Griffoli and Pauwels (2006), but are not adopted here, as this section limits itself to a mere illustration of the new panel test for heterogeneous breaks.

Second, the errors of the model are found to be cross sectionally dependent. It is necessary, therefore, to augment the proposed test in the fashion proposed in section 2.4, in order to “filter out” the common factors causing cross sectional correlations. These unknown common factors can be proxied by the cross sectional sample averages of the regressors and regressand, as proposed in Pesaran’s (2006) common correlated effect (CCE) estimator.

Finally, quarterly data were obtained from Eurostat, IMF DOTS and IFS, as in most other relevant empirical papers. The unilateral import values were obtained from IMF DOTS. All data were seasonally adjusted using the standard X.12 smoothing algorithm.

Given these modifications, the equation serving as the baseline model for the panel stability test is written as

$$\Delta V_{i,jt} \quad = \quad \alpha_{i,j} + \psi_1 \Delta Y_{it} + \psi_2 \Delta Y_{jt} + \psi_3 \Delta \xi_{i,jt} + \epsilon_{it} \tag{26}$$

where Δ indicates the first difference of the variable.

Results from this paper’s proposed panel test for heterogeneous breaks are presented below.[6] For simplicity – again because this section merely aims to be an illustration – only one potential break point is considered, in 1998 Q1, one year prior to the actual adoption of the euro. This is to take into account the extent to which agents are forward looking, as well as directly test the findings of Micco et al. (2003) and Flam and Nordström (2003) who find a “euro effect” as early as 1998. Furthermore, results

6 All empirical results were generated using RATS 6.30. The programming code is available upon request. Estimation results and other specifications of the regression equation are covered in Mancini-Griffoli and Pauwels (2006).

are presented for $\bar{S}^0$ statistics found by sampling from different date ranges in the pre-break sample. This is to gauge the sensitivity of the test to variations in data and to serial correlation in the errors: on the one hand, the closer are the sampling dates in the $\bar{S}^0$ and $\bar{S}^1$ statistics, the greater are the chances of disturbances due to serial correlation. On the other hand, the earlier is the pre-break sampling date, the more disturbances could arise from less reliable data. Thus, results are presented for four "test samples", each including a different pre-break sampling date: 1980 Q2, 1985 Q1, 1987 Q1 and 1990 Q1. Results are presented in Table 5.

The first general pattern that emerges from glancing at the results across the various test samples, is that there indeed appears to be a break in the relation between trade and its explanatory variables in 1998 Q1. Indeed, most test samples have consistent periods over which the null of stability is rejected.

Second, and more specifically, the degree with which the null is rejected – if at all – is sensitive to the number of quarters included in the post-break sample (the length of m in the earlier derivations). The length of m is progressively increased from about 10% of the sample, which translates roughly to eight post-break observations in the current sample. When eight to 10 quarters of post break observations are included, the null is not rejected. The test results may be sensitive when the post-break observations are few. But as the post-break sample grows to cover 11 - 14 quarters of data, most test samples reject the null consistently with at least 10% significance. This is in line with the earlier Monte Carlo results showing how the test's power increases with m. Note that as more than 14 quarters of post-break data are considered, the null is no longer rejected. Thus, the break in trade due to the euro, although significant, seems to be limited in duration, lasting only slightly more than three years. In itself, this is an interesting finding, contradicting views sometimes expressed in political debates that common market effects will grow with time. It seems that the Eurozone has already reaped the benefits of the euro, at least in terms of gains in trade.

Third, results, although broadly consistent, can be somewhat sensitive to the choice of pre-break sampling dates. Indeed, there are slight variations in results across the various test samples. First, only the 1990 test sample rejects the null with 1-5% significance for 11-14 quarters. The other samples reject with 1-10% significance for a subset of these quarters. Again, these slight differences are to be expected given the noise in the data. Thus, testing for the robustness of results to the choice of pre-break sampling can be important empirically. Secondly, results with the most recent 1990 test sample are consistent and show strong significance. This highlights, once again, the relative robustness of the test to serial correlation, as mentioned in the earlier Monte Carlo results.

On the whole, the above exercise has allowed for both rigor and flexibility in testing an important policy question, and has delivered a statistically solid and relatively consistent answer; this is an improvement over previous work which, although pioneering, was clouded by somewhat ill-adapted traditional econometric techniques. What, exactly, in the new currency caused this rise in trade is another question well worth considering in further research. But at least end of sample instability tests, like the one presented here, lay solid and precise foundations for such research to continue its course.

6. Concluding Remarks

This paper builds a stability test for panel data, robust to non-normal, heteroskedastic and serially correlated errors, and, importantly, to very few datapoints after a break. Moreover, the test is specifically designed for heterogeneous breaks, whereby only some – and not all – individuals in a panel exhibit a break. The test statistic is constructed as a standardised average of independent test statistics computed for each cross section. Asymptotic results show that the test is normally distributed as per the Lindeberg-Feller central limit theorem.

Monte Carlo results show that the test performs well in terms of power and size, even when the time and individual dimensions are small. Moreover, the test performs relatively well in the presence of serial correlation in the errors, especially when the time dimension is large. These results should allow the test to be used widely in finance and economics applications. This paper explores one such application, testing for the effect of the euro's introduction on intra-Eurozone trade.

References

Andrews, D. W. K. (1993), "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica*, 61(4): 821–56.

Andrews, D. W. K. (2003), "End-of-sample Instability Tests," *Econometrica*, 71(6): 1661–94.

Andrews, D. W. K. and W. Ploberger (1994), "Optimal Tests when a Nuisance Parameter is Present only under the Alternative," *Econometrica*, 62: 1383–414.

Bai, J. (2005), "Panel Data Models with Interactive Fixed Effects," Unpublished, New York University.

Bai, J. and S. Ng (2002), "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70(11): 191–221.

Bai, J. and P. Perron (1998), "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica*, 66: 47–78.

Baldwin, R. (2006), "The Euro's Trade Effects," ECB Working Paper No.594.

Barr, D., F. Breedon and D. Miles (2003), "Life on the Outside: Economic Conditions and Prospects Outside Euroland," *Economic Policy*, 37: 573–613.

Bun, M. J. and F. J. Klaassen (2002), "Has the Euro Increased Trade?" Tinbergen Institute Discussion Paper TI 2002-108/2.

Choi, I. (2001), "Unit Root Tests for Panel Data," *Journal of International Money and Finance,* 20: 249–72.

Chow, G. C. (1960), "Tests of Equality between Sets of Coefficient in Two Linear Regressions," *Econometrica*, 28: 591–605.

Coakley, J., A. Fuertes and R. Smith (2002), "A Principal Components Approach to Cross-dependence in Panels," Unpublished manuscript, Birkbeck College, University of London.

De Nardis, S. and C. Vicarelli (2003), "Currency Unions and Trade: The Special Case of Emu," *World Review of Economics,* 139(4): 625–49.

De Sousa, L. V. (2002), "Trade Effects of Monetary Integration in Large, Mature Economies: A Primer on European Monetary Union," Kiel Working Paper No.1137.

De Wachter, S. and E. Tzavalis (2004), "Detection of Structural Breaks in Linear Dynamic Panel Data Models," Working Paper No.505, Queen Mary, University of London, Department of Economics.

Emerson, J. and C. Kao (2001), "Testing for Structural Change of a Time Trend Regression in Panel Data: Part i," *Journal of Propagations in Probability and Statistics*, 2: 57–75.

Emerson, J. and C. Kao (2002), "Testing for Structural Change of a Time Trend Regression in Panel Data: Part ii," *Journal of Propagations in Probability and Statistics*, 2: 207–50.

Flam, H. and H. Nordström (2003), "Trade Volume Effects of the Euro: Aggregate and Sector Estimates," Manuscript, Institute for International Economic Studies.

Han, A. K. and D. Park (1989), "Testing for Structural Change in Panel Data: Application to a Study of U.S. Foreign Trade in Manufacturing Goods," *Review of Economics and Statistics*, 71: 135–42.

Im, K. S., H. Pesaran and Y. Shin (2003), "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics*, 115: 53–74.

Kao, C., L. Trapani, and G. Urga (2005), "Modelling and Testing for Structural Breaks in Panels with Common and Idiosyncratic Stochastic Trends," Syracuse University, Department of Economics.

Mancini-Griffoli, T. and L. L. Pauwels (2006), "Did the Euro Affect Trade? Answers from End-of-sample Instability Tests," HEI Working Paper, Graduate Institute of International Studies, Geneva, Economics Section.

Micco, A., G. Ordoñez and E. Stein (2003), "The Currency Union Effect on Trade: Early Evidence from Emu," *Economic Policy*, 18(37): 316–56.

Nitsch, V. (2002), "Honey, I Shrunk the Currency Union Effect on Trade," *The World Economy*, 25: 457–74.

Nitsch, V. and H. Berger (2005), "Zooming Out: the Trade Effect of the Euro in Historical Perspective," CESifo Working Paper No.1435.

Pesaran, M. H. (2006), "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," *Econometrica*, 74(4): 967–1012.

Piscitelli, L. (2003), "Available from UK Treasury," Mimeo.

Rose, A. (2000), "One Market one Money: Estimating the Effect of Common Currencies on Trade," *Economic Policy*, 15(30): 7–45.

Table 1. Moments for the Distribution of the Test under the Null

				N				
Moments	$\bar{T}$	m	ρ	**20**	**40**	**60**	**80**	**100**
Mean	30	3	0.4	-0.063	0.078	-0.122	-0.135	-0.112
	50	5	0.4	-0.088	-0.086	-0.156	-0.150	-0.181
	100	10	0.4	-0.057	-0.051	-0.112	-0.063	-0.061
	100	10	0.95	-0.539	-0.746	-0.936	-1.06	-1.24
Variance	30	3	0.4	1.19	1.09	1.07	1.02	1.04
	50	5	0.4	1.18	1.076	1.06	1.02	1.05
	100	10	0.4	1.11	1.03	1.01	1.04	1.05
	100	10	0.95	0.907	0.914	0.834	0.758	0.735
Skewness	30	3	0.4	0.033	-0.095	-0.041	-0.007	0.002
	50	5	0.4	0.022	-0.005	0.011	0.036	-0.009
	100	10	0.4	0.016	-0.020	-0.004	-0.040	0.025
	100	10	0.95	0.471	0.397	0.288	0.394	0.262
Kurtosis	30	3	0.4	3.28	3.16	3.09	3.05	3.06
	50	5	0.4	3.13	2.87	2.92	3.02	3.11
	100	10	0.4	3.07	3.07	3.16	2.95	2.96
	100	10	0.95	3.11	3.05	2.98	3.52	3.11
Jarque-Bera	30	3	0.4	6.90	5.14	1.24	0.22	0.30
	50	5	0.4	1.57	1.42	0.57	0.47	1.04
	100	10	0.4	0.49	0.54	2.14	0.74	0.36
	100	10	0.95	75.08	52.71	27.62	74.39	23.97

Note: ρ is the autocorrelation coefficient, N are the total number of individuals and $\bar{T} = T_i + m_i$ where $T_i = T$ is the time dimension prior to the instability fixed for all individuals $i, m_i = m$ is the number of observations post instability fixed for all individuals i and m is set to equal 10% of $\bar{T}$. The Jarque-Bera normality test has an asymptotic χ^2_2 distribution and its critical value is 5.99 at the 5% level of significance.

Table 2. Moments under the Null When $m = \frac{1}{20} \times \bar{T}$

				N				
Moments	$\bar{T}$	m	ρ	**20**	**40**	**60**	**80**	**100**
Mean	30	6	0.4	-0.042	-0.048	-0.048	-0.102	-0.096
	50	10	0.4	-0.031	-0.059	-0.064	-0.052	-0.035
Variance	30	6	0.4	1.16	1.05	1.03	1.02	1.08
	50	10	0.4	1.20	1.07	1.06	1.01	1.02
Skewness	30	6	0.4	-0.012	-0.05	0.003	0.023	-0.005
	50	10	0.4	-0.006	0.012	-0.028	-0.038	0.008
Kurtosis	30	6	0.4	3.09	2.84	3.17	3.10	3.14
	50	10	0.4	3.12	2.89	3.13	3.10	3.04
Jarque-Bera	30	6	0.4	0.72	2.97	2.41	1.01	1.64
	50	10	0.4	1.12	1.06	1.67	1.31	2.35

Note: ρ is the autocorrelation coefficient, N are the total number of individuals and $\bar{T} = T_i + m_i$ where $T_i = T$ is the time dimension prior to the instability fixed for all individuals i, $m_i = m$ is the number of observations post instability fixed for all individuals i and m is set to equal 20% of $\bar{T}$. The Jarque-Bera normality test has an asymptotic χ^2_2 distribution and its critical value is 5.99 at the 5% level of significance.

Table 3. Size of Normal Significance Level 0.05

			N				
$\bar{T}$	m	ρ	**20**	**40**	**60**	**80**	**100**
					$m = 0.1 \times \bar{T}$		
30	3	0.4	0.069	0.061	0.064	0.054	0.053
50	5	0.4	0.070	0.056	0.055	0.057	0.058
100	10	0.4	0.065	0.047	0.055	0.053	0.054
	10	0.95	0.055	0.092	0.130	0.142	0.199
					$m = 0.2 \times \bar{T}$		
30	6	0.4	0.071	0.055	0.058	0.049	0.060
50	10	0.4	0.075	0.061	0.062	0.048	0.053

Note: ρ is the autocorrelation coefficient, N are the total number of individuals and $\bar{T} = T_i + m_i$ where $T_i = T$ is the time dimension prior to the instability fixed for all individuals i, $m_i = m$ is the number of observations post instability fixed for all individuals i and m is set to equal 10% of $\bar{T}$ and also 20% of $\bar{T}$ for $\bar{T}$ = 30, 50.

Table 4. Power of Normal Significance Level 0.05 for $\frac{N_1}{N} = c$

				N				
$\bar{T}$	m	c	ρ	**20**	**40**	**60**	**80**	**100**
						$m = 0.1 \times \bar{T}$		
30	3	0.80	0.4	0.015	0.041	0.071	0.139	0.223
	3	1	0.4	0.079	0.263	0.477	0.681	0.817
50	5	0.50	0.4	0.009	0.011	0.011	0.025	0.034
	5	0.80	0.4	0.056	0.216	0.435	0.651	0.809
	5	1	0.4	0.297	0.744	0.951	0.992	0.999
100	10	0.10	0.4	0.000	0.003	0.004	0.005	0.006
	10	0.50	0.4	0.022	0.114	0.294	0.523	0.696
	10	0.65	0.4	0.095	0.520	0.856	0.967	0.991
	10	0.80	0.4	0.347	0.911	0.991	0.999	1.00
	10	1	0.4	0.852	0.994	1.00	1.00	1.00
100	10	0.50	0.95	0.008	0.011	0.011	0.012	0.014
	10	0.80	0.95	0.075	0.203	0.368	0.542	0.705
	10	1	0.95	0.270	0.620	0.869	0.967	0.994
						$m = 0.2 \times \bar{T}$		
30	6	0.80	0.4	0.031	0.077	0.141	0.233	0.348
	6	1	0.4	0.130	0.359	0.602	0.786	0.905
50	10	0.50	0.4	0.016	0.024	0.052	0.097	0.155
	10	0.80	0.4	0.134	0.433	0.720	0.898	0.965
	10	1	0.4	0.454	0.869	0.984	1.00	1.00

Note: ρ is the autocorrelation coefficient, N are the total number of individuals and $\bar{T} = T_i + m_i$ where $T_i = T$ is the time dimension prior to the instability fixed for all individuals i, $m_i = m$ is the number of observations post instability fixed for all individuals i and m is set to equal 10% of $\bar{T}$ and also 20% of $\bar{T}$ for $\bar{T}$ = 30, 50.

Table 5. Empirical Example - Euro's Trade Effect

Pre-break sampling date	Presumed break date	Number of quarters post-break	Value of Z-statistic	p-value
1980 Q2	1998 Q1	8 (up to 2000Q1)	0.62	0.53
		9	1.06	0.29
		10	-0.84	0.40
		11	1.81	0.07 *
		12 (up to 2001 Q1)	2.37	0.02 **
		13	2.66	0.01 ***
		14	0.39	0.70
		27	1.55	0.12
1985 Q1	1998 Q1	8 (up to 2000 Q1)	1.45	0.15
		9	0.71	0.48
		10	-0.89	0.38
		11	1.73	0.08 *
		12 (up to 2001 Q1)	1.38	0.17
		13	1.34	0.18
		14	0.70	0.49
		27	1.10	0.27
1987 Q1	1998 Q1	8 (up to 2000 Q1)	1.1	0.27
		9	1.94	0.05 **
		10	-0.88	0.38
		11	1.47	0.14
		12 (up to 2001 Q1)	-0.34	0.73
		13	2.55	0.01 ***
		14	1.77	0.08 *
		27	1.39	0.16
1990 Q1	1998 Q1	8 (up to 2000 Q1)	1.84	0.07 *
		9	0.91	0.36
		10	-0.87	0.38
		11	2.10	0.04 **
		12 (up to 2001 Q1)	2.44	0.01 ***
		13	2.31	0.02 **
		14	2.17	0.03 **
		27	0.66	0.51

Note: */**/*** indicate 10%/5%/1% level of significance.

Appendix

Proof of Lemma 1 *Since S_i^{ν} $\forall$ i is calculated by treating each cross section as a univariate time-series, the properties of S_i^{ν} are identical to the properties of the S and P statistic derived in Andrews (2003) for all i. This completes the proof.* ■

Proof of Lemma 2 *The proof of part (a) is similar to the proof of Lemma 1. Using Theorem 1 in Andrews (2003), $S_{i,\infty}^{\nu}$ has a well defined distribution with finite mean and variance for all i. Given part (a) S_i^{ν} converges to a well defined distribution with finite mean and variance for all i. This completes the proof.* ■

Proof of Lemma 3 *It is sufficient to show that the following Lindeberg condition holds under the assumptions made in the paper:*

$$\lim_{N\to\infty}(\bar{\sigma}_N^2 N)^{-1}\sum_{i=1}^{N}\int_{(S_i-E(S_i))^2\geq\bar{\sigma}_N^2 N\varepsilon}(S_i-E(S_i))^2 dF_i(S_i)=0 \qquad \text{(A-1)}$$

The proof holds for both ν = 0, 1 so the ν is dropped for notation convenience. Let $\bar{\sigma}_N^2 = N^{-1}\sum_{i=1}^{N} Var(S_i)$ and $A(\varepsilon,N)$ be the set such that $A(\varepsilon,N) = \{S_i|(S_i-E(S_i))^2\geq\bar{\sigma}_N^2 N\varepsilon\}$. Define

$$\delta_i(\varepsilon,N)=\int_{A(\varepsilon,N)}(S_i-E(S_i))^2 dF_i(S_i) \qquad \text{(A-2)}$$

Since $A(\varepsilon,N)\subset\mathbb{R}^+$, it is obvious that

$$\delta_i(\varepsilon,N)<\int_{\mathbb{R}^+}(S_i-E(S_i))^2 dF_i(S_i)=\sigma_i^2 \qquad \forall i=1,...,N$$

Notice that $\delta_i(\varepsilon,N)\to 0$ and $\bar{\sigma}_N^2 N=\sum_{i=1}^{N}\sigma_i^2\to\infty$ as $N\to\infty$ $\forall\varepsilon>0$, $i=1,...,N$. Define

$$r_i(\varepsilon,N)=\frac{\delta_i(\varepsilon,N)}{\sigma_i^2} \qquad \text{(A-3)}$$

so that $r_i(\varepsilon,N)\in[0,1)$, $r_i(\varepsilon,N)\to 0$ as $N\to\infty$ and $r_i(\varepsilon,N)>r_i(\varepsilon,N+1)$, $\forall N\in\mathbb{Z}^+$ and $\forall\varepsilon>0$. Given these definitions, equation (A-1) can be rewritten as

$$\lim_{N\to\infty}(\bar{\sigma}_N^2 N)^{-1}\sum_{i=1}^{N}\int_{(S_i-E(S_i))^2\geq\bar{\sigma}_N^2 N\varepsilon}(S_i-E(S_i))^2 dF_i(S_i)=\lim_{N\to\infty}\frac{\sum_{i=1}^{N}r_i(\varepsilon,N)\sigma_i^2}{\sum_{i=1}^{N}\sigma_i^2} \qquad \text{(A-4)}$$

Since $\sigma_i^2<\infty$, $\exists\sigma_{\max}^2<\infty$ such that $\sigma_{\max}^2\geq\sigma_i^2$, $\forall i$ and $\exists\sigma_{\min}^2>0$ such that $\sigma_{\min}^2\leq\sigma_i^2$, $\forall i$. Therefore,

$$\frac{\sum_{i=1}^{N} r_i(\varepsilon,N)\sigma_i^2}{\sum_{i=1}^{N}\sigma_i^2} \le \frac{\sigma_{\max}^2}{N\sigma_{\min}^2}\sum_{i=1}^{N} r_i(\varepsilon,N)$$
$$= \frac{c}{N}\sum_{i=1}^{N} r_i(\varepsilon,N)$$

where $c = \sigma_{\max}^2/\sigma_{\min}^2$. *Let*

$$\triangle_{i,N} = \frac{r_i(\varepsilon,N+1)}{r_i(\varepsilon,N)}$$

It is clear that $\triangle_{i,N} < 1 \ \forall i,N,\varepsilon$. *Therefore,* $\exists\triangle < 1$ *such that* $\triangle \ge \triangle_{i,N} \ \forall i,N,\varepsilon$. *Hence,*

$$\frac{c}{N}\sum_{i=1}^{N} r_i(\varepsilon,N) \le \frac{c}{N}\sum_{i=1}^{N}\triangle^{N-i} r_i(\varepsilon,i)$$
$$< \frac{c}{N}\sum_{i=1}^{N}\triangle^{N-i}$$
$$= \frac{c}{N}\frac{1-\triangle^N}{1-\triangle}$$

Therefore,

$$\lim_{N\to\infty}\frac{c}{N}\frac{1-\triangle^N}{1-\triangle} = 0$$

This completes the proof. ■

Proof of Lemma 4 *Under Lemma 2, it is straightforward to show that* S_i^{ν} *is independent of* S_j^{ν}, *for* $i \ne j$. *In addition,* $\bar{S}^{\nu}$ *satisfy the Lindeberg-Feller condition as shown in Lemma 3 for* $\nu = 0, 1$. *As these satisfy the conditions required by the Lindeberg-Feller CLT, then*

$$\sqrt{N}\,\frac{\bar{S}^{\nu} - E(\bar{S}^{\nu})}{\sqrt{Var(\bar{S}^{\nu})}} \overset{A}{\sim} N(0,1)$$

This completes the proof. ■

Proof of Theorem 1 *Under Lemma 4,* $\bar{S}^0$ *and* $\bar{S}^1$ *converge to a normal distribution in probability. By construction, the* Z *statistic is the standardised difference between two random variables that are normally distributed and therefore converge to a* N *(0, 1). Under the null hypothesis* $E\left[\bar{S}^0\right] = E\left[\bar{S}^1\right]$ *and hence* Z *converges in probability to a* N *(0, 1) distribution. This completes the proof.* ■

Proof of Lemma 5 *Note that Assumption CCE 1 (a) is identical to Assumption 1 in Pesaran (2006). Moreover, Assumption CCE 1 (b) is a special case of Assumption 2 in Pesaran (2006). Assumption CCE 1 (c) is equivalent to Assumption 3 in Pesaran (2006) where Assumption 4 in Pesaran (2006) is automatically satisfied as* $\mathbf{\Theta}_i$ *is assumed to be fixed under* H_0 *and non-random under* H_1*, for all* i*. Finally, Assumption CCE 1 (d) is equivalent to Assumption 5a in Pesaran (2006). Therefore, Theorem 1 in Pesaran (2006) holds under Assumptions CCE (1) (a) to (d), and hence* $\tilde{\mathbf{\Theta}}_i$ *is consistent and asymptotically normal. This completes the proof.* ■

Proof of Theorem 2 *It is straightforward to show that the result in Lemma 5 implies Assumptions 2 and 3 using Theorem 1 in Pesaran (2006). Moreover, since the independently distributed residuals,* u_{it} *can be estimated consistently using the CCE estimator, the statistics,* S_i^{ν}*, are therefore also independently distributed* $\forall i = 1, ...n$*,* ν *= 0,1. Given Assumptions (1) - (3) and the requirement of independence for the LF-CLT are satisfied, the proof then follows the same argument from the proof of Theorem 1. This completes the Proof.* ■