

Graduate Institute of  
International Studies | Geneva

# Economics

HEI Working Paper No: 24/2006

## Stability Tests for Heterogeneous Panel Data

Felix Chan

School of Economics and Finance  
Curtin University of Technology

Tommaso Mancini-Griffoli

Paris-Jourdan Sciences Economiques (PSE),  
CEPREMAP

Laurent L. Pauwels

Hong Kong Monetary Authority and  
Graduate Institute of International Studies, Geneva

### Abstract

This paper proposes a new test for structural instability in heterogeneous panels. The test builds on the seminal work of Andrews (2003) originally developed for time series. It is robust to non-normal, heteroskedastic and serially correlated errors, and allows for the number of post break observations to be small. Importantly, the test considers the alternative 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 additional cross sectional dimension of panel data. This greatly facilitates the calculation of critical values. Monte Carlo experiments show that the test has good size and power under a wide range of circumstances. The test is then applied to investigate the effect of the Euro on trade.

© The Authors.

All rights reserved. No part of this  
paper may be reproduced without  
the permission of the authors.

# Stability Tests for Heterogeneous Panel Data

Felix Chan  
School of Economics and Finance  
Curtin University of Technology

Tommaso Mancini-Griffoli  
Paris-Jourdan Sciences Economiques (PSE),  
CEPREMAP

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

This draft, December 2006

## Abstract

This paper proposes a new test for structural instability in heterogeneous panels. The test builds on the seminal work of Andrews (2003) originally developed for time series. It is robust to non-normal, heteroskedastic and serially correlated errors, and allows for the number of post break observations to be small. Importantly, the test considers the alternative 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 additional cross sectional dimension of panel data. This greatly facilitates the calculation of critical values. Monte Carlo experiments show that the test has good size and power under a wide range of circumstances. The test is then applied to investigate the effect of the Euro on trade.

**Keywords:** Structural change, end-of-sample instability tests, heterogeneous panels, Monte Carlo, Euro effect on trade.

**JEL Codes:** C23, C52

---

\*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. This paper was written while the author 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.

# 1 Introduction

This paper proposes a method to test for structural instability among only some individuals at the end of a sample in a panel regression model. Most tests for structural breaks in the literature, like the celebrated Chow (1960) tests, those for unknown or multiple break dates in Andrews (1993), Andrews and Ploberger (1994) and Bai and Perron (1998) are appropriate when the break is relatively long lasting and happens in the middle of a sample. 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. Yet, it is often at the end of a sample that researchers and policy-makers alike are interested in testing for instability.

Andrews (2003) proposed a test for structural break which was shown to be particularly useful when the number of post-break observations is small. Monte Carlo results suggested that his test has reasonable size and power even when the number of post break observation is 1. In addition, the critical values of the test statistics is calculated by using parametric sub-sampling method and therefore the test is robust to non-normal, heteroskedastic and serial 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 an alternative hypothesis that all individuals exhibit a break, as in other tests for structural breaks in the panel literature, like in Han and Park (1989) which extends the CUSUM tests, or Emerson and Kao (2001, 2002), Bai and Ng (2004) and de Wachter and Tzavalis (2004) which build on Andrews (1993) and Andrews and Ploberger (1994). Yet, the interesting question is how to test for the 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's (2003) end of sample stability test. In particular, this paper advances a standardized  $Z$  statistic built from Andrews (2003) statistics average across individuals. Methodologically, this is similar to the approach in Im, Pesaran, and Shin (2003) which, while focussing on the different question of unit root tests, also focusses on the average of separate statistics. This paper also derives the asymptotic distribution of the test statistics using the Lindeberg-Feller Central Limit Theorem (LF-CLT). As a result, the test statistics 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 occur at the end of a given sample. Lastly, the test covers the cases of parameter heterogeneity or homogeneity

pre and post instability.

A series of Monte Carlo experiments show that the proposed structural break test performs very well in finite sample. The experiments accommodate serial correlation in the error terms with a mixture of different distributions for the innovations. The results show that even under these circumstances the LF-CLT holds and the distribution of the test is close to a standardised normal. The Monte Carlo results indicate that the test has good size and power with relatively small number of time series and cross sections with moderate level of serial correlations. In addition, the performance of the test improves as time increases for high level of serial correlation. Lastly, the test has good power and size for partial instabilities of small magnitude.

Finally, this paper applies the test to investigate the effect of the Euro on intra-Eurozone trade. The question has been at the center of lively debates in academic and policy circles alike. Yet, the papers that have tackled the issue have not provided strong empirical evidence to support the existence of such effect. This is largely due to two empirical issues: the few datapoints available after the presumed break and the heterogeneity of the trade effect over different countries. Given both these characteristics, the test introduced in this paper is particularly well adapted. The results presented at the end of this paper show that indeed, there seems to have been a break in Eurozone trade starting in 1998, as hypothesize in the trade literature.

The paper is organised as follows. Section 2 build the basic framework and model with heterogeneous coefficients that serves as a base for the stability tests. Next, the panel data stability test for partial break is presented in section 3. Some important asymptotic results are derived in section 4. This is followed by some Monte Carlo results in section 5. Finally, section 6 illustrates how the test can be put to use, to answer the question of the Euro's effect on intra-Eurozone trade. The last section concludes.

## 2 The heterogeneous panel model

Consider the following model for panel data,

$$Y_{it} = \Theta_i' \mathbf{X}_{it} + U_{it} \quad i = 1, \dots, N, \quad t = 1, \dots, T, \quad (1)$$

where  $Y_{it}$  and  $U_{it}$  are the dependent variable and the idiosyncratic shocks specific to each individual and assumed to be independently distributed of  $\mathbf{X}_{it}$ , respectively.  $\mathbf{X}_{it}$  is the  $d \times 2$  vector of regressors that stacks all right hand side variables and  $\Theta_i'$  a  $d \times 2$  vector of stacked coefficients

$$\begin{aligned} \mathbf{X}_{it} &= [\mathbf{x}_{it} \quad \mathbf{y}_{it-1}] \\ &_{(d \times 2)} \\ \Theta_i' &= \begin{bmatrix} \beta_i' & \rho_i' \end{bmatrix} \\ &_{(d \times 2)} \end{aligned}$$

where the number of regressors  $d$  is defined as  $d = k + p$  for simplicity,  $\mathbf{y}_{it-1}$  is a  $p \times 1$  vector of lagged dependent variables,  $\mathbf{x}_{it}$  is the  $k \times 1$  vector of explanatory variables,  $\beta_i$  and  $\rho_i$  are a  $k \times 1$  and a  $p \times 1$  vector of coefficients respectively.

The final setting used in the heterogeneous stability tests can be specified following the nomenclature

$$Y_{it} = \begin{cases} \Theta'_{0i} \mathbf{X}_{it} + U_{it} & t = 1, \dots, T \\ \Theta'_{1i} \mathbf{X}_{it} + U_{it} & t = T + 1, \dots, T + m \end{cases} \quad (2)$$

for individuals  $i = 1, \dots, N$ , and where  $T$  is the presumed break date and  $m$  are the number of post-break observations. For simplicity,  $\bar{T}$  is defined as  $\bar{T} = T + m$ .  $\Theta_{0i}$  is the parameter vector before the break and  $\Theta_{1i}$  is the equivalent vector after the instability.

### 3 Heterogeneous panel data stability tests

#### 3.1 The hypotheses

The test relies on the same dynamic heterogeneous setting, namely in equation (2). The test's hypotheses are

$$\begin{aligned} H_0 : \Theta_{1i} &= \Theta_{0i} \quad \forall i = 1, \dots, N \text{ and } \forall t, \\ H_1 : \Theta_{1i} &\neq \Theta_{0i} \quad \exists i \in [1, \dots, N] \text{ and for } t > T \end{aligned}$$

with  $t = 1, \dots, T, T + 1, \dots, T + m$ . Let  $N = N_0 + N_1$ , where  $N_0$  is the number of individuals such that  $\Theta_{1i} = \Theta_{0i}$  and  $N_1$  is the total number of individual that have a break ( $\Theta_{1i} \neq \Theta_{0i}$ ). The null hypothesis states that there are no structural breaks across all  $N$  individuals, whereas the alternative states that at least one individual experiences a structural break. In the case of a homogeneous panel, one could take full advantage of the panel structure, by restricting  $\Theta_{0i}$  to  $\Theta_0$  for all  $i$ , while  $\Theta_{1i}$  would remain heterogeneous. This implies that the  $\Theta_0$  can be estimated consistently with either large  $N$  or large  $T$ , whereas in the case of heterogeneous panel the consistent of  $\Theta_{0i}$  will have to rely on large  $T$  only for all  $i$ . This former approach is used in the Monte Carlo experiments in section 5.

The test requires that the proportion of individuals who experience a break relative to  $N$  tends to a non-zero positive constant as  $N \rightarrow \infty$ , that is  $\lim_{N \rightarrow \infty} (\frac{N_1}{N}) = c$ , where  $0 < c \leq 1$  as in Im, Pesaran, and Shin (2003). This assumption is needed to ensure that the test statistic generated from the sample reflects that of the population. Therefore, it is important to note that when the test rejects the null hypothesis, it implies only a proportion  $N_1 < N$  of the individuals experience a structural break. Depending on the specific of a researcher, the proportion of individuals who experience a break can be estimated by conducting the Andrews (2003) test on each individual

separately. Using multiple Andrews tests, though, is not a replacement for the panel test, most straightforwardly because it only uses time series variation in its estimation, instead of benefitting from cross sectional movements as in the panel setting suggested. Furthermore, the asymptotic properties of the Andrews test do not warrant inference on the proportion of individuals exhibiting a break in the population, contrarily to this paper's panel test.

### 3.2 Panel data stability test

The test standardised  $Z$  statistic essentially amounts to comparing two average statistics,  $\bar{S}$ , taken from a pre-break subsample and the post break sample. The construction of the  $\bar{S}$  for both a pre-break subsample and the post-break sample require to compute the Andrews (2003) test statistic  $N$  times for each individual time series. The individual computation of the  $S$  statistic is identical to the initial time series end-of-sample instability test proposed by Andrews (2003). For  $m \geq d$ , the  $S_i$  statistics for each individual  $i$  is written as

$$S_i^0 = S_{i,1}^m = S_i \left( \hat{\Theta}_{i,1}^T, \hat{\Sigma}_{i,1}^{T+m} \right) \quad (3)$$

$$S_i^1 = S_{i,T+1}^{T+m} = S_i \left( \hat{\Theta}_{i,1}^{T+m}, \hat{\Sigma}_{i,1}^{T+m} \right) \quad (4)$$

for the a pre-break sample and the post-break sample respectively. The average statistic

$$\bar{S}^0 = N^{-1} \sum_{i=1}^N S_i^0 \quad (5)$$

$$\bar{S}^1 = N^{-1} \sum_{i=1}^N S_i^1 \quad (6)$$

the test statistics amount to summing each individual  $S$  statistic obtained from running the test on the separate time-series.  $\bar{S}^0$  is used to standardise the proposed test statistic as explained further on.

The individual test statistics,  $S_i^\nu$  for  $i = 1, \dots, N$  and  $\nu = 0, 1$  is computed as in Andrews (2003)

$$S_i(\Theta_i, \Sigma_i) = A_i(\Theta_i, \Sigma_i)' V_i^{-1}(\Sigma_i) A_i(\Theta_i, \Sigma_i), \quad (7)$$

$$A_i(\Theta_i, \Sigma_i) = \mathbf{X}_{i,k}^{\prime l} \Sigma_i^{-1} \left( \mathbf{Y}_{i,k}^l - \mathbf{X}_{i,k}^l \Theta_i \right), \quad (8)$$

$$V_i(\Sigma_i) = \mathbf{X}_{i,k}^{\prime l} \Sigma_i^{-1} \mathbf{X}_{i,k}^l \quad (9)$$

for all  $i = 1, \dots, N$  and where  $\mathbf{X}_{i,k}^l$  is the explanatory variables for all  $i$  starting from the time index  $k$  to  $l$ . For example, in order to calculate  $S_i^1$ , the post-break sample spanning from  $k = T + 1$  to  $l = T + m$  is used. The estimated time-series covariance matrix derived in Andrews (2003) are

used as a weight matrix, which estimates the individual  $i$ 's variances and autocovariances. The covariance matrix is

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

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

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

The coefficient vector  $\widehat{\Theta}_{i,1}^{T+m}$  is the least square estimates of  $\Theta$  for individual  $i$  over the full temporal sample. The covariance matrix is computed for each individual and used respectively in each individual test. Hence, the cross section covariance are not estimated. This test requires the assumption of independence across individuals. This does not limit the test in its practical application, as there are various methods to control for such cross section dependencies and thus satisfy the independency assumption.<sup>1</sup> However, the effects of such correction method on the performance of the proposed test is unknown and is an area for further research. Lastly, if  $m \leq d$ , the projection matrix collapses to a  $m \times m$  identity matrix and the  $S_i(\Theta_i, \Sigma_i)$  statistic becomes

$$P_i(\Theta_i, \Sigma) = \mathbf{U}_{i,k}^{\prime l} \Sigma_i^{-1} \mathbf{U}_{i,k}^l \quad (10)$$

The standardised test statistic can be written as

$$Z = \frac{(\bar{S}^1 - \bar{S}^0)}{\sqrt{\widehat{Var}(\bar{S}^1 - \bar{S}^0)}} \quad (11)$$

$\widehat{Var}(\bar{S}^1 - \bar{S}^0)$  is the estimated variance of the difference of the average statistics since variances of the individual statistics are unknown. The sample size used for  $\bar{S}^0$  is the same as  $\bar{S}^1$ . It is recommended to use the first  $m$  observations to estimate  $\bar{S}^0$  in order to maximise the distance between the subsamples allowing to minimise the potential impact of serial correlation in the errors. The sample size used to calculate the  $\bar{S}^0$  can be increased if needed, notwithstanding the potentially increasing problem of serial correlation. This is essentially an empirical issue, and any subsample selection bias related to  $\bar{S}^0$  should diminish as  $N$  increases, as discussed in the following section 4. The computation of the  $S$  statistics can be simplified in the case of homogeneous panel where  $\Theta_0$  can be estimated consistently by using the large cross sectional dimension.  $\Theta_0$  is estimated once to construct the test and does not need to be estimate for each cross section.

<sup>1</sup>It is possible to use the method introduced by Pesaran (2006a) to filter out the dependencies. Pesaran (2006b) proposes a modified ADF test for panel and a modified IPS test, which control also for cross-sectional dependence in the spirit of Pesaran (2006a).

## 4 Asymptotic results

### 4.1 Assumptions

This section provides the asymptotic properties of the proposed test. Define the data set as a sequence of random variables  $\{\mathbf{W}_{0i,t}\}$  where  $\{\mathbf{Y}_i, \mathbf{X}_i\} \subset \{\mathbf{W}_{0i,t}\}$ . Let  $B(\boldsymbol{\Theta}_{0i}, \epsilon_T)$  be a ball centered around  $\boldsymbol{\Theta}_{0i}$  with radius  $\epsilon_T > 0$  as in Andrews (2003). For  $m > d$ , the following assumptions for the  $S_i^\nu$ ,  $\nu = 0, 1$  are:

#### Assumption 1

- (a)  $\{\mathbf{W}_{0i,t} : t \geq 1\} \forall i$ , is stationary and ergodic under both  $H_0$  and  $H_1$ .  
 (b)  $E[\mathbf{W}'_{0i,t} \mathbf{W}_{0j,s}] = 0$ , for  $i \neq j$  and  $\forall s, t$ .

#### Assumption 2

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

#### Assumption 3

- (a)  $S_i(\boldsymbol{\Theta}_i, \boldsymbol{\Sigma}_i)$  is continuously differentiable in a neighbourhood of  $(\boldsymbol{\Theta}_{0i}, \boldsymbol{\Sigma}_{0i})$  with probability one under  $H_0$  and  $H_1$ , where  $\boldsymbol{\Sigma}_{0i}$  is as in assumption 2(b).  
 (b) Let  $(\partial/\partial(\boldsymbol{\Theta}_i, \boldsymbol{\Sigma}_i^{-1}))$  denote the partial differentiation with respect to  $\boldsymbol{\Theta}_i$  and the non redundant elements of  $\boldsymbol{\Sigma}_i^{-1}$ .  $S_i$  is bounded as

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

for some  $\epsilon_T > 0$ , where  $\boldsymbol{\Sigma}_{0i}$  is as in assumption 2(b).  $N(\boldsymbol{\Sigma}_{0i})$  denotes some neighbourhood of  $\boldsymbol{\Sigma}_{0i}$

- (c) The distribution function of  $S_{i,1}(\boldsymbol{\Theta}_{0i}, \boldsymbol{\Sigma}_{0i})$  is continuous and increasing at its  $1 - \alpha$  quantile, when  $m > d$ .

#### Assumption 4

- (a)  $E[U_{i1} \mathbf{X}_{i1}] = 0$ ,  $\forall i$ .  
 (b)  $E[U_{i1}^2] < \infty$  and  $E\|\mathbf{X}_{i1}\|^{2+\delta} < \infty$  for some  $\delta > 0$  and  $\forall i$ .  
 (c)  $E[\mathbf{X}_{i1} \mathbf{X}'_{i1}]$  and  $\boldsymbol{\Sigma}_0 = E[\mathbf{U}_{i,1}^m \mathbf{U}'_{i,1}^m]$  are positive definite,  $\forall i$ .

Assumptions (1) - (4) are identical to that of Andrews (2003). The first assumption allows for both weakly dependent processes and long memory processes, but also conditional variation in all moments, including conditional heteroskedasticity. Moreover, assumption 1 (b) requires that there are no cross sectional dependencies. Assumption 2 is required to ensure the consistency of the estimators for both the coefficient vector and the variance-covariance matrix. Assumption 3 is required to ensure that the empirical



distribution of the  $S$  statistic converges 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 4 is sufficient if the test statistic is carried out with the OLS estimator.

## 4.2 Results and comments

This sub-section derives the asymptotic distribution for the  $Z$  statistic and define the properties of the  $Z$  test. The following lemma is useful for reducing the number of assumptions required as well as extending the estimating procedure beyond least squares.

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

**Proof.** See Appendix

**Remark 1** Assumption 4 is made strictly for the Least Squares estimation procedure. Lemma 1 and other subsequent results in this paper will hold for other estimators such as IV or GMM as long as the conditions in assumption 4 have been modified accordingly. For the appropriate conditions in assumption 4 for IV or GMM see Andrews (2003).

**Lemma 2** *Let  $S_{i,\infty}^\nu$  be a random variable with the same distribution as  $S_i^\nu(\Theta_{0i}, \Sigma_{0i})$ ,  $\nu = 0, 1$ . Under assumptions 1-3 and Theorem in Andrews (2003), then as  $T \rightarrow \infty$ :*

- (a)  $S_i^\nu(\hat{\Theta}_i, \hat{\Sigma}_i) \xrightarrow{d} S_{i,\infty}^\nu, \quad \forall i = 1, \dots, N$  and  $\nu = 0, 1$   
(b) *Let  $F_i^\nu$  be the distribution of  $S_i^\nu \forall i$ , then  $F_i^\nu$  is a well defined distribution with finite mean and variance.*

**Proof.** See Appendix

**Lemma 3** *Under Lemma 2,*

$$\lim_{N \rightarrow \infty} (\bar{\sigma}^2 N)^{-1} \sum_{i=1}^N \int_{(S_i^\nu - E(S_i^\nu))^2 \geq \epsilon N \bar{\sigma}^2} ((S_i^\nu - E(S_i^\nu))^2) dF_i(S_i^\nu) = 0$$

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

**Proof.** See Appendix

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

$$\sqrt{n} \bar{S}^\nu \overset{A}{\approx} N(E(\bar{S}^\nu), \text{Var}(\bar{S}^\nu))$$

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

**Proof.** See Appendix

**Theorem 1** Under Lemma 4, the  $Z$  statistic as described in equation (11)

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

has an asymptotic distribution

$$\sqrt{n} Z \overset{A}{\approx} N(0, 1)$$

**Proof.** See Appendix

**Remark 2** Assumption 1 (b) requires cross sectional independence.

**Remark 3** Lemma 2 shows that each  $S_i^\nu$  converge to a well defined distribution with finite mean and variance. This is an important result as it allows us to show in Lemma 4, that the arithmetic average of  $S_i^\nu$  ( $\bar{S}^\nu$ ) has a normal distribution. This reduces substantially the computation burden required to calculate the empirical critical values as proposed by Andrews (2003) and in the previous chapter in this context. The cross sectional dimension presented in this test allow us to use the central limit theorem to derive the asymptotic distribution of the statistic, rather than computing the joint empirical distribution from each of the  $S_i^\nu$  statistics. The large cross sectional dimension allows to compare the  $\bar{S}^\nu$  statistic to a Normal distribution for which the critical values are known. Large cross sections are common in panel data setting, providing a good stage for using this simplified test.

**Remark 4** Lemma 3 shows that Assumptions 1 - 4 is sufficient to ensure for the Lindeberg condition required by the LF-CLT. This is particular important as the proposed statistics is the average of the S-statistics as proposed in Andrews (2003) from every individual in the panel and thus the assumptions imply that the variance of the proposed statistics is not dominated by the variance of the S-statistics from any particular individual.

**Remark 5** Although  $\bar{S}^\nu$  converge 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 the  $\bar{S}^\nu$  alone. Under the null hypothesis, however, the mean of  $\bar{S}^\nu$  is the same for  $\nu = 0, 1$  and therefore the  $Z$  statistic will have mean 0. Furthermore, the variance of  $\bar{S}^1 - \bar{S}^2$  can be estimated from  $S_i^\nu$  for  $i = 1, \dots, N$  and  $\nu = 0, 1$  directly. It ensues that the  $Z$  statistic will converge to a standard normal distribution in which valid inference can be obtained.

**Remark 6** Theorem 1 and Lemma 1 - 4 hold under assumptions 1 - 4 in the case when  $m \leq d$ , with  $S_i^\nu(\Theta_i, \Sigma_i)$  replaced by  $P_i^\nu(\Theta_i, \Sigma_i)$  as defined in equation (10).

## 5 Simulations

### 5.1 Monte Carlo design

The experiment uses the linear regression model as exposed in equation (1). The number of regressors is set to  $d = 5$ , which includes a constant, but does not include a lagged dependent variable. To start with, some benchmark results are generated in order to investigate the normality, the size and power of the test. 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  $\bar{T}$  sample using the coefficient vector  $\Theta_{1i} = \Theta_{0i} = 0, \forall i$ . Secondly, the power properties of the test are examined. The alternative hypothesis of partial instability is simulated, allowing for some individuals to experience a structural break while some do not. The ratio  $\frac{N_1}{N}$  is gradually changed from .10, .50, .65, .80 and 1, in order to allow a larger proportion of the individuals to experience a structural break. Furthermore, the alternative hypothesis featuring a partial structural break is simulated,  $\Theta_0 = 0, \Theta_{1i} = \frac{1}{10} \times (1, 1, 1, 1, 1)'$ , for some  $i$  and  $\Theta_{1j} = \Theta_0 = 0$ , for some  $j \neq i$ . The magnitude of the break is very small, equal to .1. Moreover, for all the results we use  $\|\Theta_{1i}\| = \sqrt{.05}$ , 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 and  $\Theta_0 = \Theta_{0i} = 0$  prior to the instability,  $\forall i$  and  $\Theta_1 = \Theta_{1i} = \frac{1}{10} \times (1, 1, 1, 1, 1)'$  after the instability,  $\forall i$ .

For the distribution property, size and power of the test, the Monte Carlo experiments are conducted with the following settings:  $m = 10, d = 5, \bar{T} = 50$  and  $100, N = 20, 40, 60, 80$  and  $100$ , where  $\bar{T} = T + m$ . All the regressors are calculated as a trigonometric function of a set of random normal variable, which are independent and identical. The regression's error term is generated with an AR(1) process with the following autoregressive parameters:  $\rho = .4$  and  $.95$  which is homogeneous across individuals, or in other words, all individuals' errors have the same  $\rho$ . Four different types of *iid* distributions for the innovation of the error term are considered: standard  $N(0, 1)$ , a recentered and rescaled  $\chi_d^2$  and  $t_5$  with mean zero and variance one, and an uniform distribution with support  $[0, 1]$ . The different individuals have different innovation processes, such that the four distributions are intermixed evenly in the panel. The number of replications equals 2000. Note that the alternative hypothesis is fixed for all individuals, implying that we test for a break occurring at a known fixed date. All simulations were carried out using Ox 4.02.<sup>2</sup>

---

<sup>2</sup>The programming code is available upon request.

## 5.2 Monte Carlo results

### 5.2.1 Size

The first results look at the probabilities of a type I error with normal significance level of .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 hold, under moderate serial correlation and relatively small time dimension. The results are shown in table 1
2. The size of the test is relatively close to the desired value of .05 for  $N > 20$  when  $\bar{T} = 100$  and  $m = 10$  with serial correlation is moderate,  $\rho = .4$ . As shown in in table 2, the test has reasonable size when the time horizon is decreased to  $\bar{T} = 50$  and  $\bar{T} = 30$  given the DGP.
3. The normality of the distribution worsen in the presence of extreme serial correlation. The mean grows substantially, the distribution is skewed and the variance shrinks below 1, as expected. The size, on the other hand, deteriorates as the number of individuals increase such that its largest value of .199 is attained for  $N = 100$  and  $\bar{T} = 100$ . This result is expected as all individuals are required to exhibit the same high degree of serial correlation, and the problem is only aggravated as the sample feature an increasing number of individuals facing the same problem, much like an epidemic. Moreover, increasing  $\bar{T}$  from 100 to 250 observations or more improves the size, as implied by ergodicity.<sup>3</sup>

In sum, the test has reasonable size in small temporal and individual sample with moderate serial correlation. However, under extreme serial correlation the size of the test deteriorates substantially.

### 5.2.2 Power

Overall the test has good power. The power of the test is analysed for a normal significance level of .05. The most important results of the Monte Carlo experiments are as follows:

1. The power of the test varies little with  $\bar{T}$  except when it is very small ( $\bar{T} \leq 50$ ). This underlines the advantage gained by working with a panel structure. The test remains powerful even if  $\bar{T}$  is decreased to 50: when  $c = .80$  and  $N = 60$ , the power is .72.
2. The the test gains power as either  $N$  or  $c$  increases. The power of the test is above .90 when  $N$  and  $c$  are high. The power is still good when both  $c$  and  $N$  are of medium size; for example when  $c = .65$  and

---

<sup>3</sup>These specific results are available upon request.

$N = 60$  the power is .85. Moreover, the power of the test is good when  $N$  is high (100) and  $c$  is low (.50) with  $\bar{T} = 100$ . The reverse is also true: when  $N = 40$  and  $c = .80$ , the power is .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 .95, the test has power of .71 when  $N = 100$  and  $c = .80$ .

Overall, the test seems quite powerful given the DGP when serial correlation is moderate. Lastly, the power of test increases as the magnitude of the break increases. The Monte Carlo experiments in this case is available upon request.

## 6 Empirical Example

This section provides an empirical application to demonstrate the usefulness of the test. Recently, the question of whether the Euro has increased intra-Eurozone trade has captivated an important strand of the empirical trade literature and is paramount in policy debates. However, empirical evidence has been limited by scarce data and not entirely appropriate econometric techniques. Most papers in the literature have used F-tests to estimate the significance of dummy variables on the Euro, thereby relying on the small sample assumptions of normally, independently and identically (NIID) distributed error, which were not easy to verify. On the contrary, the test proposed in this paper is robust to these problems as it is especially well adapted to few datapoints after the break, and, importantly, it also addresses the very likely possibility that the Euro affected each country differently.

Since the publication of Rose (2000), the question on a common currency's effect on trade has been a focus of the trade literature. Indeed, Rose's seemingly indisputable finding that a common currency caused more than a doubling of trade naturally caused controversy. While a flurry of different empirical methodologies failed to disprove the finding, although often finding a diminished effect, the peculiar nature of the data - emanating from poor, open, remote island states - seemed to be at the origin of the effect. The introduction of the Euro was therefore seen by many as a closer to natural experiment, importantly concerning western economies, to test the celebrated effect. If not in trade, there immediately was a boom in papers attempting to evaluate the question. These include, more prominently, Micco, Ordoñez, and Stein (2003) and Flam and Nordström (2003), but also Nitsch (2002), Bun and Klaassen (2002), De Sousa (2002), Barr, Breedon, and Miles (2003), De Nardis and Vicarelli (2003), Piscitelli (2003), Nitsch and Berger (2005), Baldwin (2006), Baldwin (2006) and Mancini-Griffoli and Pauwels (2006).<sup>4</sup> On the whole, these mostly find a positive effect starting

---

<sup>4</sup>For a summary see Baldwin (2006), or Mancini-Griffoli and Pauwels (2006).

between 1998 and 1999 of the order of 10 to 20%. Only Mancini-Griffoli and Pauwels (2006), by applying a modification of the Andrews (2003) test to homogeneous panel data (considering the alternative of a common effect of the Euro across all trading partners) shows that the effect is significant, and refines the conclusion by showing that the break first occurred in the growth rate of trade around 1998, and only had noticeable repercussions on the level of trade around 2002.

This section extends the findings of Mancini-Griffoli and Pauwels (2006) to test for a break in the trade between only some of the Eurozone countries. Indeed, while Mancini-Griffoli and Pauwels (2006) can be seen as providing a methodological advancement with respect to the existing literature, the paper is limited by the assumption that all individuals (unilateral trading partners) exhibit the same break in trade due to the Euro. But intuition tells us the contrary. For instance, while it was clear that Germany was going to play a central role in the Euro project from its inception, it was uncertain whether Italy would meet the strict accession requirements 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. Thus, the particular test developed in this paper appears especially fitting.

Mancini-Griffoli and Pauwels (2006) discuss the microfoundations of the appropriate gravity-type trade regression equation, as well as the distinct estimation methods suitable for the particularities of the data. Notably, the test introduced in this paper, like the original Andrews (2003) test, calls for stationary and ergodic data prior to the presumed break point. This is found not to be the case in the data which does not reject the null of a unit root in each of its variables. Mancini-Griffoli and Pauwels (2006), where these results are presented, suggests one solution, namely to estimate the regression equation as an ECM, interpreted as the correlation between the growth rates of variables (corrected for over-differencing by the cointegrating vector). Other solutions to the problem of non-stationary data are also considered in Mancini-Griffoli and Pauwels (2006), but are not adopted here, as this section limits itself to a mere illustration of the new test for instability in panels. In summary, the regression used here is (in two stages, as is typical for an ECM):

$$\begin{aligned}
V_{i,jt} &= \alpha_{i,j} + \lambda_t + \varphi_j + \gamma_1 Y_{it} + \gamma_2 Y_{jt} + \gamma_3 \xi_{i,jt} + \epsilon_{it} \\
\Delta V_{i,jt} &= \rho \hat{\epsilon}_{it-1} + \delta_1 \Delta Y_{it} + \delta_2 \Delta Y_{jt} + \delta_3 \Delta \xi_{i,jt} + \\
&\quad + \gamma_1 \Delta V_{\bullet,t} + \gamma_2 \Delta Y_{\bullet,t} + \nu_{it}
\end{aligned} \tag{12}$$

where  $V_{i,jt}$  is the value of imports from country  $j$  to country  $i$ ,  $Y_{jt}$  and  $Y_{it}$  are nominal GDP,  $\xi_{i,jt}$  is the real exchange rate between the two countries engaged in trade,  $\epsilon_{it}$  is a regression error,  $\alpha_{i,j}$  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,  $\lambda_t$  allows for the country pair intercept to be time dependent, and  $\varphi_j$  is a country of origin dummy as used in Rose and van Wincoop (2001).

The model is estimated using difference from sample mean fixed effects and the second using pooled OLS and where the  $\bullet$  notation indicates sample average over the subscript it replaces. These additional variables are included to remove cross sectional dependence in the errors, as described in Pesaran (2006a) as the common correlated effect (CCE) estimator.<sup>5</sup> The goal is to control for the common factor causing cross correlations, but because it is unknown, Pesaran suggests it can be proxied by a linear combination of the sample averages of the regressors and regressand. Mancini-Griffoli and Pauwels (2006) show that using the additional CCE terms do indeed get rid of cross sectional dependence and prove that this condition allows for the inversion of the covariance matrix in the Andrews (2003) test statistic.

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

Recall that the  $Z$  test statistic in this paper is standardised with the  $\bar{S}^0$  average statistic calculated within the pre-instability sample. This paper suggested to estimate the latter from the very beginning of the sample. The Euro data set, however, is not very reliable over the first 4 years (16 quarters) of data, as several series are extrapolated using moving averages from yearly data to fill in some missing observations. Hence, to overcome some of these shortcomings, the benchmark  $\bar{S}^0$  is found using data starting from 1985 Q1. This still ensures a large period between the pre-break subsampling and the post-break observations thereby minimizing disturbances due to serial correlation. The particular choice of date to anchor the  $\bar{S}^0$  statistic will be tested for robustness.

Results are as shown in table 4. Several important comments are called for. First, recall that Mancini-Griffoli and Pauwels (2006) finds a break in the growth of trade (in the ECM) in 1998Q1 at the 10% significance level and that this break appears to last 6 quarters. Note that the break date is defined as the first quarter for which the null of stability is rejected for 6 quarters with at least 10% significance. Thus, it is reassuring to note that the break date seems to be robust to the alternative hypothesis of a heterogeneous break. Indeed, the null of stability is rejected, this time at the 1% level, for a break in 1998Q1 and lasting 6 quarters. Second, the break appears to last longer with this paper's test - up to 12 quarters (or

---

<sup>5</sup>This estimator is increasingly used, due to its effectiveness but also computational simplicity, as opposed to other solutions based on estimating common factor loadings such as Bai and Ng (2002). An elegant example of its implementation, as well as a clear explanation of its properties, can be found in Imbs, Mumtaz, Ravn, and Rey (2005).

3 years). Note that the length of a break is found by fixing the break date and repeating the test while adding on quarter to the post break sample period with each iteration. These results are also as expected. As different Eurozone countries reacted differently to the Euro (some exhibiting a break and some not, or perhaps a much shorter one), the alternative hypothesis of a common break across the board is restrictive and probably only fitting for a few quarters. On the contrary, the more accommodating and realistic alternative of a heterogeneous break is accepted for a longer time period. Fourth, it is encouraging to note that the test results are barely sensitive to the choice of pre-break sampling date. This is as argued in the paper and stands given the stability and ergodicity of the pre-break data. In summary, it appears that indeed, the introduction of the Euro did have a noticeable impact on intra Eurozone trade, as anticipated by the original Rose (2000) hypothesis. What, exactly, in the new currency caused this rise in trade, is another question well worth considering in other 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.

## 7 Concluding remarks

This paper built a stability test for heterogeneous panel data in the light of the IPS test for unit roots, to test the null of stability for all cross sections versus the alternative that some cross sections experience the instability and some do not. 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. The results show that the test performs relatively well in the presence of serial correlation in the errors and that the results can be improved by increasing the time dimension. These results allow the test to be used widely in finance and economics applications. This paper explored one particular example, showing the existence of a trade effect of the Euro's introduction.

## References

- ANDREWS, D. W. K. (1993): "Tests for Parameter Instability and Structural Change With Unknown Change Point," *Econometrica*, 61(4), 821–856.
- (2003): "End-of-Sample Instability Tests," *Econometrica*, 71(6), 1661–1694.



- ANDREWS, D. W. K., AND W. PLOBERGER (1994): “Optimal tests when a nuisance parameter is present only under the alternative,” *Econometrica*, 62, 1383–1414.
- BAI, J., AND S. NG (2002): “Determining the Number of Factors in Approximate Factor Models,” *Econometrica*, 70(11), 191–221.
- (2004): “A Panic Attack on Unit Roots and Cointegration,” *Econometrica*, 72(4), 1127–1177.
- 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.
- CHOW, G. C. (1960): “Tests of Equality Between Sets of Coefficient in Two Linear Regressions,” *Econometrica*, 28, 591–605.
- DE NARDIS, S., AND C. VICARELLI (2003): “Currency Unions and Trade: The Special Case of EMU,” *World Review of Economics*, 139(4), 625–649.
- 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.
- (2002): “Testing for Structural Change of a Time Trend Regression in Panel Data: Part II,” *Journal of Propagations in Probability and Statistics*, 2, 207–250.
- 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, 135–142.
- IM, K. S., H. PESARAN, AND Y. SHIN (2003): “Testing for Unit Roots in Heterogeneous Panels,” *Journal of Econometrics*, 115, 53–74.
- IMBS, J., H. MUMTAZ, M. RAVN, AND H. REY (2005): “PPP strikes back: aggregation and the real exchange rate,” *Quarterly Journal of Economics*, 120(1), 1–43.
- 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), 316–356.
- NITSCH, V. (2002): “Honey, I Shrunk the Currency Union Effect on Trade,” *The World Economy*, 25, 457–474.
- 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. (2006a): “Estimation and inference in large heterogeneous panels with a multifactor error structure,” *Econometrica*, 74(4), 967–1012.
- (2006b): “A Simple Panel Unit Root Test in the Presence of Cross Section Dependence,” *Journal of Applied Econometrics*, Forthcoming.
- 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.
- ROSE, A. K., AND E. VAN WINCOOP (2001): “National Money as a Barrier to International Trade: The Real Case for Currency Union,” *American Economic Review, Papers and Proceedings*, 91(2), 386–390.

## Appendix

**Proof of Lemma 1** *Since  $S_i^y \forall i$  is calculated by treating each cross section as a univariate time-series, the properties of  $S_i^y$  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)  $S_i^\nu$  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*

$$\lim_{N \rightarrow \infty} (\bar{\sigma}_\nu^2 N)^{-1} \int_{(S_i^\nu - E(S_i^\nu))^2 \geq \varepsilon N \bar{\sigma}_\nu^2} ((S_i^\nu - E(S_i^\nu))^2 dF_i(S_i^\nu)) = 0,$$

*under the null hypothesis, where  $\bar{\sigma}_\nu^2 = \sum_{i=1}^N \text{Var}(S_i^\nu)$ ,  $\forall i = 1, \dots, N$ ,  $\nu = 0, 1$ . Let  $\sigma_i^\nu = \text{Var}(S_i^\nu)$ , then under Lemma (1),  $0 < \sigma_i^\nu < \infty \forall i = 1, \dots, N$ ,  $\nu = 0, 1$ . That is,*

$$\sigma_i^\nu = \int_{\mathbb{R}_+} ((S_i^\nu - E(S_i^\nu))^2 dF_i(S_i^\nu)) < \infty.$$

*Let  $A$  be the set such that  $A = \{S_i^\nu : (S_i^\nu - E(S_i^\nu))^2 \geq \varepsilon N \bar{\sigma}_\nu^2\}$ . Since  $A \subset \mathbb{R}_+$ ,  $((S_i^\nu - E(S_i^\nu))^2 \geq 0$  and  $F_i(S_i^\nu) \in [0, 1]$ ,  $\forall i = 1, \dots, N$ ,  $\nu = 0, 1$  therefore*

$$0 \leq \int_A ((S_i^\nu - E(S_i^\nu))^2 dF_i(S_i^\nu)) \leq \int_{\mathbb{R}_+} ((S_i^\nu - E(S_i^\nu))^2 dF_i(S_i^\nu)).$$

*Hence,*

$$0 \leq (\bar{\sigma}_\nu^2 N)^{-1} \int_A ((S_i^\nu - E(S_i^\nu))^2 dF_i(S_i^\nu)) \leq \frac{\sigma_i^\nu}{\bar{\sigma}_\nu^2 N}.$$

*Now, as  $N \rightarrow \infty$ ,  $\bar{\sigma}_\nu^2 \rightarrow \infty$  but  $\sigma_i^\nu$  remains a fixed constant  $\forall i = 1, \dots, N$ ,  $\nu = 0, 1$ , thus*

$$\lim_{N \rightarrow \infty} \frac{\sigma_i^\nu}{\bar{\sigma}_\nu^2 N} = 0$$

*therefore*

$$\lim_{N \rightarrow \infty} (\bar{\sigma}_\nu^2 N)^{-1} \int_A ((S_i^\nu - E(S_i^\nu))^2 dF_i(S_i^\nu)) = 0$$

*$\forall i = 1, \dots, N$ ,  $\nu = 0, 1$ . This completes the proof. ■*

**Proof of Lemma 4** *Under Lemma 2 and 3, it is straightforward to show that as  $S_i^\nu$  is independent of  $S_j^\nu$ , for  $i \neq j$ , satisfying the conditions required by the Lindeberg-Feller CLT and therefore*

$$\sqrt{n} \bar{S}^\nu \stackrel{A}{\approx} N(E(\bar{S}^\nu), \text{Var}(\bar{S}^\nu))$$

*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  $E[\bar{S}^0] = E[\bar{S}^1]$  and hence  $Z$  converges in probability to a  $N(0, 1)$  distribution. This completes the proof. ■

Table 1: Moments for the distribution of the test under the null

Moments	$\bar{T}$	$\rho$	$N$				
			20	40	60	80	100
Mean	30	.4	-.005	.028	.036	.023	.003
	50	.4	-.015	-.059	-.064	-.052	-.021
	100	.4	-.057	-.051	-.112	-.063	-.117
	100	.95	-.539	-.746	-.936	-1.06	-1.24
Variance	30	.4	1.19	1.09	1.06	1.03	1.02
	50	.4	1.17	1.07	1.06	1.01	1.01
	100	.4	1.11	1.03	1.01	1.04	1.11
	100	.95	.907	.914	.834	.758	.735
Skewness	30	.4	.015	-0.07	.028	.003	.067
	50	.4	.071	.012	-.028	-.038	-.045
	100	.4	.016	-.020	-.004	-.040	.066
	100	.95	.471	.397	.288	.394	.262
Kurtosis	30	.4	3.04	2.97	3.15	2.97	3.11
	50	.4	3.37	2.89	3.13	3.10	2.81
	100	.4	3.07	3.07	3.16	2.95	2.71
	100	.95	3.11	3.05	2.98	3.52	3.11

Note:  $\rho$  is the autocorrelation coefficient,  $N$  are the total number of individuals and  $\bar{T} = T + m$  where  $T$  is the time dimension prior to the instability,  $m$  is the number of observations post instability and  $m = 10$ .

Table 2: Size of normal significance level .05

$\bar{T}$	$\rho$	$N$				
		20	40	60	80	100
30	.4	.073	.061	.057	.049	.052
	.95	.055	.092	.130	.142	.199
50	.4	.067	.061	.062	.048	.053
	.95	.065	.047	.055	.053	.059
100	.4	.065	.047	.055	.053	.059
	.95	.055	.092	.130	.142	.199

Note:  $\rho$  is the autocorrelation coefficient,  $N$  are the total number of individuals and  $\bar{T} = T + m$  where  $T$  is the time dimension prior to the instability,  $m$  is the number of observations post instability and  $m = 10$ .

Table 3: Power of normal significance level .05 for  $\frac{N_1}{N} = c$

$\bar{T}$	$c$	$\rho$	$N$				
			20	40	60	80	100
30	.80	.4	.019	.036	.060	.082	.136
	1	.4	.054	.127	.238	.355	.480
50	.50	.4	.016	.024	.052	.097	.155
	.80	.4	.134	.433	.720	.898	.965
	1	.4	.454	.869	.984	1.00	1.00
100	.10	.4	.000	.003	.004	.005	.006
	.50	.4	.022	.114	.294	.523	.696
	.65	.4	.095	.520	.856	.967	.991
	.80	.4	.347	.911	.991	.999	1.00
	1	.4	.852	.994	1.00	1.00	1.00
100	.50	.95	.008	.011	.011	.012	.014
	.80	.95	.075	.203	.368	.542	.705
	1	.95	.270	.620	.869	.967	.994

Note:  $\rho$  is the autocorrelation coefficient,  $N$  are the total number of individuals and  $\bar{T} = T + m$  where  $T$  is the time dimension prior to the instability,  $m$  is the number of observations post instability and  $m = 10$ .

Table 4: ECM model: Partial Instability (EA-EA data)

Pre-break sampling date	Break date	Break duration (quarters)	Critical value	p-value
1985 Q1	1998 Q1	6	2.84	<b>0.00</b>
		12	2.18	<b>0.03</b>
		13	1.22	0.22
		14	1.00	0.31
1987 Q1	1998 Q1	6	3.01	<b>0.00</b>
		12	2.01	<b>0.04</b>
		13	1.85	<b>0.06</b>
		14	1.43	0.15
1990 Q1	1998 Q1	6	3.78	<b>0.00</b>
		12	3.15	<b>0.00</b>
		13	1.62	0.11
		14	0.61	0.53