Journal of Time Series Econometrics

Volume 4, *Issue* 2 2012 *Article* 3

Testing for Structural Change in Heterogeneous Panels with an Application to the Euro's Trade Effect

Laurent L. Pauwels, *University of Sydney*
Felix Chan, *Curtin University*
Tommaso Mancini Griffoli, *Swiss National Bank*

Recommended Citation:
Pauwels, Laurent L.; Chan, Felix; and Mancini Griffoli, Tommaso (2012) "Testing for Structural Change in Heterogeneous Panels with an Application to the Euro's Trade Effect," *Journal of Time Series Econometrics*: Vol. 4: Iss. 2, Article 3.
DOI: 10.1515/1941-1928.1141

Testing for Structural Change in Heterogeneous Panels with an Application to the Euro's Trade Effect

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

Abstract

This paper presents a structural change test for panel data models in which the break (or the change) affects some, but not all, cross-section units in the panel. The test is robust to non-normal, heteroskedastic and autocorrelated errors, as well as end-of-sample structural change. The test amounts to computing and comparing pre- and post-break sample statistics as Chow (1960) type F statistics averaged over cross-section units. The cases of known and unknown break date are both considered. Under mild assumptions, the test has a limiting standard normal distribution as the number of cross-sections tends to infinity. Monte Carlo experiments show that the test has good size and power under a wide range of circumstances, including when the break date is unknown and differs across individual units, and when errors exhibit cross-section dependence. Finally, the test is illustrated by seeking a break in the dynamics of trade among euro area countries following the introduction of the euro.

KEYWORDS: structural break, parameter stability, cross-section dependence, common correlated effects, gravity model, euro effect on trade

Author Notes: 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, Vasilis Sarafidis 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 the first 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 second author would like to acknowledge the financial support from the Australian Research Council.

1 Introduction

There is a vast literature on testing for parameter instability or structural change in regression models. Most tests for structural change have been developed for univariate time-series models, like the popular F test of Chow (1960). This F test is modified for the cases of unknown and multiple change point or break dates in Andrews (1993), Andrews and Ploberger (1994) and Bai and Perron (1998) among others. The distribution of most of these tests is usually found using asymptotics in which the number of pre- and post-break observations goes to infinity, a requirement that does not allow for few post-break observations. Dufour et al. (1994) considers this problem and proposes several methods to obtain critical values. Each of these methods present some limitations. These limitations are discussed in Andrews (2003) which suggests an S test statistic for which the critical values are calculated using parametric subsampling methods to make the test robust to non-normal, heteroskedastic and serially correlated errors.

Unlike the time-series literature, the longitudinal or panel data literature counts very few tests for structural change. One of the earliest contributions in this literature is Han and Park (1989) which extends the traditional CUSUM-type tests to the context of panel data models. Emerson and Kao (2001, 2002) extend the maximum Wald test for a structural change of Andrews (1993) and Andrews and Ploberger (1994). Another extension of the maximum Wald test is presented in Kao et al. (2005) in the context of a cointegrated panel data model with common factors. De Wachter and Tzavalis (2004) builds a test for structural change based on Andrews (1993) specifically for an AR(1) dynamic panel data model. Such a model relies on the Arellano and Bond (1991) GMM estimator for which the cross-section dimension (N) is large and the time series dimension (T) is fixed. For a known breakpoint, the test-statistic is asymptotically χ^2 distributed with degrees of freedom depending on the model used. For an unknown breakpoint, the quantiles of the asymptotic distribution need to be computed through simulation. The literature has also investigated extensively panel stationarity tests with structural change, some notable contributions include Carrion-i-Silvestre et al. (2005) and a recent application by Basher and Carrion-i-Silvestre (2009).

To the best of the authors knowledge, all the relevant tests in the panel data literature assume that the alternative hypothesis of a structural change is homogenous for all i. The interesting alternative that only some – and not all – individual units are affected by a break is likely to be more prominent in applied work, as shocks rarely affect all individual units equally, if at all. It is also worthwhile noting that Andrews (1993), Emerson and Kao (2001, 2002), De Wachter and Tzavalis (2004) for example, require the post-break observations to be at least 15% of T, while often in empirical work the interest lies in detecting breaks occurring at the end of the sample. Moreover, Andrews' (1993) supremum tests for time series or its panel extension by Emerson and Kao (2001, 2002) necessitates tabulated critical values, Dufour et al. (1994) requires either distributional assumptions or the use of semi-nonparametric density estimation methods, and Andrews (2003) S test utilises parametric subsampling methods.

This paper introduces a Z test for structural change in panel data models, for which the break is sustained by some, but not all, cross-section units $i = 1, \ldots, N$ in the panel. This break can also occur at different dates for each i. The basic idea of the Z test is to compare cross-sectional average statistics from the pre- and post-break samples. For each sample, a cross-sectional average statistic is computed by averaging sum of squared residuals-type statistics (S_i) over N. These S_i statistics are calculated in a similar fashion to Chow's univariate F test and Andrews' (2003) S test. The Z test is then calculated as the scaled difference between the pre- and post-break cross-sectional average statistics. Methodologically, this is most similar to the approach in Im et al. (2003) which, while focussing on the different question of unit root tests, also considers an average of separate statistics.

The Z test follows a standard normal distribution as $N \to \infty$, which greatly facilitates the calculation of the critical values required to conduct inference. Unlike Chow's F test, the errors of the linear regression model need not be normal. It is only assumed that the errors have finite second moments. Furthermore, calculation of asymptotic critical values through simulations are not required. As Andrews' univariate S test, the Z test is robust to non-normal, heteroskedastic and autocorrelated errors, and breaks of short duration such as in end-of-sample structural change. Section 5 reviews simulation results showing good size and power of this paper's proposed test when post-break observations are only equal to 10% of T.

The Z test presented in this paper is applicable to the cases of known and unknown break date. When the break date is unknown, a 'supremum' Z statistic (sup Z) can be used in the spirit of Andrews (1993) and Bai and Perron (1998). The break date is found by maximising difference between the post- and pre-break sum of squared residuals statistics for each cross-section unit. The theoretical results for the Z statistic with a known break date extend to the case of unknown break date.

Overall, this paper contributes to the existing literature on structural change in panel data models in three ways: (i) it proposes a Z test for structural change in panel data models for which the break does not need to occur at all or with the same date for all i; (ii) it shows that, under a mild set of assumptions, the Z test has a limiting standard normal distribution as the cross-section dimension tends to infinity ($N \to \infty$), which greatly facilitates inference; (iii) it presents a sup Z test for structural change when the break date is unknown with theoretical properties similar to the Z test. Moreover, the tests are extended to accommodate cross-section dependence in the errors. Two approaches to "filter out" cross-sectionally dependent errors are investigated, namely the Common Correlated Effects estimator (CCE) proposed in Pesaran (2006) and the Principal Components estimator by Bai and Ng (2002), Bai (2003, 2005), and Pesaran and Kapetanios (2005).

The paper is organised as follows: Section 2 introduces the Z test for structural change in panel data models. Section 3 provides a discussion of the assumptions required for the asymptotic results to hold, and a derivation of relevant asymptotic results. Section 4 discusses how the tests can accommodate cross-sectionally dependent errors. Section 5 investigates the test's finite sample properties with Monte Carlo simulations. Simulation results are also provided when the break date is unknown and when cross-section dependent errors are filtered with CCE. The Monte Carlo simulations show that the proposed structural change test has good size and power in finite sample, and the distribution of the test is close to a standardised normal. Finally, section 6 illustrates the test by investigating a break in intra–euro area trade dynamics after the introduction of the euro, with only a few years of data following the postulated break point.

2 Structural change test

2.1 Setup

Consider the following linear heterogenous panel data (or longitudinal data) model

$$y_{it} = \Theta_i^{'}\mathbf{x}_{it} + u_{it}, \quad i = 1,2,\dots,N,\ t = 1,2,\dots,T \tag{1}$$

where y_{it} is the observation on the dependent variable for cross-section unit i at time t, $\mathbf{x}_{it} = (x_{it}^{(1)},\dots,x_{it}^{(d)})'$ is the $d \times 1$ vector of explanatory variables including intercepts and/or seasonal dummies, Θ_i is the $d \times 1$ vector of coefficients. Moreover, u_{it} are the idiosyncratic shocks specific to each individual unit and may contain fixed or random effects. u_{it} may also exhibit cross-section dependence, heteroskedasticity and autocorrelation. Heteroskedasticity and autocorrelation are explicitly taken into account in the derivation of the test statistic. Some of the forms and remedies for cross-section dependence are discussed in section 4.

The notation used throughout this paper is defined as follows. Let $t = 1,\dots,T$ with $T_{0i} \in (1,T)$, where $\{1,\dots,T_{0i}\}$ is the pre-break time span and $\{T_{0i}+1,\dots,T\}$ is post-break time span. T_{0i} is the date of the structural change and $T_{1i} = T - T_{0i}$ is the number of post-break observations. Let $\mathbf{Y}_i = (y_{i1}, y_{i2}, \dots, y_{iT})'$

be a $T \times 1$ vector and

$$\mathbf{X}_i = \begin{pmatrix} x_{i1}^{(1)} & x_{i1}^{(2)} & \dots & x_{i1}^{(d)} \\ x_{i2}^{(1)} & x_{i2}^{(2)} & \dots & x_{i2}^{(d)} \\ \vdots & \vdots & \dots & \vdots \\ x_{iT}^{(1)} & x_{iT}^{(2)} & \dots & x_{iT}^{(d)} \end{pmatrix}$$

be a $T \times d$ matrix for all i. $\mathbf{Y}_i$ contains the values of the endogenous variable and $\mathbf{X}_i$ contains the values of all the explanatory variables for the individual unit i.

The model (1) can be re-written in matrix form

$$\mathbf{Y}_i = \mathbf{X}_i \Theta_{0i} + \begin{pmatrix} \mathbf{0}_{T_{0i}} \\ \mathbf{X}_{iT_{1i}} \end{pmatrix} \delta_i + \mathbf{U}_i \tag{2}$$

where $\mathbf{0}_{T_{0i}}$ is a $T_{0i} \times d$ null matrix, $\mathbf{X}_{iT_{1i}}$ is the $T_{1i} \times d$ matrix of explanatory variables and $\mathbf{U}_i = (u_{i1}, u_{i2}, \dots, u_{iT})'$ is a $T \times 1$ vector containing the errors for i. Θ_{0i} is the parameter vector prevailing for the pre-break sample and δ_i is the difference between the pre- and post-break parameter vectors.

If $\delta_i = \mathbf{0}$ then it implies that Θ_{0i} is constant for all t and i in model (2). Thus, the null hypothesis of no structural change is,

$$\begin{aligned} H_0 : \delta_i = \mathbf{0} \quad & \text{for all } i = 1, \dots, N \text{ and } t = 1, \dots, T_{0i}, \dots, T \\ H_1 : \delta_i \neq \mathbf{0} \quad & \text{for some } i \text{ and for } t > T_{0i}. \end{aligned}$$

Let $N = N_0 + N_1$, where N_0 is the number of individual units for which $\delta_i = \mathbf{0}$ and N_1 is the number of individual units that exhibit a break ($\delta_i \neq \mathbf{0}$). The null hypothesis states that there is no structural change for all i, whereas the alternative states that a proportion of individual units (N_1) experience a structural change. The alternative requires that this proportion relative to N tends to a non-zero positive constant as $N \rightarrow \infty$. Mathematically, this implies $\lim_{N\rightarrow\infty}(\frac{N_1}{N}) = c$, 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.

2.2 Z test with known break date

This paper's proposed test for structural change essentially amounts to comparing two cross-sectional average statistics, one calculated from the pre-break sample and the other from the post-break sample. These two statistics result from averaging sum of squared residuals type statistics (S_i) over N. Similarly to Chow's (1960) F

test and Andrews' (2003) S test, the fundamental test statistic for each individual unit i is defined to be

$$S_i(\Theta_i, \Sigma_i) = A_i(\Theta_i, \Sigma_i)' [V_i(\Sigma_i)]^{-1} A_i(\Theta_i, \Sigma_i), \tag{3}$$

$$A_i(\Theta_i, \Sigma_i) = \mathbf{X}_i' \Sigma_i^{-1} (\mathbf{Y}_i - \mathbf{X}_i \Theta_i), \tag{4}$$

$$V_i(\Sigma_i) = \mathbf{X}_i' \Sigma_i^{-1} \mathbf{X}_i \tag{5}$$

for all $i = 1, \ldots, N$, with the assumption that $\Sigma_i^{-\frac{1}{2}}$ exists and is positive definite so that $\Sigma_i^{\frac{1}{2}} \Sigma_i^{\frac{1}{2}} = \Sigma_i$ and hence $Var(\Sigma_i^{-\frac{1}{2}} \mathbf{U}_i) = \sigma^2 \mathbf{I}$. The power of the test is also improved with this weighting matrix Σ_i when heteroskedasticity or autocorrelation or both are present.

There are two specific variants of S_i used in calculating the standardised Z statistic

$$S_i^0(T_{0i}) = S_i\left(\widehat{\Theta}_i, \widehat{\Sigma}_i\right) \quad \forall t \in \{1, \ldots, T_{0i}\} \tag{6}$$

$$S_i^1(T_{0i}) = S_i\left(\widehat{\Theta}_i, \widehat{\Sigma}_i\right) \quad \forall t \in \{T_{0i}+1, \ldots, T\} \tag{7}$$

$\widehat{\Theta}_i$ is the least square's estimate of Θ_{0i} using the sample spanning over $\{1, \ldots, T\}$ for all i. The statistics S_i^1 are computed using the post-break sample spanning over $\{T_{0i}+1, \ldots, T\}$, whereas the S_i^0 statistics can be calculated using observations in the pre-break sample spanning over $\{1, \ldots, T_{0i}\}$. In practice, the variance-covariance weight matrix Σ_i is seldom known and is estimated in different ways. One way to estimate Σ_i is to use Newey-West autocorrelation consistent covariance estimator (see Newey and West (1987) for details). Another way to estimate Σ_i is to follow Andrews (2003):

$$\widehat{\Sigma}_i = (T_{0i}+1)^{-1} \sum_{r=1}^{T_{0i}+1} \left(\widehat{\mathbf{U}}_{i,(r:r+)} \widehat{\mathbf{U}}_{i,(r:r+)}'\right) \tag{8}$$

where $(r : r+)$ indicates that the sample spans from r to $(r + T_{1i} - 1)$ and $\widehat{\mathbf{U}}_{i,(r:r+)} = \mathbf{Y}_{i,(r:r+)} - \mathbf{X}_{i,(r:r+)} \widehat{\Theta}_i$ is a $T_{1i} \times 1$ estimated residuals vector resulting from the i^{th} time-series regression.

The S_i^0 and S_i^1 statistics are sufficient to define the Z test for structural change in panel data models as follows:

$$Z = \frac{\left(\bar{S}^1 - \bar{S}^0\right)}{\sqrt{\widehat{Var}\left(\bar{S}^1 - \bar{S}^0\right)}} \tag{9}$$

where $\bar{S}^{\nu} = \bar{S}^{\nu}(T_{0i}) = N^{-1} \sum_{i=1}^{N} S_i^{\nu}(T_{0i})$, with $\nu = 0, 1$ are the cross-sectional 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 the Z statistic is, the more evidence there is to reject the null hypothesis in favour of the alternative. Since the variances of the cross-sectional average statistics are unknown, the Z statistic uses the estimated variance of the difference of the average statistics. The estimator $\widehat{Var}\left(\bar{S}^1 - \bar{S}^0\right)$ used in this paper is defined in proposition (1) in the next section.

2.3 $\sup Z$ test with unknown break date

The Z test proposed above assumes the break date is known for each cross-section unit. This is often not the case in practice and therefore, this section proposes an extension of the Z test for an unknown break date. The idea is to find the break date $\hat{T}_{0i}$ such that the difference between the post- and pre-break sum of squared residuals statistics (S_i^ν) is maximised for each individual unit, i. This approach is popular in time series literature, see for example Andrews (1993) and Bai and Perron (1998). The approach has two merits: (1) it maximises the power of the test by finding the most likely break date for all i in the panel and (2) it offers an alternative to choosing a date in an ad-hoc manner or based on *a priori* knowledge. The two steps to computing the $\sup Z$ statistic are:

Step 1. Find the maximum difference between the sum of squared residuals statistics:

$$\widehat{T}_{0i} = \sup_{T_{0i} \in (1,T)} \left[S_i^1\left(T_{0i}\right) - S_i^0\left(T_{0i}\right)\right] \tag{10}$$

for all $i = 1, \ldots, N$ with S_i^0 and S_i^1 as defined in equations (6) and (7).

Step 2. Calculate $\bar{S}^\nu\left(\widehat{T}_{0i}\right) = N^{-1}\sum_{i=1}^N S_i^\nu\left(\widehat{T}_{0i}\right)$ for $\nu = 0, 1$, then the $\sup Z$ test statistics can be calculated as:

$$\sup Z = \frac{\bar{S}^1\left(\widehat{T}_{0i}\right) - \bar{S}^0\left(\widehat{T}_{0i}\right)}{\sqrt{\widehat{Var}\left[\bar{S}^1\left(\widehat{T}_{0i}\right) - \bar{S}^0\left(\widehat{T}_{0i}\right)\right]}} \tag{11}$$

3 Asymptotic results

3.1 Assumptions

Notation: $||\mathbf{X}|| = (\mathrm{Tr}(\mathbf{XX}'))^{1/2}$ is the Euclidean norm of matrix $\mathbf{X}$. The operator $\xrightarrow{p}$ denotes convergence in probability and $\overset{A}{\sim}$ denotes asymptotic distribution. All asymptotics are carried out under $N \to \infty$ with either T fixed or sequentially, letting $T \to \infty$ and then $N \to \infty$. $T \to \infty$ implies that either $T_{0i} \to \infty$ with T_{1i} fixed or $T_{1i} \to \infty$ with T_{0i} fixed.

Define the data set as the outcomes of a sequence of random variables $\{\mathbf{W}_i\}$ where $\{\mathbf{Y}_i, \mathbf{X}_i\} \subset \{\mathbf{W}_i\}$. For $T_{1i} > d$, the following general set of conditions on the linear regression model are required:

Assumption 1 *Let* $\mathbf{U}_i = (u_{i1}, \ldots, u_{iT})'$ *and* $\mathbf{U}_j = (u_{j1}, \ldots, u_{jT})'$ *be independently distributed random variables for all* $i \neq j$ *and* t*, such that* $\mathbb{E}\left(\mathbf{U}_i'\mathbf{U}_j\right) = \mathbf{0}$.

Assumption 2 *The data* $\{\mathbf{W}_i\}_{t=1}^{T}$ *for all i and t are stationary and ergodic. Under* H_0*, the data are* $\{\mathbf{W}_i\}_{t=1}^{T}$ *for* $i \in \{1, ..., N\}$*, while under* H_1*,* $\{\mathbf{W}_i\}_{t=1}^{T_{1i}}$ *are some random variables with a joint distribution different from* $\{\mathbf{W}_i\}_{t=T_{0i}+1}^{T}$*. Assume also that the distribution of* $\mathbf{W}_i$ *is independent of* T*. Note that under* H_1 *the data are from a triangular array since the breakpoint is changing as* T *increases.*

Assumption 3
(a) $\|\widehat{\Theta}_i - \Theta_{0i}\| \xrightarrow{p} 0$ *as* $T \to \infty$.
(b) $\sup_{\Theta \in B(\Theta_{0i}, \varepsilon)} \|\widehat{\Sigma}_i(T) - \Sigma_{0i}\| \xrightarrow{p} 0$ *as* $T \to \infty$*, where* $B(\Theta_{0i}, \varepsilon)$ *is the* ε*-neighbourhood centered around* Θ_{0i} *with radius* $\varepsilon > 0$ *for some nonsingular matrix* Σ_{0i}.

Assumption 4 $\exists\, \delta \geq 0$ *such that* $\mathbb{E}|S_i(\Theta_{0i}, \Sigma_{0i})|^{2+\delta} < \infty$*, for all* $i = 1, \ldots, N$.

Assumption 1 can be relaxed by assuming there exists an invertible matrix $\mathbf{M}$ so that $\mathbb{E}(\mathbf{U}_i'\mathbf{M}\mathbf{U}_j) = \mathbf{0}_{T \times T}$. In practice, there are several methods available to identify $\mathbf{M}$ as discussed in section 4. Assumption 2 allows for both weakly dependent processes and long memory processes, as well as conditional variation in all moments, including conditional heteroskedasticity. Assumption 3 (a) and (b) are required to ensure the consistency of the estimators for both the coefficient vector and the variance-covariance matrix. If $\Theta_{0i} = \Theta_{0j}$ for all $i \neq j$ then convergence is provided under $N \to \infty$ as long as the rate of convergence of the estimators are as fast or faster than the rate of convergence of $\bar{S}^{\nu}$, for $\nu = 0, 1$, and T_{1i} fixed under H_0 and H_1. Note also that assumption 3 is general enough to ensure consistency of the typical within-group and between-group panel data estimators. In fact, the current setting of equation (1) allows for the inclusion of fixed effects and random effects. Assumption 4 ensures the existence of second moment for all individual test statistics (S_i^{ν}).

Under the assumption that T_{0i} is known, Assumptions (3) and (4) can be replaced by the following:

Assumption 5
(a) $\mathbb{E}[\mathbf{U}_i] = 0$ *and* $\mathbb{E}[\mathbf{U}_i|\mathbf{X}_i] = 0$*, for all i.*
(b) $\mathbb{E}\left[\mathbf{U}_i'\mathbf{U}_i\right] < \infty$ *and* $\mathbb{E}\|\mathbf{X}\|^{2+\delta} < \infty$ *for some* $\delta > 0$ *and for all i.*
(c) $\mathbb{E}\left[\mathbf{X}_i'\mathbf{X}_i\right]$ *and* Σ_0 *are positive definite, for all i.*

In this case, there is no need for sequential asymptotic as only $N \to \infty$ is required. However, sequential asymptotic is required, if T_{0i} is unknown. Specifically, Assumption 4 is implied by the following:

Assumption 6 $T^{-1}\mathbf{X}_{i0}'\Sigma_{i0}^{-1}\mathbf{X}_{i0} \xrightarrow{p} Q_1$ *and* $T^{-1}\mathbf{X}_i'\Sigma_i^{-1}\mathbf{X}_i \xrightarrow{p} Q$ *where* $\mathbf{X}_{i0}$ *is the* $T_{0i} \times d$ *sub-matrix of* $\mathbf{X}_i$ *such that* $\mathbf{X}_i = \left(\mathbf{X}_{i0}', \mathbf{X}_{i1}'\right)'$ *and* $\Sigma_{i0} = \mathbb{E}\left(\mathbf{U}_{i0}\mathbf{U}_{i0}'\right)$ *with* Q *and* Q_1 *being two positive definite matrices.*

3.2 Results and comments

This sub-section derives the asymptotic distribution for the Z statistic and defines the properties of the tests.

Lemma 1 *Under Assumptions 1 - 4,*

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

where $\bar{\sigma}_\nu^2 = N^{-1}\sum_{i=1}^{N} Var(S_i^\nu)$*,* $i = 1,...,N$*,* $\nu = 0,1$ *and for all* $\varepsilon > 0$*.*
Proof. See Appendix.

Remark 1 Lemma 1 shows that Assumptions 1 - 4 are sufficient to satisfy the Lindeberg condition required by the LF-CLT. It implies that the variance of the $\bar{S}^\nu$ statistic is not dominated by the variance of any particular S_i^ν statistics.

Lemma 2 *Under Lemma 1 the asymptotic distribution of the* $\bar{S}^\nu$ *statistic is*

$$\sqrt{N}\,\frac{\bar{S}^\nu - \bar{\mu}^\nu}{\sqrt{\bar{\sigma}_\nu^2}} \overset{A}{\sim} N(0,1)$$

with $\bar{\mu}^\nu = N^{-1}\sum_{i=1}^{N}\mathbb{E}(S_i^\nu)$*,* $\nu = 0,1$*.*
Proof. See Appendix.

Theorem 1 *Under Lemma 2, the Z statistic as described in equation (9)*

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

has an asymptotic distribution

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

Proof. See Appendix.

Proposition 1 *Under assumption 3 (b), the residuals can be transformed such that* $Cov\left(S_i^1, S_i^0\right) = 0$*, for all i. Under assumption 1,* $Cov\left(S_i^\nu, S_j^\nu\right) = 0$ *for all* $i \neq j$ *and* $\nu = 0,1$. *Let*

$$\widehat{Var}\left(\bar{S}^1 - \bar{S}^0\right) = \frac{1}{N-1}\sum_{i=1}^{N}\left[\left(S_i^1 - \bar{S}^1\right)^2 + \left(S_i^0 - \bar{S}^0\right)^2\right]$$

such that $\widehat{Var}\left(\bar{S}^1 - \bar{S}^0\right) \xrightarrow{p} Var\left(\bar{S}^1 - \bar{S}^0\right)$ *as* $N \to \infty$.
Proof. See Appendix.

Remark 2 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 $\bar{S}^0$ is the same as $\bar{S}^1$ 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 as shown in Proposition 1. It ensues that the Z statistic will converge to a standard normal distribution in which valid inference can be obtained.

Remark 3 Under assumptions 1 - 4, lemmas 1 - 2 and Theorem 1 still hold when $T_{1i} \leq d$, which implies that equations (3) - (5) simplify to

$$P_i(\Theta_i, \Sigma_i) = \left(\mathbf{Y}_i - \mathbf{X}_i\Theta_i\right)' \, \Sigma_i^{-1} \, \left(\mathbf{Y}_i - \mathbf{X}_i\Theta_i\right) \tag{12}$$

Result similar to Theorem 1 can be obtained for the $\sup Z$ test statistic as the proposition below made precise:

Proposition 2 *Under assumption 1 - 3 and 6, the test statistic,* $\sup Z$ *as defined in equation (11)*

$$\sup Z = \frac{\bar{S}^1\left(\widehat{T}_{0i}\right) - \bar{S}^0\left(\widehat{T}_{0i}\right)}{\sqrt{\widehat{Var}\left[\bar{S}^1\left(\widehat{T}_{0i}\right) - \bar{S}^0\left(\widehat{T}_{0i}\right)\right]}} \tag{13}$$

has an asymptotic distribution

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

Proof: See Appendix.

4 Cross-section dependence

Assumption 1 requires cross-section independence which may be violated in practice. This section presents a brief discussion of two common approaches to handle cross-sectionally dependent errors such that the proposed tests and their asymptotic properties remain valid. Section 5.2 follows up on these extensions by presenting simulation results in which the timing of the break is unknown and the model errors are cross-sectionally dependent.

Assuming that common (unobserved) factors are the underlying cause of cross-section dependence, equation (1) can be re-written in the following way

$$\begin{aligned} y_{it} &= \Theta_i' \mathbf{x}_{it} + u_{it}, \\ u_{it} &= \gamma_i' \mathbf{f}_t + \varepsilon_{it} \quad i = 1,\ldots,N, \quad t = 1,\ldots,T, \end{aligned} \tag{14}$$

such that $\mathbf{f}_t$ is the $l \times 1$ vector of unobserved common effects and γ_i' are the factor loadings associated with $\mathbf{f}_t$. The factor loadings are assumed to be constant over time and the structural change can only occur in the Θ_i. ε_{it} are the idiosyncratic shocks specific to each individual unit. Hence, it is possible to construct a matrix $\mathbf{M}$ so as to "filter out" the common factors $\mathbf{f}_t$ as such

$$\mathbf{M_F} = (\mathbf{I}_T - \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'),$$

with $\mathbf{F} = (\mathbf{f}_{1t}, \mathbf{f}_{2t}, \ldots, \mathbf{f}_{lt})$ a $T \times l$ matrix. Equation (14) can be rewritten with matrices as in equation (2). By pre-multiplying it with the projection matrix $\mathbf{M_F}$ to eliminate the factors as follows

$$\mathbf{M_F Y}_i = \mathbf{M_F X}_i \Theta_{0i} + \begin{pmatrix} \mathbf{0}_{T_{0i}} \\ \mathbf{M_F X}_{iT_{1i}} \end{pmatrix} \delta_i + \mathbf{M_F U}_i \tag{15}$$

such that $\mathbb{E}(\mathbf{U}_i' \mathbf{M_F U}_j) = \mathbf{0}$ for all $i \neq j$ and t. Note that equation (15) satisfies assumptions 1-4 required by Theorem 1. Hence, no additional assumption on the Z test is required when the model has cross-sectionally dependent errors.

The two common approaches to estimate the matrix $\mathbf{M_F}$ are the Common Correlated Effect (CCE) estimator proposed by Pesaran (2006) and the Principal Component approach by Bai and Ng (2002), Bai (2003) and Pesaran and Kapetanios (2005), which use observed data and not the estimated model errors. The CCE approach uses cross-sectional averages of the endogenous and explanatory variables as proxies for the common factors, $\mathbf{f}_t$. Specifically, the estimate of $\mathbf{M_F}$ can be written as $\widehat{\hat{\mathbf{M}}}_\mathbf{F} = (\mathbf{I}_T - \hat{\bar{\mathbf{F}}}(\hat{\bar{\mathbf{F}}}'\hat{\bar{\mathbf{F}}})^{-1}(\hat{\bar{\mathbf{F}}})')$, where $\hat{\bar{\mathbf{F}}} = (\bar{\mathbf{y}}_t, \bar{\mathbf{X}}_t)$ is the matrix of cross-sectional averages of the endogenous and exogenous variables. Pesaran (2006) shows that the CCE estimators are consistent. The CCE estimator is appealing as it is simple to implement in applied contexts. Another way to estimate $\mathbf{M_F}$ is to use Principal Components as presented in Coakley et al. (2002), Bai and Ng (2002) and Bai (2003). Specifically, Pesaran and Kapetanios (2005) applies the Bai (2003) procedure to the observed data so as to obtain consistent estimates of the unobserved factors. Furthermore, the number of factors (l) can also be estimated using the procedure introduced by Bai and Ng (2002).

5 Monte Carlo simulations

This section aims to provide some benchmark Monte Carlo results in order to investigate the normality, size and power of the proposed tests for structural change. Note that the purpose of the simulation is to show that the proposed tests satisfy the asymptotic properties even when subjected to strenuous conditions that often appear in empirical work. Hence, the size and power of the test are presented under the more challenging conditions, that is with only 10% of the observations after the break. When considering 20% or even 50% of the observations after the break, the size of the test seems to be unaffected while the power of the test increases when the number of observations post break increase, as shown in Tables 3 and 4 (for the 20% case). [1]

5.1 Experiment with a known break and autocorrelation

The experiment uses the following linear regression model

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

where $\mathbf{x}_{it} \sim i.i.d.N(0,1)$ the number of regressors is set to $d = 5$ including a constant. The regression's error term, u_{it} is generated with an AR(1) process

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

[1]The extended results are available upon request.

All cross-sections have the same autoregressive parameter, ρ, set to .4 and .95.

Four different types of *iid* distributions for the innovations of the error term, ε_{it}, are considered: standard normal ($N(0,1)$), a recentered and rescaled χ_5^2 and t_5 with mean zero and variance one, and an uniform distribution with support $[0,1]$. Formally,

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

Hence, different individual units have different innovation processes, such that the four distributions are intermixed evenly in the panel.

The results from the simulation exercises are presented in two parts. First, the null hypothesis is simulated in order to analyse the size of the Z test. Moreover, a discussion of the properties of the distribution of the test under the LF-CLT is provided. The simulation of the null uses the full sample T and coefficient vector $\Theta_{0i} = 0$, for all i. Second, the power properties of the test are examined by simulating the alternative hypothesis of a structural change for only a limited number of individual units (N_1). Under the alternative hypothesis $\Theta_{0i} = 0$ and $\delta_i = 0.1 \times (1,1,1,1,1)'$ for some i, and $\delta_j = 0$ for $j \neq i$. In order to examine the power of the test against the proportion of individuals that experience a structural change, five sets of Monte Carlo experiments have been conducted with the proportion, $\frac{N_1}{N} = c$, being .10, .50, .65, .80 and 1. When $\frac{N_1}{N} = 1$, the coefficient vector is homogeneous implying that each individual unit experiences a structural change.

The Monte Carlo experiments are conducted with the following settings: $T_{1i} = 0.1 \times T$, $T = 30,50,100$, $N = 20,40,60,80,100$. For simplicity, the break dates T_{0i} and post-break observations T_{1i} are known and identical for all i. The distribution, size and power of the test are also investigated when the number of post-break observations are increased to $T_{1i} = 0.2 \times T$ for $T = 30,50$. The number of replications is 2000. All simulations are carried out using Scilab 5.2.0.[2]

5.1.1 Distribution and size of the test

Table 1 shows that the Z test is close to standard normal and that the LF-CLT holds with relatively small time and cross-section dimensions and when serial correlation is moderate. The Jarque-Bera test statistics show evidence of normality at the 5

[2]The programming code is available upon request.

% level of significance. These conclusions remain the same when the number of post-break observations are increased to 20% of T instead of 10%, as summarised in Table 2. The normality of the distribution worsens, however, in the presence of very high autocorrelation ($\rho = .95$). The Jarque-Bera test for normality is rejected at the 5% level of significance.

Table 1: Moments for the distribution of the test under the null, $T_{1i} = 0.1 \times T$

				N				
Moments	T	T_{1i}	ρ	20	40	60	80	100
Mean	30	3	.4	-.063	.078	-.122	-.135	-.112
	50	5	.4	-.088	-.086	-.156	-.150	-.181
	100	10	.4	-.057	-.051	-.112	-.063	-.061
	100	10	.95	-.539	-.746	-.936	-1.06	-1.24
Variance	30	3	.4	1.19	1.09	1.07	1.02	1.04
	50	5	.4	1.18	1.08	1.06	1.02	1.05
	100	10	.4	1.11	1.03	1.01	1.04	1.05
	100	10	.95	.907	.914	.834	.758	.735
Skewness	30	3	.4	.033	-.095	-.041	-.007	.002
	50	5	.4	.022	-.005	.011	.036	-.009
	100	10	.4	.016	-.020	-.004	-.040	.025
	100	10	.95	.471	.397	.288	.394	.262
Kurtosis	30	3	.4	3.28	3.16	3.09	3.05	3.06
	50	5	.4	3.13	2.87	2.92	3.02	3.11
	100	10	.4	3.07	3.07	3.16	2.95	2.96
	100	10	.95	3.11	3.05	2.98	3.52	3.11
Jarque-Bera	30	3	.4	6.90	5.14	1.24	.22	.30
	50	5	.4	1.57	1.42	.57	.47	1.04
	100	10	.4	.49	.54	2.14	.74	.36
	100	10	.95	75.08	52.71	27.62	74.39	23.97

Note: ρ is the autocorrelation coefficient, N is the cross-section dimension, T is the time series dimension and T_{1i} is the number of observations post-break. 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 for the distribution of the test under the null, $T_{1i} = 0.2 \times T$

				N				
Moments	T	T_{1i}	ρ	20	40	60	80	100
Mean	30	6	.4	-.042	-.048	-.048	-.102	-.096
	50	10	.4	-.031	-.059	-.064	-.052	-.035
Variance	30	6	.4	1.16	1.05	1.03	1.02	1.08
	50	10	.4	1.20	1.07	1.06	1.01	1.02
Skewness	30	6	.4	-.012	-0.05	.003	.023	-.005
	50	10	.4	-.006	.012	-.028	-.038	.008
Kurtosis	30	6	.4	3.09	2.84	3.17	3.10	3.14
	50	10	.4	3.12	2.89	3.13	3.10	3.04
Jarque-Bera	30	6	.4	.72	2.97	2.41	1.01	1.64
	50	10	.4	1.12	1.06	1.67	1.31	2.35

Note: ρ is the autocorrelation coefficient, N is the cross-section dimension, T is the time series dimension and T_{1i} is the number of observations post-break. 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.

As can be seen from Table 3, the test has reasonable size when N is large and T is small, or vice-versa. The size of the test is relatively close to the desired value of .05 for large N ($N \geq 80$) and small T ($T = 30$) and $T_{i1} = 0.1 \times T$. These results hold in the presence of moderate autocorrelation ($\rho = 0.4$). The size is also close to .05 when the time horizon is increased to $T = 50$ and $T = 100$, while N is reduced to $N \geq 40$ and 20 respectively. Note that it is relatively unaffected if the number of post-break observations are increased to 20% of T. When autocorrelation is very high ($\rho = .95$), the size of the test declines considerably. Under these conditions, it would be best to increase T to improve the size of the test.

Table 3: Size of normal significance level .05

			N				
T	T_{1i}	ρ	20	40	60	80	100
					$T_{1i} = 0.1 \times T$		
30	3	.4	.069	.061	.064	.054	.053
50	5	.4	.070	.056	.055	.057	.058
100	10	.4	.065	.047	.055	.053	.054
	10	.95	.055	.092	.130	.142	.199
					$T_{1i} = 0.2 \times T$		
30	6	.4	.071	.055	.058	.049	.060
50	10	.4	.075	.061	.062	.048	.053

Note: ρ is the autocorrelation coefficient, N is the cross-section dimension, T is the time series dimension and T_{1i} is the number of observations post-break, which is set to equal 10% of T and also 20% of T for $T = 30, 50$.

5.1.2 Power of the test

Table 4 summarises the results for the power of the test. As N grows large, the power of the Z test improves even when T is small. When $T = 50$, for example, the power of the test is .65 with $N = 80$ and $c = .80$. If the number of post-break observations (T_{1i}) are increased to 20% of T, the power improves to .90. The power of the test thus improves as the number of post break observations are increased, unlike the size of the test which remains mostly unaffected.

Table 4: Power of normal significance level .05

T	T_{1i}	c	ρ	N = 20	40	60	80	100
						$T_{1i} = 0.1 \times T$		
30	3	.80	.4	.015	.041	.071	.139	.223
	3	1	.4	.079	.263	.477	.681	.817
50	5	.50	.4	.009	.011	.011	.025	.034
	5	.80	.4	.056	.216	.435	.651	.809
	5	1	.4	.297	.744	.951	.992	.999
100	10	.10	.4	.000	.003	.004	.005	.006
	10	.50	.4	.022	.114	.294	.523	.696
	10	.65	.4	.095	.520	.856	.967	.991
	10	.80	.4	.347	.911	.991	.999	1.00
	10	1	.4	.852	.994	1.00	1.00	1.00
100	10	.50	.95	.008	.011	.011	.012	.014
	10	.80	.95	.075	.203	.368	.542	.705
	10	1	.95	.270	.620	.869	.967	.994
						$T_{1i} = 0.2 \times T$		
30	6	.80	.4	.031	.077	.141	.233	.348
	6	1	.4	.130	.359	.602	.786	.905
50	10	.50	.4	.016	.024	.052	.097	.155
	10	.80	.4	.134	.433	.720	.898	.965
	10	1	.4	.454	.869	.984	1.00	1.00

Note: ρ is the autocorrelation coefficient, N is the cross-section dimension, T is the time series dimension and T_{1i} is the number of observations post-break, which is set to equal 10% of T and also 20% of T for $T = 30, 50$. Note that c is the fraction of the individual units experiencing the break and $c = \frac{N_1}{N}$

The test gains power as either N, c, T or T_{1i} increases. For instance, the power of the test is above .80 when $N \geq 60$ and $c = 1$, for $T = 30$. Even when both c and N are of medium size, for example $c = .65$ and $N = 60$, the power is .85 with a larger T ($T = 100$). Moreover, the power of the test is still good (.70) when N is large ($N = 100$) and c is low ($c = .50$) with $T = 100$. The reverse is also true: when $N = 40$ and $c = .80$, the power is .91.

The power is quite robust to autocorrelation, especially when N and c are large. Even when serial correlation is .95, for instance, the test has power of .71 when $N = 100$ and $c = .80$.

5.2 Experiment with unknown break and cross-section dependence

This set of experiments investigates the size and power of the test when there is an unknown break and the errors are cross-sectionally dependent as discussed in section 4. The Monte Carlo design uses the same linear regression model as before but the regression's error term, u_{it} is generated by

$$u_{it} = \gamma_i^{'} f_t + \varepsilon_{it} \quad t = 1, \ldots, T, \quad i = 1, \ldots, N$$

where $f_t \sim i.i.d.N(0,1)$ is the unobserved common factor and the error term is composed of the same mix of four different types of distributions as before and $\gamma_i = \sqrt{0.02}$. The test statistic handles the cross-section dependence by applying the Common Correlated Effect (CCE) estimator proposed by Pesaran (2006). The cross-sectional averages of the data are used to construct the $\widehat{\bar{\mathbf{M}}}_{\mathbf{F}}$ as shown in section 4. The alternative hypothesis is simulated using $\Theta_{0i} = (0,0,0,0,0)$ and $\delta_i = (1,1,1,1,1)^{'}$, for some i and $\delta_j = 0$, for $j \neq i$, whereas under the null hypothesis the coefficient vector is $\Theta_{0i} = 1$.

The sample sizes N and T as well as the fraction of individual units that experience the break c are incremented as in the previous experiments. The post-break observations T_{1i} are identical for all i and set to be $T_{1i} = 0.1 \times T$ as in the previous section for $T = 50, 100$ and $T_{1i} = 0.2 \times T$ for $T = 30$. The break dates T_{1i} are estimated by optimising the difference between the sum of squared residuals statistics as shown in equation (10) and the $\sup Z$ test is calculated as in (11).

5.2.1 Distribution and size of the test

Table 5 presents the moments of the distribution of the Z test. The results are comparable to the moments of the distribution in the previous experiment, as shown in Table 1, and suggest that the Z test is close to standard normal. The Jarque-Bera test does not reject the null of normality at the 5 % level, providing further evidence of normality.

Table 5: Moments for the distribution of the test under the null

Moments	T	T_{1i}	N = 20	40	60	80	100
	30	3	-.019	.010	.001	-.004	-.020
Mean	50	5	.011	.004	.010	.034	-.039
	100	10	-.031	-.0005	.004	.010	.0005
	30	3	1.21	1.15	1.21	1.26	1.31
Variance	50	5	1.18	1.16	1.05	1.08	1.18
	100	10	1.20	1.09	1.11	1.03	1.08
	30	3	.074	-.026	.014	-.098	-.021
Skewness	50	5	.051	.008	.043	-.034	-.039
	100	10	-.028	.027	-.027	-.109	-.049
	30	3	2.85	2.86	2.78	2.84	2.81
Kurtosis	50	5	2.76	2.83	2.82	3.05	3.14
	100	10	2.88	2.88	2.98	2.88	2.78
	30	3	3.71	.453	.121	6.54	.302
Jarque-Bera	50	5	1.67	.040	1.21	.911	1.28
	100	10	.526	.500	.527	8.18	1.55

Note: N is the cross-section dimension, T is the time series dimension and T_{1i} is the number of observations post-break. 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. The break date is unknown for all individual units and each break date is obtained from optimising the difference between the sum of squared residuals as shown in equation (10).

The results for the size of the test are presented in Table 6. The test has good size when $T > 30$ and $N > 40$. As N grows larger, it gets closer to the desired level of .05. Overall, the size of the test is comparable to the previous experiment. There is some size distortion, however, when $T = 30$ which does not seem to improve much as N grows large.

Finally, it seems that filtering the common factor with a CCE estimator does not affect the size of the test. Furthermore, the search for a break date does not affect the size of the test.

Table 6: Size of normal significance level .05

		N				
T	T_{1i}	20	40	60	80	100
30	3	.072	.069	.069	.077	.089
50	5	.072	.061	.053	.054	.069
100	10	.073	.064	.054	.049	.056

Note: N is the cross-section dimension, T is the time series dimension and T_{1i} is the number of observations post-break, which is set to equal 10% of T and also 20% of T for $T = 30$ only. The break date is unknown for all individual units and each break date is obtained from optimising the difference between the sum of squared residuals as shown in equation (10).

5.2.2 Power of the test

The power of the Z test increases as N grows large for $c \geq .5$ and for either $T = 50$ or $T = 100$ as summarised in Table 7. The power also improves as either T or c grow. Furthermore, in this set of Monte Carlo experiments, the test has good power even when $T = 30$ providing that $c > 0.5$. Furthermore, the power of the test is unaffected by the unobserved common factor provided that it is filtered out appropriately, as in this case with the CCE approach.

The results also imply that estimating the break date at the individual unit level increases the power of the test. The average of the estimated break dates ($\widehat{\bar{T}} = N^{-1} \sum_{i=1}^{N} \widehat{T}_i$) is reported in Table 8. Note also that the estimated break date is not affected by either the size of T or N.

Table 7: Power of normal significance level .05

T	T_{1i}	c	N = 20	40	60	80	100
					$T_{1i} = 0.2 \times T$		
30	6	.80	.992	1.00	1.00	1.00	1.00
30	6	1	.998	1.00	1.00	1.00	1.00
					$T_{1i} = 0.1 \times T$		
50	5	.50	.999	1.00	1.00	1.00	1.00
50	5	.80	1.00	1.00	1.00	1.00	1.00
50	5	1	1.00	1.00	1.00	1.00	1.00
100	10	.10	.165	.314	.321	.336	.365
100	10	.50	.800	.972	.992	.998	1.00
100	10	.65	.943	.995	.999	.999	.999
100	10	.80	.987	.999	.999	.999	.999
100	10	1	.999	1.00	.999	1.00	1.00

Note: N is the cross-section dimension, T is the time series dimension and T_{1i} is the number of observations post-break, which is set to equal 10% of T and 20% of T for $T = 30$ only. Note that c is the fraction of the individual unit experiencing the break and $c = \frac{N_1}{N}$. The break date is unknown for all individual units and each break date is obtained from optimising the difference between the sum of squared residuals as shown in equation (10).

Table 8: Average estimated break dates

T	T_{1i}	c	N = 20	40	60	80	100
					$T_{1i} = 0.2 \times T$		
30	6	.80	22.86	22.83	22.82	22.84	22.83
			(3.71)	(3.79)	(3.79)	(3.76)	(3.77)
30	6	1	22.86	22.88	22.85	22.85	22.86
			(3.76)	(3.71)	(3.77)	(3.78)	(3.77)
					$T_{1i} = 0.1 \times T$		
50	5	.50	44.10	44.07	44.05	44.04	44.02
			(1.34)	(1.29)	(1.27)	(1.24)	(1.30)
50	5	.80	44.00	43.98	43.98	43.98	43.98
			(1.41)	(1.40)	(1.37)	(1.33)	(1.30)
50	5	1	44.05	44.04	44.03	44.04	44.03
			(1.35)	(1.27)	(1.27)	(1.17)	(1.23)
100	10	.10	92.83	92.53	92.15	91.77	91.43
			(2.06)	(2.35)	(2.80)	(3.04)	(3.18)
100	10	.50	88.89	88.84	88.76	88.76	88.76
			(5.40)	(5.23)	(5.35)	(5.29)	(5.32)
100	10	.65	88.64	88.59	88.58	88.60	88.58
			(5.66)	(5.63)	(5.65)	(5.55)	(5.57)
100	10	.80	88.52	88.56	88.53	88.59	88.55
			(5.73)	(5.67)	(5.62)	(5.56)	(5.64)
100	10	1	88.88	88.90	88.86	88.87	88.86
			(5.04)	(5.02)	(5.07)	(5.08)	(5.08)

Note: The average of the estimated break dates is calculated as $\widehat{\widehat{T}} = \frac{1}{N}\sum_{i=1}^{N}\widehat{T}_{0i}$ and each break date is obtained from optimising the difference between the sum of squared residuals as shown in equation (10). The number in brackets (.) are the standard deviations for each of the average estimated break dates.

6 Empirical illustration: euro's trade effect

This section provides an empirical illustration of how this paper's proposed test can be applied in practice. The example is chosen to highlight the usefulness of the test in a case where there is an uncertain number of individual units undergoing a break, the break occurs at the end of sample and the errors are cross-sectionally correlated. The empirical question is whether or not there was a break in intra–euro area trade following the introduction of the euro. The question was at the forefront of the empirical trade literature until the financial crisis began in 2007, revived after the seminal contribution of Rose (2000), and has been discussed actively in policy circles. However, empirical evidence has been clouded by somewhat ill-suited econometric techniques. Most papers in the literature, for example Micco et al. (2003) and Flam and Nordström (2003), introduce various dummy variables in their regressions to capture the new currency's introduction.[3] Furthermore, the use of F-tests employed to evaluate the significance of the dummy coefficients rest on restrictive assumptions such as normal, homoskedastic and *i.i.d* errors.

The test developed in this paper is well suited to 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 capturing the effect of the euro. Second, the test requires very few regularity conditions. It remains asymptotically valid despite non-normal, heteroskedastic and autocorrelated errors. Third, the test makes no distributional assumptions on S_i^v for all i; only the cross-sectional average statistics are shown to be asymptotically normal as warranted by the panel's cross-section dimension. Fourth, the test explicitly allows for some individual units, and not all, to exhibit a break. This last point is relevant to the euro area example where trade among countries which were certain to enter the euro might have picked up before it did in other countries.

6.1 Regression model

The model to test for the euro's effect on trade comes from a standard trade gravity equation developed by Anderson and van Wincoop (2003):

$$IM_{i,jt} = \alpha_{i,j} + \lambda_t + \varphi_j + \psi_{i1} GDP_{it} + \psi_{i2} GDP_{jt} + \psi_{i3} RER_{i,jt} + u_{it} \qquad (16)$$

where $IM_{i,jt}$ is the value of imports from country j to country i, GDP_{jt} and GDP_{it} are nominal GDP at time t for country j and country i, respectively, to control for demand and country size effects, $RER_{i,jt}$ is the real exchange rate between the

[3]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).

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 such as common border, language, history, legal system, distance and other variables traditionally shown to matter in gravity equations. λ_t allows for the country pair intercept to be time dependent, to reflect evolving trade costs for instance, and φ_j is a country of origin dummy. The goal is to capture the "multilateral trade resistance term" first shown to be important by Anderson and van Wincoop (2003), interpreted as an average trade barrier between two countries relative to all their trading partners. u_{it} is the error term. The microfoundations of the model are discussed and summarised at some length in Mancini-Griffoli and Pauwels (2006).

It is convenient to re-write the above error term in order to capture both the observed common time dependent effects, λ_t, as well as potential unobserved common effects, $\mathbf{f}_t$. From equation (16), $u_{it} = \gamma_i'\mathbf{f}_t + \nu_{it}$ where ν_{it} is the idiosyncratic shock. The observed and unobserved common effects induce cross-section correlation of the errors. Therefore, model (16) becomes

$$\begin{aligned} IM_{i,jt} &= \alpha_{i,j} + \varphi_j + \psi_{i1} GDP_{it} + \psi_{i2} GDP_{jt} + \psi_{i3} RER_{i,jt} + \varepsilon_{it} \\ \varepsilon_{it} &= \gamma_i'\mathbf{f}_t + \lambda_t + \nu_{it} \end{aligned} \qquad (17)$$

The gravity model (17) is set up in the following way:

$$IM_{i,jt} = \left(1, 1, GDP_{it}, GDP_{jt}, RER_{i,jt}\right) \begin{pmatrix} \alpha_{i,j} \\ \varphi_j \\ \psi_{i1} \\ \psi_{i2} \\ \psi_{i3} \end{pmatrix} + \varepsilon_{it}$$

Next, the model is first-differenced

$$\Delta IM_{i,jt} = \left(\Delta GDP_{it}, \Delta GDP_{jt}, \Delta RER_{i,jt}\right) \begin{pmatrix} \psi_{i1} \\ \psi_{i2} \\ \psi_{i3} \end{pmatrix} + \Delta\varepsilon_{it}$$

which eliminates both fixed-effects, $\alpha_{i,j}$ and φ_j. First-differencing the model is also beneficial to ensure against non-stationarity. The test therefore becomes one for a break in the relation between the growth of trade and the growth of its explanatory variables. Define $\mathbf{X}_t = \left(GDP_{it}, GDP_{jt}, RER_{i,jt}\right)'$ and $\Psi_i = (\psi_{i1}, \psi_{i2}, \psi_{i3})'$, the model can be re-written similarly to equation (2)

$$\Delta IM_{i,jt} = \begin{cases} \Psi_i'\Delta\mathbf{X}_t + \Delta\varepsilon_{it} & t = \{1, \ldots, T_{0i}\} \\ (\Psi_i' + \delta_i')\Delta\mathbf{X}_t + \Delta\varepsilon_{it} & t = \{T_{0i}+1, \ldots, T\} \end{cases},$$

with the hypotheses

$$H_0 : \delta_i = \mathbf{0} \quad \text{for all } i = 1,\dots,N \text{ and } t = 1,\dots,T_{0i},\dots,T$$
$$H_1 : \delta_i \neq \mathbf{0} \quad \text{for some } i \text{ and for } t > T_{0i}.$$

The quarterly data were obtained from Eurostat, IMF DOTS and IFS, as in most other relevant empirical papers. The bilateral import values from country j to country i were obtained from IMF DOTS for 10 euro area countries. The individual units in the panel represent trading pairs, where i imports from j, rather than single countries. This means that there are $N = 90$ trading pairs for 10 euro area countries. Greece is excluded from the sample since it joined the euro only in January 2001, and Luxembourg and Belgium are grouped together as their trade data are confounded over most of the sample period. The data spans from the first quarter of 1980 until the fourth quarter of 2004 ($T = 100$). All data were seasonally adjusted using the standard X.12 smoothing algorithm.

6.2 Empirical results

The euro's trade effect is investigated in two contexts: (1) when the break date is assumed to be known and common to all trading pairs and (2) when the break date is unknown and is allowed to change for all trading pairs. Table 9 presents 4 versions of the Z-test. The first Z-statistic is computed under the assumptions that $\Sigma_i = \sigma^2\mathbf{I}$ and the errors are cross-sectionally independent. In the second version, the Z-statistic is calculated with the variance-covariance matrix ($\widehat{\Sigma}_i$) as estimated in equation (8). In the third, cross-section dependence in the errors are taken into account using CCE ($\bar{\mathbf{M}}_{\mathbf{W}}$) as in section 4. The 4^{th} Z-statistic is generated using CCE and the estimated variance-covariance matrix (8).

First, the break date is assumed to be common to all trading pairs and is set to the first quarter of 1998 (1998 Q1), one year prior to the actual adoption of the euro. This is to take into account forward-looking agents and is aligned with the findings of Micco et al. (2003) and Flam and Nordström (2003) who find a "euro effect" as early as 1998. The first general pattern that emerges from glancing at the results across the various test samples is that the evidence of a break in the relation between trade and its explanatory variables in 1998 Q1 is mixed. Two test-statistics are significant at the 1% level, one at the 5% level and the other is not significant.

Second, when the break date is unknown, each break date is estimated by optimising the difference between the sum of squared residuals as shown in equation (10). Since the euro was instituted by the provisions of the Maastricht treaty in 1992, the break search interval starts on that date and spans till the end of the sample (2004 Q4). In contrast to the known break case, the unknown break results

Table 9: euro's trade effect

Estimator(s) in S_i^v	Presumed/Estimated break date	Value of Z-statistic	p-value
	Known common break		
$S_i^v(\widehat{\Theta}_i)$	1998 Q1	3.45	<.01***
$S_i^v(\widehat{\Theta}_i, \widehat{\Sigma}_i)$	1998 Q1	3.56	<.01***
$S_i^v(\widehat{\Theta}_i, \bar{\mathbf{M}}_{\mathbf{W}})$	1998 Q1	1.23	.219
$S_i^v(\widehat{\Theta}_i, \widehat{\Sigma}_i, \bar{\mathbf{M}}_{\mathbf{W}})$	1998 Q1	2.12	.034**
	Unknown individual break		
$S_i^v(\widehat{\Theta}_i)$	1997 Q3	3.08	.002***
$S_i^v(\widehat{\Theta}_i, \widehat{\Sigma}_i)$	1997 Q2	17.33	<.01***
$S_i^v(\widehat{\Theta}_i, \bar{\mathbf{M}}_{\mathbf{W}})$	1998 Q1	2.69	.007***
$S_i^v(\widehat{\Theta}_i, \widehat{\Sigma}_i, \bar{\mathbf{M}}_{\mathbf{W}})$	1997 Q3	13.40	<.01***

Note: $S_i^v(\widehat{\Theta}_i, \widehat{\Sigma}_i)$ uses the estimated variance-covariance matrix as in equation (8) and $S_i^v(\widehat{\Theta}_i, \widehat{\Sigma}_i, \bar{\mathbf{M}}_{\mathbf{W}})$ uses the CCE estimator constructed as in section 4 with the aforementioned variance-covariance matrix. The average of the estimated break dates is calculated as $\widehat{\widehat{T}} = \frac{1}{N}\sum_{i=1}^{N}\widehat{T}_{0i}$. The break search interval starts in 1992 Q1 until 2004 Q4. */**/*** indicate 10%/5%/1% level of significance.

offer strong evidence of euro's effect on trade within euro area member countries. All tests are significant at the 1% level and are consistent with each other regardless of the type of estimator used. This underscores the advantage of using a statistical method to find a break date, instead of a priori knowledge which may prove to miss certain dynamics of the data. Average estimated break dates (calculated as $\widehat{\widehat{T}} = \sum_{i=1}^{N} \widehat{T}_{0i}$) are reported in Table 9. The reported $\widehat{\widehat{T}}$ is fairly consistent across the different estimators and with the literature. The estimated average break date is between the second quarter of 1997 and the first quarter of 1998, just preceding the introduction of the euro as found in Micco et al. (2003) and Flam and Nordström (2003) which emphasize the forward looking reactions of agents.

In sum, the above example has shown that the proposed test performs relatively consistently and adequately in an empirical context. It has allowed for rigour in testing an important policy question.

7 Concluding remarks

This paper proposes a Z test for structural change in panel data models for which some – and not all – individual units exhibit a break. The test statistic is constructed as a standardised difference between cross-sectional averages of a sum of squared residuals type statistics computed in the pre- and post-break sample. Asymptotic results show that the Z test is normally distributed according to the Lindeberg-Feller central limit theorem. The test is robust to non-normal, heteroskedastic and autocorrelated model errors and to end-of-sample structural change with few post-break observations (for example 10% of T). The Monte Carlo simulations confirm that the test performs relatively well in these circumstances. The simulations also show that the test's distribution is close to a standard normal even when the T and N dimensions are small.

Two extensions are considered. First, this paper proposes a sup Z test statistic for the case when the break date is unknown. Second, the Common Correlated Effects (CCE) estimator of Pesaran (2006) is employed to "filter out" cross-section dependence in the error term of the panel data model. Simulation results show that filtering cross-section dependence with CCE does not alter the power and the size of the sup Z test. In addition, Monte Carlo results show that estimating the break date improves the power of the test.

In the end, an illustration of the test is provided to gauge whether there was a break in intra–euro area trade following the introduction of the euro. The test proves particularly useful as there is an uncertain number of countries undergoing a break, the break occurs at the end of sample and the errors are cross-sectionally dependent.

Appendix

Proof of Proposition 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*

$$\widehat{Var}\left(\bar{S}^1 - \bar{S}^0\right) \xrightarrow{p} Var\left(\bar{S}^1 - \bar{S}^0\right)$$

as $N \to \infty$*. Under Assumption 4 (*$Var(S_i^\nu) < \infty$*,* $\forall i$*,* $\nu = 0,1$*):*

$$Var\left(\bar{S}^1 - \bar{S}^0\right) = \bar{\sigma}_1^2 + \bar{\sigma}_0^2 - 2\,\bar{\sigma}_{0,1}$$

where $\bar{\sigma}_\nu^2 = N^{-1}\sum_{i=1}^N Var(S_i^\nu)$*, for* $\nu = 0,1$ *and* $\bar{\sigma}_{0,1} = N^{-1}\sum_{i=1}^N Cov\left(S_i^1, S_i^0\right)$*. Either under the assumption of a known* Σ_i *or under assumption 3 (b) which requires that* $\widehat{\Sigma}_i^{-1}$ *is a consistent estimate of* Σ_i^{-1}*, the transformed residuals,* $\Sigma_i^{-\frac{1}{2}}\mathbf{U}_i$*, are i.i.d and therefore* $Cov\left(S_i^1, S_i^0\right) = 0$*,* $\forall i$*. Furthermore, either under assumption 1 or under the assumption of cross-section independence,* $Cov\left(S_i^\nu, S_j^\nu\right) = 0$*,* $\forall i \neq j$ *and* $\nu = 0,1$ *so that*

$$Var\left(\bar{S}^1 - \bar{S}^0\right) = \bar{\sigma}_1^2 + \bar{\sigma}_0^2$$

Let, $\hat{\bar{\sigma}}_\nu^2 = (N-1)^{-1}\sum_{i=1}^N \widehat{Var}(S_i^\nu)$*, for* $\nu = 0,1$*, such that*

$$\widehat{Var}\left(\bar{S}^1 - \bar{S}^0\right) = \hat{\bar{\sigma}}_1^2 + \hat{\bar{\sigma}}_0^2$$

since $Cov(S_i^0, S_i^1) = 0$*,* $\forall i$ *and* $Cov(S_i^\nu, S_j^\nu) = 0$*,* $\forall i \neq j$ *and* $\nu = 0,1$ *under the above stated assumptions. This above expression can be easily computed from the data as*

$$\widehat{Var}\left(\bar{S}^1 - \bar{S}^0\right) = \frac{1}{N-1}\left[\sum_{i=1}^N \left(S_i^1 - \bar{S}^1\right)^2 + \sum_{i=1}^N \left(S_i^0 - \bar{S}^0\right)^2\right]$$

$\widehat{Var}\left(\bar{S}^1 - \bar{S}^0\right)$ *is a consistent estimator because by Weak Law of Large Numbers*

$$\hat{\bar{\sigma}}_1^2 \xrightarrow{p} \bar{\sigma}_1^2$$
$$\hat{\bar{\sigma}}_0^2 \xrightarrow{p} \bar{\sigma}_0^2$$

and by preservation of convergence for continuous transformation

$$\hat{\bar{\sigma}}_1^2 + \hat{\bar{\sigma}}_0^2 \xrightarrow{p} \bar{\sigma}_1^2 + \bar{\sigma}_0^2$$

This completes the proof. ■

Proof of Lemma 1 *Following White (1999) (see also Loeve (1977) pp. 292-294), it is sufficient to show that the following Lindeberg condition holds under the assumptions 2 - 4:*

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

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

$$\delta_i(\varepsilon,N)=\int_{A(\varepsilon,N)}(S_i-\mathbb{E}(S_i))^2 dF_i(S_i). \tag{A-2}$$

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

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

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

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

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

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

Let $\sigma_{\max}^2=\max\{\sigma_i^2\}$ and $\sigma_{\min}^2=\min\{\sigma_i^2\}$. Since $0<\sigma_i^2<\infty$, $\forall i$ and therefore $0<\sigma_{\max}^2,\sigma_{\min}^2<\infty$. Hence,

$$\begin{aligned}\frac{\sum_{i=1}^N r_i(\varepsilon,N)\sigma_i^2}{\sum_{i=1}^N\sigma_i^2}&\leq\frac{\sigma_{\max}^2}{N\sigma_{\min}^2}\sum_{i=1}^N r_i(\varepsilon,N)\\&=\frac{c}{N}\sum_{i=1}^N r_i(\varepsilon,N)\end{aligned}$$

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

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

Let $\triangle = \max_{i,N,\varepsilon}\{\triangle_{i,N}\}$ *where* $\triangle$ *is not a monotonic function of N and since* $\triangle_{i,N} < 1$ $\forall i, N, \varepsilon$, *therefore* $\triangle < 1$. *Hence,*

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

Therefore,

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

This completes the proof. ■

Proof of Lemma 2 *Under Assumption 1,* S_i^ν *is independent of* S_j^ν, *for* $i \neq j$ *and* $\nu = 0, 1$. *In addition, under Lemma 1 the average statistics* $\bar{S}^\nu$ *satisfy the Lindeberg-Feller condition for* $\nu = 0, 1$. *Since these two conditions are required by the Lindeberg-Feller CLT, then*

$$\sqrt{N}\,\frac{\bar{S}^\nu - \bar{\mu}^\nu}{\sqrt{\bar{\sigma}^2_\nu}} \overset{A}{\sim} N(0,1)$$

This completes the proof. ■

Proof of Theorem 1 *Under Lemma 2,* $\bar{S}^0$ *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 of* $\bar{\mu}^0 = \bar{\mu}^1$. *This completes the proof.* ■

Proof of Proposition 2 *The proof of Proposition 2 follows the same arguments as the proof of Theroem 1. Since Assumptions 1 - 4 are sufficient to ensure Lemma 1, it is therefore sufficient to show that Assumption 6 implies Assumption 4 in the case of unknown break date. This can be achieved by first establishing equivalence between the* $S_i^\nu\left(\widehat{T}_{0i}\right)$ *test statistic and the* $LM_T(\pi)$ *statistic as defined in Andrews*

(1993). Since $LM_T(\pi)$ converges to a tied-down Bessel process as shown in Andrews (1993), it implies $\mathbb{E}|S_i^{\nu}|^{2+\delta} < \infty$ for some $\delta \geq 0$ as $T \to \infty$.

Since the proof holds for all individuals, i, the subscript i will be dropped for notation convenience. To establish equivalence, define

$$\begin{aligned} m_{0T} &= \mathbf{X}_0 \Sigma_0^{-1} \mathbf{U}_0 \\ m_{1T} &= \mathbf{X}_1 \Sigma_1^{-1} \mathbf{U}_1 \end{aligned}$$

with

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_0 \\ \mathbf{X}_1 \end{pmatrix} \qquad \mathbf{U} = \begin{pmatrix} \mathbf{U}_0 \\ \mathbf{U}_1 \end{pmatrix}$$
$$\Sigma = \begin{pmatrix} \mathbb{E}\mathbf{U}_0\mathbf{U}_0' & 0 \\ 0 & \mathbb{E}\mathbf{U}_1\mathbf{U}_1' \end{pmatrix} = \begin{pmatrix} \Sigma_0 & 0 \\ 0 & \Sigma_1 \end{pmatrix}$$

such that $\mathbf{X}_0$ and $\mathbf{X}_1$ are $T_0 \times d$ and $T_1 \times d$ matrices, respectively. Similar definition applies to $\mathbf{U}_0$ and $\mathbf{U}_1$. Note that,

$$\begin{aligned} m_T(\Theta) &= m_{0T}(\Theta) + m_{1T}(\Theta) \\ &= \mathbf{X}'\Sigma^{-1}(\mathbf{Y} - \mathbf{X}\Theta) = \mathbf{X}'\Sigma^{-1}\mathbf{U}. \end{aligned}$$

Moreover, define

$$D = T^{-1} Var[m_T(\Theta)] \quad and \quad E = T^{-1}\mathbb{E}\frac{\partial m_T(\Theta)}{\partial \Theta}.$$

Following equation (4.4) in Andrews (1993), the $LM_T(\pi)$ statistic is defined to be

$$\frac{T}{\pi(1-\pi)} m_{0T}(\Theta,\pi)' \hat{D}^{-1}\hat{E}\left(\hat{E}'D^{-1}\hat{E}\right)^{-1}\hat{E}'\hat{D}^{-1} m_{0T}(\Theta,\pi)$$

with $\pi = T_0/T$. Note that the sample version of D and E, namely, $\hat{D}$ and $\hat{E}$ are identical in this case. That is $\hat{D} = \hat{E} = T^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{X}$ given the definition of $m_T(\Theta)$. In this case, the $LM_T(\pi)$ statistic simplifies to

$$\begin{aligned} LM_T(\pi) &= \frac{T}{\pi(1-\pi)} m_{0T}(\Theta,\pi)' \hat{E}^{-1} m_{0T}(\Theta,\pi) \\ &= \frac{1}{\pi(1-\pi)} (\mathbf{Y}_0 - \mathbf{X}_0\Theta)' \Sigma_0^{-1}\mathbf{X}_0 \left(\mathbf{X}'\Sigma^{-1}\mathbf{X}\right)^{-1} \mathbf{X}_0'\Sigma_0^{-1}(\mathbf{Y}_0 - \mathbf{X}_0\Theta) \\ &= \frac{S^0}{\pi(1-\pi)}. \end{aligned}$$

The same argument holds for S^1 by considering $LM_T(1-\pi)$ and replacing m_{0T} by m_{1T}. This completes the proof. ■

References

Anderson, J. E. and van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, 93(1):170–192.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61(4):821–856.

Andrews, D. W. K. (2003). End-of-sample instability tests. *Econometrica*, 71(6):1661–1694.

Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, 62:1383–1414.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies*, 58:277–297.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(1):135–171.

Bai, J. (2005). Panel data models with interactive fixed effects. Unpublished, New York University.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(11):191–221.

Bai, J. and Perron, P. (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., Breedon, F., and Miles, D. (2003). Life on the outside: economic conditions and prospects outside euroland. *Economic Policy*, 37:573–613.

Basher, S. A. and Carrion-i-Silvestre, J. L. (2009). Price level convergence, purchasing power parity and multiple structural breaks in panel data analysis: An application to u.s. cities. *Journal of Time Series Econometrics*, 1(1).

Bun, M. J. and Klaassen, F. J. (2002). Has the euro increased trade? Tinbergen Institute Discussion Paper TI 2002-108/2.

Carrion-i-Silvestre, J. L., del Barrio-Castro, T., and López-Bazo, E. (2005). Breaking the panels: An application to the gdp per capita. *Econometrics Journal*, 8:159–175.

Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Finance*, 20:249–272.

Chow, G. C. (1960). Tests of equality between sets of coefficient in two linear regressions. *Econometrica*, 28:591–605.

Coakley, J., Fuertes, A., and Smith, R. (2002). A principal components approach to cross-dependence in panels. Unpublished manuscript, Birkbeck College, University of London.

De Nardis, S. and Vicarelli, C. (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 Tzavalis, E. (2004). Detection of structural breaks in linear dynamic panel data models. Working Paper No. 505, Queen Mary, University of London, Department of Economics.

Dufour, J. M., Ghysels, E., and Hall, A. (1994). Generalised predictive tests and structural change analysis in econometrics. *International Economic Review*, 35:199–229.

Emerson, J. and Kao, C. (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 Kao, C. (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 Nordström, H. (2003). Trade volume effects of the euro: Aggregate and sector estimates. Manuscript, Institute for International Economic Studies.

Han, A. K. and Park, D. (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., Pesaran, H., and Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115:53–74.

Kao, C., Trapani, L., and Urga, G. (2005). Modelling and testing for structural breaks in panels with common and idiosyncratic stochastic trends. Syracuse University, Department of Economics.

Loeve, M. (1977). *Probability Theory*. Springer-Verlag, New York.

Mancini-Griffoli, T. and Pauwels, L. L. (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., Ordoñez, G., and Stein, E. (2003). The currency union effect on trade: Early evidence from emu. *Economic Policy*, 18(37):316–356.

Newey, W. and West, K. (1987). A semi positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55:703–708.

Nitsch, V. (2002). Honey, i shrunk the currency union effect on trade. *The World Economy*, 25:457–474.

Nitsch, V. and Berger, H. (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.

Pesaran, M. H. and Kapetanios, G. (2005). Alternative approaches to estimation and inference in large multifactor panels: Small sample results with an application to modelling of asset returns. CESifo Working Papers No. 1416.

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.

White, H. (1999). *Asymptotic Theory for Econometricians*. Academic Press.