

UNIVERSITY OF EXETER

Economics Department Discussion Papers Series

ISSN 1473 – 3307

Tests for Cointegration with Structural Breaks Based on Subsamples

James Davidson and Andrea Monticini

Paper number 07/04

URL: http://business-school.exeter.ac.uk/economics/papers/

Tests for Cointegration with Structural Breaks Based on Subsamples

James Davidson* and Andrea Monticini
University of Exeter

Revised, August 2007

Abstract

This paper considers tests for cointegration with allowance for structural breaks, using the extrema of residual-based tests over subsamples of the data. One motivation for the approach is to formalize the practice of data snooping by practitioners, who may examine subsamples after failing to find a predicted cointegrating relationship. Valid critical values for such multiple testing situations may be useful. The methods also have the advantage of not imposing a form for the alternative hypothesis, in particular slope vs. intercept shifts and single versus multiple breaks, and being comparatively easy to compute. A range of alternative subsampling procedures, including sample splits, incremental and rolling samples are tabulated and compared experimentally. Shiller's annual stock prices and dividends series provide an illustration.

*Corresponding author, james.davidson@exeter.ac.uk.

1 Introduction

It is not an uncommon dilemma in econometric practice to test for cointegrating relations that economic theory predicts should exist, and find that the null hypothesis of noncointegration is not rejected. Inevitably, structural change may be suspected to play a part in such outcomes. Since the correct modelling of data containing conjectured structural breaks is potentially an elaborate and tricky undertaking, there is a need for easily implemented tests of the null hypothesis of non-cointegration, against alternatives allowing cointegrating relations subject to breaks or intermittency. It must be assumed that the dates of breaks are unknown.

One method adopted in the literature has been to run regressions containing dummy variables for each possible break point within a feasible interval of the sample (say, $[0.15T]$ to $[0.85T]$, where T is sample size) and tabulate the extrema of suitable test statistics under the null hypothesis. This is the approach undertaken by Zivot and Andrews (1992) for tests of univariate integration (the I(1) hypothesis), and by Gregory and Hansen (1996) for tests of cointegration. In these schemes the whole sample is used for each statistic evaluation, although the switching parameters are nonetheless estimated from the subsamples, in effect. It is necessary to anticipate the type of break that may exist – a switch of mean, or of the slope coefficients, or to noncointegration. Discontinuities in the dummy variables present complications in deriving asymptotic distributions for such extremum statistics, as detailed by Zivot and Andrews (1992). Qu (2007) also considers this problem, from a slightly different perspective, and investigates a nonparametric procedure based on the Breitung (2002) test of I(1).

The approach explored in the present paper is to compute the usual test statistics from subsamples of the data. The breaks hypothesis asserts that there are periods in which fixed cointegrating relationships hold. The possibility of a consistent test requires that these periods must grow in length with the overall sample, and hence, the detection of a relationship should be possible by choosing an appropriate subsample for the test. In its most general form, this approach suggests some sort of incremental or rolling sequence of subsamples. This idea is investigated in the context of unit root testing by Banerjee, Lumsdaine and Stock (1992) for example. A similar approach to our own is developed in Kim (2003) although, as we remark in the sequel, Kim's asymptotic analysis appears inadequate to establish his results.

The appeal of this approach is threefold. First computational simplicity, since it only involves evaluating one of the standard cointegration statistics from the literature at each step. Second, there is no need to estimate the model under the alternative, so that the type of break, and the number of breaks, do not need to be specified. Third, following from this, there is the potential to detect breaks anywhere in the sample. For the case of a single break, at least, we suggest methods which place no limitation on the location of the break date. While statistics based on dummy shift variables may have power advantages against specific alternatives, our hope is to develop a test with general diagnostic applications. In other words, our tests can be computed routinely, and a rejection will simply point to the existence of some relationship worthy of further investigation by specific methods.

Another way to view our proposals is as a way of formalizing the practice of data snooping by testing subsamples. Suppose a practitioner inspects his/her data, conjectures the existence of a break at a particular date, and then recomputes the cointegration test for the subsample either preceding or following this date. This procedure would of course contaminate the inferences based on the test outcome, particularly if repeated with a succession of conjectured breakpoints. However, if we tabulate the null distribution of the extremum of test statistic values over a specified set of subsamples, these critical values will place valid bounds on test size. Since the extremum statistic exceeds (absolutely) any of an arbitrary sequence of cointegration tests based on snooping subsample regressions, its tabulation provides a valid test based on the criterion

"one or more of the subsamples yields a rejection".

One very simple procedure that involves no snooping is to split the sample into equal halves, and take the extremum of the cointegration statistics for each half. This technique can of course be implemented very easily, given access to the tabulation of the extremum of two independent statistics. Suppose we can assume at most one break, under the alternative. Then, this must occur in either the first half or the second half of the sample. Unless the series are actually non-cointegrated under a regime persisting from more than half the sample, even this simple procedure must yield a consistent test for cointegration.

A more ambitious scheme is to compute the statistic incrementally, and tabulate the extremum over all subsamples defined by initial or terminal observations in given ranges. In the case of a single break, note that the indicated range of terminal dates to consider is either $[T/2]+1, ..., T$, with 1 for the initial date, or alternatively initial dates $1, ..., [T/2]$, with T for the terminal date. Again, since a single break must occur in either the first half or the second half of the sample, there should be no need to consider other ranges than these two. In other words, to detect a single break there is no advantage in starting the recursion, either forwards or backwards, on a subsample of fewer than half the observations. The situation is different for alternatives with breaks at two or more dates. Then we may face a trade-off between the advantages of considering relatively short subsamples, and the disadvantages of this technique when the overall sample is not large, so that asymptotic approximations are correspondingly poor.

The paper is organized as follows. Section 2 sets out the framework of the analysis and Section 3 introduces the alternative tests under consideration. Section 4 derives their asymptotic distributions under the null hypothesis that the data are non-cointegrated I(1) processes. Tabulations computed by Monte Carlo for the cases of one and two regressors are reported. Section 5 investigates the performance of the tests in simulations of a range of bivariate alternatives. One feature of our experimental design is that we randomize the location of the breaks, drawing these from the uniform [0,1] distribution as a proportion of sample size. Therefore, our power comparisons can be viewed as integrating out the break locations and reporting an average performance over possible break patterns. Since there is no reason to think that the location of a break is likely to have a systematic relationship with the span of the sample in applications, this scheme offers the most useful comparison of alternative techniques. Section 6 applies the techniques to Shiller's well-known 1871-2004 annual stock prices and dividends series, and Section 7 concludes the paper.

2 Models with Structural Shifts

Let $\boldsymbol{x}_t = (x_{1t}, \boldsymbol{x}'_{2t})'$ be a p-vector I(1) process, such that

$$\boldsymbol{x}_t = \boldsymbol{x}_0 + \sum_{s=1}^{t} \boldsymbol{u}_s$$

where $E|\boldsymbol{x}_0| < \infty$, $E(\boldsymbol{u}_t) = \mathbf{0}$, $\mathbf{\Sigma} = E(\boldsymbol{u}_t \boldsymbol{u}'_t)$ and $\mathbf{\Lambda} = \sum_{j=1}^{\infty} E(\boldsymbol{u}_t \boldsymbol{u}'_{t-j}) < \infty$, such that

$$T^{-1} E(\boldsymbol{x}_T - \boldsymbol{x}_0)(\boldsymbol{x}_T - \boldsymbol{x}_0)' \rightarrow \mathbf{\Omega} = \mathbf{\Sigma} + \mathbf{\Lambda} + \mathbf{\Lambda}'.$$

A question routinely at issue in econometrics is whether $\mathbf{\Omega}$ is singular, in which case the process said to be cointegrated in the sense of Engle and Granger (1987). By implication there exists a vector $\boldsymbol{\beta}_0$ such that

$$z_t = (x_{1t} - x_{10}) - \boldsymbol{\beta}'_0 (\boldsymbol{x}_{2t} - \boldsymbol{x}_{20})$$

$$= x_{1t} - \alpha_0 - \boldsymbol{\beta}_0' \boldsymbol{x}_{2t} \tag{2.1}$$

is an I(0) process with a mean of zero, where $\alpha_0 = x_{10} - \boldsymbol{\beta}_0' \boldsymbol{x}_{20}$, a constant under the distribution conditional on $\boldsymbol{x}_0$. The normalization on the element x_{1t} is arbitrary but it's convenient to assume that $\text{rank}(\boldsymbol{\Omega}_{22}) = p - 1$ in the partition

$$\boldsymbol{\Omega} = \begin{bmatrix} \omega_{11} & \boldsymbol{\omega}_{12} \\ \boldsymbol{\omega}_{12}' & \boldsymbol{\Omega}_{22} \end{bmatrix}. \tag{2.2}$$

Strictly, our characterization of cointegration is that there exists a partition of the data having these properties.[1] Note that non-cointegration, in this sense, implies $z_t \sim \text{I}(1)$ for all choices of $\boldsymbol{\beta}_0$. The case with $\boldsymbol{\beta}_0 = \mathbf{0}$ and $x_{1t} \sim \text{I}(0)$ is not cointegration as usually understood, although it does of course imply $\boldsymbol{\Omega}$ singular, having first row and column zero.

A case where cointegration fails to hold as defined, but may be said to exist more generally, is where

$$z_t = x_{1t} - (\alpha_0 + \alpha_1 \varphi_t) - (\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \varphi_t)' \boldsymbol{x}_{2t} \tag{2.3}$$

defines an I(0) zero mean sequence, where

$$\varphi_t = \begin{cases} 1, & t \in A \\ 0, & t \notin A \end{cases}$$

and A is a specified subset of the sample, generally consisting of blocks of contiguous observations. Consider for clarity the hypothesis of a single break, so that $A = \{[Tr] + 1, \ldots, T\}$ for $0 < r < 1$. It is easy to elaborate the following story to multiple breaks without altering the essentials.

Suppose first that $\boldsymbol{\beta}_1 \neq \mathbf{0}$. Then we may be able to say that

$$([Tr_1])^{-1} E(\boldsymbol{x}_{[Tr]} - \boldsymbol{x}_0)(\boldsymbol{x}_{[Tr]} - \boldsymbol{x}_0)' \to \boldsymbol{\Omega}_1 \tag{2.4a}$$

$$(T - [Tr_1])^{-1} E(\boldsymbol{x}_T - \boldsymbol{x}_{[Tr]})(\boldsymbol{x}_T - \boldsymbol{x}_{[Tr]})' \to \boldsymbol{\Omega}_2 \tag{2.4b}$$

but $\boldsymbol{\Omega}_1 \neq \boldsymbol{\Omega}_2$, and although both of $\boldsymbol{\Omega}_0$ and $\boldsymbol{\Omega}_1$ are singular,

$$T^{-1} E(\boldsymbol{x}_T - \boldsymbol{x}_0)(\boldsymbol{x}_T - \boldsymbol{x}_0)' \to \boldsymbol{\Omega} = r\boldsymbol{\Omega}_1 + (1 - r)\boldsymbol{\Omega}_2$$

is in general nonsingular. The covariance structure of the multivariate process changes at date $[Tr]$ with an accompanying intercept shift of magnitude

$$\alpha_1 = -\boldsymbol{\beta}_1' \boldsymbol{x}_{20}.$$

We may call this a regime shift. By implication, the standard cointegration tests are inconsistent in this case, notwithstanding that a form of cointegration exists.

Alternatively we may have what Gregory and Hansen (1996) call a level shift. One way to understand this is an autonomous shift in the cointegrating intercept at date $[Tr]$. However, it may be useful to think of it as arising when a shock hitting the system at date $[Tr]$ breaks the covariance structure, temporarily but with permanent effect. Writing from (2.1) the implicit relation

$$u_{1t} = \boldsymbol{\beta}_0' \boldsymbol{u}_{2t} + \Delta z_t$$

[1] A singular $\boldsymbol{\Omega}_{22}$ implies the independent existence of cointegration among the elements of $\boldsymbol{x}_{2t}$. This is a complication we avoid by assumption, but in the context of the present analysis we might meet it by conducting the analysis on $\boldsymbol{x}_{2t}$, and on subsets of the data generally. Any cointegrated collection of time series contains one or more 'irreducible' subsets of cointegrating rank 1; see Davidson (1998) for details.

observe that under cointegration the random processes driving the system are linked implicitly into a long-run singular relationship by the 'over-differenced' residual component. A shock to an element of $\boldsymbol{u}_t$ which is not linked to the other elements though the cointegrating relation shifts the relation permanently. In other words, if we define a process

$$\delta_t = \begin{cases} \alpha_1, & t = [Tr] + 1, \\ 0, & \text{otherwise}, \end{cases}$$

the difference equation

$$u_{1t} = \boldsymbol{\beta}_0' \boldsymbol{u}_{2t} + \Delta z_t + \delta_t$$

has solution

$$z_t = \begin{cases} x_{1t} - \alpha_0 - \boldsymbol{\beta}_0' \boldsymbol{x}_{2t} & t \leq [Tr], \\ x_{1t} - (\alpha_0 + \alpha_1) - \boldsymbol{\beta}_0' \boldsymbol{x}_{2t} & t > [Tr]. \end{cases}$$

A succession of such shocks would destroy the cointegating relationship since they would integrate to an additional stochastic trend, but an isolated shock effects a level shift. Unlike the regime shift case, the process defined in (2.1) has bounded variance in the limit, but under the distribution conditional on $\{\delta_t\}$ it is nonstationary with time-varying mean. Since the break is of small order relative to $\boldsymbol{x}_t$, note that $\boldsymbol{\Omega}_1 = \boldsymbol{\Omega}_2$ (singular) in (2.4). In this sense, the process cointegrates normally and conventional cointegration tests are nominally consistent. However, these are tabulated under the stationarity assumption, and may have low power in finite samples where the break may effectively mimic an I(1) component.

One further case that we may wish consider is where (say) $\boldsymbol{\Omega}_0$ is singular but $\boldsymbol{\Omega}_1$ is nonsingular, so that cointegration exists in only a portion of the sample. The main point to emphasize here is that this is distinct from the case $\boldsymbol{\beta}_1 = -\boldsymbol{\beta}_0$. In the latter case there is no cointegration as such in the second period, but x_{1t} becomes a stationary process, and $\boldsymbol{\Omega}_1$ is singular through having first row and column zero. By contrast, the former case is one in which there exists no stationary linear combination when $t \in A$. In the error-correction representation of the system which we invoke in Section 5, where z_t assumes the role of an error-correction term, we may represent this case by writing

$$z_t = (1 - \varphi_t)(x_{1t} - \alpha_0 - \boldsymbol{\beta}_0' \boldsymbol{x}_{2t}). \tag{2.5}$$

3 Tests for Noncointegration

Gregory and Hansen (1996) attack the problem of detecting cointegration in the presence of single breaks by fitting the breaking-cointegration models to the data, for all choices of r_1 in a suitable interval, and tabulating the extremum of their resulting sequence of cointegration tests statistics under the null hypothesis of noncointegration (and hence, no breaks). By contrast, we derive statistics that can be applied to the usual putative cointegrating regression (2.1), and hence avoid the need to specify the form of the breaks model. All our tests should all be able to detect cointegration with a single break consistently, so long as the series are cointegrated on both sides of the break. Power to detect intermittent cointegration, where the relation breaks down entirely in one regime, may also be available subject to the location of the break. The tests can also have power against some multiple-break alternatives. At a minimum, what is required for consistency of at least one test of the type proposed here is that there should exist a segment of the sample of length $O(T)$ satisfying a cointegrating relation.

For expositional simplicity, let t_T stand initially for the Dickey-Fuller statistic without augmentation for autocorrelation corrections, as would be appropriate for the case where $\boldsymbol{\Lambda} = \boldsymbol{0}$. This was one of the procedures first suggested by Engle and Granger (1987) and its asymptotic

properties of this test are well-known; see e.g. Engle and Yoo (1987), Phillips and Ouliaris (1990) or for a general exposition, Davidson (2000) Chapter 15.3. Consider the subsample statistics defined for $\lambda_1 \in [0,1)$ and $\lambda_2 \in (\lambda_1, 1]$ as

$$t_T(\lambda_1,\lambda_2) = \frac{\sum_{t=[T\lambda_1]+1}^{[T\lambda_2]} \hat{z}_{t-1}(\lambda_1,\lambda_2)\Delta\hat{z}_t(\lambda_1,\lambda_2)}{s(\lambda_1,\lambda_2)\left(\sum_{t=[T\lambda_1]+1}^{[T\lambda_2]} \hat{z}_{t-1}(\lambda_1,\lambda_2)^2\right)^{1/2}} \tag{3.1}$$

where

$$\widehat{z}_t(\lambda_1,\lambda_2) = x^*_{1t}(\lambda_1,\lambda_2) - \hat{\boldsymbol{\beta}}(\lambda_1,\lambda_2)'\boldsymbol{x}^*_{2t}(\lambda_1,\lambda_2). \tag{3.2}$$

for $t = [T\lambda_1]+1,\ldots,[T\lambda_2]$, and

$$s^2(\lambda_1,\lambda_2) = \frac{1}{[T\lambda_2]-[T\lambda_1]} \sum_{t=[T\lambda_1]+1}^{[T\lambda_2]} \widehat{e}_t(\lambda_1,\lambda_2)^2 \tag{3.3}$$

where $\widehat{e}_t(\lambda_1,\lambda_2)$ is the residual from the regression of $\Delta\widehat{z}_t(\lambda_1,\lambda_2)$ on $\widehat{z}_{t-1}(\lambda_1,\lambda_2)$. The data subsamples, expressed for convenience in (subsample-) mean deviation form, are

$$\boldsymbol{x}^*_t(\lambda_1,\lambda_2) = \boldsymbol{x}_t - \frac{1}{[T\lambda_2]-[T\lambda_1]} \sum_{s=[T\lambda_1]+1}^{[T\lambda_2]} \boldsymbol{x}_s, \quad t = [T\lambda_1]+1,\ldots,[T\lambda_2].$$

Also,

$$\hat{\boldsymbol{\beta}}(\lambda_1,\lambda_2) = \left(\sum_{t=[T\lambda_1]+1}^{[T\lambda_2]} \boldsymbol{x}^*_{2t}(\lambda_1,\lambda_2)\boldsymbol{x}^*_{2t}(\lambda_1,\lambda_2)'\right)^{-1} \sum_{t=[T\lambda_1]+1}^{[T\lambda_2]} \boldsymbol{x}^*_{2t}(\lambda_1,\lambda_2)x^*_{1t}(\lambda_1,\lambda_2). \tag{3.4}$$

To allow for the presence of short-run autocorrelation in the differenced variables, the DF statistic might be augmented in the usual way by projecting $\hat{z}_{t-1}$ onto lags of $\Delta\hat{z}_t$. However, this ADF statistic presents some difficulties for asymptotic analysis (see Phillips and Ouliaris 1990) and a more amenable approach is to adopt the Phillips-Perron (1988) non-parametric correction

$$C_l(\lambda_1,\lambda_2) = \frac{1}{[T\lambda_2]-[T\lambda_1]} \sum_{j=1}^{l(\lambda_1,\lambda_2)} w_{jl} \sum_{t=[T\lambda_1]+j}^{[T\lambda_2]} \widehat{e}_t(\lambda_1,\lambda_2)\widehat{e}_{t-j}(\lambda_1,\lambda_2)$$

where, adapting the Newey and West (1987) formulation for example, we could set $w_{lj} = 1 - j/(1+l(\lambda_1,\lambda_2))$ and $l(\lambda_1,\lambda_2) = O(T^{1/3})$. The subsample Phillips-Perron (PP) statistic is

$$\begin{aligned}\hat{Z}_T(\lambda_1,\lambda_2) &= \frac{\sum_{t=[T\lambda_1]+1}^{[T\lambda_2]} \left(\hat{z}_{t-1}(\lambda_1,\lambda_2)\Delta\hat{z}_t(\lambda_1,\lambda_2) - C_l(\lambda_1,\lambda_2)\right)}{S_l^2(\lambda_1,\lambda_2)\left(\sum_{t=[T\lambda_1]+1}^{[T\lambda_2]} \hat{z}_{t-1}(\lambda_1,\lambda_2)^2\right)^{1/2}} \\ &= \frac{s^2(\lambda_1,\lambda_2)}{S_l^2(\lambda_1,\lambda_2)} t_T(\lambda_1,\lambda_2) - \frac{([T\lambda_1]-[T\lambda_2])C_l(\lambda_1,\lambda_2)}{S_l^2(\lambda_1,\lambda_2)\left(\sum_{t=[T\lambda_1]+1}^{[T\lambda_2]} \hat{z}_{t-1}(\lambda_1,\lambda_2)^2\right)^{1/2}}\end{aligned} \tag{3.5}$$

where $S_l^2(\lambda_1,\lambda_2)^2 = s^2(\lambda_1,\lambda_2) + 2C_l(\lambda_1,\lambda_2)$. This statistic is asymptotically equivalent to $t_T(\lambda_1,\lambda_2)$ in case $\boldsymbol{u}_t$ is serially uncorrelated.

We adopt $q_T(\lambda_1,\lambda_2)$ in the sequel as a generic notation for a cointegration statistic, for which specific cases such as $t_T(\lambda_1,\lambda_2)$ and $\hat{Z}_T(\lambda_1,\lambda_2)$ can be substituted in formulae. We consider the

class of tests based on extreme values of subsample statistics over alternative of sets of (λ_1, λ_2) values. Here are some cases.

$$\{\lambda_1, \lambda_2\} \in \Lambda_S = \{\{0, \tfrac{1}{2}\}, \{\tfrac{1}{2}, 1\}\} \tag{3.6a}$$

$$\{\lambda_1, \lambda_2\} \in \Lambda_{0f} = \{0, [\lambda_0, 1]\} \tag{3.6b}$$

$$\{\lambda_1, \lambda_2\} \in \Lambda_{0b} = \{[0, 1 - \lambda_0], 1\} \tag{3.6c}$$

$$\{\lambda_1, \lambda_2\} \in \Lambda_{0R} = \{[0, 1 - \lambda_0], \lambda_0 + \lambda_1]\} \tag{3.6d}$$

Thus, Λ_S represents a simple split sample, with just two elements. Λ_{0f} and Λ_{0b} are each defined for a constant $\lambda_0 > 0$, and define forwards- and backwards-running incremental samples of minimum length $[T\lambda_0]$ and maximum length T. The case Λ_{0R} defines rolling samples of fixed length $[T\lambda_0]$. In addition, there are the cases

$$\Lambda_S^* = \Lambda_S \cup \{0, 1\} \tag{3.7a}$$

$$\Lambda_{0R}^* = \Lambda_{0R} \cup \{0, 1\} \tag{3.7b}$$

where the sets of subsamples are augmented by the full sample. The cases Λ_{0f} and Λ_{0b} already contain the elements $\{0, 1\}$, so we don't need to make their inclusion explicit in these cases.

The tests we consider are therefore of three types.[2] The split-sample tests are

$$Q_S = \min_{\{\lambda_1, \lambda_2\} \in \Lambda_S} q_T(\lambda_1, \lambda_2) \tag{3.8}$$

$$Q_S^* = \min_{\{\lambda_1, \lambda_2\} \in \Lambda_S^*} q_T(\lambda_1, \lambda_2). \tag{3.9}$$

The incremental tests take the form

$$Q_I(\lambda_0) = \inf_{\lambda \in \Lambda_{0f} \cup \Lambda_{0b}} q_T\left(\lambda_1, \lambda_2\right). \tag{3.10}$$

and the rolling sample tests are of the forms

$$Q_R(\lambda_0) = \inf_{\lambda \in \Lambda_{0R}} q_T\left(\lambda_1, \lambda_2\right) \tag{3.11}$$

$$Q_R^*(\lambda_0) = \inf_{\lambda \in \Lambda_{0R}^*} q_T\left(\lambda_1, \lambda_2\right) \tag{3.12}$$

4 Asymptotic Analysis

Define $\boldsymbol{X}_T(r) = T^{-1/2}\boldsymbol{x}_{[Tr]}$, and $\boldsymbol{J}_T(r) = T^{-1}\sum_{t=2}^{[Tr]} \boldsymbol{x}_{t-1}\boldsymbol{u}_t'$. Let $D[0,1]^m$ denote the space of m-dimensional cadlag functions on the unit interval, that is to say, functions having the property of right-continuity and a left limit at each point of $(0, 1]$. Also let $\boldsymbol{B}$ be a p-vector of Brownian motions on $[0, 1]$ with $\boldsymbol{B}(0) = \boldsymbol{0}$ a.s., and covariance matrix $\boldsymbol{\Omega}$, and let $\boldsymbol{J}$ denote a $p \times p$ matrix-valued process defined by $\boldsymbol{J}(r) = \int_0^r \boldsymbol{B} d\boldsymbol{B}' + r\boldsymbol{\Lambda}$ for $r \in [0, 1]$. Our fundamental assumptions will be the following.

Assumption 1

$$(\boldsymbol{X}_T, \operatorname{Vec} \boldsymbol{J}_T) \xrightarrow{d} (\boldsymbol{B}, \operatorname{Vec} \boldsymbol{J}) \tag{4.1}$$

where ' $\xrightarrow{d}$ *' denotes joint weak convergence with respect to the Skorokhod metric on* $D[0,1]^{p(1+p)}$.

[2] The authors' Ox code to compute these tests based on PP and ADF statistics is available at http://www.timeseriesmodelling.com

Assumption 2

$$\frac{1}{T}\sum_{j=1}^{l(0,1)} w_{lj} \sum_{t=1+j}^{T} \boldsymbol{u}_t \boldsymbol{u}'_{t-j} \xrightarrow{pr} \boldsymbol{\Lambda}$$

*where $l()$ and w_{lj} are defined in (**??**) and '$\xrightarrow{pr}$' denotes convergence in probability.*

Setting high-level assumptions avoids the question of specifying sufficient conditions on the underlying discrete process $\{\boldsymbol{u}_t\}$, although for the benefit of practitioners it can suffice to say that such conditions match those needed for conventional cointegration analysis. A large literature dealing with these questions exists, and we prefer to focus attention on the specific problem of the limiting behaviour of our extremum statistics. Note that although the sequence elements $\boldsymbol{X}_T$ and $\operatorname{Vec}\boldsymbol{J}_T$ are cadlag, the limit processes $\boldsymbol{B}$ and $\operatorname{Vec}\boldsymbol{J}$ are elements respectively of $C[0,1]^p$ and $C[0,1]^{p^2}$, with probability 1, where $C[0,1]^m \subset D[0,1]^m$ is the space of m-dimensional continuous functions on the unit interval (see e.g. Kurtz 2001, Corollary 5.4). In $C[0,1]$ the Skorokhod topology is equivalent to the uniform topology and, as pointed out by Billingsley (1968, page 112), convergence to a continuous limit process in the Skorokhod topology is equivalent to uniform convergence.

Results on stochastic integral convergence (i.e., the case $\boldsymbol{J}_T$) are commonly given in the literature in point-wise form, by considering the random matrix $\boldsymbol{J}_T(r)$ for some r, typically $r = 1$. See for example Chan and Wei (1988), and De Jong and Davidson (2000), for results of this type. However, Hansen (1992) deploys results from Kurtz and Protter (1991) (see also Kurtz and Protter 1995 for a detailed exposition of the theory) which establish weak convergence of the cadlag process $\boldsymbol{J}_T$ in the sense asserted in (4.1). The conditions cited for Hansen's (1992) Theorem 4.1 are sufficient for (4.1) to hold. An alternative approach allowing different lower-level assumptions might be to take a pointwise convergence result in combination with an argument showing tightness of the sequence $\{\operatorname{Vec}\boldsymbol{J}_T\}$, by verifying a condition such as Billingsley (1968) Theorem 15.6, for example. We do not pursue these extensions here, however.

Given Assumptions 1 and 2, the asymptotics of our procedures under the null hypothesis can be derived as extensions of the results developed by Zivot and Andrews (1992) and Gregory and Hansen (1996). In the presence of autocorrelation ($\boldsymbol{\Lambda} \neq \boldsymbol{0}$) we follow the latter authors in working with the Phillips-Perron cointegration statistic (3.5). First, we establish the limiting distributions of the statistics $\hat{Z}_T(\lambda_1, \lambda_2)$ under the null hypothesis. The key point to be established is that for given values λ_1 and λ_2, these are continuous functionals of the limit processes specified in (4.1).

It is convenient to adapt the approach and notation of Davidson (2000), Chapter 15, modified to the subsample setup. Let $\boldsymbol{W}$ denote a standard p-vector Brownian motion having variance matrix $\boldsymbol{I}_p$. Then, define

$$\boldsymbol{\xi}(\lambda_1, \lambda_2) = (1, \ -\boldsymbol{\zeta}(\lambda_1, \lambda_2)')' \quad (p \times 1) \tag{4.2}$$

where

$$\boldsymbol{\zeta}(\lambda_1, \lambda_2) = \left(\int_{\lambda_1}^{\lambda_2} \boldsymbol{W}_2^* \boldsymbol{W}_2^{*\prime} dr \right)^{-1} \int_{\lambda_1}^{\lambda_2} \boldsymbol{W}_2^* W_1^* dr. \tag{4.3}$$

and we define (omitting the dependence on λ_1, λ_2 for brevity)

$$\boldsymbol{W}^*(r) = \boldsymbol{W}(r) - \boldsymbol{W}(\lambda_1) - \int_{\lambda_1}^{\lambda_2} \boldsymbol{W} ds, \qquad \lambda_1 \leq r \leq \lambda_2 \tag{4.4}$$

Similarly, let

$$\int_{\lambda_1}^{\lambda_2} \boldsymbol{W}^* d\boldsymbol{W}' = \int_{\boldsymbol{\lambda_1}}^{\boldsymbol{\lambda_2}} (\boldsymbol{W}(r) - \boldsymbol{W}(\lambda_1)) d\boldsymbol{W}' - (\boldsymbol{W}(\boldsymbol{\lambda}_2) - \boldsymbol{W}(\boldsymbol{\lambda}_1)) \int_{\lambda_1}^{\lambda_2} \boldsymbol{W} ds. \tag{4.5}$$

If $\boldsymbol{W} = \boldsymbol{L}'^{-1}\boldsymbol{B}$ where $\boldsymbol{\Omega} = \boldsymbol{L}'\boldsymbol{L}$, further note that

$$\int_{\boldsymbol{\lambda_1}}^{\boldsymbol{\lambda_2}} \boldsymbol{W} d\boldsymbol{W}' = \boldsymbol{L}^{-1}[\boldsymbol{J}(\lambda_2) - \boldsymbol{J}(\lambda_1) - (\lambda_2 - \lambda_1)\boldsymbol{\Lambda}]\boldsymbol{L}'^{-1}. \tag{4.6}$$

Therefore, adapting the standard development, such as that given in Davidson (2000) leading to equation (15.3.20), for example, we may assert that under Assumptions 1 and 2,

$$\begin{aligned} \hat{Z}_T(\lambda_1, \lambda_2) &\xrightarrow{d} \tau(\lambda_1, \lambda_2) \\ &= \frac{\boldsymbol{\xi}(\lambda_1, \lambda_2)' \int_{\lambda_1}^{\lambda_2} \boldsymbol{W}^* d\boldsymbol{W} \boldsymbol{\xi}(\lambda_1, \lambda_2)}{\sqrt{\boldsymbol{\xi}(\lambda_1, \lambda_2)' \boldsymbol{\xi}(\lambda_1, \lambda_2)} \sqrt{\boldsymbol{\xi}(\lambda_1, \lambda_2)' \int_{\lambda_1}^{\lambda_2} \boldsymbol{W}^* \boldsymbol{W}^{*\prime} dr \boldsymbol{\xi}(\lambda_1, \lambda_2)}} \end{aligned} \tag{4.7}$$

where the equality defines the limit random variable $\tau(\lambda_1, \lambda_2)$. Note that marginal distribution of $\tau(\lambda_1, \lambda_2)$ is independent of nuisance parameters, and in particular, does not depend on λ_1 and λ_2 on account of the self-similarity of Brownian motion. The range of integration determines the variances of the stochastic integrals appearing in the expression as a function of (λ_1, λ_2), but these scale factors cancel in the ratio.

Therefore, the problem is to establish, for specified sets Λ, the weak convergence

$$\inf_{\{\lambda_1, \lambda_2\} \in \Lambda} \hat{Z}_T(\lambda_1, \lambda_2) \xrightarrow{d} \inf_{\{\lambda_1, \lambda_2\} \in \Lambda} \tau(\lambda_1, \lambda_2). \tag{4.8}$$

Theorem 4.1 *Under Assumptions 1 and 2, the weak convergence specified in (4.8) holds for the cases where Λ is one of Λ_S, Λ_S^*, $\Lambda_{0I} = \Lambda_{0f} \cup \Lambda_{0b}$, Λ_{0R} and Λ_{0R}^*, for any $\lambda_0 > 0$.*

The proof is given in the Appendix. Note that the limiting distribution in (4.7) is shared by the subsample PP and ADF statistics, as shown by Phillips and Ouliaris (1990), and we take this result as given under our assumptions. In practice, Theorem 4.1, which simply establishes the extension from the pointwise to the uniform case, applies equally to either test. The result could be extended to other statistics, such as the trace test, with minor modifications. We remark that the analysis of Kim (2003) claims results similar to our own, citing theorems given in Chan and Wei (1988) as sufficient authority. However, our proof makes clear that these results do not suffice here.[3]

We have tabulated large sample critical values for these distributions by simulation. The results are reported in Table 1 for the cases of one and two regressors, with and without linear trend, and sample size $T = 1000$. The extremum statistics based on $t_T(\lambda_1, \lambda_2)$ in (3.1) were used for speed of calculation, since the limiting distributions are as shown in (4.8) when the generated series are Gaussian random walks. The tables are constructed in each case from $40,000$ replications. To construct these tables, and also to compute the simulations of the next section, the extremum statistics were in practice calculated by evaluating the statistics for every fifth case of the subsamples in the specified ranges, and returning the minimum of these values. This is done to reduce the computing time required, but given the continuity of the statistics on $C[0,1]$ as functions of λ_1 and λ_2, as established in Theorem 4.1, the approximations involved should be small in a large sample.

[3]Theorem 2.3 of Chan and Wei (1988) gives a form of continuous mapping theorem for a process $Z_n(t) = \int_0^t f(Y_n(u))du$ where Y_n is a vector converging to Brownian motion. However, note that the numerator of (4.7) is not of this form, even when generalized to a variable initial time. Theorem 2.4 of the same source gives a result for stochastic integral convergence, but this is a pointwise result specifying a *fixed* terminal date $t = 1$. Thus, the stochastic integral convergence proved relates to a random variable $\int_0^1 HdW$ where W is a Brownian motion – not an a.s. continuous stochastic process on the unit interval, for which the result is required here.

Regressors	**Type**	λ_0	**50%**	**10%**	**5%**	**2.5%**	**1%**
1	DF	–	−2.059	−3.053	−3.358	−3.614	−3.903
	Q_S	–	−2.493	−3.356	−3.610	−3.851	−4.120
	Q_S^*	–	−2.617	−3.463	−3.718	−3.938	−4.228
	$Q_I(\lambda_0)$	0.5	−3.224	−4.067	−4.327	−4.554	−4.846
		0.35	−3.388	−4.194	−4.452	−4.667	−4.935
		0.2	−3.562	−4.325	−4.568	−4.767	−5.032
		0.1	−3.702	−4.433	−4.648	−4.863	−5.143
	$Q_R(\lambda_0)$	0.5	−3.344	−4.143	−4.392	−4.614	−4.864
	$Q_R^*(\lambda_0)$	0.5	−3.363	−4.152	−4.402	−4.623	−4.873
1 + Trend	DF		−2.552	−3.503	−3.793	−4.051	−4.358
	Q_S	–	−2.959	−3.791	−4.061	−4.297	−4.578
	Q_S^*	–	−3.082	−3.909	−4.165	−4.399	−4.666
	$Q_I(\lambda_0)$	0.5	−3.673	−4.480	−4.745	−4.956	−5.221
		0.35	−3.840	−4.602	−4.860	−5.071	−5.329
		0.2	−3.999	−4.735	−4.969	−5.177	−5.435
	$Q_R(\lambda_0)$	0.5	−3.781	−4.563	−4.803	−5.017	−5.294
	$Q_R^*(\lambda_0)$	0.5	−3.794	−4.563	−4.803	−5.042	−5.294
2	DF	–	−2.069	−3.054	−3.361	−3.608	−3.915
	Q_S	–	−2.478	−3.355	−3.618	−3.867	−4.175
	Q_S^*	–	−2.614	−3.466	−3.726	−3.963	−4.258
	$Q_I(\lambda_0)$	0.5	−3.220	−4.079	−4.341	−4.571	−4.854
		0.35	−3.387	−4.200	−4.460	−4.679	−4.950
		0.2	−3.559	−4.323	−4.565	−4.780	−5.050
	$Q_R(\lambda_0)$	0.5	−3.350	−4.154	−4.405	−4.636	−4.888
	$Q_R^*(\lambda_0)$	0.5	−3.360	−4.164	−4.405	−4.636	−4.888
2+Trend	DF	–	−2.549	−3.502	−3.801	−4.057	−4.342
	Q_S	–	−2.954	−3.795	−4.053	−4.301	−4.560
	Q_S^*	–	−3.090	−3.912	−4.165	−4.397	−4.660
	$Q_I(\lambda_0)$	0.5	−3.680	−4.502	−4.755	−4.956	−5.240
		0.35	−3.845	−4.623	−4.858	−5.073	−5.339
		0.2	−4.007	−4.746	−4.973	−5.179	−5.445
	$Q_R(\lambda_0)$	0.5	−3.780	−4.569	−4.797	−5.024	−5.281
	$Q_R^*(\lambda_0)$	0.5	−3.799	−4.578	−4.811	−5.025	−5.288

Table 1: Critical Values for Dickey-Fuller/Phillips-Perron Extremum Tests

Since the observations are partitioned according to fixed proportions of of the sample size, establishing the consistency of these tests is a fairly straightforward extension of the standard arguments of Phillips and Ouliaris (1990) and other authors, provided that the passage to the limit is specified appropriately. Consider the sets Λ, each consisting of a collection of subintervals of $[0,1]$. Also consider a sequence $\{A_T,\ t \geq 1\}$ where A_T is any such interval, and write TA_T for the corresponding set of integers from $\{1,\ldots,T\}$.

Assumption 3 *Given a sequence of samples* $\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_T;\ T \geq 1\}$, $\exists\ T_0 < \infty$ *such that for all* $T \geq T_0$ *the observations* $\{\boldsymbol{x}_t : t \in TA_T\}$ *form a cointegrated sequence, where* $\{A_T, t \geq 1\}$ *is a sequence of intervals of* $[0,1]$ *having positive width.*

Let Λ denote one of the sets Λ_S, Λ_S^*, $\Lambda_{0I} = \Lambda_{0f} \cup \Lambda_{0b}$, Λ_{0R} and Λ_{0R}^*, as specified in Theorem 4.1, and Q_Λ the corresponding infimum statistic, as defined by one of (3.8), (3.9), (3.10), (3.11) and (3.12).

Theorem 4.2 *If Assumptions 1, 2 and 3 hold, then the subsample test* Q_Λ *is consistent if and only if* $\exists\ \{\lambda_1,\lambda_2\} \in \Lambda$ *such that* $[\lambda_1,\lambda_2] \subseteq A_{T_1}$ *and* $T_0 \leq T_1 < \infty$.

This consistency result calls only for one of the tested subsets to be contained in a cointegrated subset, and in particular, it shows consistency for the cases of 'normal' cointegration ($\boldsymbol{\Omega}_1 = \boldsymbol{\Omega}_2$), broken cointegration ($\boldsymbol{\Omega}_1 \neq \boldsymbol{\Omega}_2$, both singular) and partial cointegration ($\boldsymbol{\Omega}_1 \neq \boldsymbol{\Omega}_2$, one nonsingular), subject to Assumption 3 holding. In the case of a single break, such that $A_T = [0, r_T]$ and/or $A_T = [r_T, 1]$ for $0 \leq r_T \leq 1$, where cointegration holds in both parts of the sample, note that all the tests are consistent for *any* sequence $\{r_T\}$.

If there is non-cointegration in one of the two parts, say $[0, r_T]$ the fact that an observation $\boldsymbol{x}_t$ cannot change its 'status' as the sample increases makes it necessary to specify how this increase takes place. The usual construction, by adding observations to the end of the sample, is not very useful because it implies $r_T \to 0$, so either no tests are consistent in the case $A_T = [0, r_T]$, or all are consistent in the case $A_T = [r_T, 1]$. It is better to consider an ensemble of increasing samples with $r_T = r$, all T, and in this case we can say that that Q_S and Q_S^* are consistent in the case $r \geq \frac{1}{2}$, whereas $Q_I(\lambda_0)$ is consistent for $r \geq \lambda_0$ when $[0, r]$ is cointegrating, and $r \leq (1 - \lambda_0)$ when $[r, 1]$ is cointegrating. The case of two or more breaks with cointegration holding in all subsamples can always be subsumed in these latter cases, because there always exists a single division into a cointegrating subsample and its noncointegrating complement. However, $Q_R(\lambda_0)$ and $Q_R^*(\lambda_0)$ are consistent, as the other tests are not, for the case where only the subsample $[r_1, r_2]$ is cointegrating., for $0 < r_1 < r_2 < 1$.

5 Monte Carlo Experiments

With a variety of tests and possible break alternatives to compare, we are restricted in the number of cases that can feasibly be studied by simulation. We decided to limit these to bivariate models, and to consider the single sample size $T = 200$. The justification for the latter decision is that we are interested primarily in the *relative* performance of tests, both to each other and to the conventional (full sample) cointegration test. These relative performances are not likely to depend unexpectedly on sample size, even if the absolute powers do so. Our software is available to readers interested in more detailed comparisons.

The cointegrated series were generated by a vector error correction model with the structure

$$\begin{aligned} \Delta x_{1t} &= \gamma_1 z_{t-1} + u_{1t} \\ \Delta x_{2t} &= \gamma_2 z_{t-1} + u_{2t} \end{aligned}$$

where u_{1t} and u_{2t} are independent standard normals, and in any period t

$$z_t = x_{1t} - \alpha - \beta x_{2t}. \tag{5.1}$$

In this setup,

$$z_t = (1 + \gamma_1 - \beta\gamma_2)z_{t-1} + u_{1t} - \beta u_{2t}. \tag{5.2}$$

and the model is therefore cointegrating if $\beta \neq 0$ and $-2 < \gamma_1 - \gamma_2\beta < 0$.

We consider two experimental models. Model A (recursive dynamics) has $\gamma_1 < 0$ and $\gamma_2 = 0$, while Model B (endogenous dynamics) has $\gamma_1 < 0$ and $\gamma_2 = -\gamma_1$. We allow structural changes by replacing α and β by

$$\begin{aligned} \alpha_T(r) &= \alpha_0 + \alpha_1\varphi_{[Tr]} \\ \beta_T(r) &= \beta_0 + \boldsymbol{\beta}_1\varphi_{[Tr]} \end{aligned}$$

where

$$\varphi_{[Tr]} = \begin{cases} 1 & r_1 \leq r \leq r_2 \\ 0 & \text{otherwise.} \end{cases}$$

Thus, the model allows either one break ($0 < r_1 < 1$, $r_2 = 1$) or two breaks. ($0 < r_1 < r_2 < 1$). In all the experiments, $\alpha_0 = 0$ and $\beta_0 = 1$. We also consider the case of partial non-cointegration, where $z_t = (1 - \varphi_{[Tr]})(x_{2t} - \alpha - \beta x_{1t})$.

An important and insufficiently noted issue with tests of cointegration is that that these are always procedures involving a pre-test. They are incorrectly sized unless the series are actually I(1), and a preliminary test of this hypothesis is routine in practice. There is a particular hazard of over-rejection if the normalized variable in the regression used to form a residual-based statistic is in fact stationary. In our experimental model this is well illustrated by the case where $\beta = 0$ in (5.1) and the cointegration test is conducted by regressing x_{1t} onto x_{2t}. Of course, the problem becomes one of potential under-sizing if the roles of regressor and regressand are interchanged in this case.

More generally, a significant proportion of series generated by models with breaks of slope appear 'stationary' in spite of the presence of the unit root, reflecting the nonlinearity of the process. The conventional PP or ADF test may well reject, despite the fact that no (single) cointegrating relationship exists in the sample. Our experiments are constructed to mimic good practice in this regard, by conducting a conventional preliminary test of I(1) on both the candidate series. The variable having the larger I(1) test statistic is always designated the regressand for the cointegration test. These considerations apart, our reported rejection frequencies are true power estimates, since the p-values have been computed from empirical distributions tabulated from 10,000 replications of the null hypothesis $\gamma_1 = \gamma_2 = 0$, using the same setup and sample size, $T = 200$. Note that the critical values in Table 1 are not used.

The relative performance of the tests under break alternatives will of course depend on the positions of the breaks. It is reasonable to assume that in repeated applications, the break points should be treated as if uniformly distributed on $[0, 1]$. In other words, we have no prior belief that one particular break position or duration is of greater practical importance than another. Therefore, rather than pick fixed values for repeated experiments, a neater presentation of the evidence is achieved by drawing r_1 or (r_1, r_2) at random from $U[0, 1]$ in each replication. The rejection frequencies can be viewed as marginal probabilities with respect to this distribution. In the two-break case, two drawings are made independently and the smaller and larger of these assigned to r_1 and r_2 respectively.

Each of the break models has been simulated for four values of γ_1, representing different degrees of persistence of the cointegrating relation according to (5.2). The five break cases

	γ	PP	Q_S^*	$Q_I(0.5)$	$Q_I(0.35)$	$Q_I(0.2)$	Q_R^*	I(1)
Case 1	0.05	0.090	0.065	0.060	0.059	0.059	0.052	0.008
No Break	0.1	0.448	0.264	0.208	0.173	0.140	0.099	0.016
	0.2	0.994	0.956	0.889	0.847	0.777	0.476	0.031
	0.5	1	1	1	1	1	1	0.046
Case 2	0.05	0.047	0.048	0.047	0.048	0.050	0.044	0.002
1 Intercept Break	0.1	0.102	0.115	0.141	0.125	0.114	0.093	0.003
$\alpha_1 = 10$	0.2	0.246	0.433	0.621	0.599	0.527	0.371	0.006
$r_1 \sim U[0,1]$	0.5	0.356	1	1	1	1	0.999	0.009
Case 3	0.05	0.045	0.051	0.051	0.051	0.054	0.041	0.002
2 Intercept Breaks	0.1	0.106	0.092	0.098	0.087	0.083	0.076	0.003
$\alpha_1 = 10$	0.2	0.216	0.298	0.341	0.337	0.305	0.284	0.007
$r_1, r_2 \sim U[0,1]$	0.5	0.308	0.655	0.675	0.835	0.895	0.811	0.008
Case 4	0.05	0.045	0.041	0.040	0.041	0.040	0.038	0.002
1 Slope Break	0.1	0.206	0.130	0.120	0.103	0.084	0.065	0.003
$\beta_1 = 1$	0.2	0.611	0.624	0.652	0.591	0.516	0.340	0.007
$r_1 \sim U[0,1]$	0.5	0.767	1	1	1	1	0.999	0.010
Case 5	0.05	0.029	0.048	0.045	0.046	0.050	0.045	0.002
Cointegration	0.1	0.102	0.087	0.092	0.082	0.077	0.071	0.007
Break	0.2	0.236	0.302	0.324	0.297	0.266	0.211	0.009
$r_1 \sim U[0,1]$	0.5	0.301	0.638	0.625	0.703	0.735	0.618	0.015

Table 2: Powers for Model A

simulated are: 1) No breaks, standard cointegration model; 2) a single break in intercept, with $a_T(r) = 10$ for $r > r_1$, 0 otherwise. 3) A double break in intercept, with $a_T(r) = 10$ for $r_1 < r < r_1$ and 0 otherwise. 4) a single break in slope, with $\beta_T(r) = 2$ for $r > r_1$, 1 otherwise; 5) a break in cointegration, with $z_t = 0$ for $r > r_1$. Tables 2 and 3 compare six different tests for these cases: the standard PP test; the Q_S^* split sample test defined by (3.9); the incremental test $Q_I(\lambda_0)$ defined by (3.10) for the case $\lambda_0 = 0.5$, 35 and 0.2; and the rolling sample test Q_R^* defined by(3.12), again for the case $\lambda_0 = 0.5$. All the latter tests are based on the PP statistic computed using the Parzen kernel and the plug-in bandwidth formula proposed by Newey and West (1993), although without the pre-whitening step suggested by those authors.[4] The last column in each table, headed I(1), shows proportion of replications in which the PP tests of I(1) for both variables reject at the nominal 5% level. Where this number is significant, as it is in a few cases, the power estimates need to be interpreted with appropriate caution.

A noteworthy feature of these experiments is the fact that the ordinary PP test often has good power to reject non-cointegration even when this is subject to breaks. This is particularly true in Case 4, the slope break case. In interpreting these results it is necessary to check the last column to see whether this power is spurious in the sense of the previous paragraph, although it is the fact that it rejects at a comparable rate to the extremum tests that is of interest. However, in the other cases the extremum tests often display a substantial power advantage. Comparing the alternatives, Q_R^* generally performs the poorest but there does not appear to be an overall winner. The best strategy must clearly depend on how many breaks there are, but $Q_I(0.35)$ looks like a good bet overall.

[4]Experimentally there appeared a risk of over-correction, resulting in reduced test power. Note that the tests are all correctly sized by construction, with the same bandwidth settings used in all runs.

	γ	PP	Q_S^*	$Q_I(0.5)$	$Q_I(0.35)$	$Q_I(0.2)$	Q_R^*	I(1)
Case 1	0.05	0.332	0.196	0.187	0.145	0.121	0.093	0.058
No Break	0.1	0.952	0.854	0.750	0.692	0.617	0.351	0.150
	0.2	1	1	0.999	0.999	0.998	0.932	0.273
	0.5	1	1	1	1	1	1	0.371
Case 2	0.05	0.053	0.068	0.083	0.080	0.079	0.070	0.014
1 Intercept Break	0.1	0.109	0.198	0.324	0.295	0.251	0.173	0.033
$\alpha_1 = 10$	0.2	0.184	0.673	0.852	0.853	0.820	0.668	0.024
$r_1 \sim U[0,1]$	0.5	0.366	1	1	1	1	1	0.038
Case 3	0.05	0.058	0.069	0.067	0.066	0.065	0.064	0.012
2 Intercept Breaks	0.1	0.148	0.222	0.251	0.247	0.227	0.201	0.029
$\alpha_1 = 10$	0.2	0.197	0.524	0.554	0.649	0.643	0.630	0.060
$r_1, r_2 \sim U[0,1]$	0.5	0.307	0.650	0.682	0.861	0.942	0.808	0.135
Case 4	0.05	0.287	0.180	0.165	0.145	0.122	0.088	0.052
1 Slope Break	0.1	0.762	0.676	0.635	0.589	0.525	0.324	0.119
$\beta_1 = 1$	0.2	0.946	0.978	0.973	0.969	0.953	0.825	0.185
$r_1 \sim U[0,1]$	0.5	0.996	1	1	1	1	1	0.371
Case 5	0.05	0.073	0.078	0.072	0.068	0.070	0.065	0.014
Cointegration	0.1	0.189	0.262	0.291	0.271	0.235	0.183	0.021
Break	0.2	0.259	0.568	0.566	0.613	0.609	0.543	0.034
$r_1 \sim U[0,1]$	0.5	0.311	0.662	0.652	0.750	0.840	0.648	0.061

Table 3: Powers for Model B

6 An Empirical Example

Campbell and Shiller (1987) show that the efficient markets hypothesis (EMH) implies stock prices and dividends should be cointegrated. It is nonetheless not uncommon to observe episodes of rapid price increases unconnected with dividend growth, followed by subsequent crashes. The 'dot-com' boom of the late 1990s is perhaps the best known, as well as the most extreme, of recent examples. It is well known that non-fundamental stock price increases and crashes can be integrated into present value models by removing the transversality condition guaranteeing a unique solution, and such outcomes can explain the presence of stochastic rational bubbles on the stock market. Shiller's US annual data on real stock prices and dividends, 1871-2004, are shown in Figure 1 [5]. The necessity of triming off some observations before the predictions of the EMH appear supported by these data may be no surprise. However, the act of trimming the sample in the light of test outcomes inevitably contaminates the inferences. Our procedures avoid the data-snooping problem by providing critical values applying to any of the specified subsamples.

Table 4 shows both the traditional residual-based tests from the regression of prices on dividends, and their counterparts based on the extremum statistics proposed here. The table shows the values of the statistics and, in parentheses, the upper bounds on p-values that can be determined from the tabulated critical values. Note that both the ADF and PP tests fail to reject the null hypothesis of no cointegration at conventional significance levels, whereas all the tests based on subsamples reject at the 0.025 level or better. This is an example of the typical situation where the economic theory predicts a cointegrating relation and the traditional tests do not confirm the theory. Our tabulations can then provide a check on the validity of censoring the sample; a rejection indicates to the researcher the existence of some relationship worthy of

[5] The data are provided by Robert Shiller on his webpage at: *http://www.econ.yale.edu/~shiller/*

Type	λ_0	ADF Test	Phillips-Perron
Standard	–	-2.81 (< 1)	-2.52 (< 1)
Q_S	–	-4.46 (< 0.01)	-4.41 (< 0.01)
Q_S^*	–	-4.46 (< 0.01)	-4.41 (< 0.01)
$Q_I(\lambda_0)$	0.5	-4.90 (< 0.01)	-4.89 (< 0.01)
	0.35	-4.93 (< 0.025)	-4.90 (< 0.025)
	0.2	-4.93 (< 0.025)	-4.90 (< 0.025)
$Q_R(\lambda_0)$	0.5	-4.72 (< 0.025)	-4.73 (< 0.025)
$Q_R^*(\lambda_0)$	0.5	-4.72 (< 0.025)	-4.73 (< 0.025)

Table 4: Real Prices and Dividends 1871-2004: p-value bounds in parentheses

further investigation.

The properties of the test procedures are illustrated in a different way in Figure 2, which shows the values of the ADF statistic for each subsample between 1871-1935 (around half the sample) and 1871-2004. The two broken lines mark the 5% critical values of, respectively, the usual Dickey-Fuller distribution (1 regressor + intercept) and the $Q_I(0.5)$ statistic from the tabulation in Table 1. One thing apparent from this plot, in addition to the gross shift in 1996 and subsequently, is the evidence for a smaller but apparently permanent break in the relation around 1959. These changes are also apparent in the plot of the incremental slope and intercept coefficients in the cointegrating regressions, shown in Figure 3, together with the two-standard error bands. (Note that these series are not shown to scale, but are normalized as a proportion of their terminal values for easy visual comparison.) We conclude that there is strong evidence from these data of the existence of cointegration subject to breaks. While further investigations are needed to determine the character of these shifts with confidence, given that the slope and intercept are highly correlated, these are plausibly breaks in the slope. Without offering any theories or speculations, we will simply draw attention to the coincidence between the fact that the first break, at least, does not appear to affect the power of the regular ADF test very severely, and the evidence on test performance in Case 4 of Tables 2 and 3.

7 Concluding Remarks

Although a wider range of examples and experiments might be required to put the efficacy of these subsample tests into a clearer perspective, our results suggest that the methods can often succeed in revealing the existence of relations subject to breaks. We should emphasize again that, given the extremum tabulations, there is no need to actually compute the full incremental or rolling procedures in order to perform valid tests. If these critical values are exceeded by any member of a sequence of subsample tests drawn from the corresponding set from (3.6) or (3.7), we are entitled to reject the null hypothesis at the corresponding significance level. However, we also point out that a rejection of the null hypothesis on our test can only be the first step in the investigation of the relationships. It does not, in particular, provide a consistent estimator of the break point(s), although a plot of the statistic values as functions of λ_1 and/or λ_2 would doubtless provide a very useful informal guide.

We note in conclusion that the tests might in principle be implemented with a range of alternative q_T statistics, such as the modified Dickey-Fuller of Elliott, Rothenberg and Stock (1996) or the Bartlett-corrected ADF proposed by Johansen (2004). The method might likewise be adapted to to the fractional ADF as in Dolado, Gonzalo and Mayoral (2002), or Lobato and Velasco (2006a), although a strategy for supplementary parameter estimation would need

to be considered for these cases. It might even be combined (at some computational cost) with bootstrap procedures such as Davidson (2002, 2006). Extending the present theory, based on the asymptotic properties of the statistics in question, appears a reasonably strsaightforward exercise for future work, but it appears to us plausible that ar least their *relative* performances with respect to different alternative hypotheses would be similar to the cases examined here.

A Appendix: Proofs

Proof of Theorem 4.1.

The proof adapts the methods developed by Zivot and Andrews (1992) and Gregory and Hansen (1996). The first step is to note that according to (4.3), (4.2) and (4.7), the weak limit of the subsample test statistic can be written as

$$\tau(\lambda_1, \lambda_2) = g(m_1(\lambda_1, \lambda_2), m_2(\lambda_1, \lambda_2))$$

where $g(\cdot, \cdot) : \mathbb{R}^{p^2} \times \mathbb{R}^{p^2} \mapsto \mathbb{R}$ is continuous in its arguments, and

$$m_1(\lambda_1, \lambda_2, \boldsymbol{J}) = \text{Vec} \int_{\lambda_1}^{\lambda_2} \boldsymbol{W}^* d\boldsymbol{W}$$
$$m_2(\lambda_1, \lambda_2, \boldsymbol{B}) = \text{Vec} \int_{\lambda_1}^{\lambda_2} \boldsymbol{W}^* \boldsymbol{W}^{*\prime} dr.$$

In turn, $m_1 : [0,1]^2 \times C[0,1]^{p^2} \mapsto \mathbb{R}^{p^2}$ and $m_2 : [0,1]^2 \times C[0,1]^p \mapsto \mathbb{R}^{p^2}$ are continuous functionals of their arguments, where continuity is defined with respect to the uniform metric. Specifically it follows from (4.2)–(4.6) that small shifts in λ_1 and λ_2, with given $\boldsymbol{J}$, and also small uniform changes in $\boldsymbol{J}$ with λ_1 and λ_2 fixed, both lead to correspondingly small changes in m_1; and similarly for m_2 with respect to λ_1, λ_2 and $\boldsymbol{B}$.

We next have to consider the various extremum statistics defined by the sets (3.6) and (3.7). Consider two pairs of points $(\lambda_1^a, \lambda_2^a)$ and $(\lambda_1^b, \lambda_2^b)$, and denote the corresponding cases by τ_a and τ_b. Since

$$\min\{x, y\} = \tfrac{1}{2}(y - x - |y - x|) + x$$

is a continuous function of its arguments, we see that

$$\min\{\tau(0, \tfrac{1}{2}), \tau(\tfrac{1}{2}, 1)\} = \min_{(\lambda_1, \lambda_2) \in \Lambda_s} \{\tau(\lambda_1, \lambda_2)\}$$

is a continuous functional of $(\boldsymbol{B}, \text{Vec}\, \boldsymbol{J})$, and the same result extends by iteration to

$$\min_{(\lambda_1, \lambda_2) \in \Lambda_s^*} \{\tau(\lambda_1, \lambda_2)\}.$$

Next, consider the cases Λ_{0f} and Λ_{0b}. Here, one of the two real arguments is held fixed, λ_1 in the first case and λ_2 in the second case. The argument of Zivot and Andrews (1992) Lemma A4 can be invoked here, which says that if (say) τ and $\tilde{\tau}$ represent the function evaluated two different points of $C[0,1]^{p(1+p)}$, then

$$\left| \inf_{\Lambda_{0f}} \tau(0, \lambda_2) - \inf_{\Lambda_{0f}} \tilde{\tau}(0, \lambda_2) \right| \leq \sup_{\Lambda_{0f}} |\tau(0, \lambda_2) - \tilde{\tau}(0, \lambda_2)| \tag{A-1}$$

That is, the difference between the infima of two functions that are uniformly close is correspondingly small. It follows that we may treat $\inf_{\Lambda_{0f}} \tau(0, \lambda_2)$ as a continuous functional of $(\boldsymbol{B}, \text{Vec}\, \boldsymbol{J})$,

with the same result holding for $\inf_{\Lambda_{0b}} \tau(\lambda_1, 1)$, and hence extending to $\inf_{\lambda \in \Lambda_{0f} \cup \Lambda_{0b}} \tau(\lambda_1, \lambda_2)$ by the preceding argument. The cases Λ_{0r} and Λ^*_{0r}, follow in the same manner.

We have therefore established that in the limit the alternative extremum statistics are continuous functionals of $(\boldsymbol{B}, \operatorname{Vec} \boldsymbol{J})$. Applying Assumptions 1 and 2 and the continuous mapping theorem completes the proof. ∎

Proof of Theorem 4.2.

By definition of the sets A_T, if $t \in T_1 A_{T_1}$ then $t \in TA_T$ for all $T > T_1$. In other words, the sequence of intervals $\{TA_T\}$ is monotone non-decreasing. It follows that if $[\lambda_1, \lambda_2] \subseteq A_{T_1}$ then $[T\lambda_1, T\lambda_2] \subseteq TA_T$.for $T > T_1$. Defining $t^* = t - T\lambda_1 + 1$ and $T^* = [T(\lambda_2 - \lambda_1)]$, it follows that the samples $\{\boldsymbol{x}_{t^*} : 1 \leq t^* \leq T^*\}$ are cointegrated for each T^*, and increasing in size at the rate $O(T)$. It follows that if q^*_T denotes a consistent cointegration statistic computed from these samples, $|q^*_T| = O_p(T^{1/2})$ by the arguments of Phillips and Ouliaris (1990) inter alia, and the test rejects with probability tending to 1. However,

$$q^*_T = q_T(\lambda_1, \lambda_2) \geq \inf_{\{\lambda_1, \lambda_2\} \in \Lambda} q_T(\lambda_1, \lambda_2) = Q_\Lambda.$$

It follows that the test based on Q_Λ is consistent. ∎

References

Banerjee, A. , R. L. Lumsdaine and J. H. Stock (1992) Recursive and sequential tests of the unit root and trend-break hypotheses: theory and international evidence. *Journal of Business and Economic Statistics* 10(3) 271-287.

Billingsley, P. (1968) *Convergence of Probability Measures* Wiley

Breitung, J. (2002), Nonparametric tests for unit roots and cointegration, *Journal of Econometrics* 108, 343–363.

Campbell, J., Shiller, R (1987) Cointegration and Tests of Present Value Models, *Journal of Political Economy*, 95, 1062–1088.

Chan, N.H. and C.Z. Wei (1988) Limiting distributions of least-squares estimates of unstable autoregressive processes. *Annals of Statistics* 16, 367-401.

Davidson, J. (1998) Structural relations, cointegration and identification: some simple results and their application. *Journal of Econometrics* 87, 87–113.

Davidson, J., 2000. *Econometric Theory*, Oxford: Blackwell Publishers.

Davidson, J., 2002. A Model of Fractional Cointegration, and Tests for Cointegration Using the bootstrap, *Journal of Econometrics*, 110, 105-112.

Davidson, J., 2006. Alternative Bootstrap Procedures for Testing Cointegration in Fractionally Integrated Processes, *Journal of Econometrics*, 133, 741-777.

De Jong, R. M. and J. Davidson (2000) The functional central limit theorem and convergence to stochastic integrals I: the weakly dependent case. *Econometric Theory* 16, 5, 621–642

Dolado, J., Gonzalo, J., Mayoral, L. (2002) A fractional Dickey-Fuller test for unit roots, *Econometrica*, 5, 1963-2006.

Dolado, J., Gonzalo, J., Mayoral, L. (2005) What is What?: A Simple Time-Domain Test of Long-memory vs. Structural Breaks, Working Paper at http://dolado.blogspot.com/

Elliott, G., T. J. Rothenberg and J. H. Stock (1996) Efficient tests for an autoregressive unit root, *Econometrica* 64, 813-836

Engle, R. F., Granger, C. W. J. (1987) Cointegration and error correction: representation, estimation, and testing, *Econometrica*, 55, 251-276.

Engle, R. F., Yoo, S. (1987) Forecasting and testing in co-integrated systems, *Journal of Econometrics*, 35, 143-159.

Gregory, A. W., Hansen, B. E. (1996) Residual-based tests for cointegration in models with regime shifts, *Journal of Econometrics*, 70, 99-126.

Hansen, B. E. (1992) Convergence to stochastic integrals for dependent heterogeneous processes, *Econometric Theory* 8, 489-500.

Kim, J.-Y. (2003) Inference on segmented cointegration, *Econometric Theory* 19, 620-639

Kurtz, T. G. (2001) *Lectures in Stochastic Analysis*, Department of Mathematics and Statistics, University of Wisconsin, at http://www.math.wisc.edu/~kurtz/735/main735.pdf

Kurtz, T. G. and P. E. Protter (1991) Weak limit theorems for stochastic integrals and stochastic differential equations, *Annals of Probability* 19, 1035-1070.

Kurtz, T.G. and and P. E. Protter (1995) Weak convergence of stochastic integrals and differential equations, in *Probabilistic Models for Nonlinear Partial Differential Equations: Lectures Given at the 1st Session of the Centro Internazionale Matematico Estivo* (C.I.M.E.), by K. Graham,, T.G. Kurtz, S. Meleard P.E. Protter, M. Pulverenti, D. Talay. Lecture Notes in Mathematics 1627, Springer Verlag.

Johansen, S. (2004) A small sample correction of the Dickey-Fuller test. In A. Welfe (ed): *New Directions in Macromodelling,* Chapter 3, 49-68, Elsevier.

Lobato, I. N. and C. Velasco (2006a) Efficient t-tests for fractional unit roots, Working Paper

Lobato, I. N. and C. Velasco (2006b) Efficient Wald tests for fractional unit roots, Working Paper

Newey, W. K., and K. D. West (1987a). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55, 703-8.

Phillips, P. C. B., Ouliaris, S. (1990) Asymptotic properties of residual based tests for cointegration, *Econometrica*, 1, 165-193.

Phillips, P.C. B. and Perron, P. (1988). Testing for a unit root in time series regression, *Biometrika*, 75, 335-346.

Qu, Z. (2007) Searching for cointegration in a dynamic system. Working paper, University of Illinois at Urbana-Champaign.

Zivot, E amd D. W. K. Andrews (1992) Further evodence on the great crash, the oil price shock and the unit root hypothesis. *Journal of Business and Economic Statistics* 10 (3) 251-270.

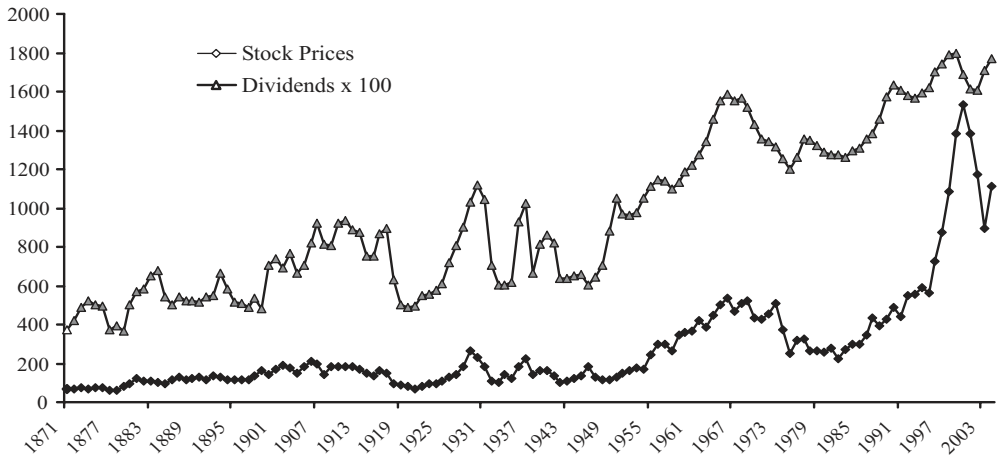


Figure 1. The Data Set

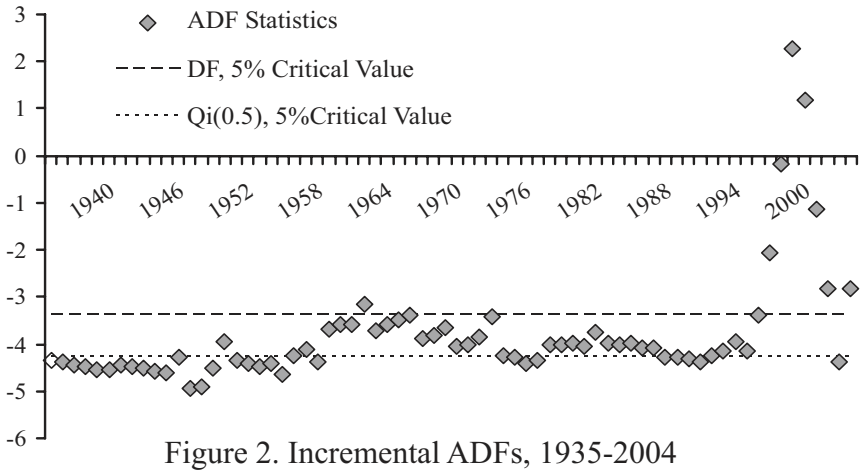


Figure 2. Incremental ADFs, 1935-2004

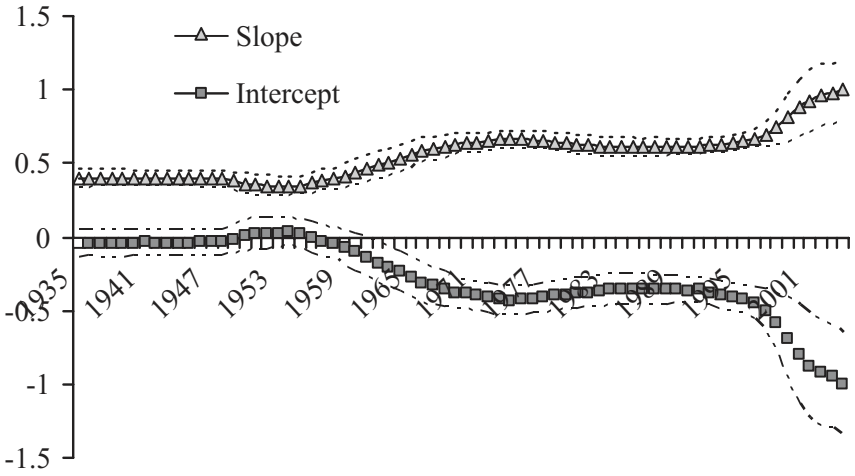


Figure 3. Incremental Regression Coefficients and 2SE Bands 1935-2004 (rescaled).