Rochester Center for Economic Research

Residual-Based Tests for Cointegration in Models with Regime Shifts

Gregory, Allan W. and Bruce E. Hansen

Working Paper No. 335
November 1992

University of Rochester

Residual-Based Tests for Cointegration in Models with Regime Shifts

Allan W. Gregory
and
Bruce E. Hansen

Rochester Center for Economic Research
Working Paper No. 335

Residual-Based Tests for Cointegration in Models with Regime Shifts

Allan W. Gregory and Bruce E. Hansen*

November 1992

Abstract

In this paper we examine tests for cointegration which allow for the possibility of regime shifts. We propose $ADF-$, $Z_\alpha-$, and Z_t- type tests designed to test the null of no cointegration against the alternative of cointegration in the presence of a possible regime shift. In particular we consider cases where the intercept and/or slope coefficients have a single break of unknown timing. A formal proof is provided for the limiting distributions of the various tests for the regime shift model (both a level and slope change). Critical values are calculated for the tests by simulation methods and a simple Monte Carlo experiment is conducted to evaluate finite sample performance. In the limited set of experiments, we find that the tests can detect cointegrating relations when there is a break in the intercept and/or slope coefficient. For these same experiments, the power of the conventional ADF test with no allowance for regime shifts falls sharply. As an illustration we test for structural breaks in the U.S. long-run money-demand equation using annual and quarterly data.

Key Words: level shift, regime shift, cointegration, Brownian motion

JEL Classification Numbers: C12, C15, C22, C52

* Gregory, Queen's University and Hansen, University of Rochester. We thank Gregor Smith for helpful comments and Kim Cook for excellent research assistance. The first author acknowledges financial support from the Social Sciences and Humanities Research Council of Canada. The second author thanks the National Science Foundation for financial support.

1. Introduction

It is now routine for researchers to test for cointegration when working with multivariate time series. The most widely applied tests are residual-based ones in which the null hypothesis of no cointegration is tested against the alternative that the relation is cointegrated in the sense of Engle and Granger (1987), meaning that a linear combination of the integrated variables has a stationary distribution. A large sample distribution theory for this class of tests has been studied by Phillips and Ouliaris (1990). Rejection of the null hypothesis in this context implies the strong result that the variables are cointegrated. Acceptance of the null hypothesis is often taken as evidence of the lack of cointegration.

While these tests and the associated distributional theory are appropriate for the precise question of no cointegration versus cointegration, there are many related questions which may appear quite similar, but actually require a different set of tests and distributional theory. In this paper we are concerned with the possibility of a more general type of cointegration, where the cointegrating vector is allowed to change at a single unknown time during the sample period. While our null hypothesis (no cointegration) is the same, our alternative hypothesis is different than the conventional tests. Indeed, we extend the class of models under consideration, since our alternative hypothesis contains the Engle-Granger model as a special subcase.

The motivation for the class of tests considered here derives from the conventional notion of regime change. In some empirical exercises, a researcher may wish to entertain the possibility that the series are cointegrated, in the sense that a linear combination of the non-stationary variables is stationary, but that this linear combination (the cointegrating vector) has shifted at one (unknown) point in the sample. In this context, the standard tests for cointegration are not appropriate, since they presume that the cointegrating vector is time invariant under the alternative hypothesis. A new class of residual-based tests for cointegration are needed which include in the alternative hypothesis the models considered here.

Specifically, we propose extensions of the ADF, Z_α, and Z_t tests for cointegration. Our tests allow for a regime shift in either the intercept alone or the entire coefficient vector, and are non-informative with respect to the timing of the regime shift. This prevents informal data analysis (such as the visual examination of time series plots) from contaminating the choice of breakpoint. The tests of this paper can be viewed as multivariate extensions of the univariate tests of Perron (1989), Zivot and Andrews (1992) and Banerjee, Lumsdaine, and Stock (1992). These papers tested the null of a unit root in a univariate time series against the alternative of stationarity, while allowing for a structural break in the deterministic component of the series. In fact, the results of these papers can be viewed as a special case of our results, when the number of stochastic regressors is taken to be zero.

The asymptotic distributions of the test statistics are derived. We find that the asymptotic distributions of the proposed test statistics are free of nuisance parameter dependencies, other than the number of stochastic and deterministic regressors. The distributional theory is more involved than the theory for the conventional cointegration model (see Phillips and Ouliaris, 1990) due to the inclusion of dummy variables and the explicit minimization over the set of possible breakpoints. It should be emphasized that our results are more general than the other results which have appeared in the literature on breaking trends and unit roots. Zivot and Andrews (1992) provided a fully rigorous proof for the simple Dickey-Fuller statistic in the univariate case, under the assumption of iid innovations. In contrast, we examine the more cumbersome Phillips Z_α and Z_t tests in the multivariate case, while allowing for general forms of serial correlation in the innovations through the use of mixing conditions.

Since there are no closed-formed solutions for the limiting distributions, critical values for up to four regressors are calculated for the tests by simulation methods. Also, since the computational requirements from recursive calculations are extremely high, preventing simulations with large sample sizes, we follow MacKinnon (1991) and estimate response surfaces to approximate the appropriate critical values. We evaluate the finite sample performance of the tests using Monte Carlo methods based upon the experimental design of Engle and Granger (1987). In a limited set of experiments, we find the tests can detect cointegrating relations when there is a break in the intercept and/or slope coefficient. For these same experiments, the power of the conventional ADF test with no allowance for regime shifts falls sharply.

The tests of the present paper are clearly useful in helping lead an applied researcher to a correct model specification. Many researchers start a cointegration analysis with the usual augmented Dickey-Fuller (ADF) test, and proceed only if the statistic rejects the null of no cointegration. If the model is indeed cointegrated with a one-time regime shift in the cointegrating vector, the standard ADF test may not reject the null and the researcher will falsely conclude that there is no long-run relationship. Indeed, Gregory and Nason (1992) have shown that the power of the conventional ADF test falls sharply in the presence of a structural break. In contrast, if the tests of the present paper are employed, there is a better chance of rejecting the null hypothesis, leading to a correct model formulation.

The tests of this paper are complementary to those of Hansen (1992a). In that paper, Hansen developed a series of tests of the hypothesis of time invariance of the coefficients of a cointegrating relation. His null hypothesis is Engle-Granger cointegration, while our null hypothesis is no cointegration. Hansen's tests are best viewed as specification tests for the Engle-Granger cointegration model. In contrast, the tests of this paper are tests for cointegration, and are therefore best viewed as pre-tests akin to the conventional residual-based cointegration tests.

As an illustration of the techniques, we test for structural breaks in the U.S. long-run money-demand equation using annual and quarterly data. Our results are consistent with those of Gregory and Nason (1992), who found evidence of instability in the long-run relationship.

The organization of the paper is as follows. In Section 2 we develop several single-equation regression models which allow for cointegration with structural change. In Section 3 we describe various tests for the null of no cointegration with power against the structural change alternatives outlined in Section 2. Section 4 contains the asymptotic distribution theory for the tests. Critical values are calculated using simulation methods. Section 5 assess the finite sample properties of the structural change tests in a simple Monte Carlo experiment and Section 6 examines an illustrative empirical example based upon money demand. Finally in Section 7 we close with some concluding remarks and suggest several directions for future research.

2. Model

In this section we develop single-equation regression models which allow for cointegration with structural change. The observed data is $y_t = (y_{1t}, y_{2t})$, where y_{1t} is real-valued and y_{2t} is an m-vector. We commence with the standard model of cointegration with no structural change.

Model 1: Standard Cointegration

$$y_{1t} = \mu + \alpha^{\top} y_{2t} + e_t, \; t = 1, ..., n, \tag{2.1}$$

where y_{2t} is $I(1)$ and e_t is $I(0)$. In this model the parameters μ and α describe the m-dimensional hyperplane towards which the vector process y_t tends over time. Engle and Granger (1987) describe cointegration as a useful model for "long-run equilibrium".

In many cases, if model 1 is to capture a long-run relationship, we will want to consider μ and α as time invariant. But in other applications, it may be desirable to think of cointegration as holding over some (fairly long) period of time, and then shifting to a new "long-run" relationship. We treat the timing of this shift as unknown . The structural change would be reflected in changes in the intercept μ and/or changes to the slope α.

To model structural change, it is useful to define the dummy variable:

$$\varphi_t = \begin{cases} 0, & \text{if } t \leq [n\,\tau] \\ 1, & \text{if } t > [n\,\tau] \end{cases}$$

where the unknown parameter $\tau \in (0, 1)$ denotes the (relative) timing of the change point, and [ ] denotes integer part.

Structural change can take several forms. A simple case is that there is a level shift in the cointegrating relationship, which can be modeled as a change in the intercept μ, while the slope coefficients α are held constant. This implies that the equilibrium equation has shifted in a parallel fashion. We call this a *level shift* model denoted by C.

Model 2: Level Shift (C)

$$y_{1t} = \mu_1 + \mu_2 \varphi_{t\tau} + \alpha^{\top} y_{2t} + e_t, \; t = 1, ..., n. \tag{2.2}$$

In this parameterization μ_1 represents the intercept before the shift, and μ_2 represents the change in the intercept at the time of the shift. We can also introduce a time trend into the level shift model.

Model 3: Level Shift with Trend (C/T)

$$y_{1t} = \mu_1 + \mu_2 \varphi_{t\tau} + \beta t + \alpha^\top y_{2t} + e_t, \ t = 1, ..., n. \tag{2.2}$$

Another possible structural change allows the slope vector to shift as well. This permits the equilibrium relation to rotate as well as shift parallel. We call this the *regime shift* model.

Model 4: Regime Shift (C/S)

$$y_{1t} = \mu_1 + \mu_2 \varphi_{t\tau} + \alpha_1^\top y_{2t} + \alpha_2^\top y_{2t} \varphi_{t\tau} + e_t, \ t = 1, ..., n. \tag{2.3}$$

In this case μ_1 and μ_2 are as in the level shift model, α_1 denotes the cointegrating slope coefficients before the regime shift and α_2 denotes the change in the slope coefficients.

There are certainly other candidate models which might be used to analyze structural shifts in cointegrated relationships. For instance we could have a shift in the trend function as well as a regime shift.

Model 5: Regime Shift with a Shift in Trend (C/S/T)

$$y_{1t} = \mu_1 + \mu_2 \varphi_{t\tau} + \beta_1 t + \beta_2 t \varphi_{t\tau} + \alpha_1^\top y_{2t} + \alpha_2^\top y_{2t} \varphi_{t\tau} + e_t, \ t = 1, ..., n. \tag{2.5}$$

Here β_1 represents the slope of the trend before the structural break and β_2 represents the change.

The standard methods to test the null hypothesis of no cointegration (derived in the context of model 1) are residual-based. The candidate cointegrating relation is estimated by ordinary least squares (OLS), and a unit root test is applied to the regression errors. In principle the same approach could be used for testing models 2-5, if the timing of the regime shift τ were known *a priori*. We take the view that such break points are unlikely to be known in practice without some appeal to the data. Indeed much of the debate about whether there was a regime shift in U.S. GNP around 1929 (as identified in Perron, 1989) can only be resolved conditional on the data (see Banerjee, Lumsdaine, and Stock, 1992; Cristiano, 1992; and Zivot and Andrews, 1992). Similar problems occur in testing for regime shifts in cointegrated models and so we develop tests procedures that do not require information regarding the timing of or indeed the occurrence of a break.

This completes the description of the structural change models under cointegration. In the next section we analyze some tests designed to detect cointegration in the possible presence of such breaks.

3. Testing the Null of No Cointegration

Hansen (1992a) constructed tests of model 1 against the alternative of model 4. A statistically significant test statistic in this context would be taken as evidence against the standard cointegration model in favor of the regime shift model. However before calculating such a test, an applied econometrician might wish to apply a conventional test for cointegration, such as the ADF test, in the context of model 1. If the true process is represented by model 4, and not model 1, the distributional theory used to assess the significance of the ADF test statistic is not same. In this section we develop such tests of the null of cointegration against the alternatives in models 2-4.

Define the innovation vector

$$u_t = \Delta y_t,$$

its cumulative process

$$S_t = \sum_{i=1}^{t} u_i,$$

(so $y_t = y_0 + S_t$), and its long-run variance

$$\Omega = \lim_{n\to\infty} \frac{1}{n} E S_n S_n^{\top}.$$

When u_t is covariance stationary, Ω is proportional to the spectral density matrix evaluated at the zero frequency.

Our null hypothesis is that model 1 holds, with $e_t \equiv I(1)$. This has the implication that $\Omega > 0$. We include this aspect of the null hypothesis in the following regularity conditions:

Assumptions:

(a) $\{u_t\}$ is mean-zero and strong mixing with mixing coefficients of size $\frac{-p\beta}{(p-\beta)}$, and $E|u_t|^p < \infty$ for some $p > \beta > \frac{5}{2}$.

(b) The matrix Ω exists with finite elements and $\Omega > 0$.

(c) y_0 is a random vector with $E|y_0| < \infty$.

The solution we adopt to handling regime shifts is similar to that of Banerjee, Lumsdaine, and Stock (1992) and Zivot and Andrews (1992). We compute the cointegration test statistic for each possible regime shift $\tau \in T$, and take the smallest value (the largest negative value) across all possible break points. In principle the set T can be any compact subset of $(0, 1)$. In practice, it will need to be small enough so that all of the statistics discussed here can be calculated. For example $T = (.15, .85)$ seems a reasonable suggestion, following the earlier literature. Although T contains an uncountable number of points, all the statistics that we consider are step functions on T, taking jumps only on the points $\{\frac{i}{n}, \, i \; integer\}$. For computational purposes, the test statistic is computed for each break point in the interval $([.15n], [.85n])$.

We now describe the computation of the test statistics. For each τ, estimate one of the models 2-4 (depending upon the alternative hypothesis under consideration) by OLS, yielding the residual $\hat{e}_{t\tau}$. The subscript τ on the residuals denotes the fact that the residual sequence depends on the choice of change point τ. From these residuals, calculate the first-order serial correlation coefficient

$$\hat{\rho}_\tau = \frac{\sum_{t=1}^{n-1} \hat{e}_{t\tau}\hat{e}_{t+1\tau}}{\sum_{t=1}^{n-1} \hat{e}_{t\tau}^2}.$$

The Phillips (1987) test statistics are formed using a bias-corrected version of the first-order serial correlation coefficient. Define the second- stage residuals

$$\hat{\nu}_{t\tau} = \hat{e}_{t\tau} - \hat{\rho}_\tau \hat{e}_{t-1\tau}.$$

The correction involves the following estimate of a weighted sum of autocovariances

$$\hat{\lambda}_\tau = \sum_{j=1}^{M} w(\frac{j}{M})\, \hat{\gamma}_\tau(j),$$

where $M = M(n)$ is the bandwidth number selected so that $M \to \infty$ and $\frac{M}{n^5} = O(1)$, the kernel weights $w(\cdot)$ satisfy the standard conditions for spectral density estimators, and

$$\hat{\gamma}_\tau(j) = \frac{1}{n} \sum_{t=j+1}^{n} \hat{\nu}_{t-j\tau} \hat{\nu}_{t\tau}.$$

The estimate of the long-run variance of $\hat{\nu}_t$ is

$$\hat{\sigma}_\tau^2 = \hat{\gamma}_\tau(0) + 2\hat{\lambda}_\tau.$$

In this paper the long-run variance is estimated using a prewhitened quadratic spectral kernel with a first-order autoregression for the prewhitening and an automatic bandwidth estimator (see Andrews, 1991 and Andrews and Monahan, 1992 for details).

The bias-corrected first-order serial correlation coefficient estimate is given by

$$\hat{\rho}_\tau^* = \frac{\sum_{t=1}^{n-1} (\hat{e}_{t\tau} \hat{e}_{t+1\tau} - \hat{\lambda}_\tau)}{\sum_{t=1}^{n-1} \hat{e}_{t\tau}^2}.$$

The Phillips test statistics can be written as

$$Z_\alpha(\tau) = n(\hat{\rho}_\tau^* - 1)$$

$$Z_t(\tau) = \frac{(\hat{\rho}_\tau^* - 1)}{\hat{s}_\tau}, \qquad \hat{s}_\tau^2 = \frac{\hat{\sigma}_\tau^2}{\sum_1^{n-1} \hat{e}_{t\tau}^2}.$$

The final statistic we discuss is the augmented Dickey-Fuller (ADF) statistic. This is calculated by regressing $\Delta\hat{e}_{t\tau}$ upon $\hat{e}_{t-1\tau}$ and $\Delta\hat{e}_{t-1\tau}, ..., \Delta\hat{e}_{t-K\tau}$ for some suitably chosen lag truncation K. The ADF statistic is the $t-$statistic for the regressor $\hat{e}_{t-1\tau}$. We denote this by

$$ADF(\tau) = tstat(\hat{e}_{t-1\tau}).$$

These test statistics are now standard tools for the analysis of cointegrating regressions without regime shifts. Our statistics of interest, however, are the *smallest* values of the above statistics, across all values of $\tau \in T$. We examine the smallest values since small values of the test statistics constitute evidence against the null hypothesis. These test statistics are

$$Z_\alpha^* = \inf_{\tau \in T} Z_\alpha(\tau) \tag{3.1}$$

$$Z_t^* = \inf_{\tau \in T} Z_t(\tau) \tag{3.2}$$

$$ADF^* = \inf_{\tau \in T} ADF(\tau). \tag{3.3}$$

4. Asymptotic Distributions

We follow much of the recent literature and give asymptotic distributions for the test statistics which are expressed as functionals of Brownian motions. This gives simple expressions for the limit distributions. Since they are not given in closed-form, however, we use simulation methods to obtain critical values.

Since interest from an applied perspective is likely to be concentrated on models 2 and 4, we will only provide the limiting distributions for those two. Our formal proofs are also limited to the Z_α^* and Z_t^* tests. We expect that the limiting distribution of ADF^* is identical to Z_t^*. Models 3 can be shown to have analogous expressions. A formal proof of the limiting distribution for the distribution of the tests of model 4 is contained in the appendix.

Theorem 1

If the test statistics are constructed using the residuals from model 2, then under the null hypothesis

$$Z_\alpha^* \to_d \inf_{\tau \in T} \frac{\int_0^1 W_\tau^* dW_\tau^*}{\int_0^1 W_\tau^{*2}},$$

$$Z_t^* \to_d \inf_{\tau \in T} \frac{\int_0^1 W_\tau^* dW_\tau^*}{[\int_0^1 W_\tau^{*2}]^{1/2} [\kappa_\tau^{*\top} D \kappa_\tau^*]^{1/2}},$$

where

$$W_\tau(r) = W_1(r) - \int_0^1 W_1 W_{2\tau}^{*\top} [\int_0^1 W_{2\tau}^* W_{2\tau}^{*\top}]^{-1} W_{2\tau}^*(r),$$

$$W_{2\tau}^*(r) = [W_2^\top(r),\ 1,\ \varphi_\tau(r)]^\top,$$

$$\varphi_\tau(r) = \{r \geq \tau\},$$

$$W(r) = \begin{pmatrix} W_1(r) \\ W_2(r) \end{pmatrix} \begin{array}{c} 1 \\ m \end{array} = BM(I_{m+1}),$$

$$\kappa_\tau^* = \begin{pmatrix} 1 \\ -[\int_0^1 W_{2\tau}^* W_{2\tau}^{*\top}]^{-1} \int_0^1 W_{2\tau}^* W_1 \end{pmatrix},$$

$$D = \begin{array}{c} \\ m+1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cc} m+1 & 2 \end{array} \\ \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \end{array}.$$

Theorem 2

If the test statistics are constructed using the residuals from model 4, then under the null hypothesis

$$Z_\alpha^* \rightarrow_d \inf_{\tau \in T} \frac{\int_0^1 W_\tau dW_\tau}{\int_0^1 W_\tau^2},$$

$$Z_t^* \rightarrow_d \inf_{\tau \in T} \frac{\int_0^1 W_\tau dW_\tau}{[\int_0^1 W_\tau^2]^{1/2} [\kappa_\tau^\top D_\tau \kappa_\tau]^{1/2}},$$

where

$$W_\tau(r) = W_1(r) - \int_0^1 W_1 W_{2\tau}^\top [\int_0^1 W_{2\tau} W_{2\tau}^\top]^{-1} W_{2\tau}(r),$$

$$W_{2\tau}(r) = [W_2^\top(r),\ 1,\ W_2^\top(r)\varphi_\tau(r),\ \varphi_\tau(r)]^\top,$$

$$\varphi_\tau(r) = \{r \geq \tau\},$$

$$W(r) = \begin{pmatrix} W_1(r) \\ W_2(r) \end{pmatrix} \begin{array}{c} 1 \\ m \end{array} = BM(I_{m+1}),$$

$$\kappa_\tau = \begin{pmatrix} 1 \\ -[\int_0^1 W_{2\tau} W_{2\tau}^\top]^{-1} \int_0^1 W_{2\tau} W_1 \end{pmatrix},$$

$$D_\tau = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & I_m & 0 & (1-\tau)I_m & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & (1-\tau)I_m & 0 & (1-\tau)I_m & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Calculating Critical Values

One standard way in which critical values have been obtained in situations where no closed-form expressions exist is to simulate a test statistic for a large sample size for a large number of replications. In the present example, we are unable to use large sample sizes since the recursive calculations which are required over the sample are particularly time-consuming on even fast computers. For instance, with $n = 300$ and 10,000 replications it took a GATEWAY2000 486/33C over a week to do the relevant calculations for the one regressor case ($m = 1$). To reduce the computational requirements we adopt a procedure due to MacKinnon (1991). Using 10,000 replications for each sample sizes $n =$ 50, 100, 150, 200, 250, 300, we obtain critical values, $Crt(n, p, m)$, where p is the percent quantile and m is the number of regressors in the equation (excluding a constant and/or trend). We then estimate by ordinary least squares for each p and m the response surface

$$Crt(n, p, m) = \psi_0 + \psi_1 n^{-1} + error.$$

Various other functional relations were tried (involving $n^{-1/2}$, n^{-2}, $n^{-3/2}$) but this one appeared to have the best fit (R^2's were generally over .98). The asymptotic critical value is taken to be the OLS estimate $\hat{\psi}_0$. Results for $p =$ .01, .025, .05, .10, and .975 and $m =$ 1, 2, 3, and 4 are presented in Tables 1A-1D. The symbols C, C/T, and C/S refer to breaks in the constant (C), and slope coefficient (S) as defined in models 2-4. We also report the OLS standard errors of $\hat{\psi}_0$ in parentheses. These should be interpreted with some skepticism since there is obvious heteroskedasticity in the errors and more importantly, while this specification of the critical values seemed best overall, estimates of ψ_0 changed on occasion by a factor of two under alternative specifications.

5. A Simple Monte Carlo Experiment

In order to gauge the finite sample properties of the proposed test, we conduct a simple Monte Carlo experiment based upon the design of Granger and Engle (1987) and also used in Banerjee, Dolado, Hendry and Smith (1986). The model in the absence of structural change is:

$$y_{1t} = 1 + 2\,y_{2t} + \epsilon_t, \quad \epsilon_t = \rho\epsilon_{t-1} + \vartheta_t, \quad \vartheta_t \sim NID(0,1),$$

$$y_{1t} = y_{2t} + \eta_t, \quad \eta_t = \eta_{t-1} + \omega_t, \quad \omega_t \sim NID(0,1),$$

where y_{2t} is scalar ($m = 1$). We first consider the size of the various tests with $\rho = 1$, so that the null of no cointegration is true. In Table 2 we report the rejection frequencies in 1000 replications at the 5 percent level of significance (we also did one and ten percent and these are available upon request) using critical values from Table 1A. Two sample sizes ($n = 50$ and 100) are considered. ADF^*, Z_t^*, Z_α^* are the test statistics defined in equations (3.1) - (3.3) respectively and ADF is the usual augmented Dickey Fuller statistic ($\tau = 1$). The symbols C and C/T for the usual ADF refer to regressions with a constant and a constant and a trend. For ADF^* and ADF the lag length K is selected on the basis of a $t - test$ following the procedure outline in Perron and Vogelsang (1992). That is, K is chosen such that the estimated coefficient on the last included lag of the first difference is significant at the 5 percent level of significance using asymptotic normal critical values. Typically over all the experiments including those with regime shifts, K was 0 or 1 but never greater than 3.

Three general results from the size simulations are: (i) ADF^* and ADF have size near their nominal values; (ii) the Z_t^* is biased away from the null hypothesis for both sample sizes; and (iii) Z_α^* for $n = 50$ is biased towards the null but is close to asymptotic size by sample size 100. The size distortion in Z_t^* makes power comparisons problematic in the exercises below.

Turning to power in the usual cointegration setting (model 1) we let $\rho = 0$ (Table 3) and $\rho = .5$ (Table 4). As expected the best power is obtained for the conventional ADF test that makes no allowance for structural breaks (clearly the size distortion in Z_t^* clouds the interpretation of power) . On the other hand, the power loss from faulty inclusion (i.e. including unnecessary regressors to capture breaks that do not exist) is not that large. Nevertheless such rejections might mistakenly lead a researcher to believe that there is a regime shift when in fact there is a single cointegrating relation. As we would expect, all tests have lower power when the error in the cointegrating regression is serially correlated (see Gregory, 1991). In the case of no serial correlation for $n = 50$ we have rejection frequencies of over 90 percent for all tests except Z_α. Rejection frequencies fall dramatically with serial correlation (Table 4) at the same sample size (in the 40-50 percent range for ADF^* and Z_t^* and no more than 12 percent for Z_α^*). There is also the tendency with more serial correlation for the relative power decline to be larger for the structural break tests for cointegration than the standard ADF tests.

We postulate a simple structural break for the intercept, slope and then intercept and slope together (Table 5-7 respectively)

$$y_t = \gamma_t + \alpha_t\, x_t + \epsilon_t, \quad \begin{cases} \gamma_t = \gamma_1,\ \alpha_t = \alpha_1 & \text{if } t \le [\tau\, n] \\ \gamma_t = \gamma_2,\ \alpha_t = \alpha_2, & \text{if } t > [\tau\, n] \end{cases}$$

$$\epsilon_t \;=\; .5\epsilon_{t-1} \;+\; \vartheta_t$$

$$y_t = x_t + \eta_t, \quad \eta_t = \eta_{t-1} + \omega_t.$$

The two errors ϑ_t and ω_t are uncorrelated and distributed as NID(0,1). Since in applied work the errors are likely to be serially correlated, we make the break experiments similar in structure to those in Table 4. A break point occurs at $\tau = .25$, .5, and .75.

In Table 5 we investigate the ability of the tests to detect cointegration in the presence of a level shift with $\gamma_1 = 1$, $\gamma_2 = 4$, $\alpha_2 = 2$, and $\alpha_2 = 2$. With only an intercept change we find one important difference between the rejection frequencies in Table 4 (with no structural break) and Table 5 (where the intercept shifts): the rejection frequency for the conventional ADF test has fallen substantially under breaks. For instance at $n = 50$ the ADF test that includes an intercept (C) has fallen from 70 percent rejection frequency to only about 20 percent, regardless of where the break occurs. In contrast, the rejection frequencies for all three of our cointegration tests for all alternative models 2-4 are very close to those obtained with time invariant relations. As we might expect the best power results are for the test based on the level shift model 2 (C). In other experiments with smaller shifts in α_2 (not shown but available upon request) the rejection frequencies for the conventional ADF are close to those of Table 4.

In Table 6 the slope changes but the intercept is fixed over the sample: $\gamma_1 = 1$, $\gamma_2 = 1$, $\alpha_2 = 2$, and $\alpha_2 = 4$. First compared to the tests with no structural break (Table 4) the rejection frequencies have fallen considerably at $n = 50$. This changes at $n = 100$ for tests designed to detect a break in the intercept/slope (C/S). These rejection frequencies are high (over 90 percent for ADF^* and Z_t^* regardless of when the break occurs and 80 percent for Z_α^*) especially when compared to the standard ADF tests (rejection frequencies have fallen to 40 percent for breaks less than half the sample size and 70 percent for breaks at $\tau = .75$). Clearly conventional cointegration tests that do not take into account a regime shift when indeed there has been a break can be misleading. Table 6 also shows that power falls when the trend regressors are incorrectly included in the test relation (by over 10 percent for ADF^* and Z_T^* and 30 percent for Z_α^*).

Finally in Table 7 we combine the structural breaks with both a shift in the intercept and slope: $\gamma_1 = 1$, $\gamma_2 = 2$, $\alpha_2 = 2$, and $\alpha_2 = 4$ (what we have called a regime shift). These experiments are very close to those in Table 6 with very high rejection frequencies for the tests that allow for breaks in the constant and slope (C/S). Interestingly in both Tables 6 and 7 there is no difference in rejection frequencies with respect to the timing of the break.

In Table 8 we present some results on the average estimated break points for the cointegration tests together with their standard error based upon the simulations in Table 7 (regime shift). Consistent with our results in Table 7 we see that the test C/S estimate the break point more accurately with a much smaller standard deviation. The evidence also suggests that breaks that occur in the latter half of the sample are better estimated. For the experiment with $\tau = .5$ and $n = 100$ from Table 7 we graph in Figure 1 the estimated density (using a kernel estimator) for ADF^*, Z_t^*, and Z_α^*. From this figure it is clear that the three tests produce quite similar distributions for $\hat{\tau}$ and that the distributions are skewed to the right.

6. An Illustrative Example

The question of whether the long-run money-demand equation is stable has received ample attention over the years. Two recent discussions are Lucas (1988) and Stock and Watson (1991). As a practical example, we consider this question in the context of our new tests using annual data (1901-1985) and quarterly data (1960:1-1990:4). The long-run money-demand relation with no structural breaks may be written as:

$$ln(m_t) - ln(p_t) = \alpha + \gamma_1 ln(y_t) + \gamma_2 r_t + e_t.$$

The annual data are from Lucas (1988) and m is $M1$, p is the implicit price deflator, y is the real net national product and r is the six-month commercial paper rate. The quarterly data are form the CITIBASE tape; the series used are GMPY, FXGM3, GMPY82, FM1, GNNP, GNNP82 and are seasonally adjusted. This specification is identical to Lucas (1988) and Stock and Watson (1991).

With this same data, Gregory and Nason (1992) found that all of Hansen 's (1992c) tests for structural breaks detected a break for both the monthly and quarterly data and that the conventional ADF tests indicated that the null of no cointegration could be rejected for the annual data but not for the quarterly. In Table 9 we report the test statistics for our tests as well as the estimated break point (in parentheses). In addition, we calculate the test statistic for the conventional ADF test. For the conventional ADF test the lag length for K (again selected on the basis of a t-test as outlined in the Section 5) is 1 (annual) and 4 (quarterly) for both the C and C/T tests. The lags selected for ADF^* tests for the C, C/T and C/S are (2,2,0) for the annual and (0,0,0) for the monthly respectively.

Examining first the annual data, we find that the null hypothesis of no cointegration is rejected (at the 5 percent level) by our new tests using the C and C/T type formulations, but not using the C/S formulation. Since the conventional ADF test rejects the same null, it would be inappropriate to conclude from this piece of information alone that there is indeed a structural break, since a conventional cointegrated system could produce this same set of results. The evidence does suggest, however, that there is some sort of long-run cointegrating relationship between these variables. For this same model and data, Gregory and Nason (1992) were able to reject the hypothesis of a constant coefficient cointegrating relationship in favor of a one-time regime shift. Taken together, these tests suggest that a model of cointegration subject to a regime shift is a better description of the data than the conventional cointegration model. The estimated break point is roughly at half of the sample (1944). In Figure 2 we graph the $ADF(\tau)$ using the annual data for C, C/T, and C/S over the truncated sample. Clearly there is a well-defined single minimum for all three of these tests.

Turning to the quarterly data, we first notice that the conventional ADF tests fail to reject the null of no cointegration. Thus some applied researchers might conclude that there is not sufficient evidence to pursue the possibility of a long-run relationship. The tests which allow for only a level shift (the C and C/T tests) also fail to reject the null. We find, however, that the null is rejected at the 5 percent level by the most general alternative (the C/S test) which allows for both the intercept and the slope coefficient to shift. For this data set and model, allowing for the possibility of a regime shift has an important effect upon our conclusions regarding the long-run relationship between the series. The break point for the C/S tests which rejects the null is again at the middle of the sample $\hat{\tau} = .52$.

In Table 10 for both data sets we present the OLS estimates (not corrected for bias) for the money-demand estimates as well as the estimated change for models 1 - 4. Overall our results reinforce the conclusions of Gregory and Nason (1992); namely that there is evidence against the stability of the U.S. long-run money-demand equation.

7. Final Remarks

The concept of cointegration, originally formulated by Engle and Granger (1987), is that over the long-run, a special linear combination of nonstationary variables may be stationary, and thus mean-reverting. The idea was that these special linear combinations (the cointegrating vector) should reflect some sort of economic fundamentals. Upon reflection, and in the course of empirical work, it is becoming clear that applied economists may be interested in allowing the cointegrating relationships to change over time. In order for the concept of cointegration to retain significant empirical context, empirical work will have to restrict the types of structural change permitted. The most basic type of structural change is the one-time regime shift model, in which the parameters are permitted to change at one time in the sample. Other models are of course possible, but may require more careful analysis.

If structural change is to be entertained in cointegrated models, applied economists need appropriate test statistics to determine if there is any evidence for such a model. The standard testing procedure is to set up the null of no cointegration against the alternative of cointegration, so rejection is considered evidence in favor of the model. In this paper we extend this family of test statistics, by setting the alternative hypothesis to be cointegration, while allowing for a one-time regime shift of unknown timing. Rejection of the null hypothesis, therefore, provides evidence in favor of this specification.

It is important to note, however, that this type of hypothesis test does not provide much evidence concerning the question of whether or not there was a regime shift, since the alternative hypothesis contains as a special case the standard model of cointegration with no regime shift. Instead, appropriate statistics for testing the hypothesis of no regime shift against the alternative of a regime shift, for a cointegrated regression model, are given in Hansen (1992a). The test statistics of the present paper complement those of Hansen (1992a), and both types of test statistics are likely to be useful to an economist interested in the possibility of structural change in a potentially cointegrated regression model.

As illustrated in our analysis of the money demand relationship in the previous section, we believe that empirical investigations will be best served by using a number of complementary statistical tests. One difficult task for the applied researcher is to juggle these separate pieces of the puzzle, but we can offer a few suggestions. The standard ADF statistic and our ADF^* statistics both test the null of no cointegration, so rejection by either statistic implies that there is some long-run relationship in the data. If the standard ADF statistic does not reject, but the ADF^* does, this implies that structural change in the cointegrating vector may be important. If both the ADF and the ADF^* reject, no inference that structural change has occurred is warranted from this piece of information alone, since the ADF^* statistic is powerful against conventional cointegration. In this event, the tests of Hansen (1992a) are useful to determine whether the cointegrating relationship has been subject to a regime shift. Unfortunately at this stage we have little guidance as to how to control for Type I error under such procedures which suggests that additional Monte Carlo work would be worthwhile.

The analysis of this paper has been confined to the question of developing residual-based tests for cointegration in the presence of a regime shift. We have not addressed the issue of efficient estimation of a cointegrated model in the presence of a regime shift, and leave this subject for future research. We have also confined our analysis to the residual-based testing methodology originally advocated by Engle and Granger (1987) and investigated by Phillips and Ouliaris (1990). It would be interesting and useful to develop an analogous set of test statistics using the likelihood ratio testing approach advocated by Johansen (1988, 1991).

References

Andrews, D.W.K. (1990), Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica*, 59,817-858.

Andrews, D.W.K. and J.C. Monahan (1992), An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator, *Econometrica*, 60,953-966.

Banerjee, A., J.J. Dolado, D.F. Hendry, and G.W. Smith (1986) Exploring Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence, *Oxford Bulletin of Economics and Statistics*, 48, 253-277.

Banerjee, A., R. Lumsdaine, and J. 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, 271-287.

Billingsley, P. (1968), *Convergence of Probability Measures*, John Wiley Publishers, New York.

Christiano, L.J. (1992), Searching for a Break in GNP, *Journal of Business and Economic Statistics*, 10, 237-250.

Engle, R.F. and C.W.J. Granger (1987) Co-integration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, 55, 251-276.

Gregory, A.W. (1991) Testing for Cointegration in Linear Quadratic Models, unpublished manuscript, Queen's University, November.

Gregory, A.W. and J.M. Nason (1992) Testing for Structural Breaks in Cointegrated Relationships, unpublished manuscript, Queen's University, February.

Hansen, B.E. Strong Laws for Dependent Heterogeneous Processes, *Econometric Theory*, 7, 213-221.

Hansen, B.E. (1992a), Tests for Parameter Instability in Regressions with I(1) Processes, *Journal of Business and Economic Statistics*, 10, 321-335.

Hansen, B.E. (1992b) Convergence to Stochastic Integrals for Dependent Heterogeneous Processes, *Econometric Theory*, 8, 489-500.

Hansen, B.E. (1992c) Consistent Covariance Matrix Estimation for Dependent Heterogeneous Processes, *Econometrica*, 60, 967-972.

Herrndorf, N. (1984), A Functional Central Limit Theorem for Weakly Dependent Sequences of Random Variables, *The Annals of Probability*, 12, 829-839.

Johansen, S. (1988), Statistical Analysis of Cointegrating Vectors, *Journal of Economic Dynamics and Control*, 12, 231-254.

Johansen, S. (1991), Estimation and Hypothesis Testing of Cointegration Vectors in the Presence of Linear Trend, *Econometrica*, 59, 1551-1580.

Lucas, R.E. (1988) Money Demand in the United States: A Quantitative Review, *Carnegie-Rochester Conference Series on Public Policy*, 29, 137-168.

MacKinnon, J.G. (1991), Critical Values for Cointegration Tests, in *Long-Run Economic Relationships: Readings in Cointegration* edited by R.F. Engle and C.W.J. Granger, Oxford Press, 267-276.

Perron, P. (1989), The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis, *Econometrica*, 57, 1361-1401.

Perron, P. and T. J. Vogelsang (1992), Nonstationarity and Level Shifts with an Application to Purchasing Power Parity, *Journal of Business and Economic Statistics*, 10, 301-320.

Phillips, P.C.B. (1987), Time Series regression with a Unit Root, *Econometrica*, 55,277-301.

Phillips, P.C.B. and S.N. Durlauf (1986), Multiple Time Series Regression with Integrated Processes, *Review of Economic Studies*, 53, 473-495.

Phillips, P.C.B. and B.E. Hansen (1990), Statistical Inference in Instrumental Variables Regression with I(1) Processes, *Review of Economic Studies*, 57, 99-125.

Phillips, P.C.B. and S. Ouliaris (1990), Asymptotic Properties of Residual Based Tests for Cointegration, *Econometrica*, 58, 165-193.

Stock, J.H. and M. W. Watson (1991), A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems, Federal Reserve Bank of Chicago Working Paper No. 91-3, Chicago.

Zivot, E. and D.W.K. Andrews (1992), Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis, *Journal of Business and Economic Statistics*, 10, 251-270.

Appendix: Mathematical Proofs

We will rigorously prove Theorem 2. The proof of Theorem 1 is omitted , since it is quite similar. To simplify the presentation assume that $y_0 = 0$. Throughout the presentation, $\Longrightarrow$ denotes weak convergence of the associate probability measures, and $\{A\}$ denotes the indicator function for event A ($\{A\} = 1$ when A is true, and $\{A\} = 0$ when A is false).

A.1 Definitions

It turns out to be convenient to partition the random vectors in various ways in different parts of the argument. We lay out these partitions here for a point of reference. Partition $S_t = (S_{1t},\ S_{2t}^{\top})^{\top}$ in conformity with u_t. All of the statistics can be written as functions of sample moments of the $(2m+3) - vector$

$$X_{t\tau}^* = \begin{pmatrix} S_t \\ 1 \\ S_{2t}\varphi_{t\tau} \\ \varphi_{t\tau} \end{pmatrix}.$$

Define the subvectors $X_t = (S_t^{\top}, 1)^{\top}, X_{1t} = S_{1t}, X_{2t} = (S_{2t}^{\top}, 1)^{\top}, X_{2t\tau} = X_{2t}\varphi_{t\tau}, X_{1t\tau}^* = S_{1t}$, and $X_{2t\tau}^* = (X_{2t}^{\top},\ X_{2t\tau}^{\top})^{\top}$. This allows $X_{t\tau}^*$ to be partitioned several ways, including

$$X_{t\tau}^* = \begin{pmatrix} X_{1t\tau}^* \\ X_{2t\tau}^* \end{pmatrix} \begin{matrix} 1 \\ 2m+2 \end{matrix} = \begin{pmatrix} X_t \\ X_{2t\tau} \end{pmatrix} \begin{matrix} m+2 \\ m+1 \end{matrix} = \begin{pmatrix} X_{1t} \\ X_{2t} \\ X_{2t\tau} \end{pmatrix} \begin{matrix} 1 \\ m+1 \\ m+1 \end{matrix}.$$

A.2 Moment Matrices

The starting point for the asymptotic analysis is the multivariate invariance principle

$$n^{-1/2}S_{[nr]} \quad \Longrightarrow \quad B(r) \equiv BM(\Omega). \qquad (A1)$$

This holds under our assumptions as shown by Herrndorf (1984) (see Phillips and Durlauf, 1986, for the extension to the vector case). Define $\delta_n = diag(n^{-1/2}I_m,\ 1)$ and $X(r) = (B(r)^{\top},\ 1)^{\top}$. (A1) implies

$$\delta_n X_{[nr]} \quad \Longrightarrow \quad X(r) \qquad (A2)$$

and

$$\frac{1}{n}\delta_n \sum_{t=1}^{[n\tau]} X_t X_t^\top \delta_n \quad \Longrightarrow \quad \int_0^\tau XX^\top, \tag{A3}$$

by the continuous mapping theorem (see Billingsley, 1968, Theorem 5.1).

Define $\varphi_\tau = \{r \geq \tau\}$ and $X_\tau(r) = X(r)\varphi_\tau(r)$, where $\{\,\}$ denotes the indicator function. Partition $X = (X_1,\, X_2^\top)^\top$, $X_\tau = (X_{1\tau}, X_{2\tau}^\top)^\top$, and $\delta_n = diag(\delta_{1n},\, \delta_{2n})$ conformably with S_t.

From (A3) we obtain

$$\begin{aligned} \frac{1}{n}\delta_{2n}\sum_{t=1}^{n} X_{2t\tau}X_{2t\tau}^\top \delta_{2n} &= \frac{1}{n}\delta_{2n}\sum_{t=[n\tau]}^{n} X_{2t}X_{2t}^\top \delta_{2n} \\ &\Longrightarrow \int_\tau^1 X_2 X_2^\top \\ &= \int_0^1 X_{2\tau}X_{2\tau}^\top. \end{aligned} \tag{A4}$$

Now define $\delta_n^* = diag(\delta_n,\, \delta_{2n})$ and $X_\tau^*(r) = (X(r)^\top,\, X_{2\tau}(r)^\top)^\top$. (A3) and (A4) combine to yield the moment matrix for the entire data vector

$$\frac{1}{n}\delta_n^* \sum_{t=1} X_{t\tau}^* X_{t\tau}^{*\top} \delta_n^* \quad \Longrightarrow \quad \int_o^1 X_\tau^* X_\tau^{*\top} \tag{A5}$$

A.3 Least Squares Coefficient Process

The regressors in model 3 are the elements of the vector $X_{2t\tau}^*$ and the dependent variable is $y_{1t} = X_{1t\tau}^*$. Partition $X_\tau^* = (X_{1\tau}^*, X_{2\tau}^{*\top})^\top$ and $\delta_n^* = diag(\delta_{1n}^*, \delta_{2n}^*)$ conformably. Note that $\delta_{1n}^* = n^{-1/2}$. Define $\hat{\theta}_\tau = (\hat{\alpha}_1^\top,\, \hat{\mu}_1,\, \hat{\alpha}_2,\, \hat{\mu}_2)$ as the least square estimator of model 3 for each τ. It follows from (A5) and our definitions that

$$\begin{aligned} n^{-1/2}\delta_{2n}^{*^{-1}}\hat{\theta}_\tau &= [n^{-1}\delta_{2n}^*\sum_{t=1}^{n} X_{2t\tau}^* X_{2t\tau}^{*\top}\delta_{2n}^*][n^{-1}\delta_{2n}^*\sum_{t=1}^{n} X_{2t\tau}^* X_{1t\tau}^{*\top}\delta_{1n}^*] \\ &\Longrightarrow [\int_0^1 X_{2\tau}^* X_{2\tau}^{*\top}]^{-1}[\int_0^1 X_{2\tau}^* X_{1\tau}^*] \end{aligned} \tag{A6}$$

Set $\hat{\eta}_\tau = n^{-1/2}\delta_n^{*-1}(1,\, -\hat{\theta}_\tau^\top)^\top = (1, -\hat{\delta}_{2n}^{*-1}\hat{\theta}_\tau^\top)^\top$. (A6) implies that

$$\hat{\eta}_\tau \quad \Longrightarrow \quad \begin{pmatrix} 1 \\ -[\int_0^1 X_{2\tau}^* X_{2\tau}^{*\top}]^{-1}[\int_0^1 X_{2\tau}^* X_{1\tau}] \end{pmatrix} = \eta_\tau. \tag{A7}$$

A.4 Convergence to the Stochastic Integral Process

Under our assumptions we have weak convergence to the stochastic integral

$$\frac{1}{n}\sum_{t=1}^{[n\tau]} S_t u_{t+1}^\top \quad \Longrightarrow \quad \int_0^\tau B\, dB^\top \, + \, \tau\Lambda_u \tag{A8}$$

where

$$\Lambda \, = \, \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E(S_t u_{t+1}^\top),$$

as shown by Hansen (1992b, Theorem 4.1). Note that $\Delta X_t \, = \, (u_t^\top, \, 0)^\top$. Thus

$$\begin{aligned} \delta_n \sum_{t=1}^{[n\tau]} X_t \Delta X_{t+1}^\top \delta_n \; &= \; [n^{-1/2}\delta_n \sum_{t=1}^{[n\tau]} X_t u_{t+1}^\top, \, 0] \\ &\Longrightarrow [\int_0^\tau X\, dB^\top \, + \, \tau \begin{bmatrix}\Lambda_u \\ 0\end{bmatrix}, \, 0] \\ &= \int_0^\tau X\, dX^\top \, + \, \tau\Lambda \end{aligned} \tag{A9}$$

where

$$\Lambda \, = \, \begin{pmatrix} \Lambda_u & 0 \\ 0 & 0 \end{pmatrix}.$$

(Note that $dX \, = \, (dB^\top, \, d1) \, = \, (dB^\top, \, 0)$).

Partition

$$\Lambda \, = \, \begin{array}{c c} & \begin{array}{cc} 1 & m+1 \end{array} \\ \begin{array}{c} 1 \\ m+1 \end{array} & \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix} \end{array}$$

and set

$$\Lambda_{2.} \, = \, (\Lambda_{21}\, \Lambda_{22}), \;\; and \;\; \Lambda_{.2} \, = \, \begin{bmatrix}\Lambda_{12} \\ \Lambda_{22}\end{bmatrix}.$$

We can now see that

$$\begin{aligned} \delta_{2n} \sum_{t=1}^{n} X_{2t\tau} \Delta X_{t+1}^\top \delta_n \; &= \; \delta_{2n} \sum_{t=[n\tau]}^{n} X_{2t} \Delta X_{t+1}^\top \delta_n \\ &\Longrightarrow \; \int_0^1 X_{2\tau}\, dX^\top \, + \, (1-\tau)\Lambda_{2.}. \end{aligned} \tag{A10}$$

(A9) and (A10) together yield

$$\delta_n^* \sum_{t=1}^{n} X_{t\tau}^* \Delta X_{t+1}^\top \delta_n \quad \Longrightarrow \quad \int_0^1 X_\tau^*\, dX^\top \, + \, \begin{bmatrix} \Lambda \\ (1-\tau)\Lambda_{2.} \end{bmatrix} \tag{A11}$$

The process $X_{2t\tau} = X_{2t}\varphi_{t\tau}$ has differences

$$\Delta X_{2t\tau} = X_{2t}\varphi_{t\tau} - X_{2t-1}\varphi_{t-1\tau} = X_{2t-1}\Delta\varphi_{t\tau} + \Delta X_{2t}\varphi_{t\tau} \tag{A12}$$

where

$$\Delta\varphi_{t\tau} = \varphi_{t\tau} - \varphi_{t-1\tau} = \{t = [n\tau]\}.$$

To derive the large sample counterpart we need to define the differential $d\varphi_\tau$. Since $\varphi_\tau(r)$ is a step function with a jump at τ, $d\varphi_\tau$ is naturally defined as the dirac function with the property that

$$\int_a^b f\, d\varphi_\tau = \lim_{z\uparrow\tau} f(z), \quad if\ a < \tau < b,$$

for all functions $f(\cdot)$ with left-limits. Note that $\varphi_\tau\, d\varphi_\tau = 0$ and $X_\tau\, d\varphi_\tau = 0$, since the left-limit of φ_τ at τ is 0.

We can define the differential $dX_{2\tau}$ by

$$dX_{2\tau} = d(X_2\varphi_\tau) = \varphi_\tau\, dX_2 = X_2\, d\varphi_\tau. \tag{A13}$$

We have the relationships

$$\int_0^1 X\, dX_{2\tau}^\top = \int_0^1 X\varphi_\tau\, dX_2^\top + \int_0^1 XX_2^\top\, d\varphi_\tau = \int_0^1 X_\tau\, dX_2^\top + X(\tau)X_2(\tau)^\top, \tag{A14}$$

and

$$\int_0^1 X_{2\tau}\, dX_{2\tau}^\top = \int_0^1 X_{2\tau}\varphi_\tau\, dX_2^\top + \int_0^1 X_{2\tau}X_2^\top\, d\varphi_\tau = \int_0^1 X_{2\tau}\, dX_2^\top. \tag{A15}$$

Using (A12),

$$X_t\Delta X_{2t+1\tau}^\top = X_tX_{2t}^\top\Delta\varphi_{t+1\tau} + X_t\Delta X_{2t+1}^\top\varphi_{t+1\tau}$$

so

$$\begin{aligned}\delta_n\sum_{t=1}^n X_t\Delta X_{2t+1\tau}^\top\delta_{2n} &= \delta_n\sum_{t=1}^n X_tX_{2t}^\top\Delta\varphi_{t+1\tau}\delta_{2n} + \delta_n\sum_{t=1}^n X_t\Delta X_{2t+1}\varphi_{t+1\tau}\delta_{2n} \quad (A16)\\ &= \delta X_{[n\tau]-1}X_{2[n\tau]-1}^\top\delta_{2n} + \delta_n\sum_{t=[n\tau]-1}^n X_t\Delta X_{2t+1}^\top\delta_{2n}\\ &\Longrightarrow X(\tau)X_2(\tau)^\top = \int_0^1 X_\tau\, dX_2^\top = (1-\tau)\Lambda_{.2}\\ &= \int_0^1 X\, dX_{2\tau}^\top = (1-\tau)\Lambda_{.2}\end{aligned}$$

by (A14).

Using (A12) and the facts $\varphi_{t\tau}\Delta\varphi_{t+1\tau} = 0$ and $\varphi_{t\tau}\varphi_{t+1\tau} = \varphi_{t\tau}$, we have

$$\begin{aligned} X_{2t\tau}\Delta X_{2t+1\tau}^{\top} &= X_{2t}X_{2t}^{\top}\varphi_{t\tau}\varphi_{t+1\tau} + X_{2t}\Delta X_{2t+1}^{\top}\varphi_{t\tau}\varphi_{t+1\tau} \\ &= X_{2t}\Delta X_{2t+1}^{\top}\varphi_{t\tau}. \end{aligned}$$

Therefore

$$\begin{aligned} \delta_{2n}\sum_{t=1}^{n} X_{2t\tau}\Delta X_{2t+1\tau}^{\top}\delta_{2n} &= \delta_{2n}\sum_{t=[n\tau]}^{n} X_{2t}\Delta X_{2t+1}^{\top}\delta_{2n}, \\ &\Longrightarrow \int_0^1 X_{2\tau}\, dX_2^{\top} + (1-\tau)\Lambda_{22}, \\ &= \int_0^1 X_{2\tau}\, dX_{2\tau}^{\top} + (1-\tau)\Lambda_{22}. \end{aligned} \tag{A17}$$

by (A15). Putting (A16) and (A17) together we obtain

$$\delta_n^*\sum_{t=1}^{n} X_{t\tau}^*\Delta X_{2t+1\tau}^{\top}\delta_{2n} \quad \Longrightarrow \quad \int_0^1 X_\tau^*\, dX_{2\tau}^{\top} + \begin{bmatrix}(1-\tau)\Lambda_{.2} \\ (1-\tau)\Lambda_{22}\end{bmatrix}. \tag{A18}$$

Setting $dX_\tau^{*\top} = (dX^{\top},\, dX_{2\tau}^{\top})^{\top}$, and combining (A11) with (A18) we obtain our goal

$$\delta_n^*\sum_{t=1}^{n} X_{t\tau}^{*\top}\Delta X_{t+1\tau}^{*\top}\delta_n^* \quad \Longrightarrow \quad \int_0^1 X_\tau^*\, dX_\tau^{*\top} + \Lambda_\tau, \tag{A19}$$

where

$$\Lambda = \begin{pmatrix} \Lambda & (1-\tau)\Lambda_{.2} \\ (1-\tau)\Lambda_{2.} & (1-\tau)\Lambda_{22} \end{pmatrix}.$$

A.5 Serial Correlation Coefficient Estimate

Note that $\hat{e}_{t\tau} = \sqrt{n}\hat{\eta}_\tau^{\top}\delta_n^* X_{t\tau}^*$ so using (A5) and (A7)

$$n^{-2}\sum_{t=1}^{n}\hat{e}_{t\tau}^2 = \hat{\eta}_\tau^{\top}\delta_n^*\frac{1}{n}\sum_{t=1}^{n} X_{t\tau}^* X_{t\tau}^{*\top}\delta_n^*\hat{\eta}_\tau \quad \Longrightarrow \quad \eta_\tau^{\top}\int_0^1 X_\tau^* X_\tau^{*\top}\eta_\tau = \int_0^1 X_{\tau\eta}^{*2}, \tag{A20}$$

where

$$X_{\tau\eta}^*(r) = \eta_\tau^{\top}X_\tau^*(r) = X_{1\tau}^*(r) - [\int_0^1 X_{1\tau}^* X_{2\tau}^{*\top}][\int_0^1 X_{2\tau}^* X_{2\tau}^{*\top}]^{-1}X_{2\tau}^*(r), \tag{A21}$$

is the stochastic process in $(r,\, \tau)$ obtained by projecting $X^*_{1\tau}(r)$ orthogonal to the process $X^*_{2\tau}(r)$. Similarly using (A19),

$$n^{-1}\sum_{t=1}^{n} \hat{e}_{t\tau}\Delta\hat{e}_{t+1\tau} \;=\; \hat{\eta}_\tau^\top \delta^*_n \sum_{t=1}^{n} X^*_{t\tau}\Delta X^{*\top}_{t+1\tau}\delta^*_n\hat{\eta}_\tau \tag{A22}$$

$$\Longrightarrow \quad \eta_\tau^\top [\int_0^1 X^*_\tau\, dX^{*\top}_\tau \;+\; \Lambda_\tau]\eta_\tau$$

$$= \int_0^1 X^*_{\tau\eta}\, dX^*_{\tau\eta} \;+\; \eta_\tau^\top\Lambda_\tau\eta_\tau.$$

Therefore

$$n(\hat{\rho}_\tau - 1) \;=\; \frac{n^{-1}\sum_{t=1}^{n}\hat{e}_{t\tau}\Delta\hat{e}_{t+1}}{n^{-2}\sum_{t=1}^{n}\hat{e}^2_{t\tau}} \tag{A23}$$

$$\Longrightarrow \quad \frac{\int_0^1 X^*_{\tau\eta}\, dX^*_{\tau\eta} \;+\; \eta_\tau^\top\Lambda_\tau\eta_\tau}{\int_0^1 X^{*^2}_{\tau\eta}}$$

by the continuous mapping theorem (if $\int_0^1 X^{*^2}_{\tau\eta} \;>\; 0\; a.s.$, as we discuss shortly).

Partition $B \;=\; (B_1,\, B_2^\top)^\top$ and

$$\Omega \;=\; \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix},$$

in conformity with S_t. Define $\sigma \;=\; [\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}]^{1/2}$, and set

$$W_1(r) \;=\; \sigma^{-1}[B_1 \;-\; \Omega_{12}\Omega_{22}^{-1}B_2(r)]$$

$$W_2(r) \;=\; \Omega_{22}^{-1/2}B_2(r),$$

so that $W_1 \;\equiv\; BM(1)$ is independent of $W_2 \;\equiv\; BM(I_m)$.

Note that $X_{1\tau} \;=\; B_1$ and $X^*_{2\tau} \;=\; [B_2^\top(r),\, 1,\, B_2^\top(r)\varphi_\tau(r),\, \varphi_\tau(r)]^\top$. Since (A21) defines $X^*_{\tau\eta}$ by projection, and B_2 is an element of $X^*_{2\tau}$, $X^*_{\tau\eta}$ can be equivalently written as the process σW_1 projected orthogonal to $X^*_{2\tau}$. Furthermore, the space spanned by $X^*_{2\tau}$ is the same as that spanned by

$$W_{2\tau}(r) \;=\; [W_2^\top(r),\, 1,\, W_2^\top(r)\varphi_\tau(r),\, \varphi_\tau(r)]^\top.$$

We summarize this discussion by rewriting (A21) as

$$X^*_{\tau\eta}(r) \;=\; \sigma W_\tau(r) \;=\; \sigma[W_1(r) \;-\; \int_0^1 W_1 W_{2\tau}^\top\{\int_0^1 W_{2\tau}W_{2\tau}^\top\}^{-1}W_{2\tau}(r)]. \tag{A24}$$

Thus (A23) can be written as

$$n(\hat{\rho}_\tau - 1) \implies \frac{\int_0^1 W_\tau \, dW_\tau + \sigma^{-2}\eta_\tau^\top \Lambda_\tau \eta_\tau}{\int_0^1 W_\tau^2}. \tag{A25}$$

This limit distribution is only well-defined if $\int_0^1 W_\tau^2 > 0$ *a.s.* . The argument of Phillips and Hansen (1990), Lemma A.2 carries over to the present case without important modification.

Due to the presence of nuisance parameters in (A25), the uncorrected serial correlation coefficient estimate will not itself be useful as a test statistic. The main contribution to our theoretical derivation is the implication of (A25) that

$$\sup_\tau |n(\hat{\rho}_\tau - 1)| = O_p(1). \tag{A26}$$

A.6 Covariance Estimation

In Hansen (1992c, Theorem 1), it was shown under our assumptions that

$$\sum_{j=1}^{M} w(\frac{j}{M})\hat{\Gamma}(j) \longrightarrow_p \Lambda_u$$

where

$$\hat{\Gamma}(j) = \frac{1}{n}\sum_{t=1+j}^{n} u_{t-j}u_t^\top.$$

We will need the stronger result

$$\sum_{j=1}^{M} w(\frac{j}{M})\hat{\Gamma}_\tau(j) \longrightarrow_p \tau\Lambda_u, \tag{A27}$$

uniformly in $\tau \in [0,1]$, where

$$\hat{\Gamma}_\tau(j) = \frac{1}{n}\sum_{t=1+j}^{[n\tau]} u_{t-j}u_t^\top.$$

It turns out that a slight modification of the proof in Hansen (1992c) yields (A27). We provide a sketch of this modification. Theorem 1 in Hansen (1992c) was due to uniformly bounding a moment of the centered estimators $|\hat{\Gamma}_\tau(j) - E\hat{\Gamma}_\tau(j)|$, which was accomplished by appealing to an inequality in Hansen (1991, Lemma 2). The latter inequality is actually a maximal inequality, allowing the centered estimators $|\hat{\Gamma}(j) - E\hat{\Gamma}(j)|$ to be similarly bounded. The rest of the argument carries over and (A27) follows.

It follows directly from (A27) that

$$\sum_{j=1}^{n} w(\frac{j}{M})\delta_n \sum_t \Delta X_{t-j} \Delta X_t^\top \delta_n \quad \longrightarrow_p \quad \Lambda. \tag{A28}$$

Using (A12),

$$\begin{aligned}
&\sum_{j=1}^{M} w(\frac{j}{M})\delta_{2n} \sum_t \Delta X_{2t-j\tau} \Delta X_t^\top \delta_n \\
&= \sum_{j=1}^{M} w(\frac{j}{M})\delta_{2n} \sum_t [X_{2t-1-j}\Delta\varphi_{t-j\tau} + \Delta X_{2t-j}\varphi_{t-j\tau}]\Delta X_t^\top \delta_n \qquad (A29)\\
&= \sum_{j=1}^{M} w(\frac{j}{M})\delta_{2n} X_{2[n\tau]-1} \Delta X_{[n\tau]+j}^\top \delta_n \\
&+ \sum_{j=1}^{M} w(\frac{j}{M})\delta_{2n} \sum_{t=[n\tau]+j}^{n} \Delta X_{2t-j} \Delta X_t^\top \delta_n \\
&\longrightarrow_p \quad (1-\tau)\Lambda_{2.},
\end{aligned}$$

uniformly in τ, since

$$\begin{aligned}
|\sum_{j=1}^{M} w(\frac{j}{M})\delta_{2n} X_{2[n\tau]-1} \Delta X_{[n\tau]+j}^\top \delta_n| &\le \sup_{0\le\tau\le1} |\delta_{2n} X_{2[n\tau]}| n^{-1/2} \sup_{1\le t\le n} |\sum_{j=1}^{M} u_{t+j}| \\
&\le O_p(1) n^{-1/2} \sup_{1\le t\le n} |\sum_{j=1}^{M} u_{t+j}| \\
&= o_p(1),
\end{aligned}$$

the final equality due to the following inequality for α-mixing processes (Hansen, 1991, Corollary 3):

$$E[\sup_{1\le t\le n} |\sum_{j=1}^{M} u_{t+j}|]^2 = O(M).$$

Similarly,

$$\sum_{j=1}^{M} w(\frac{j}{M})\delta_2 \sum_t \Delta X_{t-j}\Delta X_{2t\tau}^{\top}\delta_{2n} \quad \longrightarrow_p \quad (1 - \tau)\Lambda_{.2}, \tag{A30}$$

uniformly in τ. Finally using (A12),

$$\begin{aligned} &\sum_{j=1}^{M} w(\frac{j}{M})\delta_{2n} \sum_t \Delta X_{2t-j\tau}\Delta X_{2t\tau}^{\top}\delta_{2n} \qquad (A31) \\ &= \sum_{j=1}^{M} w(\frac{j}{M})\delta_{2n} \sum_t [X_{2t-j\tau}\Delta\varphi_{t-j\tau} + \Delta X_{2t-j}\varphi_{t-j\tau}][X_{2t-1\tau}^{\top}\Delta\varphi_{t\tau} + \Delta X_{2t}^{\top}\varphi_{t\tau}]\delta_{2n} \\ &= \sum_{j=1}^{M} w(\frac{j}{M})\delta_{2n} \sum_{t=[n\tau]+j}^{n} \Delta X_{2t-j}\Delta X_{2t}^{\top}\delta_{2n} + o_p(1) \\ &\longrightarrow_p (1 - \tau)\Lambda_{22} \end{aligned}$$

uniformly in τ. (A28)-(A31) yield

$$\sum_{j=1}^{M} w(\frac{j}{M})\delta_n^* \sum_t \Delta X_{t-j\tau}^* \Delta X_{t\tau}^{*\top}\delta_n^* \quad \longrightarrow_p \quad \Lambda_\tau \tag{A32}$$

uniformly in τ.

A7. Bias Correction

The Phillips' statistics are constructed using the statistic $\hat{\lambda}_\tau$. We are now in a position to give its limiting distribution. First note that since $\hat{e}_{t\tau} = \sqrt{n}\hat{\eta}_\tau^{\top}\delta_n^* X_{t\tau}^*$,

$$\begin{aligned} &\sum_{j=1}^{M} w(\frac{j}{M})\Delta\hat{e}_{t-j\tau}\Delta\hat{e}_{t\tau} \qquad (A33) \\ &= \hat{\eta}_\tau^{\top} \sum_{j=1}^{M} w(\frac{j}{M})\delta_n^* \sum_t \Delta X_{t-j\tau}^*\Delta X_{t\tau}^{*\top}\delta_n^*\hat{\eta}_\tau \\ &\Longrightarrow \quad \eta_\tau^{\top}\Lambda_\tau\eta_\tau. \end{aligned}$$

by (A32) and (A7).

Now note that $\hat{\nu}_{t\tau} = \Delta\hat{e}_{t\tau} - (\hat{\rho} - 1)\hat{e}_{t-1\tau}$, so

$$|\sum_{j=1}^{M} w(\frac{j}{M})\frac{1}{n}\sum_t \Delta\hat{e}_{t-j\tau}[\Delta\hat{e}_{t\tau} - \hat{\nu}_{t\tau}]|$$

$$
\begin{aligned}
&= |\sum_{j=1}^{M} w(\frac{j}{M})\frac{1}{n}\sum_t \Delta\hat{e}_{t-j\tau}\hat{e}_{t-1\tau}(\hat{\rho} - 1)| \\
&\leq \sum_{j=1}^{M} |\frac{1}{n}\sum_t \Delta\hat{e}_{t-j\tau}\hat{e}_{t-1\tau}|\, |\hat{\rho}_\tau - 1| \\
&\leq \frac{M}{n}\max_{j\leq M}|\frac{1}{n}\sum_t \Delta\hat{e}_{t-j\tau}\hat{e}_{t-1\tau}|\ \sup_\tau n|\hat{\rho}_\tau - 1| \\
&\leq \frac{M}{\sqrt{n}}[n^{-1}\sum_t \Delta\hat{e}_t^2]^{1/2}[n^{-2}\sum_t \hat{e}_t^2]^{1/2}O_p(1) \\
&\leq \frac{M}{\sqrt{n}}\,O_p(1) \\
&= o_p(1)
\end{aligned}
$$

where we have used (A26), Holder's inequality, and the assumption that $\frac{M}{\sqrt{n}} \to 0$. Combined with (A33), it follows that

$$
\sum_{j=1}^{M} w(\frac{j}{M})\frac{1}{n}\sum_t \Delta\hat{e}_{t-j\tau}\hat{\nu}_{t\tau} \quad \Longrightarrow \quad \eta_\tau^\top \Lambda_\tau \eta_\tau.
$$

Similarly, we can replace $\Delta\hat{e}_{t-j\tau}$ with $\hat{\nu}_{t-j\tau}$, and we conclude that

$$
\hat{\lambda}_\tau = \sum_{j=1}^{M} w(\frac{j}{M})\frac{1}{n}\sum_t \hat{\nu}_{t-j\tau}\hat{\nu}_{t\tau} \quad \Longrightarrow \quad \eta_\tau^\top \Lambda_\tau \eta_\tau. \tag{A34}
$$

A.8 Z_α Process

From (A20),(A22), and (A34) we obtain

$$
\begin{aligned}
Z_\alpha(\tau) &= \frac{n^{-1}\sum_{t=1}^{n}\hat{e}_{t\tau}\Delta\hat{e}_{t+1\tau} - \hat{\lambda}_\tau}{n^{-2}\sum_{t=1}^{n}\hat{e}_{t\tau}^2} \\
&\Longrightarrow \frac{\int_0^1 X_{\tau\eta}^*\, dX_{\tau\eta}^*}{\int_0^1 X_{\tau\eta}^{*2}} = \frac{\int_0^1 W_\tau\, dW_\tau}{\int_0^1 W_\tau^2}.
\end{aligned} \tag{A35}
$$

A.9 Long-Run Variance Estimate

We can define a matrix Ω_τ as we defined Λ_τ, but using the matrix Ω instead of Λ. Writing this out explicitly,

$$
\Omega_\tau = \begin{pmatrix}
\Omega_{11} & \Omega_{12} & 0 & (1-\tau)\Omega_{12} & 0 \\
\Omega_{21} & \Omega_{22} & 0 & (1-\tau)\Omega_{22} & 0 \\
0 & 0 & 0 & 0 & 0 \\
(1-\tau)\Omega_{21} & (1-\tau)\Omega_{22} & 0 & (1-\tau)\Omega_{22} & 0 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

We can factor this matrix as

$$\Omega_\tau = \Omega_e^\top D_\tau \Omega_e$$

where

$$D\tau = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & I_m & 0 & (1-\tau)I_m & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & (1-\tau)I_m & 0 & (1-\tau)I_m & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$\Omega_e = \begin{pmatrix} \sigma & 0 & 0 & 0 & 0 \\ \Omega_{22}^{-1/2}\Omega_{21} & \Omega_{22}^{1/2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \Omega_{22}^{1/2} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

A straightforward extension of (A34) is

$$\hat{\sigma}_\tau^2 \implies \eta_\tau^\top \Omega_\tau \eta_\tau = (\Omega_e \eta_\tau)^\top D_\tau (\Omega_e \eta_\tau). \tag{A36}$$

Note that

$$\eta_\tau = \begin{pmatrix} 1 \\ -[\int_0^1 X_{2\tau}^* X_{2\tau}^{*\top}]^{-1}[\int_0^1 X_{2\tau}^* \{\sigma W_1 + X_{2\tau}^{*\top} \left[\begin{smallmatrix}\Omega_{22}^{-1}\Omega_{21} \\ 0\end{smallmatrix}\right]\}] \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ -[\int_0^1 X_{2\tau}^* X_{2\tau}^{*\top}]^{-1}[\int_0^1 X_{2\tau}^* \sigma W_1] - \left[\begin{smallmatrix}\Omega_{22}^{-1}\Omega_{21} \\ 0\end{smallmatrix}\right] \end{pmatrix}.$$

Thus

$$\Omega_e \eta_\tau = \begin{pmatrix} \sigma & 0 \\ \left[\begin{smallmatrix}\Omega_{22}^{-1/2}\Omega_{21} \\ 0\end{smallmatrix}\right] & I_{m+2} \end{pmatrix} \begin{pmatrix} 1 \\ -[\int_0^1 W_{2\tau} W_{2\tau}^\top]^{-1}[\int_0^1 W_{2\tau}\sigma W_1] - \left[\begin{smallmatrix}\Omega_{22}^{-1/2}\Omega_{21} \\ 0\end{smallmatrix}\right] \end{pmatrix}$$

$$= \sigma \begin{pmatrix} 1 \\ -[\int_0^1 W_{2\tau} W_{2\tau}^\top]^{-1}[\int_0^1 W_{2\tau} W_1] \end{pmatrix} = \sigma \kappa_\tau,$$

say. Thus we can write $\eta_\tau^\top \Omega_\tau \eta_\tau = \sigma^2 \kappa_\tau^\top D_\tau \kappa_\tau$, and reexpress (A36) as

$$\hat{\sigma}^2 \implies \sigma^2 \kappa_\tau^\top D_\tau \kappa_\tau \tag{A37}$$

Z_t Process

From (A20),(A35), and (A37), we obtain

$$
\begin{aligned}
Z_t(\tau) &= Z_\alpha(\tau)[\frac{n^{-2}\sum_{t=1}^{n}\hat{e}_{t\tau}^2}{\hat{\sigma}_\tau^2}]^{1/2} \qquad (A38)\\
&\Longrightarrow \frac{\int_0^1 X_{\tau\eta}^* \, dX_{\tau\eta}^*}{\int_0^1 X_{\tau\eta}^{*2}}[\frac{\int_0^1 X_{\tau\eta}^{*2}}{\sigma^2\kappa_\tau^\top D_\tau\kappa_\tau}]^{1/2} \\
&= \frac{\int_0^1 W_\tau \, dW_\tau}{[\int_0^1 W_\tau^2]^{1/2}[\kappa_\tau^\top D_\tau\kappa_\tau]^{1/2}}.
\end{aligned}
$$

A.12 Proof of Theorem 2

Theorem 2 follows from (A35), (A38), (A39), and the continuous mapping theorem. The supremum mapping is indeed continuous since all the limit processes are *a.s.* continuous.

Table 1A: Asymptotic Critical Values m = 1 (One Regressor)

Level	.01	.025	.05	.10	.975
ADF^*, Z_t^*					
C	-5.13 (.02)	-4.83 (.01)	-4.61 (.01)	-4.34 (.01)	-2.25 (.01)
C/T	-5.45 (.03)	-5.21 (.02)	-4.99 (.01)	-4.72 (.01)	-2.72 (.02)
C/S	-5.47 (.04)	-5.28 (.01)	-4.95 (.01)	-4.68 (.01)	-2.55 (.01)
Z_α^*					
C	-50.07 (.29)	-45.01 (.34)	-40.48 (.21)	-36.19 (.18)	-10.63 (.09)
C/T	-57.28 (.33)	-52.09 (.34)	-47.96 (.23)	-43.22 (.24)	-15.90 (.14)
C/S	-57.17 (.50)	-51.32 (.30)	-47.04 (.31)	-41.85 (.23)	-13.15 (.07)

NOTE:

These critical values are based on the response surface

$$Crt = \psi_0 + \psi_1 \, n^{-1} + error,$$

where Crt is the critcal value obtained from 10,000 replications at sample size n = 50, 100, 150, 200, 250, and 300. The asymptotic critical value is the ordinary least squares estimate (OLS) of ψ_0. OLS standard errors of this estimate are in parentheses. ADF^*, Z_t^*, Z_α^* are the test statistics defined in equations (3.1)- (3.3) respectively. The symbols C, C/T, and C/S refer to models 2 -4 respectively.

Table 1B: Asymptotic Critical Values m = 2 (Two Regressors)

Level	.01	.025	.05	.10	.975
ADF^*, Z_t^*					
C	-5.44	-5.16	-4.92	-4.69	-2.61
	(.02)	(.01)	(.01)	(.01)	(.01)
C/T	-5.80	-5.51	-5.29	-5.03	-3.01
	(.02)	(.01)	(.01)	(.01)	(.01)
C/S	-5.97	-5.73	-5.50	-5.23	-3.12
	(.04)	(.03)	(.02)	(.01)	(.01)
Z_α^*					
C	-57.01	-51.41	-46.98	-42.49	-14.27
	(.41)	(.17)	(.27)	(.24)	(.08)
C/T	-64.77	-58.57	-53.92	-48.94	-19.19
	(.87)	(.42)	(.31)	(.19)	(.12)
C/S	-68.21	-63.28	-58.33	-52.85	-19.72
	(.79)	(.65)	(.59)	(.43)	(.14)

NOTE:

These critical values are based on the response surface

$$Crt = \psi_0 + \psi_1\, n^{-1} + error,$$

where Crt is the critcal value obtained from 10,000 replications at sample size n = 50, 100, 150, 200, 250, and 300. The asymptotic critical value is the ordinary least squares estimate (OLS) of ψ_0. OLS standard errors of this estimate are in parentheses. ADF^*, Z_t^*, Z_α^* are the test statistics defined in equations (3.1)- (3.3) respectively. The symbols C, C/T, and C/S refer to models 2 -4 respectively.

Table 1C: Asymptotic Critical Values m = 3 (Three Regressors)

Level	.01	.025	.05	.10	.975
ADF^*, Z_t^*					
C	-5.77 (.02)	-5.50 (.02)	-5.28 (.02)	-5.02 (.02)	-2.96 (.01)
C/T	-6.05 (.03)	-5.79 (.02)	-5.57 (.01)	-5.33 (.01)	-3.33 (.01)
C/S	-6.51 (.02)	-6.23 (.02)	-6.00 (.01)	-5.75 (.01)	-3.65 (.01)
Z_α^*					
C	-63.64 (.27)	-57.96 (.21)	-53.58 (.23)	-48.65 (.20)	-18.20 (.11)
C/T	-70.27 (.59)	-64.26 (.40)	-59.76 (.30)	-54.94 (.23)	-22.72 (.11)
C/S	-80.15 (.58)	-73.91 (.61)	-68.94 (.53)	-63.42 (.43)	-26.64 (.06)

NOTE:

These critical values are based on the response surface

$$Crt = \psi_0 + \psi_1 \, n^{-1} + error,$$

where Crt is the critcal value obtained from 10,000 replications at sample size n = 50, 100, 150, 200, 250, and 300. The asymptotic critical value is the ordinary least squares estimate (OLS) of ψ_0. OLS standard errors of this estimate are in parentheses. ADF^*, Z_t^*, Z_α^* are the test statistics defined in equations (3.1)- (3.3) respectively. The symbols C, C/T, and C/S refer to models 2 -4 respectively.

Table 1D: Asymptotic Critical Values m = 4 (Four Regressors)

Level	.01	.025	.05	.10	.975
ADF^*, Z_t^*					
C	-6.05	-5.80	-5.56	-5.31	-3.26
	(.01)	(.01)	(.01)	(.01)	(.01)
C/T	-6.36	-6.07	-5.83	-5.59	-3.59
	(.03)	(.01)	(.01)	(.01)	(.01)
C/S	-6.92	-6.64	-6.41	-6.17	-4.12
	(.03)	(.01)	(.01)	(.00)	(.01)
Z_α^*					
C	-70.18	-64.41	-59.40	-54.38	-22.04
	(.54)	(.31)	(.28)	(.24)	(.07)
C/T	-76.95	-70.56	-65.44	-60.12	-26.46
	(.96)	(.61)	(.49)	(.29)	(.12)
C/S	-90.35	-84.00	-78.52	-72.56	-33.69
	(1.1)	(.79)	(.71)	(.60)	(.28)

NOTE:

These critical values are based on the response surface

$$Crt = \psi_0 + \psi_1 \, n^{-1} + error,$$

where Crt is the critcal value obtained from 10,000 replications at sample size n = 50, 100, 150, 200, 250, and 300. The asymptotic critical value is the ordinary least squares estimate (OLS) of ψ_0. OLS standard errors of this estimate are in parentheses. ADF^*, Z_t^*, Z_α^* are the test statistics defined in equations (3.1)- (3.3) respectively. The symbols C, C/T,C/S, and $C/S/T$ refer to breaks in the constant (C), trend (T), and slope coeffcient (S) as defined in models 2 -5 respectively.

Table 2: Size Comparisons

$$y_t = 1 + 2\,x_t + \epsilon_t, \quad \epsilon_t = \epsilon_{t-1} + \vartheta_t, \quad \vartheta_t \sim NID(0,1)$$

$$y_t = x_t + \eta_t, \quad \eta_t = \eta_{t-1} + \omega_t, \quad \omega_t \sim NID(0,1)$$

	$n = 50$	$n = 100$
<u>ADF^*</u>		
C	.07	.06
C/T	.09	.06
C/S	.09	.06
<u>Z_t^*</u>		
C	.12	.08
C/T	.14	.10
C/S	.12	.09
<u>Z_α^*</u>		
C	.00	.03
C/T	.00	.03
C/S	.00	.03
<u>ADF</u>		
C	.03	.04
C/T	.05	.04

NOTE: Rejection frequencies at the five percent level of significance using critical values from Table 1 in 1000 replications. ADF^*, Z_t^*, Z_α^* are the test statistics defined in equations (3.1)- (3.3) respectively and ADF is the usual augmented Dickey Fuller statistic with $\tau = 1$. C, C/T, and C/S refer to models 2-4 respectively and C and C/T for the usual ADF refer to regressions with a constant and a constant and a trend. For ADF^* and ADF the lag length M is set on the basis of a $t - test$ (see text).

Table 3: Power Comparisons

$$y_t = 1 + 2\,x_t + \epsilon_t, \quad \epsilon_t \sim NID(0,1),$$

$$y_t = x_t + \eta_t, \quad \eta_t = \eta_{t-1} + \omega_t, \quad \omega_t \sim NID(0,1)$$

	$n = 50$	$n = 100$
$\underline{ADF^*}$		
C	.96	1.0
C/T	.91	1.0
C/S	.95	1.0
$\underline{Z_t^*}$		
C	.99	1.0
C/T	.97	1.0
C/S	.99	1.0
$\underline{Z_\alpha^*}$		
C	.78	1.0
C/T	.37	1.0
C/S	.53	1.0
$\underline{ADF}$		
C	.95	1.0
C/T	.91	.99

NOTE: Rejection frequencies at the five percent level of significance using critical values from Table 1 in 1000 replications. ADF^*, Z_t^*, Z_α^* are the test statistics defined in equations (3.1)- (3.3) respectively and ADF is the usual augmented Dickey Fuller statistic with $\tau = 1$. C, C/T, and C/S refer to models 2-4 respectively and C and C/T for the usual ADF refer to regressions with a constant and a constant and a trend. For ADF^* and ADF the lag length M is set on the basis of a $t - test$ (see text).

Table 4: Power Comparisons with Serial Correlated Errors

$$y_t = 1 + 2\,x_t + \epsilon_t, \quad \epsilon_t = .5\,\epsilon_{t-1} + \vartheta_t, \quad \vartheta_t \sim NID(0,1)$$

$$y_t = x_t + \eta_t, \quad \eta_t = \eta_{t-1} + \omega_t, \quad \omega_t \sim NID(0,1)$$

	$n = 50$	$n = 100$
$\underline{ADF^*}$		
C	.44	.96
C/T	.39	.92
C/S	.39	.94
$\underline{Z_t^*}$		
C	.57	.98
C/T	.50	.96
C/S	.50	.97
$\underline{Z_\alpha^*}$		
C	.12	.93
C/T	.02	.80
C/S	.03	.86
ADF		
C	.70	.99
C/T	.45	.96

NOTE: Rejection frequencies at the five percent level of significance using critical values from Table 1 in 1000 replications. ADF^*, Z_t^*, Z_α^* are the test statistics defined in equations (3.1)- (3.3) respectively and ADF is the usual augmented Dickey Fuller statistic with $\tau = 1$. C, C/T, and C/S refer to models 2-4 respectively and C and C/T for the usual ADF refer to regressions with a constant and a constant and a trend. For ADF^* and ADF the lag length M is set on the basis of a $t - test$ (see text).

Table 5: Structural Break in the Intercept

$$y_t = \alpha_t + 2\,x_t + \epsilon_t, \quad \begin{bmatrix} \alpha_t = 1,\ t \leq [\tau\, T] \\ \alpha_t = 4,\ t > [\tau\, T] \end{bmatrix}$$

$$\epsilon_t \;=\; .5\epsilon_{t-1} \;+\; \vartheta_t, \quad \vartheta_t \sim NID(0,1)$$

$$y_t = x_t + \eta_t, \quad \eta_t = \eta_{t-1} + \omega_t, \quad \omega_t \sim NID(0,1)$$

	$n = 50$			$n = 100$		
τ	.25	.50	.75	.25	.50	.75
$\underline{ADF^*}$						
C	.39	.43	.49	.96	.97	.97
C/T	.32	.39	.46	.91	.93	.94
C/S	.29	.34	.40	.92	.93	.91
$\underline{Z_t^*}$						
C	.52	.54	.57	.97	.98	.98
C/T	.46	.51	.58	.94	.96	.97
C/S	.43	.45	.46	.95	.95	.96
$\underline{Z_\alpha^*}$						
C	.10	.10	.12	.92	.92	.93
C/T	.02	.02	.04	.77	.81	.83
C/S	.03	.03	.02	.80	.80	.83
ADF						
C	.24	.18	.18	.62	.51	.52
C/T	.23	.28	.33	.68	.82	.83

NOTE: Rejection frequencies at the five percent level of significance using critical values from Table 1 in 1000 replications. See Table 2 for a definition of the other symbols used.

Table 6: Structural Break in the Slope

$$y_t = 1 + \theta_t \, x_t + \epsilon_t, \quad \begin{bmatrix} \theta_t = 2, \; t \leq [\tau \, T] \\ \theta_t = 4, \; t > [\tau \, T] \end{bmatrix}$$

$$\epsilon_t \; = \; .5\epsilon_{t-1} \; + \; \vartheta_t, \quad \vartheta_t \; \sim \; NID(0,1)$$

$$y_t = x_t + \eta_t, \quad \eta_t = \eta_{t-1} + \omega_t, \quad \omega_t \; \sim \; NID(0,1)$$

	$n = 50$			$n = 100$		
τ	.25	.50	.75	.25	.50	.75
<u>ADF^*</u>						
C	.19	.28	.36	.45	.54	.83
C/T	.20	.27	.38	.39	.53	.77
C/S	.34	.41	.37	.95	.94	.93
<u>Z_t^*</u>						
C	.33	.38	.49	.53	.63	.88
C/T	.33	.38	.49	.49	.64	.84
C/S	.50	.49	.46	.97	.96	.96
<u>Z_α^*</u>						
C	.06	.06	.10	.37	.45	.78
C/T	.02	.02	.03	.25	.36	.64
C/S	.04	.03	.03	.85	.82	.80
ADF						
C	.23	.22	.40	.38	.40	.67
C/T	.15	.20	.41	.34	.39	.74

NOTE: Rejection frequencies at the five percent level of significance using critical values from Table 1 in 1000 replications. See Table 2 for a definition of the other symbols used.

Table 7: Structural Break in the Intercept and Slope

$$y_t = \alpha_t + \theta_t\, x_t + \epsilon_t, \quad \begin{bmatrix} \alpha_t = 1,\ \theta_t = 2,\ t \le [\tau\, T] \\ \alpha_t = 2,\ \theta_t = 4,\ t > [\tau\, T] \end{bmatrix}$$

$$\epsilon_t \;=\; .5\epsilon_{t-1} \;+\; \vartheta_t, \quad \vartheta_t \;\sim\; NID(0,1)$$

$$y_t = x_t + \eta_t, \quad \eta_t = \eta_{t-1} + \omega_t, \quad \omega_t \sim NID(0,1)$$

	$n = 50$			$n = 100$		
τ	.25	.50	.75	.25	.50	.75
$\underline{ADF^*}$						
C	.18	.26	.34	.46	.54	.82
C/T	.17	.25	.36	.37	.52	.77
C/S	.34	.36	.34	.95	.93	.93
$\underline{Z_t^*}$						
C	.31	.35	.48	.56	.63	.90
C/T	.30	.38	.48	.50	.62	.85
C/S	.48	.47	.45	.98	.96	.96
$\underline{Z_\alpha^*}$						
C	.06	.06	.09	.39	.46	.79
C/T	.01	.01	.02	.26	.37	.66
C/S	.03	.02	.02	.85	.82	.81
ADF						
C	.24	.22	.43	.42	.40	.66
C/T	.14	.19	.43	.37	.42	.74

NOTE: Rejection frequencies at the five percent level of significance using critical values from Table 1 in 1000 replications. See Table 2 for a definition of the other symbols used.

Table 8: Estimating Break Point

	$n = 50$			$n = 100$		
τ	.25	.50	.75	.25	.50	.75
<u>ADF^*</u>						
C	.55	.66	.66	.49	.66	.71
	(.22)	(.18)	(.22)	(.24)	(.18)	(.20)
C/T	.55	.59	.59	.50	.62	.65
	(.21)	(.20)	(.21)	(.22)	(.19)	(.22)
C/S	.43	.60	.66	.32	.56	.72
	(.19)	(.14)	(.19)	(.13)	(.10)	(.15)
<u>Z_t^*</u>						
C	.49	.63	.61	.45	.66	.70
	(.23)	(.58)	(.24)	(.23)	(.18)	(.21)
C/T	.52	.58	.59	.50	.63	.66
	(.22)	(.20)	(.22)	(.22)	(.19)	(.22)
C/S	.40	.57	.62	.32	.55	.71
	(.18)	(.15)	(.21)	(.13)	(.10)	(.16)
<u>Z_α^*</u>						
C	.49	.63	.61	.46	.66	.70
	(.23)	(.20)	(.22)	(.23)	(.18)	(.21)
C/T	.53	.59	.60	.50	.63	.66
	(.21)	(.20)	(.22)	(.22)	(.19)	(.22)
C/S	.40	.57	.52	.32	.56	.71
	(.18)	(.15)	(.20)	(.13)	(.10)	(.17)

NOTE: Average estimated break point and its standard error in parentheses. The results are for the experiment described Table 6.

Table 9 Testing for Regime Shifts in U.S. Money Demand

$$ln(m_t) - ln(p_t) = \alpha + \gamma_1 ln(y_t) + \gamma_2 r_t + e_t$$

	Annual Data 1901-1985		Quaterly Data 1960:1-1990:4	
	Test Stat	Break Point	Test Stat	Break Point
$\underline{ADF^*}$				
C	-5.43**	(.49)	-3.73	(.48)
C/T	-5.66**	(.50)	-4.63	(.85)
C/S	-5.01	(.49)	-5.70**	(.52)
$\underline{Z_t^*}$				
C	-5.33**	(.50)	-3.57	(.85)
C/T	-5.85**	(.49)	-4.66	(.85)
C/S	-5.38*	(.49)	-5.82**	(.52)
$\underline{Z_\alpha^*}$				
C	-46.39*	(.50)	-22.94	(.85)
C/T	-52.52*	(.49)	-36.38	(.85)
C/S	-44.36	(.50)	-53.57*	(.52)
$\underline{ADF}$				
C	-4.55**	(–)	-2.03	(–)
C/T	-4.37**	(–)	-1.89	(–)

NOTE: These are the test statistics where an * and ** indicates significance at the ten and five percent respectively. Beside these in parentheses are the estimated break points. The annual data are from Lucas (1988) and m is $M1$, p is the implicit price deflator, y is the real net national product and r is the six-month commercial paper rate. The quarterly data are form the CITIBASE tape; the series used are GMPY, FXGM3, GMPY82, FM1, GNNP, GNNP82 and are seasonally adjusted.

Table 10 Parameter Estimates in U.S. Money Demand using ADF^*

$$ln(m_t) - ln(p_t) = \alpha + \beta t + + \gamma_1 ln(y_t) + \gamma_2 r_t + e_t$$

	$\hat{\alpha}$	$\Delta\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}_1$	$\Delta\hat{\gamma}_1$	$\hat{\gamma}_2$	$\Delta\hat{\gamma}_2$
Annual Data							
Model 1	1.85	-	-	.96	-	-.08	-
Model 2 (C)	1.73	-.12	-	.98	-	-.08	-
Model 3 (C/T)	3.04	-.180	.012	.59	-	-.07	-
Model 4 (C/S)	1.81	5.40	-	.98	-1.03	-.09	.08
Quarterly Data							
Model 1	3.22	-	-	.40	-	-.022	-
Model 2 (C)	4.86	.17	-	.18	-	-.01	-
Model 3 (C/T)	1.55	.20	-.003	.63	-	-.008	-
Model 4 (C/S)	4.86	.20	-	.18	.25	-.01	-.003

NOTE: These are the ordinary least squares estimates without any bias corrections. The annual data are from Lucas (1988) and m is $M1$, p is the implicit price deflator, y is the real net national product and r is the six-month commercial paper rate. The quarterly data are form the CITIBASE tape; the series used are GMPY, FXGM3, GMPY82, FM1, GNNP, GNNP82 and are seasonally adjusted.

Figure 1: Estimating Break Points

T = 100

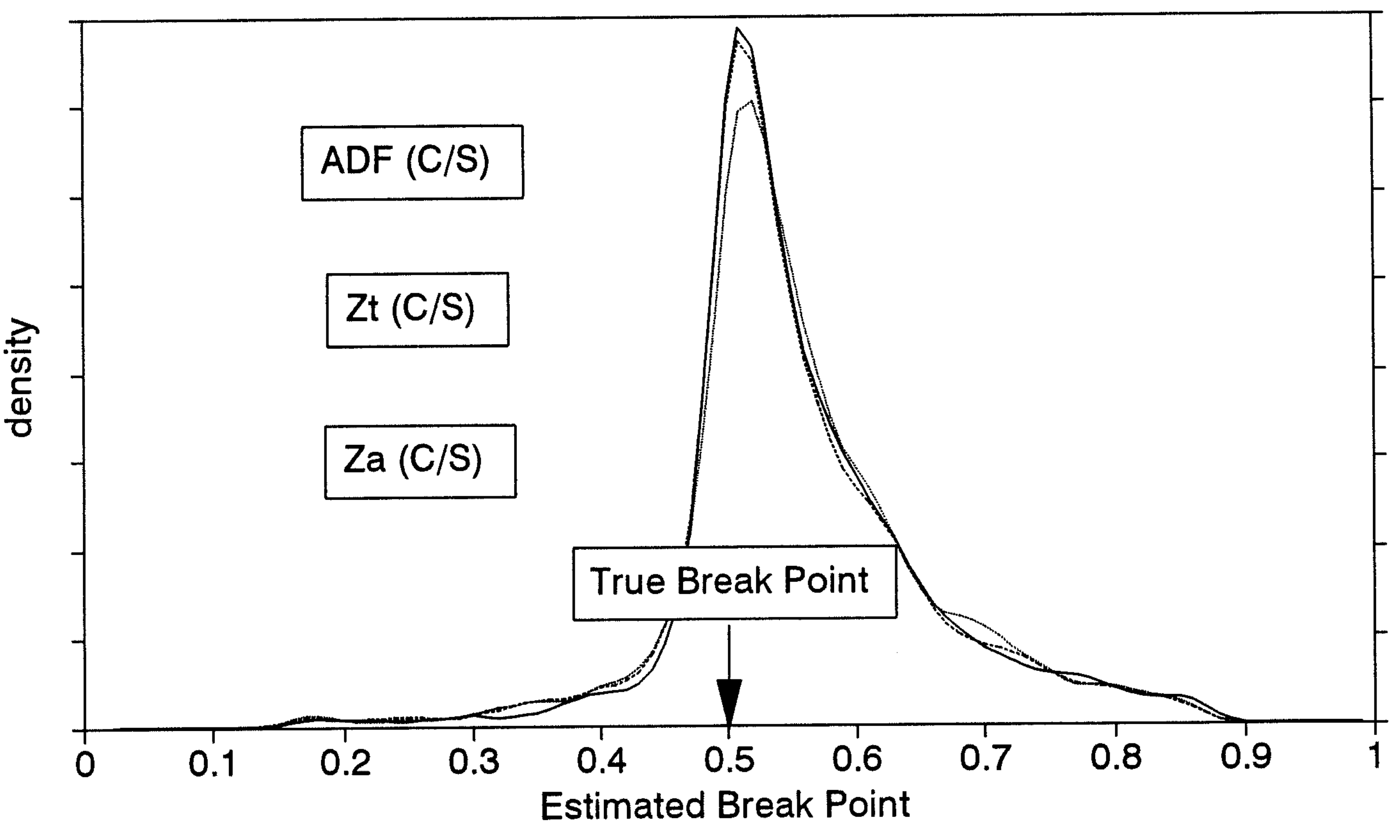

Figure 2: Regime Shifts with ADF*

U.S. Annual Money Demand (1901-1985)