Cointegration Tests in the Presence of Structural Breaks

Julia Campos, Neil R. Ericsson, and David F. Hendry [1]

1 Introduction

Structural breaks in stationary time series can induce apparent unit roots, as shown by Perron (1989) analytically and empirically and by Hendry and Neale (1991) via Monte Carlo. Consequently, tests of unit roots have low power when applied to series with structural breaks. Conversely, such series mimic series with actual unit roots. Using Monte Carlo, this paper investigates the power of several cointegration tests when the marginal process of one of the variables in the cointegrating relationship contains a structural break. The detectability of the structural break itself is also examined, both by classical constancy tests and by recently introduced tests for invariance. The data include stationary and non-stationary series with breaks. Calculation and analysis of results employ recently developed Monte Carlo techniques, as described in Hendry (1984) and Hendry and Neale (1987, 1990).

A structural break has little effect on the size of the cointegration tests studied. However, the break does affect the power of cointegration tests when the process generating the data does not have a common factor. Specifically, tests based on estimated error correction models generally are more powerful than Engle and Granger's (1987) commonly used "two-step" procedure employing the Dickey-Fuller unit root test. The error-correction-based test uses available information more efficiently than the Dickey-Fuller test, paralleling Kremers, Ericsson, and Dolado's (1992) results without a structural break.

Section 2 describes the data generation process for the Monte Carlo study, the test statistics considered, and some of their analytical properties. Section 3 presents the experimental design and simulation techniques. Section 4 examines the detectability of stationarity and of the break in the marginal process. Section 5 interprets the Monte Carlo results on the cointegration tests. Section 6 concludes. The appendix derives asymptotic properties of the unit root estimator when the underlying process has a break.

2 The Data Generation Process, the Test Statistics, and Some Analytical Properties

This section describes the data generation process for the Monte Carlo simulation (Section 2.1), the test statistics studied (Section 2.2), and some analytical properties of those statistics (Section 2.3). The data generation process (DGP) is a first-order bivariate vector autoregressive process with a possible structural break in one of the variables' processes. The test statistics are Dickey-Fuller (static regression) and error correction t-ratios, and each is considered with and without knowledge of the cointegrating vector (if it exists). Together, the DGP and the test statistics delimit the scope of the analysis. Some analytical properties of the test statistics prove helpful in interpreting the Monte Carlo evidence.

---

[1] The first and third authors are professors at the Universidad de Alicante, Alicante, Spain, and at Nuffield College, Oxford, England respectively; the second author is a staff economist in the Division of International Finance, Federal Reserve Board, Washington, D.C., U.S.A. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting those of the Board of Governors of the Federal Reserve System or other members of its staff. We are grateful to Peter Boswijk, Jean-Marie Dufour, Andreas Fischer, Eric Ghysels, Peter Phillips, Carmela Quintos, and two referees for helpful comments, and to the U.K. Economic and Social Research Council for financial support to the third author under grant B00220012. All numerical results were obtained using PCNAIVE Version 6.01 and PcGive Version 7.00; cf. Hendry and Neale (1990) and Doornik and Hendry (1992).

2.1 The Data Generation Process

The DGP is a linear first-order vector autoregression with normal disturbances, Granger causality in only one direction, and a possible structural break in the strictly exogenous process. For expositional convenience, this DGP is written as a conditional error correction model (1) and a marginal process (2):

$$\Delta y_t = a\Delta z_t + b(y - \lambda z)_{t-1} + \epsilon_t \tag{1}$$

$$z_t = \rho z_{t-1} + \delta D_t + u_t, \tag{2}$$

where

$$\left[\begin{array}{c} \epsilon_t \\ u_t \end{array}\right] \sim IN\left(\left[\begin{array}{c} 0 \\ 0 \end{array}\right], \left[\begin{array}{cc} \sigma\epsilon^2 & 0 \\ 0 & \sigma_u^2 \end{array}\right]\right) \qquad t = 1, ..., T, \tag{3}$$

and

$$D_t = \begin{cases} 1 & \text{if } t = T_0 + 1, \ldots, T_1 \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

$\mathcal{L}$ is the lag operator, Δ is the first-difference operator $1 - \mathcal{L}$, T is the sample size, and $T_0 + 1$ and T_1 are the beginning and the end of the break ($1 \leq T_0 \leq T_1 \leq T$). Note that $a = \lambda$ implies a valid common factor restriction. That is, (1) becomes $[1 - (1 + b)\mathcal{L}]y_t = \lambda[1 - (1 + b)\mathcal{L}]z_t + \epsilon_t$ when $a = \lambda$ to create a common factor between y_t and z_t of $[1 - (1 + b)\mathcal{L}]$; see Sargan (1980).

For $Y = \exp(y)$ and $Z = \exp(z)$, a is the short-run elasticity of Y with respect to Z, and λ is the long-run elasticity (provided $b \neq 0$). The parameter b is the error correction coefficient in the conditional model of y_t, given lagged y and current and lagged z; and ϵ_t and u_t are the disturbances in this conditional-marginal factorization. By a suitable scaling of z_t (and without loss of generality), $\lambda = 1$: that is, the cointegrating vector for $(y_t : z_t)'$ is $(1 : -1)$ if y_t and z_t are cointegrated. Knowing that $\lambda = 1$ and imposing λ at that value for estimation is *with* loss of generality, so Section 2.2 considers tests that utilize and that ignore this information.

The parameter space is restricted to $\{0 \leq a \leq 1,\ -1 \leq b \leq 0\}$. In many empirical studies, $a \approx 0.5$ and $b \approx -0.1$, with $\sigma_u^2 \geq \sigma\epsilon^2$. That is, the short-run elasticity (a) is smaller than the long-run elasticity (unity), adjustment to remaining disequilibria is slow, and the innovation error variance for the regressor process is larger than that of the conditional process. Also, z_t is assumed weakly exogenous for the parameters in the conditional model (1); see Engle, Hendry, and Richard (1983) and Johansen (1992a). Section 3.1 gives the precise experimental design.

Four types of processes for z_t can arise from (2), and they are denoted Cases I–IV.

Case I: z_t has a unit root ($\rho = 1$) but no break ($\delta = 0$). The variables y_t and z_t are integrated of order one [denoted I(1)] and are cointegrated if $-1 \leq b < 0$. Banerjee, Dolado, Hendry, and Smith (1986) and Kremers, Ericsson, and Dolado (1992) *inter alia* analyze the properties of various cointegration estimators and test statistics under this DGP, both analytically and by Monte Carlo.

Case II: z_t is stationary ($|\rho| < 1$) with no break ($\delta = 0$). Then, y_t is I(1) if $b = 0$, and y_t and z_t are jointly stationary with an error correction representation if $-1 \leq b < 0$. Davidson, Hendry, Srba, and Yeo (1978) and Davidson and Hendry (1981) provide some asymptotic and Monte Carlo evidence on the properties of the error correction test statistic.

Case III: z_t is I(1) ($\rho = 1$) and has a break ($\delta \neq 0$). If the break is large enough, z_t may appear to be an I(2) process. Likewise, a trend-stationary variable with a break in the slope may appear I(2): the difference of the series has a break in mean. While Case III is of potential empirical interest, this paper only briefly considers it, in Section 4. However, see Johansen (1992b, 1992c) for theoretical and empirical analyses of I(2) variables.

Case IV: z_t is stationary ($|\rho| < 1$) and has a break ($\delta \neq 0$). This case is the primary focus of this paper. From extensive Monte Carlo evidence in Hendry and Neale (1991), z_t may well appear to have a unit root when standard unit root tests are applied, and the break may be difficult to detect (see also Section 4 below). With z_t behaving like a unit root process with no break (Case I), we conjectured that cointegration tests involving such a z_t process would behave as in Case I. For the most part, this conjecture appears correct. However, the imposition of a common factor restriction by Engle-Granger cointegration tests plays an even larger role than anticipated, as shown both analytically (Section 2) and in the Monte Carlo (Section 5).

As implied by Kremers, Ericsson, and Dolado (1992, Section 5), the logical issues arising from common factor restrictions apply to processes more general than (1)–(4). Specifically, the cointegrating vector or vectors may enter more than one equation (i.e., no weak exogeneity); and a constant term, seasonal dummies, additional variables, additional lags, and error autocorrelation may be included. Likewise, generalizations of the Dickey-Fuller statistic such as the augmented Dickey-Fuller (ADF), $Z\alpha$, and Z_t statistics still impose a common factor restriction and so suffer from a loss of information when that restriction is invalid. With more general DGPs, the distributions of some statistics are more complicated, so this paper focuses on the bivariate case.

Before describing the test statistics in detail, it should be emphasized that either Case I or Case IV may characterize empirical time series, and it may be extremely difficult to distinguish between the two cases in practice. To illustrate, consider real per capita expenditure on non-durables and services (CN) and real per capita disposable income (INC) in Venezuela, graphed in logs in Figure 1 as cn and inc. Initially, both series grow smoothly. Then, income jumps by over 30% in 1974 from increased petroleum revenues, remains relatively constant through 1981, and during the LDC debt crisis of the early 1980s falls to approximately its 1973 level. Expenditure parallels or lags behind income through 1981, but falls only slightly during 1981–1985. Campos and Ericsson (1988) find that each series appears I(1) empirically, and that expenditure and income are cointegrated, provided that inflation and liquidity effects are properly accounted for.

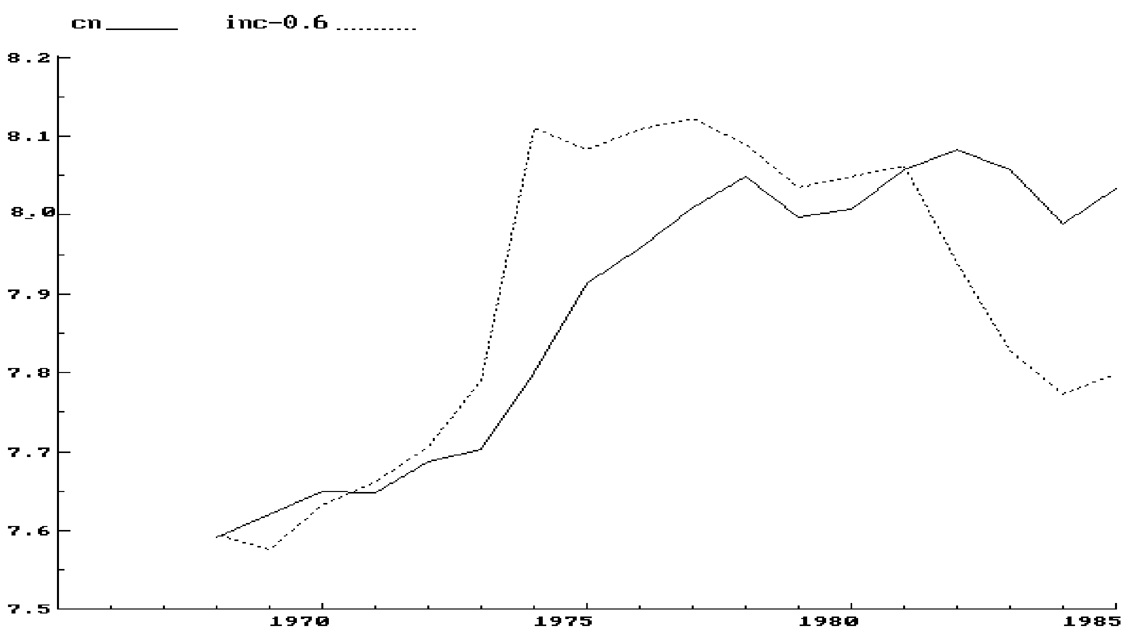

Figure 1: Venezuela data

Still, the expenditure and income series in Figure 1 could easily have been generated by a stationary process with breaks in 1974 and 1981. Figure 2 plots the series y and z generated by (1)–(4) under Case IV,[2] and these series bear a marked resemblance to cn and inc in Figure 1. Yet, the power of a Dickey–Fuller test to detect stationarity in such z_t is less than its size; and the power of the Chow test to detect the break in z_t is about 10% (at a nominal 5% level), even when the break point is known. Such similarities in time series motivate the Monte Carlo study below.

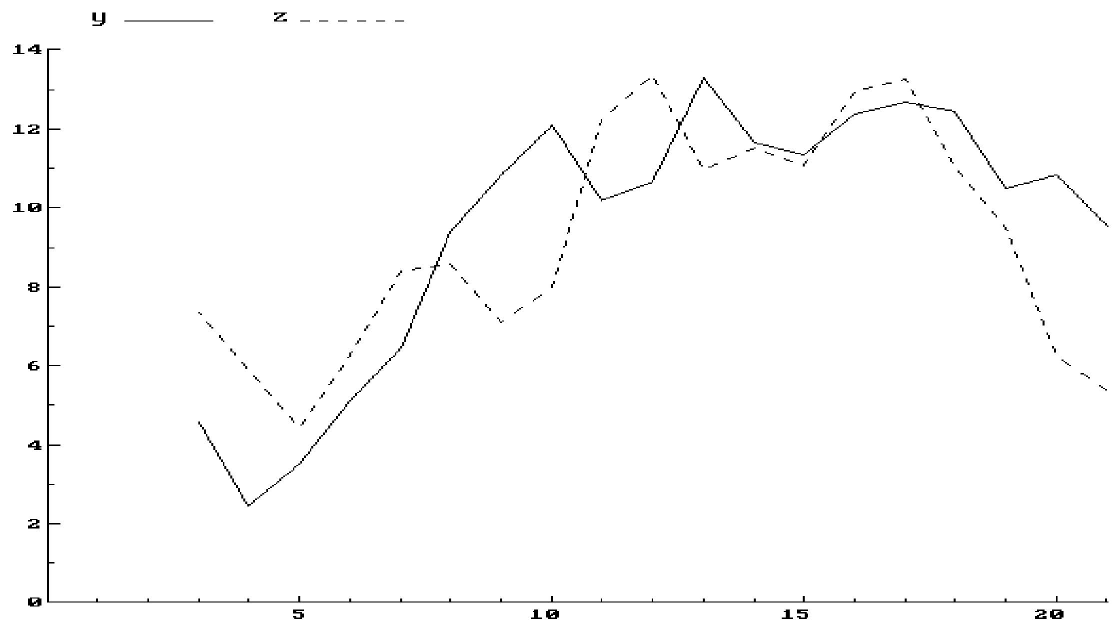

Figure 2: Artificial data

2.2 The Test Statistics

This subsection describes the test statistic for a unit root in the marginal process (2) and the test statistics for the cointegration of y_t and z_t.

2.2.1 Marginal Processes

Frequently, investigators pre-test for unit roots in univariate autoregressive representations. Many tests exist: see Dickey and Fuller (1979, 1981), Phillips and Perron (1988), and Banerjee, Dolado, Galbraith, and Hendry (1993) *inter alia*. While only the Dickey-Fuller t-statistic (denoted t_{DF}) is considered in this paper, some of its properties are generic to unit root statistics. The asymptotic distribution of t_{DF} is well-known for difference-stationary processes such as Case I. As shown in (15), its distribution is different when a break occurs.

2.2.2 Cointegrated Processes

Using the dynamic bivariate process (1)–(4), this paper focuses on the relative merits of the two-step Engle-Granger and single-step dynamic-model procedures for testing for the existence of cointegration.

[2] Specifically, $a = 0$, $b = -0.1$, $\rho = 0.8$, $\delta = 1$, $\sigma\epsilon = \sigma_u = 1$, $T = 100$, $T_0 = 25$, and $T_1 = 75$. To ensure comparable samples for Figures 1 and 2, the Monte Carlo series y_t and z_t were averaged over four periods and selectively sampled, extracting every fourth (averaged) observation. Figure 2 plots observations 3 through 21 of the resulting series: breaks occur in observations 7 and 19. The graph of the selectively sampled unaveraged series is similar to Figure 2. Calculations for size and power use the full sample ($T = 100$).

See Engle and Granger (1987) on the former and Banerjee, Dolado, Hendry, and Smith (1986) *inter alia* on the latter. The former is characterized by a Dickey-Fuller (DF) statistic used to test for the existence of a unit root in the residuals of a static cointegrating regression. The latter is based upon the t-ratio of the coefficient on the error correction term in a dynamic model reparameterized as an error correction mechanism (ECM), noting that cointegration implies and is implied by an ECM. This t-ratio is denoted the ECM statistic. Each statistic may utilize or ignore knowledge about the value of the cointegrating vector (i.e., that $\lambda = 1$).[3] This subsection describes these four test statistics and clarifies analytical relationships between the test statistics; Section 2.3 presents some of their asymptotic properties.

The variables y_t and z_t are cointegrated or not, depending upon whether $b < 0$ or $b = 0$. Thus, tests of cointegration rely upon some estimate of b. The four statistics considered here are all t-ratios on regression estimates of b. Let us denote those t-ratios by t_{ECMk}, t_{ECMu}, t_{DFk}, and t_{DFu}, where the subscripts ECM and DF denote error correction model and Dickey-Fuller, and k and u indicate that the cointegrating vector is assumed known or unknown. These t-ratios on b are derived from the regressions:

$$\Delta y_t = \kappa + a\Delta z_t + b(y-z)_{t-1} + \epsilon_{kt}, \tag{5}$$

$$\Delta y_t = \kappa + a\Delta z_t + b(y-z)_{t-1} + cz_{t-1} + \epsilon_{ut}, \tag{6}$$

$$\begin{aligned} y_t &= \kappa_0 + z_t + w_{kt} \\ \Delta(y - \tilde{\kappa}_0 - z)_t &= \kappa_1 + b(y - \tilde{\kappa}_0 - z)_{t-1} + e_{kt}, \end{aligned} \tag{7}$$

and

$$\begin{aligned} y_t &= \kappa_0 + \lambda z_t + w_{ut} \\ \Delta(y - \tilde{\kappa}_0 - \tilde{\lambda} z)_t &= \kappa_1 + b(y - \tilde{\kappa}_0 - \tilde{\lambda} z)_{t-1} + e_{ut}, \end{aligned} \tag{8}$$

respectively, where κ, κ_0, κ_1, and c are constants; and a tilde $\tilde{}$ denotes an estimate from the static regression in (8). The regression errors for t_{ECMk} and t_{ECMu} are ϵ_{kt} and ϵ_{ut} respectively. Those for t_{DFk} and t_{DFu} are e_{kt} and e_{ut}, with $\tilde{w}_{kt}$ and $\tilde{w}_{ut}$ being the associated static-regression residuals on which t_{DFk} and t_{DFu} are based.

The statistics t_{DFk} and t_{ECMk} are used to test the null hypothesis that $b = 0$ in (1), i.e., that y and z are *not* cointegrated with a cointegrating vector $(1 : -1)$. The statistics t_{DFu} and t_{ECMu} ignore that $\lambda = 1$, and so are (implicitly or explicitly) used to test the null hypothesis that y and z are *not* cointegrated with an arbitrary cointegrating vector $(1 : -\lambda)$.

Both asymptotic and finite sample properties of the statistics can be better understood by examining the relationship between the DF regression equation (7) and the ECM regression equation (5). To do so, subtract Δz_t from both sides of the conditional DGP (1) and rearrange:

$$\begin{aligned} \Delta(y-z)_t &= b(y-z)_{t-1} + [(a-1)\Delta z_t + \epsilon_t] \\ &= b(y-z)_{t-1} + e_t, \end{aligned} \tag{9}$$

where the disturbance e_t is:

$$e_t = (a-1)\Delta z_t + \epsilon_t. \tag{10}$$

[3] A priori knowledge of the cointegrating vector frequently arises in economic modeling: for instance, of (logs of) consumers' expenditure and disposable income, of wages and prices, of money and income, and of the exchange rate and foreign and domestic price levels.

Equation (7) is (9) with $e_{kt} = e_t$ and an explicit constant term. The statistic t_{DFk} ignores potential information contained in Δz_t. Equivalently, (7) imposes the restriction that the short- and long-run elasticities are equal ($a = 1$), or that there is a common factor in the relationship between y_t and z_t.

For the DGP (1)–(4), the error e_t in (10) is white noise because both Δz_t and ϵ_t are white noise. Thus, the ADF, $Z\alpha$, and Z_t statistics offer no additional benefit in this case. The problem is the common factor restriction, which highlights the difference between white noise and an innovation. As discussed in Kremers, Ericsson, and Dolado (1992), the ADF, $Z\alpha$, and Z_t statistics impose a common factor restriction.

A useful measure of the information ignored by t_{DFk} is:

$$q \quad = \quad -(a-1)s, \tag{11}$$

where $s = \sigma_u/\sigma\epsilon$. That is, q^2 is the variance of $(a-1)\Delta z_t$ relative to that of ϵ_t. Also, $q^2 = \mathcal{R}^2/(1-\mathcal{R}^2)$, where $\mathcal{R}^2$ is the population R^2 for Δw_{kt} (or e_t in (10)) regressed on Δz_t when $b = 0$. The value of q directly affects the distribution of t_{ECMk}; see Section 2.3. See Kremers, Ericsson, and Dolado (1992) for details on the restriction $a = 1$, and Hendry and Mizon (1978) and Sargan (1964, 1980) on common factors.

The relationship between (8) and (6) (the regressions with unknown λ) parallels that between (7) and (5). Equation (1) can be transformed to:

$$\Delta(y-\tilde{\lambda}z)_t \quad = \quad b(y-\tilde{\lambda}z)_{t-1} + [(a-1)\Delta z_t + (1-\tilde{\lambda})\Delta z_t + b(\tilde{\lambda}-1)z_{t-1} + \epsilon_t], \tag{12}$$

which is (8) where the disturbance e_{ut} is:

$$e_{ut} \quad = \quad (a-1)\Delta z_t + (1-\tilde{\lambda})\Delta z_t + b(\tilde{\lambda}-1)z_{t-1} + \epsilon_t, \tag{13}$$

and the constant in (12) is implicit. Thus, t_{DFu} is affected not only by the common factor restriction, but also by the discrepancy between estimated and actual values of the cointegrating vector. While “super consistent”, the static-regression estimate $\tilde{\lambda}$ may have poor finite sample properties, thereby affecting the properties of t_{DFu}; see Banerjee, Dolado, Hendry, and Smith (1986).

2.3 Properties of the Test Statistics

This subsection describes properties of the test statistic for a unit root in the marginal process and properties of the cointegration test statistics.

2.3.1 Marginal Processes

Dickey and Fuller (1979, 1981), Phillips (1987b, 1988), and Banerjee, Dolado, Galbraith, and Hendry (1993) derive the asymptotic distributions of t_{DF} under Cases I and II; and they are identical to the distributions of t_{DFk}, as described in Section 2.3.2 below. The appendix derives properties of t_{DF} under Case III, providing a baseline for interpreting the Monte Carlo results, which are for Case IV.

Under Case III, z_t has a unit root ($\rho = 1$) and a break ($\delta \neq 0$). Suppose an investigator, unaware of the break, estimates

$$\Delta z_t \quad = \quad \mu + \phi z_{t-1} + \xi_t \tag{14}$$

in order to test for a unit root ($\phi = 0$). The probability limit of the least squares estimator $(\hat{\mu} : \hat{\phi})$ of $(\mu : \phi)$ is:

$$
\begin{aligned}
\operatorname{plim}\begin{bmatrix}\hat{\mu}\\ T\hat{\phi}\end{bmatrix} &= \operatorname{plim}\begin{bmatrix} 1 & T^{-2}\sum_{1}^{T} z_{t-1} \\ T^{-2}\sum_{1}^{T} z_{t-1} & T^{-3}\sum_{1}^{T} z_{t-1}^{2}\end{bmatrix}^{-1} \operatorname{plim}\begin{bmatrix}\delta K + O_p(T^{-1/2}) \\ \frac{1}{2}\delta^2 K^2 + O_p(T^{-1/2})\end{bmatrix} \\
&= \begin{bmatrix} 1 & \frac{1}{2}\delta K(K+2M) \\ \frac{1}{2}\delta K(K+2M) & \frac{1}{3}\delta^2 K^2 (K+3M)\end{bmatrix}^{-1} \begin{bmatrix}\delta K \\ \frac{1}{2}\delta^2 K^2\end{bmatrix} \\
&= \tfrac{1}{2} H^{-1} \begin{bmatrix}\delta K(M+K/6) \\ 1-(K+2M)\end{bmatrix},
\end{aligned}
\tag{15}
$$

where K is the length of the break $(T_1 - T_0)/T$, M is the time after the break $(T - T_1)/T$, and $H = (1 - M - K)M + K(4 - 3K)/12$. As found in additional Monte Carlo simulations not reported here, the estimated means of $\hat{\mu}$ and $\hat{\phi}$ appear to match closely the analytical values in (15).

Four implications of (15) are of interest. First, the estimated intercept has a nonzero population value, which is proportional to δ. Second, the unit root estimator $\hat{\phi} + 1$ differs from unity by only $O_p(T^{-1})$. Third, the corresponding discrepancy does not depend on δ (to $o_p(T^{-1})$) and is relatively negligible, especially when $(K + 2M)$ is close to unity. Fourth, a break in the marginal process (2) affects the cointegration statistic t_{DFk} and the unit root statistic t_{DF} for z_t similarly when the common factor restriction is not valid for t_{DFk} ($a \neq 1$). For t_{DFk}, the equation being estimated is:

$$
\begin{aligned}
\Delta(y-z)_t &= b(y-z)_{t-1} + e_{kt} \\
&= b(y-z)_{t-1} + [(a-1)\Delta z_t + \epsilon_t] \\
&= b(y-z)_{t-1} \\
&\qquad + [(a-1)\delta D_t + (a-1)(\rho-1)z_{t-1} + (a-1)u_t + \epsilon_t]
\end{aligned}
\tag{16}
$$

where the constant is implicit and the term in square brackets is the error. Under Case III, $\rho = 1$, so the error is comprised of a break $(a-1)\delta D_t$ and white noise $(a-1)u_t + \epsilon_t$, paralleling the regression (14) under the DGP (2) with $\rho = 1$ and $\delta \neq 0$. Under Case IV, $|\ \rho\ | < 1$, so the error also includes z_{t-1}, which has a break in it. Thus, whenever the marginal process has a break and the conditional process does not have a common factor, the regression for t_{DFk} induces a break in the error, which may reduce the power of the corresponding cointegration test. By contrast, cointegration tests based on t_{ECMk} and t_{ECMu} do not suffer from this problem because their corresponding regressions are still properly specified.

2.3.2 Cointegrated Processes

Asymptotic distributions of all four cointegration test statistics are known for Case I (z_t with a unit root and no break), providing a baseline for the Monte Carlo study. Distributions under Cases II–IV are viewed as variants, with Case IV being the particular focus of Section 5. Because (1)–(4) has a unit root under Case I, distributional results involve Wiener processes. That said, t_{ECMk} is approximately normally distributed for large q (when $a \neq 1$). Derivation of the asymptotic distributions appear in Dickey and Fuller (1979, 1981) and Phillips (1987b, 1988) (for t_{DFk}); Banerjee, Dolado, Hendry, and Smith (1986) and Kremers, Ericsson, and Dolado (1992) (for t_{ECMk}); Engle and Granger (1987) and

Phillips and Ouliaris (1990) (for t_{DFu}); and Park and Phillips (1988, 1989), Kiviet and Phillips (1992), Banerjee and Hendry (1992), Boswijk (1992), Banerjee, Dolado, Galbraith, and Hendry (1993), and this subsection (for t_{ECMu}).

For expositional convenience, we adopt certain notational conventions concerning Brownian motion (or Wiener) processes. Consider a normal, independently and identically distributed variable $\eta_t, t = 1, ..., T$: that is, $\eta_t \sim IN(0, \sigma\eta^2)$. Here, η_t is usually either e_t, ϵ_t, or u_t. Define $B_{T,\eta}(r)$ as the partial sum $\sum_1^{[Tr]} \eta_t/(T\sigma\eta^2)^{1/2}$, where r lies in $[0,1]$, and $[Tr]$ is the integer part of Tr. As discussed in Phillips (1987b), $B_{T,\eta}(r)$ converges weakly to a standardized Wiener process, denoted $B\eta(r)$. For simplicity of notation, the argument r is suppressed, as is the range of integration over r when that range is $[0,1]$. Thus, integrals such as $\int_0^1 B\eta(r)^2 dr$ are written as $\int B\eta^2$. The symbol " $\Rightarrow$ " denotes weak convergence of the associated probability measures as the sample size $T \to \infty$. See Billingsley (1968) and Banerjee, Dolado, Galbraith, and Hendry (1993) for further discussion. Derivations are presented below for cases without an intercept in the regression, for simplicity. When a constant term is included, the Brownian motions should be thought of as being "de-meaned", corresponding to a prior regression that eliminates a constant. Mann and Wald's (1943) order notation is used where needed.

The null hypothesis is *no* cointegration ($b = 0$). The alternative hypothesis is cointegration ($b < 0$), and is characterized as a local alternative with:

$$b \quad = \quad e^{\gamma/T} - 1 \quad \approx \quad \gamma/T, \tag{17}$$

where γ is a negative fixed scalar. Equation (17) parallels the usual Pitman-type local alternative except that, in order to obtain statistics of $O_p(1)$, b differs from the null by $O_p(T^{-1})$ rather than by $O_p(T^{-1/2})$. Conveniently, distributions under the null hypothesis are obtained by setting $\gamma = 0$. The generalization of $B\eta(r)$ under this local alternative is the diffusion process:

$$\begin{aligned} K\eta(r) &= \int_0^r e^{(r-j)\gamma} dB\eta(j) \\ &= B\eta(r) + \gamma \int_0^r e^{(r-j)\gamma} B\eta(j) dj, \end{aligned} \tag{18}$$

where $K\eta(r)$ is an implicit function of γ; see Phillips (1987b). If $\gamma = 0$, then $K\eta(r) = B\eta(r)$. As with $B\eta$, the argument r in $K\eta(r)$ and the limits of integration are dropped if no ambiguity arises from doing so.

Under the local alternative, t_{DFk} is distributed as:

$$t_{DFk} \quad \Rightarrow \quad \gamma(\textstyle\int K_e^2)^{1/2} \; + \; \dfrac{\int K_e dB_e}{\sqrt{\int K_e^2}}, \tag{19}$$

which simplifies to the Dickey-Fuller distribution:

$$t_{DFk} \quad \Rightarrow \quad \frac{\int B_e dB_e}{\sqrt{\int B_e^2}} \tag{20}$$

under the null hypothesis.

Under the local alternative, the ECM statistic t_{ECMk} is distributed as:

$$\begin{aligned} t_{ECMk} \quad \Rightarrow \quad & \gamma(1+q^2)^{1/2}(\textstyle\int K_e^2)^{1/2} \\ & + \frac{(a-1)\int K_u dB\epsilon + s^{-1}\int K\epsilon dB\epsilon}{\sqrt{(a-1)^2 \int K_u^2 + 2(a-1)s^{-1}\int K_u K\epsilon + s^{-2}\int K\epsilon^2}}. \end{aligned} \tag{21}$$

When $a = 1$, (21) simplifies to the DF distributions (19) (for $\gamma \neq 0$) and (20) (for $\gamma = 0$).

For $a \neq 1$, (21) can be reparameterized in terms of γ and q exclusively:

$$t_{ECMk} \quad \Rightarrow \quad \gamma(1+q^2)^{1/2}(\int K_e^2)^{1/2} \;+\; \frac{\int K_u dB\epsilon - q^{-1}\int K\epsilon dB\epsilon}{\sqrt{\int K_u^2 - 2q^{-1}\int K_u K\epsilon + q^{-2}\int K\epsilon^2}}. \tag{22}$$

For large q, (22) is approximately a standardized normal distribution:

$$t_{ECMk} \quad \Rightarrow \quad N\left(\gamma(1+q^2)^{1/2}(\int K_u^2)^{1/2}, 1\right) \;+\; O_p(q^{-1}), \tag{23}$$

conditional on the process for u_t. Under the null hypothesis, (23) simplifies to:

$$t_{ECMk} \quad \Rightarrow \quad N(0,1) + O_p(q^{-1}). \tag{24}$$

The approximation in q is "small-σ" in nature; cf. Kadane (1970, 1971). Thus, as q varies from small to large, the asymptotic distribution of t_{ECMk} shifts from the DF distribution to the normal distribution.

The asymptotic powers of t_{DFk} and t_{ECMk} are determined by (19) and (22). When $q = 0$, the two tests have the same power. When q is sufficiently large, t_{ECMk} has (arbitrarily) greater power than t_{DFk}. That discrepancy arises because t_{DFk} is ignoring substantial information on Δz_t, whereas t_{ECMk} uses that information to obtain a more precise estimate of b.

The asymptotic distributions of t_{DFu} and t_{ECMu} resemble those of t_{DFk} and t_{ECMk}, but are somewhat more complicated because the hypothesized cointegrating vector is estimated. Phillips and Ouliaris (1990) derive the asymptotic distribution of t_{DFu} under the null hypothesis that $b = 0$.

The test of $b = 0$ using t_{ECMu} is known to be similar when the conditioning variable z_t is strongly exogenous for the regression parameters $(a, b, c, \sigma\epsilon^2)$; see Kiviet and Phillips (1992), Banerjee and Hendry (1992), and Banerjee, Dolado, Galbraith, and Hendry (1993). The null limiting distribution of t_{ECMu} is implicit as a special case of Park and Phillips (1988, Theorem 4.1a; 1989, Theorem 4.1b) and Boswijk (1992, Theorem 4.2, p. 85). The explicit representation provides insights not apparent from their general formulae, so it is derived here.

The DGP is (1)–(4) with $b = 0$ under Case I (i.e., with $\rho = 1$ and $\delta = 0$):

$$\begin{array}{rcll} \Delta y_t &=& a\Delta z_t + \epsilon_t & \qquad \epsilon_t \sim IN\left(0, \sigma\epsilon^2\right) \\ \Delta z_t &=& u_t & \qquad u_t \sim IN\left(0, \sigma_u^2\right). \end{array} \tag{25}$$

The estimated model is:

$$\Delta y_t \quad = \quad a\Delta z_t + b(y-z)_{t-1} + cz_{t-1} + \nu_t, \tag{26}$$

which is (6) without a constant term.

The rescaled parameter estimates from (26) are:

$$\begin{bmatrix} \sqrt{T}\,(\hat{a}-a) \\ T\hat{b} \\ T\hat{c} \end{bmatrix} = \begin{bmatrix} T^{-1}\sum_1^T u_t^2 & T^{-3/2}\sum_1^T w_{t-1}u_t & T^{-3/2}\sum_1^T z_{t-1}u_t \\ T^{-3/2}\sum_1^T w_{t-1}u_t & T^{-2}\sum_1^T w_{t-1}^2 & T^{-2}\sum_1^T w_{t-1}z_{t-1} \\ T^{-3/2}\sum_1^T z_{t-1}u_t & T^{-2}\sum_1^T w_{t-1}z_{t-1} & T^{-2}\sum_1^T z_{t-1}^2 \end{bmatrix}^{-1} \cdot \begin{bmatrix} T^{-1/2}\sum_1^T u_t\epsilon_t \\ T^{-1}\sum_1^T w_{t-1}\epsilon_t \\ T^{-1}\sum_1^T z_{t-1}\epsilon_t \end{bmatrix} \tag{27}$$

where, from (25), w_t is the random walk:

$$\begin{aligned} w_t &= y_t - z_t \\ &= w_{t-1} + e_t, \end{aligned} \quad (28)$$

and $e_t = (a-1)u_t + \epsilon_t$, as in (10). Asymptotic distributions of the elements in (27) involve the Brownian motion processes B_u, $B\epsilon$, and B_e, where the last is related to the first two by:

$$\sigma_e B_e = \sigma\epsilon B\epsilon + (a-1)\sigma_u B_u. \quad (29)$$

Also, $\sigma_e^2 = \sigma\epsilon^2 + (a-1)^2\sigma_u^2$. Hence, by partitioned inversion in (27):

$$T\hat{b} \Rightarrow \frac{\sigma\epsilon}{\sigma_e} \cdot \frac{(\int B_u^2)(\int B_e dB\epsilon) - (\int B_u B_e)(\int B_u dB\epsilon)}{(\int B_u^2)(\int B_e^2) - (\int B_u B_e)^2}. \quad (30)$$

Thus, the limiting distribution of t_{ECMu} is:

$$\begin{aligned} t_{ECMu} &\Rightarrow \frac{\int B_e dB\epsilon - (\int B_e B_u)(\int B_u^2)^{-1}(\int B_u dB\epsilon)}{\sqrt{\int B_e^2 - (\int B_u B_e)^2 (\int B_u^2)^{-1}}} \\ &= \frac{\int B\epsilon dB\epsilon - (\int B\epsilon B_u)(\int B_u^2)^{-1}(\int B_u dB\epsilon)}{\sqrt{\int B\epsilon^2 - (\int B\epsilon B_u)^2 (\int B_u^2)^{-1}}}, \end{aligned} \quad (31)$$

using (29). Because the asymptotic distribution of t_{ECMu} does not depend on a, σ_u, or $\sigma\epsilon$, it is invariant to those parameters, and hence tests based on t_{ECMu} are similar.

An alternative expression for (31) is enlightening. Consider the regression

$$y_t = \beta z_t + \zeta_t \quad (32)$$

"linking" the levels of y_t and z_t. Let R denote the limiting form of Yule's correlation for that regression when $a = b = 0$ (i.e., under the null of no relation between y_t and z_t) and that correlation is not adjusted for sample means. From Phillips (1986), R is:

$$\mathsf{R} = \frac{\int B\epsilon B_u}{\sqrt{(\int B_u^2)(\int B\epsilon^2)}}. \quad (33)$$

Thus, schematically (31) is:

$$t_{ECMu} \Rightarrow \frac{DF - \mathsf{R} \cdot N(0,1)}{\sqrt{1 - \mathsf{R}^2}}, \quad (34)$$

where DF and $N(0,1)$ denote random variables with Dickey-Fuller and standardized normal distributions. This demonstrates that t_{ECMu} and the Dickey-Fuller statistic have different limiting distributions, so their critical values need separate tabulation. Also, (34) mirrors Park and Phillips's (1988) Lemma 5.6, in which a related problem is addressed and R is a constant.

3 Experimental Design and Simulation

This section presents the experimental design and Monte Carlo simulation of the cointegration test statistics.

3.1 Experimental Design

To analyze the size and power of the cointegration tests in the presence of a structural break, two sets of Monte Carlo experiments were conducted with (1)–(4) as the DGP. The first set is a "broad" design, aimed at highlighting the effects of common factors, cointegration, and breaks over a range of sample sizes for estimation. The second set focuses on how the lack of a common factor affects the cointegration tests. Without loss of generality, $\sigma\epsilon^2 = 1$ and $\lambda = 1$. Thus, the experimental design variables are the parameters (a, b, s, ρ, δ) (noting that s now is σ_u), the estimation sample size T_e, the break points T_0 and T_1, and the full sample size T.

The first set of experiments is a full factorial design of:

$$\begin{aligned} a &= (1.0 \text{ [a common factor: } q = 0],\ 0.0 \text{ [no common factor: } q = s]) \\ b &= (0.0 \text{ [no cointegration]},\ -0.1 \text{ [cointegration]}) \\ s &= 1.0 \\ \rho &= (1.0 \text{ [integration]},\ 0.8 \text{ [stationarity]}) \\ \delta &= (0.0 \text{ [no break]},\ 1.0 \text{ [a break of size } s]) \\ T_e &= 10, 11, 12, \ldots, T-2, T-1, T \\ T_0 &= 25 \\ T_1 &= 75 \\ T &= 100, \end{aligned} \qquad (35)$$

resulting in 1456 experiments. For both sets of experiments, new z's were generated for each replication, and the number of replications per experiment was $P = 10,000$.

The parameter values were chosen with the following in mind. For $a = 1$ (and so $q = 0$), the common factor restriction holds, so the distributions of the DF and ECM statistics should resemble each other. For $a = 0$, the common factor restriction is violated, but $q = s = 1$, which is a "moderately small" value. The two values of b, 0.0 and -0.1, imply lack of and existence of cointegration respectively, although, in the latter case, the corresponding root of the system is still large: 0.9. The root of the marginal process (ρ) is either unity or large but stationary. The break in the marginal process is either zero or unity (i.e., $1 \cdot \sigma_u$), where the latter value is rather small by empirical standards. However, for the purposes of this paper, unity seemed appropriate because it is small enough to make its detection by standard Chow tests difficult.

The estimation sample size includes all integers in the interval $[10, 100]$, providing small, medium, and large values. To ensure that most estimation sample sizes included a break, $T_0 = 25$, with $T_1 = 75$ so as to maximize the power of constancy tests over the full sample; see Hendry and Neale (1991). For a given length of break $(T_1 - T_0)/T$, the particular choice of T_0 and T_1 matters little for the power of the full-sample constancy tests.

While the generated data have two breaks, recursive testing allows examining estimation samples with zero breaks ($T_e < T_0$), one "permanent" break ($T_0 \le T_e < T_1$), and two breaks ($T_1 \le T_e \le T$). In this design, two breaks are equivalent to a single "temporary" break. While the number of observations generated (T) and the estimation period (T_e) are identical in the analytical derivations above, distinguishing between T and T_e is necessary in recursive Monte Carlo. Below, the context clarifies whether "sample size" means T or T_e. Recursive testing also allows examining situations in which the break is at the very end of the estimation sample.

Critical values are all at the 5% level. For t_{DFk} and t_{DFu}, the values are calculated from MacKinnon's (1991, Table 1) response surfaces with $N = 1$ and $N = 2$ respectively (MacKinnon's "N"), "with a constant but no trend" for both. The correct critical values for t_{ECMk} depend upon q as well as T_e, and those for t_{ECMu} depend upon T_e. Both sets of critical values could be simulated. However, q may not be known in practice, and even asymptotic critical values for t_{ECMu} have not yet been simulated, so "safe" critical values may be constructed by assuming that t_{ECMk} and t_{ECMu} have pure Dickey-Fuller-type distributions (i.e., $q = 0$). Thus, the critical values for t_{ECMk} and t_{ECMu} used here are the same as those for t_{DFk} and t_{DFu}.

Because q is such an important parameter and because $q \leq 1$ in (35) is relatively small by empirical standards, the second set of experiments considers the effect of $q = 3$, albeit in a more limited design:

$$\begin{aligned}
a &= 0.0 \text{ [no common factor: } q = s\text{]} \\
b &= (0.0 \text{ [no cointegration]}, \ -0.1 \text{ [cointegration]}) \\
s &= 3.0 \\
\rho &= 0.8 \text{ [stationarity]} \\
\delta &= 3.0 \text{ [a break of size } s\text{]} \\
T_e &= 10, 11, 12, \ldots, T-2, T-1, T \\
T_0 &= 25 \\
T_1 &= 75 \\
T &= 100, \qquad (36)
\end{aligned}$$

resulting in 182 experiments. The values of s and δ are equal in order to keep the time series properties of z the same as in (35).

3.2 Simulation

Simulation proceeded as follows. Random numbers for ϵ_t and u_t were generated by multiplicative congruential generators and transformed to a normal distribution by Box and Muller's (1958) method. The first twenty observations of each replication from (1)–(4) were discarded in order to attenuate the effects of initial values in stationary relations (such as in $y_t - z_t$ when $b < 0$). For a particular experiment, P replications were generated, with a statistic lying in its critical region S of P times (S dependent upon the statistic). The fraction of rejections S/P is an unbiased Monte Carlo estimate of the underlying rejection frequency (e.g., of size or power).

Recursive algorithms exist for the statistics t_{DFk}, t_{ECMk}, and t_{ECMu}, providing a computationally efficient means for their calculation over the full range of sample sizes $T_e = 10, 11, \ldots, 99, 100$ for any given set of values of the other experimental design variables.[4] Thus, a replication of size $T = 100$ was generated, and the statistics were calculated recursively on that sample for all estimation sample sizes. Such re-use of the sample greatly reduces Monte Carlo variation for different values of T_e; see Hammersley and Handscomb (1964). Further, calculation of all test statistics on the same sample reduces Monte Carlo variation for the differences in properties across statistics. Graphical (rather than tabular) analysis of the Monte Carlo rejection frequencies is highly desirable, given the large number of experiments; cf. Ericsson (1991). Graphical analysis also corresponds to a (pseudo-) nonparametric estimation of the size and power functions of the tests.

4 Post-simulation Analysis: the Marginal Process

This section briefly examines one unit-root test statistic and two constancy test statistics on the marginal process for z_t. The unit root statistic is the Dickey-Fuller statistic with a constant term (t_{DF}), which is the t-ratio on ϕ in (14), noting that the break dummy D_t is explicitly excluded. The first constancy statistic is Chow's (1960) predictive failure statistic applied to (14) and is denoted F_{CHOW}. The second

[4] The statistic t_{DFu} cannot be calculated recursively, so its properties are considered for $T_e = 100$ only. Also, D_t was perturbed by a small error in order to permit recursive estimation for $T_e \leq 25$ of equations including D_t.

constancy statistic is the t-ratio on the least squares estimate of δ in the correctly specified marginal process (albeit with a constant term estimated):

$$\Delta z_t = \mu + \phi z_{t-1} + \delta D_t + u_t, \tag{37}$$

and is denoted t_δ. This last statistic is also a statistic for testing the invariance of (14) to D_t, as discussed in Engle and Hendry (1993). Critical values are at the 5% level and are taken from MacKinnon (1991) for t_{DF}, the F distribution for F_{CHOW}, and the t distribution for t_δ.

Figures 3, 4, and 5 plot the estimated rejection frequencies of t_{DF}, F_{CHOW}, and t_δ respectively. The four lines on each graph correspond to Case I ($\rho = 1,\ \delta = 0$: —), Case II ($\rho = 0.8,\ \delta = 0$: – –), Case III ($\rho = 1,\ \delta = 1$: $\cdots\cdots$), and Case IV ($\rho = 0.8,\ \delta = 1$: - - -).

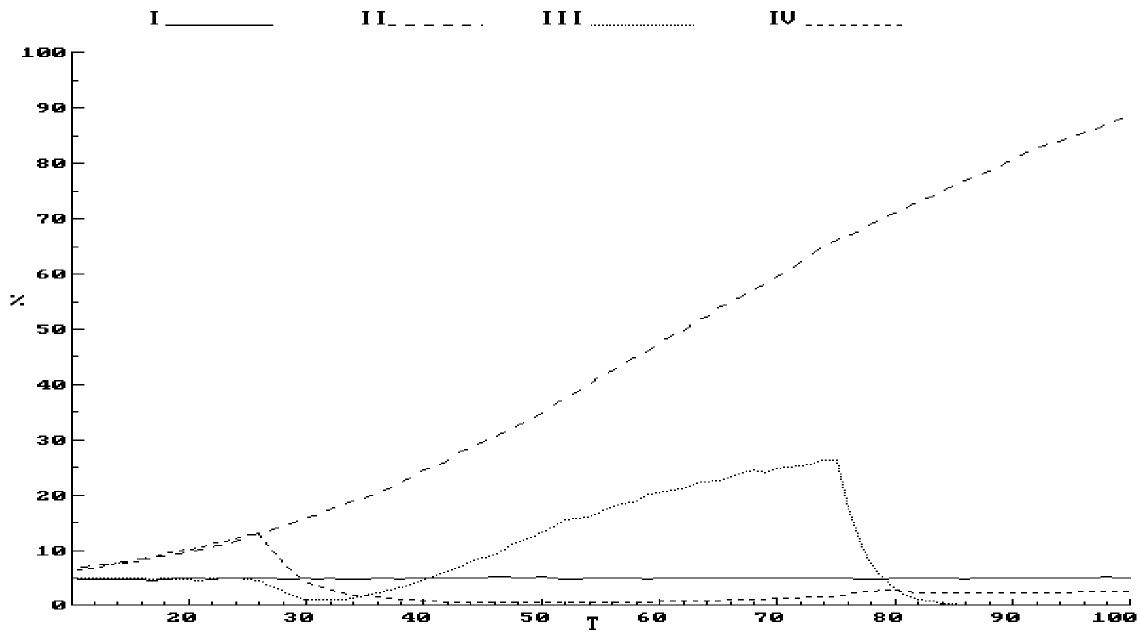

Figure 3: Rejection frequencies of t_{DF}

From Figure 3, the estimated size of t_{DF} (Case I) is close to 5% for all sample sizes. When $\rho = 0.8$ without a break (Case II), the power increases monotonically from about 7% (at $T_e = 10$) to 90% (at $T_e = 100$). When a break is added to that stationary series (Case IV), the power falls from 13% at $T_e = 25$ to less than 1% by $T_e = 40$, increasing to only around $2\frac{1}{2}$% by $T_e = 100$. As examined in greater detail by Hendry and Neale (1991), even small breaks can dramatically reduce the power of unit root tests. The size is also affected by breaks, noting that the rejection frequency for Case III varies between 0% and 25%, depending upon what fraction of the sample includes the break.

From Figure 4, the estimated size of F_{CHOW} (Cases I and II) is about 7–9% for small samples, tending to its nominal 5% value by $T_e = 100$. For stationary z_t with a break, the power is never higher than 12%, even for a sample split at $T_e = 25$ where the first break occurs. The power is somewhat higher for non-stationary z_t, but still never exceeds 30%. In essence, a break of one standard deviation over half the sample is small and hard to detect, in spite of its consequences on the unit root test.

The statistic t_δ should provide a highly powerful test, given that the dates and nature of the break are treated as known. However, controlling its size is problematic, as Figure 5 documents. Intuitively, D_t behaves like an integrated process and so the distribution of t_δ is affected by the correlation between z_{t-1}

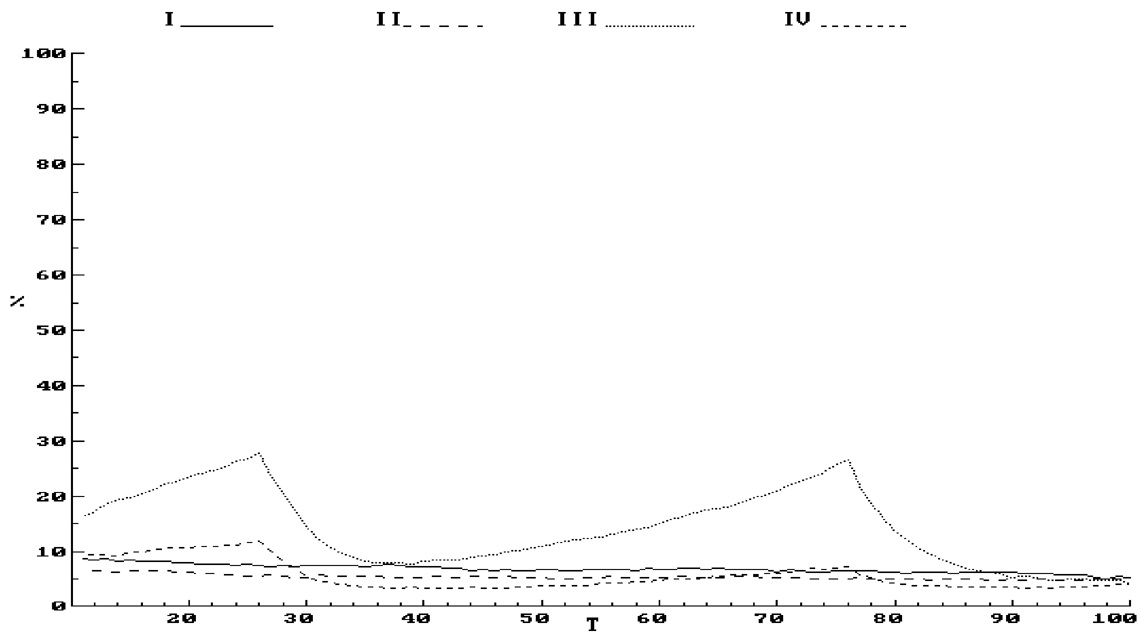

Figure 4: Rejection frequencies of F_{CHOW}

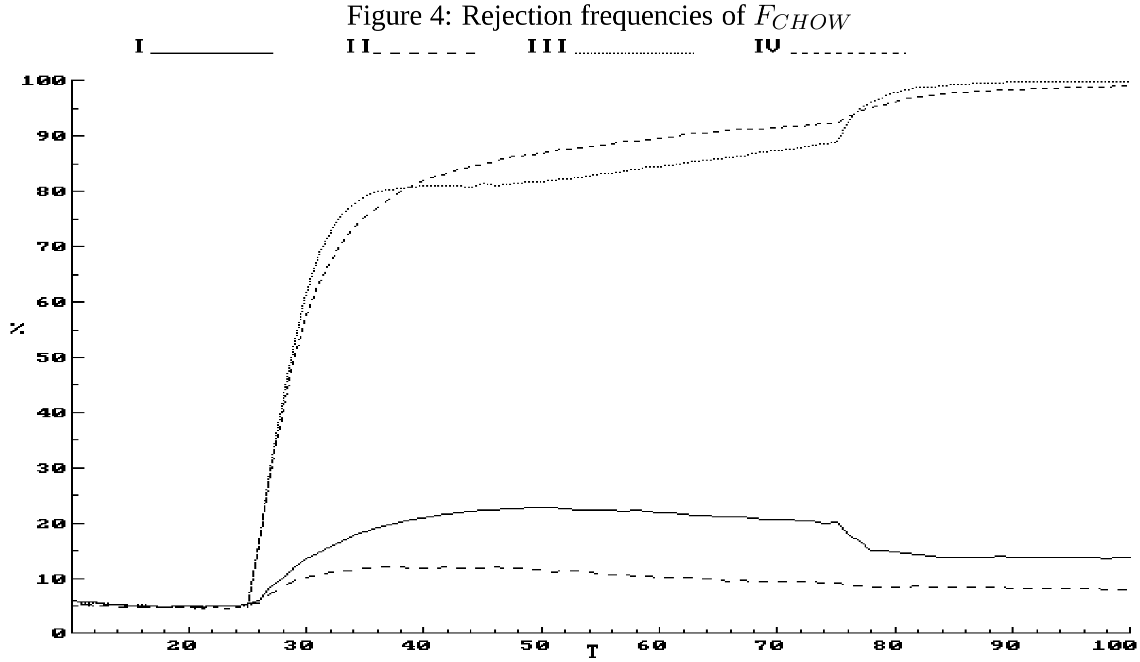

Figure 5: Rejection frequencies of t_δ

and D_t in (37). This effect is lessened but not eliminated when z_t has a stationary root.[5] The "power" of t_δ appears impressive, but must be treated with caution, given the large distortion to size.

To summarize, the break in the marginal process is difficult to detect with the Chow statistic, yet it dramatically reduces the power of the Dickey-Fuller statistic for detecting a stationary root. A stationary process with a break is virtually observationally equivalent to a unit root process with no break.[6]

[5] Note, however, that if the root of z_t were treated as known, the distribution of $t\delta$ would be exactly a t.

[6] Faust (1993) formally establishes the near observational equivalence of trend-stationary and difference-stationary processes. His framework also may help establish a parallel result here.

5 Post-simulation Analysis: Tests of Cointegration

This section examines the cointegration statistics ($t_{ECMk}, t_{DFk}, t_{ECMu}, t_{DFu}$) as defined by (5)–(8). Rejection frequencies are the primary focus, but means, standard deviations, and overall distributional properties are also considered.

Figures 6–9 and 10–13 plot rejection frequencies by the four cointegration tests for the first set of experiments. These rejection frequencies are under the hypotheses of no cointegration (Figures 6–9) and cointegration (Figure 10–13), and correspond to size and power, provided the correct critical values are used. Figures 6–9 plot estimated sizes for ($a = 1,\ \delta = 0$), ($a = 1,\ \delta = 1$), ($a = 0,\ \delta = 0$), and ($a = 0,\ \delta = 1$) respectively: that is, for DGPs with and without a common factor in the conditional process and with and without a break in the marginal process. Figures 10–13 present the corresponding plots for estimated powers. The primary interest here is in discerning the differences between Cases I and IV, so $\rho = 1$ when $\delta = 0$ and $\rho = 0.8$ when $\delta = 1$.

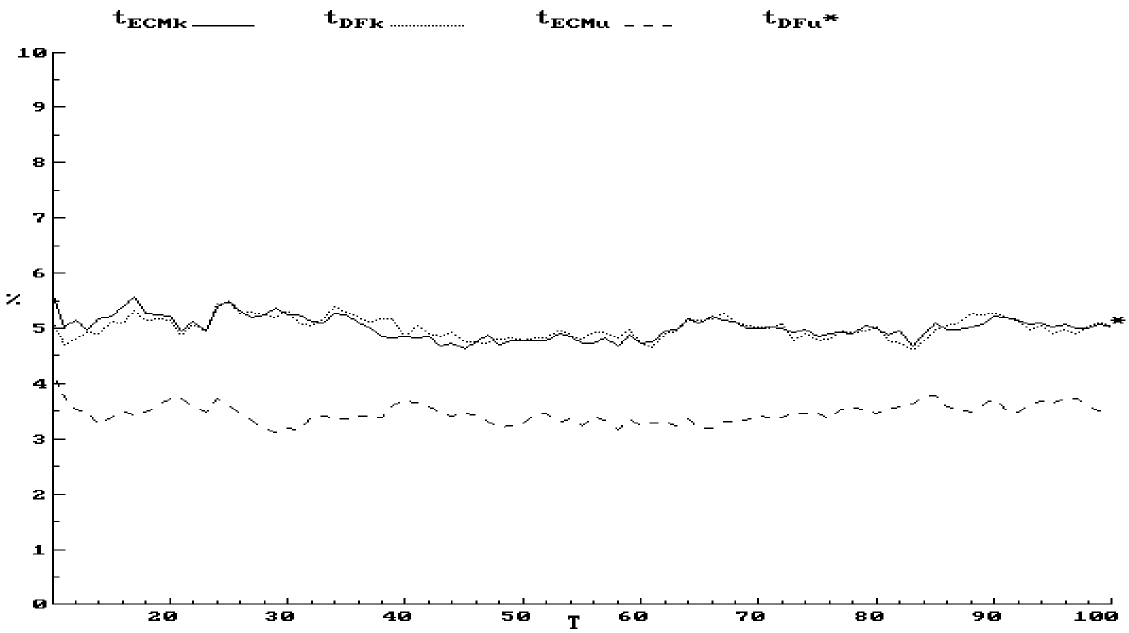

Figure 6: Null rejection frequencies $a = 1,\ \delta = 0$

From Figures 6 and 7, the estimated sizes for t_{ECMk} and t_{DFk} are both approximately 5%, which follows from the asymptotic equivalence of the two statistics when there is a common factor ($a = 1$). The size of t_{ECMu} is around 3% because of the conservative choice of using MacKinnon's critical values. All three sizes are virtually unaffected by the sample size T_e, confirming the accuracy of MacKinnon's response surfaces for the critical values. The estimated size of t_{DFu} is available at only $T_e = 100$, and is approximately 5%.

Invalidity of the common factor restriction (Figures 8 and 9) clearly affects t_{ECMk} and t_{DFk}. As anticipated from the asymptotics with no break, the rejection frequencies for t_{ECMk} are below 5% (typically, between 2% and 4%), while those for t_{ECMu} are unchanged (at 3%) from simulations with a common factor. The rejection frequency of t_{DFk} is about 5% in Figure 8, in line with its invariance to the existence or lack of a common factor when there is no break; cf. Kremers, Ericsson, and Dolado (1992). However, its rejection frequency is not invariant to the lack of a common factor when there is a break. The residual in the estimated equation for t_{DFk} involves Δz_t, which includes D_t and a stationary error: see (9). The average size for t_{DFk} is about $6\frac{1}{2}\%$, which is substantially higher than the sizes for

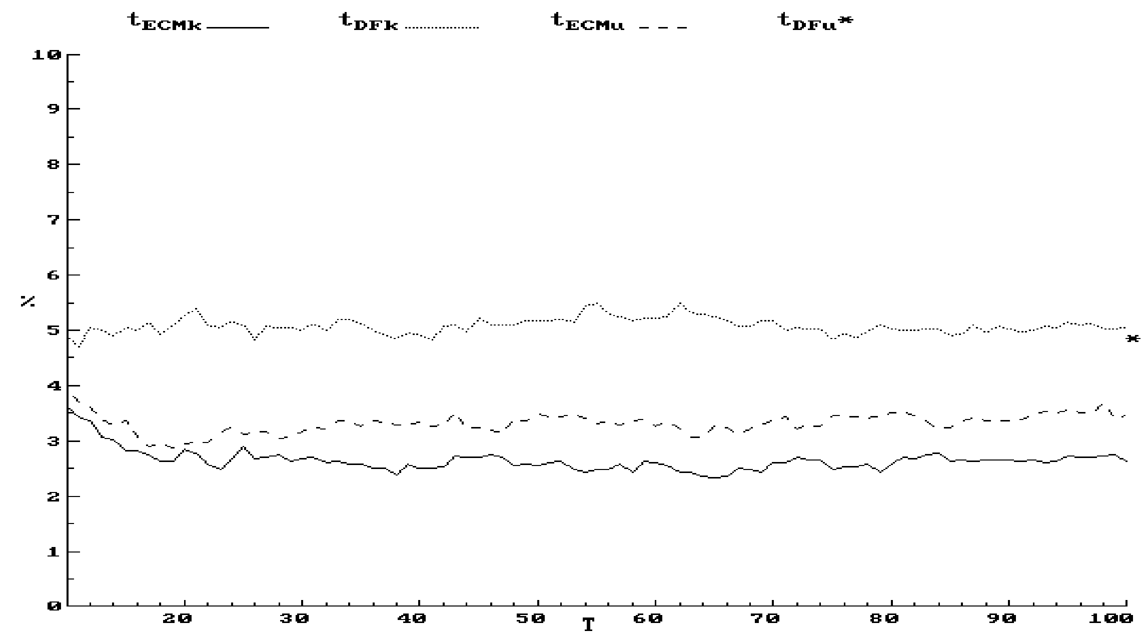

Figure 7: Null rejection frequencies $a = 1,\ \delta = 1$

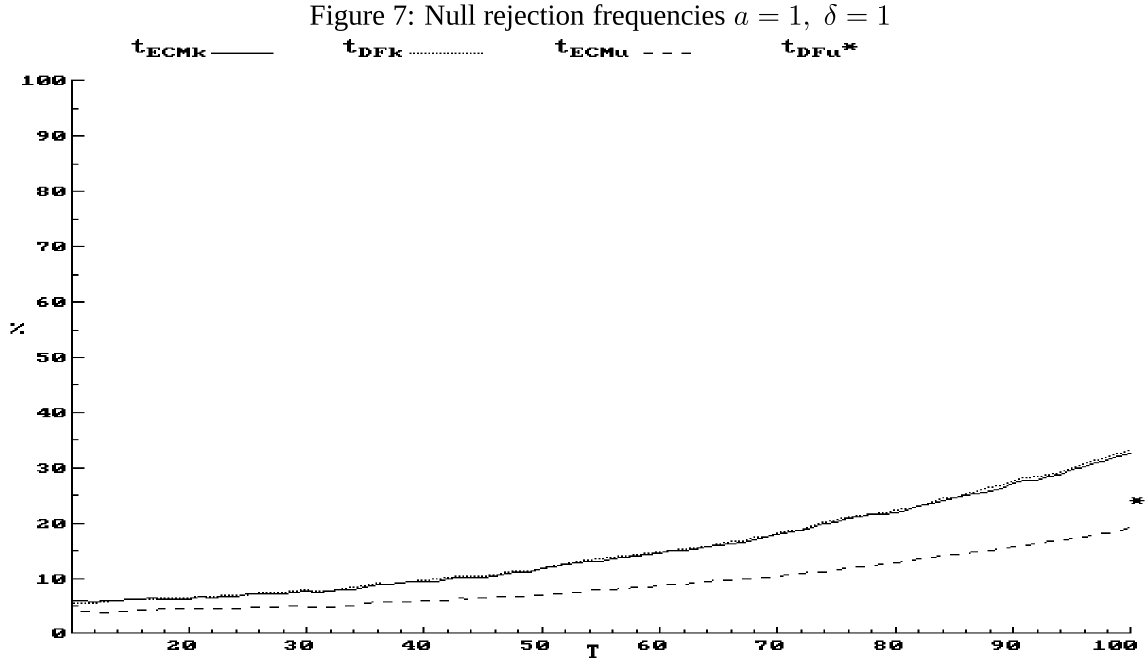

Figure 8: Null rejection frequencies $a = 0,\ \delta = 1$

t_{ECMk} and t_{ECMu}.

Clearly, both the DF and ECM tests are less than perfect. However, noting that the actual size is defined as the maximum rejection frequency over the parameter space for the null hypothesis, the ECM tests at least control for size whereas the DF test does not. Also, because t_{ECMu} appears invariant to the break and is invariant to a, better critical values for it are easily calculated by Monte Carlo. When there is no common factor, t_{DFk} is not invariant to the break (see Figure 9), so useful critical values for t_{DFk} are problematic to obtain. That lack of invariance is even more apparent for powers and for larger q, as seen below.

In Figures 10 and 11, the DGP has a common factor, and y_t and z_t are cointegrated. As under the null

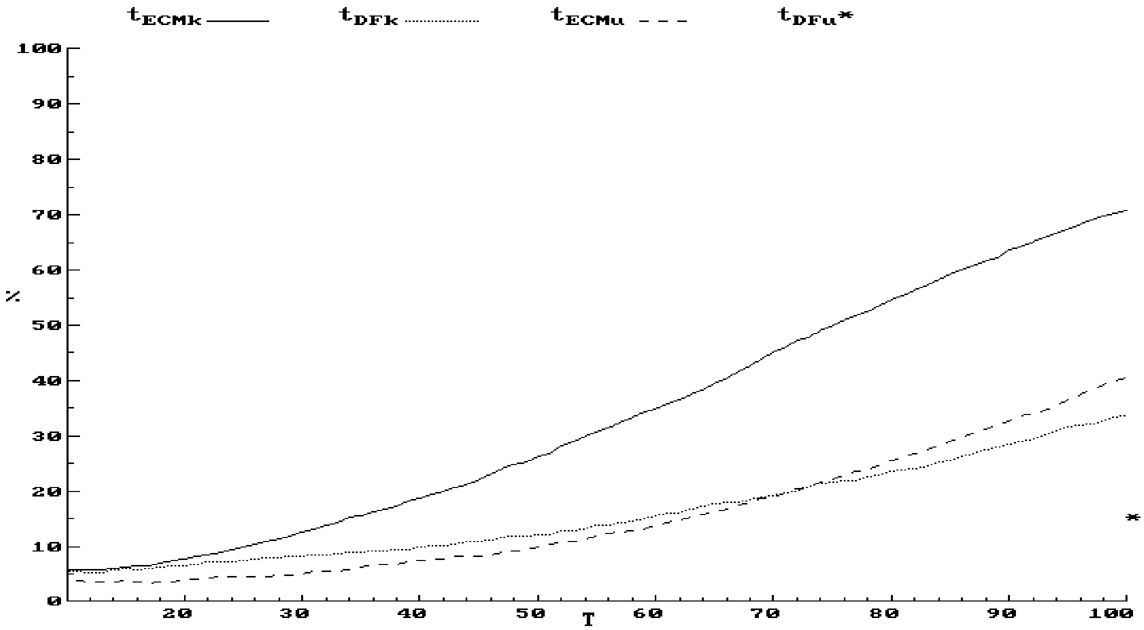

Figure 9: Null rejection frequencies $a = 0,\ \delta = 0$

of no cointegration, the presence or lack of a break has no effect on the test statistics; and the estimated powers for t_{DFk} and t_{ECMk} are virtually identical, ranging from 5% at $T_e = 10$ to 33% at $T_e = 100$. The rejection frequency of t_{ECMu} is somewhat less, and unsurprisingly so because its rejection frequency under the null is less than 5% and because it ignores λ being unity.

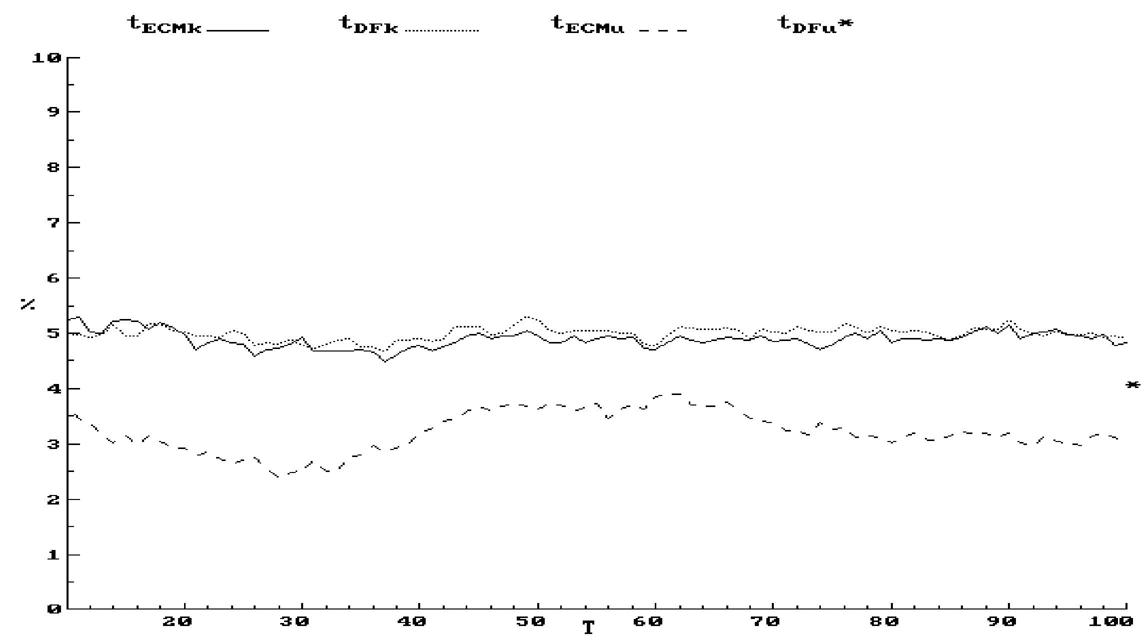

Figure 10: Non-null rejection frequencies $a = 1,\ \delta = 0$

When there is no common factor and no break (Figure 12), the power of t_{ECMk} substantially dominates that of t_{DFk}, with the former increasing to 70% by $T_e = 100$. Even t_{ECMu} does better than t_{DFk} at moderate to large samples, in spite of ignoring the value of the cointegrating vector. In fact, the power of t_{DFk} is invariant across Figures 10, 11, and 12. By contrast, the powers of t_{ECMk} and t_{ECMu}

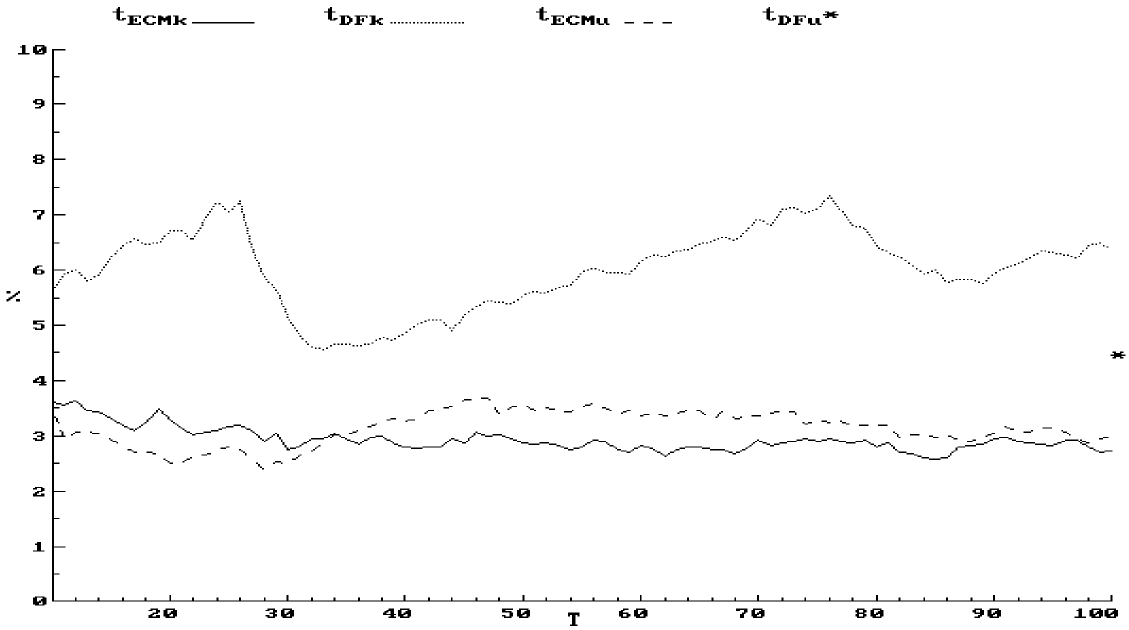

Figure 11: Non-null rejection frequencies $a = 1,\ \delta = 1$

increase when the common factor is invalid, as predicted by theory.

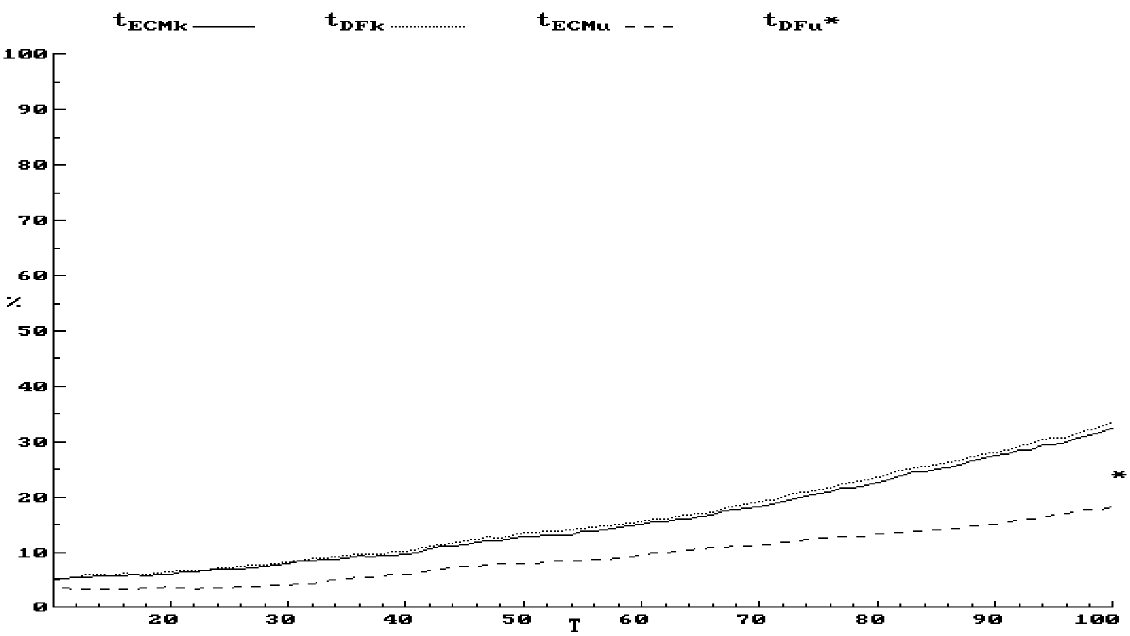

Figure 12: Non-null rejection frequencies $a = 0,\ \delta = 0$

With a break but no common factor (Figure 13), the power of t_{ECMk} exceeds that of t_{DFk} except for very small samples, where both powers appear approximately equal to size. Discrepancies between their powers at larger sample sizes appear smaller than without a break, but this is probably a spurious result due to inadequate control of the rejection frequency of t_{DFk} under the null hypothesis (see Figure 9).

Figures 14 and 15 plot estimated sizes and powers for the second set of experiments. The DGP has no common factor and a large break ($\delta = 3$), so these figures are qualitatively similar to Figures 9 and 13, but effects of the break are more pronounced from having a larger q. The rejection frequency

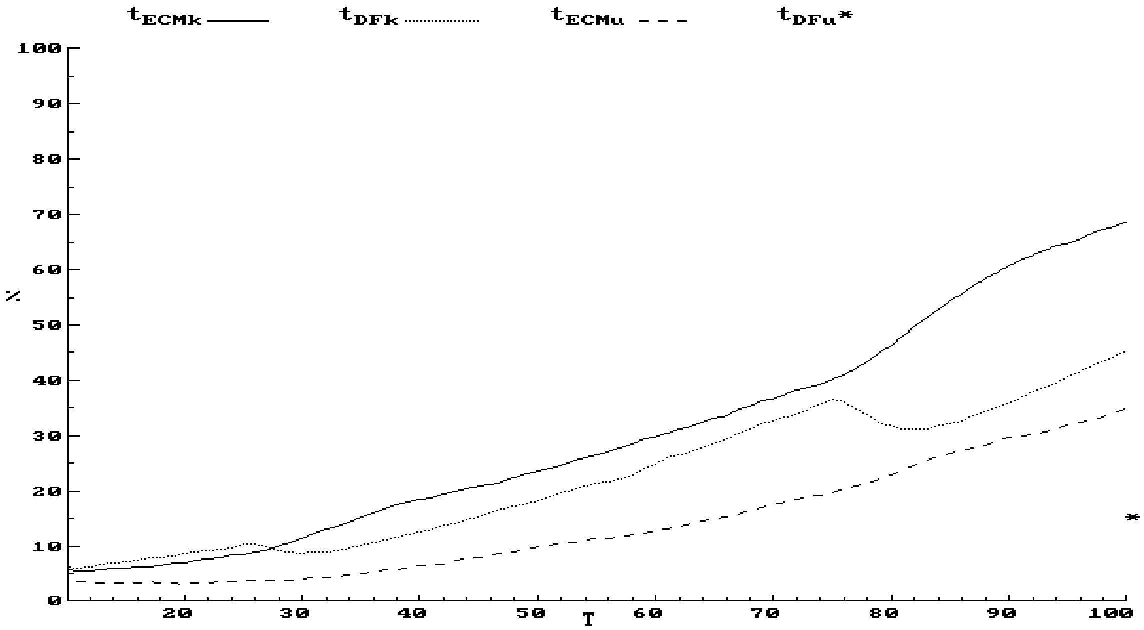

Figure 13: Non-null rejection frequencies $a = 0,\ \delta = 1$

of t_{ECMk} in Figure 14 is even smaller than in Figure 9, as predicted by the asymptotic distribution of t_{ECMk} shifting towards a normal distribution as q increases. Rejection frequencies for t_{DFk} resemble those in Figure 9, but with a larger range, 2–11%. The distribution of t_{ECMu} still appears invariant to the break.

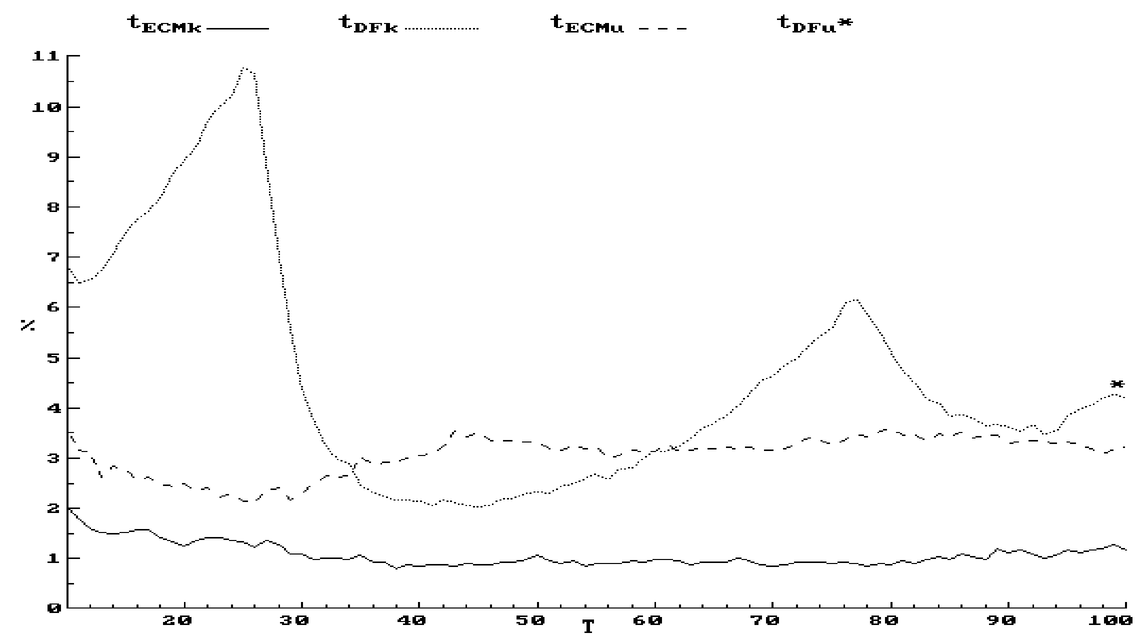

Figure 14: Null rejection frequencies, second set

Because of the greater information content in Δz_t, the powers for t_{ECMk} and t_{ECMu} in Figure 15 increase more rapidly with T_e than their powers in Figure 13 do. The "power" of t_{DFk} has more pronounced dips after the breaks occur than in Figure 13, and is somewhat inflated because its rejection frequency under the null hypothesis is inadequately controlled. Even so, the power of t_{ECMk} dominates

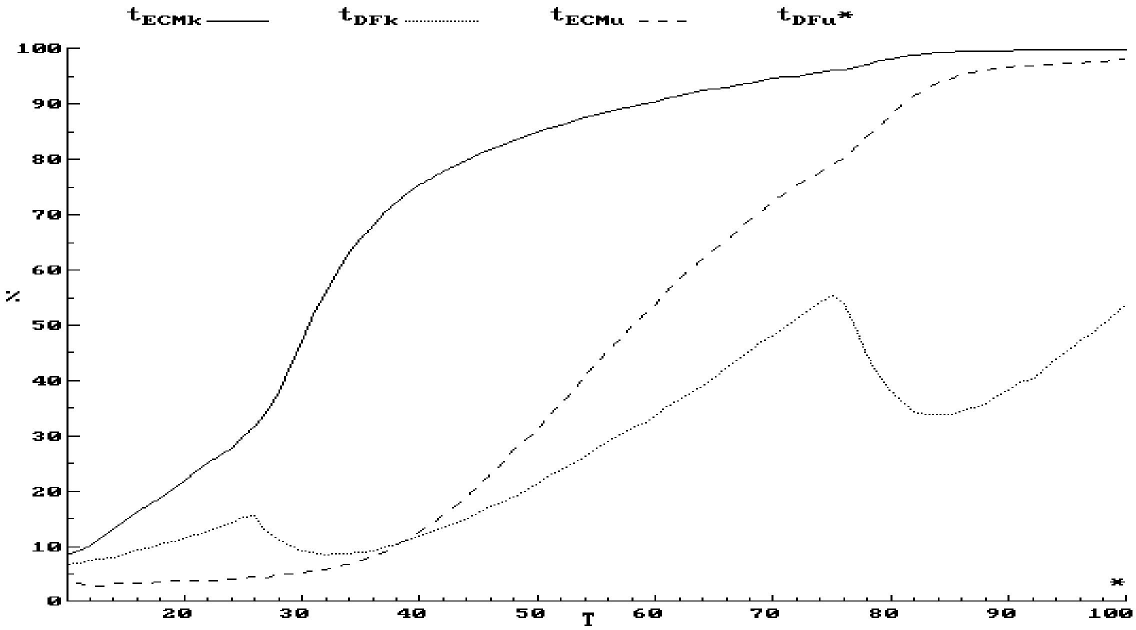

Figure 15: Non-null rejection frequencies, second set

that of t_{DFk} for all sample sizes, as does the power of t_{ECMu} for $T_e > 40$. By contrast, the power of t_{DFu} is less than its size at $T_e = 100$. Figures 13 and 15 also show how a test's power may either increase (t_{ECMk} and t_{ECMu}) or decrease (t_{DFk}) when the estimation sample is lengthened to include the first observation of a break.

Table 1.
Estimated Rejection Frequencies for Cointegration Tests

Figure	a	δ	t_{DFk}	t_{ECMk}	t_{DFu}	t_{ECMu}
				Size		
3a	1	0	5.0	5.0	5.0	3.5
3b	1	1	5.0	4.9	4.1	3.2
3c	0	0	5.1	2.6	4.8	3.3
3d	0	1	[4.6, 7.4]	2.9	4.3	3.1
5a	0	3	[2.0, 10.8]	1.1	4.5	3.1
				Power		
4a	1	0	[5.5, 33.2]	[5.8, 32.7]	22.8	[3.9, 19.2]
4b	1	1	[5.1, 33.5]	[5.2, 32.4]	22.2	[3.2, 18.3]
4c	0	0	[5.3, 33.7]	[5.7, 70.7]	14.6	[3.4, 40.4]
4d	0	1	[6.1, 45.4]	[5.5, 68.7]	15.5	[3.2, 34.8]
5b	0	3	[6.8, 55.3]	[8.4, 99.9]	2.9	[2.8, 98.1]

Notes: Single values are means; values in square brackets indicate the range over T_e; all values are percentages. The size and power for t_{DFu} were estimated for $T_e = 100$ only.

Table 1 summarizes the results in Figures 3–15 by listing the approximate values or (where appropriate) ranges of rejection frequencies for all four tests under the various parameterizations: with and without a common factor ($a = 1$ and $a = 0$); and with no break ($\delta = 0$), a small break ($\delta = 1$), and a large break ($\delta = 3$). With a common factor, DF and ECM tests perform similarly. Without a common factor, ECM tests are preferred, having higher power and better control of size.

Campos, Ericsson, and Hendry (1993, Appendix B) examine other distributional aspects of the cointegration test statistics, graphing their histograms, their means, and the means of the corresponding estimated coefficients. With the exception of the means of the statistics themselves, there is little to distinguish the properties of the test statistics across DGPs (e.g., with or without a break) and even across test statistics for a given DGP.

For each test statistic and DGP, the recursively estimated Monte Carlo mean of the estimate of b is calculated with plus-or-minus twice its (average) estimated standard error, as would be computed by a regression package, and plus-or-minus twice the Monte Carlo standard deviation, reflecting the actual sampling distribution of the estimator of b. The Monte Carlo standard deviation is always larger than the estimated standard error for these experiments and test statistics.

As predicted by theory, the means of t_{DFk} and t_{ECMu} appear invariant to q when there is no break. The mean of t_{ECMk} shifts toward zero as q increases (under the null), and becomes more negative as q increases (under the alternative).

The histograms of all four statistics are obtained for $T_e = 100$ only. While 10,000 replications is only a moderate number of replications for examining full distributional properties of the statistics, their distributions overall appear normal, in line with Banerjee and Dolado's (1988) result that the Dickey-Fuller distribution is well approximated by a normal distribution with a negative mean.

While the Monte Carlo study in this paper is limited by a relatively small experimental design for a, s, and b, both asymptotic theory and these simulations point to the advantages of the ECM statistics for empirically common values of a and s. Control of size appears relatively straightforward for the ECM statistics in the presence of breaks. Tests with t_{ECMu} are insensitive to breaks under the null hypothesis, and MacKinnon's Dickey-Fuller critical values provide a "safe" choice for t_{ECMk} and t_{ECMu}. Further, the power of the ECM statistics commonly exceeds that of DF statistics for empirically interesting parameter values.

Recursive algorithms helped reduce Monte Carlo imprecision across statistics and across (econometric) estimation sample sizes, with graphical analysis providing a clear, simple summary of a vast array of estimated sizes and powers. Recursive procedures are also appealing empirically. Because statistics are affected by the accrual of information over time, full-sample and partial-sample inferences may differ, especially with breaks. Recursive estimation and testing offer a window on those effects.

6 Summary and Remarks

Testing for cointegration has become an important facet of the empirical analysis of economic time series. Various tests have been proposed and widely applied, but most distributional results rest on the assumption of unit root processes with no structural breaks. Even so, regime shifts and structural breaks are empirically and economically plausible, as indicated by extensive discussion of the Lucas (1976) critique. Using Monte Carlo methodology, this paper examines the finite sample properties of four common tests of cointegration in the presence of a structural break.

When conditioning is valid, Dickey-Fuller statistics used to test for cointegration have no particular advantage over their ECM counterparts; and there is much to gain from using the latter when the common factor restriction is invalid, and especially so if a break occurs as well. These differences arise because the DF statistic ignores potentially valuable information by imposing a possibly invalid common factor restriction. Because common factor restrictions are generic to univariate-based tests of cointegration, these results should hold for the augmented Dickey-Fuller statistic, Sargan and Bhargava's (1983) statistic, Phillips's (1987a) $Z\alpha$ and Z_t statistics, and generalizations thereon by Phillips and Perron (1988) and Gregory and Hansen (1992). The problem is in using these univariate-based statistics to test a multivariate hypothesis, and not in the statistics themselves.

Conversely, maximum likelihood procedures such as those developed by Johansen (1988, 1991, 1992a), Johansen and Juselius (1990), Phillips (1991), and Boswijk (1992) do not impose common factor restrictions and so can have more desirable properties. Some caveats apply in practice. First, systems procedures may require modeling the break itself, and that may be difficult. Second, while conditional modeling is often simpler than dealing with complete systems, the assumed weak exogeneity may be invalid, implying trade-offs between conditional modeling and systems modeling. Third, even in conditional models, general dynamics may not be sufficient to account for breaks. If breaks occur in the cointegrating vector itself, the Lagrange multiplier statistic of Quintos and Phillips (1992) may help detect them. With or without breaks anywhere in the DGP, properly accounting for the dynamic relationship between variables can be critical in testing for a long-run relationship between them.

Appendix: Breaks and the Distribution of the Unit Root Estimator

This Appendix derives asymptotic properties of the unit root estimator when the underlying process has a break. The asymptotic distribution of t_{DFk} with no break was solved by Dickey and Fuller (1979).

Consider the DGP for z_t in (2) under Case III:

$$\Delta z_t \quad = \quad \delta D_t + u_t \qquad u_t \sim IN(0, \sigma_u^2), \tag{38}$$

where $z_0 = 0$. Let K, L, and M be the length of the break $(T_1 - T_0)/T$, the time until the end of the break T_1/T, and the time after the break $1 - L$ respectively, all relative to the time period T. An investigator, unaware of the break, estimates

$$\Delta z_t \quad = \quad \mu + \phi z_{t-1} + \xi_t \tag{39}$$

in order to test for a unit root ($\phi = 0$). This section derives large sample properties of:

$$\begin{bmatrix} \hat{\mu} \\ \\ \hat{\phi} \end{bmatrix} = \begin{bmatrix} T & \sum_1^T z_{t-1} \\ \\ \sum_1^T z_{t-1} & \sum_1^T z_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_1^T \Delta z_t \\ \\ \sum_1^T z_{t-1}\Delta z_t \end{bmatrix}, \tag{40}$$

the least squares estimator of $(\mu : \phi)$ in (39), when (38) holds for fixed nonzero δ and K. Here and below, summations are over t unless otherwise indicated.

Evaluation of the four different summations in (40) is required. Without loss of generality, set $\sigma_u^2 = 1$, so δ is measured in standard deviations of u_t. Also, all summations utilize an explicit representation for z_t:

$$z_t \quad = \quad \sum_{j=1}^{t} \Delta z_j \quad = \quad \delta \sum_{j=1}^{t} D_j + \sum_{j=1}^{t} u_j \quad = \quad \delta I_t + h_t, \tag{41}$$

where

$$I_t = \begin{cases} 0 & \text{if } t = 1, \ldots, T_0 \\ t - T_0 & \text{if } t = T_0 + 1, \ldots, T_1 \\ TK & \text{if } t = T_1 + 1, \ldots, T \end{cases} \tag{42}$$

and

$$h_t \;=\; \sum_{j=1}^{t} u_j \qquad t = 1, \ldots, T, \tag{43}$$

with h_t being a random walk.

Evaluating (41) at $t = T$ obtains the summation $\sum_1^T \Delta z_t$:

$$\sum_1^T \Delta z_t \;=\; z_T \;=\; \underset{O_p(T)}{\delta TK} \;+\; \underset{O_p(T^{1/2})}{h_T,} \tag{44}$$

where orders of magnitude appear below the component terms in the last line. While terms such as δTK are non-stochastic, it is convenient (and legitimate) to use probabilistic orders throughout.

The summation $\sum_1^T z_{t-1}$ is $\left(\sum_1^T z_t\right) - z_T + z_0$, where $\sum_1^T z_t$ can be obtained by using (41) and evaluating the summation of I_t over the three subsamples.

$$\begin{aligned}
\sum_1^T z_t &= \delta \sum_1^T I_t + \sum_1^T h_t \\
&= 0 + \delta \sum_{T_0+1}^{T_1} I_t + \delta \sum_{T_1+1}^{T} I_t + \sum_1^T h_t \\
&= \delta \sum_{T_0+1}^{T_1} (t - T_0) + \delta \sum_{T_1+1}^{T} TK + \sum_1^T h_t \\
&= \delta \sum_{j=1}^{TK} j + \delta \sum_{j=1}^{TM} TK + \sum_1^T h_t \\
&= \tfrac{1}{2}\delta T^2 K[(K + T^{-1}) + 2M] + \sum_1^T h_t \\
&= \underset{O_p(T^2)}{\tfrac{1}{2}\delta T^2 K[K + 2M]} \;+\; \underset{O_p(T^{3/2})}{\sum_1^T h_t} \;+\; O_p(T).
\end{aligned} \tag{45}$$

Because z_T is $O_p(T)$ and $z_0 = 0$, the summation $\sum_1^T z_{t-1}$ is (45), to $O_p(T)$.

The summation $\sum_1^T z_{t-1}^2$ is obtained in a similar fashion.

$$
\begin{aligned}
\sum_1^T z_t^2 &= \delta^2 \sum_1^T I_t^2 + \sum_1^T h_t^2 + 2\delta \sum_1^T I_t h_t \\
&= \delta^2 [\sum_{T_0+1}^{T_1} I_t^2 + \sum_{T_1+1}^{T} I_t^2] + \sum_1^T h_t^2 + 2\delta[\sum_{T_0+1}^{T_1} I_t h_t + TK \sum_{T_1+1}^{T} h_t] \\
&= \delta^2 [\sum_{j=1}^{TK} j^2 + T^2K^2TM] + \sum_1^T h_t^2 + 2\delta[\sum_{j=1}^{TK} j h_{T_0+j} + TK \sum_{j=1}^{TM} h_{T_1+j}] \\
&= \underbrace{\tfrac{1}{3}\delta^2 T^3 K^2 [K+3M]}_{O_p(T^3)} \\
&\quad + \underbrace{2\delta[\sum_{j=1}^{TK} j h_{T_0+j} + TK \sum_{j=1}^{TM} h_{T_1+j}]}_{O_p(T^{5/2})} + O_p(T^2) \qquad (46)
\end{aligned}
$$

Because z_T^2 is $O_p(T^2)$, the summation $\sum_1^T z_{t-1}^2$ is (46), to $O_p(T^2)$.

Noting (38), the summation $\sum_1^T z_{t-1}\Delta z_t$ is obtained by evaluating each of the summations $\sum_1^T z_{t-1}u_t$ and $\sum_1^T z_{t-1}D_t$:

$$
\begin{aligned}
\sum_1^T z_{t-1}u_t &= \delta \sum_1^T I_{t-1}u_t + \sum_1^T h_{t-1}u_t \\
&= \underbrace{\delta[\sum_{j=1}^{TK} j u_{T_0+j} + TK \sum_{j=1}^{TM} u_{T_1+j}]}_{O_p(T^{3/2})} + \underbrace{\sum_1^T h_{t-1}u_t}_{O_p(T)} + O_p(T^{1/2}), \qquad (47)
\end{aligned}
$$

and

$$
\begin{aligned}
\sum_1^T z_{t-1}D_t &= \delta \sum_1^T I_{t-1}D_t + \sum_1^T h_{t-1}D_t \\
&= \delta \sum_{j=1}^{TK} j + \sum_{j=1}^{TK} h_{T_0+j-1} + O_p(T) \\
&= \underbrace{\tfrac{1}{2}\delta T^2 K^2}_{O_p(T^2)} + \underbrace{\sum_{j=1}^{TK} h_{T_0+j-1}}_{O_p(T^{3/2})} + O_p(T). \qquad (48)
\end{aligned}
$$

From (47) and (48), it follows that:

$$\begin{aligned} \sum_{1}^{T} z_{t-1}\Delta z_t &= \sum_{1}^{T} z_{t-1}u_t + \delta \sum_{1}^{T} z_{t-1}D_t \\ &= \underset{O_p(T^2)}{\tfrac{1}{2}\delta^2 T^2 K^2} \\ &\quad + \underset{O_p(T^{3/2})}{\delta[\sum_{j=1}^{TK}(h_{T_0+j-1} + ju_{T_0+j}) + TK\sum_{j=1}^{TM} u_{T_1+j}]} \quad + \quad O_p(T) \end{aligned} \tag{49}$$

The probability limit of (40) can now be evaluated. Pre-multiplying (40) by $\begin{bmatrix} 1 & 0 \\ 0 & T \end{bmatrix}$ and substituting (44), (45), (46), and (49) into that equation obtains (15) in the text. The limiting *distribution* of (40) is somewhat complicated. Because $(\hat{\mu} : T\hat{\phi})$ has a nonzero plim, the stochastic components of the first as well as the second matrix on the right-hand side of (40) must be taken into account. We plan to derive that limiting distribution in due course.

References

Banerjee, A. and J.J. Dolado (1988) "Tests of the Life Cycle-Permanent Income Hypothesis in the Presence of Random Walks: Asymptotic Theory and Small-sample Interpretations", *Oxford Economic Papers*, 40, 4, 610–633.

Banerjee, A., J.J. Dolado, J.W. Galbraith, and D.F. Hendry (1993) *Co-integration, Error Correction, and the Econometric Analysis of Non-stationary Data*, Oxford, Oxford University Press.

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, 3, 253–277.

Banerjee, A. and D.F. Hendry (1992) "Testing Integration and Cointegration: An Overview", *Oxford Bulletin of Economics and Statistics*, 54, 3, 225–255.

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

Boswijk, H.P. (1992) *Cointegration, Identification, and Exogeneity*, Amsterdam, Tinbergen Institute Research Series No. 37.

Box, G.E.P. and M.E. Muller (1958) "A Note on the Generation of Random Normal Deviates", *Annals of Mathematical Statistics*, 29, 2, 610–611.

Campos, J. and N.R. Ericsson (1988) "Econometric Modeling of Consumers' Expenditure in Venezuela", International Finance Discussion Paper No. 325, Board of Governors of the Federal Reserve System, Washington, D.C.

Campos, J., N.R. Ericsson, and D.F. Hendry (1993) "Cointegration Tests in the Presence of Structural Breaks", International Finance Discussion Paper No. 440, Board of Governors of the Federal Reserve System, Washington, D.C.

Chow, G.C. (1960) "Tests of Equality between Sets of Coefficients in Two Linear Regressions", *Econometrica*, 28, 3, 591–605.

Davidson, J.E.H. and D.F. Hendry (1981) "Interpreting Econometric Evidence: The Behaviour of Consumers' Expenditure in the UK", *European Economic Review*, 16, 1, 177–192.

Davidson, J.E.H., D.F. Hendry, F. Srba, and S. Yeo (1978) "Econometric Modelling of the Aggregate Time-series Relationship between Consumers' Expenditure and Income in the United Kingdom", *Economic Journal*, 88, 352, 661–692.

Dickey, D.A. and W.A. Fuller (1979) “Distribution of the Estimators for Autoregressive Time Series with a Unit Root”, *Journal of the American Statistical Association*, 74, 366, 427–431.

Dickey, D.A. and W.A. Fuller (1981) “Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root”, *Econometrica*, 49, 4, 1057–1072.

Doornik, J.A. and D.F. Hendry (1992) *PcGive Version 7: An Interactive Econometric Modelling System*, Version 7.00, Oxford, Institute of Economics and Statistics, University of Oxford.

Engle, R.F. and C.W.J. Granger (1987) “Co-integration and Error Correction: Representation, Estimation, and Testing”, *Econometrica*, 55, 2, 251–276.

Engle, R.F. and D.F. Hendry (1993) “Testing Super Exogeneity and Invariance in Regression Models”, *Journal of Econometrics*, 56, 1/2, 119–139.

Engle, R.F., D.F. Hendry, and J.-F. Richard (1983) “Exogeneity”, *Econometrica*, 51, 2, 277–304.

Ericsson, N.R. (1991) “Monte Carlo Methodology and the Finite Sample Properties of Instrumental Variables Statistics for Testing Nested and Non-nested Hypotheses”, *Econometrica*, 59, 5, 1249–1277.

Faust, J. (1993) “Near Observational Equivalence and Unit Root Processes: Formal Concepts and Implications”, International Finance Discussion Paper No. 447, Board of Governors of the Federal Reserve System, Washington, D.C.

Gregory, A.W. and B.E. Hansen (1992) “Testing for Regime Shifts in Cointegrated Models”, mimeo, presented at the conference “Recent Developments in the Econometrics of Structural Change”, C.R.D.E., University of Montreal, Montreal, Canada, October 2–3, 1992.

Hammersley, J.M. and D.C. Handscomb (1964) *Monte Carlo Methods*, London, Chapman and Hall.

Hendry, D.F. (1984) “Monte Carlo Experimentation in Econometrics”, Chapter 16 in Z. Griliches and M.D. Intriligator (eds.) *Handbook of Econometrics*, Amsterdam, North-Holland, Volume 2, 937–976.

Hendry, D.F. and G.E. Mizon (1978) “Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England”, *Economic Journal*, 88, 351, 549–563.

Hendry, D.F. and A.J. Neale (1987) “Monte Carlo Experimentation using PC-NAIVE” in T.B. Fomby and G.F. Rhodes, Jr. (eds.) *Advances in Econometrics*, Greenwich, Connecticut, JAI Press, Volume 6, 91–125.

Hendry, D.F. and A.J. Neale (1990) *PC-NAIVE: An Interactive Program for Monte Carlo Experimentation in Econometrics*, Version 6.01, Oxford, Institute of Economics and Statistics and Nuffield College, University of Oxford (documentation by D.F. Hendry, A.J. Neale, and N.R. Ericsson).

Hendry, D.F. and A.J. Neale (1991) “A Monte Carlo Study of the Effects of Structural Breaks on Tests for Unit Roots”, Chapter 8 in P. Hackl and A.H. Westlund (eds.) *Economic Structural Change: Analysis and Forecasting*, Berlin, Springer-Verlag, 95–119.

Johansen, S. (1988) “Statistical Analysis of Cointegration Vectors”, *Journal of Economic Dynamics and Control*, 12, 2/3, 231–254.

Johansen, S. (1991) “Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models”, *Econometrica*, 59, 6, 1551–1580.

Johansen, S. (1992a) “Cointegration in Partial Systems and the Efficiency of Single-equation Analysis”, *Journal of Econometrics*, 52, 3, 389–402.

Johansen, S. (1992b) “A Representation of Vector Autoregressive Processes Integrated of Order 2”, *Econometric Theory*, 8, 2, 188–202.

Johansen, S. (1992c) “Testing Weak Exogeneity and the Order of Cointegration in UK Money Demand Data”, *Journal of Policy Modeling*, 14, 3, 313–334.

Johansen, S. and K. Juselius (1990) “Maximum Likelihood Estimation and Inference on Cointegration — With Applications to the Demand for Money”, *Oxford Bulletin of Economics and Statistics*, 52, 2, 169–210.

Kadane, J.B. (1970) “Testing Overidentifying Restrictions When the Disturbances Are Small”, *Journal of the American Statistical Association*, 65, 329, 182–185.
Kadane, J.B. (1971) “Comparison of k-Class Estimators When the Disturbances Are Small”, *Econometrica*, 39, 5, 723–737.
Kiviet, J.F. and G.D.A. Phillips (1992) “Exact Similar Tests for Unit Roots and Cointegration”, *Oxford Bulletin of Economics and Statistics*, 54, 3, 349–367.
Kremers, J.J.M., N.R. Ericsson, and J.J. Dolado (1992) “The Power of Cointegration Tests”, *Oxford Bulletin of Economics and Statistics*, 54, 3, 325–348.
Lucas, Jr., R.E. (1976) “Econometric Policy Evaluation: A Critique” in K. Brunner and A.H. Meltzer (eds.) *Carnegie-Rochester Conference Series on Public Policy*, Volume 1, *Journal of Monetary Economics*, supplement, 19–46.
MacKinnon, J.G. (1991) “Critical Values for Cointegration Tests”, Chapter 13 in R.F. Engle and C.W.J. Granger (eds.) *Long-run Economic Relationships: Readings in Cointegration*, Oxford, Oxford University Press, 267–276.
Mann, H.B. and A. Wald (1943) “On Stochastic Limit and Order Relationships”, *Annals of Mathematical Statistics*, 14, 3, 217–226.
Park, J.Y. and P.C.B. Phillips (1988) “Statistical Inference in Regressions with Integrated Processes: Part 1”, *Econometric Theory*, 4, 3, 468–497.
Park, J.Y. and P.C.B. Phillips (1989) “Statistical Inference in Regressions with Integrated Processes: Part 2”, *Econometric Theory*, 5, 1, 95–131.
Perron, P. (1989) “The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis”, *Econometrica*, 57, 6, 1361–1401.
Phillips, P.C.B. (1986) “Understanding Spurious Regressions in Econometrics”, *Journal of Econometrics*, 33, 3, 311–340.
Phillips, P.C.B. (1987a) “Time Series Regression with a Unit Root”, *Econometrica*, 55, 2, 277–301.
Phillips, P.C.B. (1987b) “Towards a Unified Asymptotic Theory for Autoregression”, *Biometrika*, 74, 3, 535–547.
Phillips, P.C.B. (1988) “Regression Theory for Near-integrated Time Series”, *Econometrica*, 56, 5, 1021–1043.
Phillips, P.C.B. (1991) “Optimal Inference in Cointegrated Systems”, *Econometrica*, 59, 2, 283–306.
Phillips, P.C.B. and S. Ouliaris (1990) “Asymptotic Properties of Residual Based Tests for Cointegration”, *Econometrica*, 58, 1, 165–193.
Phillips, P.C.B. and P. Perron (1988) “Testing for a Unit Root in Time Series Regression”, *Biometrika*, 75, 2, 335–346.
Quintos, C.E. and P.C.B. Phillips (1992) “Parameter Constancy in Cointegrating Regressions”, mimeo, Department of Economics, Yale University, New Haven, Connecticut.
Sargan, J.D. (1964) “Wages and Prices in the United Kingdom: A Study in Econometric Methodology” in P.E. Hart, G. Mills, and J.K. Whitaker (eds.) *Econometric Analysis for National Economic Planning*, Colston Papers, Vol. 16, London, Butterworths, 25–63 (with discussion); reprinted in D.F. Hendry and K.F. Wallis (eds.) (1984) *Econometrics and Quantitative Economics*, Oxford, Basil Blackwell, 275–314.
Sargan, J.D. (1980) “Some Tests of Dynamic Specification for a Single Equation”, *Econometrica*, 48, 4, 879–897.
Sargan, J.D. and A. Bhargava (1983) “Testing Residuals from Least Squares Regression for Being Generated by the Gaussian Random Walk”, *Econometrica*, 51, 1, 153–174.