The Power of Single Equation Tests for Cointegration when the Cointegrating Vector is Prespecified*

by

Eric Zivot
University of Washington

November 10, 1995
Revised: November 20, 1996

Abstract

In this paper I present an alternative derivation of the asymptotic distribution of Kremers, Ericsson and Dolado's (1992) conditional ECM based t-test for cointegration with a single prespecified cointegrating vector. This alternative distribution, which is identical to the distribution of Hansen's (1995) covariate augmented t-test for a unit root, is valid for weakly exogenous regressors and depends on a consistently estimable nuisance parameter that takes on values in the unit interval. I show analytically, using asymptotic power functions based on near-cointegrated alternatives, that the ECM t-test with a prespecified cointegrating vector can have much higher power than single equation tests for cointegration based on estimating the cointegrating vector. I also characterize situations in which the ECM t-test computed with a misspecified cointegrating vector will have high power.

JEL CLASSIFICATIONS: C22, C51.

KEYWORDS: cointegration, common factor, error correction model, local power, misspecification, near-cointegration, strong exogeneity, weak exogeneity.

---

*Department of Economics, University of Washington, Box 353330, Seattle, WA 98195-3330. Phone: (206) 543-6715; FAX: (206) 685-7477; email: ezivot@u.washington.edu. A Preliminary version of this paper was presented at the Seventh World Congress of the Econometric Society in Tokyo, Japan and at the NBER-NSF Time Series Conference in Cambridge, MA (November, 1995). Partial research support was generously provided by the University of Washington Graduate Research Fund. The author thanks Neil Ericsson, Charles Engel, Bruce Hansen, Charles Nelson, Richard Startz and Timothy Vogelsang for helpful comments and suggestions. The usual disclaimer applies.

1. Introduction

The single equation conditional error correction model (ECM) based t-test for no-cointegration imposing a prespecified cointegrating vector, proposed by Kremers, Ericsson and Dolado (1992), hereafter KED, has not been used empirically because its asymptotic distribution depends on a nuisance parameter that can take on any positive number and is valid only for strongly exogenous regressors. In this paper, I give an alternative representation of the asymptotic distribution for the ECM t-test that is a mixture of a Dickey-Fuller unit root distribution and a standard normal distribution. This mixture distribution depends on a consistently estimable nuisance parameter, ρ^2, that takes on values in the unit interval and describes the long-run contribution of the short-run dynamics to the fit of the ECM regression. This result makes the test feasible for empirical purposes and, additionally, is valid for non-strongly exogenous regressors. It turns out that the asymptotic distribution of the ECM t-test is identical to the asymptotic distribution of Hansen's (1995) covariate augmented t-test for a unit root. In addition, the single equation tests presented herein can be thought of as conditional versions of some of the system-based ECM tests for cointegration described in Horvath and Watson (1995).

I derive analytic power functions for the ECM t-test based on near cointegrated alternatives and show that if ρ^2 is small: (1) The power of the ECM t-test can be arbitrarily larger than the power of the ADF t-test based on a prespecified cointegrating vector; (2) At 50% power, the difference between the ECM t-test with a correctly specified cointegrating vector and an ECM t-test based on estimating the cointegrating vector corresponds to a sample size increase of up to 220%; (3) At 50% power, the implied sample size increase from using an ECM t-test based on estimating the cointegrating vector versus the Engle-Granger residual ADF t-test is about 170%. (4) The power of the ECM t-test when the prespecified cointegrating vector is misspecified is still considerably larger than the power of the ECM t-test using an estimated cointegrating vector for moderate degrees of misspecification. These results emphasize that imposing a prespecified cointegrating vector and correctly modeling the short-run dynamics can have an enormous impact on the power of tests for cointegration.

The plan of the paper is as follows. Section 2 reviews the relationship between cointegration, error correction models and single equation conditional error correction models. In section 3, I discuss the test regressions used to compute the ECM t-test for no-cointegration with a prespecified cointegrating vector and I derive the asymptotic distribution of the t-test under the null of no-cointegration and under near cointegrated alternatives. In section 4, I compare the local power of the ECM t-test when the cointegrating vector is prespecified with the power of the ECM t-test when the cointegrating vector is estimated from the data. Section 5 considers the effects on local power of misspecifying the cointegrating vector. Concluding remarks are given in section 6. Proofs of important results are relegated to the appendix.

I use the following notational conventions throughout the paper. I use the symbol " $\Rightarrow$ " to signify weak convergence, the symbol " $\equiv$ " to signify equality in distribution and the inequality " > 0 " to signify positive definite when applied to matrices. *I(d)* denotes integrated of order *d*. $BM(\Omega)$ refers to a Brownian motion with covariance matrix Ω. Brownian motions $B(r)$ on [0,1] are frequently written as B to achieve notational economy and I often write integrals with respect to Lebesgue measure such as $\int_0^1 B(s)ds$ more simply as $\int_0^1 B$.

2. Cointegration and Conditional Error Correction Models

In this paper, I consider the following single equation conditional ECM with a prespecified cointegrating vector $\alpha = (1, -\beta')'$:

$$\Delta y_{1t} = \mu_{1\cdot 2} + \tau_{1\cdot 2}t + \phi'\Delta y_{2t} + \delta_{1\cdot 2}(y_{1t-1} - \beta' y_{2t-1}) + C_{11}(L)\Delta y_{1t-1} + C_{12}(L)'\Delta y_{2t-1} + \eta_t \qquad (1)$$

where $C_{11}(L)$ and $C_{12}(L)$ are lag polynomials of orders l and p, y_{2t} is an $(n-1)$-dimensional $I(1)$ vector time series and η_t is an innovation process with respect to $\{y_{2t}, y_{1t-j}, y_{2t-j}, j=1,2,...\}$ with variance $\omega_{\eta\eta}$. Equation (1) is a general specification of the type of single equation ECMs discussed at length in Banerjee, Dolado, Galbraith and Hendry (1993), henceforth BDGH, and employed in many empirical studies using the "LSE" or "Hendry" methodology. To interpret (1), think of the data in logs so that the elements of β represent "long-run elasticities" of y_1 with respect to the elements of y_2 and the elements of ϕ represent "short-run elasticities". As discussed in Boswijk (1994), (1) is stable and $y_t = (y_{1t}, y_{2t}')'$ is cointegrated with cointegrating vector $\alpha = (1, -\beta')'$ if the roots of the characteristic equation

$$\psi(z) = (1 - z)(1 - C_{11}(z)) - \delta_{1\cdot 2}z = 0$$

lie outside the unit circle. In this case the cointegrating relationship represents the long-run equilibrium relationship. The model is unstable and there is no long-run equilibrium if there is a root on the unit circle, in which case $\delta_{1\cdot 2} = 0$.

The conditional ECM (1) can be thought of as having been derived from a VAR(p) model for the $(n \times 1)$ vector y_t by conditioning on Δy_{2t}. The VAR formulation is useful for illustrating several concepts that are important for testing the cointegration hypothesis so I will digress for a moment on the relationship between the VAR and the conditional ECM. Let y_t follow the augmented *VAR(p)* process

$$y_t = d_t + x_t \tag{2a}$$

$$\Pi(L)x_t = \epsilon_t, \tag{2b}$$

where d_t represents deterministic terms, $\Pi(L) = I_n - \sum_1^p \Pi_i L^i$ and $\epsilon_t \sim i.i.d.\ N(0,\Sigma)$. To isolate the long-run components it is useful to decompose (2b) as

$$\Delta x_t = \Pi x_{t-1} + \Gamma(L)\Delta x_{t-1} + \epsilon_t, \tag{3}$$

where $\Pi = -\Pi(1)$, $\Gamma(L) = \sum_1^{p-1}\Gamma_i L^{i-1}$ and $\Gamma_i = -\sum_{i+1}^p \Pi_j$. Further, assume that $x_t \sim I(1)$ and Π has rank 1 so that x_t is cointegrated with a single cointegrating vector which is assumed to be of the form $\alpha' x_t = x_{1t} - \beta' x_{2t} \sim I(0)$. Given that Π has rank 1, it can be expressed as

$$\Pi = \delta\alpha' = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \left(1,\ -\beta'\right),$$

where δ_1 *is* (1×1) and δ_2 and β are $((n-1) \times 1)$ vectors, respectively. Then (3) may be rewritten as the vector error correction model (VECM)

$$\Delta x_t = \delta\alpha' x_{t-1} + \Gamma(L)\Delta x_{t-1} + \epsilon_t. \tag{4}$$

Let $d_t = \gamma + \theta t$, for example, and substitute (2a) into (2b) to give VECM representation for y_t

$$\Delta y_t = \mu + \tau t + \delta\alpha' y_{t-1} + \Gamma(L)\Delta y_{t-1} + \epsilon_t, \tag{5}$$

where $\mu = (I_n - \Gamma(1))\theta + \delta\alpha'\theta - \delta\alpha'\gamma$ and $\tau = \delta\alpha'\theta$. Partitioning (4) with respect to y_{1t} and y_{2t} gives the system of equations

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \tau_1 \\ \tau_2 \end{pmatrix} \cdot t + \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \left(y_{1t-1} - \beta' y_{2t-1} \right) + \begin{bmatrix} \Gamma_{11}(L) & \Gamma'_{12}(L) \\ \Gamma_{21}(L) & \Gamma_{22}(L) \end{bmatrix} \begin{bmatrix} \Delta y_{1t-1} \\ \Delta y_{2t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix}. \tag{6}$$

In the VECM (6), $y_t \sim I(1)$ and is not cointegrated if Π has rank zero which implies that $\delta = 0$.

Hence no-cointegration imposes n zero restrictions on δ.

Let $I_{t-1} = \sigma(\Delta y_{1t-1}, \Delta y_{1t-2},..., \Delta y_{1t-p+1}, \Delta y_{2t-1}, \Delta y_{2t-2},..., \Delta y_{2t-p+1}, \; w_{t-1})$. Using the normality assumption, conditional on Δy_{2t} and I_{t-1}, Δy_{1t} is normally distributed with conditional mean and variance given by

$$E[\Delta y_{1t}/\Delta y_{2t}, I_{t-1}] = \mu_{1\cdot 2} + \tau_{1\cdot 2}t + \delta_{1\cdot 2}(y_{1t-1} - \beta' y_{2t-1}) + \phi'\Delta y_{2t} + C_{11}(L)\Delta y_{1t-1} + C_{12}'(L)\Delta y_{2t-1} ,$$

$$var(\Delta y_{1t}/\Delta y_{2t}, I_{t-1}) = \sigma_{11.2} = \sigma_{11} - \sigma_{21}'\Sigma_{22}^{-1}\sigma_{21},$$

where $\phi = \Sigma_{22}^{-1}\sigma_{21}$, $\mu_{1\cdot 2} = \mu_1 - \phi'\mu_2$, $\tau_{1\cdot 2} = \tau_1 = \phi'\tau_2$, $\delta_{1\cdot 2} = \delta_1 - \phi'\delta_2$, $C_{11}(L) = \Gamma_{11}(L) - \phi'\Gamma_{21}(L)$ and $C_{12}'(L) = \Gamma_{12}'(L) - \phi'\Gamma_{22}(L)$. As an alternative to the unconditional system (6), y_t can be thought of as being generated by the conditional/marginal system

$$\Delta y_{1t} = \mu_{1\cdot 2} + \tau_{1\cdot 2}t + \delta_{1\cdot 2}(y_{1t-1} - \beta' y_{2t-1}) + \phi'\Delta y_{2t} + C_{11}(L)\Delta y_{1t-1} + C_{12}'(L)\Delta y_{2t-1} + \epsilon_{1.2t}, \qquad (7)$$

$$\Delta y_{2t} = \mu_2 + \tau_2 t + \delta_2(y_{1t-1} - \beta' y_{2t-1}) + \Gamma_{21}(L)\Delta y_{1t-1} + \Gamma_{22}(L)\Delta y_{2t-1} + \epsilon_{2t}, \qquad (8)$$

where $\epsilon_{1.2t} = \epsilon_{1t} - \phi'\epsilon_{2t}$ and

$$\begin{bmatrix}\epsilon_{1\cdot 2t}\\ \epsilon_{2t}\end{bmatrix} \sim N\left[\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\sigma_{11\cdot 2} & 0\\ 0 & \Sigma_{22}\end{pmatrix}\right].$$

Equation (7) is in the form of (1) with $\mu = \mu_{1\cdot 2}$ and $\eta_t = \epsilon_{1.2t}$.

The conditional ECM (7) is assumed to be the model of primary interest for testing the presence of cointegration with a prespecified cointegrating vector. Accordingly, it is important to discuss the conditions under which the marginal model (8) can be safely ignored when testing for cointegration using (7). The concept of weak exogeneity as defined by Engle, Hendry and Richard (1983) plays a key role in determining the consistency and power of tests for cointegration using conditional ECMs. Johansen (1992) and Urbain (1993) discuss weak exogeneity in general error correction models and the reader is referred to these articles for full details. In the present context, if y_t is cointegrated with cointegrating vector $\alpha = (1, -\beta')'$ and if y_{2t} is weakly exogenous for $\psi = (\delta_1, \beta')'$, then β and δ_1 can be efficiently estimated from the single equation conditional ECM (6). Johansen (1992) shows that y_{2t} is weakly exogenous for ψ if $\delta_2 = 0$; i.e. if the marginal equation for Δy_{2t} is not error correcting. In this case, $\delta_{1\cdot 2} = \delta_1$ and the conditional ECM for Δy_{1t} becomes

$$\Delta y_{1t} = \mu_{1\cdot 2} + \tau_{1\cdot 2}t + \delta_1(y_{1t-1} - \beta' y_{2t-1}) + \phi'\Delta y_{2t} + C_{11}(L)\Delta y_{1t-1} + C_{12}'(L)\Delta y_{2t-1} + \epsilon_{1.2t}.$$

Under weak exogeneity, therefore, testing for no cointegration only involves testing a zero restriction on the scalar parameter δ_1.

Under cointegration, if y_{2t} is not weakly exogenous for ψ then $\delta_{1\cdot 2} = \delta_1 - \phi'\delta_2$ and the hypothesis of no-cointegration requires $\delta_1 = 0$ and $\delta_2 = 0$ since $\delta_{1\cdot 2} = 0$ may occur if y_{1t} and y_{2t} are cointegrated but $\delta_1 = \phi'\delta_2$. In this latter case, the single equation conditional model does not contain all of the necessary information to test the no-cointegration hypothesis and a systems based approach, as in Johansen (1988) or Horvath and Watson (1995), is preferred. As a result, for the single equation tests analyzed in this paper it is necessary to make the additional assumption that under cointegration y_{2t} is weakly exogenous for $\psi = (\delta_1, \beta')'$.

3. Testing for Cointegration in Conditional ECMs with a Prespecified Cointegrating Vector

3.1 Test Statistics

The discussion in the previous section makes it clear that testing for cointegration in the single equation conditional ECM (1), assuming weak exogeneity under cointegration, is based on testing the hypotheses

$$H_0\text{: (no cointegration) } \delta_1 = 0 \text{ vs. } H_1\text{: (cointegration) } \delta_1 < 0.$$

KED suggested using the standard t-ratio $t^K(\hat{\delta}_1) = \hat{\delta}_1/SE(\hat{\delta}_1)$, where $\hat{\delta}_1$ is the OLS estimate of δ_1 and $SE(\hat{\delta}_1)$ is its estimated standard error. They derived the asymptotic distribution of $t^K(\hat{\delta}_1)$ under the null of no cointegration and under local-to-zero near cointegrated alternatives for a simple bivariate model with no deterministic terms or higher order dynamics. Their functional representation of the limiting distributions, however, depends on a nuisance parameter that can take on any positive value and so is difficult to use in practice.

With deterministic terms, d_t, in the representation for y_t, the specific regression equation used to estimate δ_1 depends on the nature of these terms. I consider the specification $d_t = \gamma + \theta t$, where both γ and θ are $n \times 1$ vectors. The conditional/marginal representation, under the weak exogeneity assumption, is then

$$\Delta y_{1t} = \mu_{1\cdot 2} + \tau_1 t + \delta_1(y_{1t-1} - \beta' y_{2t-1}) + \phi'\Delta y_{2t} + C_{11}(L)\Delta y_{1t-1} + C_{12}'(L)\Delta y_{2t-1} + \eta_t,$$

$$\Delta y_{2t} = \mu_2 + \Gamma_{21}(L)\Delta y_{1t-1} + \Gamma_{22}(L)\Delta y_{2t-1} + \epsilon_{2t},$$

where $\tau_1 = \delta_1\alpha'\theta$. Notice that the weak exogeneity assumption, $\delta_2 = 0$, eliminates the time trend from the marginal model for Δy_{2t}.

There are four versions of the specification $d_t = \gamma + \theta t$ that are used in empirical applications. These cases and the restrictions they imply on the trend parameters in the conditional and marginal models are summarized in table 1. In case I, $\gamma = \theta = 0$ so that $\mu = \tau = 0$. The conditional ECM to be estimated is then

$$\Delta y_{1t} = \delta_1 \alpha' y_{t-1} + \zeta' z_t + \eta_t \quad (9)$$

where $z_t' = (\Delta y_{1t-1}, \ldots, \Delta y_{1t-l}, \Delta y_{2t}', \Delta y_{2t-1}', \ldots, \Delta y_{2t-p}')$ and $\zeta' = (c_{11,1}, \ldots, c_{11,l-1}, \phi', c_{12,1}', \ldots, c_{12,p-1}')$ are $(1 \times k)$ vectors. In case II, $\gamma \neq 0$ and is unrestricted but $\theta = 0$. This implies that $\mu_2 = 0$ so that $\mu_{1\cdot 2} = \mu_1 = -\delta_1 \alpha' \gamma$ and $\tau_1 = 0$. The conditional ECM becomes

$$\Delta y_{1t} = \delta_1(\alpha' y_{t-1} - \mu^+) + \zeta' z_t + \eta_t \quad (10)$$

where $\mu^+ = \alpha' \gamma$. Here, y_t is not trending and there are no restrictions on the initial values of x_t or on the mean of the error correction term $\alpha' y_t$. It is unlikely, however, that μ^+ is known *a priori*, e.g. specified by economic theory, so it is not possible to estimate (10) directly by OLS. Moreover, under the null of no-cointegration $\alpha' y_t = \alpha' x_t + \alpha' \gamma$, which is $I(1)$ with drift, so that a constant is required in the ECM regression to obtain a similar test statistic. Therefore, the test regression is (9) with z_t augmented with a constant[1]. In case III, $\gamma \neq 0$ and $\theta \neq 0$ but is restricted by the relation $\alpha' \theta = 0$. This implies that $\mu_{1\cdot 2}$ is unrestricted, $\tau_1 = 0$ and so the conditional ECM becomes

$$\Delta y_{1t} = \mu_{1\cdot 2} + \delta_1 \alpha' y_{t-1} + \zeta' z_t + \eta_t.$$

The test regression in this case is also (9) with z_t augmented with a constant. Case IV has $\gamma \neq 0$ and $\theta \neq 0$ with no restrictions on either γ or θ. Here $\mu_{1\cdot 2}$ is unrestricted but $\tau_1 = \delta_1 \alpha' \theta$ so that the time trend is restricted to the error correction term. The conditional ECM is

$$\Delta y_{1t} = \mu_{1\cdot 2} + \delta_1(\alpha' y_{t-1} - \tau^+(t-1)) + \zeta' z_t + \eta_t \quad (11)$$

where $\tau^+ = \alpha' \theta$. As in case II, it unlikely that τ^+ is prespecified by economic theory so that (11) is not directly estimable by OLS. Also under the null of no-cointegration, $\alpha' y_t = \alpha' x_t + \alpha' \gamma + \alpha' \theta t$ and so a constant and a time trend must be included in the ECM regression to obtain a similar test. Therefore the test regression in this case is (9) where z_t is augmented with a constant and a time trend.

3.2 Asymptotic Theory Under the Null of No Cointegration

It will be useful to rewrite the conditional model for Δx_{1t} as

$$a(L)\Delta\alpha' x_t = \delta_1\alpha' x_{t-1} + b(L)'\Delta x_{2t} + \eta_t = \delta\alpha' x_{t-1} + e_t \qquad (12)$$

where $a(L) = 1 - C_{11}(L)L$, $b(L) = (\phi - \beta) + [C_{12}(L) + C_{11}(L)\beta]L$ and $e_t = b(L)'\Delta x_{2t} + \eta_t$. Define $\xi_t = (\eta_t, \Delta x_{2t}')'$ and $v_t = (\eta_t, e_t)'$. It is assumed that

$$T^{-1/2}\sum_{t=1}^{[Tr]}\xi_t \Rightarrow B_\xi(r) \equiv BM(\Omega_\xi),\; T^{-1/2}\sum_{t=1}^{[Tr]}v_t \Rightarrow B_v(r) \equiv BM(\Omega_v),$$

where $B_\xi(r) = (B_\eta(r), B_2(r)')'$ is an n-dimensional Brownian motion, $B_v(r) = (B_\eta(r), B_e(r))'$ is a bivariate Brownian motion,

$$\Omega_\xi = \begin{pmatrix} \omega_{\eta\eta} & \omega_{\eta 2}' \\ \omega_{2\eta} & \Omega_{22} \end{pmatrix},\; \Omega_v = \begin{pmatrix} \omega_{\eta\eta} & \omega_{\eta e} \\ \omega_{e\eta} & \omega_{ee} \end{pmatrix}\;,$$

with $\omega_{ee} = b(1)'\Omega_{22}b(1) + \omega_{\eta\eta} + 2b(1)'\omega_{2\eta}$ and $\omega_{\eta e} = b(1)'\omega_{2\eta} + \omega_{\eta\eta}$. In addition, define the long-run correlation parameters

$$r^2 = \frac{\omega_{\eta 2}'\Omega_{22}^{-1}\omega_{2\eta}}{\omega_{\eta\eta}},\; \rho^2 = \frac{\omega_{\eta e}^2}{\omega_{\eta\eta}\omega_{ee}},\; R^2 = \frac{\omega_{\eta\eta}}{\omega_{ee}},\; q^2 = \frac{\omega_{ee}}{\omega_{\eta\eta}} - 1\;.$$

In the above expressions, r^2 is the squared long-run multiple correlation coefficient between η_t and Δx_{2t}, ρ^2 is the squared long-run correlation between e_t and η_t, R^2 is the proportion of the long-run variance of e_t explained by η_t and q^2 measures the percentage by which the long-run variance of e_t is larger than the long-run variance of η_t. Notice that $R^2 = 1/(1 + q^2)$ and if Δx_{2t} is strongly exogenous, i.e. if Δx_{1t} does not Granger cause Δx_{2t} in the marginal model for Δx_{2t}, then $r^2 = 0$ and $\rho^2 = R^2$.

At one extreme, $\rho^2 = 1$ which implies that the error terms e_t and η_t are perfectly correlated in the long-run and have the same long-run variance. In this case, $b(L)'\Delta x_{2t}$ explains none of the long-run variability of e_t. This occurs in the conditional ECM if $b(1) = 0$. KED emphasize that $b(1) = 0$

if, for example, Δx_{1t} satisfies the generalized common factor restriction $\phi = \beta$ and $C_{12}(1) = C_{11}(1)\beta$, in which case the conditional ECM for Δx_{1t} takes the form of an ADF regression. At the other extreme, $\rho^2 = 0$ so that the long-run variance of e_t is infinitely larger than the long-run variance of η_t. This case occurs when $b(L)\Delta x_{2t}$ explains all of the long-run movement of e_t.

The condition $R^2 = \rho^2$ occurs when Δx_{2t} is long-run uncorrelated with η_t, which, Hansen (1995) (hereafter referred to as Hansen) states, should hold in a well specified dynamic regression. In the VECM set-up, however, this occurs when Δx_{2t} is weakly and strongly exogenous and a well specified conditional ECM only requires current and lagged values of Δx_{2t} as well as lagged values of Δx_{1t}. If, however, Δx_{2t} is not strongly exogenous then the long-run correlation between Δx_{2t} and η_t can be eliminated by adding leads of Δx_{2t} to the conditional ECM [2]. In this case, we define $b(L) = (\phi - \beta) + [C_{12}(L) + C_{11}(L)\beta]L + C^{12}(L^{-1})$ where $C^{12}(L^{-1})$ is a polynomial in the forward shift operator L^{-1}. Alternatively, the long-run correlation may also be eliminated using a Phillips-Hansen type nonparametric correction to the ECM as in Inder (1993).

To succinctly express the limiting distributions of the ECM-based test statistics when deterministic terms are added to the test regressions, it is useful to employ the following notation. Let $X(r)$ and $Z(r)$ denote two vector processes defined on [0,1]. Consider the continuous time regression of $X(r)$ on $Z(r)$, $X(r) = \hat{\alpha}'Z(r) + Q_{Z(r)}X(r)$ where $\hat{\alpha}$ solves $\min_{\alpha} \int_0^1 \|X(r) - \alpha' Z(r)\|^2 dr$. The continuous time regression residual, $Q_{Z(r)}X(r)$, is defined as

$$Q_{Z(r)}X(r) = X(r) - \left(\int_0^1 Z(r)X(r)dr\right)' \left(\int_0^1 Z(r)Z(r)'dr\right)^{-1} Z(r) .$$

For example, if $Z(r) = 1$ then $Q_1X(r) = X(r) - \int_0^1 X(r)dr$ which is a demeaned version of $X(r)$.

The following theorem gives the asymptotic distributions for the ECM based t-test under the null hypothesis of no-cointegration when a fixed cointegrating vector β is imposed.

Theorem 1 *In case I, if δ_1 is estimated from (9) then as $T \to \infty$*

$$t^K(\hat{\delta}_1) \Rightarrow \frac{\rho \cdot \int_0^1 W_e(r) dW_e(r)}{\left(\int_0^1 W_e(r)^2 dr \right)^{1/2}} + (1 - \rho^2)^{1/2} \cdot N(0,1) \ .$$

where $W_e(r)$ is a standard Brownian motion independent of the N(0,1) random variable. In cases II and III, if a constant is added to the ECM regression (9) then $W_e(r)$ is replaced by $Q_1 W_e(r)$; In case IV, if a constant and trend are added to the ECM regression (9) then $W_e(r)$ is replaced by $Q_{(1,r)} W_e(r)$.

Corollary *If Δy_{2t} is strongly exogenous then*

$$\frac{\rho \cdot \int_0^1 W_e(r) dW_e(r)}{\left(\int_0^1 W_e(r)^2 dr \right)^{1/2}} + (1 - \rho^2)^{1/2} \cdot N(0,1) \equiv \frac{\int_0^1 (W_\eta(r) + q \cdot W_{e \cdot \eta}(r)) dW_\eta(r)}{\left(\int_0^1 (W_\eta(r) + q \cdot W_{e \cdot \eta}(r))^2 dr \right)^{1/2}}$$

where $W_\eta(r)$ and $W_{e \cdot \eta}(r)$ are independent standard Brownian motions.

Theorem 1 shows that if a fixed cointegrating vector is imposed then the asymptotic distribution of $t^K(\hat{\delta}_1)$ depends on the nuisance parameter ρ^2 measuring the long-run contribution of Δy_{2t} to the conditional model. The asymptotic distribution is a linear combination of a Dickey-Fuller unit root distribution and a standard normal random variable. Notice that when $\rho^2 = 1$ the distribution collapses to the Dickey-Fuller unit root distribution and when $\rho^2 = 0$ the distribution reduces to a standard normal. In fact, the distribution of $t^K(\hat{\delta}_1)$ is identical to the asymptotic distribution of Hansen's covariate augmented t-test for a unit root. Simulated critical values for selected values of $\rho^2 \in (0,1)$ are given in Hansen.

If Δy_{2t} is strongly exogenous then $\rho^2 = 1/(1+q^2)$ and the Corollary shows that the asymptotic distribution of the t-test can be alternatively expressed in terms of q^2. This is the result first obtained by KED for the simple case with no deterministic terms, $n = 2$ and $p = 0$. Hence KED's result is only valid for strongly exogenous conditioning variables whereas the result presented here holds more generally.

Even though the asymptotic distribution of $t^K(\hat{\delta}_1)$ depends on the nuisance parameter ρ^2, Hansen shows it can still be used for inference since it is possible to consistently estimate ρ^2 using nonparametric techniques. For example, an estimate of ρ can be constructed from the nonparametric estimate of Ω_v:

$$\hat{\Omega}_v = \begin{pmatrix} \hat{\omega}_{\eta\eta} & \hat{\omega}_{\eta e} \\ \hat{\omega}_{e\eta} & \hat{\omega}_{ee} \end{pmatrix} = \sum_{k=-M}^{M} w\left(\frac{k}{M}\right) \frac{1}{T} \sum_{t=k+1}^{T} \hat{v}_t \hat{v}_{t-k}', \quad \hat{\rho}^2 = \frac{\hat{\omega}_{\eta e}^2}{\hat{\omega}_{\eta\eta}\hat{\omega}_{ee}}$$

where $w(\cdot)$ is a kernel weight function, M is a bandwidth parameter and $\hat{v}_t = (\hat{\eta}_t, \hat{e}_t)'$ is constructed from the parameters of the estimated ECM. Hansen suggests using the Bartlett or Parzen kernel with M determined by Andrews' (1991) automatic bandwidth selection procedure.

3.3 Asymptotic theory Under Near Cointegrated Alternatives

The asymptotic power analysis is for near cointegrated alternatives of the form

$$H_a\colon\ \delta_1 = -ca(1)/T \tag{13}$$

where c is a constant and T is the sample size. The no-cointegration null holds when $c = 0$ and holds locally as $T \to \infty$ for $c > 0$.

The asymptotic power functions for the near-cointegrated alternatives are derived using the local-to-unity asymptotics of Phillips (1987) and Chan and Wei (1987) as applied by Hansen. This theory is based on diffusion representations of continuous stochastic processes. Let $Z(r)$ be any stochastic process and let c be any constant. Then $Z^c(r)$ is defined as the solution to the stochastic differential equation $dZ^c(r) = -cZ^c(r) + dZ(r)$.

The following theorem gives the asymptotic distribution of $t^K(\hat{\delta}_1)$ under the near cointegrated alternative (13).

Theorem 2 *In case I, if* y_{1t} *is generated from (1),* Δy_{2t} *is weakly exogenous for* $\psi = (\beta', \delta_1)'$ *and (13) holds then as* $T \to \infty$

$$t^K(\hat{\delta}_1) \Rightarrow \frac{-c}{R}\left(\int_0^1 W_e^c(r)^2 dr\right)^{1/2} + \frac{\rho \cdot \int_0^1 W_e^c(r) dW_e(r)}{\left(\int_0^1 W_e^c(r)^2 dr\right)^{1/2}} + (1 - \rho^2)^{1/2} \cdot N(0,1) ,$$

where $W_e(r)$ *is a standard Brownian motion independent of the N(0,1) random variable. In cases II and III, if a constant is added to the ECM regression then* $W_e^c(r)$ *is replaced by* $Q_1 W_e^c(r)$*; In case IV, if a constant and trend are added to the ECM regression then* $W_e^c(r)$ *is replaced by* $Q_{(1,r)} W_e^c(r)$.

The local asymptotic distribution of $t^K(\hat{\delta}_1)$ is identical to the local distribution of Hansen's covariate augmented t-test. The local power of $t^K(\hat{\delta}_1)$ depends on c, ρ^2 and R^2. Two cases are of interest. In the first case, Δy_{2t} is strongly exogenous so that $r^2 = 0$ and $\rho^2 = R^2$. In the second case, Δy_{2t} is not strongly exogenous, $r^2 \neq 0$ and $\rho^2 \neq R^2$. Hansen shows that local power increases with decreases in ρ^2 and decreases with increases in the number of deterministic terms in the regression. He also shows that, for a given value of ρ^2, power is larger for smaller values of R^2 and vice-versa.

When $\rho^2 = 1$, $t^K(\hat{\delta}_1)$ behaves very much like the ADF t-test for a unit root in the cointegrating residuals $\alpha' y_t$. A comparison of the local power of $t^K(\hat{\delta}_1)$ for $\rho^2 = 1$ and $\rho^2 = 0.1$ for a given value of c gives an indication of the potential power gains from using $t^K(\hat{\delta}_1)$. Alternatively, as in Horvath and Watson (1995), one may fix the power at a given percent and compare the sample size differentials implied by the different values of c for the two test statistics. For example, figure 1 illustrates the asymptotic local power of $t^K(\hat{\delta}_1)$ for the case in which $\rho^2 = R^2$ and a constant is included in the ECM test regression. From figure 1, it can be deduced that at 50% power the potential power gain from using $t^K(\hat{\delta}_1)$ instead of the ADF t-test, for a model estimated with a constant, corresponds to a sample size increase of roughly 667%.

4. Comparison of Local Powers of ECM t-tests with β Prespecified and β Estimated.

It is of interest to compare the asymptotic local power of the conditional ECM t-test with β prespecified to a conditional ECM t-test with β estimated. This comparison will highlight the local power gains from using a test that imposes the true cointegrating vector versus a test that does not.

BDGH, building on earlier work of Banerjee, Hendry and Smith (1986) and KED, propose a simple t-test for no-cointegration in a conditional ECM with unknown β. Their approach is based on rewriting (9) as

$$\Delta y_{1t} = \delta_1 \alpha_A' y_{t-1} + \psi' y_{2t-1} + \zeta' z_t + \eta_t \tag{14}$$

where $\alpha_A = (1, -\beta_A')'$, β_A is an arbitrary $(n-1) \times 1$ vector and $\psi = \delta_1(\beta_A - \beta)$[3]. Notice that δ_1 is not affected by imposing the arbitrary error correction term so that a test for no-cointegration based on the significance of δ_1 is still, in principle, valid. Hence the t-ratio for δ_1 from this regression can be used as a test for cointegration with β unknown provided its asymptotic distribution can be determined. We denote this statistic $t^U(\hat{\delta}_1)$[4]. Using similar arguments as in KED, BDGH claim that $t^U(\hat{\delta}_1)$ will have higher power than the residual-based two-step Engle-Granger ADF t-statistic.

Boswijk (1994) derives the asymptotic null distribution of $t^U(\hat{\delta}_1)$ and shows that it is asymptotically similar only if Δy_{2t} is strongly exogenous[5]. In this case, the asymptotic null distribution is independent of ρ^2 but depends on the dimension, $n-1$, of Δy_{2t}. Banerjee, Dolado and Mestre (1994) tabulate critical values for $t^U(\hat{\delta}_1)$ for $n-1$ = 1, ..., 5 for the no-constant, constant only and constant and trend cases and show that these critical values are very similar to the critical values tabulated by Phillips-Ouliaris (1990) for residual-based tests for cointegration.

Using the results of the previous section it is straightforward to derive the asymptotic distribution of $t^U(\hat{\delta}_1)$ under the local alternative (13).

Theorem 3 *If y_{1t} is generated from (1), Δy_{2t} is strongly exogenous and (13) holds then as $T \rightarrow \infty$*

$$t^{U}(\hat{\delta}_1) \Rightarrow -c\cdot\left(\int_0^1 \left(Q_{Z(r)}[W_\eta^c(r) + q\cdot a' W_2^c(r)]\right)^2\right)^{1/2} + \frac{\int_0^1 Q_{Z(r)}[W_\eta^c(r) + q\cdot a' W_2^c(r)]dW_\eta(r)}{\left(\int_0^1 \left(Q_{Z(r)}[W_\eta^c(r) + q\cdot a' W_2^c(r)]\right)^2\right)^{1/2}},$$

where a is any (n-1) ×1 vector of unit length, $q^2 = (1-\rho^2)/\rho^2$ *and Z(r) is a stochastic process on [0,1] such that: (case I)* $Z(r) = W_2(r)$*; (cases II and III)* $Z(r) = (W_2(r)', 1)'$ *if a constant is included in (14); and (case IV)* $Z(r) = (W_2(r)', 1, r)'$ *if a constant and trend are included in(14).*

The asymptotic distribution of $t^U(\hat{\delta}_1)$ under the local alternative depends on c, ρ^2, n and the nature of the deterministic terms in the ECM regression. When $c = 0$, the distribution collapses to

$$\left(\int_0^1 (Q_{Z(r)}W_\eta(r))^2\right)^{-1/2} \int_0^1 Q_{Z(r)}W_\eta(r)dW_\eta(r)$$

which is independent of ρ^2, but dependent on the dimension of W_2, and is equivalent to the expression given in theorem 2 of Boswijk (1994).

Figures 2-4 compare the local powers of $t^U(\hat{\delta}_1)$ and $t^K(\hat{\delta}_1)$ for $\rho^2 = 0.1$, 0.5 and 0.9 with $n = 2$. For each value of ρ^2 the power of $t^K(\hat{\delta}_1)$ is well above the power of $t^U(\hat{\delta}_1)$ and the power gains are larger at smaller values of ρ^2. For example, at 50% power the power difference when no deterministic terms are included in the regression corresponds to sample size increases of roughly 220%, 75% and 56% for $\rho^2 = 0.1$, 0.5 and 0.9. When a constant is included the sample increases are 175%, 85% and 35% and when a constant and trend is included the sample size increases are 220%, 81% and 20%, respectively.

Figures 5-7 show the difference in local power between $t^U(\hat{\delta}_1)$ and $t^K(\hat{\delta}_1)$ as the dimension of Δy_{2t} increases for $\rho^2 = 0.9$, 0.5 and 0.1. For a given value of ρ^2, the power of $t^K(\hat{\delta}_1)$ is the same for all n whereas the power of $t^U(\hat{\delta}_1)$ declines as n increases. Interestingly, the power loss of $t^U(\hat{\delta}_1)$ as n increases is substantially reduced for small values of ρ^2. Comparing the power of $t^U(\hat{\delta}_1)$ at ρ^2

=0.1 and 0.9 gives an indication of the potential power increase from using $t^U(\hat{\delta}_1)$ versus the Engle-Granger residual-based ADF t-statistic. At 50% power and n=2, the potential power gain for a model estimated with a constant corresponds to a sample size increase of roughly 172%.

5. Effects on Local Power of Misspecifying the Cointegrating Vector

It is clear from the previous sections that there are potentially very large power gains associated with imposing the true value of the cointegrating vector in single equation tests for no-cointegration. However, it is not so clear what happens to the performance of $t^K(\hat{\delta}_1)$ if the wrong cointegrating vector is imposed in the estimated ECM regression. Following Horvath and Watson, I consider the behavior of $t^K(\hat{\delta}_1)$ under the local alternative (13) since under fixed cointegrated alternatives $t^K(\hat{\delta}_1)$ is an inconsistent test if the lagged error correction term is misspecified.

To simplify the analysis, let (1) represent the true model with Δy_{2t} strongly exogenous. Suppose an investigator imposes the misspecified cointegrating vector $\alpha_M = (1, -\beta_M')'$ where

$$\beta_M = \beta + a(1)^{-1}d \tag{16}$$

and d is any $(n\text{-}1)\times 1$ vector. The misspecified error correction term is then $\alpha_M' y_t = \alpha' y_t - a(1)^{-1} d' y_{2t}$. The true model may therefore be reexpressed as (14) with β_A given by (16). Notice that the misspecification of the error correction term creates additional $I(1)$ regressors in the true model (14). Under the local alternative (13), $\psi = -cd/T$ so that the coefficients on the additional $I(1)$ regressors are local-to-zero. The estimated model, however, is the misspecified model which excludes the lagged value of y_{2t}:

$$\Delta y_{1t} = \delta_1 \alpha_M' y_{t-1} + \zeta' z_t + u_t \tag{17}$$

where $u_t = \psi' y_{2t-1} + \eta_t$. The asymptotic distribution of $t^K(\hat{\delta}_1)$ computed from (17) under the local alternative (13) is given in the next theorem.

Theorem 4 *In case I, if y_{1t} is generated from (1), Δy_{2t} is strongly exogenous and (13) and (16) hold then as $T \rightarrow \infty$*

$$t^K(\hat{\delta}_1) \Rightarrow -c\cdot\left(\int_0^1 Z^c(q,s;r)^2\right)^{-1/2} - c\cdot s\cdot\frac{\int_0^1 Z^c(q,s;r)(a'W_2(r))}{\left(\int_0^1 Z^c(q,s;r)^2\right)^{1/2}} + \frac{\int_0^1 Z^c(q,s;r)dW_\eta(r)}{\left(\int_0^1 Z^c(q,s;r)^2\right)^{1/2}}$$

where $Z^c(q,s;r) = q(a'W_2^c(r)) - s(a'W_2(r)) + W_\eta^c(r)$, $q^2 = b(1)'\Omega_{22}b(1)/\omega_{\eta\eta} = (1 - \rho^2)/\rho^2$, $s^2 = d'\Omega_{22}d/\omega_{\eta\eta}$ *and* a *is any (n-1) × 1 vector of unit length. In case II, if a constant is included in the regression then* $Z^c(q,s;r)$ *is replaced by* $Q_1Z^c(q,s;r)$. *In cases III and IV, if a constant and trend are included then* $Z^c(q,s;r)$ *is replaced by* $Q_{(1,r)}Z^c(q,s;r)$.

The asymptotic distribution of $t^K(\hat{\delta}_1)$ computed from the misspecified model under the local alternative depends on the parameters c, q^2 (and hence ρ^2), s^2 and n. The parameter s is the length of d scaled by the relative variability of the long-run variances of Δy_{2t} and η_t. When s is large (large misspecification) the second term in the limiting expression for $t^K(\hat{\delta}_1)$, arising from the local-to-zero I(1) regressors that are created by misspecifying the error correction term, becomes a large positive number and reduces power relative to the correctly specified model. When $d = 0$, the estimated model is correctly specified and the distribution of $t^K(\hat{\delta}_1)$ reduces to the expression given in theorem 2. Under the null of no-cointegration, $c = 0$, the distribution collapses to the expression given in the corollary to theorem 1 with q^2 replaced by $(q - s)^2$. Notice that in case III it is necessary to include both a constant and trend in the test regression since the misspecification of the error correction term induces a deterministic trend in the model.

Figures 8-25 give the local power functions of $t^K(\hat{\delta}_1)$ and $t^U(\hat{\delta}_1)$ for a bivariate model with $d = 0, 0.1, 0.3, 0.5$; $\rho^2 = 0.9, 0.1$, $\nu = \omega_{22}/\omega_{\eta\eta} = 1, 5,$ and 10 . Power curves are given for models fitted with no constant or trend, constant only and constant and trend. The qualitative results for these three cases are similar. To interpret the degree of misspecification in β think of the data in logs with $\beta = 1$ so that the true model imposes long-run homogeneity between y_{1t} and y_{2t}. Then, for example, $d = 0.1$ corresponds to misspecifying the long-run elasticity by 10%. .

To interpret the relationship between d, s, ν and ρ^2 consider the case where $\nu = 1$ so that s

$= d$. Since ν is held fixed, changes in ρ^2 are due solely to changes in $b(1)^2$. Recall, when $b(1) = 0$ there is a common factor in the dynamics of the ECM so that large values of $b(1)$ correspond to large violations in the common factor restriction. Next consider the case where $\nu = 5$. Here the long-run variance of Δy_{2t} is five times larger than the long-run variance of η_t. The increase in ν scales up the degree of misspecification captured by d and so one may think of s as the scaled deviation from the true cointegrating vector. In this regard, the case with $\nu = 5$ and $s = 0.22$, 0.67 and 1.11 corresponds to the case with $\nu = 1$ and $d = 0.22$, 0.67 and 1.11.

Figures 8, 9, 14, 15, 20 and 21 give the local power results for $\nu = 1$. For $\rho^2 = 0.9$, the local power of $t^K(\hat{\delta}_1)$ falls as d rises. For $d = 0.1$, power is very close to the power for $d = 0$ and is uniformly above the power of $t^U(\hat{\delta}_1)$ except in the constant and trend case for $c > 12$. For $d > 0.1$ the power of $t^K(\hat{\delta}_1)$ drops precipitately and lies below the power of $t^U(\hat{\delta}_1)$ for moderate values of c. The situation for $\rho^2 = 0.1$ is much different. For $d < 0.5$, the power of $t^K(\hat{\delta}_1)$ is almost identical to the power at $d = 0$. When $d = 0.5$, however, the power starts to fall for large values of c. This makes sense since in these cases the model specification approaches one with δ_1 fixed and $t^K(\hat{\delta}_1)$ is an inconsistent test. In sum, with a strong violation of the common factor restriction and $\nu = 1$, even relatively large misspecifications of the cointegrating vector do not seriously affect the local power of $t^K(\hat{\delta}_1)$.

Next, consider the power results for $\nu = 5$ presented in figures 10, 11, 16, 17, 22 and 23. The increase in the long-run variability of Δy_{2t} scales up any misspecification in β and, consequently, the power of $t^K(\hat{\delta}_1)$ is uniformly lower relative to the case where $\nu = 1$. When $\rho^2 = 0.9$ and $d = 0.1$ ($s = 0.22$), the power of $t^K(\hat{\delta}_1)$ is now substantially lower than the power at $d=0$ and lies below the power of $t^U(\hat{\delta}_1)$ for moderate values of c. For d > 0.1, the power of $t^K(\hat{\delta}_1)$ never gets above 15%. The results are better, however, for $\rho^2 = 0.1$. Here, the power of $t^K(\hat{\delta}_1)$ for $d = 0.1$ is almost identical to the power at $d = 0$. For d > 0.1, power starts to fall for larger values of c but still remains greater than 50%, for all trend cases, at $c = 16$.

Last, figures 12, 13, 18, 19, 24 and 25 illustrate the results for $\nu = 10$. For $\rho^2 = 0.9$ only the $d = 0.1$ case with no constant or trend exhibits non-negligible power but $t^U(\hat{\delta}_1)$ dominates $t^K(\hat{\delta}_1)$ for values of c greater than seven. For $\rho^2 = 0.1$, the power of $t^K(\hat{\delta}_1)$ for $d = 0.1$ is still indistinguishable from the power for $d = 0$. The power results for $d > 0.1$ are similar to the $\nu = 5$ case. Thus, even for large values of ν the misspecified model retains high power for moderate values of d provided there is a large violation in the common factor restriction.

The preceding power analysis for a misspecified model is similar to the analysis presented in Horvath and Watson. However, they use a simple bivariate model without short-run dynamics, impose weak exogeneity and set the covariance of the errors equal to the identity matrix. In this set-up, $\rho^2 = 1$ and Horvath and Watson's ECM Wald test behaves very similarly to the ADF t-test.

6. Conclusions

In this paper I provided an alternative representation of the asymptotic distribution of KED's t-test for no-cointegration with a prespecified cointegrating vector that allows for an empirically feasible test. The test is shown to be closely related to Hansen's covariate augmented t-test for a unit root. The ECM t-test with a prespecified cointegrating vector is shown to have higher power than the ADF test as well as single equation tests that implicitly estimate the cointegrating vector. The ECM t-test is also shown to have good power even when the cointegrating vector is moderately misspecified.

The single-equation conditional ECM-based tests considered in this paper require that the cointegrating rank be one and that the integrated regressors be weakly exogenous for the long-run parameters under the alternative of cointegration. If the number of cointegrating vectors is greater than one or if weak exogeneity fails then a systems-based ECM approach as in Horvath and Watson (1995) is recommended.

7. Notes

1.Horvath and Watson (1995) found that their ECM-based tests for no-cointegration that did not impose the restriction that the constant enter into the cointegrating vector had higher power than tests that imposed the restriction.

2.This technique is used by Phillips and Loretan (1991), Saikkonnen (1991) and Stock and Watson (1994) to get efficient estimates of a cointegrating vector in the presence of long-run correlation.

3.BDGH suggest using $\beta_A = \iota$ where ι is an $(n\text{-}1) \times 1$ vector of ones. If the data are in logs, then the error correction term $y_{1t-1} - \iota' y_{2t-1}$ imposes long-run homogeneity and the term $\zeta' y_{2t-1}$ allows for any departure in long-run homogeneity.

4.This test is called the PC-GIVE unit root test in Hendry and Doornik's (1993) program PC-GIVE.

5.If Δy_{2t} is not strongly exogenous then the ECM regression may be modified with leads of Δy_{2t} or with a Phillips-Hansen type nonparametric correction to eliminate the long-run correlation between Δy_{2t} and η_t.

8. References

Andrews, D.W.K. (1991) Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica*, **59**, 817-858.

Banerjee, A., J. Dolado, and Mestre (1995) On the Power of Cointegration Tests: Dimension Invariance vs. Common Factors. Institute for Economic Research Discussion Paper #922, Queens University, Kingston, Ontario, Canada.

_____, J. Dolado, J.W. Galbraith and D.F. Hendry (1993) *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data* Advanced Texts in Econometrics, Oxford University Press.

_____, 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**, 253-77.

Boswijk, H.P. (1994) Testing for an Unstable Root in Conditional and Structural Error Correction Models. *Journal of Econometrics*, **63**, 37-60.

Campos, J., N.R. Ericsson and D.F. Hendry (1996) Cointegration Tests in the Presence of Structural Breaks. *Journal of Econometrics* **70**, 187-220.

Chan, N.H. and C.Z. Wei (1987) Asymptotic Inference for Nearly Nonstationary AR(1) Processes. *The Annals of Statistics*, **15**, 1050-1063.

Doornik, J.A. and D.F. Hendry (1994) *PcGive Professional 8.0: An Interactive Econometric Modelling System*. International Thomson Publishing, London, England.

Engle, R.F. and C.W.J. Granger (1987) Co-Integration and Error Correction: Representation, Estimation and Testing. *Econometrica*, **55**: 251-76.

_____ and Hendry, D. F., and Richard, J. -F. (1983) Exogeneity. *Econometrica*, **51**: 277-304.

Hansen, B.E. (1995) Rethinking the Univariate Approach to Unit Root Testing: Using Covariates to Increase Power. *Econometric Theory*, Vol. **11**, No. 5, 1148-1171..

Horvath, M.T. and M.W. Watson (1995) Testing for Cointegration when some of the Cointegrating Vectors are Prespecified. *Econometric Theory*, Vol. **11**, No. 5, 984-1014.

Inder, B. (1993) Estimating Long-Run Relationships in Economics: A Comparison of Different Approaches. *Journal of Econometrics*

Johansen, S.J. (1988) Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control*, **12**, 231-54.

_____ (1992) Cointegration in Partial Systems and the Efficiency of Single Equation Analysis. *Journal of Econometrics*, **52**, 389-402.

Kremers, J.J.M, N.R. Ericsson and J. Dolado (1992) The Power of Cointegration Tests. *Oxford Bulletin of Economics and Statistics*, **54**: 325-48.

Phillips, P.C.B. (1987) Toward a Unified Asymptotic Theory for Autoregression. *Biometrika*, **74**, 535-547.

_____ and M. Loretan (1991) Estimating Long-Run Economic Equilibria. *Review of Economic Studies*, **58**, 407-36.

_____ and S. Ouliaris (1990) Asymptotic Properties of Residual Based Tests for Cointegration. *Econometrica*, **58**, 165-93.

Saikkonen, P. (1991) Asymptotically Efficient Estimation of Cointegrating Regressions. *Econometric Theory*, **7**, 1-21,

Stock, J.H. and M. W. Watson (1993) A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica*, **61**, 1097-1107.

Urbain, J-P. (1993) *Exogeneity in Error Correction Models* Lecture Notes in Economics and Mathematical Systems No. **398**, Springer-Verlag, Berlin, Germany.

9. Appendix

For the proofs I require the following Lemma taken from Hansen (1995).

Lemma A1 *Let $w_t = \alpha' x_t$ be generated by (12) and assume (13) holds. Then*

1. $T^{-1/2} w_{[Tr]} \Rightarrow a(1)^{-1} B_e^c(r) \equiv a(1)^{-1} \omega_{ee}^{1/2} W_e^c(r)$,

2. $T^{-2} \sum_1^T w_{t-1}^2 \Rightarrow a(1)^{-2} \int_0^1 (B_e^c)^2 \equiv a(1)^{-2} \omega_{ee} \int_0^1 (W_e^c)^2$,

3. $T^{-1} \sum_1^T w_{t-1} \eta_t \Rightarrow a(1)^{-1} \int_0^1 B_e^c dB_\eta \equiv a(1)^{-1} (\omega_{ee}\omega_{\eta\eta})^{1/2} \left(\rho \int_0^1 W_e^c dW_e + (1 - \rho^2)^{1/2} \int_0^1 W_e^c dW_{\eta\cdot e} \right)$,

where $W_{\eta\cdot e}(r)$ is a standard Brownian motion independent of $W_e(r)$.

Proof of Theorems 1 and 2 The proofs use arguments similar to those used in the proof of theorem 2 from Hansen (1995) and are therefore omitted.

Proof of Corollary If $r^2 = 0$ then $\rho^2 = 1/(1 + q^2)$ and the bivariate Brownian motion $B_v(r) = (B_\eta(r), B_e(r))'$ may be decomposed as

$$\begin{pmatrix} B_\eta(r) \\ B_e(r) \end{pmatrix} \equiv \begin{pmatrix} \omega_{\eta\eta}^{1/2}(\rho W_e(r) + (1 - \rho^2)^{1/2} W_{\eta\cdot e}(r)) \\ \omega_{ee}^{1/2} W_e(r) \end{pmatrix} \equiv \begin{pmatrix} \omega_{\eta\eta}^{1/2} W_\eta(r) \\ \omega_{ee}^{1/2}(W_\eta(r) + q W_{e\cdot\eta}(r)) \end{pmatrix},$$

where $W_e(r)$ and $W_{\eta\cdot e}(r)$ are independent standard Brownian motions and $W_\eta(r)$ and $W_{e\cdot\eta}(r)$ are independent standard Brownian motions. The result follows by substituting the latter result into the expression given in Theorem 1. ■

Proof of Theorem 3 The proof is given for case I. The extension to the other cases is straightforward and is therefore omitted. The model (14) may be rewritten as

$$\Delta y_{1t} = \delta_1 w_{A,t-1} + \theta' z_t^* + \eta_t \qquad \text{(A1)}$$

where $z_t^{*\prime} = (x_{2t-1}', z_t')$, $\theta' = (\psi', \zeta')$ and $w_{A,t-1} = w_{t-1} + (\beta - \beta_A)' x_{2t-1}$. Let $Q_Z = I - Z(Z'Z)^{-1}Z'$ for any matrix Z of full rank, and let $W_{A,-1}$, Z^* and η denote the $T \times 1$, $T \times (k+n-1)$ and $T \times 1$ matrices of observations on $w_{A,t-1}$, z_t^* and η_t, respectively. Since $Q_{Z^*} W_{A,-1} = Q_{Z^*} W_{-1}$, partitioned regression on (A1) gives

$$\hat{\delta}_1 = \delta_1 + \left(W_{-1}'Q_{Z^*}W_{-1}\right)^{-1} W_{-1}'Q_{Z^*}\eta, \; SE(\hat{\delta}_1) = \left(\hat{\omega}_{\eta\eta}\left(W_{-1}'Q_{Z^*}W_{-1}\right)^{-1}\right)^{1/2},$$

where $\hat{\omega}_{\eta\eta} = T^{-1}\sum_1^T (\Delta y_{1t} - \hat{\delta}_1 w_{A,t-1} - \hat{\theta}' z_t^*)^2$. Define $D_T = diag(T^1 I_{n-1}, T^{1/2} I_k)$. Note that under strong exogeneity $B_e^c(r) = b(1)'B_2^c(r) + B_\eta^c(r) \equiv \omega_{\eta\eta}^{1/2}(\omega_{\eta\eta}^{-1/2} b(1)'\Omega_{22}^{1/2} W_2^c(r) + W_\eta^c(r)) \equiv \omega_{\eta\eta}^{1/2}(q\cdot a' W_2^c(r) + W_\eta^c(r))$ where $q^2 = b(1)'\Omega_{22} b(1)/\omega_{\eta\eta}$, and a is any $(n-1) \times 1$ vector of unit length. Then using Lemma A1 and the assumption that Δx_{2t} is strongly exogenous the following convergence results can be established:

$$T^{-1/2}\sum_1^{[Tr]} e_t \Rightarrow a(1)^{-1}B_e^c(r) \equiv a(1)^{-1}\omega_{\eta\eta}^{1/2}\left(qa'W_2^c(r) + W_\eta^c(r)\right), \; D_T^{-1}Z^{*\prime}Z^*D_T^{-1} \Rightarrow \begin{pmatrix} \int_0^1 B_2B_2' & 0 \\ 0 & V_Z \end{pmatrix}$$

$$T^{-1}D_T^{-1}Z^{*\prime}W_{-1} \Rightarrow \begin{pmatrix} a(1)^{-1}\int_0^1 B_2B_e^c \\ 0 \end{pmatrix}, \; D_T^{-1}Z^{*\prime}W_{-1} \Rightarrow \begin{pmatrix} \int_0^1 B_2 dB_\eta \\ N(0, \; \omega_{\eta\eta}V_Z) \end{pmatrix}.$$

where $V_Z = plim\; T^{-1}\sum_1^T z_t z_t' > 0$.

Using the above results and Lemma A1 it follows that

$$T^{-2}W_{-1}'Q_{Z^*}W_{-1} \Rightarrow a(1)^{-2}\omega_{\eta\eta}\int_0^1\left(Q_{W_2(r)}[W_\eta^c + qa'W_2^c]\right)^2,$$

$$T^{-1}W_{-1}'Q_{Z^*}\eta \Rightarrow a(1)^{-1}\omega_{\eta\eta}\int_0^1 Q_{W_2(r)}[W_\eta^c + qa'W_2^c]dW_\eta,$$

and so by the continuous mapping theorem (CMT)

$$T\hat{\delta}_1 \Rightarrow -ca(1) + \frac{a(1)\int_0^1 Q_{W_2(r)}[W_\eta^c + qa'W_2^c]dW_\eta}{\int_0^1 \left(Q_{W_2(r)}[W_\eta^c + qa'W_2^c]\right)^2},$$

$$TSE(\hat{\delta}_1) \Rightarrow \left(a(1)^{-2}\int_0^1 \left(Q_{W_2(r)}[W_\eta^c + qa'W_2^c]\right)^2\right)^{1/2}.$$

The desired result follows from the definition of the t-statistic and the CMT. ■

Proof of Theorem 4 The proof is given for case I where $d_t = 0$ and $y_t = x_t$. The extensions to the other cases are straightforward and are thus omitted. The misspecified error correction term may be rewritten as $w_{M,t} = \alpha_M'x_t = \alpha'x_t + (\beta - \beta_M)'x_{2t} = w_t + a(1)^{-1}d$ and under the local alternative (13) the true model may be expressed as

$$\Delta x_{1t} = \delta_1 w_{M,t-1} + \psi' x_{2,t-1} + \zeta' z_t + \eta_t,$$

where $\delta_1 = -a(1)c/T$ and $\psi = -cd/T$ are local to zero. The estimated model is (17) and partitioned regression gives

$$\hat{\delta}1 = \delta_1 + (W_{M,-1}'Q_Z W_{M,-1})^{-1}W_{M,-1}'Q_Z X_{2,-1}\psi + (W_{M,-1}'Q_Z W_{M,-1})^{-1}W_{M,-1}'Q_Z\eta,$$

$$SE(\hat{\delta}_1) = \left(\hat{\omega}_{\eta\eta}(W_{M,-1}'Q_Z W_{M,-1})^{-1}\right)^{1/2}, \; \hat{\omega}_{\eta\eta} = T^{-1}\sum_1^T(\Delta x_{1t} - \hat{\delta}_1 w_{M,t-1} - \hat{\zeta}'z_t)^2.$$

Define the stochastic process $Z^c(q,s;r) = q(a'W_2^c(r)) - s(a'W_2(r)) + W_\eta(r)$ where $q^2 = b(1)'\Omega_{22}b(1)/\omega_{\eta\eta}$, $s^2 = d'\Omega_{22}d/\omega_{\eta\eta}$ and a is any $(n-1) \times 1$ vector of unit length. Then using Lemma A1 and the assumption of strong exogeneity the following convergence results can be established:

$$T^{-1/2}w_{M,[Tr]} \Rightarrow a(1)^{-1}\omega_{\eta\eta}^{1/2}Z^c(q,s;r), \;\; T^{-2}\sum_1^T w_{M,t-1}^2 \Rightarrow a(1)^{-2}\omega_{\eta\eta}\int_0^1 Z^c(q,s)^2$$

$$T^{-1}\sum_1^T w_{M,t-1}\eta_t \Rightarrow a(1)^{-1}\omega_{\eta\eta}\int_0^1 Z^c(q,s)dW_\eta, \; T^{-2}\sum_1^T w_{M,t-1}x_{2t-1}' \Rightarrow a(1)^{-1}\omega_{\eta\eta}^{1/2}\int_0^1\left(\Omega_{22}^{1/2}W_2 Z^c(q,s)\right)'.$$

Using the above results it follows that

$$T^{-2}W_{M,-1}'Q_Z W_{M,-1} = T^{-2}\sum_1^T w_{M,t-1}^2 + o_p(1) \Rightarrow a(1)^{-2}\omega_{\eta\eta}\int_0^1 Z^c(q,s)^2,$$

$$T^{-1}W_{M,-1}^{\prime}Q_Z X_{2,-1}\psi = -cT^{-2}\sum_1^T w_{M,t-1}x_{2,t-1}d + o_p(1) \Rightarrow -ca(1)^{-1}\omega_{\eta\eta}^{1/2}d^{\prime}\Omega_{22}^{1/2}\int_0^1 Z^c(q,s)W_2,$$

$$T^{-1}W_{M,-1}^{\prime}Q_Z\eta = T^{-1}\sum_1^T w_{M,t-1}\eta_t + o_p(1) \Rightarrow a(1)^{-1}\omega_{\eta\eta}\int_0^1 Z^c(q,s)dW_\eta.$$

and so by the CMT

$$T\hat{\delta}_1 \Rightarrow -ca(1) - ca(1)s\frac{\int_0^1 Z^c(q,s)(a^{\prime}W_2)}{\int_0^1 Z^c(q,s)^2} + a(1)\frac{\int_0^1 Z^c(q,s)dW_\eta}{\int_0^1 Z^c(q,s)^2},$$

$$TSE(\hat{\delta}_1) \Rightarrow a(1)\left(\int_0^1 Z^c(q,s)^2\right)^{-1/2}.$$

The desired result follows from the definition of the t-statistic and the CMT. ■

Table 1

Trend Parameters under Cointegration and Weak Exogeneity

$y_t = d_t + x_t$, $d_t = \gamma + \theta t$, $\delta_2 = 0$ (weak exogeneity)

	Case I	Case II	Case III	Case IV
Trend Parameter	$\gamma=\theta=0$	$\gamma \neq 0, \theta = 0$ γ unrestricted	$\gamma \neq 0, \theta \neq 0, \alpha'\theta = 0$	$\gamma \neq 0, \theta \neq 0$ γ, θ unrestricted
μ	0	$\delta\alpha'\gamma$	$(I-\Gamma(1))\theta - \delta\alpha'\gamma$	$(I-\Gamma(1))\theta + \delta\alpha'\theta - \delta\alpha'\gamma$
μ_1	0	$\delta_1\alpha'\gamma$	$(1-\Gamma_{11}(1))\theta_1 - \Gamma_{12}'\theta_2 - \delta_1\alpha'\gamma$	$(1-\Gamma_{11}(1))\theta_1 - \Gamma_{12}'\theta_2 + \delta_1\alpha'\theta - \delta_1\alpha'\gamma$
μ_2	0	0	$(I-\Gamma_{22}(1))\theta_2 - \Gamma_{21}(1)\theta_1$	$(I-\Gamma_{22}(1))\theta_2 - \Gamma_{21}(1)\theta_1$
$\mu_{1\cdot2}$	0	μ_1	$\mu_1 - \phi'\mu_2$	$\mu_1 - \phi'\mu_2$
τ	0	0	0	$-\delta\alpha'\theta$
τ_1	0	0	0	$-\delta_1\alpha'\theta$
τ_2	0	0	0	0
$\tau_{1\cdot2}$	0	0	0	τ_1