Parameter Constancy in Cointegrating Regressions

CARMELA E. QUINTOS  
Department of Economics, Yale University, P.O. Box 2125 Yale Station,  
New Haven, Connecticut 06520-2125, USA  
PETER C. B. PHILLIPS  
Cowles Foundation for Research in Economics, Yale University, P.O. Box 2125 Yale Station,  
New Haven, Connecticut 06520-2125, USA

Abstract: This paper proposes an approach to testing for coefficient stability in cointegrating regressions in time series models. The test statistic considered is the one-sided version of the Lagrange Multiplier (LM) test. Its limit distribution is non-standard but is nuisance parameter free and can be represented in terms of a stochastic bridge process which is tied down like a Brownian bridge but relies on a random rather than a deterministic fraction to do so. The approach provides a test of the null hypothesis of cointegration against specific directions of departure from the null; subset coefficient stability tests are also available. A small simulation studies the size and power properties of these tests and an empirical illustration to Australian data on consumption, disposable income, inflation and money is provided.

JEL Classification System-Numbers: C22, C32

1 Introduction

A considerable number of statistical procedures that test for the presence of cointegration are available in the literature. The most commonly used methods in empirical research are the residual-based approaches, such as the augmented Dickey-Fuller test of Dickey-Fuller (1979, 1981) and the semiparametric Z-tests developed by Phillips (1987), and Johansen’s (1988) likelihood ratio test for cointegration in vector autoregressions. The asymptotic properties of these tests are now well known, and the reader is referred to Phillips and Ouliaris (1990) for a detailed study of the various residual-based tests. All of these tests work from a null hypothesis of no cointegration. This approach has been criticized (see, for example, Engle and Yoo (1991)) because it is usually the alternative (i.e., cointegration) that is of primary interest in applications and the mechanism by which classical hypothesis testing is carried out ensures that the null hypothesis (here no cointegration) is accepted unless there is strong evidence against it.

In this paper we propose an approach to testing for parameter constancy and, hence, cointegration in regressions with nonstationary regressors. Our tests are extensions of the Lagrange Multiplier (LM) and locally best invariant (LBI) tests that are presently used in the literature to test for parameter constancy.

---

Footnote: 1 Our thanks go to Bruce Hansen for sending us a copy of his related work (1992a) and for the use of his GAUSS procedure for computing the fully modified least squares estimator.
in regressions with stationary regressors—see King-Hillier (1985), Leybourne-McCabe (1989) and Nabeya-Tanaka (1988). In our case, the parameter that is the focus of attention is a cointegrating coefficient, or as the case may be, a cointegrating submatrix in multivariate regressions. The approach allows for random evolution of the parameter(s) under the alternative hypothesis. The subset tests are especially useful because they provide a means of isolating the source of cointegration failure that is not available in residual-based tests.

The limit theory of our tests is developed and is characterized by a stochastic bridge process that resembles a Brownian bridge. However, the limiting bridge process in this case is tied down at its upper end by a random (as opposed to a deterministic) fraction of the interval [0, 1]. The random fraction is itself a stochastic process whose features depend on the stochastic properties and the number of cointegrating regressors.

The present paper is related to recent work by Tanaka (1993) and Hansen (1992a) that came to our attention when our work was nearly completed. Tanaka applies the LBI test to the coefficient of the nonstationary regressor that causes the failure of cointegration under the alternative, whereas we take the coefficient to be random and apply an LM test to the variance of the error process that governs the random parameter. In this sense, our approach accords more with the formulation of parameter constancy tests in the literature. Hansen (1992a) develops a limit theory for the LM test that corresponds with our own approach. His test of constancy applies to the full matrix of cointegrating coefficients and his test for cointegration is based on the constancy of the intercept in the regression. Our LM tests of coefficient constancy are developed and applied to submatrices of the cointegrating matrix as well as the full cointegrating matrix. The advantage of this formulation is that in empirical work, tests of the constancy of individual coefficients can be conducted, thereby isolating the variables that are responsible for failure in the null hypothesis.

The plan of the paper is as follows. Section 2 outlines the models under consideration and the assumptions required for the construction of our tests. The formulation of the test and its limiting distribution are given in Section 3. Section 4 presents multivariate extensions, specializations to cases where the regressors contain trends and subset tests. Some Monte Carlo experiments including size calculations and power comparisons are given in Section 5. Section 6 reports a brief empirical illustration of our methods and Section 7 gives some concluding remarks.

2 Models and Assumptions

Let \( \{x_t, y_t\} \) be \( I(1) \) processes and consider the following varying coefficient regression

\[
y_t = \beta_t x_t + \varepsilon_{0t}
\]  

(1)
\( A x_t = \varepsilon_t \) \hspace{2cm} (2)

\( \beta_t = \beta_{t-1} + \eta_t, \quad \eta_t \equiv \text{iid } N(0, \Sigma) \) \hspace{2cm} (3)

where we have assumed that the parameter variation follows a Gaussian random walk process, allowing the model to accommodate fundamental changes in structure over time. For reasons that will be made clear later we require that the initial condition \( \beta_0 \) be nonzero. The initial conditions for \( \{x_t, y_t\} \) do not affect the subsequent theory so we allow \( (x_0, y_0) \) to be any random variable.

Throughout the paper, we assume the sequence \( \{\varepsilon_t\} = (\varepsilon_{0t}, \varepsilon_{1t})' \) has mean zero and satisfies an invariance principle. More precisely, it is assumed that for \( r \in [0, 1] \)

\[
S_{[m]} = n^{-1/2} \sum_{i=1}^{[m]} \varepsilon_i \overset{d}{\to} B(r) = \begin{pmatrix} B_0(r) \\ B_1(r) \end{pmatrix} \hspace{2cm} (4)
\]

where \( B(r) \) is a scaled Brownian motion and \( B_0 \) and \( B_1 \) need not be independent. In (4) and elsewhere in the paper, \([m]\) denotes the largest integer that does not exceed \( m \).

The invariance principle (4) holds under very general conditions and is valid for the sequence \( \{\varepsilon_t\} \) driven by a large class of models with varying degrees of heterogeneity and memory restrictions. Explicit conditions under which (4) holds are discussed in detail in earlier work by Phillips-Durlauf (1986) and Phillips-Solo (1992). When \( \{\varepsilon_t\}^\infty \) is weakly stationary as would be the case when (1) and (2) form a cointegrated system, the following mixing and moment conditions, for example, are sufficient

**Assumption 1:**

i) \( E|\varepsilon_1|^p < \infty \) for some \( 2 \leq \rho < \infty \)

ii) either \( \sum_{m=1}^{\infty} \phi_m^{1-1/\rho} < \infty \), or \( \rho > 2 \) and \( \sum_{m=1}^{\infty} \phi_m^{1-2/\rho} < \infty \), where \( \phi_m \) and \( \sigma_m \) are the uniform mixing and \( \alpha \)-mixing coefficients for \( \{\varepsilon_t\} \) respectively.

This ensures that

\[
\Psi = \lim_{n \to \infty} E(n^{-1} S_n S_n') = E(\varepsilon_{11})' + \sum_{k=2}^{\infty} \{E(\varepsilon_{1k} \varepsilon_{k1}') + E(\varepsilon_{k1} \varepsilon_{1k}')\}
\]

is finite. When we work with Assumption 1, we presume none of the common exogeneity conditions for \( x_t \) and we allow for contemporaneous correlation of the form \( E(x_t \varepsilon_{0t}) \neq 0 \).

The long-run covariance matrix \( \Psi \) of the limit Brownian motion \( B \) in (4) is partitioned as

\[
\Psi = \begin{pmatrix} \Psi_{00} & \Psi_{01} \\ \Psi_{10} & \Psi_{11} \end{pmatrix} = \lim_{n \to \infty} n^{-1} E(\Sigma_n \varepsilon_t)(\Sigma_n \varepsilon_t)'
\]

\[
= \Sigma + A + A'
\]
where

\[ \Sigma = \lim_{n \to \infty} n^{-1} \Sigma_1^n E(e_i e'_i) = \begin{pmatrix} \sigma^2 & \Sigma_{01} \\ \Sigma_{10} & \Sigma_{11} \end{pmatrix} \]

\[ A = \lim_{n \to \infty} n^{-1} \sum_{i=2}^{n} \sum_{j=1}^{i-1} E(e_i e'_i) = \begin{pmatrix} \lambda_{00} & A_{01} \\ A_{10} & A_{11} \end{pmatrix} \]

with the partitions made conformable to that of the long-run variance matrix $\Psi$. If we require \{\epsilon_i\} to have continuous spectral density matrix $f_\epsilon(\lambda)$ then $\Psi$ is simply $f_\epsilon(\cdot)$ evaluated at the zero frequency, i.e. $\Psi = 2\pi f_\epsilon(0)$, the long-run covariance matrix of $\epsilon_i$. If (1) and (2) form a cointegrated system, then $\Psi$ is singular although we require $\Psi_1$ to be invertible. The model for which $\Psi_1$ is singular presents additional complications (see Phillips (1991b)) and will be considered in later work. For later discussion we define

\[ \gamma = \Sigma + A = \begin{pmatrix} Y_{00} & Y_{01} \\ Y_{10} & Y_{11} \end{pmatrix} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix}. \quad (6) \]

If \{\epsilon_i\} is iid then of course $\Psi = \Sigma = \gamma$.

As a direct extension of (1) we also consider time series \{\gamma_i\} that are generated by

\[ y_i = \beta_i x_i + \Gamma h_i + \epsilon_{0i}, \quad (1') \]

where $h_i$ is a deterministic function of time of order $p$. For example, $h_i$ may consist of a constant ($p = 0$) or a time trend ($p = 1$).

The assumptions that allow $h_i$ suitably standardize by a power of $n^{-1/2}$ to converge to a limit function $h(r)$ are discussed in Phillips and Hansen (1990) and Hansen (1990). Here we simply assume the existence of a diagonal matrix of weights $\gamma$ such that $\gamma_n h [n r] \to h(r)$ uniformly in $r$. For example if $h_i$ consists of a simple time trend ($i$), $\gamma_n = n^{-1}$ and $h(r) = r$.

In order to understand the LM approach to testing parameter constancy we first derive the test and its limit distribution under a set of simplifying assumptions. Specifically, we strengthen Assumption 1 to:

**Assumption 2:**

i) \{\epsilon_i\} = iid

ii) \{x_i\} is strictly exogenous.

As we shall see, extending the results to models that allow for stationary errors requires (possibly semiparametric) corrections to remove the bias introduced by the serial correlation and endogeneity of \{\epsilon_i\} and \{x_i\}. 
3 The LM Test Statistics and their Limiting Distributions

We are interested in testing for the constancy of the coefficient $\beta_0$ in (1) and (1').
In order to understand the behavior of the LM test, we begin with a single $I(1)$ regressor $x_t$ and denote by $\psi$ its long-run covariance matrix given in (5). The multivariate extension for $x_t$ of dimension $k$ will be dealt with later in the paper in which case we will retain the matrix notation $\Psi$ for the long-run covariance matrix. By backward substitution of (3) we write model (1) as

$$y_t = \beta_0 x_t + \sigma^2 \eta_t + \varepsilon_t$$

and (1') as

$$y_t = \beta_0 x_t + \Gamma h_t + w_t$$

then a test for constancy reduces to a zero restriction on the variance of the innovations that drive the random parameter, $\sigma^2$. We write this as

$$H_0: \sigma^2 = 0$$

which we test against the alternative $H_a: \sigma^2 > 0$. Throughout the paper we assume that $\{x_t\}$ satisfies Assumptions 1 or 2 so that under $H_0$, (7) is a constant coefficient cointegrating regression and (7)' is a constant coefficient cointegrating regression around a trend. Under $H_a$, (7) and (7)' are not cointegrated systems since $w_t$ is non-stationary. Thus a test for constancy of the long-run coefficient can be interpreted as testing the null of cointegration.

Before we derive the LM test statistic we summarize some limit theory for the least squares estimators in (7) $\hat{\beta}_0$ and (7)' $\hat{\beta}_0'$. All integrals, such as $\int B^2$, are taken with the limits $[0, 1]$ in what follows to simplify notation.

Lemma 1: (Park and Phillips (1988)). For the model (7) under $H_0$

(i) $n(\hat{\beta}_0 - \beta) \overset{d}{\rightarrow} (\int B^2)^{-1}(\int B_1 dB_0 + \tau_{10})$

where $\tau_{10} = Y_{10}$, as defined in (6) and $B_1(r) = (\sigma^2)^{1/2}v_t$, $B_0 = (\sigma^2)^{1/2}w_0$ where $v_t$ and $w_0$ are independent standard Brownian motions.

Lemma 2: (Phillips and Hansen (1990)). For the model (7)'

i) $n(\hat{\beta}_0 - \beta) \overset{d}{\rightarrow} (\int B_{1h}^2)^{-1}(\int B_{1h} dB_0 + \tau_{10})$

where $B_{1h} = B_1(r) - \int B_1 h'(\int h'h')^{-1}(H) dB_0 + \tau_{10}$

where $h(r) = h(r) - (\int hB_1)(\int B_1^2)^{-1}B_1(r)$ and $\tau_{10} = -\int hB_1r(\int B_1^2)^{-1}r_{10}$.

$\square$
To derive the LM test we define \( y = (y_1, \ldots, y_n)' \), \( x = (x_1, \ldots, x_n)' \), and \( w = (w_1, \ldots, w_n)' \). The residual vector is \( w = y - \beta_0 x \) and under Assumption 2 the likelihood conditional on \( x \) is

\[
\mathcal{L} = -\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln |\Omega| - \frac{1}{2} w' \Omega^{-1} w
\]

(9)

where \( \Omega = \text{var}(w) = \sigma_0^2 I_n + \sigma_0^2 D_x L L' D_x' \), \( D_x = \text{diag}(x_1, \ldots, x_n) \),

\[
L = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}
\]

and \( LL' \) is the Choleski decomposition of \( V = \{\min(i, j)\}_{n \times n} \). The score vector with respect to \( \sigma_0^2 \) is

\[
x_i = -\frac{1}{2} \text{tr}(\Omega^{-1} A) + \frac{1}{2} w' \Omega^{-1} A \Omega^{-1} w
\]

(10)

where we have set \( A = D_x L L' D_x' \). For testing purposes we concentrate on the stochastic part of the score \( s_i^* = w' \Omega^{-1} A \Omega^{-1} w \) and observe that under \( H_0 \), \( s_i^* = (\sigma_0^2)^{-2} w' A w \) and \( \text{var}(s_i^*) = (\sigma_0^2)^{-2} \text{tr}(A^2) \). Excluding the scale factor \( \text{tr}(A^2) \) the LM test is given by

\[
LM = s_i^*/[\text{var}(s_i^*)]^{1/2} = \tilde{w}' D_x L L' D_x' \tilde{w}/\tilde{\sigma}_0^2
\]

(11)

where \( \tilde{w} = y - \tilde{\beta}_0 x \) and \( "\cdots" \) denotes the restricted ML estimator. Observe that the test has the same form as the locally best invariant test of Nabeya and Tanaka (1988) and of Leybourne and McCabe (1989). This is so because we are working with the conditional likelihood so the variance of the error structure, \( \Omega \), is of the same form in either case. However, the asymptotics of the test are quite different as we now show.

The asymptotic distribution of the LM test is derived by noticing that premultiplying a vector by \( L' \) has the effect of forming a backwards cumulative sum. Tests based on either the forward or backward sums are equal in distribution and we can write

\[
\tilde{w}' D_x L L' D_x' \tilde{w} = \sum_{k=1}^{n} [\Sigma^k x_s \tilde{w}_s - \Sigma^k x_s \tilde{w}_s]^2
\]

\[
= \sum_{k=1}^{n} [\Sigma^{k-1} x_s \tilde{w}_s]^2
\]

(12)

since \( \Sigma^k x_s \tilde{w}_s = 0 \) by the OLS normal equations. Set \([nr] = k - 1\) and using Lemma 1 we have
\[
\tilde{z}_n(r) = n^{-1} \sum_{i}^{n} x_i \tilde{w}_i = n^{-1} \sum_{i}^{n} x_i (y_i - \tilde{\beta}_0 x_i) \\
= n^{-1} \sum_{i}^{n} x_i w_i - n(\tilde{\beta}_0 - \beta_0)n^{-2} \sum_{i}^{n} x_i^2 \\
\overset{d}{\to} \int_{0}^{r} B_1 \, dB_0 - \int_{0}^{r} \frac{\sigma_1^2}{\sigma_0^2} \frac{1}{\sqrt{\int_{0}^{r} V_1 \, dW_0}} \int_{0}^{r} B_1 \, dB_0 \\
= (\sigma_1^2/\sigma_0^2)^{1/2} \left( \int_{0}^{r} V_1 \, dW_0 - \int_{0}^{r} \frac{V_1^2}{\int_{0}^{r} V_1^2} \int_{0}^{r} V_1 \, dW_0 \right)
\]

(13)

where \( V_1 \) and \( W_0 \) are independent standard Brownian motion. Squaring \( \tilde{z}_n(r) \) and dividing by consistent estimates of \( \sigma_1^2 \) and \( \sigma_0^2 \) yields an LM statistic with a limiting distribution that is free of nuisance parameters. The results are summarized in the following theorem.

**Theorem 3.1:** The LM statistic for parameter constancy that tests a null of cointegration in model (7) satisfies

\[
LM_1 = n^{-3} \bar{w}' D_x L' D_x \bar{w}/\hat{\sigma}_0^2 \overset{d}{\to} \int_{0}^{1} W_{\tilde{r}}(r)^2 \, dr,
\]

where

\[
W_{\tilde{r}}(r) = \int_{0}^{r} V_1 \, dW_0 - \int_{0}^{r} \frac{V_1^2}{\int_{0}^{r} V_1^2} \int_{0}^{r} V_1 \, dW_0,
\]

\[
\hat{\sigma}_0^2 = n^{-1} \bar{w}' \bar{w} \overset{p}{\to} \sigma_0^2,
\]

and

\[
\hat{\sigma}_1^2 = n^{-1} \sum_{i}^{n} A x_i^2 \overset{p}{\to} \sigma_1^2.
\]

This is a simple 'cointegrated system' extension of results in Nabeya-Tanaka (1988), Leybourne-McCabe (1989) and Kwiatkowski-Phillips-Schmidt-Shin (1992), where the limit distribution of the LM test is shown to be a functional of a Brownian bridge. The nature of the above extension is that, in a cointegrated system, the behavior of the test asymptotically has the form of a generalized Brownian bridge that is tied down at the fraction \( r = 1 \) but is scaled by a random fraction. Note that if \( V_1 = 1 \) we would have \( W_{\tilde{r}}(r) = \int_{0}^{r} dW_0 - r \int_{0}^{1} dW_0 = W_0(r) - r W_0(1) \), a Brownian bridge with deterministic fraction \( r \). We call \( W_{\tilde{r}}(r) \) a \( V \)-based Brownian bridge which, like a Brownian bridge, is tied down at \( r = 1 \) to zero (i.e. \( W_{\tilde{r}}(1) = 0 \)) but has a random fraction \( J(r) = \int_{0}^{r} V_1^2/\int_{0}^{1} V_1^2 \).

The effect of the additional random walk component \( V_1 \) and random fraction \( J(r) \) on the LM test is reflected in the critical values obtained from the distributions constructed from either functionals. This is illustrated in Table 1 where it is shown that the LM test behaves very differently when it is constructed from a \( V \)-based and a standard Brownian bridge. The critical values given there are the
Table 1. Upper critical values of $LM_1$ test

<table>
<thead>
<tr>
<th></th>
<th>Size $\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.01</td>
</tr>
<tr>
<td>$v$-based</td>
<td>.4898</td>
</tr>
<tr>
<td>standard</td>
<td>.7868</td>
</tr>
</tbody>
</table>

Notes: Data generated with iid $N(0,1)$ errors.
Replication times $= 10,000$ and sample size $= 250$.

upper tail percentage points of the asymptotic density, computed by Monte Carlo simulations using $n = 250$.

Now suppose we are interested in testing for cointegration in a model with an included trend. When using residual-based tests in testing the null of no cointegration (see Phillips and Ouliaris (1990) and Hansen (1992b) for a study of these tests) this is of course equivalent to estimating the trend coefficient and then using it to form the residuals. In our case we write the residuals as $\hat{r}_t^* = y_t - \hat{\beta}_0 x_t - \hat{h}_t y_t$ and using Lemma 2 we get

$$\hat{\xi}_n^*(r) = n^{-1} \sum_{i} x_t w_i - n(\hat{\beta}_0 - \beta_0) n^{-2} \sum_{i} x_i^2 - n^{-1}(\hat{\beta}_0 - \beta_0) n^{-2} \sum_{i} x_i h_i y_i$$

where $Q(r) = (B_1(r), h(r))$. Note that

$$B_0(1) = \int_0^1 B_1 \, dB_0 - \left( \int_0^1 B_1 \, Q \right) \left( \int_0^1 Q Q' \right)^{-1} \int_0^1 Q \, dB_0 = \int_0^1 B_1 \, dB_0 = 0,$$

since $B_1(r) = B_1(r) - \int_0^1 B_1 \, Q \left( \int_0^1 Q Q' \right)^{-1} Q(r) = 0$ is the $L_2$ projection residual of $B_1$ on the space orthogonal to $Q(r)$ (which of course, contains $B_1(r)$ in its span) – see Phillips (1988a) and Park-Phillips (1988) for a discussion of $L_2$ Hilbert projections in this context. Thus $B_0(r)$ is a tied-down process.

The asymptotic behavior of parameter constancy tests in a model with integrated and trend regressors is summarized in the following Theorem.

**Theorem 3.2:** The $LM$ statistic for parameter constancy that tests the null of cointegration with trend behaves in the limit as

$$LM_2^* = n^{-3} \hat{\lambda}^w \hat{D}_1 LL'D_1 \hat{\lambda}^w / \hat{\sigma}^2 \hat{\sigma}_0^2 \to \int_0^1 W_0(r)^2 \, dr.$$

Consistent estimators $\hat{\sigma}^2$ and $\hat{\sigma}_0^2$ are constructed as in Theorem 3.1. $\square$

The results in Theorems 3.1 and 3.2 hold even with the addition of a stationary, zero mean exogenous regressor. This is because the coefficient of the $I(1)$
regressor, $\hat{\beta}_0$ (under $H_0$) is estimated consistently at rate $O_p(n^{-1})$ as opposed to the coefficient of the stationary component which is estimated consistently at rate $O_p(n^{-1/2})$. To see this consider the model

$$ y_t = \hat{\beta}_0 x_t + \pi z_t + \epsilon_{1t} $$
$$ Ax_t = \epsilon_{1t}, \quad z_t = \epsilon_{2t} $$

where we assume $\{\epsilon_t\} = (\epsilon_{0}, \epsilon_{1t}, \epsilon_{2r})$ is iid and $z_t$ and $x_t$ are strictly exogenous. Under $H_0$, the asymptotic behavior of the LS estimator $\hat{\pi}$ is

$$ n^{1/2}(\hat{\pi} - \pi) \overset{d}{\rightarrow} N(0, \sigma_0^2 \sigma_{22}^{-1}, \sigma_{22}^{-1}) $$

The regression residual is $\hat{w}_t = w_t - (\hat{\beta}_0 - \beta_0) x_t - (\hat{\pi} - \pi) z_t$ and since the last term is $O_p(n^{-1/2})$ we have

$$ \hat{\zeta}_n(t) \overset{d}{\rightarrow} \int_0^t B_i \, dB_0 - \frac{1}{\int_0^t B_i^2} \int_0^t B_i \, dB_0 $$

as for $\zeta_n$ in model (1), Theorem 3.1. This result is in contrast to that of Leybourne–McCabe (1989) who consider the case with stationary regressors, in which case the distribution of the test depends on the dimensions of the exogenous regressors as well. By similar calculations, $\hat{\zeta}_n$ converges weakly to the same limits as those of $\zeta^*_n$ used in Theorem 3.2.

3 Time Series Extensions

Extensions to the time series case with the error process $\{\epsilon_t\}$ satisfying Assumption 1 is straightforward. It requires the use of an optimal estimator such as the fully modified least squares estimator developed by Phillips and Hansen (1990). Define

$$ \hat{\gamma} = y - \hat{\psi}_{11} \hat{\psi}_{11}^{-1} Ax $$
$$ w_{0-1} = w - \hat{\psi}_{11} \hat{\psi}_{11}^{-1} \epsilon_1. $$

The fully modified $OLS$ estimator in the single equation (7) (under $H_0$) is given by

$$ \hat{\beta}_0 = (X'X)^{-1}(X'\hat{\gamma} - n\delta^+) $$

where $\delta^+ = \hat{\gamma}_i - \hat{\psi}_{11}^{-1} \hat{\psi}_{10}^{'}$. From Phillips and Hansen the corrected estimator (16) has a compound normal limit distribution free of nuisance parameters

$$ n(\hat{\beta}_0 - \beta_0) \overset{d}{\rightarrow} \left( \int B_i^2 \right)^{-1} \left( \int B_i \, dB_{0-1} \right) $$

where $B_0 = B_0 - \hat{\psi}_{11} \hat{\psi}_{11}^{-1} B_i$ and the limit process $B_i$ and $B_{0-1}$ are independent of each other.
The limit theory for the LM test is derived following the arguments of the previous section using the optimal estimator (16). Write the corrected residual \( \tilde{w}_s \) as

\[
\tilde{w}_s^+ = \tilde{y}_s^+ - \tilde{\beta}_0^+ x_s = w_s - (\tilde{\beta}_0^+ - \beta_0) x_s - \tilde{\psi}_1 \hat{\psi}_1^{-1} \Delta x_s
\]

and using (17),

\[
\zeta_n^+(r) = n^{-1} \sum_{t=1}^{[nr]} x_t \tilde{w}_s^+
\]

\[
= n^{-1} \sum_{t=1}^{[nr]} x_t w_s - n(\tilde{\beta}_0^+ - \beta_0)n^{-2} \sum_{t=1}^{[nw]} x_t^2 - \tilde{\psi}_1 \hat{\psi}_1^{-1} n^{-1} \sum_{t=1}^{[nr]} (x_t \Delta x_t)
\]

\[
\overset{d}{\to} \left\{ \int_0^r B_1 \ dW_{0,1} - \frac{f_0 B_1^2}{\int_0^1 B_1^2} \int_0^r B_1 \ dW_{0,1} \right\} + r \delta^*
\]

\[
= (\psi_{0,1}, \psi_1)^{1/2} W_{\psi^+}(r) + r \delta^*
\]

where \( B_1(r) = \psi_{1}^{1/2} V_1(r) \). Observe that the limit of \( \zeta_n^+(r) \) contains a nuisance parameter \( \delta^* \). This leads us to suggest the following modifications to the cumulative sum \( \zeta_n^+ \). Let

\[
\tilde{\zeta}_n^+(r) = n^{-1} \sum_{t=1}^{[nr]} x_t \tilde{w}_s^+ = \zeta_n^+(r) - r \delta^*
\]

then we have a nuisance parameter free distribution,

\[
\tilde{\zeta}_n^+(r) \overset{d}{\to} (\psi_{0,1}, \psi_1)^{1/2} W_{\psi^+}(r) - \int_0^r V_1 \ dW_{0,1} - \frac{f_0 B_1^2}{\int_0^1 B_1^2} \int_0^r V_1 \ dW_{0,1}
\]

(20)

This gives us an extension of Theorem 3.1 to more general models.

**Corollary 3.1:** The limiting distribution of the LM statistic that tests the null of cointegration is

\[
LM_1^+ = n^{-3} \tilde{w}^{+++} D_s LL' D_s \tilde{w}^{+++} / \hat{\psi}_{0,1} \hat{\psi}_1 \overset{d}{\to} \int_0^1 W_{\psi^+}(r)^2 \ dr
\]

where \( D_s \tilde{w}^{+++} \) and \( W_{\psi^+}(r) \) are given in (19) and (20) respectively and \( \hat{\psi}_{0,1} \) and \( \hat{\psi}_1 \) are consistent semiparametric estimators for \( \psi_{0,1} \) and \( \psi_1 \) (see Phillips (1987), Newey-West (1987) and Andrews (1991) for a discussion of possible estimators). 

The result of Theorem 3.2 can also be extended to time series models by similar arguments. We will not go through the derivations but give the extensions in the results below. We use the fully modified LS estimators of the parameters in (7'). These are:
\[
\hat{\beta}_0^+ = \left( \sum_{i=1}^{n} x_i^2 \right)^{-1} \left( \sum_{i=1}^{n} x_i \hat{y}_i^+ - n \delta^+ \right) \quad \text{and} \quad \\
\hat{I}^+ = \left( \sum_{i=1}^{n} h_i h_i \right)^{-1} \left( \sum_{i=1}^{n} h_i \hat{y}_i^+ - n \delta^k \right)
\]

where \( x_h = x_i - \sum_{i=1}^{n} x_i h_i \left( \sum_{i=1}^{n} h_i h_i \right)^{-1} h_i \), \( h_i = h_i - \sum_{i=1}^{n} h_i x_i \left( \sum_{i=1}^{n} x_i^2 \right)^{-1} x_i \), and \( \delta^h = \hat{\delta}_0^{10} - \hat{\psi}_0 \hat{\psi}_1 \hat{\psi}_1^{-1} \hat{\delta}_1^{10} \).

Corollary 3.2: Let \( \hat{\omega}_s^{**} = \hat{y}_s^+ - \hat{\beta}_0^+ x_s - \hat{I}^+ h_s \) and set
\[
\zeta_n^{***}(r) = n^{-1} \left[ x_i \hat{\omega}_s^{***} \right] - n^{-1} \left[ x_i \hat{\omega}_s^{**} \right] - r \delta^+.
\]

Then we have
\[
LM_2^{**} = n^{-3} \hat{\omega}^{***} D_x LL'D_x \hat{\omega}^{***} / \hat{\psi}_{0,1} \hat{\psi}_1 \rightarrow_d \int_0^1 W_Q^+(r)^2 dr
\]

where \( W_Q^+(r) = \int_0^r V_1 dW_{0,1} - (\int_0^r V_1 Q (\int_0^r Q)^{-1} \int_0^r Q) dW_{0,1} \).

Observe that the extension to the time series case consists of correcting for both serial correlation and endogeneity in \( \hat{\beta}_0 \) and \( D_x \hat{\omega} \), and by replacing \( \hat{\sigma}_0^2 \) and \( \hat{\sigma}^2 \) by the long-run conditional variance \( \hat{\psi}_{0,1} \) and long-run variance \( \hat{\psi}_1 \) respectively.

4 Multivariate Extensions

We now extend the results of the previous section to the case where the regressor \( \{x_i\} \) is a column vector of order \( k \), \( \{x_i\} = (x_{1i}, \ldots, x_{ki}) \). Consider first the case where the innovations are iid. We have

\[
y_i = \beta' x_i + \varepsilon_{0i} \quad (1)''
\]

\[
\Delta x_i = \varepsilon_{1i}, \quad \varepsilon_{1i} \equiv iid N(0, \Sigma_{11}) \quad (2)''
\]

\[
\hat{\beta}_i = \hat{\beta}_{i-1} + \eta_i, \quad \eta_i \equiv iid N(0, \Sigma_{\eta}) \quad (3)''
\]

which we write as

\[
y_i = \beta_0' x_i + \omega_i \quad (7)''
\]

with \( \beta_0 = (\beta_{01}, \ldots, \beta_{0k})' \) and
\[ w_i = \left( \sum_{k} \eta_k \right) x_i + e_{oi} = x_i'(I_k, I_k, I_k, \ldots, I_k, 0, \ldots, 0) \eta + e_{oi} \]

\[ = D_x \eta + e_{oi} \tag{22} \]

where \( \eta = (\eta_1, \ldots, \eta_k)' \), \( D_x = \text{diag} (x_1', x_2', \ldots, x_k') \) and \( L = \begin{bmatrix} I_k & 0 & \cdots & 0 \\ I_k & I_k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ I_k & I_k & \cdots & I_k \end{bmatrix} \).

The likelihood is the same as before but now \( \Omega = \text{var}(w) = \sigma_0^2 I_n + D_x \Sigma L(I_n \otimes \Sigma \otimes L)' D_x \).

The differential of the likelihood is

\[ d \mathcal{L} = -(1/2) \text{tr}(\Omega^{-1} d \Omega) + (1/2) w'(\Omega^{-1} d \Omega \Omega^{-1} w - \text{tr}(L d \Sigma_w)) \tag{23} \]

with

\[ d \Omega = D_x \Sigma L(I_n \otimes \Sigma \otimes L)' D_x \tag{24} \]

and \( L \) is the matrix of Lagrange multipliers. Evaluating at the restricted elements and noting that \( \Omega = \sigma_0^2 I \) we have

\[ d \mathcal{L} = -(1/2 \sigma_0^2) \text{tr}(D_x \Sigma L(I_n \otimes \Sigma \otimes L)' D_x) \]

\[ + (1/2 \sigma_0^2) \hat{w}' D_x \Sigma L(I_n \otimes \Sigma \otimes L)' D_x \hat{w} - \text{tr}(L d \Sigma_w) = 0. \tag{25} \]

Writing \( L \) in block partition form,

\[ L = [L_{11}, L_{12}, \ldots, L_{1n}] \tag{26} \]

we can show that

\[ L = -(1/2 \sigma_0^2) \sum_{i=1}^{n} L_{i1} L_{i2} D_x L_{i1} + (1/2 \sigma_0^2) \sum_{i=1}^{n} L_{i1} Q_{i1} \hat{w}' D_x L_{i1} \]

and thus concentrating on the stochastic part and excluding constants we have

\[ \text{tr} L = (1/2 \sigma_0^2) \hat{w}' \hat{D}_x L \hat{L} \hat{D}_x \hat{w}. \tag{28} \]

Observe that (28) is equivalent in form to the \( LM \) statistic based on the score \( s^*_x \) given in (11). The limit theory follows through except that here premultiplying by \( L' \) picks up the \( k \)'th block whereas postmultiplying by \( L \) picks up the \( r \)'th element. Since the distribution is the same in either case we have

\[ n^{-3} \hat{w}' \hat{D}_x L \hat{L} \hat{D}_x \hat{w} = \sum_{j=1}^{k} \left[ \Sigma_{i=1}^{n} \hat{w}_x x_{ij} \right] \left[ \Sigma_{i=1}^{n} x_{ij} \hat{w}_x \right] \]

\[ \xrightarrow{d} \sigma_0^2 \text{tr} \left\{ \Sigma_{i=1}^{n} \int_{0}^{1} W_{r}(r) \left[ W_{r}(r)' \right] dr \right\} \]

where \( W_{r}(r) = \int_{0}^{r} V_i dW_0 - \int_{0}^{r} V_i V_i'(e_{oi} + \Sigma_{i=1}^{n} \hat{w}_x x_{ij} \hat{w}_x) V_i dW_0. \) To obtain a nuisance parameter free limit theory we set up
\[ LM_3 = n^{-3} \tilde{w}' \mathcal{D}_z \mathcal{L}(I_n \otimes \tilde{\Sigma}_{11}^{-1}) \mathcal{L}' \mathcal{D}_z \tilde{w}/\hat{a}_0^2 \]
\[
\overset{d}{\rightarrow} \text{tr} \left\{ \int_0^1 W_{r_1}(r) W_{r_1}(r)' \, dr \right\}
\]

where \( \hat{a}_0^2 = n^{-1} \tilde{w}' \tilde{w} \overset{p}{\rightarrow} \sigma_0^2 \) and \( \tilde{\Sigma}_{11} = n^{-1} \sum_{i=1}^n (Ax_i Ax_i)' \overset{p}{\rightarrow} \Sigma_{11} \). The limit distribution depends only on the rank of \( V_1 \).

Extending the results to the general time series case and to models with deterministic trends proceeds as before. The notation follows that of Section 3.2 except that the correction now takes the form

\[
n^{-1} \sum_{i=1}^n x_i \tilde{w}_i^{++} = n^{-1} \sum_{i=1}^{m_n} x_i \tilde{w}_i^+ - r \tilde{\delta}^+ \text{ where } \tilde{\delta}^+ = \tilde{r}_1 \left[ \begin{array}{c} 1 \\ \tilde{r}_1^{-1} \tilde{r}_1 \end{array} \right].
\]

The multivariate extensions for Corollaries 3.1 and 3.2 are as follows

**Theorem 4.1:** The LM \( k \)-parameter constancy test for the null of cointegration in a model without trend is

\[
LM_1^* = n^{-3} \tilde{w}_i^{++} \mathcal{D}_z \mathcal{L}(I_n \otimes \tilde{\Psi}_1^{-1}) \mathcal{L}' \mathcal{D}_z \tilde{w}_i^{++}/\hat{\Psi}_{0,1}^{-1} \overset{d}{\rightarrow} \text{tr} \left\{ \int_0^1 W_{r_1}(r) W_{r_1}(r)' \, dr \right\}
\]

where \( W_{r_1}(r) = \int_0^r V_1 \, dW_{0,1} - \left( \int_0^r V_1 \right) \left( \int_0^r V_1 \right)^{-1} \int_0^r V_1 \, dW_{0,1}. \)

**Theorem 4.2:** The LM \( k \)-parameter constancy test for the null of cointegration in a model with trend is

\[
LM_2^* = n^{-3} \tilde{w}_i^{++} \mathcal{D}_z \mathcal{L}(I_n \otimes \tilde{\Psi}_1^{-1}) \mathcal{L}' \mathcal{D}_z \tilde{w}_i^{++}/\hat{\Psi}_{0,1}^{-1}
\]
\[
\overset{d}{\rightarrow} \text{tr} \left\{ \int_0^1 W_{Q_1}(r) W_{Q_1}(r)' \, dr \right\}
\]

where \( W_{Q_1}(r) = \int_0^r V_1 \, dW_{0,1} - \left( \int_0^r V_1 \right) \left( \int_0^r V_1 \right)^{-1} \int_0^r V_1 \, dW_{0,1}. \)

Observe that the limit distributions are nuisance parameter free and depend only on the number of regressors. The latter test, \( LM_2^* \), also depends on the order \( (p) \) of the trend regressors. In fact, for many applications of parameter constancy tests as tests for cointegration, the limit distribution of the LM statistic will depend only on these regressor counts. If the regressors are not cointegrated, then the rank of \( V_1 \) corresponds to the number of the regressors \( \{ x_t \} \). If we allow cointegration among regressors, then \( \Psi_1 \) is singular and the situation is a little more complex. It may be handled by the methods recently developed in Phillips (1991).
4.1 Testing when Regressors Contain Trends

The test can also be applied to models where regressors are allowed to contain both deterministic and stochastic trends. For example, consider the regression equation (1)" with the regressors \( \{x_i\} \) generated by

\[
x_i = Ah_{1i} + x_1^o, \quad x_1^o = \sum_{i=1}^{t} e_{1i} + x_0.
\]

(2)"

Here \( x_1^o \) is a \( k \) dimensional vector of stochastic trends, \( h_{1i} \) is a \( p \) dimensional vector of deterministic trends and \( A \) is a \( k \times p \) matrix of parameters. As in the previous section, convergence of the deterministic trends \( h_{1[\tau]} \) to the limit vector \( h_{1i} (r) \) will be given by a diagonal matrix of weights \( \gamma_{1n} \) such that \( \gamma_{1n} h_{1[\tau]} \rightarrow h_{1i} (r) \) uniformly in \( r \).

To separate the trends driving \( \{x_i\} \), we define an orthogonal matrix \( J = [J_1, J_2] \) where \( J_1 = A(A' A)^{-1/2}, \) \( J_2 = A_\perp \) so that \( J' J = I_k \). Furthermore, we denote by \( Z \), the variables transformed by \( J \),

\[
Z_i = J' x_i = \begin{pmatrix} (A'A)^{1/2} h_{1i} + J_1 x_1^o \\ J_2 x_1^o \end{pmatrix} = \begin{pmatrix} Z_{1i} \\ Z_{2i} \end{pmatrix},
\]

where \( Z_{1i} \) and \( Z_{2i} \) are of dimension \( p \) and \( k - p \) respectively. Then we can rewrite (7)" as

\[
y_i = \beta_0 (J J') x_i + w_i = C_1 Z_{1i} + C_2 Z_{2i} + w_i = Z_i' C + w_i
\]

(7)"

where \( C = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} J_1' \beta_0 \\ J_2' \beta_0 \end{pmatrix} \). The variables \( Z_{1i} \) and \( Z_{2i} \) contain the deterministic and stochastic trends respectively since, given the weight matrix

\[
I_n = \begin{pmatrix} I_{1n} \\ I_{2n} \end{pmatrix} = \begin{pmatrix} \gamma_{1n} (A'A)^{-1/2} & 0 \\ 0 & 1/\sqrt{n} \end{pmatrix},
\]

we have

\[
I_n Z_{[\tau]} = \begin{pmatrix} \gamma_{1n} (A'A)^{-1/2} Z_{1[\tau]} \\ 1/\sqrt{n} Z_{2[\tau]} \end{pmatrix} \rightarrow \begin{pmatrix} h_1 (r) \\ B_1 (r) \end{pmatrix} = \bar{Q}(r)
\]

(30)

where \( \bar{B}_1 (r) = J_2' B_1 (r) \). Using (30), we find that the coefficients in the regression (7)" are distributed as follows.

Lemma 3:

i) \( n^{1/2} I_{1n}^{-1}(\bar{C} - C_1) \stackrel{d}{\rightarrow} (\int h_1^+ h_1^+)\int (J_1^+ dB_0 + \bar{Y}_{10}) \)

where \( h_1^+ = h_1 (r) - \int h_1 \bar{B}_1 (r) (J_1^+ \bar{B}_1) (r) \) and \( \bar{Y}_{10} = -\int h_1 \bar{B}_1 (r) (J_1^+ \bar{B}_1) \int J_2' Y_{10} \).
ii) \( n^{1/2} I_n^{-1} (C_2 - C_2) \xrightarrow{d} (\int \overline{B}_1 \overline{B}_1^* \int dB_0 + J_2 Y_{10}) \)

where \( \overline{B}_1^* = \overline{B}_1(r) - \int \overline{B}_1 h_1(\int h_1 h_1)^{-1} h_1(r) \).

The limit theory for the LM test is derived following the arguments of the previous section using the asymptotics of Lemma 3. Write the correct residual from (7) as \( \tilde{w}_s = w_s - Z_s (C - C) - \Delta X_s \tilde{\phi}_1^{-1} \tilde{\phi}_{10} \) and our score is now

\[
n^{-1/2} n \sum \tilde{w}_s z_s = n^{-1/2} \sum \Gamma_n z_s w_s - n^{-1} \sum \Gamma_n z_s z_s \Gamma_n^{-1} n^{-1/2} (C - C)
- \frac{\delta}{n} \int \frac{\overline{O}_0}{\overline{O}_0} d\overline{B}_{0,1} + r \left( 0 \right)
\]

To eliminate the nuisance parameter \( J_2 \delta^* \) at the limit we set \( z_s = \Gamma_n z_s \) and write \( \sum \tilde{Z}_s \tilde{w}_s^* = \sum \tilde{Z}_s \tilde{w}_s^* - r \left( 0 \right) \) as before. Then the LM test is

\[
LM = n^{-2} \tilde{w}_s z_s \Lambda(L_n \otimes \tilde{\phi}_1^{-1}) L_s \tilde{w}_s / \psi_{0,1}
\]

where \( \tilde{\phi}_1 \) and \( \psi_{0,1} \) are semiparametric estimators of the long-run variances of the \( \{x_t\} \) generated by (2) and are constructed after first detrending the series \( \{x_t\} \). Methods of estimating these long-run variances are now well-known and are discussed in detail, for example, in Andrews (1991). The limit distribution in this case depends only on the \( p \) and \( k - p \) deterministic and stochastic trends contained in \( \overline{O}(r) \).

4.2 Subset Testing

Suppose that instead of testing for the constancy of the \( k \)-parameter vector \( \beta \), we test for the constancy of a subset \( \beta_k \) of the coefficients where \( 1 \leq k < k \). This is useful in empirical cases where cointegration among three or more variables fails and one is interested in which coefficient "caused" the failure. Consider for example the following system

\[
y_t = \beta_{1t} x_{at} + \beta_{2t} x_{bt} + \epsilon_{0t}
\]
\[ Ax_t = \varepsilon_{1t}, \quad \varepsilon_{1t} = (\varepsilon_{a1}, \varepsilon_{b1}) \]

\[ \beta_{at} = \beta_{at-1} + \eta_{at} \]  

which we can write as

\[ y_t = \beta_0 x_{aat} + \beta_b x_{bt} + w_{at}, \quad w_{at} = (\Sigma_1^{\eta_{aat}}) x_{aat} + \varepsilon_{0t}. \]  

The multivariate invariance principle is \( n^{-1/2} \sum_{\ell=1}^a d B_\ell(r) = (B_\ell(r), B_k(r)) \) where the partition is conformable with that of \( \{x_t\} \). Because \( \{\varepsilon_{0t}\} \) is stationary we have the cointegrating vector \( (1, -\beta_0, -\beta_b) \) under the null of constant coefficients. With the alternative \( \Sigma_n > 0 \), rejection of the null indicates that the failure of cointegration is caused by the instability of the \( k_1 \) parameters \( \beta_k \).

The \( LM \) test as before is

\[ LM_a = n^{-3} \sum_{\ell=1}^a \hat{x}^{+ \ell} D_{x\ell} L(I_2 \otimes \hat{\Psi}_{\ell1}) L D_{x\ell} \hat{x}^{+ \ell} / \hat{\Psi}_{01} \]  

with \( n^{-3} \sum_{\ell=1}^a \hat{x}^{+ \ell} D_{x\ell} L(I_2 \otimes \hat{\Psi}_{\ell1}) L D_{x\ell} \hat{x}^{+ \ell} / \hat{\Psi}_{01} \) and \( \hat{x}^{+ \ell} = \hat{x}_{\ell0} - \hat{\Psi}_{\ell1} \hat{\Psi}_{01}^{-1} \hat{x}_{\ell0} \). The limit theory is easy to deduce along the lines of our earlier analysis and is given by

\[ LM_a \overset{d}{\to} \text{tr} \left\{ \int_0^1 W_{y_{a}}^+(r) W_{y_{a}}^+(r)^\prime \, dr \right\}. \]

where \( W_{y_{a}}^+(r) = \int_0^1 V_a dW_{0.1} - (\int_0^1 V_a V_{1}^{\prime}) (\int_0^1 V_1 V_1^{\prime})^{-1} (\int_0^1 V_1 dW_{0.1}) \). Here the limit theory depends only on the rank of the \( I(1) \) regressors \( (k) \) and the rank of the subset of the regressors being tested \( (k_1) \). Observe that at \( r = 1 \) we have \( W_{y_{a}}^+(r) = \int_0^1 V_a dW_{0.1} - (\int_0^1 V_a V_{1}^{\prime}) (\int_0^1 V_1 V_1^{\prime})^{-1} (\int_0^1 V_1 dW_{0.1}) = \int_0^1 V_a dW_{0.1} = 0 \), since \( V_1 V_1^{\prime} \) is orthogonal to \( V_{1}^{\prime} \). Thus, \( W_{y_{a}}^+(r) \) is a tied-down process.

In summary, we have considered models with and without trend for the case of univariate and multivariate regressors. In both cases, the null that is of interest is cointegration or the stability of the coefficients of the \( I(1) \) regressors. When the model contains deterministic trends, the version of the \( LM \) test that is to be used depends on the treatment of the trends. If interest is in a test of a null of stochastic cointegration, the appropriate test is Corollary 3.2 or Theorem 4.2 for the univariate and multivariate cases, respectively. If the null of interest is deterministic cointegration, then the version of the test given in Section 4.1 is appropriate.

5 Limit Theory Under Local Alternatives and Some Simulation Results

It is important to know if the empirical \( LM \) tests will have good power properties under general alternative hypotheses. One important criterion is test con-
sistency. Clearly the $L M$ tests considered above diverge when $\sigma_n^2 > 0$ (or $\Sigma_n > 0$) so the tests are consistent. The behavior of the tests under local alternatives can also be considered. Let the sequence of local alternatives for the regression coefficient be given by

$$H_a: \beta_t = \beta_0 + \gamma_t/n$$  \hspace{1cm} (36)

where $\gamma_t$ can be any constant or a random variable that is $O(p(1))$. The specification (36) can be interpreted as a condition of near cointegration as discussed in Phillips (1988b) and Tanaka (1993). Because the parameter of interest in our case is the variance term $\sigma_n^2$, it is convenient to let

$$\gamma_t = n^{-1/2} \sum_t \eta_t.$$  \hspace{1cm} (37)

With this specification for $\gamma_t$, (36) is in fact a triangular array (i.e., $\gamma_t = \gamma_t^{(n)}$, $\beta_t = \beta_t^{(n)}$), but we will not overburden the notation to accommodate this fact but instead continue to work with single indexed functions like $\beta_t$. The model under the alternative is

$$y_t = \beta_t x_t + \varepsilon_{0t}$$

$$- \beta_0 x_t + n^{-1/2} \sum \eta_t x_t + \varepsilon_{0t}$$

$$= \beta_0 + n^{-1/2} \sum \eta_t x_t + \varepsilon_{0t}. \hspace{1cm} (38)$$

Now $n^{-1/2} \sum \eta_t x_t \xrightarrow{d} \mathcal{B}_1(r)$ and $n^{-1} \sum \eta_t x_t \varepsilon_{0t} \xrightarrow{d} B_1(r) \mathcal{B}_1(r) \equiv G_{13}^B(r)$, say. Now write $G_{13}^B(r) = (\sigma_1^2 \sigma_n^2)^{1/2} W_1(r) W_3(r) = \sigma_1^2 \sigma_n^2 G_{13}^n(r)$. It follows that $\beta$ is still consistent since (taking the innovations to be iid)

$$n(\hat{\beta} - \beta_0) = n^{-1/2} \sum x_t \varepsilon_{0t}^{-1} n \sum \eta_t^2 x_t (\varepsilon_{0t}^{-1} n^{-1} \sum \eta_t x_t + \varepsilon_{0t})$$

$$= n^{-1/2} \sum x_t \varepsilon_{0t}^{-1} n \sum \eta_t^2 x_t (n^{-1} \sum \eta_t x_t + n^{-1} \sum \eta_t x_t \varepsilon_{0t})$$

$$\xrightarrow{d} (\int B_1^2)^{-1} (\int \beta_0 G_{13}^B + \int \beta_1 dB_t). \hspace{1cm} (39)$$

But our $L M$ test is now a functional of

$$1 \sum x_t \varepsilon_{0t} = n^{-1} \sum \frac{\eta_t x_t}{\sqrt{n}} - n(\hat{\beta}_0 - \beta_0) n^{-2} \sum x_t^2$$

$$\xrightarrow{d} \left( \int_0^1 B_t G_{13}^B + \int_0^1 B_t G_{13}^B + (\sigma_1^2 \sigma_n^2)^{1/2} W_{60w},(r) \right)$$

$$= \sigma_n^2 (\sigma_1^2)^{1/2} G_{13}^n + (\sigma_1^2 \sigma_n^2)^{1/2} W_{60w},(r) \hspace{1cm} (40)$$

which is well defined and distinct from the null distribution provided $\sigma_n^2 \neq 0$. When $\sigma_n^2 = 0$ the statistic converges to the distribution of the score under the
null. In the last line of (40), $G_{1,3, W_1}$ and $W_{0, W_1}^{W}(r)$ are bridge processes constructed from $G_{0, W_1}(r)$ and $W_0(r)$ using the Brownian motion $W_t(r)$.

In order to examine the performance of the $LM$ test in finite samples we performed some small-scale simulations for the simplest case of iid errors. The first experiment checked the performance of the $LM$ test under the null of no parameter variation. From Table 2 we see that the $LM$ test has only minor size distortion in samples of size $n = 100, 150, 200$.

The second experiment checked the empirical power of the $LM$ test. The sample size used was 250 and critical values based on Table 1 were employed. Table 3 shows the frequency of rejection of the null using the upper 1%, 5% and 10% asymptotic critical values. It is clear that it takes a relatively large variance ($\sigma^2_u$) for the test to attain high power, although it has at least 50% power at the 10% and 5% levels with $\sigma^2_u = 2$ and $\sigma^2_u = 4$, respectively. This evidence suggests that the upper 5% and 10% tails of the density may be better suited for use in empirical testing.

Tables 4.A (a)–(d) give the 1%, 5%, 10% and 20% asymptotic critical values of the test for the model

$$y_t = \mu + \beta x_t + \epsilon_t$$

(41)
Table 4A.

(a) Critical values at the 1% level

<table>
<thead>
<tr>
<th>Total regressors</th>
<th>Number of subset regressors</th>
<th>k = 1</th>
<th>k = 2</th>
<th>k = 3</th>
<th>k = 4</th>
<th>k = 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2.422</td>
<td>.3652</td>
<td>.3339</td>
<td>.4200</td>
<td>.5344</td>
</tr>
<tr>
<td>k = 1</td>
<td></td>
<td>.2362</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 2</td>
<td></td>
<td>.2107</td>
<td>.3225</td>
<td>.4243</td>
<td>.5086</td>
<td>.6109</td>
</tr>
<tr>
<td>k = 3</td>
<td></td>
<td>.1900</td>
<td>.2946</td>
<td>.4041</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 4</td>
<td></td>
<td>.1942</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Critical values at the 5% level

<table>
<thead>
<tr>
<th>Total regressors</th>
<th>Number of subset regressors</th>
<th>k = 1</th>
<th>k = 2</th>
<th>k = 3</th>
<th>k = 4</th>
<th>k = 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>.1052</td>
<td>.1747</td>
<td>.2266</td>
<td>.2784</td>
<td></td>
</tr>
<tr>
<td>k = 1</td>
<td></td>
<td>.0989</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 2</td>
<td></td>
<td>.0891</td>
<td>.1503</td>
<td>.2652</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 3</td>
<td></td>
<td>.0834</td>
<td>.1433</td>
<td>.2038</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 4</td>
<td></td>
<td>.0819</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(c) Critical values at the 10% level

<table>
<thead>
<tr>
<th>Total regressors</th>
<th>Number of subset regressors</th>
<th>k = 1</th>
<th>k = 2</th>
<th>k = 3</th>
<th>k = 4</th>
<th>k = 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>.0661</td>
<td>.1172</td>
<td>.1548</td>
<td>.1939</td>
<td></td>
</tr>
<tr>
<td>k = 1</td>
<td></td>
<td>.0625</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 2</td>
<td></td>
<td>.0568</td>
<td>.1074</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 3</td>
<td></td>
<td>.0519</td>
<td>.1013</td>
<td>.1474</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 4</td>
<td></td>
<td>.0501</td>
<td>.0952</td>
<td>.1400</td>
<td>.1843</td>
<td></td>
</tr>
<tr>
<td>k = 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(d) Critical values at the 20% level

<table>
<thead>
<tr>
<th>Total regressors</th>
<th>Number of subset regressors</th>
<th>k = 1</th>
<th>k = 2</th>
<th>k = 3</th>
<th>k = 4</th>
<th>k = 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>.0380</td>
<td>.0730</td>
<td>.1000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 1</td>
<td></td>
<td>.0351</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 2</td>
<td></td>
<td>.0320</td>
<td>.0663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 3</td>
<td></td>
<td>.0291</td>
<td>.0623</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 4</td>
<td></td>
<td>.0373</td>
<td>.0582</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k = 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Number of iterations = 20,000, n = 500.

where \( \mu \) is the intercept and the dimension of \( x \) is \( k (k = 1, 2, 3, 4, 5) \). When the \( \{x_t\} \) (for \( k < 1 \)) are allowed to contain both deterministic and stochastic trends, the critical values of the test differ and are given in Table 4B. Table 4C contains the critical values for \( \{x_t\} \) being I(1) without drift but with the regression model (41) containing a trend. The sample size used was 500 and the number of iterations = 20,000.
Table 4B. Upper critical values with trends in regressors

(a) Critical values at the 1% level

| Total regressors |  
|------------------|------------------|------------------|------------------|------------------|------------------|
|                  | $k = 1$ | $k = 2$ | $k = 3$ | $k = 4$ | $k = 5$ |
| $k = 2$          | 0.2356  | 0.2888  |        |        |        |
| $k = 3$          | 0.2166  | 0.3484  | 0.4034 |        |        |
| $k = 4$          | 0.1940  | 0.3197  | 0.4357 | 0.4987 |        |
| $k = 5$          | 0.1836  | 0.3019  | 0.4071 | 0.5096 | 0.5655 |

(b) Critical values at the 5% level

| Total regressors |  
|------------------|------------------|------------------|------------------|------------------|------------------|
|                  | $k = 1$ | $k = 2$ | $k = 3$ | $k = 4$ | $k = 5$ |
| $k = 2$          | 0.0982  | 0.1463  |        |        |        |
| $k = 3$          | 0.0923  | 0.1661  | 0.2019 |        |        |
| $k = 4$          | 0.0832  | 0.1511  | 0.2168 | 0.2528 |        |
| $k = 5$          | 0.0789  | 0.1440  | 0.2078 | 0.2669 | 0.3023 |

(c) Critical values at the 10% level

| Total regressors |  
|------------------|------------------|------------------|------------------|------------------|------------------|
|                  | $k = 1$ | $k = 2$ | $k = 3$ | $k = 4$ | $k = 5$ |
| $k = 2$          | 0.0606  | 0.0918  |        |        |        |
| $k = 3$          | 0.0561  | 0.1091  | 0.1377 |        |        |
| $k = 4$          | 0.0513  | 0.1000  | 0.1475 | 0.1759 |        |
| $k = 5$          | 0.0485  | 0.0960  | 0.1415 | 0.1861 | 0.2143 |

(d) Critical values at the 20% level

| Total regressors |  
|------------------|------------------|------------------|------------------|------------------|------------------|
|                  | $k = 1$ | $k = 2$ | $k = 3$ | $k = 4$ | $k = 5$ |
| $k = 2$          | 0.0342  | 0.0571  |        |        |        |
| $k = 3$          | 0.0311  | 0.0654  | 0.0856 |        |        |
| $k = 4$          | 0.0285  | 0.0605  | 0.0924 | 0.1109 |        |
| $k = 5$          | 0.0267  | 0.0570  | 0.0878 | 0.1169 | 0.1358 |

Notes: Number of iterations = 20,000, $n = 500$. 
Table 4.C. Upper critical values with trend in regression

(a) Critical values at the 1% level

<table>
<thead>
<tr>
<th>Total regressors</th>
<th>k = 1</th>
<th>k = 2</th>
<th>k = 3</th>
<th>k = 4</th>
<th>k = 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>k = 1</td>
<td>.4412</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>k = 2</td>
<td>.3455</td>
<td>.5511</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>k = 3</td>
<td>.2643</td>
<td>.4304</td>
<td>.5689</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>k = 4</td>
<td>.2077</td>
<td>.3257</td>
<td>.4443</td>
<td>.5597</td>
<td>—</td>
</tr>
<tr>
<td>k = 5</td>
<td>.1818</td>
<td>.2803</td>
<td>.3739</td>
<td>.4722</td>
<td>.5685</td>
</tr>
</tbody>
</table>

(b) Critical values at the 5% level

<table>
<thead>
<tr>
<th>Total regressors</th>
<th>k = 1</th>
<th>k = 2</th>
<th>k = 3</th>
<th>k = 4</th>
<th>k = 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>k = 1</td>
<td>.1980</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>k = 2</td>
<td>.1544</td>
<td>.2704</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>k = 3</td>
<td>.1233</td>
<td>.2158</td>
<td>.3040</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>k = 4</td>
<td>.1002</td>
<td>.1747</td>
<td>.2453</td>
<td>.3141</td>
<td>—</td>
</tr>
<tr>
<td>k = 5</td>
<td>.0850</td>
<td>.1493</td>
<td>.2106</td>
<td>.2665</td>
<td>.3228</td>
</tr>
</tbody>
</table>

(c) Critical values at the 10% level

<table>
<thead>
<tr>
<th>Total regressors</th>
<th>k = 1</th>
<th>k = 2</th>
<th>k = 3</th>
<th>k = 4</th>
<th>k = 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>k = 1</td>
<td>.1252</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>k = 2</td>
<td>.0996</td>
<td>.1872</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>k = 3</td>
<td>.0809</td>
<td>.1535</td>
<td>.2187</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>k = 4</td>
<td>.0673</td>
<td>.1245</td>
<td>.1792</td>
<td>.2327</td>
<td>—</td>
</tr>
<tr>
<td>k = 5</td>
<td>.0571</td>
<td>.1069</td>
<td>.1537</td>
<td>.2008</td>
<td>.2451</td>
</tr>
</tbody>
</table>

(d) Critical values at the 20% level

<table>
<thead>
<tr>
<th>Total regressors</th>
<th>k = 1</th>
<th>k = 2</th>
<th>k = 3</th>
<th>k = 4</th>
<th>k = 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>k = 1</td>
<td>.0723</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>k = 2</td>
<td>.0581</td>
<td>.1205</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>k = 3</td>
<td>.0485</td>
<td>.0994</td>
<td>.1476</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>k = 4</td>
<td>.0413</td>
<td>.0828</td>
<td>.1226</td>
<td>.1624</td>
<td>—</td>
</tr>
<tr>
<td>k = 5</td>
<td>.0353</td>
<td>.0713</td>
<td>.1069</td>
<td>.1402</td>
<td>.1744</td>
</tr>
</tbody>
</table>

Notes: Number of iterations = 20,000, n = 500.
6 Empirical Application

In this section we apply our testing method to Australian aggregate macro-economic data and study the long-run form of the aggregate Australian consumption function following Phillips (1992). Our LM statistic is presented, together with other classical statistical procedures, to test for the long-run relationships between private consumption expenditure, household disposable income and real wealth.\(^2\) Comparison between outcomes of these tests facilitates a study of the sensitivity of our LM procedure in detecting departures from the null.

6.1 Description and Characteristics of the Data

The data set we use covers the period 1965(1)–1988(4) and consists of nominal private consumption expenditure (C), household disposable income (YD), the consumer price deflator (P) and real liquid assets (M3/P). Real variables are constructed from the data and are denoted by lower case letters: \(c = C/P = \) real consumption expenditure; \(yd = YD/P = \) real household disposable income; and \(m3 = M3/P = \) real money stock. All variables are seasonally adjusted and the constant price series are at average 1984/1985 prices.

Figures 1(i)–(ii) graph these series (in log levels) to illustrate their relationship with each other over time. Clearly, consumption and income move closely together, although the nominal series appear to move closer than the real series. The real money stock, m3, in Figure 1(ii) exhibits divergent behavior midway through and towards the end of the sample.

Tests were conducted to determine the presence of stochastic trends in the data. Table 5 reports the results for the Augmented Dickey-Fuller (ADF) test and Phillips (1987) \(Z_a\) and \(Z_t\) tests for both the nominal and real series with a constant and trend in the regression equations (i.e. trend degree = 1). The null that is tested is that of a unit root and in all cases, for the consumption, income and wealth series at lags = 2, 5, the null cannot be rejected at the 5% level. These results corroborate Phillips’ finding of a unit root in sample data on consumption and income over the longer period 1959(3)–1988(4).

\(^2\) Phillips conducted a similar analysis and compared classical statistical procedures to the Phillips–Ploberger (1991) Bayesian posterior odds test.
Fig. 1.
698

C. E. Quintos and P. C. B. Phillips

Table 5. Unit root tests

<table>
<thead>
<tr>
<th>Test</th>
<th>lag</th>
<th>ln(C)</th>
<th>ln(YD)</th>
<th>ln(c)</th>
<th>ln(yd)</th>
<th>ln(m3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF</td>
<td>2</td>
<td>-1.51</td>
<td>-1.22</td>
<td>-1.95</td>
<td>-1.89</td>
<td>-1.59</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-2.21</td>
<td>-1.63</td>
<td>-2.31</td>
<td>-1.64</td>
<td>-1.55</td>
</tr>
<tr>
<td>Z_a</td>
<td>2</td>
<td>-3.03</td>
<td>-2.74</td>
<td>-2.86</td>
<td>-3.67</td>
<td>-6.69</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-3.76</td>
<td>-3.49</td>
<td>-2.49</td>
<td>-3.77</td>
<td>-6.46</td>
</tr>
<tr>
<td>Z_t</td>
<td>2</td>
<td>-1.31</td>
<td>-1.16</td>
<td>-1.56</td>
<td>-1.78</td>
<td>-1.69</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-1.44</td>
<td>-1.31</td>
<td>-1.54</td>
<td>-1.79</td>
<td>-1.66</td>
</tr>
</tbody>
</table>

5% Critical values: $Z_a = -20.84; ADF, Z_t = -3.46$. (Critical values from Phillips-Perron (1988)).

6.2 The Aggregate Australian Consumption Function

The presence of deterministic and stochastic trends in the time series suggests that the long-run form of the Australian consumption function should exhibit co-movement between the determinants of consumption behavior and consumption expenditure. The simplest specification of the long-run consumption function under the permanent income hypothesis is the consumption-income relation

$$\ln(C) = \alpha + \beta \ln(YD) + u_t$$  \hspace{1cm} (42.N)

or its real equivalent

$$\ln(c) = \alpha + \beta \ln(yd) + u_t$$  \hspace{1cm} (42.R)

If the permanent income hypothesis were true, consumption and income should be cointegrated and the residuals $u_t$ in (42) should exhibit stationary behavior. In fact, an earlier study by Hall-Trevor (1991) found that the null hypothesis of no cointegration could not be rejected for aggregate real variables (42.R) but could be rejected for nominal variables (42.N), leading them to conclude that

"... for real consumer expenditure an aggregate equation should not be estimated.... In contrast, results for the nominal variables suggest that it would be appropriate to estimate the aggregate model."\(^3\)

However, Phillips (1991a) found somewhat different results. Using Bayesian and classical procedures, his conclusion was that both nominal and real variables do not exhibit long-run cointegrating relationships. Following earlier work on alternative forms of the consumption function that incorporate the effects of

inflation (notably Hendry-von Ungern-Sternberg (1981) and Anstic-Gray-Pagan (1983)), Phillips used the equation
\[
\ln(C) = \alpha + \beta \ln(YD^*) + u_t. \tag{43}
\]
Here, the variable \(YD^*\) is the "perceived income" or disposable income adjusted for the loss in liquid wealth due to inflation \(\pi\):
\[
YD^* = YD \{1 - \phi(\bar{\pi} * M3/YD)\} = YD \{1 - \text{loss}\}, \tag{44}
\]
and the parameter \(\phi\) is introduced to allow for scale effects resulting from an inappropriate choice of an inflation measure \(\bar{\pi}\) (= average inflation rate over the past four quarters) or for choosing \(M3\) as a measure of liquid assets. Taking the logarithm of (44) the approximation to (43) is given by
\[
\ln(C) = \alpha + \beta \ln(YD) + \gamma \text{loss} + u_t, \tag{45.N}
\]
\[
\ln(c) = \alpha + \beta \ln(yd) + \gamma \text{loss} + u_t, \tag{45.R}
\]
which correspond to equations (33) and (34) in Phillips (1992). For both equations, using classical and Bayesian procedures, Phillips found that these augmented regression equations for consumption are cointegrating. Table 6 reports the long-run parameter estimates of the regression equations using fully modified least squares. We see that the loss measure is significant in explaining consumption in both equations \(t\)-ratio = \(-6.34\) and \(-3.13\) for the real and

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Nominal</th>
<th>(t)-Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>(42.N)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d = -0.224)</td>
<td></td>
<td>(-7.52)</td>
</tr>
<tr>
<td>(\phi = 1.012)</td>
<td></td>
<td>(324.86)</td>
</tr>
<tr>
<td>(45.N)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d = -0.229)</td>
<td></td>
<td>(-5.78)</td>
</tr>
<tr>
<td>(\phi = 1.024)</td>
<td></td>
<td>(171.68)</td>
</tr>
<tr>
<td>(\gamma = -1.8e-06)</td>
<td></td>
<td>(-3.13)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Real</th>
<th>(t)-Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>(42.R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d = -0.271)</td>
<td></td>
<td>(-0.47)</td>
</tr>
<tr>
<td>(\phi = 1.016)</td>
<td></td>
<td>(18.07)</td>
</tr>
<tr>
<td>(45.R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d = -1.087)</td>
<td></td>
<td>(-3.91)</td>
</tr>
<tr>
<td>(\phi = 1.104)</td>
<td></td>
<td>(39.4)</td>
</tr>
<tr>
<td>(\gamma = -4.69e-06)</td>
<td></td>
<td>(-6.34)</td>
</tr>
<tr>
<td>(46.R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d = -1.34)</td>
<td></td>
<td>(-3.63)</td>
</tr>
<tr>
<td>(\phi = 0.9895)</td>
<td></td>
<td>(15.18)</td>
</tr>
<tr>
<td>(\gamma = -3.69e-06)</td>
<td></td>
<td>(-4.96)</td>
</tr>
<tr>
<td>(\delta = 0.134)</td>
<td></td>
<td>(1.65)</td>
</tr>
</tbody>
</table>
(i) Nominal consumption and perceived income

(ii) Real consumption and perceived income

Fig. 2.
nominal equations respectively). Furthermore, although the marginal propensity to consume out of disposable income is not significantly different from unity in (42), the propensity to consume is more sensitive to shocks in inflation – adjusted income as the marginal propensity to consume out of income is greater in (45). The behavior of perceived income $Y_D$ is graphed together with consumption in Figure 2(i)–(ii) in log levels. Clearly inflationary effects introduce more sample variability into income and support the conclusion that inflation plays a useful role in relating consumption behavior to income expenditure in the long-run.

In order to determine the sensitivity of our tests to departures from the null of cointegration, we report the results of our $LM_3$ test applied to equations (42) and (45) in the upper panel of Table 7. Hansen’s (1992a) parameter constancy test on the intercept term is denoted by $LM_{i,i}$ and is also reported in the table. The lower panel presents the results of the other classical procedures that test the null of no cointegration, and the estimated autoregressive coefficients in the residuals are denoted by $\hat{\rho}_n$ and $\hat{\rho}_r$ for the nominal and real regressions, respectively. For the first consumption equation (42), the $ADF$, $Z_{q}$ and $Z_{v}$ tests do not reject the null of no cointegration for both the nominal and real variables corroborating Phillips’ finding. The $LM$ tests for the null of cointegration present conflicting results. Our $LM_3$ test confirms cointegration in the nominal case but rejects cointegration in the real case (this corroborates the Hall-Trevor result). The $LM_{i,i}$ test supports the presence of cointegration in both cases. Thus the empirical evidence is mixed. One possible explanation for these mixed findings

<table>
<thead>
<tr>
<th>Test</th>
<th>42.N</th>
<th>42.R</th>
<th>5% cv</th>
<th>45.N</th>
<th>45.R</th>
<th>5% cv</th>
<th>46.R</th>
<th>5% cv</th>
</tr>
</thead>
<tbody>
<tr>
<td>$LM_3$</td>
<td>0.03</td>
<td>0.18*</td>
<td>0.05</td>
<td>0.14</td>
<td>0.15</td>
<td>0.14</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>$LM_{i.i}$</td>
<td>0.32</td>
<td>0.29</td>
<td>0.62</td>
<td>0.25</td>
<td>0.78</td>
<td>0.21</td>
<td>0.90</td>
<td></td>
</tr>
</tbody>
</table>

**Table 7. Tests for cointegration**

<table>
<thead>
<tr>
<th>Test-lag</th>
<th>42.N</th>
<th>42.R</th>
<th>5% cv</th>
<th>45.N</th>
<th>45.R</th>
<th>5% cv</th>
<th>46.R</th>
<th>5% cv</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ADF$ - 2</td>
<td>-2.07</td>
<td>-1.98</td>
<td>-3.40</td>
<td>-4.05*</td>
<td>-3.8**</td>
<td>-3.84</td>
<td>-3.46</td>
<td>-4.22</td>
</tr>
<tr>
<td>5</td>
<td>-1.45</td>
<td>-1.39</td>
<td>-3.40</td>
<td>-3.56**</td>
<td>-3.01</td>
<td>-3.84</td>
<td>-2.79</td>
<td>-4.22</td>
</tr>
<tr>
<td>$Z_v$ - 2</td>
<td>-2.33</td>
<td>-2.17</td>
<td>-3.40</td>
<td>-4.39*</td>
<td>-4.01*</td>
<td>-3.84</td>
<td>-3.69</td>
<td>-4.22</td>
</tr>
<tr>
<td>5</td>
<td>-2.42</td>
<td>-2.26</td>
<td>-3.40</td>
<td>-4.69*</td>
<td>-4.29*</td>
<td>-3.84</td>
<td>-4.00**</td>
<td>-4.22</td>
</tr>
<tr>
<td>$\hat{\rho}$</td>
<td>0.86</td>
<td>0.88</td>
<td>0.62</td>
<td>0.67</td>
<td>0.70</td>
<td>0.70</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* = reject null at 5% level
** = reject null at 10% level
is that the statistics test different null hypotheses. The classical tests supports the null of no cointegration unless there is strong evidence to the contrary, whereas the \( LM \) tests support the null of cointegration unless there is strong evidence against parameter constancy. Note that the \( LM_3 \) coefficient test shows more discriminatory capability than the \( LM-i \) intercept test, since it supports a deviation from the null of parameter constancy in the real regression equation.\(^4\)

The suggestion that perhaps it is perceived income rather than disposable income that is more relevant to consumption decisions leads us to test for a cointegrating relation in equation (45). From Table 7 we see that the \( LM \) and classical tests for both nominal and real variables agree with Phillips' earlier result that consumption in the long-run is affected by perceived income. These findings suggest that equation (45) better describes a long-run relationship for both nominal and real variables between consumption and income rather than the relation (42). Figures 3(i)–(ii) plot the residuals from the two regressions. The behavior of the residuals for the nominal and real case are similar over the sample period in Figure 3(ii) but differ substantially in Figure 3(i). Thus, the cointegration found for the nominal case (45.N) should also hold in the real case (45.R), and the outcome from Table 7 seems to support this visual evidence.

A final empirical formulation is to alter the steady-state solution by the addition of the wealth effect in the regression equation

\[
\ln(c) = \alpha + \beta \ln(y_d) + \gamma \text{ loss} + \delta \ln(m_3) + u_r. \tag{46.R}
\]

An interesting outcome of Phillips' analysis is that although (46.R) is cointegrating, Bayesian posteriors support the specification of (45.R) to (46.R) and show an appreciable probability of misspecification in equation (46.R) (i.e. no cointegration among the four variables). Our test results support this analysis and are reported in the last panel of Table 7. Both the \( LM \) tests cannot reject the null of cointegration, while the other tests cannot reject the null of no cointegration at the 5\% level, reflecting, in part, the problem with classical hypothesis – testing (\( \hat{\rho}_R = .702 \)). We perform a subset test on the parameter \( \delta \) to find out whether the conflicting results presented above can be explained by the instability in the real money stock. The value of our \( LM_3 \) test equals .02, and with critical value = .0923 at the 5\% level (see Table 4.B(b)) we cannot confirm any instability of the real money stock coefficient. Still, the inclusion of an additional, insignificant (\( t\)-ratio = 1.65) variable in a cointegrating equation carries some cost, especially in small samples where the additional variable in the regression introduces additional uncertainty in the regression residuals. This is, in fact, illustrated in Figure 3(iii), where the behavior of the residuals from (46.R) exhibit more variance over the sample period and differ substantially from the residuals of (45.R)

\(^4\) When the test was applied to the full sample data, the \( LM \) test could not reject the hypothesis of cointegration for the real variables. One possible explanation is that the noise in the series is toward the middle of the sample, and extending the data to accommodate the periods with less noise gives the test lesser power to detect deviations from the null of stability.
Fig. 3. Cointegrating regression residuals
in some periods (notably 1974–1976). These results support Phillips' conclusion that the form of the consumption function is best described by equation (45).\footnote{We have also considered an alternative formulation where a trend term is included in regressions (42), (45) and (46.R) so that we test for cointegration between the stochastic components only. But the coefficients of income ($\beta$) in the regressions with the real variables ranged from .60 to .70. We prefer the economically more plausible coefficients that are closer to 1.}

7 Conclusion

This paper has developed tests for parameter constancy in cointegrating regressions. These tests provide an alternative approach to testing for cointegration in time series regression models. The tests considered are the one-sided version of the $LM$ test commonly used to test for coefficient stability in regression equations. A limit theory for the tests is developed allowing for deterministic trends and multivariate regressors. The distributions are non-standard but are free of nuisance parameters and in the multivariate case, depend only on the rank of certain regressors. A small Monte Carlo study suggests that the $LM$ test is
adequate for use in empirical testing with sample sizes greater than 100 and asymptotic sizes chosen at the 5% or 10% levels. The tests are easy to use and have optimal asymptotic properties, being in fact extensions of the locally best invariant tests for parameter constancy used in models with fixed regressors.

The tests are constructed so that the null under test is the hypothesis of cointegration. This formulation is useful in many applications because it is the hypothesis of cointegration rather than that of no cointegration that is often of primary interest. Such a formulation accords satisfactorily with classical hypothesis testing, since the hypothesis of cointegration is sharp in our formulation, whereas the hypothesis of no cointegration is composite, including a vast class of integrated processes. Formal statistical tests of the form suggested in this paper are helpful because they enable researchers to detect the presence of cointegration and help to determine the sources of cointegration failure when the tests reject.

The theory presented in this paper deals with the case where the stochastic regressors are not themselves cointegrated. Extension of the theory to the case where the regressors are cointegrated is more complex. Both they have been worked out and will be presented in later work.

References

Chu CS James, White H (1990) Testing for structural change in some simple time series models. Manuscript, University of California, San Diego
Davies RB (1987) Hypothesis testing when a nuisance parameter is present only under the alternative. Biometrika 74: 33–43
Hansen BE (1990) Testing for structural change of unknown form in models with non-stationary regressors. Manuscript, University of Rochester


Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economics time series have a unit root?. Journal of Econometrics 54: 154–178


