Bounds Testing Approaches to the Analysis of Long Run Relationships*

M. Hashem Pesaran
Trinity College, Cambridge

Yongcheol Shin
Department of Economics, University of Edinburgh

Richard J. Smith
Department of Economics, University of Bristol

February 1999

Abstract

This paper develops a new approach to the problem of testing the existence of a long-run level relationship between a dependent variable and a set of regressors, when it is not known with certainty whether the underlying regressors are trend- or first-difference stationary. The proposed tests are based on standard F- and t- statistics used to test the significance of the lagged levels of the variables in a first-difference regression. The asymptotic distributions of these statistics are non-standard under the null hypothesis that there exists no level relationship between the dependent variable and the included regressors, irrespective of whether the regressors are I(0) or I(1). Two sets of asymptotic critical values are provided: One set assuming that all the regressors are *I(1)*, and another set assuming that they are all *I(0)*. These two sets of critical values provide a band covering all possible classifications of the regressors into *I(0)*, *I(1)* or mutually cointegrated. Accordingly, various bounds testing procedures are proposed. It is shown that the proposed tests are consistent, and their asymptotic distribution under the null and suitably defined local alternatives are derived. The empirical relevance of the bounds procedures are demonstrated by a re-examination of the earnings equation included in the UK Treasury macroeconometric model. This is a particularly relevant application as there is considerable doubt concerning the order of integration of the variables such as the unemployment rate, the union strength and the wedge between the "real product wage" and the "real consumption wage" that enter the earnings equation.

JEL Classification: C12, C22, C32.
Key Words: Long-Run Relationship, Unrestricted Error Correction Model, Cointegration, Unit Roots, Bounds Tests, Critical Value Bounds, Asymptotic Local Power, Earnings Equation.

*We are grateful to Michael Binder, Clive Granger, Brian Henry, Joon-Yong Park and Ron Smith for helpful comments. Partial financial support from the ESRC (grant No. R000233608) and the Isaac Newton Trust of Trinity College, Cambridge, is gratefully acknowledged.

1 Introduction

Over the past decade considerable attention has been paid in empirical economics to testing for the existence of long-run relations, mainly using cointegration techniques. There have been two main approaches: the two-step residual-based procedure for testing the null of no-cointegration (see Engle and Granger (1987), and Phillips and Ouliaris (1990)), and the system-based reduced rank regression approach due to Johansen (1991,1995). There are also other procedures such as the variable addition approach of Park (1990), the residual-based procedure for testing the null of cointegration by Shin (1994), and the stochastic common trends (system) approach of Stock and Watson (1988).

All these testing procedures require the underlying variables to be integrated of order 1; or $I(1)$. This inevitably involves a certain degree of pre-testing, thus introducing a further degree of uncertainty into the analysis of long-run relations. (See, for example, Cavanagh *et al.* (1995)).

In this paper we propose a new approach to testing for the existence of a long-run relationship which is applicable irrespective of whether the underlying regressors are $I(0)$, $I(1)$ or mutually cointegrated. The statistic underlying our procedure is the familiar Wald or F-statistic in a generalized Dicky-Fuller type regression used to test the significance of lagged levels of the variables under consideration in an unrestricted error correction regression. We show that the asymptotic distributions of both statistics are non-standard under the null hypothesis that there exists no relationship between the levels of the included variables; irrespective of whether the regressors are $I(0)$, $I(1)$ or mutually cointegrated. We establish that the proposed test is consistent, and derive its asymptotic distribution under the null and suitably defined local alternatives, again for a mixture of $I(0)/I(1)$ set of regressors.

We provide two sets of asymptotic critical values for the two polar cases: one which assumes that all the regressors are $I(1)$, and the other assuming that they are $I(0)$. Since these two sets of critical values provide *critical value bounds* for all classifications of the regressors into $I(1)$ and/or $I(0)$, we propose a bounds testing procedure. If the computed Wald or F-statistic falls outside the critical value bounds, a conclusive inference can be drawn without needing to know whether the underlying regressors are $I(1)$, cointegrated amongst themselves or individually $I(0)$. However, if the Wald or F-statistic falls inside the critical values' band, inference would be inconclusive and knowledge of the order of the integration of the underlying variables will be needed before conclusive inferences can be made. We also apply the bounds procedure to the cointegration test proposed in Banerjee, Dolado and Mestre (1998) which is based on the t-ratio of the coefficient of the lagged dependent variable in an augmented autoregressive distributed lag (ARDL) model. We derive the asymptotic distribution of this t-statistic both in the case where all the regressors are $I(1)$, the case considered by these authors; and when one or more of the regressors are individually $I(0)$, or are mutually cointegrated. We provide the relevant critical value bounds for this t-test as well.

The empirical relevance of the bounds procedure is demonstrated by a re-examination of the earnings equation included in the UK Treasury macroeconometric model. This is a particularly relevant application as there is considerable doubt concerning the order of integration of the variables such as the degree of unionization of the work force, the replacement ratio (unemployment benefit-wage ratio) and the wedge between the "real product wage" and the "real consumption wage" that typically enter the earnings equation. There is another consideration in the choice of this application. Under the influence of the seminal contributions of Phillips (1958) and Sargan (1964) econometric analysis of wages and earnings has played an important role in the development of time series econometrics in the UK. The work of Sargan is particularly noteworthy as it is one of the first to articulate and apply the error correction mechanism to wage rate determination. Sargan, however, did not consider the problem of testing the existence of a long-run relationship between real wages and its determinants (which he considered to be the unemployment rate, the index of labour productivity in manufacturing, the ratio of consumption expenditures at market prices to consumption expenditures at factor costs, and a linear time trend).

The long-run *level* relationship underlying Treasury's earning equation relates real average earnings of the private sector to labour productivity, the unemployment rate, an index of union density, a wage variable (comprising a tax wedge and an import price wedge) and the replacement ratio (defined as the ratio of the unemployment benefit to the wage rate). These are the variables predicted by the bargaining theory of wage determination reviewed, for example, in Layard, Nickell and Jackman (1991). We estimated a number of ARDL models in these five variables and found that once a sufficiently high order is selected for the conditional model, the hypothesis that there exists no long-run level relationship between these variables is rejected; irrespective of whether they are $I(0)$ or $I(1)$. Having established the existence of a long-run level relationship between these variables, we then use the ARDL modelling approach advanced in Pesaran and Shin (1999) to estimate our preferred error correction model of average earnings. In our analysis the identification problem discussed by Manning (1993) is approached by assuming that the level of the unemployment rate enters the wage equation, but not *vice versa*. This assumption, of course, does not preclude the rate of change of earnings entering the unemployment equation, or there being other long-run level relationships between the remaining four variables. Our approach accommodates both possibilities.

The plan of the paper is as follows: Section 2 sets out the underlying model and addresses the issues involved in testing for the existence of long-run level relationships. Section 3 considers the Wald statistic (or the F-statistic) for testing the hypothesis that there exists no long-run level relation between the variables under consideration and derives the associated asymptotic theory. Section 4 discusses the power properties of the proposed test. Section 5 describes the empirical application. Section 6 provides some concluding remarks.

The following notations will be used. The symbol $\Rightarrow$ signifies "weak convergence in probability measure," $\mathbf{I}_m$ "an identity matrix of order m," $I(d)$ "integrated of order d," $O_P(K)$ "of the same order as K in probability" and $o_P(K)$ "of smaller order than K in probability".

2 The Underlying VAR Model and Assumptions

Let $\{\mathbf{z}_t\}_{t=1}^{\infty}$ denote a $(k+1)$-vector random process. The data generating process for $\{\mathbf{z}_t\}_{t=1}^{\infty}$ is the vector autoregressive model of order p (VAR(p)):

$$\Phi(L)(\mathbf{z}_t - \boldsymbol{\mu} - \boldsymbol{\gamma} t) = \boldsymbol{\varepsilon}_t,\ t = 1, 2, ..., \tag{2.1}$$

where L is the lag operator, $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ are unknown $(k+1)$-vectors of intercepts and trend coefficients, the $(k+1, k+1)$ matrix lag polynomial $\Phi(L) = \mathbf{I}_{k+1} - \sum_{i=1}^{p} \Phi_i L^i$ with $\{\Phi_i\}_{i=1}^{p}$ $(k+1, k+1)$ matrices of unknown coefficients. The properties of the $(k+1)$-vector error process $\{\boldsymbol{\varepsilon}_t\}_{t=1}^{\infty}$ are given in Assumption 2 below. All the analysis of this paper is conducted given the initial observations $\mathbf{Z}_0 \equiv (\mathbf{z}_{1-p}, ..., \mathbf{z}_0)$. We make the following assumptions.

Assumption 1. The roots of $|\mathbf{I}_{k+1} - \sum_{i=1}^{p} \Phi_i z^i| = 0$ are either outside the unit circle $|z| = 1$ or satisfy $z = 1$.

Assumption 2. The vector error process $\{\boldsymbol{\varepsilon}_t\}_{t=1}^{\infty}$ is $IN(\mathbf{0}, \Omega)$, Ω positive definite.

Assumption 1 permits the elements of $\mathbf{z}_t$ to be $I(1)$, $I(0)$ or cointegrated but excludes the possibility of seasonal unit roots and explosive roots.[1] Assumption 2 may be relaxed somewhat to permit $\{\boldsymbol{\varepsilon}_t\}_{t=1}^{\infty}$ to be a conditionally mean zero and homoskedastic process; see, for example, Pesaran, Shin and Smith (1998, Assumption 4.1).

We may re-express the lag polynomial $\Phi(L)$ in vector error correction form; *viz.*

$$\Phi(L) \equiv -\Pi L + \Gamma(L)(1 - L). \tag{2.2}$$

In (2.2), the long-run multiplier matrix is defined by

$$\Pi \equiv -\left(\mathbf{I}_m - \sum_{i=1}^{p} \Phi_i\right) \tag{2.3}$$

and the short-run response matrix lag polynomial $\Gamma(L) \equiv \mathbf{I}_m - \sum_{i=1}^{p-1} \Gamma_i L^i$, $\Gamma_i = -\sum_{j=i+1}^{p} \Phi_j$, $i = 1, ..., p-1$. Hence, the VAR(p) model (2.1) may be rewritten in vector error correction form as

$$\Delta \mathbf{z}_t = \mathbf{a}_0 + \mathbf{a}_1 t + \Pi \mathbf{z}_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta \mathbf{z}_{t-i} + \boldsymbol{\varepsilon}_t, \quad t = 1, 2, ..., \tag{2.4}$$

where $\Delta \equiv 1 - L$ is the difference operator,

$$\mathbf{a}_0 \equiv -\Pi \boldsymbol{\mu} + (\Gamma + \Pi)\boldsymbol{\gamma},\ \mathbf{a}_1 \equiv -\Pi \boldsymbol{\gamma}, \tag{2.5}$$

and the sum of the short-run coefficient matrices

$$\Gamma \equiv \mathbf{I}_m - \sum_{i=1}^{p-1} \Gamma_i = -\Pi + \sum_{i=1}^{p} i \Phi_i. \tag{2.6}$$

[1]Assumptions 5a and 5b below further restrict the maximal order of integration of $\{\mathbf{z}_t\}_{t=1}^{\infty}$ to unity.

As detailed in Pesaran, Shin and Smith (1998, Section 2), if $\boldsymbol{\gamma} \neq \mathbf{0}$, the restrictions (2.5) on the trend coefficients $\mathbf{a}_1$ in (2.4) ensure that the deterministic trending behaviour of the level process $\{\mathbf{z}_t\}_{t=1}^{\infty}$ is invariant to the (cointegrating) rank of Π; a similar result holds for the intercept of $\{\mathbf{z}_t\}_{t=1}^{\infty}$ if $\boldsymbol{\mu} \neq \mathbf{0}$ and $\boldsymbol{\gamma} = \mathbf{0}$.

The interest of this paper concerns the conditional modelling of the *scalar* process y_t given the k-vector $\mathbf{x}_t$ and the past values $\{\mathbf{z}_{t-i}\}_{i=1}^{t-1}$ and $\mathbf{Z}_0$, where we have partitioned $\mathbf{z}_t = (y_t, \mathbf{x}_t')'$. Partitioning the error term $\boldsymbol{\varepsilon}_t$ conformably with $\mathbf{z}_t = (\mathbf{y}_t', \mathbf{x}_t')'$ as $\boldsymbol{\varepsilon}_t = (\varepsilon_{yt}, \boldsymbol{\varepsilon}_{xt}')'$ and its variance matrix as

$$\Omega = \begin{pmatrix} \omega_{yy} & \boldsymbol{\omega}_{yx} \\ \boldsymbol{\omega}_{xy} & \Omega_{xx} \end{pmatrix}$$

we may express ε_{yt} conditionally in terms of $\boldsymbol{\varepsilon}_{xt}$ as

$$\varepsilon_{yt} = \boldsymbol{\omega}_{yx}\Omega_{xx}^{-1}\boldsymbol{\varepsilon}_{xt} + u_t, \tag{2.7}$$

where $u_t \sim IN(0, \omega_{uu})$, $\omega_{uu} \equiv \omega_{yy} - \boldsymbol{\omega}_{yx}\Omega_{xx}^{-1}\boldsymbol{\omega}_{xy}$ and u_t is independent of $\boldsymbol{\varepsilon}_{xt}$. Substitution of (2.7) into (2.4) together with a similar partitioning of the parameter vectors and matrices $\mathbf{a}_0 = (a_{y0}, \mathbf{a}_{x0}')'$, $\mathbf{a}_1 = (a_{y1}, \mathbf{a}_{x1}')'$, $\Pi = (\boldsymbol{\pi}_y', \Pi_x')'$, $\Gamma = (\boldsymbol{\gamma}_y', \Gamma_x')'$, $\Gamma_i = (\boldsymbol{\gamma}_{yi}', \Gamma_{xi}')'$, $i = 1, ..., p-1$, provides a conditional model for Δy_t in terms of $\mathbf{z}_{t-1}, \Delta\mathbf{x}_t, \Delta\mathbf{z}_{t-1}, \Delta\mathbf{z}_{t-2}, ...$; *viz.*

$$\Delta y_t = c_0 + c_1 t + \boldsymbol{\pi}_{y.x}\mathbf{z}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i'\Delta\mathbf{z}_{t-i} + \boldsymbol{\omega}'\Delta\mathbf{x}_t + u_t, \tag{2.8}$$

$t = 1, 2, ...,$ where $\boldsymbol{\omega} \equiv \Omega_{xx}^{-1}\boldsymbol{\omega}_{xy}$, $c_0 \equiv a_{y0} - \boldsymbol{\omega}'\mathbf{a}_{x0}$, $c_1 \equiv a_{y1} - \boldsymbol{\omega}'\mathbf{a}_{x1}$, $\boldsymbol{\psi}_i' \equiv \boldsymbol{\gamma}_{yi} - \boldsymbol{\omega}'\Gamma_{xi}$, $i = 1, ..., p-1$, and $\boldsymbol{\pi}_{y.x} \equiv \boldsymbol{\pi}_y - \boldsymbol{\omega}'\Pi_x$. The deterministic relations (2.5) are modified to

$$c_0 = -\boldsymbol{\pi}_{y.x}\boldsymbol{\mu} + (\boldsymbol{\gamma}_{y.x} + \boldsymbol{\pi}_{y.x})\boldsymbol{\gamma},\ c_1 = -\boldsymbol{\pi}_{y.x}\boldsymbol{\gamma}, \tag{2.9}$$

where $\boldsymbol{\gamma}_{y.x} \equiv \boldsymbol{\gamma}_y - \boldsymbol{\omega}'\Gamma_x$.

We now partition the long-run multiplier matrix Π conformably with $\mathbf{z}_t = (\mathbf{y}_t', \mathbf{x}_t')'$ as

$$\Pi = \begin{pmatrix} \pi_{yy} & \boldsymbol{\pi}_{yx} \\ \boldsymbol{\pi}_{xy} & \Pi_{xx} \end{pmatrix}.$$

The next assumption is critical for the analysis of this paper.

Assumption 3. The k-vector $\boldsymbol{\pi}_{xy} = \mathbf{0}$.

Under Assumption 3,

$$\Delta\mathbf{x}_t = \mathbf{a}_{x0} + \mathbf{a}_{x1}t + \Pi_{xx}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1}\Gamma_{xi}\Delta\mathbf{z}_{t-i} + \boldsymbol{\varepsilon}_{xt}, \tag{2.10}$$

$t = 1, 2, ...$ Therefore, we may regard the process $\{\mathbf{x}_t\}_{t=1}^{\infty}$ as *long-run forcing* for $\{y_t\}_{t=1}^{\infty}$ as there is no feedback from the *level* of y_t in (2.10); see Granger and Lin (1995).[2] Assumption

[2]Note that this restriction does not preclude $\{\mathbf{y}_t\}_{t=1}^{\infty}$ being *Granger-causal* for $\{\mathbf{x}_t\}_{t=1}^{\infty}$ in the *short-run*.

3 ensures that there exists *at most* one long-run level relationship between y_t and $\mathbf{x}_t$ which includes *both* y_t *and* $\mathbf{x}_t$, irrespective of the level of integration of the process $\{\mathbf{x}_t\}_{t=1}^{\infty}$; see (2.13) below.

Under Assumption 3, the conditional error correction model (2.8) now becomes

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t, \tag{2.11}$$

$t = 1, 2, ...,$ where

$$c_0 = -(\pi_{yy}, \boldsymbol{\pi}_{yx.x})\boldsymbol{\mu} + [\boldsymbol{\gamma}_{y.x} + (\pi_{yy}, \boldsymbol{\pi}_{yx.x})]\boldsymbol{\gamma},\ c_1 = -(\pi_{yy}, \boldsymbol{\pi}_{yx.x})\boldsymbol{\gamma} \tag{2.12}$$

and $\boldsymbol{\pi}_{yx.x} \equiv \boldsymbol{\pi}_{yx} - \boldsymbol{\omega}'\Pi_{xx}$.[3]

The next assumption together with Assumptions 5a and 5b below which constrain the maximal order of integration of the system (2.11) and (2.10) to be unity defines the cointegration properties of the system.

Assumption 4. The matrix Π_{xx} has rank r, $0 \leq r \leq k$.

Therefore, under Assumption 4, from (2.10), we may express Π_{xx} as

$$\Pi_{xx} = \boldsymbol{\alpha}_{xx}\boldsymbol{\beta}_{xx}',$$

where $\boldsymbol{\alpha}_{xx}$ and $\boldsymbol{\beta}_{xx}$ are both (k, r) matrices of full column rank; see, for example, Engle and Granger (1987) and Johansen (1991). If the maximal order of integration of the system (2.11) and (2.10) is unity, under Assumptions 1, 3 and 4, the process $\{\mathbf{x}_t\}_{t=1}^{\infty}$ is mutually cointegrated of order r, $0 \leq r \leq k$. However, in contradistinction to, for example, Banerjee *et al.* (1998) who fix $r = 0$, we do not wish to impose an *a priori* specification of r. When $\boldsymbol{\pi}_{xy} = \mathbf{0}$ *and* $\Pi_{xx} = \mathbf{0}$, then $\mathbf{x}_t$ is weakly exogenous for the coefficients π_{yy} and $\boldsymbol{\pi}_{yx.x} = \boldsymbol{\pi}_{yx}$ in (2.11); see, for example, Johansen (1995, Theorem 8.1, p.122). Moreover, in the more general case where Π_{xx} is non-zero, as π_{yy} and $\boldsymbol{\pi}_{yx.x} = \boldsymbol{\pi}_{yx} - \boldsymbol{\omega}'\Pi_{xx}$ are variation-free from the parameters in (2.10), $\mathbf{x}_t$ is also weakly exogenous for the parameters of (2.11).

Note that under Assumption 4 the maximal cointegrating rank of the long-run multiplier matrix Π for the system (2.11) and (2.10) is $r + 1$ and the minimal cointegrating rank of Π is r. The next assumptions provide the conditions for the maximal order of integration of the system (2.11) and (2.10) to be unity. Firstly, we consider the requisite conditions for the case in which $rank(\Pi) = r$. In this case, under Assumptions 1, 3 and 4, $\pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx} - \boldsymbol{\phi}'\Pi_{xx} = \mathbf{0}'$ for some k-vector $\boldsymbol{\phi}$. Note that $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ implies the latter condition. Thus, under Assumptions 1, 3 and 4, the long-run multiplier matrix Π has rank r and is given by

$$\Pi = \begin{pmatrix} 0 & \boldsymbol{\pi}_{yx} \\ \mathbf{0} & \Pi_{xx} \end{pmatrix}.$$

[3] Pesaran, Shin and Smith (1998) and Harbo *et al.* (1998) consider a similar model but where $\mathbf{x}_t$ are $I(1)$; that is, under the additional assumption $\Pi_{xx} = \mathbf{0}$.

Hence, we may express $\Pi = \boldsymbol{\alpha}\boldsymbol{\beta}'$ where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $(k+1, r)$ matrices of full column rank and

$$\boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{\alpha}_{yx} \\ \boldsymbol{\alpha}_{xx} \end{pmatrix}, \boldsymbol{\beta} = \begin{pmatrix} \mathbf{0}' \\ \boldsymbol{\beta}_{xx} \end{pmatrix}.$$

Let the columns of the $(k+1, k-r+1)$ matrices $(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})$ and $(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})$, where $\boldsymbol{\alpha}_y^{\perp}$, $\boldsymbol{\beta}_y^{\perp}$ and $\boldsymbol{\alpha}^{\perp}$, $\boldsymbol{\beta}^{\perp}$ are respectively $(k+1)$-vectors and $(k+1, k-r)$ matrices, denote bases for the orthogonal complements of respectively $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$; in particular, $(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\boldsymbol{\alpha} = \mathbf{0}$ and $(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})'\boldsymbol{\beta} = \mathbf{0}$.

Assumption 5a. If $rank(\Pi) = r$, the matrix $(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\Gamma(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})$ is full rank $k-r+1$, $0 \leq r \leq k$.

Cf. Johansen (1991, Theorem 4.1, p.1559).

Secondly, if the long-run multiplier matrix Π has rank $r+1$, then under Assumptions 1, 3 and 4, $\pi_{yy} \neq 0$ and Π may be expressed as $\Pi = \boldsymbol{\alpha}_y\boldsymbol{\beta}_y' + \boldsymbol{\alpha}\boldsymbol{\beta}'$, where $\boldsymbol{\alpha}_y = (\alpha_{yy}, \mathbf{0}')'$ and $\boldsymbol{\beta}_y = (\beta_{yy}, \boldsymbol{\beta}_{yx}')'$ are $(k+1)$-vectors, the former of which preserves Assumption 3. For this case, the columns of $\boldsymbol{\alpha}^{\perp}$ and $\boldsymbol{\beta}^{\perp}$ form respective bases for the orthogonal complements of $(\boldsymbol{\alpha}_y, \boldsymbol{\alpha})$ and $(\boldsymbol{\beta}_y, \boldsymbol{\beta})$; in particular, $\boldsymbol{\alpha}^{\perp\prime}(\boldsymbol{\alpha}_y, \boldsymbol{\alpha}) = \mathbf{0}$ and $\boldsymbol{\beta}^{\perp\prime}(\boldsymbol{\beta}_y, \boldsymbol{\beta}) = \mathbf{0}$.

Assumption 5b. If $rank(\Pi) = r+1$, the matrix $\boldsymbol{\alpha}^{\perp\prime}\Gamma\boldsymbol{\beta}^{\perp}$ is full rank $k-r$, $0 \leq r \leq k$.

Assumptions 1, 3, 4 and 5a and 5b permit the two polar cases for the $\{\mathbf{x}_t\}_{t=1}^{\infty}$ process. Firstly, if $\{\mathbf{x}_t\}_{t=1}^{\infty}$ is an $I(0)$ vector process, then Π_{xx}, and, hence, $\boldsymbol{\alpha}_{xx}$ and $\boldsymbol{\beta}_{xx}$, are nonsingular. Secondly, if $\{\mathbf{x}_t\}_{t=1}^{\infty}$ is an $I(1)$ vector process, then $\Pi_{xx} = \mathbf{0}$, and, hence, $\boldsymbol{\alpha}_{xx}$ and $\boldsymbol{\beta}_{xx}$ are also null matrices.

Therefore, under Assumptions 1, 3, 4 and 5b, it immediately follows from (2.11) that, if $\pi_{yy} \neq 0$ and $\boldsymbol{\pi}_{yx.x} \neq \mathbf{0}$, there exists a non-degenerate *long-run level relationship* between y_t and $\mathbf{x}_t$ defined by

$$y_t = \theta_0 + \theta_1 t + \boldsymbol{\theta}\mathbf{x}_t + v_t, \; t = 1, 2, ..., \tag{2.13}$$

where $\theta_0 \equiv -c_0/\pi_{yy}$, $\theta_1 \equiv -c_1/\pi_{yy}$ and $\boldsymbol{\theta} \equiv -\boldsymbol{\pi}_{yx.x}/\pi_{yy}$ and $\{v_t\}$ is a zero mean stationary process.[4] Note that the long-run level relationship (2.13) between y_t and $\mathbf{x}_t$, $t = 1, 2, ...$, may be *degenerate* in the sense that $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ is possible as well as the non-degenerate $\boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$. The former possibility is somewhat of an anomaly from an applied perspective as the differenced variable Δy_t then depends on its own lagged level y_{t-1} in the ECM (2.11) but *not* on the lagged levels $\mathbf{x}_{t-1}$ of the forcing variables, $t = 1, 2, ...$ In this case there are no long-run effects running from $\mathbf{x}_t$ to y_t and the long-run relationship of the model only involves y_t and possibly a deterministic trend, $t = 1, 2, ...$

In order to test for the absence of a long-run level relationship between y_t and $\mathbf{x}_t$, the method adopted in this paper is to examine the *joint* hypothesis $\pi_{yy} = 0$ *and* $\boldsymbol{\pi}_{yx.x} = \mathbf{0}$ in

[4] Using (A.1) in Appendix A it is easily seen that $(\pi_{yy}, \boldsymbol{\pi}_{yx.x})\mathbf{z}_t = (\pi_{yy}, \boldsymbol{\pi}_{yx.x})(\boldsymbol{\mu} + \boldsymbol{\gamma}t + \mathbf{C}^*(L)\boldsymbol{\varepsilon}_t)$. Hence, $v_t \equiv (1, \boldsymbol{\theta})\mathbf{C}^*(L)\boldsymbol{\varepsilon}_t$.

the ECM (2.11).[5] In contradistinction, the approach of Banerjee *et al.* (1998) may be simply described in terms of (2.11) using Assumption 5b:

$$\begin{aligned}\Delta y_t &= c_0 + c_1 t + \alpha_{yy}(\beta_{yy} y_{t-1} + \boldsymbol{\beta}'_{yx}\mathbf{x}_{t-1}) + (\boldsymbol{\alpha}_{yx} - \boldsymbol{\omega}'\boldsymbol{\alpha}_{xx})\boldsymbol{\beta}'_{xx}\mathbf{x}_{t-1} \\ &\quad + \sum_{i=1}^{p-1} \boldsymbol{\psi}'_i \Delta\mathbf{z}_{t-i} + \boldsymbol{\omega}'\Delta\mathbf{x}_t + u_t. \end{aligned} \quad (2.14)$$

Banerjee *et al.* (1998) test for the exclusion of y_{t-1} in (2.14) when $r = 0$, that is, $\boldsymbol{\beta}_{xx} = \mathbf{0}$ in (2.14) or $\Pi_{xx} = \mathbf{0}$ in (2.10) and, thus, $\{\mathbf{x}_t\} \sim I(1)$; cf. Harbo *et al.* (1998) and Pesaran, Shin and Smith (1998). Effectively, therefore, Banerjee *et al.* (1998) consider the hypothesis $\alpha_{yy} = 0$ (or $H_0^{\pi_{yy}} : \pi_{yy} = 0$).[6] More generally though, when $0 < r \leq k$, we require the imposition of the subsidiary hypothesis $\boldsymbol{\alpha}_{yx} - \boldsymbol{\omega}'\boldsymbol{\alpha}_{xx} = \mathbf{0}'$; that is, the limiting distribution of the Banerjee *et al.* (1998) test is obtained under the *joint* hypothesis $\pi_{yy} = 0$ *and* $\boldsymbol{\pi}_{yx.x} = \mathbf{0}$, in (2.11).

In the following sections of the paper, we focus on (2.11) and differentiate between five cases of interest delineated according to the specification of the deterministics; *viz.*

- **Case I:** (No Intercepts; No Trends.) $c_0 = 0$ and $c_1 = 0$. That is, $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\gamma} = \mathbf{0}$. Hence, the ECM (2.11) becomes

$$\Delta y_t = \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}'_i \Delta\mathbf{z}_{t-i} + \boldsymbol{\omega}'\Delta\mathbf{x}_t + u_t. \quad (2.15)$$

- **Case II:** (Restricted Intercepts; No Trends.) $c_0 = -(\pi_{yy}, \boldsymbol{\pi}_{yx.x})\boldsymbol{\mu}$ and $c_1 = 0$. Here, $\boldsymbol{\gamma} = \mathbf{0}$. The ECM is

$$\Delta y_t = \pi_{yy}(y_{t-1} - \mu_y) + \boldsymbol{\pi}_{yx.x}(\mathbf{x}_{t-1} - \boldsymbol{\mu}_x) + \sum_{i=1}^{p-1} \boldsymbol{\psi}'_i \Delta\mathbf{z}_{t-i} + \boldsymbol{\omega}'\Delta\mathbf{x}_t + u_t, \quad (2.16)$$

  where $\boldsymbol{\mu} \equiv (\mu_y, \boldsymbol{\mu}'_x)'$ is partitioned conformably with $\mathbf{z}_t = (y_t, \mathbf{x}'_t)'$.

- **Case III:** (Unrestricted Intercepts; No Trends.) $c_0 \neq 0$ and $c_1 = 0$. Again, $\boldsymbol{\gamma} = \mathbf{0}$. In this case, the intercept restriction $c_0 = -(\pi_{yy}, \boldsymbol{\pi}_{yx.x})\boldsymbol{\mu}$ is ignored and the ECM estimated is

$$\Delta y_t = c_0 + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}'_i \Delta\mathbf{z}_{t-i} + \boldsymbol{\omega}'\Delta\mathbf{x}_t + u_t. \quad (2.17)$$

---

[5] This joint hypothesis may also be justified by the application of Roy's union-intersection principle to tests of $\pi_{yy} = 0$ in (2.11) given $\boldsymbol{\pi}_{yx.x}$. Let $W_{\pi_{yy}}(\boldsymbol{\pi}_{yx.x})$ be the Wald statistic for testing $\pi_{yy} = 0$ for a given value of $\boldsymbol{\pi}_{yx.x}$. The test $\max_{\boldsymbol{\pi}_{yx.x}} W_{\pi_{yy}}(\boldsymbol{\pi}_{yx.x})$ is identical to the Wald test of $\pi_{yy} = 0$ *and* $\boldsymbol{\pi}_{yx.x} = \mathbf{0}$ in (2.11).

[6] Partitioning $\Gamma_{xi} = (\boldsymbol{\gamma}_{xy,i}, \Gamma_{xx,i})$, $i = 1, ..., p-1$, conformably with $\mathbf{z}_t = (y_t, \mathbf{x}'_t)'$, Banerjee *et al.* (1998) also set $\boldsymbol{\gamma}_{xy,i} = \mathbf{0}$, $i = 1, ..., p-1$, which implies $\boldsymbol{\gamma}_{xy} = \mathbf{0}$, where $\Gamma_x = (\boldsymbol{\gamma}_{xy}, \Gamma_{xx})$; that is, Δy_t does *not Granger cause* $\Delta\mathbf{x}_t$.

- **Case IV:** (Unrestricted Intercepts; Restricted Trends.) $c_0 \neq 0$ and $c_1 = -(\pi_{yy}, \boldsymbol{\pi}_{yx.x})\boldsymbol{\gamma}$. Thus

$$\Delta y_t = c_0 + \pi_{yy}(y_{t-1} - \gamma_y t) + \boldsymbol{\pi}_{yx.x}(\mathbf{x}_{t-1} - \boldsymbol{\gamma}_x t) + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t, \tag{2.18}$$

where $\boldsymbol{\gamma} \equiv (\gamma_y, \boldsymbol{\gamma}_x')'$ is partitioned conformably with $\mathbf{z}_t = (y_t, \mathbf{x}_t')'$.

- **Case V:** (Unrestricted Intercepts; Unrestricted Trends.) $\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 \neq \mathbf{0}$. Here, the deterministic trend restriction $c_1 = -(\pi_{yy}, \boldsymbol{\pi}_{yx.x})\boldsymbol{\gamma}$ is ignored and the ECM estimated is

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x} \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t. \tag{2.19}$$

It should be emphasised that the DGPs for Cases II and III are treated as identical as are those for Cases IV and V. However, as in the test for a unit root proposed by Dickey and Fuller (1979) compared with that of Dickey and Fuller (1981) for univariate models, estimation and hypothesis testing in Cases III and V proceed ignoring the constraints linking respectively the intercept and trend coefficient vectors, c_0 and c_1, to the parameter vector $(\pi_{yy}, \boldsymbol{\pi}_{yx.x})$ whereas Cases II and IV fully incorporate the restrictions in (2.12).

In the following exposition, we concentrate on Case IV, that is, (2.18), which may be specialised to yield the remainder.

3 Bounds Tests for a Long-Run Level Relationship

In this section we develop bounds procedures for testing for the existence of a long-run level relationship between the levels of y_t and $\mathbf{x}_t$, $t = 1, 2, ...$, using (2.15)-(2.19); see (2.13). The approach taken here, cf. Engle and Granger (1987) and Banerjee *et al.* (1998), is to test for the *absence* of any long-run relationship between the levels of y_t and $\mathbf{x}_t$, $t = 1, 2, ...$; that is, the exclusion of the lagged level variables y_{t-1} and $\mathbf{x}_{t-1}$ in (2.15)-(2.19). Consequently, we define the constituent null hypotheses

$$H_0^{\pi_{yy}} : \pi_{yy} = 0, H_0^{\boldsymbol{\pi}_{yx.x}} : \boldsymbol{\pi}_{yx.x} = \mathbf{0}'$$

and alternative hypotheses

$$H_1^{\pi_{yy}} : \pi_{yy} \neq 0, H_1^{\boldsymbol{\pi}_{yx.x}} : \boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'.$$

Hence, the joint null hypothesis of interest in (2.15)-(2.19) is given by:

$$H_0 = H_0^{\pi_{yy}} \cap H_0^{\boldsymbol{\pi}_{yx.x}} \tag{3.1}$$

and the alternative hypothesis is correspondingly stated as:

$$H_1 = H_1^{\pi_{yy}} \cup H_1^{\boldsymbol{\pi}_{yx.x}}. \tag{3.2}$$

As indicated in Section 2, not only does the alternative hypothesis H_1 of (3.2) cover the case of interest in which $\boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$ but also permits $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$; cf. (2.11). That is, the possibility of a *degenerate* long-run relationship between the levels of y_t and $\mathbf{x}_t$, $t = 1, 2, ...$, is admitted under H_1 of (3.2).

For ease of exposition, we again consider Case IV and rewrite model (2.18) in matrix notation as

$$\Delta\mathbf{y} = \boldsymbol{\iota}_T c_0 + \mathbf{Z}_{-1}^{*}\boldsymbol{\pi}_{y.x}^{*} + \Delta\mathbf{Z}_{-}\boldsymbol{\psi} + \mathbf{u}, \tag{3.3}$$

where $\boldsymbol{\iota}_T$ is a T-vector of ones, $\Delta\mathbf{y} \equiv (\Delta y_1, ..., \Delta y_T)'$, $\Delta\mathbf{X} \equiv (\Delta\mathbf{x}_1, ..., \Delta\mathbf{x}_T)'$, $\Delta\mathbf{Z}_{-i} \equiv (\Delta\mathbf{z}_{1-i}, ..., \Delta\mathbf{z}_{T-i})'$, $i = 1, ..., p-1$, $\boldsymbol{\psi} \equiv (\boldsymbol{\omega}', \boldsymbol{\psi}_1', ..., \boldsymbol{\psi}_{p-1}')'$, $\Delta\mathbf{Z}_{-} \equiv (\Delta\mathbf{X},\ \Delta\mathbf{Z}_{-1}, ..., \Delta\mathbf{Z}_{1-p})$, $\mathbf{Z}_{-1}^{*} \equiv (\boldsymbol{\tau}_T,\ \mathbf{Z}_{-1})$, $\boldsymbol{\tau}_T \equiv (1, ..., T)'$, $\mathbf{Z}_{-1} \equiv (\mathbf{z}_0, ..., \mathbf{z}_{T-1})'$, $\mathbf{u} \equiv (u_1, ..., u_T)'$ and

$$\boldsymbol{\pi}_{y.x}^{*} = \begin{pmatrix} -\boldsymbol{\gamma}' \\ \mathbf{I}_{k+1} \end{pmatrix} \begin{pmatrix} \pi_{yy} \\ \boldsymbol{\pi}_{yx.x}' \end{pmatrix}.$$

The least squares (LS) estimator of $\boldsymbol{\pi}_{y.x}^{*}$ is given by:

$$\hat{\boldsymbol{\pi}}_{y.x}^{*} \equiv \left(\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_{-}}\bar{\mathbf{Z}}_{-1}^{*}\right)^{-1}\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_{-}}\widetilde{\Delta\mathbf{y}}, \tag{3.4}$$

where $\bar{\mathbf{Z}}_{-1}^{*} \equiv \bar{\mathbf{P}}_{\iota}\mathbf{Z}_{-1}^{*}$, $\widetilde{\Delta\mathbf{Z}}_{-} \equiv \bar{\mathbf{P}}_{\iota}\Delta\mathbf{Z}_{-}$, $\widetilde{\Delta\mathbf{y}} \equiv \bar{\mathbf{P}}_{\iota}\Delta\mathbf{y}$, $\bar{\mathbf{P}}_{\iota} \equiv \mathbf{I}_T - \boldsymbol{\iota}_T(\boldsymbol{\iota}_T'\boldsymbol{\iota}_T)^{-1}\boldsymbol{\iota}_T'$ and $\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_{-}} \equiv \mathbf{I}_T - \widetilde{\Delta\mathbf{Z}}_{-}\left(\widetilde{\Delta\mathbf{Z}}_{-}'\widetilde{\Delta\mathbf{Z}}_{-}\right)^{-1}\widetilde{\Delta\mathbf{Z}}_{-}'$. The Wald and the F-statistics for testing the null hypothesis H_0 of (3.1) against the alternative hypothesis H_1 of (3.2) are respectively:

$$W \equiv \hat{\boldsymbol{\pi}}_{y.x}^{*\prime}\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_{-}}\bar{\mathbf{Z}}_{-1}^{*}\hat{\boldsymbol{\pi}}_{y.x}^{*}/\hat{\omega}_{uu},\ F \equiv \frac{W}{k+2}, \tag{3.5}$$

where

$$\hat{\omega}_{uu} \equiv \frac{1}{T-m}\sum_{t=1}^{T}\tilde{u}_t^2, \tag{3.6}$$

$m \equiv (k+1)(p+1)+1$ is the number of estimated coefficients and $\tilde{u}_t$, $t = 1, 2, ..., T$, are the least squares (LS) residuals from (3.3).

The next theorem presents the asymptotic null distribution of the Wald statistic; the limit behaviour of the F-statistic is a simple corollary and is not presented here or subsequently. Let $\mathbf{W}_{k-r+1}(a) \equiv (W_u(a), \mathbf{W}_{k-r}(a)')'$ denote a $(k-r+1)$-dimensional standard Brownian motion partitioned into the scalar and $(k-r)$-dimensional sub-vector independent standard Brownian motions $W_u(a)$ and $\mathbf{W}_{k-r}(a)$, $a \in [0,1]$. We will also require the corresponding de-meaned $(k-r+1)$-vector standard Brownian motion

$$\bar{\mathbf{W}}_{k-r+1}(a) \equiv \mathbf{W}_{k-r+1}(a) - \int_0^1 \mathbf{W}_{k-r+1}(a)da, \tag{3.7}$$

and de-meaned and de-trended $(m-r)$-vector standard Brownian motion

$$\hat{\mathbf{W}}_{k-r+1}(a) \equiv \bar{\mathbf{W}}_{k-r+1}(a) - 12\left(a - \frac{1}{2}\right)\int_0^1 \left(a - \frac{1}{2}\right)\bar{\mathbf{W}}_{k-r+1}(a)da, \tag{3.8}$$

and their respective partitioned counterparts $\tilde{\mathbf{W}}_{k-r+1}(a) = (\tilde{W}_u(a), \tilde{\mathbf{W}}_{k-r}(a)')'$, and $\hat{\mathbf{W}}_{k-r+1}(a) = (\hat{W}_u(a), \hat{\mathbf{W}}_{k-r}(a)')'$, $a \in [0,1]$.

Theorem 3.1 *(Limiting Distribution of W.) If Assumptions 1-4 and 5a hold, then under $H_0 : \pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ of (3.1), as $T \to \infty$, the asymptotic distribution of the Wald statistic W of (3.5) has the representation*

$$W \Rightarrow \mathbf{z}_r'\mathbf{z}_r + \int_0^1 d\mathrm{W}_u(a)\mathbf{F}_{k-r+1}(a)' \left(\int_0^1 \mathbf{F}_{k-r+1}(a)\mathbf{F}_{k-r+1}(a)'da\right)^{-1} \int_0^1 \mathbf{F}_{k-r+1}(a)d\mathrm{W}_u(a), \tag{3.9}$$

where $\mathbf{z}_r \sim N(\mathbf{0}, \mathbf{I}_r)$ is distributed independently of the second term in (3.9) and

$$\mathbf{F}_{k-r+1}(a) = \left\{ \begin{array}{cc} \mathbf{W}_{k-r+1}(a) & \textbf{Case I} \\ (\mathbf{W}_{k-r+1}(a)', 1)' & \textbf{Case II} \\ \tilde{\mathbf{W}}_{k-r+1}(a) & \textbf{Case III} \\ (\tilde{\mathbf{W}}_{k-r+1}(a)', a - \frac{1}{2})' & \textbf{Case IV} \\ \hat{\mathbf{W}}_{k-r+1}(a) & \textbf{Case V} \end{array} \right\},$$

$r = 0, ..., k$, and Cases I-V are defined in (2.15)-(2.19), $a \in [0,1]$.

The asymptotic distribution of the Wald statistic W of (3.5) depends on the dimension and cointegration rank of the forcing variables $\{\mathbf{x}_t\}$, k and r respectively. In Case IV, referring to (2.14), the first component in (3.9), $\mathbf{z}_r'\mathbf{z}_r \sim \chi^2(r)$, corresponds to testing for the exclusion of the r-dimensional stationary vector $\boldsymbol{\beta}_{xx}'\mathbf{x}_{t-1}$, that is, the hypothesis $\boldsymbol{\alpha}_{yx} - \boldsymbol{\omega}'\boldsymbol{\alpha}_{xx} = \mathbf{0}'$, whereas the second term in (3.9), which is a non-standard Dickey-Fuller unit-root distribution, corresponds to testing for the exclusion of the $(k-r+1)$-dimensional $I(1)$ vector $(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})'\mathbf{z}_{t-1}$ and, in Cases II and IV, the intercept and time-trend respectively or, equivalently, $\alpha_{yy} = 0$.

We specialise Theorem 3.1 to the two polar cases in which, firstly, the process for the forcing variables $\{\mathbf{x}_t\}$ is integrated of order zero, that is, $r = k$ and Π_{xx} is of full rank, and, secondly, the $\{\mathbf{x}_t\}$ process is not mutually cointegrated, $r = 0$, and, hence, the $\{\mathbf{x}_t\}$ process is integrated of order one.

Corollary 3.1 *(Limiting Distribution of W if $\{\mathbf{x}_t\} \sim I(0)$.) If Assumptions 1-4 and 5a hold and $r = k$, that is, $\{\mathbf{x}_t\} \sim I(0)$, then under $H_0 : \pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ of (3.1), as $T \to \infty$, the asymptotic distribution of the Wald statistic W of (3.5) has the representation*

$$W \Rightarrow \mathbf{z}_k'\mathbf{z}_k + \frac{\left(\int_0^1 F(a)d\mathrm{W}_u(a)\right)^2}{\left(\int_0^1 F(a)^2 da\right)}, \tag{3.10}$$

where $\mathbf{z}_k \sim N(\mathbf{0}, \mathbf{I}_k)$ is distributed independently of the second term in (3.10) and

$$F(a) = \left\{ \begin{array}{cc} \mathrm{W}_u(a) & \textbf{Case I} \\ (\mathrm{W}_u(a), 1)' & \textbf{Case II} \\ \tilde{\mathrm{W}}_u(a) & \textbf{Case III} \\ (\tilde{\mathrm{W}}_u(a), a - \frac{1}{2})' & \textbf{Case IV} \\ \hat{\mathrm{W}}_u(a) & \textbf{Case V} \end{array} \right\},$$

$r = 0, ..., k$, where Cases I-V are defined in (2.15)-(2.19), $a \in [0,1]$.

Corollary 3.2 *(Limiting Distribution of W if $\{\mathbf{x}_t\} \sim I(1)$.) If Assumptions 1-4 and 5a hold and $r = 0$, that is, $\{\mathbf{x}_t\} \sim I(1)$, then under $H_0 : \pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ of (3.1), as $T \to \infty$, the asymptotic distribution of the Wald statistic W of (3.5) has the representation*

$$W \Rightarrow \int_0^1 d\mathrm{W}_u(a)\mathbf{F}_{k+1}(a)' \left(\int_0^1 \mathbf{F}_{k+1}(a)\mathbf{F}_{k+1}(a)'da \right)^{-1} \int_0^1 \mathbf{F}_{k+1}(a)d\mathrm{W}_u(a),$$

where $\mathbf{F}_{k+1}(a)$ is defined in Theorem 3.1 for Cases I-V, $a \in [0,1]$.

See also Boswijk (1992).

In practice, however, it is unlikely that one would possess *a priori* knowledge of the rank r of Π_{xx}; that is, the cointegration rank of the forcing variables $\{\mathbf{x}_t\}$ or, more particularly, whether $\{\mathbf{x}_t\} \sim I(0)$ or $\{\mathbf{x}_t\} \sim I(1)$. Long-run analysis of (2.15)-(2.19) predicated on a prior determination of the cointegration rank r in (2.10) is prone to the possibility of a pre-test specification error; see, for example, Cavanagh *et al.* (1995). However, it may be shown by simulation that the asymptotic critical values obtained from Corollaries 3.1 ($r = k$ and $\{\mathbf{x}_t\} \sim I(0)$) and 3.2 ($r = 0$ and $\{\mathbf{x}_t\} \sim I(1)$) provide lower and upper bounds respectively for those corresponding to the general case considered in Theorem 3.1 when the cointegration rank of the forcing variables $\{\mathbf{x}_t\}$ process is $0 \leq r \leq k$.[7] Hence, these two sets of critical values provide *critical value bounds* covering all possible classifications of $\{\mathbf{x}_t\}$ into $I(0)$, $I(1)$ and mutually cointegrated processes. Therefore, Tables C1.i-C1.v provide two sets of asymptotic critical values for the F-statistics covering Cases I-V; one set assuming that the forcing variables $\{\mathbf{x}_t\}$ are $I(0)$ and the other assuming that $\{\mathbf{x}_t\}$ are $I(1)$.[8]

Hence, we suggest a *bounds procedure* to test $H_0 : \pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ of (3.1), that is, the absence of a long-run level relationship between y_t and $\mathbf{x}_t$, within the ECMs (2.15)-(2.19). If the computed Wald or F- statistics fall outside the critical value bounds, a conclusive decision results without needing to know the cointegration rank r of the $\{\mathbf{x}_t\}$ process. If, however, the Wald or F- statistic fall within these bounds, inference would be inconclusive. In such circumstances, knowledge of the cointegration rank r of the forcing variables $\{\mathbf{x}_t\}$ is required to proceed further.

The ECM (2.15)-(2.19), derived from the underlying VAR(p) model (2.4), may also be interpreted as an autoregressive distributed lag model of orders $(p, p, ..., p)$ (ARDL$(p, ..., p)$). However, one could also allow for differential lag lengths on the lagged variables y_{t-i} and $\mathbf{x}_{t-i}$ in (2.4) to arrive at, for example, an ARDL$(p, q_1, q_2, ..., q_k)$ without affecting the asymptotic results derived in this section. Hence, our approach is quite general in the sense that one can use a flexible choice for the dynamic lag structure in (2.15)-(2.19) as well as allowing for short-run feedbacks from the lagged dependent variables, Δy_{t-i}, $i = 1, ..., p$, to $\Delta \mathbf{x}_t$ in (2.10). Moreover, within the single equation context, the above analysis is more general than the cointegration analysis of partial systems carried out by Boswijk (1992, 1995), Harbo *et al.* (1998), Johansen (1992, 1995), Pesaran, Shin and Smith (1998) and Urbain (1992), where it is assumed in addition that $\Pi_{xx} = \mathbf{0}$ or $\{\mathbf{x}_t\} \sim I(1)$ in (2.10).

[7] The critical values of the Wald and F- statistics in the general case (not reported here) may be computed via stochastic simulations with different combinations of values for k and $0 \leq r \leq k$.

[8] The critical values for the Wald version of the bounds test are given by $k+1$ times the critical values of the F-test in Cases I, III and V, and $k+2$ times in cases II and IV .

To conclude this section, we re-consider the approach of Banerjee *et al.* (1998). There are three scenarios for the deterministics given by (2.15), (2.17) and (2.19) respectively. Note that the restrictions on the deterministics' coefficients (2.12) are ignored in Cases II and IV and, thus, Cases II and IV are now subsumed by Cases III and V respectively. To summarise, the three cases considered are

- **Case I:** (No Intercepts; No Trends.) $c_0 = 0$ and $c_1 = 0$. The ECM estimated is

$$\Delta y_t = \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x} \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t. \tag{3.11}$$

- **Case III:** (Unrestricted Intercepts; No Trends.) $c_0 \neq 0$ and $c_1 = 0$. The ECM estimated is

$$\Delta y_t = c_0 + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x} \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t. \tag{3.12}$$

- **Case V:** (Unrestricted Intercepts; Unrestricted Trends.) $\mathbf{c}_0 \neq \mathbf{0}$ and $\mathbf{c}_1 \neq \mathbf{0}$. The ECM estimated is

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x} \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t, \tag{3.13}$$

As noted below (2.14), the implicit hypothesis $\boldsymbol{\alpha}_{yx} - \boldsymbol{\omega}' \boldsymbol{\alpha}_{xx} = \mathbf{0}'$ is also imposed but not tested; that is, the limiting distributional results given below are also obtained under the joint hypothesis $H_0 : \pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ of (3.1). Banerjee *et al.* (1998) test $\alpha_{yy} = 0$ (or $H_0^{\pi_{yy}} : \pi_{yy} = 0$) *via* the exclusion of y_{t-1} in Cases I, III and V; that is, (3.11), (3.12) and (3.13). For example, in Case V, they consider the t-statistic

$$t_{\pi_{yy}} = \frac{\hat{\mathbf{y}}_{-1}' \bar{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_-, \hat{\mathbf{X}}_{-1}} \widehat{\Delta \mathbf{y}}}{\hat{\omega}_{uu}^{1/2} \left(\hat{\mathbf{y}}_{-1}' \bar{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_-; \hat{\mathbf{X}}_{-1}} \hat{\mathbf{y}}_{-1} \right)^{1/2}},$$

where $\hat{\omega}_{uu}$ is defined in (3.6), $\widehat{\Delta \mathbf{y}} \equiv \bar{\mathbf{P}}_{\boldsymbol{\iota}_T, \boldsymbol{\tau}_T} \Delta \mathbf{y}$, $\hat{\mathbf{y}}_{-1} \equiv \bar{\mathbf{P}}_{\boldsymbol{\iota}_T, \boldsymbol{\tau}_T} \mathbf{y}_{-1}$, $\mathbf{y}_{-1} \equiv (y_0, ..., y_{T-1})'$, $\hat{\mathbf{X}}_{-1} \equiv \bar{\mathbf{P}}_{\boldsymbol{\iota}_T, \boldsymbol{\tau}_T} \mathbf{X}_{-1}$, $\mathbf{X}_{-1} \equiv (\mathbf{x}_0, ..., \mathbf{x}_{T-1})'$, $\widehat{\Delta \mathbf{Z}}_- \equiv \bar{\mathbf{P}}_{\boldsymbol{\iota}_T, \boldsymbol{\tau}_T} \Delta \mathbf{Z}_-$, $\bar{\mathbf{P}}_{\boldsymbol{\iota}_T, \boldsymbol{\tau}_T} \equiv \bar{\mathbf{P}}_{\boldsymbol{\iota}_T} - \bar{\mathbf{P}}_{\boldsymbol{\iota}_T} \boldsymbol{\tau}_T (\boldsymbol{\tau}_T' \bar{\mathbf{P}}_{\boldsymbol{\iota}_T} \boldsymbol{\tau}_T)^{-1} \boldsymbol{\tau}_T' \bar{\mathbf{P}}_{\boldsymbol{\iota}_T}$, $\bar{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_-, \hat{\mathbf{X}}_{-1}} = \bar{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_-} - \bar{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_-} \hat{\mathbf{X}}_{-1} (\hat{\mathbf{X}}_{-1}' \bar{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_-} \hat{\mathbf{X}}_{-1})^{-1} \hat{\mathbf{X}}_{-1}' \bar{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_-}$ and $\bar{\mathbf{P}}_{\widehat{\Delta \mathbf{Z}}_-} \equiv \mathbf{I}_T - \widehat{\Delta \mathbf{Z}}_- (\widehat{\Delta \mathbf{Z}}_-' \widehat{\Delta \mathbf{Z}}_-)^{-1} \widehat{\Delta \mathbf{Z}}_-'$.

Theorem 3.2 *(Limiting Distribution of the t-statistic for $H_0^{\pi_{yy}} : \pi_{yy} = 0$.) If Assumptions 1-4 and 5a hold and $\boldsymbol{\gamma}_{xy} = \mathbf{0}$, where $\Gamma_x = (\boldsymbol{\gamma}_{xy}, \Gamma_{xx})$, then under $H_0 : \pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ of (3.1), as $T \to \infty$, the asymptotic distribution of the t-statistic for $H_0^{\pi_{yy}} : \pi_{yy} = 0$ has the representation*

$$\int_0^1 d\mathrm{W}_u(a) F_{k-r}(a) \left(\int_0^1 F_{k-r}(a)^2 da \right)^{-1/2}, \tag{3.14a}$$

where

$$F_{k-r}(a) = \left\{ \begin{array}{ll} \mathrm{W}_u(a) - \int_0^1 \mathrm{W}_u(a)\mathbf{W}_{k-r}(a)'da \left(\int_0^1 \mathbf{W}_{k-r}(a)\mathbf{W}_{k-r}(a)'da\right)^{-1} \mathbf{W}_{k-r}(a) & \textbf{Case I} \\ \tilde{\mathrm{W}}_u(a) - \int_0^1 \tilde{\mathrm{W}}_u(a)\bar{\mathbf{W}}_{k-r}(a)'da \left(\int_0^1 \bar{\mathbf{W}}_{k-r}(a)\bar{\mathbf{W}}_{k-r}(a)'da\right)^{-1} \bar{\mathbf{W}}_{k-r}(a) & \textbf{Case III} \\ \hat{\mathrm{W}}_u(a) - \int_0^1 \hat{\mathrm{W}}_u(a)\hat{\mathbf{W}}_{k-r}(a)'da \left(\int_0^1 \hat{\mathbf{W}}_{k-r}(a)\hat{\mathbf{W}}_{k-r}(a)'da\right)^{-1} \hat{\mathbf{W}}_{k-r}(a) & \textbf{Case V} \end{array} \right\},$$

$r = 0, ..., k$, *and Cases I, III and V are defined in (3.11), (3.12) and (3.13) respectively,* $a \in [0, 1]$.

The form of the asymptotic representation (3.14a) is similar to that of a Dickey-Fuller test for a unit root except that the standard Brownian motion $W_u(a)$ is replaced by the residual from an asymptotic regression of $W_u(a)$ on the independent $(k-r)$-vector standard Brownian motion $\mathbf{W}_{k-r}(a)$ (or their de-meaned and de-meaned and de-trended counterparts). As is emphasised in the Proof of Theorem 3.2 given in Appendix A, if the asymptotic analysis is conducted under $H_0^{\pi_{yy}} : \pi_{yy} = 0$ only, the resultant limit distribution depends on the nuisance parameter $\boldsymbol{\omega} - \boldsymbol{\phi}$, where, under Assumption 5a, $\boldsymbol{\alpha}_{yx} - \boldsymbol{\phi}'\boldsymbol{\alpha}_{xx} = \mathbf{0}'$. Moreover, if Δy_t is allowed to Granger-cause $\Delta \mathbf{x}_t$, that is, $\boldsymbol{\gamma}_{xy,i} \neq \mathbf{0}$ for some $i = 1, ..., p-1$, then the limit distribution also is dependent on the nuisance parameter $\boldsymbol{\gamma}_{xy}/(\gamma_{yy} - \boldsymbol{\phi}'\boldsymbol{\gamma}_{xy})$; see Appendix A.

Similarly, to the analysis following Theorem 3.1, we detail the limiting distribution of the t-statistic for $\pi_{yy} = 0$ in the two polar cases in which the forcing variables $\{\mathbf{x}_t\}$ are integrated of order zero and one respectively.

Corollary 3.3 *(Limiting Distribution of the t-statistic for* $H_0^{\pi_{yy}} : \pi_{yy} = 0$ *if* $\{\mathbf{x}_t\} \sim I(0)$.*) If Assumptions 1-4 and 5a hold and* $r = k$, *that is,* $\{\mathbf{x}_t\} \sim I(0)$, *then under* $H_0 : \pi_{yy} = 0$ *and* $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ *of (3.1), as* $T \to \infty$, *the asymptotic distribution of the t-statistic for* $H_0^{\pi_{yy}} : \pi_{yy} = 0$ *has the representation*

$$\int_0^1 d\mathrm{W}_u(a)F(a) \left(\int_0^1 F(a)^2 da\right)^{-1/2},$$

where

$$F(a) = \left\{ \begin{array}{ll} \mathrm{W}_u(a) & \textbf{Case I} \\ \tilde{\mathrm{W}}_u(a) & \textbf{Case III} \\ \hat{\mathrm{W}}_u(a) & \textbf{Case V} \end{array} \right\},$$

and Cases I, III and V are defined in (3.11), (3.12) and (3.13) respectively, $a \in [0, 1]$.

Corollary 3.4 *(Limiting Distribution of the t-statistic for* $H_0^{\pi_{yy}} : \pi_{yy} = 0$ *if* $\{\mathbf{x}_t\} \sim I(1)$.*) If Assumptions 1-4 and 5a hold and* $r = 0$, *that is,* $\{\mathbf{x}_t\} \sim I(1)$, *then under* $H_0 : \pi_{yy} = 0$ *and* $\boldsymbol{\pi}_{yx.x} = \mathbf{0}'$ *of (3.1), as* $T \to \infty$, *the asymptotic distribution of the t-statistic for* $H_0^{\pi_{yy}} : \pi_{yy} = 0$ *has the representation*

$$\int_0^1 d\mathrm{W}_u(a)F_k(a) \left(\int_0^1 F_k(a)^2 da\right)^{-1/2},$$

where

$$F_k(a) = \left\{ \begin{array}{ll} \mathrm{W}_u(a) - \int_0^1 \mathrm{W}_u(a)\mathbf{W}_k(a)'da \left(\int_0^1 \mathbf{W}_k(a)\mathbf{W}_k(a)'da\right)^{-1} \mathbf{W}_k(a) & \textbf{Case I} \\ \tilde{\mathrm{W}}_u(a) - \int_0^1 \tilde{\mathrm{W}}_u(a)\tilde{\mathbf{W}}_k(a)'da \left(\int_0^1 \tilde{\mathbf{W}}_k(a)\tilde{\mathbf{W}}_k(a)'da\right)^{-1} \tilde{\mathbf{W}}_k(a) & \textbf{Case III} \\ \hat{\mathrm{W}}_u(a) - \int_0^1 \hat{\mathrm{W}}_u(a)\hat{\mathbf{W}}_k(a)'da \left(\int_0^1 \hat{\mathbf{W}}_k(a)\hat{\mathbf{W}}_k(a)'da\right)^{-1} \hat{\mathbf{W}}_k(a) & \textbf{Case V} \end{array} \right\},$$

and Cases I, III and V are defined in (3.11), (3.12) and (3.13) respectively, $a \in [0, 1]$.

As above, it may be shown by simulation that the asymptotic critical values obtained from Corollaries 3.3 ($r = k$ and $\{\mathbf{x}_t\} \sim I(0)$) and 3.4 ($r = 0$ and $\{\mathbf{x}_t\} \sim I(1)$) provide lower and upper bounds respectively for those corresponding to the general case considered in Theorem 3.2. Hence, a *bounds procedure* for testing $H_0^{\pi_{yy}} : \pi_{yy} = 0$ based on these two polar cases may be implemented as described above based on the t-statistic for the exclusion of y_{t-1} in the ECMs (3.11), (3.12) and (3.13) without prior knowledge of the cointegrating rank r; see Tables C2.i, C2.iii and C2.v for Cases I, III and V respectively.

4 The Asymptotic Power of the Bounds Procedure

This section firstly demonstrates that the proposed bounds testing procedure described in Section 3 is consistent. Secondly, it derives the asymptotic distribution of the Wald statistic under a sequence of local alternatives.

In the discussion of the consistency of the bounds test procedure, because the rank of the long-run multiplier matrix Π may be either r or $r + 1$ under the alternative hypothesis $H_1 = H_1^{\pi_{yy}} \cup H_1^{\pi_{yx.x}}$ of (3.2) where $H_1^{\pi_{yy}} : \pi_{yy} \neq 0$ and $H_1^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$, it is necessary to deal with these two possibilities. Firstly, under $H_1^{\pi_{yy}} : \pi_{yy} \neq 0$, the rank of Π is $r + 1$ so Assumption 5b applies; in particular, $\alpha_{yy} \neq 0$. Secondly, under $H_0^{\pi_{yy}} : \pi_{yy} = 0$, the rank of Π is r so Assumption 5a applies; in this case, $H_1^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$ holds and, in particular, $\boldsymbol{\alpha}_{yx} - \boldsymbol{\omega}'\boldsymbol{\alpha}_{xx} \neq \mathbf{0}'$.

Theorem 4.1 *(Consistency of the Bounds Test Procedure under $H_1^{\pi_{yy}}$.) If Assumptions 1-4 and 5b hold, then under $H_1^{\pi_{yy}} : \pi_{yy} \neq 0$ of (3.2) the Wald statistic W (3.5) is consistent against $H_1^{\pi_{yy}} : \pi_{yy} \neq 0$ in Cases I-V defined in (2.15)-(2.19).*

Theorem 4.2 *(Consistency of the Bounds Test Procedure under $H_1^{\pi_{yx.x}} \cap H_0^{\pi_{yy}}$.) If Assumptions 1-4 and 5a hold, then under $H_1^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$ of (3.2) and $H_0^{\pi_{yy}} : \pi_{yy} = 0$ of (3.1) the Wald statistic W (3.5) is consistent against $H_1^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$ in Cases I-V defined in (2.15)-(2.19).*

Hence, combining Theorems 4.1 and 4.2, the bounds procedure of section 3 based on the Wald statistic W (3.5) defines a consistent test of $H_0 = H_0^{\pi_{yy}} \cap H_0^{\pi_{yx.x}}$ of (3.1) against $H_1 = H_1^{\pi_{yy}} \cup H_1^{\pi_{yx.x}}$ of (3.2). This result holds irrespective of whether the forcing variables $\{\mathbf{x}_t\}$ are $I(0)$, $I(1)$ or mutually cointegrated.

We now turn to consider the asymptotic distribution of the Wald statistic (3.5) under a suitably specified sequence of local alternatives. Recall that under Assumption 5b

$$\boldsymbol{\pi}_{y.x}[= (\pi_{yy}, \boldsymbol{\pi}_{yx.x})] = (\alpha_{yy}\beta_{yy}, \alpha_{yy}\boldsymbol{\beta}'_{xy} + (\boldsymbol{\alpha}_{yx} - \boldsymbol{\omega}'\boldsymbol{\alpha}_{xx})\boldsymbol{\beta}'_{xx}).$$

Consequently, we define the sequence of local alternatives

$$H_{1T} : \boldsymbol{\pi}_{y.xT}[= (\pi_{yyT}, \boldsymbol{\pi}_{yx.xT})] = (T^{-1}\alpha_{yy}\beta_{yy}, T^{-1}\alpha_{yy}\boldsymbol{\beta}'_{xy} + T^{-1/2}(\boldsymbol{\delta}_{yx} - \boldsymbol{\omega}'\boldsymbol{\delta}_{xx})\boldsymbol{\beta}'_{xx}). \tag{4.1}$$

Hence, under Assumption 3, defining

$$\Pi_T \equiv \begin{pmatrix} \pi_{yyT} & \boldsymbol{\pi}_{yxT} \\ \mathbf{0} & \Pi_{xxT} \end{pmatrix},$$

and recalling $\Pi = \boldsymbol{\alpha}\boldsymbol{\beta}'$, where $(1, -\boldsymbol{\omega}')\boldsymbol{\alpha} = \boldsymbol{\alpha}_{yx} - \boldsymbol{\omega}'\boldsymbol{\alpha}_{xx} = \mathbf{0}'$, we have

$$\Pi_T - \Pi = T^{-1}\boldsymbol{\alpha}_y\boldsymbol{\beta}'_y + T^{-1/2}\begin{pmatrix} \boldsymbol{\delta}_{yx} \\ \boldsymbol{\delta}_{xx} \end{pmatrix}\boldsymbol{\beta}'. \tag{4.2}$$

In order to detail the limit distribution of the Wald statistic under the sequence of local alternatives H_{1T} of (4.1), it is necessary to define the $(k-r+1)$-dimensional Ornstein-Uhlenbeck process $\mathbf{J}^*_{k-r+1}(a) = (J^*_u(a), \mathbf{J}^*_{k-r}(a)')'$ which obeys the stochastic integral and differential equations

$$\mathbf{J}^*_{k-r+1}(a) = \mathbf{W}_{k-r+1}(a) + \mathbf{a}\mathbf{b}'\int_0^a \mathbf{J}^*_{k-r+1}(r)dr,$$

and

$$d\mathbf{J}^*_{k-r+1}(a) = d\mathbf{W}_{k-r+1}(a) + \mathbf{a}\mathbf{b}'\mathbf{J}^*_{k-r+1}(a)da,$$

where $\mathbf{W}_{k-r+1}(a)$ is a $(k-r+1)$-dimensional standard Brownian motion and

$$\mathbf{a} = [(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\Omega(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})]^{-1/2}(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\boldsymbol{\alpha}_y,$$

$$\mathbf{b} = [(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\Omega(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})]^{1/2}[(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})'\Gamma(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})]^{-1}(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})'\boldsymbol{\beta}_y,$$

together with the de-meaned and de-meaned and de-trended counterparts $\bar{\mathbf{J}}^*_{k-r+1}(a) = (\tilde{J}^*_u(a), \bar{\mathbf{J}}^*_{k-r}(a)')'$ and $\hat{\mathbf{J}}^*_{k-r+1}(a) = (\hat{J}^*_u(a), \hat{\mathbf{J}}^*_{k-r}(a)')'$ partitioned similarly, $a \in [0,1]$. See, for example, Johansen (1995, Chapter 14, pp.201-210).

Theorem 4.3 *(Limiting Distribution of W under H_{1T}.) If Assumptions 1-4 and 5a hold, then under $H_{1T} : \pi_{y.x} = T^{-1}\alpha_{yy}\boldsymbol{\beta}'_y + T^{-1/2}(\boldsymbol{\delta}_{yx} - \boldsymbol{\omega}'\boldsymbol{\delta}_{xx})\boldsymbol{\beta}'$ of (4.1), as $T \to \infty$, the asymptotic distribution of the Wald statistic W of (3.5) has the representation*

$$W \Rightarrow \mathbf{z}'_r\mathbf{z}_r + \int_0^1 d\mathrm{J}^*_u(a)\mathbf{F}_{k-r+1}(a)'\left(\int_0^1 \mathbf{F}_{k-r+1}(a)\mathbf{F}_{k-r+1}(a)'da\right)^{-1}\int_0^1 \mathbf{F}_{k-r+1}(a)d\mathrm{J}^*_u(a), \tag{4.3}$$

where $\mathbf{z}_r \sim N(\mathbf{Q}^{1/2}\boldsymbol{\eta}, \mathbf{I}_r)$, $\mathbf{Q}[= \mathbf{Q}^{1/2\prime}\mathbf{Q}^{1/2}] = p\lim_{T\to\infty}\left(T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\right)$, $\boldsymbol{\eta} \equiv (\boldsymbol{\delta}_{yx} - \boldsymbol{\omega}'\boldsymbol{\delta}_{xx})'$, *is distributed independently of the second term in (4.3) and*

$$\mathbf{F}_{k-r+1}(a) = \left\{\begin{array}{cc} \mathbf{J}_{k-r+1}^*(a) & \textbf{Case I} \\ (\mathbf{J}_{k-r+1}^*(a)', 1)' & \textbf{Case II} \\ \bar{\mathbf{J}}_{k-r+1}^*(a) & \textbf{Case III} \\ (\bar{\mathbf{J}}_{k-r+1}^*(a)', a - \frac{1}{2})' & \textbf{Case IV} \\ \hat{\mathbf{J}}_{k-r+1}^*(a) & \textbf{Case V} \end{array}\right\},$$

$r = 0, ..., k$, *and Cases I-V are defined in (2.15)-(2.19),* $a \in [0, 1]$.

The first component of (4.3) $\mathbf{z}_r'\mathbf{z}_r$ is non-central chi-square distributed with r degrees of freedom and non-centrality parameter $\boldsymbol{\eta}'\mathbf{Q}\boldsymbol{\eta}$ and corresponds to the local alternative $H_{1T}^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.xT} = T^{-1/2}(\boldsymbol{\delta}_{yx} - \boldsymbol{\omega}'\boldsymbol{\delta}_{xx})\boldsymbol{\beta}_{xx}'$ under $H_0^{\pi_{yy}} : \pi_{yy} = 0$. The second term in (4.3) is a non-standard Dickey-Fuller unit-root distribution under the local alternative $H_{1T}^{\pi_{yy}} : \pi_{yyT} = T^{-1}\alpha_{yy}\beta_{yy}$ and $\boldsymbol{\delta}_{yx} - \boldsymbol{\omega}'\boldsymbol{\delta}_{xx} = \mathbf{0}'$. Note that under H_0 of (3.1), that is, $\alpha_{yy} = 0$ and $\boldsymbol{\delta}_{yx} - \boldsymbol{\omega}'\boldsymbol{\delta}_{xx} = \mathbf{0}'$, the limiting representation (4.3) reduces to (3.9) as should be expected.

5 An Application: UK Earnings Equation

In this section we provide a re-examination of the earnings equation included in the UK Treasury macroeconometric model described in Chan, Savage and Whittaker (1995, CSW). The theoretical basis of Treasury's earnings equation is the bargaining model advanced in Nickell and Andrews (1983) and reviewed, for example, in Layard, Nickell and Jackman (1991, Chapter 2). The theoretical derivation of the earnings equation is based on a Nash bargaining framework where the firms and the unions set wages to maximize a weighted average of the firm's profits and the union's utility. Following Darby and Wren-Lewis (1993), the theoretical real wage equation underlying Treasury's earnings equation is given by

$$w_t = \frac{Prod_t}{1 + \frac{f(UR_t)(1-RR_t)}{Union_t}}, \tag{5.4}$$

where w_t is the real wage, $Prod_t$ is labour productivity, RR_t is the replacement ratio defined as the ratio of unemployment benefit to the wage rate, $Union_t$ is a measure of "union power", and $f(UR_t)$ is the probability of a union member becoming unemployed, which is assumed to be an increasing function of the unemployment rate, UR_t. The econometric specification is based on a log-linearized version of (5.4) after allowing for the wedge effect that takes account of the difference between the "real product wage" which is the focus of the firm's decision, and the "real consumption wage" which is the focus of the union.[9] The theoretical arguments for a possible long-run wedge effect on real wages is mixed and, as emphasized by CSW, whether such long-run effects are present is an empirical matter. The change in the unemployment rate (ΔUR_t) is also included in the Treasury's wage equation. CSW cite two

[9]The wedge effect is further decomposed into a tax wedge and an import price wedge in the Treasury model, but this decomposition is not pursued here.

different theoretical rationale for the inclusion of ΔUR_t in the wage equation: the differential moderating effects of long-term and short-term unemployed on real wages, and the 'insider-outsider' theories which argue that only rising unemployment will be effective in significantly moderating wage demands. See Blanchard and Summers (1986) and Lindbeck and Snower (1989). The ARDL model and its associated unrestricted error correction formulation that we shall be using automatically allow for such effects.

Following the modelling approach proposed in this paper we start from the maintained assumption that the time series properties of the key variables in Treasury's earnings equation can be well approximated by a log-linear $VAR(p)$ model, augmented with appropriate deterministic components such as intercepts or time trends. To ensure comparability of our results with those of the Treasury, the replacement ratio is not included in the analysis. CSW (p. 50) report that "... it has not proved possible to identify a significant effect from the replacement ratio, and this had to be omitted from our specification." Also, as in CSW, we include two dummy variables to take account of the effects of incomes policies on average earnings. These dummy variables are defined by

$$\begin{aligned} D7475_t &= 1 \text{ during the 8 quarters of 1974-75 and zero elsewhere,} \\ D7579_t &= 1 \text{ during the 20 quarters of 1975-79 and zero elsewhere.} \end{aligned}$$

The asymptotic theory developed in the paper is not affected by the inclusion of such "one-off" dummy variables.[10] Let

$$\mathbf{z}_t = (w_t, Prod_t, UR_t, Wedge_t, Union_t)' = (w_t, \mathbf{x}_t')',$$

then using the analysis of Section 2 the conditional model of interest can be written as

$$\Delta w_t = c_0 + c_1 t + c_2 D7475_t + c_3 D7579_t + \pi_{ww} w_{t-1} + \boldsymbol{\pi}_{wx.x}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + \boldsymbol{\delta}' \Delta \mathbf{x}_t + u_t. \tag{5.5}$$

Under the assumption that lagged real wages, w_{t-1}, do not enter the sub-VAR model for $\mathbf{x}_t$, the above real wage equation is identified and can be estimated consistently by the OLS.[11] Notice, however, that this assumption does not rule out the inclusion of *lagged changes* in real wages in the unemployment or productivity equations, for example. The exclusion of the *level* of real wages from the unemployment and productivity equations is an identification requirement and allows us to identify the bargaining theory of wages from other alternatives, such as the efficiency wage theory which postulates that labour productivity is partly

[10] The asymptotic theory and the associated critical values must, however, be modified in the case of the dummy variables where the fraction of periods in which the dummy variables are non-zero does not tend to zero with the sample size, T.

[11] See Assumption 3 and the discussion that follows it. Notice that by construction the contemporaneous effects, $\Delta \mathbf{x}_t$, are uncorrelated with the disturbance term, u_t, and instrumental variable estimation which has been particularly popular in the empirical literature on the wage equation is not needed. In fact, given the unrestricted nature of the lag distribution of the conditional model, (5.5), it is difficult to find suitable instruments: namely variables that are not already included in the model, which are uncorrelated with u_t *and* at the same time have a reasonable degree of correlation with the variables that are included in (5.5).

determined by the *level* of real wages.[12] It is clear that the bargaining theory as set out in CSW and the efficiency wage theory can not be entertained simultaneously, at least not in the long run. Our framework does, of course, allow for changes in real wages to affect labour productivity or the unemployment rate.

The above specification is also based on the assumption that the disturbances, u_t, are serially uncorrelated. It is therefore important that p, the order of the underlying VAR, is selected appropriately. There is a delicate balance between choosing p to be sufficiently large to mitigate the residual serial correlation problem, and at the same time sufficiently small so that the model is not unduly over-parameterized, particularly in view of the limited time series data which is available.

Finally, a decision must be made concerning the time trend in (5.5) and whether its coefficient should be restricted.[13] This issue can only be settled in light of the particular sample period under consideration. The time series data we shall be using are quarterly, cover the period 1970q1-1997q4, and are seasonally adjusted (when relevant).[14] To ensure comparability of the estimation results for different choices of p, we carried out all the estimations over the period 1972q1-1997q4 (a total of 104 quarters), and reserved the first 8 observations for the construction of lagged variables.

The five variables in the earnings equation were constructed from primary sources in the following manner:

$$\begin{aligned}
w_t &= \ln\left(\frac{ERPR_t}{PYNONG_t}\right), \\
Prod_t &= \ln\left(\frac{YPROM_t + 278.29 * YMF_t}{EMF_t + ENMF_t}\right), \\
UR_t &= \ln\left(\frac{100 * ILOU_t}{ILOU_t + WFEMP_t}\right), \\
Wedge_t &= \ln(1 + TE_t) + \ln(1 - TD_t) - \ln\left(\frac{RPIX_t}{PYNONG_t}\right), \\
Union_t &= \ln(UDEN),
\end{aligned}$$

where $ERPR_t$ is average private sector earnings per employee (£), $PYNONG_t$ is the non-oil non-government GDP deflator, $YPROM_t$ is output in the private, non-oil, non-manufacturing, and public traded sectors at constant factor cost (£million,1990), YMF_t is manufacturing output index adjusted for stock changes (1990=100), EMF_t and $ENMF_t$ are respectively employment in UK manufacturing and non-manufacturing sectors (thousands), $ILOU_t$ is the International Labour Office (ILO) measure of unemployment (thousands), $WFEMP_t$ is total employment (thousand), TE_t is the average employers National Insurance contribution rate, TD_t is the average direct tax rate on employment incomes, $RPIX_t$ is the Retail Price Index excluding mortgage payments, and $UDEN_t$ is union density (used to proxy union power) and measured as union membership as a percentage of employment.

[12] For a discussion of the issues that surround the identification of the wage equation see Manning (1993).

[13] See, for example, Pesaran, Shin and Smith (1998), and the discussion at the end of Section 2.

[14] We are grateful to Andrew Gurney and Rod Whittaker for providing us with the data. For further details about the sources and the descriptions of the variables see Chan *et al.* (1995, pages 46-51, and page 11 of its Annex).

5.1 Empirical Results

The time series plots of real wages (average earnings) and the productivity variable clearly show steadily rising trends with real wages growing at a slightly faster rate than productivity. This suggests, at least initially, that the linear trend need to be included in the real wage equation (5.5). Also the application of the unit root tests to the five variables, perhaps not surprisingly, yields mixed results; with strong evidence in favour of the unit root hypothesis only in the case of the real wage and the productivity variables. This does not, of course, necessarily mean that the other three variables (UR, $Wedge$, and $Union$) are not likely to have any long-run impacts on real wages. Following the methodology developed in this paper it is possible to test the existence of a long-run real wage equation involving all the five variables irrespective of whether they are $I(0)$, $I(1)$, or mutually cointegrated.[15]

To determine the appropriate lag length, p, and whether a deterministic linear trend is also required in addition to the productivity variable, we estimated the conditional model (5.5) by the OLS with and without a linear time trend, for $p = 1, 2, ..., 7$. As pointed out earlier all the regressions were computed over the same period, 1972q1-1997q4. We found that lagged changes of the productivity variable, $\Delta Prod_{t-1}$, $\Delta Prod_{t-2}$,..., were not significant (either singly or jointly) in any of the regressions. Therefore, for the sake of parsimony and to avoid unnecessary over-parameterization we decided to re-estimate the regressions without these lagged variables, but including the lagged changes of all the other variables. Table 1 gives the Akaike's Information and Schwarz's Bayesian Criteria, denoted respectively by AIC and SBC, and the Lagrange multiplier (LM) statistics for testing the hypothesis of residual serial correlation of order 1 and 4. These are denoted by $\chi^2_{SC}(1)$ and $\chi^2_{SC}(4)$, respectively.

As to be expected the lag order selected by the AIC, namely $\widehat{p}_{aic} = 6$ irrespective of whether a deterministic trend term included in the model or not, is much larger than the lag order selected by the SBC. This latter criterion estimates p to be only 1 when the model contains a trend and 4 when it does not. The χ^2_{SC} statistics also suggest using a relatively high lag order: 4 or more. In view of the importance of the assumption of serially uncorrelated errors for the validity of the bounds test, it seems prudent to select p to be either 5 or 6.[16] A higher lag order does not seem necessary. Nevertheless, in what follows for completeness we report the test results for $p = 4, 5$ and 6. The results in Table 1 also show that there is little to choose between the conditional model with and without a linear deterministic trend.

Table 2 gives the values of the F- and t-statistics for testing the existence of a long-run earnings equation under 3 different cases depending on whether the model contains a linear trend and whether the trend coefficients are restricted. See Sections 3 for a detailed discussion of these cases.

[15] The view that long-run relationships could exist only among variables that are integrated of order 1 or higher is implicit in much of the empirical literature on cointegration.

[16] In the Treasury model different lag orders are chosen for different variables. The highest lag order selected is 4; applied to the log of the price deflator and the wedge variable. The estimation period of the earnings equation in the Treasury model is 1971q1-1994q3.

Table 1*

Statistics for Selecting the Lag Order of the Earnings Equation

	With Deterministic Trends					Without Deterministic Trends			
p	AIC	SBC	$\chi^2_{SC}(1)$	$\chi^2_{SC}(4)$		AIC	SBC	$\chi^2_{SC}(1)$	$\chi^2_{SC}(4)$
1	319.33	302.14	16.86†	35.89†		317.51	301.64	18.38†	34.88†
2	324.25	301.77	2.16	19.71†		323.77	302.62	1.98	21.52†
3	321.51	293.74	0.52	17.07†		320.87	294.43	1.56	19.35†
4	334.37	301.31	3.48**	7.79**		335.37	303.63	3.41**	7.13
5	335.84	297.50	0.03	2.50		336.49	299.47	0.03	2.15
6	337.06	293.42	0.85	3.58		337.03	294.72	0.99	3.99
7	336.96	288.04	0.17	2.20		336.85	289.25	0.09	0.64

* Notes: p is the lag order of the conditional model (5.5), with zero restrictions on the coefficients of lagged changes in the productivity variable. $AIC_p = LL_p - s_p$ and $SBC_p = LL_p - \frac{s_p}{2}\ln(T)$, are the Akaike and Schwarz Information Criteria, where LL_p is the maximized log-likelihood value of the model, p is the lag order, s_p is the number of freely estimated coefficients, and T is the sample size. $\chi^2_{SC}(1)$ and $\chi^2_{SC}(4)$ are the LM statistics for testing residual serial correlations of orders 1 and 4. The symbols †, $*$, and $**$ represent significance at 1% or less, 5% or less, and 10% or less, respectively.

Table 2*

F- and t- Statistics for Testing the Existence of a Long-Run Earnings Equation

p	With Deterministic Trends			Without Deterministic Trends	
	F_{IV}	F_V	t_V	F_{III}	t_{III}
4	2.99^a	2.34^a	-2.26^a	3.63^b	-3.02^b
5	4.42^b	3.96^b	-2.83^a	5.23^c	-4.00^c
6	4.78^c	3.59^b	-2.44^a	5.42^c	-3.48^b

* Notes: p is the lag order of the underlying model. See also the notes to Table 1. F_{IV} is the F-statistic for testing zero restrictions on the coefficients of the lagged level variables and the trend term in (5.5). F_V is the F-statistic for testing zero restrictions on the coefficients of the lagged level variables in (5.5). F_{III} is the F-statistic for testing zero restrictions on the coefficients of the lagged level variables in (5.5) without the trend term. t_V and t_{III} are the t-ratios of the coefficient of w_{t-1} in (5.5) with and without a deterministic linear trend. a denotes that the statistic lies below the 95% lower bound, b denotes it falls within the 95% bounds, and c denotes that it falls outside the 95% upper bound.

The various statistics in Table 2 need to be compared with the critical value bounds provided in Tables C1 and C2. First consider the bounds F test. For the model with a deterministic trend, F_V is the standard F-statistic for testing the restrictions $\pi_{ww} = 0$ and $\boldsymbol{\pi}_{wx.x} = \mathbf{0}$, while F_{IV} is the standard F-statistic for testing $\pi_{ww} = 0$, $\boldsymbol{\pi}_{wx.x} = \mathbf{0}$, and $c_1 = 0$, in (5.5). As has been argued in Pesaran, Shin and Smith (1998), the statistic F_{IV} which sets the trend coefficient to zero under the null of no level long-run relationship is more appropriate than F_V which does not impose such a restriction. Notice that when the trend coefficients are not restricted, (5.5) implies a quadratic trend in real wages under the null hypothesis of $\pi_{ww} = 0$ and $\boldsymbol{\pi}_{wx.x} = \mathbf{0}$, which is not plausible. The critical value bounds for the statistics F_{IV} and F_V are given in Tables C1.iv and C1.v, respectively. Since the model contains 4 regressors, the 95% critical value bounds are (3.66, 4.76) and (3.47, 4.57) for F_{IV} and F_V , respectively. The test outcome critically depends on the choice of the lag order, p. For $p = 4$ the hypothesis that there exists no long-run earnings equation is not rejected at the 95% level, irrespective of whether the regressors are $I(0)$ or $I(1)$. For $p = 5$ the bounds test is inconclusive. For $p = 6$ (selected by the AIC) the statistic F_V is still inconclusive, but $F_{IV} = 4.78$ just lies outside the 95% critical value bounds and rejects the null hypothesis that there exists no long-run earnings equation, irrespective of whether the regressors are $I(0)$ or $I(1)$.[17] This conclusion is confirmed even more conclusively when the bounds F-test is applied to the earnings equations without a linear trend. The relevant test statistic is F_{III}, and its associated 95% critical value bounds are (2.86, 4.01).[18] For $p = 4$, $F_{III} = 3.63$, and the test result is inconclusive. But for $p = 5$ and 6 the values of F_{III} are 5.23 and 5.42 and the hypothesis of no long-run earnings equation is conclusively rejected.

The results from the application of the bounds t-tests to the earnings equations are less clear cut and do not allow the imposition of the trend restrictions discussed above. The two t-statistics reported in Table 2, t_V and t_{III}, are the t-ratios of the OLS estimate of π_{ww} in (5.5), with and without a linear time trend, respectively.[19] The 95% critical value bounds for t_{III} and t_V tests are (-2.86, -3.99) and (-3.41, -4.36).[20] Therefore, when a linear trend is included in the model the bounds t-test does not reject the null hypothesis even for $p = 5$ or 6. But when the trend term is excluded the null hypothesis is just rejected for $p = 5$.

Overall, the test results support the existence of a long-run earnings equation when a sufficiently high lag order is selected and when the statistically insignificant deterministic trend term is excluded from the conditional model. Such a specification is in accord with the evidence on the performance of the alternative conditional models set out in Table 1, and in the remainder of this section we focus our attention on this specification and provide estimates of the long-run coefficients and the short-run dynamics based on a conditional earnings equation with $p = 6$, but without a deterministic time trend. For this model, using the ARDL approach to the estimation of the long-run relations discussed in Pesaran and Shin (1999), we obtain the following level long-run earnings equation[21]

[17] The same conclusion is also reached for $p = 7$.

[18] See Table C1.iii.

[19] Notice also that lagged changes in the productivity variable are excluded.

[20] See Tables C2.iii and C2.v, for $k = 4$ under the columns headed 95%.

[21] Notice that the ARDL approach advanced in Pesaran and Shin (1999) is applicable irrespective of whether the regressors are $I(0)$ or $I(1)$.

$$w_t = \underset{(0.0526)}{1.063\, Prod_t} \quad \underset{(0.0398)}{-0.114\, UR_t} \quad \underset{(0.0297)}{-0.998\, Wedge_t} \quad \underset{(0.377)}{+1.495\, Union_t} \quad \underset{(0.274)}{+2.714} \quad +\hat{v}_t, \tag{5.6}$$

where $\hat{v}_t$ is the error-correction term. The standard errors of the long-run estimates are given in brackets.[22] All the long-run estimates are highly significant and have the expected signs. The coefficients of the productivity and the wedge variables do not differ significantly from unity. In Treasury's earnings equation the long-run coefficient of the productivity variable is imposed to be unity, and the above estimates can be viewed as providing empirical support for such an *a priori* restriction. Our long-run estimates of the effects of the unemployment rate and the union variable on real wages (namely -0.114 and 1.495) are also in line with Treasury estimates of -.09 and 1.31.[23] The main difference between the two sets of estimates concerns the long-run coefficient of the wedge variable. We obtain a much larger estimate, almost twice as much as the estimate obtained by the Treasury.

The error correction regression associated with the above (level) long-run relationship is given in Table 3.[24] These estimates provide further direct evidence on the complicated dynamics that seem to exist between real wage movements and its main determinants.[25] All the five lagged changes in real wages are statistically significant, further justifying the choice of $p = 6$. The error correction coefficient is estimated to be -0.235 (0.0675),[26] which is reasonably large and highly significant. The auxiliary equation of the autoregressive part of the model has the real roots 0.9258, and -0.8931, and two pairs of complex roots with moduli, 0.7853, and 0.5951; thus suggesting an initially cyclical real wage process which slowly converges towards its equilibrium given by (5.6).[27] Despite the many insignificant coefficients that are retained in this error correction specification, the regression fits reasonably well and satisfies the diagnostic tests for non-normal errors and heteroskedasticity. However, it fails the functional form misspecification at the 5% level; which may be suggestive of some non-linear effects or asymmetries in the adjustment of real wage process that our linear specification is incapable of taking into account.[28] Recursive estimation of the

[22] The long-run estimates and their standard errors are computed using *Microfit 4.0.* See Pesaran and Pesaran (1997).

[23] CSW do not report standard errors for the long-run estimates of the Treasury earnings equation.

[24] In practice it may be desirable also to derive a more parsimonious error correction model by imposing a unit long-run coefficient on the productivity variable and by dropping lagged changes with (jointly) statistically insignificant coefficients. But for our purposes this does not seem to be necessary.

[25] The standard errors of the estimates reported in Table 3 allow for the uncertainty associated with the estimation of the long-run coefficients. This is important in the present application where it is not known with certainty whether the regressors are $I(0)$ or $I(1)$. Only in the case where it is not known for sure that all the regressors are $I(1)$ and cointegrated would it be reasonable in large samples to treat the estimates of the long-run level coefficients as known; on the grounds of their super-consistency.

[26] The error correction coefficient in the Treasury's earnings equation is estimated to be -0.1848 (0.0528), which is quite a bit smaller than our estimate. (See p. 11 in Annex of CSW.) This seems to be due to the shorter lag lengths used in the estimation of the Treasury's equation rather than the fact that it has been estimated over a shorter time period: 1971q1-1994q3. Notice also that the t-ratio reported for this coefficient does not have the standard t-distribution.

[27] The complex roots are $0.3406 \pm 0.7076i$, and $-0.2016 \pm 0.5859i$, where $i = \sqrt{-1}$.

[28] The error correction regression in Table 3 also passes the residual serial correlation test. However, the

error correction model also suggests that the regression coefficients are generally stable over the sample period. The cumulative sum and cumulative sum of squares plots based on the recursive residuals are given in Figures 1 and 2 and do not show evidence of statistically significant breaks. However, these tests are known to have low powers and are likely to have missed some important breaks. Overall, the error correction earnings equation presented in Table 3 has a number of desirable features and provides a sound basis for further research.

model was chosen specifically to meet this test, and should not therefore be given any extra credits for passing the serial correlation test!

Table 3*

Error Correction Form of the Earnings Equation

(Dependent variable: Δw_t, Estimation Period: 1972q1-1997q4)

Regressor	Coefficient	Standard Error	p-value
$\hat{v}_{t-1}$	-.2347	.0675	N/A
Δw_{t-1}	-.4546	.1024	.000
Δw_{t-2}	-.3629	.1157	.002
Δw_{t-3}	-.5560	.1122	.000
Δw_{t-4}	-.1930	.1063	.074
Δw_{t-5}	-.1957	.0925	.038
$\Delta Prod_t$	.3037	.1024	.004
ΔUR_t	.0011	.0090	.899
ΔUR_{t-1}	.0164	.0140	.244
ΔUR_{t-2}	.0054	.0131	.684
ΔUR_{t-3}	.0286	.0123	.024
ΔUR_{t-4}	.0266	.0131	.046
ΔUR_{t-5}	.0016	.0130	.903
$\Delta Wedge_t$	-.2904	.0583	.000
$\Delta Wedge_{t-1}$	-.0594	.0661	.372
$\Delta Wedge_{t-2}$	-.1143	.0687	.101
$\Delta Wedge_{t-3}$	-.2045	.0698	.005
$\Delta Wedge_{t-4}$	-.0594	.0653	.366
$\Delta Wedge_{t-5}$	.0141	.0610	.818
$\Delta Union_t$	-1.3090	.8899	.146
$\Delta Union_{t-1}$	-2.8188	.9120	.003
$\Delta Union_{t-2}$	-.2278	.9351	.808
$\Delta Union_{t-3}$	-.0263	.8532	.976
$\Delta Union_{t-4}$	-1.7437	.8296	.039
$\Delta Union_{t-5}$	-.5100	.5868	.388
intercept	.6371	.1891	.001
$D7475_t$	.0288	.0090	.002
$D7579_t$	.0195	.0072	.008

$$\bar{R}^2 = .5473, \quad \hat{\sigma} = .0084, \quad AIC = 337.03, \quad SBC = 294.72,$$
$$\chi^2_{SC}(4) = 3.99[.408], \quad \chi^2_{FF}(1) = 5.83[.016], \quad \chi^2_N(2) = 0.82[.663], \quad \chi^2_H(1) = 0.22[.638]$$

* Notes: The error correction term, $\hat{v}_{t-1}$, is defined by (5.6). The regression is based on the conditional model (5.5) with $p = 6$, but excluding lagged changes in the productivity variable. $\bar{R}^2$ is the adjusted squared multiple correlation coefficient, $\hat{\sigma}$ is the standard error of the regression, AIC and SBC are the Akaike and Schwarz Information Criteria, $\chi^2_{SC}(4)$, $\chi^2_{FF}(1)$, $\chi^2_N(2)$, and $\chi^2_H(1)$ are the Chi-squared statistics for tests of residual serial correlation, functional form mis-specification, non-normal errors and hetroskedasticity, respectively. For the details of these diagnostic tests see, for example, Pesaran and Pesaran (1997, Ch. 18).

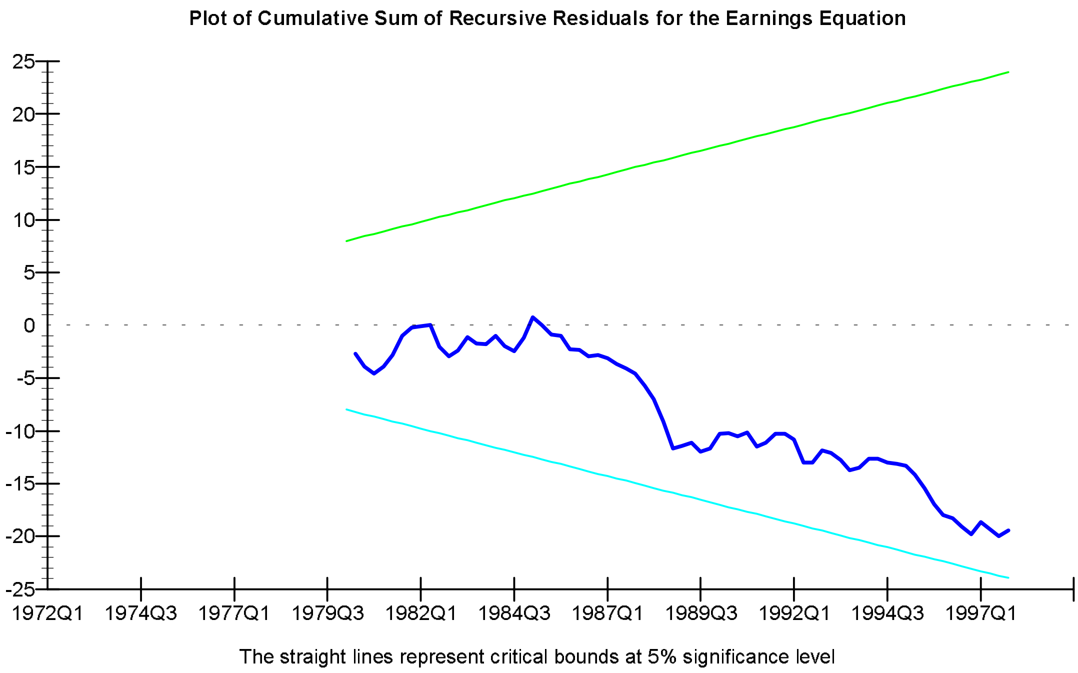

Figure 1

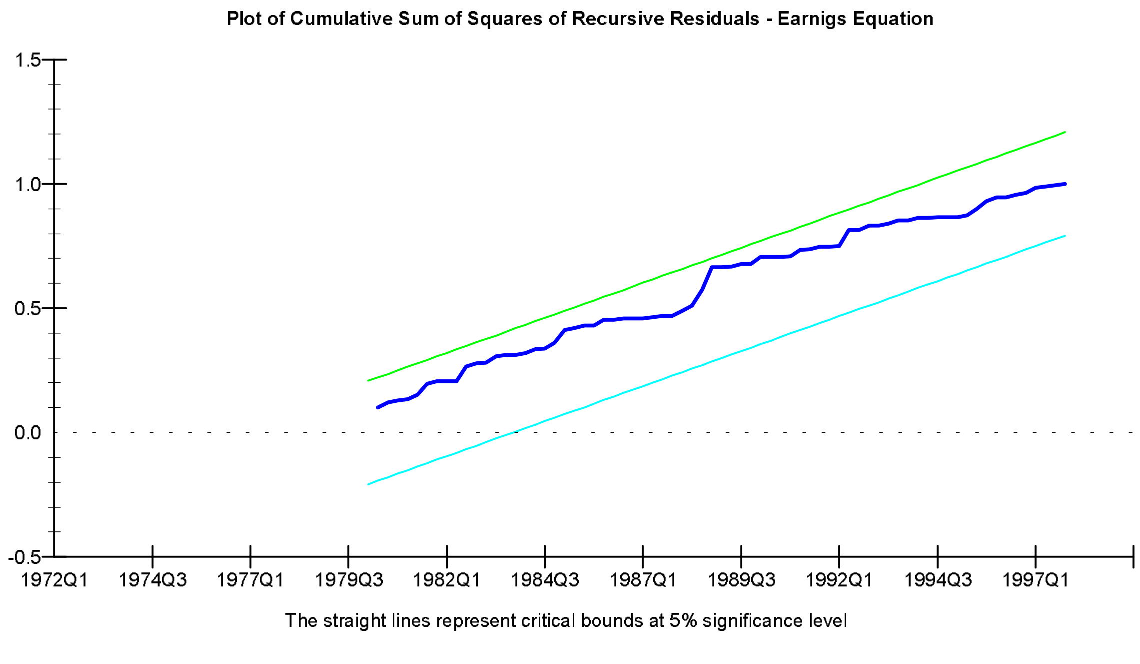

Figure 2

6 Concluding Remarks

Empirical analysis of long run relationships has been an integral part of time series econometrics and pre-dates the recent literature on unit roots and cointegration[29] However, the emphasis of this early literature was on the estimation of long-run relationships and did not address the testing problem. The cointegration literature attempts to fill this vacuum, but under the relatively restrictive assumption that the regressors, $\mathbf{x}_t$, entering the long-run determination of the dependent variable of interest, y_t, are all integrated of order 1 or more. In this paper we show that the problem of testing the existence of a long-run relationship between y_t and $\mathbf{x}_t$ continues to be present and furthermore is non-standard even if *all* the regressors under consideration are $I(0)$. This is because under the null hypothesis that there exists no long-run relationship between y_t and $\mathbf{x}_t$, the y_t process will be $I(1)$, irrespective of whether the regressors are $I(0)$, $I(1)$ or mutually cointegrated. The asymptotic theory developed in this paper provides a framework for testing the existence of a single long-run level relationship between y_t and $\mathbf{x}_t$ when it is not known with certainty whether the regressors are $I(0)$, $I(1)$ or mutually cointegrated.[30] In this framework it is not necessary for the order of integration of the underlying regressors to be ascertained first, before the existence of a long-run relationship between y_t and $\mathbf{x}_t$ can be tested; and therefore unlike the cointegration analysis is not subject to this particular pre-testing problem. The application of the proposed bounds testing procedure to the UK earnings equation highlights this point, where one need not take a position as to whether the rate of unemployment or the union density variable, for example, are $I(1)$ or $I(0)$. It is, however, worth emphasizing that the test developed in this paper is not appropriate in situations where there may be more than one long-run *level* relationship involving y_t. Extending our approach to deal with such cases is relatively straightforward, but involves further theoretical developments and requires computation of new critical values.

[29] For an excellent review of this early literature see Hendry, Pagan and Sargan (1984).

[30] Clearly, the system approach developed by Johansen (1991, 1995) can also be applied to a set of variables containing possibly a mixture of $I(0)$ and $I(1)$ regressors. But in such cases the result of the trace or the maximum eigenvalue tests will be difficult to interpret; as it will not be possible to identify whether the reduced rank outcome (if any) is indicative of the existence of long run relationships or is due to the presence of $I(0)$ regressors in the model.

Appendix A: Proofs for Section 3

We confine the main proof of Theorem 3.1 to that for Case IV and briefly detail the alterations necessary for the other cases. Under Assumptions 1-4 and 5a, the process $\{\mathbf{z}_t\}_{t=1}^{\infty}$ has the infinite moving-average representation

$$\mathbf{z}_t = \boldsymbol{\mu} + \boldsymbol{\gamma} t + \mathbf{C}\mathbf{s}_t + \mathbf{C}^*(L)\boldsymbol{\varepsilon}_t, \tag{A.1}$$

where the partial sum $\mathbf{s}_t \equiv \sum_{i=1}^{t} \boldsymbol{\varepsilon}_i$, $\Phi(z)\mathbf{C}(z) = \mathbf{C}(z)\Phi(z) = (1-z)\mathbf{I}_{k+1}$, $\Phi(z) \equiv \mathbf{I}_{k+1} - \sum_{i=1}^{p} \Phi_i z^i$, $\mathbf{C}(z) \equiv \mathbf{I}_{k+1} + \sum_{i=1}^{\infty} \mathbf{C}_i z^i = \mathbf{C} + (1-z)\mathbf{C}^*(z)$, $t = 1, 2...$; see Johansen (1991) and Pesaran, Shin and Smith (1998). Note that $\mathbf{C} = (\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})[(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\Gamma(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})]^{-1}(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'$; see Johansen (1991, (4.5), p.1559). Define the $(k+2, r)$ and $(k+2, k-r+1)$ matrices $\boldsymbol{\beta}_*$and $\boldsymbol{\delta}$:

$$\boldsymbol{\beta}_* \equiv \begin{pmatrix} -\boldsymbol{\gamma}' \\ \mathbf{I}_{k+1} \end{pmatrix} \boldsymbol{\beta}, \boldsymbol{\delta} \equiv \begin{pmatrix} -\boldsymbol{\gamma}' \\ \mathbf{I}_{k+1} \end{pmatrix} (\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp}),$$

where $(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})$ is a $(k+1, k-r+1)$ matrix whose columns are a basis for the orthogonal complement of $\boldsymbol{\beta}$. Hence, $(\boldsymbol{\beta}, \boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})$ is a basis for $\mathcal{R}^{k+1}$. Let $\boldsymbol{\xi}$ be the $(k+2)$-unit vector $(1, \mathbf{0}')'$. Then, $(\boldsymbol{\beta}_*, \boldsymbol{\xi}, \boldsymbol{\delta})$ is a basis for $\mathcal{R}^{k+2}$. It therefore follows that

$$T^{-1/2}\boldsymbol{\delta}'\mathbf{z}^*_{[Ta]} = T^{-1/2}(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})'\boldsymbol{\mu} + T^{-1/2}(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})'\mathbf{C}\mathbf{s}_{[Ta]} + (\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})'T^{-1/2}\mathbf{C}^*(L)\boldsymbol{\varepsilon}_{[Ta]}$$

$$\Rightarrow (\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})'\mathbf{C}\mathbf{B}_{k+1}(a),$$

where $\mathbf{z}^*_t = (t, \mathbf{z}'_t)'$, $\mathbf{B}_{k+1}(a)$ is a $(k+1)$-vector Brownian motion with variance matrix Ω and $[Ta]$ denotes the integer part of Ta, $a \in [0, 1]$; see Phillips and Solo (1992, Theorem 3.15, p.983). Also

$$T^{-1}\boldsymbol{\xi}'\mathbf{z}^*_t = T^{-1}t \Rightarrow a.$$

Similarly, noting that $\boldsymbol{\beta}'\mathbf{C} = \mathbf{0}$, we have

$$\boldsymbol{\beta}'_*\mathbf{z}^*_t = \boldsymbol{\beta}'\boldsymbol{\mu} + \boldsymbol{\beta}'\mathbf{C}^*(L)\boldsymbol{\varepsilon}_t = O_P(1).$$

Hence, from Phillips and Solo (1992, Theorem 3.16, p.983), defining $\bar{\mathbf{Z}}^*_{-1} \equiv \bar{\mathbf{P}}_{\iota}\mathbf{Z}^*_{-1}$ and $\widetilde{\Delta\mathbf{Z}}_{-} \equiv \bar{\mathbf{P}}_{\iota}\Delta\mathbf{Z}_{-}$, it follows that

$$T^{-1}\boldsymbol{\beta}'_*\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_* = O_P(1), T^{-1}\boldsymbol{\beta}'_*\bar{\mathbf{Z}}^{*\prime}_{-1}\widetilde{\Delta\mathbf{Z}}_{-} = O_P(1), T^{-1}\widetilde{\Delta\mathbf{Z}}'_{-}\widetilde{\Delta\mathbf{Z}}_{-} = O_P(1),$$

$$T^{-1}\mathbf{B}'_T\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_* = O_P(1), T^{-1}\mathbf{B}'_T\bar{\mathbf{Z}}^{*\prime}_{-1}\widetilde{\Delta\mathbf{Z}}_{-} = O_P(1), \tag{A.2}$$

where $\mathbf{B}_T \equiv \left(\boldsymbol{\delta}, T^{-1/2}\boldsymbol{\xi}\right)$. Similarly, defining $\bar{\mathbf{u}} \equiv \bar{\mathbf{P}}_{\iota}\mathbf{u}$,

$$T^{-1/2}\boldsymbol{\beta}'_*\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{u}} = O_P(1), T^{-1/2}\widetilde{\Delta\mathbf{Z}}'_{-}\bar{\mathbf{u}} = O_P(1). \tag{A.3}$$

Cf. Johansen (1991, Lemma A.3, p.1569) and Johansen (1995, Lemma 10.3, p.146).

The next result follows from Phillips and Solo (1992, Theorem 3.15, p.983); cf. Johansen (1991, Lemma A.3, p.1569) and Johansen (1995, Lemma 10.3, p.146) and Phillips and Durlauf (1986).

Lemma A.1 *Let* $\mathbf{B}_T \equiv \left(\boldsymbol{\delta}, T^{-1/2}\boldsymbol{\xi}\right)$ *and define* $\mathbf{G}(a) = (\mathbf{G}_1(a)', \mathbf{G}_2(a))'$, *where* $\mathbf{G}_1(a) \equiv (\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})'\mathbf{C}\bar{\mathbf{B}}_{k+1}(a)$, $\bar{\mathbf{B}}_{k+1}(a)[= (\tilde{B}_1(a)', \bar{\mathbf{B}}_k(a)')'] = \mathbf{B}_{k+1}(a) - \int_0^1 \mathbf{B}_{k+1}(a)da$, *and* $\mathbf{G}_2(a) \equiv a - \frac{1}{2}$, $a \in [0, 1]$. *Then*

$$T^{-2}\mathbf{B}'_T\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{Z}}^*_{-1}\mathbf{B}_T \Rightarrow \int_0^1 \mathbf{G}(a)\mathbf{G}(a)'da,$$

$$T^{-1}\mathbf{B}'_T\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{u}} \Rightarrow \int_0^1 \mathbf{G}(a)d\tilde{\mathrm{B}}^*_u(a),$$

where $\tilde{\mathrm{B}}^*_u(a) \equiv \tilde{\mathrm{B}}_1(a) - \boldsymbol{\omega}'\bar{\mathbf{B}}_k(a)$ *and* $\bar{\mathbf{B}}_k(a) = (\tilde{\mathrm{B}}_1(a), \bar{\mathbf{B}}_k(a)')'$, $a \in [0, 1]$.

Proof of Theorem 3.1: Under H_0 of (3.1), the Wald statistic W of (3.5) can be written as

$$\hat{\omega}_{uu}W = \bar{\mathbf{u}}'\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\left(\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\right)^{-1}\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}$$

$$= \bar{\mathbf{u}}'\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\mathbf{A}_T\left(\mathbf{A}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\mathbf{A}_T\right)^{-1}\mathbf{A}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}.$$

where $\mathbf{A}_T \equiv T^{-1/2}\left(\boldsymbol{\beta}_*, T^{-1/2}\mathbf{B}_T\right)$. Consider the matrix $\mathbf{A}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\mathbf{A}_T$. It follows from (A.2) and Lemma A.1 that

$$\mathbf{A}_T'\bar{\mathbf{Z}}'_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}\mathbf{A}_T = \begin{pmatrix} T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_* & \mathbf{0}' \\ \mathbf{0} & T^{-2}\mathbf{B}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{Z}}^*_{-1}\mathbf{B}_T \end{pmatrix} + o_P(1). \quad \text{(A.4)}$$

Next, consider $\mathbf{A}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}$. From (A.3) and Lemma A.1,

$$\mathbf{A}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}} = \begin{pmatrix} T^{-1/2}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}} \\ T^{-1}\mathbf{B}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{u}} \end{pmatrix} + o_P(1). \quad \text{(A.5)}$$

Finally, the estimator (3.6) for the error variance ω_{uu},

$$\hat{\omega}_{uu} = (T-m)^{-1}\left[\bar{\mathbf{u}}'\bar{\mathbf{u}} - \bar{\mathbf{u}}'\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\mathbf{A}_T(\mathbf{A}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\mathbf{A}_T)^{-1}\mathbf{A}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}\right]$$

$$= (T-m)^{-1}\bar{\mathbf{u}}'\bar{\mathbf{u}} + o_P(1) = \omega_{uu} + o_P(1). \quad \text{(A.6)}$$

From (A.4)-(A.6) and Lemma A.1,

$$W = T^{-1}\bar{\mathbf{u}}'\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\left(T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\right)^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}/\omega_{uu}$$

$$+T^{-2}\bar{\mathbf{u}}'\bar{\mathbf{Z}}^*_{-1}\mathbf{B}_T\left[T^{-2}\mathbf{B}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{Z}}^*_{-1}\mathbf{B}_T\right]^{-1}\mathbf{B}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{u}}/\omega_{uu} + o_P(1). \quad \text{(A.7)}$$

We consider each of the terms in the representation (A.7) in turn. A central limit theorem allows us to state

$$(T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*)^{-1/2}T^{-1/2}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}/\omega_{uu}^{1/2} \Rightarrow \mathbf{z}_r \sim N(\mathbf{0}, \mathbf{I}_r).$$

Hence, the first term in (A.7) converges in distribution to $\mathbf{z}_r'\mathbf{z}_r$, a chi-square random variable with r degrees of freedom; that is,

$$T^{-1}\bar{\mathbf{u}}'\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*(T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*)^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}/\omega_{uu} \Rightarrow \mathbf{z}_r'\mathbf{z}_r \sim \chi^2(r). \quad \text{(A.8)}$$

From Lemma A.1, the second term in (A.7) weakly converges to

$$\int_0^1 d\tilde{B}_u^*(a)\mathbf{G}_{k+1}(a)'\left(\int_0^1 \mathbf{G}_{k+1}(a)\mathbf{G}_{k+1}(a)'dr\right)^{-1}\int_0^1 \mathbf{G}_{k+1}(a)d\tilde{B}_u^*(a)/\omega_{uu},$$

which, as $\mathbf{C} = (\boldsymbol{\beta}_y^\perp, \boldsymbol{\beta}^\perp)[(\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)'\Gamma(\boldsymbol{\beta}_y^\perp, \boldsymbol{\beta}^\perp)]^{-1}(\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)'$, may be expressed as

$$\int_0^1 d\tilde{B}_u^*(a)\begin{pmatrix} (\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)'\bar{\mathbf{B}}_{k+1}(a) \\ a-\frac{1}{2} \end{pmatrix}'\left(\int_0^1 \begin{pmatrix} (\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)'\bar{\mathbf{B}}_{k+1}(a) \\ a-\frac{1}{2} \end{pmatrix}\begin{pmatrix} (\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)'\bar{\mathbf{B}}_{k+1}(a) \\ a-\frac{1}{2} \end{pmatrix}' da\right)^{-1}$$

$$\times \int_0^1 \begin{pmatrix} (\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)'\bar{\mathbf{B}}_{k+1}(a) \\ a-\frac{1}{2} \end{pmatrix} d\tilde{B}_u^*(a)/\omega_{uu}.$$

Now, noting that under H_0 of (3.1) we may express $\boldsymbol{\alpha}_y^{\perp} = (1, -\boldsymbol{\omega}')'$ and $\boldsymbol{\alpha}^{\perp} = (\mathbf{0}, \boldsymbol{\alpha}_{xx}^{\perp\prime})'$ where $\boldsymbol{\alpha}_{xx}^{\perp\prime}\boldsymbol{\alpha}_{xx} = \mathbf{0}$, we define the $(k-r+1)$-vector of independent de-meaned standard Brownian motions,

$$\bar{\mathbf{W}}_{k-r+1}(a)[= (\tilde{W}_u(a), \bar{\mathbf{W}}_{k-r}(a)')'] \equiv [(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\Omega(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})]^{-1/2}(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\bar{\mathbf{B}}_{k+1}(a)$$
$$= \begin{pmatrix} \omega_{uu}^{-1/2}\tilde{B}_u(a) \\ \left(\boldsymbol{\alpha}_{xx}^{\perp\prime}\Omega_{xx}\boldsymbol{\alpha}_{xx}^{\perp}\right)^{-1/2}\boldsymbol{\alpha}_{xx}^{\perp\prime}\bar{\mathbf{B}}_k(a) \end{pmatrix},$$

where $\tilde{B}_u^*(a) = \tilde{B}_1(a) - \boldsymbol{\omega}'\bar{\mathbf{B}}_k(a)$ is independent of $\bar{\mathbf{B}}_k(a)$ and $\bar{\mathbf{B}}_{k+1}(a) \equiv (\tilde{B}_1(a), \bar{\mathbf{B}}_k(a)')'$ is partitioned according to $\mathbf{z}_t = (y_t, \mathbf{x}_t')'$, $a \in [0,1]$. Hence, the second term in (A.7) has the following asymptotic representation

$$\int_0^1 d\tilde{W}_u(a)\begin{pmatrix} \bar{\mathbf{W}}_{k-r+1}(a) \\ a-\frac{1}{2}\end{pmatrix}'\left(\int_0^1 \begin{pmatrix} \bar{\mathbf{W}}_{k-r+1}(a) \\ a-\frac{1}{2}\end{pmatrix}\begin{pmatrix} \bar{\mathbf{W}}_{k-r+1}(a) \\ a-\frac{1}{2}\end{pmatrix}' da\right)^{-1}\int_0^1 \begin{pmatrix} \bar{\mathbf{W}}_{k-r+1}(a) \\ a-\frac{1}{2}\end{pmatrix} d\tilde{W}_u(a). \tag{A.9}$$

Note that $d\tilde{W}_u(a)$ in (A.9) may be replaced by $dW_u(a)$, $a \in [0,1]$. Combining (A.8) and (A.9) gives the result of Theorem 3.1.

For the remaining cases, we need only make minor modifications to the proof for Case IV. In Case I, $\boldsymbol{\delta} = (\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})$ with $\left(\boldsymbol{\beta}, \boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp}\right)$ a basis for $\mathcal{R}^{k+1}$ and $\mathbf{B}_T = \boldsymbol{\delta}$. For Case II, where $\mathbf{Z}_{-1}^* = (\boldsymbol{\iota}_T, \mathbf{Z}_{-1}')'$, we have $\boldsymbol{\beta}_* = \begin{pmatrix} -\boldsymbol{\mu}' \\ \mathbf{I}_{k+1}\end{pmatrix}\boldsymbol{\beta}$ and, consequently, we define $\boldsymbol{\xi}$ as in Case IV, $\boldsymbol{\delta} = \begin{pmatrix} -\boldsymbol{\mu}' \\ \mathbf{I}_{k+1}\end{pmatrix}(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})$ and $\mathbf{B}_T = (\boldsymbol{\delta}, \boldsymbol{\xi})$. Case III is similar to Case I as is Case V. ■

Proof of Corollary 3.1: Follows immediately from Theorem 3.1 by setting $r = k$. ■

Proof of Corollary 3.2: Follows immediately from Theorem 3.1 by setting $r = 0$. ■

Proof of Theorem 3.2: We provide a proof for Case V which may be simply adapted for Cases I and III. To emphasise the potential dependence of the limit distribution on nuisance parameters, the proof is initially conducted under Assumptions 1-4 together with Assumption 5a which implies $H_0^{\pi_{yy}} : \pi_{yy} = 0$ but not necessarily $H_0^{\pi_{yx.x}} : \boldsymbol{\pi}_{yx.x} = \mathbf{0}'$; in particular, note that we may write $\boldsymbol{\alpha}_y^{\perp} = (1, -\boldsymbol{\phi}')'$ for some k-vector $\boldsymbol{\phi}$. The t-statistic for $H_0^{\pi_{yy}} : \pi_{yy} = 0$ may be expressed as the square root of

$$\widehat{\Delta\mathbf{y}}'\bar{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}}\hat{\mathbf{Z}}_{-1}\mathbf{A}_T\left(\mathbf{A}_T'\hat{\mathbf{Z}}_{-1}'\bar{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-}\hat{\mathbf{Z}}_{-1}\mathbf{A}_T\right)^{-1}\mathbf{A}_T'\hat{\mathbf{Z}}_{-1}'\bar{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}}\widehat{\Delta\mathbf{y}}/\hat{\omega}_{uu} \tag{A.10}$$

where $\mathbf{A}_T \equiv T^{-1/2}\left(\boldsymbol{\beta}, T^{-1/2}\mathbf{B}_T\right)$ and $\mathbf{B}_T = (\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})$. Note that only the diagonal element of the inverse in (A.10) corresponding to $\boldsymbol{\beta}_y^{\perp}$ is relevant which implies that we only need to consider the blocks $T^{-2}\mathbf{B}_T'\hat{\mathbf{Z}}_{-1}'\bar{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-}\hat{\mathbf{Z}}_{-1}\mathbf{B}_T$ and $T^{-1}\mathbf{B}_T'\hat{\mathbf{Z}}_{-1}'\bar{\mathbf{P}}_{\widehat{\Delta\mathbf{Z}}_-,\hat{\mathbf{X}}_{-1}}\widehat{\Delta\mathbf{y}}$ in (A.10). Therefore, using (A.2) and (A.3), (A.10) is asymptotically equivalent to

$$T^{-1}\hat{\mathbf{u}}'\bar{\mathbf{P}}_{\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp}}\hat{\mathbf{Z}}_{-1}\mathbf{B}_T\left(T^{-2}\mathbf{B}_T'\hat{\mathbf{Z}}_{-1}'\hat{\mathbf{Z}}_{-1}\mathbf{B}_T\right)^{-1}T^{-1}\mathbf{B}_T'\hat{\mathbf{Z}}_{-1}'\bar{\mathbf{P}}_{\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp}}\hat{\mathbf{u}}/\omega_{uu}, \tag{A.11}$$

where $\bar{\mathbf{P}}_{\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp}} \equiv \mathbf{I}_T - \hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp}(\boldsymbol{\beta}_{xx}^{\perp\prime}\hat{\mathbf{X}}_{-1}'\hat{\mathbf{X}}_{-1}\boldsymbol{\beta}_{xx}^{\perp})^{-1}\boldsymbol{\beta}_{xx}^{\perp\prime}\hat{\mathbf{X}}_{-1}'$. Now,

$$T^{-1/2}\boldsymbol{\beta}_{xx}^{\perp\prime}\hat{\mathbf{x}}_{[Ta]} \Rightarrow (\mathbf{0}, \boldsymbol{\beta}_{xx}^{\perp\prime}\boldsymbol{\beta}_{xx}^{\perp})[(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\Gamma(\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})]^{-1}(\boldsymbol{\alpha}_y^{\perp}, \boldsymbol{\alpha}^{\perp})'\hat{\mathbf{B}}_{k+1}(a)$$
$$= (\boldsymbol{\beta}_{xx}^{\perp\prime}\boldsymbol{\beta}_{xx}^{\perp})[\boldsymbol{\alpha}_{xx}^{\perp\prime}(\Gamma_{xx} - \boldsymbol{\lambda}_{xy}^{\phi}\boldsymbol{\gamma}_{yx.x}^{\phi})\boldsymbol{\beta}_{xx}^{\perp}]^{-1}\boldsymbol{\alpha}_{xx}^{\perp\prime}\hat{\mathbf{B}}_k^{\phi}(a),$$

where for convenience, but without loss of generality, we have set $\boldsymbol{\beta}_y^{\perp} = (\beta_{yy}^{\perp}, \mathbf{0}')'$, $\boldsymbol{\lambda}^{\phi} \equiv \boldsymbol{\gamma}_{xy}/\gamma_{yy.x}^{\phi}$, $\gamma_{yy.x}^{\phi} \equiv \gamma_{yy} - \boldsymbol{\phi}'\boldsymbol{\gamma}_{xy}$, $\boldsymbol{\gamma}_{yx.x}^{\phi} \equiv \boldsymbol{\gamma}_{yx} - \boldsymbol{\phi}'\Gamma_{xx}$ and $\hat{\mathbf{B}}_k^{\phi}(a) \equiv \hat{\mathbf{B}}_k(a) - \boldsymbol{\lambda}^{\phi}\hat{B}_u^{\phi}(a)$, $\hat{B}_u^{\phi}(a) \equiv \hat{B}_1(a) - \boldsymbol{\phi}'\hat{\mathbf{B}}_k(a)$, $a \in [0,1]$. Hence, (A.11) weakly converges to

$$\left[\int_0^1 \hat{B}_u^{\phi}(a)dW_u(a) - \left(\int_0^1 \hat{B}_u^{\phi}(a)\hat{\mathbf{B}}_k^{\phi}(a)'da\right)\boldsymbol{\alpha}_{xx}^{\perp}\left[\boldsymbol{\alpha}_{xx}^{\perp\prime}\left(\int_0^1 \hat{\mathbf{B}}_k^{\phi}(a)\hat{\mathbf{B}}_k^{\phi}(a)'da\right)\boldsymbol{\alpha}_{xx}^{\perp}\right]^{-1}\boldsymbol{\alpha}_{xx}^{\perp\prime}\left(\int_0^1 \hat{\mathbf{B}}_k^{\phi}(a)dW_u(a)\right)\right]^2$$

$$\div\left[\int_0^1 \hat{B}_u^\phi(a)^2 da - \left(\int_0^1 \hat{B}_u^\phi(a)\hat{\mathbf{B}}_k^\phi(a)'da\right)\boldsymbol{\alpha}_{xx}^\perp\left[\boldsymbol{\alpha}_{xx}^{\perp\prime}\left(\int_0^1 \hat{\mathbf{B}}_k^\phi(a)\hat{\mathbf{B}}_k^\phi(a)'da\right)\boldsymbol{\alpha}_{xx}^\perp\right]^{-1}\boldsymbol{\alpha}_{xx}^{\perp\prime}\left(\int_0^1 \hat{\mathbf{B}}_k^\phi(a)\hat{B}_u^\phi(a)da\right)\right].$$

Under the conditions of the theorem, $\boldsymbol{\phi} = \boldsymbol{\omega}$ and $\boldsymbol{\lambda}^\phi = \mathbf{0}$ and, therefore, $\hat{B}_u^\phi(a)[= \hat{B}_u^*(a)] = \omega_{uu}^{1/2}\hat{W}_u(a)$ and $\boldsymbol{\alpha}_{xx}^{\perp\prime}\hat{\mathbf{B}}_k^\phi(a)[= \boldsymbol{\alpha}_{xx}^{\perp\prime}\hat{\mathbf{B}}_k(a)] = \left(\boldsymbol{\alpha}_{xx}^{\perp\prime}\Omega_{xx}\boldsymbol{\alpha}_{xx}^\perp\right)^{1/2}\hat{\mathbf{W}}_{k-r}(a)$, $a \in [0,1]$. ■

Proof of Corollary 3.3: Follows immediately from Theorem 3.2 by setting $r = k$. ■

Proof of Corollary 3.4: Follows immediately from Theorem 3.2 by setting $r = 0$. ■

Appendix B: Proofs for Section 4

Proof of Theorem 4.1: Again, we consider Case IV; the remaining Cases I-III and V may be dealt with similarly. Under $H_1^{\pi_{yy}} : \pi_{yy} \neq 0$, Assumptions 5b holds and, thus, $\Pi = \boldsymbol{\alpha}_y\boldsymbol{\beta}_y' + \boldsymbol{\alpha}\boldsymbol{\beta}'$ where $\boldsymbol{\alpha}_y = (\alpha_{yy}, \mathbf{0}')'$ and $\boldsymbol{\beta}_y = (\beta_{yy}, \boldsymbol{\beta}_{yx}')'$; see above Assumption 5b. Under Assumptions 1-4 and 5b, the process $\{\mathbf{z}_t\}_{t=1}^\infty$ has the infinite moving-average representation

$$\mathbf{z}_t = \boldsymbol{\mu} + \boldsymbol{\gamma}t + \mathbf{C}\mathbf{s}_t + \mathbf{C}^*(L)\boldsymbol{\varepsilon}_t,$$

but where now $\mathbf{C} \equiv \boldsymbol{\beta}^\perp[\boldsymbol{\alpha}^{\perp\prime}\Gamma\boldsymbol{\beta}^\perp]^{-1}\boldsymbol{\alpha}^{\perp\prime}$. We re-define $\boldsymbol{\beta}_*$and $\boldsymbol{\delta}$ as the $(k+2, r+1)$ and $(k+2, k-r)$ matrices

$$\boldsymbol{\beta}_* \equiv \begin{pmatrix} -\boldsymbol{\gamma}' \\ \mathbf{I}_{k+1} \end{pmatrix}(\boldsymbol{\beta}_y, \boldsymbol{\beta}), \boldsymbol{\delta} \equiv \begin{pmatrix} -\boldsymbol{\gamma}' \\ \mathbf{I}_{k+1} \end{pmatrix}\boldsymbol{\beta}^\perp,$$

where $\boldsymbol{\beta}^\perp$ is a $(k+1, k-r)$ matrix whose columns are a basis for the orthogonal complement of $(\boldsymbol{\beta}_y, \boldsymbol{\beta})$. Hence, $(\boldsymbol{\beta}_y, \boldsymbol{\beta}, \boldsymbol{\beta}^\perp)$ is a basis for $\mathcal{R}^{k+1}$ and, thus, $(\boldsymbol{\beta}_*, \boldsymbol{\xi}, \boldsymbol{\delta})$ a basis for $\mathcal{R}^{k+2}$, where again $\boldsymbol{\xi}$ is the $(k+2)$-unit vector $(1, \mathbf{0}')'$. It therefore follows that

$$T^{-1/2}\boldsymbol{\delta}'\mathbf{z}_{[Ta]}^* = T^{-1/2}\boldsymbol{\beta}^{\perp\prime}\boldsymbol{\mu} + T^{-1/2}\boldsymbol{\beta}^{\perp\prime}\mathbf{C}\mathbf{s}_{[Ta]} + \boldsymbol{\beta}^{\perp\prime}T^{-1/2}\mathbf{C}^*(L)\boldsymbol{\varepsilon}_{[Ta]}$$

$$\Rightarrow \boldsymbol{\beta}^{\perp\prime}\mathbf{C}\mathbf{B}_{k+1}(a).$$

Also, as above, $T^{-1}\boldsymbol{\xi}'\mathbf{z}_t^* = T^{-1}t \Rightarrow a$ and $\boldsymbol{\beta}_*'\mathbf{z}_t^* = (\boldsymbol{\beta}_y, \boldsymbol{\beta})'\boldsymbol{\mu} + (\boldsymbol{\beta}_y, \boldsymbol{\beta})'\mathbf{C}^*(L)\boldsymbol{\varepsilon}_t = O_P(1)$.

The Wald statistic (3.5) multiplied by $\hat{\omega}_{uu}$ may be written as

$$\bar{\mathbf{u}}'\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}^*\mathbf{A}_T\left(\mathbf{A}_T'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}^*\mathbf{A}_T\right)^{-1}\mathbf{A}_T'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}$$

$$+2\boldsymbol{\lambda}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}} + \boldsymbol{\lambda}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}^*\boldsymbol{\lambda}_*, \tag{B.1}$$

where $\boldsymbol{\lambda}_* \equiv \boldsymbol{\beta}_*(\boldsymbol{\alpha}_y, \boldsymbol{\alpha})'(1, -\boldsymbol{\omega}')'$, $\mathbf{A}_T \equiv T^{-1/2}\left(\boldsymbol{\beta}_*, T^{-1/2}\mathbf{B}_T\right)$ and $\mathbf{B}_T \equiv \left(\boldsymbol{\delta}, T^{-1/2}\boldsymbol{\xi}\right)$. Note that (A.6) continues to hold under $H_1^{\pi_{yy}} : \pi_{yy} \neq 0$. A similar argument to that in the Proof of Theorem 3.1 demonstrates that the first term in (B.1) divided by ω_{uu} has the limiting representation

$$\mathbf{z}_{r+1}'\mathbf{z}_{r+1} + \int_0^1 dW_u(a)\mathbf{F}_{k-r}(a)'\left(\int_0^1 \mathbf{F}_{k-r}(a)\mathbf{F}_{k-r}(a)'da\right)^{-1}\int_0^1 \mathbf{F}_{k-r}(a)dW_u(a), \tag{B.2}$$

where $\mathbf{z}_{r+1} \sim N(\mathbf{0}, \mathbf{I}_{r+1})$, $\mathbf{F}_{k-r}(a) = (\bar{\mathbf{W}}_{k-r}(a)', a - \frac{1}{2})'$ and $\bar{\mathbf{W}}_{k-r}(a) \equiv \left(\boldsymbol{\alpha}_{xx}^{\perp\prime}\Omega_{xx}\boldsymbol{\alpha}_{xx}^\perp\right)^{-1/2}\boldsymbol{\alpha}_{xx}^{\perp\prime}\bar{\mathbf{B}}_k(a)$ is a $(k-r)$-vector of de-meaned independent standard Brownian motions independent of the standard Brownian motion $W_u(a)$, $a \in [0,1]$; cf. (3.9). Now, $\int_0^1 \mathbf{F}_{k-r}(a)dW_u(a)$ is mixed normal with conditional variance matrix $\int_0^1 \mathbf{F}_{k-r}(a)\mathbf{F}_{k-r}(a)'da$. Therefore, the second term in (B.2) is unconditionally distributed as a $\chi^2(k-r)$ random variable and is independent of the first term; cf. (A.4). Hence, the first term in (B.1) divided by ω_{uu} has a limiting $\chi^2(k+1)$ distribution.

The second term in (B.1) may be written as

$$2(1,-\boldsymbol{\omega}')(\boldsymbol{\alpha}_y,\boldsymbol{\alpha})\boldsymbol{\beta}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}} = 2T^{1/2}(1,-\boldsymbol{\omega}')(\boldsymbol{\alpha}_y,\boldsymbol{\alpha})\left(T^{-1/2}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}\right) = O_P(T^{1/2}), \tag{B.3}$$

and the third term as

$$(1,-\boldsymbol{\omega}')(\boldsymbol{\alpha}_y,\boldsymbol{\alpha})\boldsymbol{\beta}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*(\boldsymbol{\alpha}_y,\boldsymbol{\alpha})'(1,-\boldsymbol{\omega}')'$$

$$= T(1,-\boldsymbol{\omega}')(\boldsymbol{\alpha}_y,\boldsymbol{\alpha})(T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*)(\boldsymbol{\alpha}_y,\boldsymbol{\alpha})'(1,-\boldsymbol{\omega}')' = O_P(T), \tag{B.4}$$

as $T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*$ converges in probability to a positive definite matrix. Moreover, as $(1,-\boldsymbol{\omega}')(\boldsymbol{\alpha}_y,\boldsymbol{\alpha}) \neq \mathbf{0}'$ under $H_1^{\pi_{yy}}: \pi_{yy} \neq 0$, the Theorem is proved. ■

Proof of Theorem 4.2: A similar decomposition to (B.1) for the Wald statistic (3.5) holds under $H_1^{\boldsymbol{\pi}_{yx.x}} \cap H_0^{\pi_{yy}}$ except that $\boldsymbol{\beta}_*$ and $\boldsymbol{\delta}$ are now as defined in the Proof of Theorem 3.1. Although $H_0^{\pi_{yy}}: \pi_{yy} = 0$ holds, we have $H_1^{\boldsymbol{\pi}_{yx.x}}: \boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$. Therefore, as in Theorem 3.2, note that we may write $\boldsymbol{\alpha}_y^{\perp} = (1,-\boldsymbol{\phi}')'$ for some k-vector $\boldsymbol{\phi} \neq \boldsymbol{\omega}$. Consequently, the first term divided by ω_{uu} may be written as

$$T^{-1}\bar{\mathbf{u}}'\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\left(T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\right)^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}/\omega_{uu}$$

$$+T^{-2}\bar{\mathbf{u}}'\bar{\mathbf{Z}}_{-1}^{*}\mathbf{B}_T\left[T^{-2}\mathbf{B}_T'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{Z}}_{-1}^{*}\mathbf{B}_T\right]^{-1}\mathbf{B}_T'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{u}}/\omega_{uu} + o_P(1); \tag{B.5}$$

cf. (A.7). As in the Proof of Theorem 3.1, the first term of (B.5) has the limiting representation $\mathbf{z}_r'\mathbf{z}_r$ where $\mathbf{z}_r \sim N(\mathbf{0},\mathbf{I}_r)$; cf. (3.9). The second term of (B.5) has the limiting representation

$$\int_0^1 d\tilde{B}_u^*(a)\begin{pmatrix}\tilde{B}_u^{\phi}(a)\\ \boldsymbol{\alpha}_{xx}^{\perp\prime}\bar{\mathbf{B}}_k(a)\\ a-\frac{1}{2}\end{pmatrix}'\left(\int_0^1\begin{pmatrix}\tilde{B}_u^{\phi}(a)\\ \boldsymbol{\alpha}_{xx}^{\perp\prime}\bar{\mathbf{B}}_k(a)\\ a-\frac{1}{2}\end{pmatrix}\begin{pmatrix}\tilde{B}_u^{\phi}(a)\\ \boldsymbol{\alpha}_{xx}^{\perp\prime}\bar{\mathbf{B}}_k(a)\\ a-\frac{1}{2}\end{pmatrix}'da\right)^{-1}$$

$$\times\int_0^1\begin{pmatrix}\tilde{B}_u^{\phi}(a)\\ \boldsymbol{\alpha}_{xx}^{\perp\prime}\bar{\mathbf{B}}_k(a)\\ a-\frac{1}{2}\end{pmatrix}d\tilde{B}_u^*(a)/\omega_{uu} = O_P(1),$$

where $\tilde{B}_u^{\phi}(a) \equiv \tilde{B}_1(a) - \boldsymbol{\phi}'\bar{\mathbf{B}}_k(a)$, $a \in [0,1]$; cf. Proof of Theorem 3.2. The second term of (B.1) becomes

$$2(1,-\boldsymbol{\omega}')\boldsymbol{\alpha}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}} = 2T^{1/2}(1,-\boldsymbol{\omega}')\boldsymbol{\alpha}\left(T^{-1/2}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}\right) = O_P(T^{1/2}),$$

and the third term

$$(1,-\boldsymbol{\omega}')\boldsymbol{\alpha}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*\boldsymbol{\alpha}'(1,-\boldsymbol{\omega}')'$$

$$= T(1,-\boldsymbol{\omega}')\boldsymbol{\alpha}(T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_*)\boldsymbol{\alpha}'(1,-\boldsymbol{\omega}')' = O_P(T).$$

The Theorem follows as $(1,-\boldsymbol{\omega}')\boldsymbol{\alpha} \neq \mathbf{0}'$ under $H_0^{\pi_{yy}}: \pi_{yy} = 0$ and $H_1^{\boldsymbol{\pi}_{yx.x}}: \boldsymbol{\pi}_{yx.x} \neq \mathbf{0}'$. ■

Proof of Theorem 4.3: We concentrate on Case IV; the remaining Cases I-III and V are proved by a similar argument. Let $\{\mathbf{z}_{tT}\}_{t=1}^T$ denote the process under H_{1T} of (4.1), $T = 1, 2, \ldots$. Hence, $\Phi(L)(\mathbf{z}_{tT} - \boldsymbol{\mu} - \boldsymbol{\gamma}t) = \boldsymbol{\xi}_{tT}$, where $\boldsymbol{\xi}_{tT} \equiv (\Pi_T - \Pi)[\mathbf{z}_{(t-1)T} - \boldsymbol{\mu} - \boldsymbol{\gamma}(t-1)] + \boldsymbol{\varepsilon}_t$ and $\Pi_T - \Pi$ is given in (4.2). Therefore, $\Delta(\mathbf{z}_{tT} - \boldsymbol{\mu} - \boldsymbol{\gamma}t) = \mathbf{C}\boldsymbol{\xi}_{tT} + \mathbf{C}^*(L)\Delta\boldsymbol{\xi}_{tT}$, $\mathbf{C}(z) = \mathbf{C} + (1-z)\mathbf{C}^*(z)$ and $\mathbf{C} = (\boldsymbol{\beta}_y^{\perp},\boldsymbol{\beta}^{\perp})[(\boldsymbol{\alpha}_y^{\perp},\boldsymbol{\alpha}^{\perp})'\Gamma(\boldsymbol{\beta}_y^{\perp},\boldsymbol{\beta}^{\perp})]^{-1}(\boldsymbol{\alpha}_y^{\perp},\boldsymbol{\alpha}^{\perp})'$, and, thus,

$$[\mathbf{I}_{k+1} - (\mathbf{I}_{k+1} + T^{-1}\mathbf{C}\boldsymbol{\alpha}_y\boldsymbol{\beta}_y')L](\mathbf{z}_{tT} - \boldsymbol{\mu} - \boldsymbol{\gamma}t) = \mathbf{C}\boldsymbol{\varepsilon}_{tT} + \mathbf{C}^*(L)\Delta\boldsymbol{\xi}_{tT}, \tag{B.6}$$

where

$$\boldsymbol{\varepsilon}_{tT} \equiv T^{-1/2}\begin{pmatrix}\boldsymbol{\delta}_{yx}\\ \boldsymbol{\delta}_{xx}\end{pmatrix}\boldsymbol{\beta}'[\mathbf{z}_{(t-1)T} - \boldsymbol{\mu} - \boldsymbol{\gamma}(t-1)] + \boldsymbol{\varepsilon}_t,$$

$t = 1, ..., T$, $T = 1, 2, ...$ Inverting (B.6) yields

$$\mathbf{z}_{tT} = (\mathbf{I}_{k+1} + T^{-1}\mathbf{C}\boldsymbol{\alpha}_y\boldsymbol{\beta}_y')^s(\mathbf{z}_{sT} - \boldsymbol{\mu} - \boldsymbol{\gamma}s) + \boldsymbol{\mu} + \boldsymbol{\gamma}t + \sum_{i=0}^{s-1}(\mathbf{I}_{k+1} + T^{-1}\mathbf{C}\boldsymbol{\alpha}_y\boldsymbol{\beta}_y')^i[\mathbf{C}\boldsymbol{\varepsilon}_{(t-i)T} + \mathbf{C}^*(L)\Delta\boldsymbol{\xi}_{(t-i)T}];$$

note that $\Delta\boldsymbol{\xi}_{tT} = (\Pi_T - \Pi)\Delta[\mathbf{z}_{(t-1)T} - \boldsymbol{\mu} - \boldsymbol{\gamma}(t-1)] + \Delta\boldsymbol{\varepsilon}_t$. It therefore follows that

$$T^{-1/2}\boldsymbol{\delta}'\mathbf{z}^*_{[Ta]T} \Rightarrow (\boldsymbol{\beta}_y^{\perp}, \boldsymbol{\beta}^{\perp})'\mathbf{C}\mathbf{J}_{k+1}(a),$$

where $\boldsymbol{\delta}$ is defined above Lemma A.1 and $\mathbf{z}^*_{tT} = (t, \mathbf{z}'_{tT})'$, $\mathbf{J}_{k+1}(a) \equiv \int_0^a \exp\{\boldsymbol{\alpha}_y\boldsymbol{\beta}_y'\mathbf{C}(a-r)\}d\mathbf{B}_{k+1}(r)$ is an Ornstein-Uhlenbeck process and $\mathbf{B}_{k+1}(a)$ is a $(k+1)$-vector Brownian motion with variance matrix Ω, $a \in [0, 1]$; cf. Johansen (1995, Theorem 14.1, p.202).

Similarly to (A.4),

$$\mathbf{A}_T'\bar{\mathbf{Z}}_{-1}'\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}\mathbf{A}_T = \begin{pmatrix} T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_* & \mathbf{0}' \\ \mathbf{0} & T^{-2}\mathbf{B}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{Z}}^*_{-1}\mathbf{B}_T \end{pmatrix} + o_P(1).$$

Therefore, the expression (B.1) for the Wald statistic (3.5) multiplied by $\hat{\omega}_{uu}$ is revised to

$$\hat{\omega}_{uu}W = T^{-1}\widetilde{\Delta\mathbf{y}}'\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\left(T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\right)^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\widetilde{\Delta\mathbf{y}}$$

$$+T^{-2}\widetilde{\Delta\mathbf{y}}'\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\mathbf{B}_T\left[T^{-2}\mathbf{B}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{Z}}^*_{-1}\mathbf{B}_T\right]^{-1}\mathbf{B}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\widetilde{\Delta\mathbf{y}} + o_P(1). \tag{B.7}$$

The first term in (B.7) may be written as

$$T^{-1}\bar{\mathbf{u}}'\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\left(T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\right)^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}$$

$$+2T^{-1}\bar{\mathbf{u}}'\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\left(T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\right)^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\pi}^{*\prime}_{yT}$$

$$+T^{-1}\boldsymbol{\pi}^*_{yT}\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\left(T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\right)^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\pi}^{*\prime}_{yT}, \tag{B.8}$$

where $\boldsymbol{\pi}^*_{yT} \equiv T^{-1}\alpha_{yy}\boldsymbol{\beta}_{y*}' + T^{-1/2}(\boldsymbol{\delta}_{yx} - \boldsymbol{\omega}'\boldsymbol{\delta}_{xx})\boldsymbol{\beta}_*'$. Defining $\boldsymbol{\eta} \equiv (\boldsymbol{\delta}_{yx} - \boldsymbol{\omega}'\boldsymbol{\delta}_{xx})'$, consider

$$T^{-1/2}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\pi}^{*\prime}_{yT} = T^{-1/2}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}(\boldsymbol{\beta}_{y*}\alpha_{yy}T^{-1} + \boldsymbol{\beta}_*\boldsymbol{\eta}T^{-1/2})$$

$$= T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\boldsymbol{\eta} + o_P(1),$$

where we have made use of $T^{-1/2}\boldsymbol{\beta}_{y*}'\mathbf{z}^*_{[Ta]T} \Rightarrow \boldsymbol{\beta}_y'\mathbf{C}\mathbf{J}_{k+1}(a)$. Therefore, (B.8) divided by ω_{uu} may be re-expressed as

$$\left[\left(T^{-1/2}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}\right) + \mathbf{Q}\boldsymbol{\eta}\right]'\mathbf{Q}^{-1}\left[\left(T^{-1/2}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{u}}\right) + \mathbf{Q}\boldsymbol{\eta}\right]/\omega_{uu} + o_P(1)$$

$$= \mathbf{z}_r'\mathbf{z}_r + o_P(1), \tag{B.9}$$

where $\mathbf{Q} \equiv p\lim_{T\to\infty}\left(T^{-1}\boldsymbol{\beta}_*'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\right)$ and $\mathbf{z}_r \sim N(\mathbf{Q}^{1/2}\boldsymbol{\eta}, \mathbf{I}_r)$.

As $\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\widetilde{\Delta\mathbf{y}} = \bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}(\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\pi}^{*\prime}_{yT} + \bar{\mathbf{u}})$, $T^{-1}\mathbf{B}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\widetilde{\Delta\mathbf{y}} = T^{-1}\mathbf{B}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}(\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\pi}^{*\prime}_{yT} + \bar{\mathbf{u}})$. Consider the second term in (B.7), in particular, $T^{-1}\mathbf{B}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\pi}^{*\prime}_{yT}$ which after substitution for $\boldsymbol{\pi}^*_{yT}$ becomes

$$T^{-2}\mathbf{B}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_{y*}\alpha_{yy} + T^{-3/2}\mathbf{B}_T'\bar{\mathbf{Z}}^{*\prime}_{-1}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}^*_{-1}\boldsymbol{\beta}_*\boldsymbol{\eta}$$

$$= T^{-2}\mathbf{B}_T'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\bar{\mathbf{Z}}_{-1}^{*}\boldsymbol{\beta}_{y*}\alpha_{yy} + o_P(1)$$

$$\Rightarrow \int_0^1 \begin{pmatrix} (\boldsymbol{\beta}_y^\perp, \boldsymbol{\beta}^\perp)'\mathbf{C}\bar{\mathbf{J}}_{k+1}(a) \\ a - \frac{1}{2} \end{pmatrix} \bar{\mathbf{J}}_{k+1}(a)'\mathbf{C}'\boldsymbol{\beta}_y\alpha_{yy}da.$$

Therefore,

$$T^{-1}\mathbf{B}_T'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\widetilde{\Delta\mathbf{y}} \Rightarrow \int_0^1 \begin{pmatrix} (\boldsymbol{\beta}_y^\perp, \boldsymbol{\beta}^\perp)'\mathbf{C}\bar{\mathbf{J}}_{k+1}(a) \\ a - \frac{1}{2} \end{pmatrix} (\omega_{uu}^{1/2}d\tilde{W}_u(a) + \bar{\mathbf{J}}_{k+1}(a)'\mathbf{C}'\boldsymbol{\beta}_y\alpha_{yy}da).$$

Consider

$$\bar{\mathbf{J}}_{k-r+1}^*(a)[= (\tilde{J}_u^*(a), \bar{\mathbf{J}}_{k-r}^*(a)')'] \equiv [(\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)'\Omega(\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)]^{-1/2}(\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)'\bar{\mathbf{J}}_{k+1}(a)$$

$$= \begin{pmatrix} \omega_{uu}^{-1/2}\tilde{J}_u(a) \\ \left(\boldsymbol{\alpha}_{xx}^{\perp\prime}\Omega_{xx}\boldsymbol{\alpha}_{xx}^\perp\right)^{-1/2}\boldsymbol{\alpha}_{xx}^{\perp\prime}\bar{\mathbf{J}}_k(a) \end{pmatrix},$$

where $\tilde{J}_u(a) = \tilde{J}_1(a) - \boldsymbol{\omega}'\bar{\mathbf{J}}_k(a)$ is independent of $\bar{\mathbf{J}}_k(a)$ and $\bar{\mathbf{J}}_{k+1}(a) \equiv (\tilde{J}_1(a), \bar{\mathbf{J}}_k(a)')'$, $a \in [0,1]$. Now, $\bar{\mathbf{J}}_{k-r+1}^*(a)$ satisfies the stochastic integral and differential equations

$$\bar{\mathbf{J}}_{k-r+1}^*(a) = \bar{\mathbf{W}}_{k-r+1}(a) + \mathbf{ab}'\int_0^a \bar{\mathbf{J}}_{k-r+1}^*(r)dr,$$

and

$$d\bar{\mathbf{J}}_{k-r+1}^*(a) = d\bar{\mathbf{W}}_{k-r+1}(a) + \mathbf{ab}'\bar{\mathbf{J}}_{k-r+1}^*(a)da,$$

where

$$\mathbf{a} = [(\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)'\Omega(\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)]^{-1/2}(\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)'\boldsymbol{\alpha}_y,$$

$$\mathbf{b} = [(\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)'\Omega(\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)]^{1/2}[(\boldsymbol{\beta}_y^\perp, \boldsymbol{\beta}^\perp)'\Gamma(\boldsymbol{\alpha}_y^\perp, \boldsymbol{\alpha}^\perp)]^{-1}(\boldsymbol{\beta}_y^\perp, \boldsymbol{\beta}^\perp)'\boldsymbol{\beta}_y;$$

cf. Johansen (1995, Theorem 14.4, p.207). Note that the first element of $\bar{\mathbf{J}}_{k-r+1}^*(a)$ satisfies

$$\tilde{J}_u^*(a) = \tilde{W}_u(a) + \omega_{uu}^{-1/2}\alpha_{yy}\mathbf{b}'\int_0^a \bar{\mathbf{J}}_{k-r+1}^*(r)dr,$$

and

$$d\tilde{J}_u^*(a) = d\tilde{W}_u(a) + \omega_{uu}^{-1/2}\alpha_{yy}\mathbf{b}'\bar{\mathbf{J}}_{k-r+1}^*(a)da.$$

Therefore,

$$T^{-1}\mathbf{B}_T'\bar{\mathbf{Z}}_{-1}^{*\prime}\bar{\mathbf{P}}_{\widetilde{\Delta\mathbf{Z}}_-}\widetilde{\Delta\mathbf{y}} \Rightarrow \int_0^1 \begin{pmatrix} (\boldsymbol{\beta}_y^\perp, \boldsymbol{\beta}^\perp)'\mathbf{C}\bar{\mathbf{J}}_{k+1}(a) \\ a - \frac{1}{2} \end{pmatrix} \omega_{uu}^{1/2}d\tilde{J}_u^*(a).$$

Hence, the second term in (B.7) weakly converges to

$$\omega_{uu}\int_0^1 d\tilde{J}_u^*(a)\mathbf{F}_{k-r+1}(a)'\left(\int_0^1 \mathbf{F}_{k-r+1}(a)\mathbf{F}_{k-r+1}(a)'da\right)^{-1}\int_0^1 \mathbf{F}_{k-r+1}(a)d\tilde{J}_u^*(a), \tag{B.10}$$

where $\mathbf{F}_{k-r+1}(a) = (\bar{\mathbf{J}}_{k-r+1}^*(a)', a - \frac{1}{2})'$.

Combining (B.9) and (B.10) gives the result stated in Theorem 4.3 as $\hat{\omega}_{uu} - \omega_{uu} = o_P(1)$ under H_{1T} of (4.1) and noting $d\tilde{J}_u^*(a)$ may be replaced by $dJ_u^*(a)$. ■

Table C1. Critical Value Bounds for the F-Statistic

Testing for the Existence of A Long-Run Relationship*

Table C1.i: Case I with no intercept and no trend

	90%		95%		97.5%		99%		mean		variance	
k	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$
0	3.00	3.00	4.20	4.20	5.47	5.47	7.17	7.17	1.16	1.16	2.32	2.32
1	2.44	3.28	3.15	4.11	3.88	4.92	4.81	6.02	1.08	1.54	1.08	1.73
2	2.17	3.19	2.72	3.83	3.22	4.50	3.88	5.30	1.05	1.69	0.70	1.27
3	2.01	3.10	2.45	3.63	2.87	4.16	3.42	4.84	1.04	1.77	0.52	0.99
4	1.90	3.01	2.26	3.48	2.62	3.90	3.07	4.44	1.03	1.81	0.41	0.80
5	1.81	2.93	2.14	3.34	2.44	3.71	2.82	4.21	1.02	1.84	0.34	0.67
6	1.75	2.87	2.04	3.24	2.32	3.59	2.66	4.05	1.02	1.86	0.29	0.58
7	1.70	2.83	1.97	3.18	2.22	3.49	2.54	3.91	1.02	1.88	0.26	0.51
8	1.66	2.79	1.91	3.11	2.15	3.40	2.45	3.79	1.02	1.89	0.23	0.46
9	1.63	2.75	1.86	3.05	2.08	3.33	2.34	3.68	1.02	1.90	0.20	0.41
10	1.60	2.72	1.82	2.99	2.02	3.27	2.26	3.60	1.02	1.91	0.19	0.37

Table C1.ii: Case II with retricted intercept and no trend

	90%		95%		97.5%		99%		mean		variance	
k	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$
0	7.60	7.60	9.20	9.20	10.79	10.79	12.88	12.88	4.07	4.07	7.07	7.07
1	4.54	5.27	5.43	6.24	6.27	7.19	7.41	8.37	2.53	3.03	2.28	2.81
2	3.51	4.46	4.13	5.16	4.73	5.84	5.50	6.67	2.02	2.69	1.23	1.71
3	2.97	4.00	3.49	4.58	3.94	5.10	4.56	5.83	1.76	2.52	0.82	1.22
4	2.65	3.71	3.07	4.19	3.46	4.65	3.95	5.24	1.61	2.41	0.60	0.94
5	2.43	3.50	2.78	3.94	3.16	4.35	3.57	4.84	1.51	2.34	0.48	0.76
6	2.27	3.36	2.60	3.75	2.91	4.13	3.29	4.56	1.44	2.29	0.39	0.64
7	2.16	3.25	2.45	3.61	2.74	3.95	3.07	4.39	1.38	2.26	0.33	0.56
8	2.06	3.17	2.34	3.50	2.59	3.80	2.91	4.19	1.34	2.23	0.29	0.50
9	1.98	3.08	2.24	3.39	2.47	3.69	2.76	4.05	1.31	2.21	0.25	0.44
10	1.92	3.02	2.16	3.32	2.38	3.58	2.63	3.94	1.28	2.19	0.23	0.40

Table C1.iii: Case III with unrestricted intercept and no trend

	90%		95%		97.5%		99%		mean		variance	
k	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$
0	6.58	6.58	8.21	8.21	9.80	9.80	11.79	11.79	3.05	3.05	7.07	7.07
1	4.04	4.78	4.94	5.73	5.77	6.68	6.84	7.84	2.03	2.52	2.28	2.89
2	3.17	4.14	3.79	4.85	4.41	5.52	5.15	6.36	1.69	2.35	1.23	1.77
3	2.72	3.77	3.23	4.35	3.69	4.89	4.29	5.61	1.51	2.26	0.82	1.27
4	2.45	3.52	2.86	4.01	3.25	4.49	3.74	5.06	1.41	2.21	0.60	0.98
5	2.26	3.35	2.62	3.79	2.96	4.18	3.41	4.68	1.34	2.17	0.48	0.79
6	2.12	3.23	2.45	3.61	2.75	3.99	3.15	4.43	1.29	2.14	0.39	0.66
7	2.03	3.13	2.32	3.50	2.60	3.84	2.96	4.26	1.26	2.13	0.33	0.58
8	1.95	3.06	2.22	3.39	2.48	3.70	2.79	4.10	1.23	2.12	0.29	0.51
9	1.88	2.99	2.14	3.30	2.37	3.60	2.65	3.97	1.21	2.10	0.25	0.45
10	1.83	2.94	2.06	3.24	2.28	3.50	2.54	3.86	1.19	2.09	0.23	0.41

Table C1.iv: Case IV with unrestricted intercept and restricted trend

	90%		95%		97.5%		99%		mean		variance	
k	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$
0	10.75	10.75	12.57	12.57	14.27	14.27	16.51	16.51	6.35	6.35	10.72	10.72
1	6.07	6.74	7.02	7.73	7.94	8.74	9.15	10.09	3.67	4.15	3.18	3.72
2	4.50	5.35	5.17	6.15	5.82	6.88	6.65	7.80	2.78	3.43	1.63	2.13
3	3.71	4.68	4.23	5.29	4.75	5.85	5.38	6.54	2.33	3.06	1.04	1.45
4	3.22	4.24	3.66	4.76	4.07	5.23	4.57	5.90	2.07	2.84	0.74	1.10
5	2.90	3.94	3.28	4.39	3.63	4.82	4.09	5.40	1.89	2.70	0.57	0.88
6	2.67	3.72	3.00	4.13	3.32	4.51	3.73	5.02	1.76	2.60	0.46	0.72
7	2.49	3.57	2.81	3.94	3.10	4.29	3.46	4.75	1.67	2.52	0.39	0.62
8	2.36	3.44	2.65	3.79	2.91	4.11	3.26	4.52	1.60	2.46	0.33	0.54
9	2.26	3.32	2.53	3.66	2.77	3.96	3.06	4.33	1.54	2.41	0.29	0.48
10	2.16	3.24	2.41	3.55	2.64	3.84	2.93	4.19	1.49	2.38	0.26	0.43

Table C1.v: Case V with unrestricted intercept and unrestricted trend

	90%		95%		97.5%		99%		mean		variance	
k	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$
0	9.81	9.81	11.64	11.64	13.36	13.36	15.73	15.73	5.33	5.33	11.35	11.35
1	5.59	6.26	6.56	7.30	7.46	8.27	8.74	9.63	3.17	3.64	3.33	3.91
2	4.19	5.06	4.87	5.85	5.49	6.59	6.34	7.52	2.44	3.09	1.70	2.23
3	3.47	4.45	4.01	5.07	4.52	5.62	5.17	6.36	2.08	2.81	1.08	1.51
4	3.03	4.06	3.47	4.57	3.89	5.07	4.40	5.72	1.86	2.64	0.77	1.14
5	2.75	3.79	3.12	4.25	3.47	4.67	3.93	5.23	1.72	2.53	0.59	0.91
6	2.53	3.59	2.87	4.00	3.19	4.38	3.60	4.90	1.62	2.45	0.48	0.75
7	2.38	3.45	2.69	3.83	2.98	4.16	3.34	4.63	1.54	2.39	0.40	0.64
8	2.26	3.34	2.55	3.68	2.82	4.02	3.15	4.43	1.48	2.35	0.34	0.56
9	2.16	3.24	2.43	3.56	2.67	3.87	2.97	4.24	1.43	2.31	0.30	0.49
10	2.07	3.16	2.33	3.46	2.56	3.76	2.84	4.10	1.40	2.28	0.26	0.44

* The critical values are computed via stochastic simulations using $T = 1,000$ and 40,000 replications for the F statistic for testing $\boldsymbol{\varphi} = \mathbf{0}$ in the following regressions: $\Delta y_t = \boldsymbol{\varphi}'\mathbf{z}_{t-1} + \mathbf{a}'\mathbf{w}_t + \xi_t$, $t = 1, 2, ..., T$, where $\mathbf{x}_{t-1} = (x_{1,t-1}, \dots, x_{k,t-1})'$,

$$\left\{ \begin{array}{lc} \mathbf{z}_{t-1} = \left(y_{t-1}, \mathbf{x}'_{t-1}\right)', \ \mathbf{w}_t = \varnothing & \text{Case I} \\ \mathbf{z}_{t-1} = \left(y_{t-1}, \mathbf{x}'_{t-1}, 1\right)', \ \mathbf{w}_t = \varnothing & \text{Case II} \\ \mathbf{z}_{t-1} = \left(y_{t-1}, \mathbf{x}'_{t-1}\right)', \ \mathbf{w}_t = 1 & \text{Case III} \\ \mathbf{z}_{t-1} = \left(y_{t-1}, \mathbf{x}'_{t-1}, t\right)', \ \mathbf{w}_t = 1 & \text{Case IV} \\ \mathbf{z}_{t-1} = \left(y_{t-1}, \mathbf{x}'_{t-1}\right)', \ \mathbf{w}_t = (1, t)' & \text{Case V} \end{array} \right\}.$$

$\mathbf{y}$ and $\mathbf{x}$ are generated as $y_t = y_{t-1} + \varepsilon_{1t}$, and $\mathbf{x}_t = \mathbf{P}\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_{2t}$, for $t = 1, ..., T$, where $y_0 = 0$, $\mathbf{x}_0 = 0$ and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \boldsymbol{\varepsilon}'_{2t})'$ are drawn from the $(k+1)$-dimensional independent standard normal distributions. When $\mathbf{x}_t$ is an $I(1)$ vector, we set $\mathbf{P} = \mathbf{I}_k$, but $\mathbf{P} = 0$ when $\mathbf{x}_t$ is an $I(0)$ vector. The critical values for $k = 0$ correspond to the squares of the critical values of the Dickey-Fuller (1979) unit root t statistics for Cases I, III and V, while they match with those in Dickey-Fuller (1984) unit root F statistics for Cases II and IV. The columns headed "$I(0)$" refer to the lower critical values bound obtained when $\mathbf{x}_t$ is an $I(0)$ vector, while the columns headed "$I(1)$" refer to the upper bound obtained when $\mathbf{x}_t$ is an $I(1)$ vector.

Table C2. Critical Value Bounds of the t-Statistic

Testing for the Existence of A Long-Run Relationship*

Table 2.i: Case I with no intercept and no trend

	90%		95%		97.5%		99%		mean		variance	
k	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$
0	-1.62	-1.62	-1.95	-1.95	-2.24	-2.24	-2.58	-2.58	-0.42	-0.42	0.98	0.98
1	-1.62	-2.28	-1.95	-2.60	-2.24	-2.90	-2.58	-3.22	-0.42	-0.98	0.98	1.12
2	-1.62	-2.68	-1.95	-3.02	-2.24	-3.31	-2.58	-3.66	-0.42	-1.39	0.98	1.12
3	-1.62	-3.00	-1.95	-3.33	-2.23	-3.64	-2.58	-3.97	-0.42	-1.71	0.98	1.09
4	-1.62	-3.26	-1.95	-3.60	-2.23	-3.89	-2.58	-4.23	-0.42	-1.98	0.98	1.07
5	-1.62	-3.49	-1.95	-3.83	-2.23	-4.12	-2.59	-4.44	-0.42	-2.22	0.98	1.05
6	-1.62	-3.70	-1.95	-4.04	-2.23	-4.34	-2.58	-4.67	-0.42	-2.43	0.98	1.04
7	-1.62	-3.90	-1.95	-4.23	-2.23	-4.54	-2.58	-4.88	-0.42	-2.63	0.98	1.04
8	-1.62	-4.09	-1.95	-4.43	-2.24	-4.72	-2.59	-5.07	-0.42	-2.81	0.98	1.04
9	-1.62	-4.26	-1.94	-4.61	-2.24	-4.89	-2.58	-5.25	-0.42	-2.98	0.98	1.04
10	-1.62	-4.42	-1.95	-4.76	-2.24	-5.06	-2.58	-5.44	-0.42	-3.15	0.98	1.03

Table C2.iii: Case III with unrestricted intercept and no trend

	90%		95%		97.5%		99%		mean		variance	
k	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$
0	-2.57	-2.57	-2.86	-2.86	-3.13	-3.13	-3.43	-3.43	-1.53	-1.53	0.71	0.71
1	-2.57	-2.91	-2.86	-3.22	-3.13	-3.50	-3.42	-3.82	-1.53	-1.80	0.71	0.81
2	-2.57	-3.21	-2.86	-3.53	-3.13	-3.80	-3.43	-4.10	-1.53	-2.04	0.72	0.86
3	-2.57	-3.46	-2.86	-3.78	-3.13	-4.05	-3.43	-4.37	-1.53	-2.26	0.72	0.89
4	-2.57	-3.66	-2.86	-3.99	-3.13	-4.26	-3.43	-4.60	-1.53	-2.47	0.72	0.91
5	-2.57	-3.86	-2.87	-4.19	-3.13	-4.46	-3.43	-4.79	-1.53	-2.65	0.72	0.92
6	-2.57	-4.04	-2.87	-4.38	-3.13	-4.66	-3.43	-4.99	-1.52	-2.83	0.72	0.93
7	-2.57	-4.23	-2.86	-4.57	-3.13	-4.85	-3.43	-5.19	-1.52	-3.00	0.72	0.94
8	-2.57	-4.40	-2.87	-4.72	-3.13	-5.02	-3.43	-5.37	-1.52	-3.16	0.72	0.96
9	-2.57	-4.56	-2.86	-4.88	-3.13	-5.18	-3.42	-5.54	-1.52	-3.31	0.72	0.96
10	-2.57	-4.69	-2.86	-5.03	-3.12	-5.34	-3.43	-5.68	-1.52	-3.46	0.72	0.96

Table C2.v: Case V with unrestricted intercept and unrestricted trend

	95%		95%		97.5%		99%		mean		variance	
k	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$	$I(0)$	$I(1)$
0	-3.13	-3.13	-3.41	-3.41	-3.66	-3.66	-3.97	-3.97	-2.18	-2.18	0.57	0.57
1	-3.13	-3.40	-3.41	-3.69	-3.65	-3.96	-3.96	-4.26	-2.18	-2.37	0.57	0.67
2	-3.13	-3.63	-3.41	-3.95	-3.66	-4.20	-3.96	-4.53	-2.18	-2.55	0.57	0.74
3	-3.13	-3.84	-3.41	-4.16	-3.65	-4.42	-3.96	-4.73	-2.18	-2.72	0.57	0.79
4	-3.13	-4.04	-3.41	-4.36	-3.65	-4.62	-3.96	-4.96	-2.18	-2.89	0.57	0.82
5	-3.13	-4.21	-3.41	-4.52	-3.65	-4.79	-3.96	-5.13	-2.18	-3.04	0.57	0.85
6	-3.13	-4.37	-3.41	-4.69	-3.65	-4.96	-3.96	-5.31	-2.18	-3.20	0.57	0.87
7	-3.13	-4.53	-3.41	-4.85	-3.65	-5.14	-3.96	-5.49	-2.18	-3.34	0.57	0.88
8	-3.13	-4.68	-3.41	-5.01	-3.65	-5.30	-3.96	-5.65	-2.17	-3.49	0.57	0.90
9	-3.13	-4.82	-3.41	-5.15	-3.65	-5.44	-3.96	-5.79	-2.17	-3.62	0.57	0.91
10	-3.13	-4.96	-3.41	-5.29	-3.64	-5.59	-3.97	-5.94	-2.17	-3.75	0.57	0.92

* The critical values are computed via stochastic simulations using $T = 1,000$ and 40,000 replications for the t-statistic for testing $\phi = 0$ in the following regressions: $\Delta y_t = \phi y_{t-1} + \boldsymbol{\delta}'\mathbf{x}_{t-1} + \mathbf{a}'\mathbf{w}_t + \xi_t$, $t = 1, 2, ..., T$, where $\mathbf{x}_{t-1} = (x_{1,t-1}, \dots, x_{k,t-1})'$, and

$$\left\{ \begin{array}{cc} \mathbf{w}_t = \varnothing & \text{Case I} \\ \mathbf{w}_t = 1 & \text{Case III} \\ \mathbf{w}_t = (1,t)' & \text{Case V} \end{array} \right\}.$$

$\mathbf{y}$ and $\mathbf{x}$ are generated as $y_t = y_{t-1} + \varepsilon_{1t}$, and $\mathbf{x}_t = \mathbf{P}\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_{2t}$, for $t = 1, ..., T$, where $y_0 = 0$, $\mathbf{x}_0 = 0$ and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \boldsymbol{\varepsilon}'_{2t})'$ are drawn from the $(k+1)$-dimensional independent standard normal distributions. When $\mathbf{x}_t$ is an $I(1)$ vector, we set $\mathbf{P} = \mathbf{I}_k$, but $\mathbf{P} = 0$ when $\mathbf{x}_t$ is an $I(0)$ vector. The critical values for $k = 0$ correspond to those of the Dickey-Fuller (1979) unit root t statistics. The columns headed "$I(0)$" refer to the lower critical values bound obtained when $\mathbf{x}_t$ is an $I(0)$ vector, while the columns headed "$I(1)$" refer to the upper bound obtained when $\mathbf{x}_t$ is an $I(1)$ vector.

References

[1] Banerjee, A., J. Dolado and R. Mestre (1998), "Error-correction Mechanism Tests for Cointegration in Single-equation Framework," *Journal of Time Series Analysis*, 19, 267-283.

[2] Blanchard, O.J. and L. Summers (1986), "Hysteresis and the European Unemployment Problem," *NBER Macroeconomics Annual*, 15-78.

[3] Boswijk, P. (1992), *Cointegration, Identification and Exogeneity: Inference in Structural Error Correction Models*, Tinbergen Institute Research Series.

[4] Boswijk, H.P. (1995), "Efficient Inference on Cointegration Parameters in Structural Error Correction Models," *Journal of Econometrics*, 69, 133-158.

[5] Cavanaugh, C.L., G. Elliott and J.H. Stock(1995), "Inference in Models with Nearly Integrated Regressors," *Econometric Theory*, 11, 1131-1147.

[6] Chan A., D. Savage and R. Whittaker (1995), "The New Treasury Model," Government Economic Series Working Paper No. 128, (Treasury Working Paper No. 70).

[7] Darby J. and S. Wren-Lewis (1993), "Is There a Cointegrating Vector for UK Wages," *Journal of Economic Studies*, 20, 87-115.

[8] Dickey, D.A. and W.A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427-431.

[9] Dickey, D.A. and W.A. Fuller (1981), "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, 49, 1057-1072.

[10] Engle, R.F. and C.W.J. Granger (1987), "Cointegration and Error Correction Representation: Estimation and Testing," *Econometrica*, 55, 251-276.

[11] Granger, C.W.J., and J.-L. Lin (1995), "Causality in the Long Run," *Econometric Theory*, 11, 530-536.

[12] Hendry, D.F., A. R. Pagan and J. D. Sargan (1984), "Dynamic Specification", in *Handbook of Econometrics*, Vol II, (ed.) Z. Griliches and M. D. Intriligator, 1023-1100, Elsevier: Amsterdam.

[13] Harbo, I., S. Johansen, B. Nielsen and A. Rahbek (1998), "Asymptotic Inference on Cointegrating Rank in Partial Systems," *Journal of Business Economics and Statistics*, 16, 388-399.

[14] Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59, 1551-80.

[15] Johansen, S. (1992), "Cointegration in Partial Systems and the Efficiency of Single-Equation Analysis," *Journal of Econometrics*, 52, 389-402.

[16] Johansen, S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.* Oxford University Press: Oxford.

[17] Layard R., S. Nickell and R. Jackman (1991), *Unemployment: Macroeconimic Performance and the Labour Market*, Oxford University Press, Oxford.

[18] Lindbeck, A. and D. Snower (1989), *The Insider Outsider Theory of Employment and Unemployment*, MIT Press: Cambridge, Mass.

[19] Manning, A. (1993), "Wage Bargaining and the Phillips Curve: The Identification and Specification of Aggregate Wage Equations," *Economic Journal*, 103, 98-118.

[20] Nickell, S. and M. Andrews (1983), "Real Wages and Employment in Britain," *Oxford Economic Papers, 35, 183-206.*

[21] Park, J.Y. (1990), "Testing for Unit Roots by Variable Addition," *Advances in Econometrics: Cointegration, Spurious Regressions and Unit Roots*, eds. T.B. Fomby and R.F. Rhodes, JAI Press, Greenwich.

[22] Pesaran, M.H. and B. Pesaran (1997), *Working with Microfit 4.0: Interactive Econometric Analysis*, Oxford University Press, Oxford.

[23] Pesaran, M.H. and Y. Shin (1999), "An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis," *Centennial Volume of Ragnar Frisch,* eds. S. Strom, A. Holly and P. Diamond, Cambridge University Press, Cambridge (forthcoming).

[24] Pesaran, M.H., Y. Shin and R. Smith (1998), "Structural Analysis of Vector Error Correction Models with Exogenous $I(1)$ Variables," *mimeo*, University of Cambridge.

[25] Phillips, A.W. (1958), "The Relationship between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957," *Economica*, 25, 283-299.

[26] Phillips, P.C.B. and S. Durlauf (1986), "Multiple Time Series with Integrated Variables," *Review of Economic Studies,* 53, 473-496.

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

[28] Phillips, P.C.B., and V. Solo (1992), "Asymptotics for Linear Processes," *Annals of Statistics*, 20, 971-1001.

[29] Sargan J.D. (1964), "Real Wages and Prices in the U.K.," *Econometric Analysis of National Economic Planning*, eds. P.E. Hart, G. Mills and J.K. Whittaker, Macmillan, New York. Reprinted in Hendry, D.F., and K.F. Wallis (eds.), *Econometrics and Quantitative Economics*, Basil Blackwell: Oxford, 1984, pp. 275-314.

[30] Shin, Y. (1994), "A Residual-Based Test of the Null of Cointegration Against the Alternative of No Cointegration," *Econometric Theory*, 10, 91-115.

[31] Stock, J. and M.W. Watson (1988), "Testing for Common Trends," *Journal of the American Statistical Association*, 83, 1097-1107.

[32] Urbain, J.P. (1992), "On Weak Exogeneity in Error Correction Models," *Oxford Bulletin of Economics and Statistics*, 52, 187-202.